

**Tumour Heterogeneity During the Progression of
Metastatic Breast Cancer**

and

**Anti-tumour Effects of the Novel FAK Inhibitor BI 853520 in
Breast Cancer**

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Stefanie Nicole Tiede

aus Deutschland

Basel, 2020

Originaldokument gespeichert auf dem Dokumentenserver der Universität Basel
edoc.unibas.ch

Dieses Werk ist lizenziert unter einer [Creative Commons Namensnennung 4.0
International Lizenz](https://creativecommons.org/licenses/by/4.0/).

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von:

Prof. Dr. Gerhard Christofori

Prof. Dr. Mohamed Bentires-Alj

Basel, den 19.11.2019

Prof. Dr. Martin Spiess

Dekan

Summary

Over the past decades, reasonable progress has been made in the understanding of breast cancer biology and the treatment of the primary tumour. However, the molecular contribution of multiple cancer cell clones on the various steps of tumour progression is still poorly understood. Aside from that, standard of care treatments, like the chemotherapeutic reagents cyclophosphamide and docetaxel, are rarely able to cure breast cancer patients and the overall survival rates for metastatic disease remain poor. Only in 2018, more than 620,000 women lost their lives to breast cancer, mostly due to the presence of tumour heterogeneity, an emerging drug resistance and the formation of secondary lesions. This exemplifies the unmet medical need to an in-depth understanding of tumour heterogeneity during the progression of metastatic breast cancer to finally develop new targeted therapies for this presently incurable disease.

The first project has aimed to assess clonal heterogeneity during tumour progression using the MMTV-PyMT mouse model of metastatic breast cancer expressing the Confetti lineage reporter. For this purpose, mammary epithelial cells have been induced to express one of the four Confetti reporter fluorescent proteins. The outgrowth of clonal cell populations has been analysed when the maximum tumour volume comprising all stages (normal, hyperplasia, adenoma, carcinoma, pulmonary metastases) had been reached. The Confetti lineage tracing system initially visualized the emergence of clonal heterogeneity, which culminated in clonal restriction during carcinogenesis and pointed towards a polychromatic metastatic spread. Laser capture microdissection, RNA sequencing and comparative gene expression analysis of various clonal lesions indicated a substantial level of heterogeneity across and also within the various stages of tumour progression. This intra-stage tumour heterogeneity manifested by differences in proliferation, oxidative phosphorylation and cell death and could also be observed in human breast cancer biopsies. This novel understanding of clonal variation and intra-stage heterogeneity needs to be implemented in diagnosis and therapeutic options.

In the past years, rising efforts have been made to develop agents targeting molecules and signalling pathways that are specifically present in breast cancer cells. Previous studies have linked an overexpression of focal adhesion kinase (FAK) – a cytoplasmic tyrosine kinase – with the initiation and progression of a wide variety of

Summary

malignancies, including breast cancer. This correlation of FAK and cancer, together with its role in cell migration, invasion, and proliferation, propose FAK as an attractive target for cancer therapy. In collaboration with a pharmaceutical company, we have assessed and characterized the therapeutic potential and the biological effects of BI 853520, a novel, potent and selective small chemical inhibitor of FAK, *in vitro* and in several preclinical mouse models of breast cancer. We observed a significant reduction in primary tumour growth driven by an anti-proliferative effect of BI 853520. In contrast, dissecting its influence on metastasis revealed heterogenous effects at different levels of the metastatic cascade. Hence, manipulation of FAK activity with the novel FAK-inhibitor BI 853520 offers a promising anti-tumour approach for breast cancer therapy.

In summary, my Ph.D. work delivered new insights into:

- The existence of an intra-stage tumour heterogeneity, which is conferred by clonal variations in proliferation, oxidative phosphorylation and cell death. This novel understanding of an intra-stage heterogeneity could have a significant impact on a patient's diagnosis and therapeutic response and should be implemented in clinical decision-making.
- The therapeutic potential and biological effects of the novel FAK-inhibitor BI 853520 *in vitro* and in preclinical mouse models of breast cancer. This highlighted BI 853520 as a promising anti-proliferative approach for cancer therapy.

“One general law, leading to the advancement of all organic beings,
namely, multiply, vary, let the strongest live and the weakest die.”

Charles Darwin, *The Origin of Species*

Table of Contents

Summary

1. GENERAL INTRODUCTION	1
1.1 The Mammary Gland	1
1.1.1 Murine Mammary Gland Development	1
1.1.2 Mammary Gland Structure	2
1.1.3 Hierarchy of the Murine Mammary Gland	2
1.2 Breast Cancer	3
1.2.1 Breast Cancer Subtypes	4
1.2.2 Metastatic cascade	4
1.2.3 Breast Cancer Heterogeneity	6
1.2.3.1 Intratumoural Genetic Heterogeneity and Patterns of Clonal Evolution in Breast Cancer	8
1.2.3.2 Transcriptomic Heterogeneity in Breast Cancer	9
1.2.3.3 Whole-omics Heterogeneity in Breast Cancer	10
1.2.3.4 Epithelial-to-Mesenchymal Transition-driven Breast Cancer Cell Plasticity	11
1.2.3.5 De-differentiation-driven Breast Cancer Cell Plasticity	12
1.2.3.6 Heterogeneity in Cell Proliferation	14
1.2.3.7 Interclonal Cooperativity in Breast Cancer	15
1.2.3.8 Temporal and spatial heterogeneity	17
1.2.3.9 Breast Cancer Heterogeneity and its Clinical Consequences	18
1.2.4 Models for Breast Cancer Heterogeneity	20
1.2.4.1 The MMTV-PyMT Mouse Model of Metastatic Breast Cancer	21
1.2.4.2 Lineage Tracing Studies	22
2. AIM OF THE STUDY	25
3. RESULTS	26
3.1 Multi-colour lineage tracing reveals intra-stage proliferative heterogeneity during mammary tumour progression	26
3.1.1 Abstract	27
3.1.2 Introduction	27
3.1.3 Results	28
3.1.3.1 Confetti lineage tracing in murine mammary tumours	28
3.1.3.2 Tumour progression is associated with clonal dominance	31
3.1.3.3 Polyclonal metastatic spread	33
3.1.3.4 Substantial heterogeneity across and even within tumour stages	35
3.1.3.5 Intra-tumour stage heterogeneity is due to differences in cell proliferation	38
3.1.3.6 Progression to malignancy is driven by unique pathways	41
3.1.4 Discussion	43
3.1.5 Methods	45
3.1.6 Supplemental Information	53

Table of Contents

3.2 The FAK inhibitor BI 853520 exerts anti-tumor effects in breast cancer	61
3.2.1 Abstract	62
3.2.2 Introduction	62
3.2.3 Results	63
3.2.3.1 The FAK-I BI 853520 represses Y397-FAK autophosphorylation.....	63
3.2.3.2 BI 853520 represses tumor cell proliferation and invasion only in 3D culture....	64
3.2.3.3 BI 853520 attenuates primary tumor growth	67
3.2.3.4 BI 853520 restrains tumor cell proliferation.....	70
3.2.3.5 Heterogeneous effects of BI 853520 at different stages of the metastatic cascade	74
3.2.3.6 E-cadherin status-dependent effects by BI 853520	78
3.2.4 Discussion.....	81
3.2.5 Materials and Methods.....	83
3.2.6 Supplemental information	93
4. CONCLUDING REMARKS AND FUTURE PERSPECTIVES.....	103
5. CONTRIBUTION TO OTHER PROJECTS.....	105
6. REFERENCES.....	106
7. ACKNOWLEDGEMENTS.....	118

Co-Authorship in Published Projects

1. General Introduction

1.1 The Mammary Gland

1.1.1 Murine Mammary Gland Development

Human and murine mammary glands constantly undergo dynamic remodelling processes in response to its physiologic development, the oestrogen cycle and within pregnancy (**Figure 1**). From embryonic day 13.5 on (E13.5) in the mouse, the placode-descending mammary bud infiltrates into the underlying mesenchyme. By E15.5, these buds start to sprout to form small glands in the nascent mammary fat pad. Postnatally, the murine mammary gland remains in a quiescent state until puberty starts at 3 weeks of age. Here, the mammary gland branches into the fat pad in response to oestrogen and forms the terminal end buds, which is followed by a progesterone-mediated side-branching. In pregnancy, an interplay between oestrogen, progesterone and prolactin controls alveolar expansion and in the late stages of pregnancy, prolactin induces milk secretion. Following lactation, the mammary gland involutes to its primitive state^{1,2}.

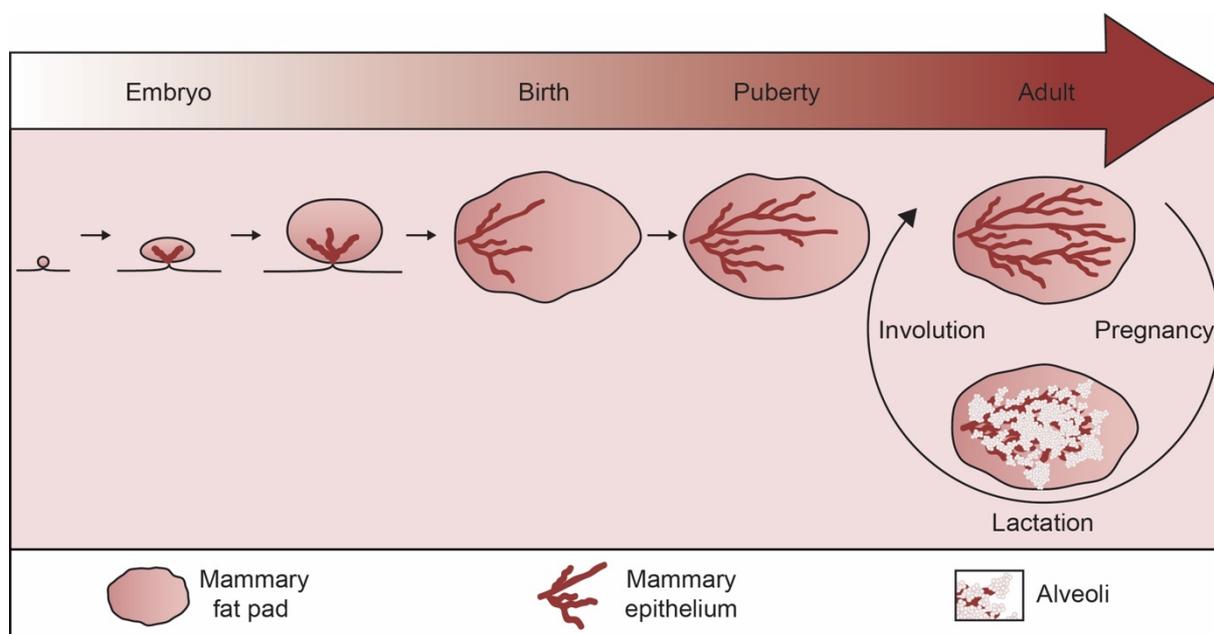


Figure 1: Remodelling processes in the mammary gland. During development, the mammary bud infiltrates into the underlying tissue and sprouts to form small glands. At puberty, the mammary epithelium branches into the mammary fat pad and forms terminal end buds. During pregnancy in adulthood, alveoli expand and milk is secreted. Following lactation, the mammary gland involutes to its original state (adapted from ²).

1.1.2 Mammary Gland Structure

The adult mammary gland mainly consists of two epithelial cell types embedded in stromal cells: an inner layer of luminal cells and an outer layer of myoepithelial cells (**Figure 2**). Luminal cells lining the ducts and alveoli express keratin 8/18 (K8/18) and/or hormone receptors (oestrogen- and/or progesterone receptors, ER/PR, respectively), whereas the basal, contractile myoepithelial cells lining on top of the basement membrane express keratin 5/14 (K5/14), and/or smooth muscle actin and/or p63^{1,2}.

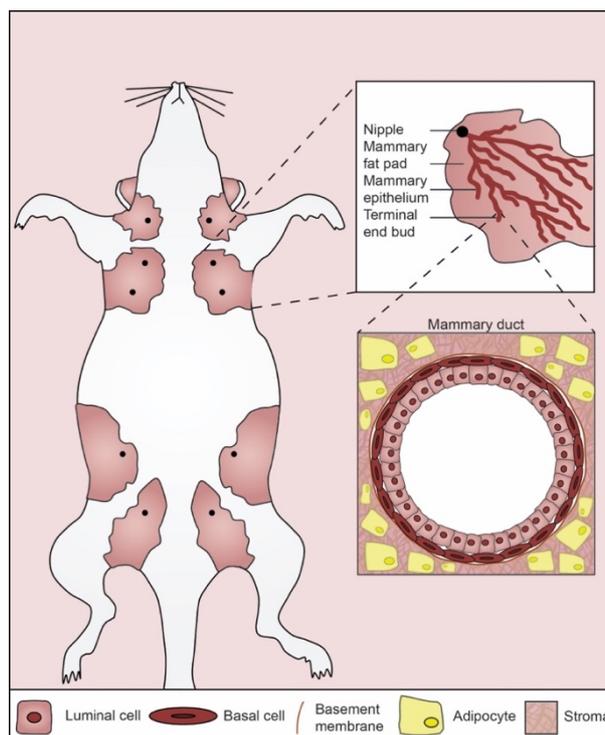


Figure 2: Schematic representation of the murine mammary gland. The mammary epithelium branches from the nipple into the mammary fat pad and forms terminal end buds (upper magnification). The mammary duct is embedded in an adipocyte-containing stroma and forms a bi-layered structure, comprising an inner luminal layer and an outer basal layer. A basement membrane is lining the outer basal layer (lower panel).

1.1.3 Hierarchy of the Murine Mammary Gland

The embryonic murine mammary gland consists of multipotent mammary stem cells (MaSC), which give rise to luminal and epithelial progenitor cells³. However, the hierarchical organization of the adult murine mammary gland is still not fully understood. K8/18 and K5/14-targeted lineage tracing studies have indicated the existence of multipotent progenitor cells during embryonic development that give rise to postnatal, long-lived, unipotent MaSCs for the luminal, as well as the myoepithelial lineage. These stem cells, with a capacity to self-renew, form their respective differentiated counterparts during puberty and adult homeostasis^{3,4}. In contrast to this,

Scheele et al. and Rios et al. have proposed a model in which a bipotent MaSC orchestrates ductal expansion during puberty and homeostasis^{5, 6}. These conflicting interpretations might be due to differences in the time of induction, labelling efficiencies, mouse models and strains⁷. Thus, further extensive research is needed to unravel the hierarchy of the physiologic mammary gland to ultimately understand the occurrence of breast cancer.

1.2 Breast Cancer

Every year, around 2.1 million women worldwide are affected by breast cancer⁸. Classically, in a simplified view, breast cancer progression is initiated by an uncontrolled proliferation of mammary epithelial cells, resulting in the formation of a premalignant, atypical hyperplasia (**Figure 3**). Transformation events, such as the loss of tumour-suppressor genes, an abnormal expression of oncogenes (like the epidermal growth factor receptor 2 (HER2/neu)) and genetic instability further accelerate progression to adenoma and pre-invasive carcinoma *in situ*, which are still confined by the basement membrane⁹⁻¹¹. Once the basement membrane is infiltrated, breast cancer cells are able to invade into the surrounding tissue (invasive carcinoma) and spread to distant organs, such as the lung, liver, bone and brain^{12, 13}.

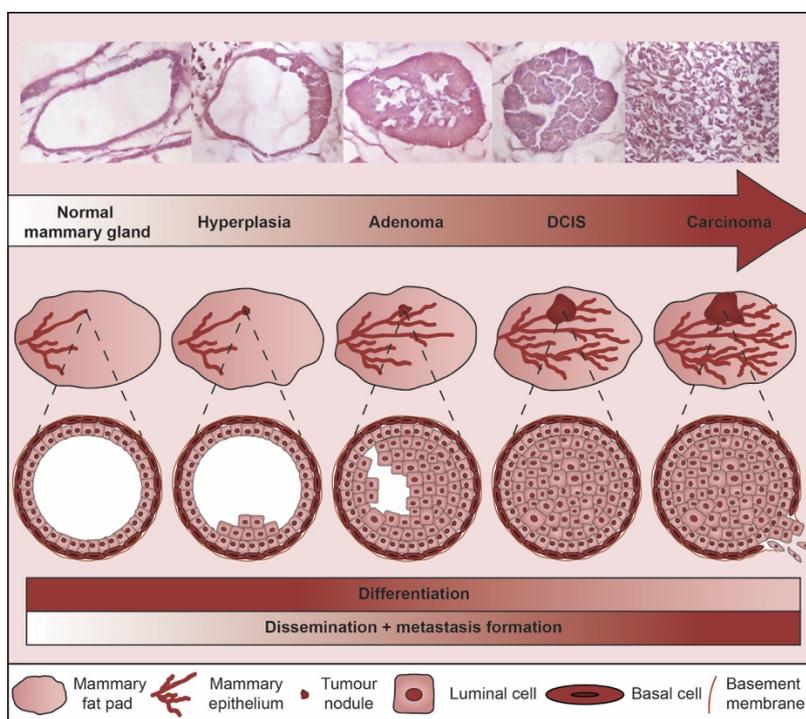


Figure 3: Haematoxylin- and eosin stained representative images (upper panels) and schematic model (lower panels) of a stepwise breast cancer progression. The zoom-ins emphasize the various stages during breast carcinogenesis, which progresses from a normal mammary gland, into hyperplasia, which is followed by the formation of an adenoma, ductal carcinoma *in situ* (DCIS) and the infiltration into the underlying basement membrane, resulting in an invasive carcinoma. During this process, cells gradually de-differentiate and gain disseminating- and seeding capacities (adapted from ^{9, 11}).

1.2.1 Breast Cancer Subtypes

The breast cancer subtype is determined by a combinatorial evaluation of histological- and immunopathological parameters and molecular profiling. Depending on the location and the invasion pattern of breast cancer cells (limited to or beyond the basement membrane), the primary tumour is graded as ductal carcinoma *in situ* (DCIS), lobular carcinoma *in situ* (LCIS), ductal carcinoma no special type (NST) or invasive lobular carcinoma (ILC). The expression pattern of ER, PR, HER2/neu and the proliferation marker Ki67, together with further genomic and transcriptomic profiling, classify the primary tumour in a specific molecular subtype^{2, 14-17}.

In 2000, DNA microarray profiling on 65 specimens of 42 different individuals has demonstrated for the first-time distinctive gene expression patterns for four breast cancer subtypes, called normal breast, luminal-like, HER2+ and basal-like, representing the presence of an enormous intertumoural heterogeneity (differences between patients)¹⁸. Ten years later, a claudin-low subtype has been added to this classification that specifically expresses epithelial to mesenchymal transition (EMT) markers, stromal- and immune related signatures as well as stem cell markers^{19, 20}.

Currently, there are six breast cancer subtypes: normal breast-like, luminal A (ER+/PR+), luminal B (ER+/PR+/HER2+), HER2-enriched (HER2+), basal-like and claudin-low. This applied classification impacts on a patient's prognosis and possible treatment options. Since the luminal subtypes of breast cancer are characterized by the expression of hormone receptors, these are mainly treated by endocrine therapy, such as tamoxifen or aromatase inhibitors, and display a relatively good prognosis. HER2-enriched breast cancers are specifically targeted by the HER2-targeting antibody trastuzumab or lapatinib, a HER2 kinase inhibitor. The absence of specific targetable molecules in triple-negative breast cancer (TNBC, including basal-like and claudin-low) makes a systemic chemotherapeutic treatment inevitable and exhibits a poor therapy response, fast clinical progression and a poor prognosis^{13, 14, 21}.

1.2.2 Metastatic cascade

Already in 1978, Fidler has elucidated the sequential multi-step process of the metastatic cascade²² (**Figure 4**). In invasive carcinomas, the underlying basement membrane is breached, and breast cancer cells invade into the surrounding tumour-associated stroma and the adjacent tissue parenchyma. Here, breast cancer cells can follow – depending on microenvironmental conditions – different modes of local

invasion: an ameboid migration (depending on Rho/ROCK activity); a mesenchymal, single-cell invasion pattern or a cohesive, multicellular, so-called collective invasion mode²³ (**Figure 5, yellow**). In the process of intravasation, locally disseminated cells enter tumour-draining blood vessels and/or lymphatics. Once in the circulation, circulating tumour cells (CTCs) travel as single CTCs, or as CTC-clusters or aggregations of CTCs and immune cells. Within minutes, these breast cancer cells arrest and/or home in secondary sites based on an interplay of various factors, e.g. physical characteristics, such as the vascular flow pattern, the diameter of the capillaries and breast cancer cells, the deformability capacity of the breast cancer cells, but also seed- and soil traits proposed by Paget in 1889^{24, 25}. At the distant organ site, tumour cells either grow into and rupture the capillaries (termed as intravascular proliferation²⁶) or actively extravasate into the underlying parenchyma. In the case of breast cancer, circulating breast cancer cells first reach the capillary beds of the lungs, where they need to beat the tight endothelium and basement membrane in the lung capillary walls. A balance between proliferation, dormancy and cell death determines whether a solitary cell is able to form a small pre-angiogenic metastasis or not. Ongoing proliferation together with the formation of new blood vessels (neoangiogenesis), a process called colonization, allow the emergence of larger, vascularized metastases, which become clinically, systemically detectable.

This metastatic process is very inefficient; it has been estimated that 0.01 % of tumour cells are capable of establishing metastases. Besides immune surveillance, obstacles including mechanical forces, niche constraints, nutrient supply and stress- and death signals need to be overcome in order to establish overt metastases. Here, (epi)genetic alterations and the later-mentioned co-option with the adjacent microenvironment dictate a tumour cell's capability to adapt and – with this – foster carcinogenesis and metastatic capacity^{12, 27-30}.

However, this complex process of tumour progression from a premalignant disease to invasive carcinoma is still not fully understood and the overall survival rates for metastatic disease remain poor. Only in 2018, over 620,000 women lost their lives to breast cancer, accounting for around 15 % of all cancer-related deaths among women⁸.

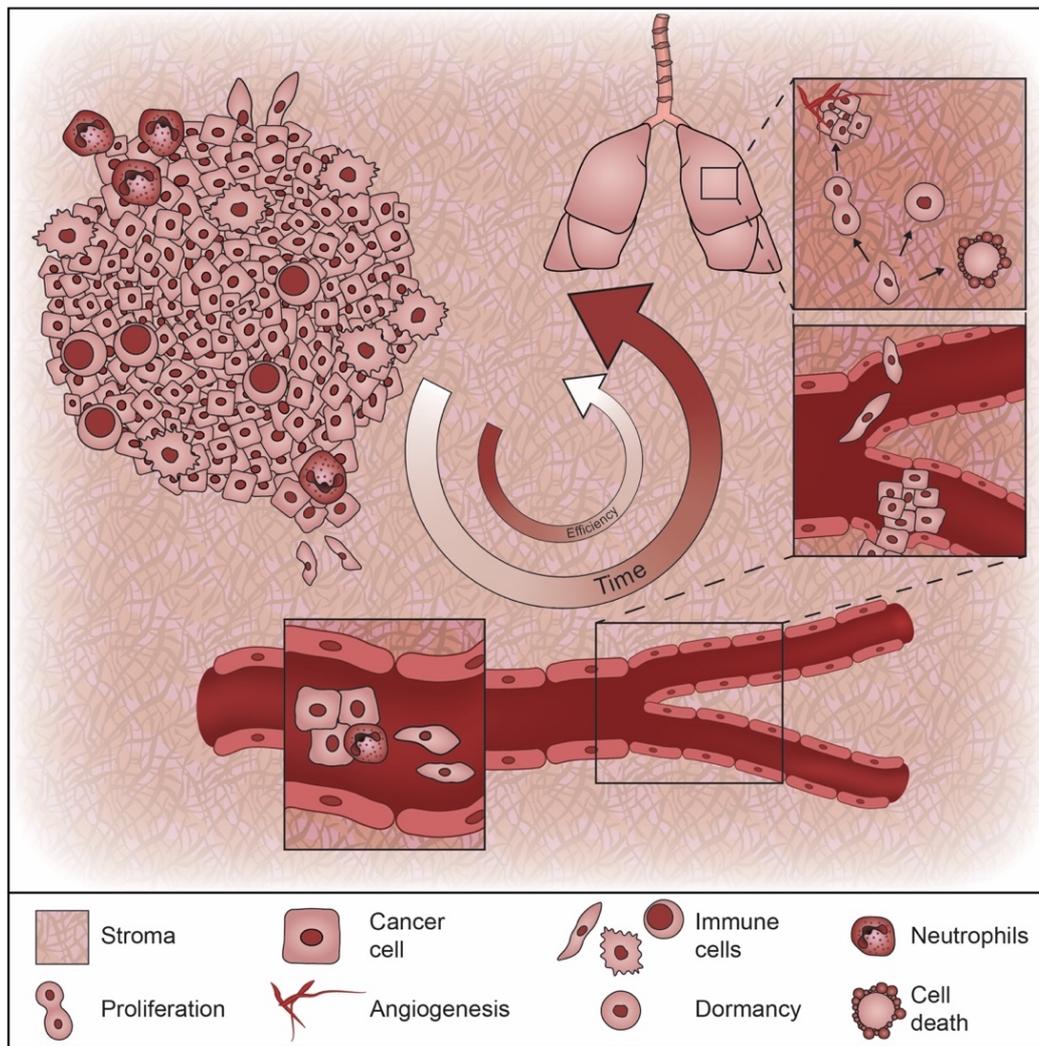


Figure 4: The metastatic cascade. Cohesive, multicellular- or single cancer cells invade and disseminate into the surrounding stroma, intravasate into the blood circulation and either rupture the capillaries by intravascular proliferation or actively extravasate into the underlying parenchyma (magnification). At distant sites, cancer cells either undergo cell death and dormancy or – in the process of colonization – proliferate and form larger, vascularized metastases.

1.2.3 Breast Cancer Heterogeneity

These cancer-related deaths are mainly due to the presence of distant lesions and therapy resistance, which are in part mediated by an extensive tumour heterogeneity³¹. Breast cancer heterogeneity is affected by differences in the following intrinsic parameters (**Figure 5**):

- Migration and invasion capacities
- Genetic profile
- Transcriptomic landscape
- Interplay between the genome, transcriptome and epigenome
- Range of cell plasticity and
- Proliferation capacities

Besides these cell-autonomous traits, also extrinsic factors in the tumour microenvironment affect breast cancer heterogeneity, such as the vasculature and tumour cell interactions with

- Fibroblasts
- Macrophages
- Neutrophils
- T-cells and others

This plethora of intrinsic and extrinsic factors vary spatially within a patient's primary tumour, metastases and between primary tumour and secondary lesions, referred to as intratumoural heterogeneity. But the intrinsic and extrinsic factors may also change stochastically and temporally during breast cancer progression, as described in the following sections^{15, 32-36}.

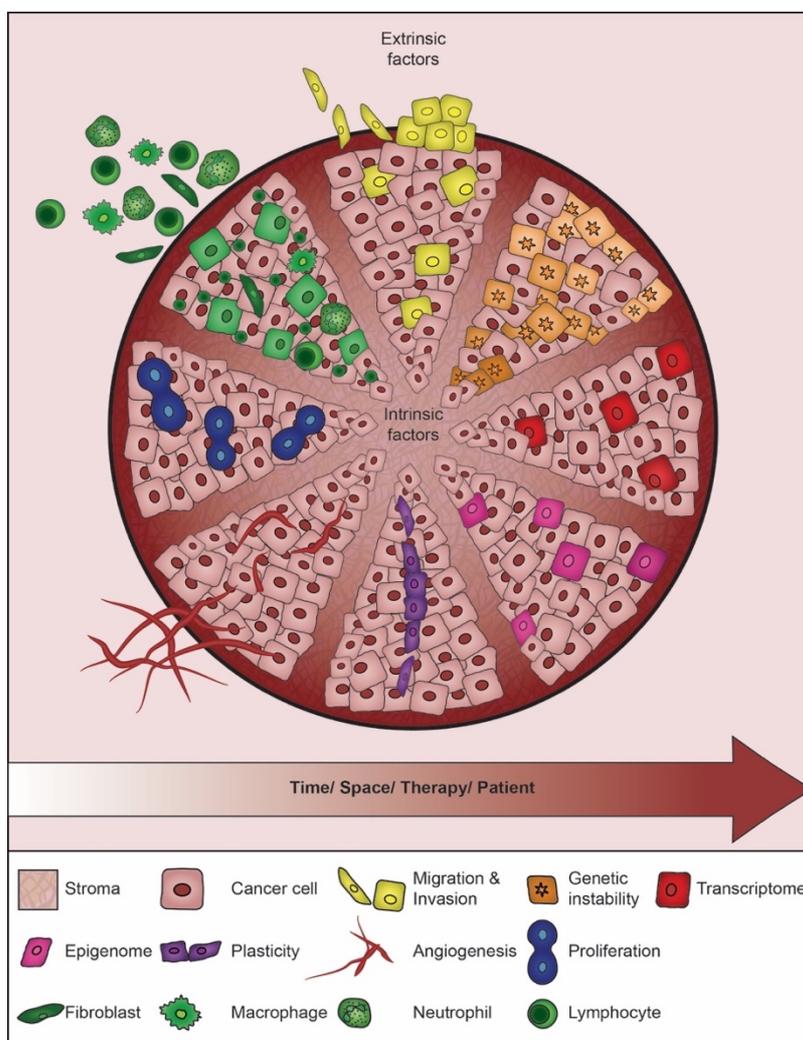


Figure 5: Intratumoural breast cancer heterogeneity and its regulation by intrinsic- and extrinsic factors. Breast cancer cells vary in migration- and invasion capabilities, genetic instability, their transcriptome and epigenome, their plasticity-capacity and their proliferation rate. Extrinsic, heterogeneous factors in the tumour microenvironment, such as the vasculature and residing immune cells can also affect intratumoural heterogeneity. Intratumoural heterogeneity changes in time and space, upon therapy and from patient to patient (intertumoural heterogeneity) (adapted from ³⁴).

1.2.3.1 Intratumoural Genetic Heterogeneity and Patterns of Clonal Evolution in Breast Cancer

Whole genome- and exome sequencing of human breast cancer samples has unveiled an extensive intratumoural genetic heterogeneity^{21, 37-40} (**Figure 5, orange**). This somatic heterogeneity ultimately results in the formation of subclones with different biological capabilities, which will follow Darwin's principle of "survival of the fittest". Already in 1976, Nowell has described the idea of a clonal evolution, where subclones with a growth advantage in a certain ecosystem will persist, survive and eventually expand, whereas less fit subclones that are unable to compete against the others, will extinct^{41, 42}.

DNA barcoding experiments with human breast tumour cells in xenograft models have demonstrated this complex heterogeneity of unchanging, enlarging and diminishing clonal populations and the resulting dramatic change in clone number, size and the emergence of new subclones within and between samples^{43, 44}. Similarly, in the majority of breast cancer xenografts undergoing deep genome and single-cell sequencing, a minor, pre-existing subclone expands and dominates the bulk tumour⁴⁵. Multi-region DNA sequencing of primary and metastatic pancreatic and renal cell tumour samples, followed by analysis of genomic rearrangements and point mutations, have identified distinct phylogenetic trees that describe the clonal evolution in these type of cancers⁴⁶⁻⁴⁸.

For breast cancer progression, genomic profiling inferred following patterns of primary tumour and metastasis evolution: a simple linear evolution model, an early metastatic spread model with parallel evolution and a late seeding pattern from a single or multiple subclone(s) within the primary tumour^{49, 50}. In a recent study by Casasent et al., topographical genomic copy number profiling of 1293 single cells from DCIS tissue from 10 breast cancer patients has proposed the so-called "multi-clonal invasion" model, which highly resembles a mixture between the simple linear evolution model and an early dissemination accompanied with a parallel evolution. Here, genomic evolution within the ducts gives rise to multiple tumour subpopulations, which subsequently co-migrate into the underlying tissue to form the invasive carcinoma⁵¹. By using genomic profiling and bioinformatic algorithms, Nik-Zainal et al. have reconstructed the evolutionary tracks of 21 breast cancer cases^{52, 53}. Here, due to the early occurrence of driver mutations and disruption of cancer genes, genetic variation arises, selective pressures are shifted, and clones expand. In contrast to these

examples of an early dissemination, sequencing of 299 samples from 170 metastatic or relapsed breast cancer patients, has highlighted a late metastatic spread and a continuous acquisition of alterations⁵⁴.

These obviously contrasting evolutionary paths might be due to distinct molecular patterns at different tumour stages in the various breast cancer subtypes, as observed by Wang et al. In this study, whole genome and exome single cell sequencing has detected a stable aneuploidy at the early stages of ER+ and TNBC progression, which was followed by gradually arising point mutations. While the ER+ subtype exhibited less mutations, the TNBC harboured a 13-fold increased mutation rate relative to normal cells, unveiling different modes of clonal evolution⁵⁵.

Regardless of an evolutionary path a tumour has taken, genetic heterogeneity, varying selection pressures and subclonal diversification enhance a primary tumour's robustness and ultimately impinge on a patients' prognosis, therapy response and clinical progression.

1.2.3.2 Transcriptomic Heterogeneity in Breast Cancer

Besides genomic heterogeneity, previous studies exploiting microarray and RNA-sequencing analyses have pointed towards an enormous heterogeneity in the transcriptomic profile of primary and distant breast cancer lesions (**Figure 5, red**).

In 2002, microarray analysis of 117 breast tumours has been used to decipher three gene expression signatures that potentially predict therapy response: a signature indicative of poor prognosis, which harboured genes important for cell cycle progression, invasion, metastasis formation and angiogenesis; an ER signature, included K18, BCL2, ERBB3 and ERBB4 expression and a signature for carriers of BRCA1 mutations, a well-known tumour-suppressor gene⁵⁶. Similarly, genome-wide gene expression analysis on lymph-node negative breast cancer patients has revealed gene expression patterns that hint towards an increased risk of distant recurrence⁵⁷. Importantly, these gene expression signatures can be exploited to predict an individual patients' treatment response and finally optimize therapy.

Furthermore, gene expression profiling of various cell types of normal breast tissue, DCIS and invasive breast cancer has shown extensive gene expression changes in secreted proteins and receptors not only in tumour cells, but also in stromal cells. For example, overexpression of the chemokines CXCL14 in myoepithelial cells and CXCL12 in myofibroblasts stimulate the proliferation, migration and invasion

capacities of epithelial cancer cells in a paracrine manner⁵⁸. Recently, a similar heterotypic interaction has been highlighted also during the progression of metastatic breast cancer. Here, single-cell RNA sequencing on CTC-neutrophil clusters has demonstrated the cell adhesion molecule VCAM1 as an important mediator of CTC-neutrophil interaction and neutrophil-derived cytokines (IL-6 and IL-1 β) as important stimuli to enhance cell-cycle progression and metastatic seeding⁵⁹.

Lastly, there is also a tremendous level of heterogeneity in organ-tropism: *in vivo* selection and transcriptomic analysis of bone and lung-homing breast cancer cell lines have determined numerous genes that are needed for specific organ-tropisms and site-specific outgrowths^{60, 61}.

Taken together, these and other studies show that transcriptomic heterogeneity increases a cell's adaptivity to constitutively changing constraints, which ultimately fosters breast cancer progression.

1.2.3.3 Whole-omics Heterogeneity in Breast Cancer

Technological advances and the development of novel whole-omics profiling techniques offer the possibility to better comprehend the full extent of tumour heterogeneity^{34, 62} (**Figure 5, orange, red, magenta**).

Around 2010, first attempts have been made to simultaneously sequence the genome and the transcriptome of an ER+ metastatic lobular breast cancer and two other human breast cancer samples^{63, 64}. In 2012, Curtis et al. have described the genomic and transcriptomic landscape of 2,000 human breast cancer samples by using an integrated analysis of copy number variations and gene expression patterns³⁹. Similarly, at the same time, the Cancer Genome Atlas Network has combined various methods to comprehend the molecular portraits of human breast tumours. Using a genomic DNA copy number array, in combination with DNA methylation analysis, exome sequencing, mRNA arrays, microRNA sequencing and a reverse-phase protein assay, the four distinct breast cancer subclasses were described in a more comprehensive and holistic view, revealing a tremendous amount of variation even within each subtype³⁸.

In a recent study by Obradovic et al., transcriptomic, proteomic and phospho-proteomic profiling of 17 patient-derived xenograft models and transplanted human cell lines and their corresponding metastatic lesions have indicated an increased stress response and glucocorticoid receptor activity during tumorigenesis, resulting in an

increased colonization capacity and a decreased survival⁶⁵. Interestingly, Gkoutela et al. have utilized whole-genome bisulfite sequencing in combination with RNA-sequencing to elucidate the methylation landscape of single CTCs and CTC clusters. In contrast to single CTCs, CTC clusters exhibited an augmented metastatic capacity and hypomethylated (thus active) binding sites in embryonic stem cell genes (OCT4, SOX2, NANOG, SIN3A), which can be therapeutically targeted by Na⁺/K⁺-ATPase inhibitors⁶⁶. Similarly, ATAC, RNA and ChIP-sequencing of fetal mammary cells, adult basal cells, hormone receptor negative (HR⁻) luminal cells and HR⁺ luminal cells have disclosed the transcription factor Sox10 as a plasticity factor in adult basal cells, that confers multi-lineage differentiation potential and invasion capacities, ultimately contributing to carcinogenesis⁶⁷. Remarkably, single cell DNA, RNA, and exome deep sequencing of 20 TNBC patients pre, mid and post neoadjuvant chemotherapy have unravelled the presence of pre-existing, adaptive genomic alterations and further acquired transcriptional reprogramming during the evolution of chemoresistance⁶⁸.

All these studies hint towards an enormous complexity of genetic, transcriptomic, epigenetic and proteomic variations, which need to be further delineated⁴⁹.

1.2.3.4 Epithelial-to-Mesenchymal Transition-driven Breast Cancer Cell Plasticity

All of a cell's intrinsic characteristics, including genetic, transcriptomic, epigenetic and proteomic traits collectively determine the phenotypic state of a cell. This phenotypic state is not a fixed determinant, but may change in time and space, which reflects the plasticity of a cell^{33, 69} (**Figure 5, purple**).

An important example of cell plasticity is epithelial-mesenchymal plasticity, which is governed by the process of an EMT and its reverse process, a mesenchymal-to-epithelial transition. During an EMT, epithelial markers, such as E-cadherin, EpCAM and ZO-1 are lost, while mesenchymal markers and transcription factors, like N-cadherin, vimentin, fibronectin, Snail1/2, Twist1 and Zeb1/2 are newly expressed⁷⁰. This leads to a disassembly of epithelial cell-cell junctions (tight, adherens and gap junctions and desmosomes) and changes an apical-basal polarity into a front-rear polarity, conferring a cell with migratory- and invasive capabilities⁷¹. In addition, EMT has been associated with increased stemness and chemoresistance⁷¹⁻⁷³.

Generally speaking, an EMT is a transient and very heterogenous process, that is transcriptionally orchestrated in a stepwise manner by complex gene regulatory

networks and specialized signalling hubs. RNAi-mediated functional screening and transcriptomic profiling of normal murine mammary gland cells undergoing an EMT have discovered an epistatic, hierarchical regulatory network consisting of diverse transcription factors and micro-RNAs that orchestrate the multiple steps of an EMT⁷⁴. A recent study by Pastushenko et al. has used mouse models of skin squamous cell carcinoma and breast cancer to study distinct transition states occurring during EMT *in vivo* and to elucidate their specific functional characteristics^{75, 76}. Here, the authors identified six populations within an EpCAM⁺ population. Depending on the state, a cell varies with regard to its range of plasticity, its invasive potential and its metastatic capacity⁷⁷. Similarly, single-cell RNA sequencing of matched primary and metastatic head and neck squamous cell carcinomas of 18 patients has uncovered partial EMT cells at the leading edge of primary tumours, which predict poor pathologic features and distant seeding⁷⁸.

Next to these examples of a classic, transcriptional regulation of EMT, a recent study by Aiello et al. has unravelled a post-transcriptional regulation of EMT by sole protein relocalization⁷⁹. In well-differentiated pancreatic ductal adenocarcinoma (PDAC), internalization of epithelial proteins renders PDAC cells in a final, partial EMT state, where epithelial and mesenchymal genes as well as cell-cell contacts are retained. However, poorly-differentiated PDAC cells undergo the well-known, complete EMT program, in which epithelial genes are transcriptionally repressed. Depending on the EMT program, PDAC cells disseminate as collective strands or single cells. These EMT programs are conserved in several carcinomas, including breast cancer and confer tumour heterogeneity.

Despite its supposed association with tumour progression, the consequences of a spontaneous EMT on metastasis formation and chemoresistance are not well defined. Intriguingly, a recent study by Ishay-Ronen et al. has found that an EMT-induced cell plasticity can be exploited to enforce a trans-differentiation of breast cancer cells into post-mitotic cancer-derived adipocytes, thereby inhibiting breast cancer progression and distant seeding⁸⁰.

1.2.3.5 De-differentiation-driven Breast Cancer Cell Plasticity

However, cancer cell plasticity cannot only be elicited by an EMT, but also by (de-)differentiation events that can ultimately result in, for example, the emergence of

stem-like and cancer stem cells (CSCs), low and high-burden secondary lesions, asymmetric cell divisions and vascular mimicry (**Figure 5, purple**).

In 2015, two studies have elucidated the effect of PIK3CA mutations – one of the most common mutations in breast cancer – on mammary tumours. Remarkably, the PIK3CA^{H1047R} mutation induces Lgr5⁺ basal and K8⁺ luminal cells to de-differentiate into multi-potent stem-like cells, which on its own can again give rise to either Lgr5⁺ basal or K8⁺ luminal cells. The tumour cell of origin determines in the end tumour aggressiveness^{81, 82}.

In addition to a mutation-induced de-differentiation process, stochastic state transitions can also equilibrate cancer cell populations. Fluorescence-activated cell sorting of human breast cancer cell lines into specific differentiation states with distinct functional capacities (stem-like, basal and luminal), followed by cell expansion, have revealed a stochastic transition of each subpopulation in one or the other. This process of a dynamic and reversible, bidirectional interconversion, also known as Markov model, is in stark contrast to the classical CSC model, in which a self-renewing CSC gives rise to one CSC and one non-CSC^{83, 84}. In accordance with the Markov model, Chaffer et al. have shown bivalent/poised chromatin marks at the Zeb1 promoter, a well-known EMT-inducing transcription factor, that – once activated by microenvironmental cues, such as the cytokine TGF β – converts basal CD44^{low} non-CSCs into basal CD44^{hi} CSCs, which further enhance breast tumorigenesis and metastasis formation⁸⁵.

There is also a considerable heterogeneity in the differentiation of metastatic lesions of human triple-negative, basal-like breast cancer cells: low-burden metastatic cells have a high expression in genes regulating stemness, anti-apoptosis, dormancy and EMT, while high-burden metastatic cells are more heterogenous, have a differentiated/luminal gene signature, a great proliferation capacity and a decrease in genes regulating dormancy and EMT. Here, experimental evidence has suggested that metastasis is initiated by stem-like cells, which differentiate into luminal-like, more advanced secondary lesions⁸⁶.

Furthermore, in TNBC, asymmetric divisions yielding K18⁺/K14⁻ luminal-like as well as K18⁺/K14⁺ progenitor-like cells provoke differentiation transitions and a phenotypic cellular heterogeneity, in which K18⁺/K14⁺ progenitor-like cells are associated with an increased tumorigenic capacity⁸⁷.

Interestingly, breast cancer cells not only de-differentiate into stem-like or more progenitor-like cells, but also into cells that confer microenvironmental stimuli. For example, expression of epiregulin, the enzyme cyclooxygenase COX2 and the matrix metalloproteinases MMP1- and 2 in human breast cancer cells facilitates neoangiogenesis and their intra and extravasation capacities for distant seeding⁸⁸. In a study by Wagenblast et al., retroviral barcoding and subsequent mammary fat pad injection of 4T1 breast cancer cells have deciphered distinct, specialized clones that either dominate the primary tumour, or have the capacity to disseminate or that harbour a distinct organ-tropism. Clones exhibiting the capacity to intravasate into the vasculature specifically expressed two releasing proteins (Serpine2 and Slpi) to establish vascular networks and to accommodate a suitable perfusion. By doing so, a proper nutrient supply is ensured, and metastatic spread is enhanced. This phenomenon is also known as vascular mimicry and displays an additional outcome of cancer cell plasticity⁸⁹ (**Figure 5, vasculature**).

1.2.3.6 Heterogeneity in Cell Proliferation

Aside from plasticity-driven heterogeneity, there is considerable variation in proliferation properties in cells within the primary tumour, in CTCs and even in secondary lesions (**Figure 5, blue**).

By titrating the numbers of implanted murine lymphoma cells, Kelly et al. have demonstrated that lymphoma tumour growth does not require the existence of rare cancer stem cells, but that lymphoma cells with and without characteristic stem cell markers were able to form tumours⁹⁰. Examining this heterogeneity in cell proliferation further, Roesch et al. have characterized within rapidly proliferating melanoma cells a slow-cycling subpopulation, which expresses the H3K4 demethylase JARID1B, and which is substantial for tumour growth⁹¹. Further studies hint towards an increased invasion capacity, metastatic seeding and drug resistance of slow-cycling subpopulations in melanoma and other cancer types⁹²⁻⁹⁴.

There is a similar difference in the proliferation potential of single CTCs vs. CTC clusters: while single CTCs do not possess open binding sites in proliferation-associated transcription factors, CTC clusters are enriched in proliferation-related genes and display an increased expression of the proliferation marker Ki67⁶⁶.

Even metastatic lesions exhibit these differences in proliferation: while micrometastases maintain a delicate balance in proliferation, apoptosis and dormancy,

evolving macrometastases undergo rapid cycles of proliferation and form large, clinically relevant metastases^{27, 30}.

This heterogeneity in cell proliferation needs to be taken into account in order to eradicate the whole tumour mass⁹⁵⁻⁹⁷ and will be further contemplated in the first part of my Ph.D. thesis.

1.2.3.7 Interclonal Cooperativity in Breast Cancer

Due to cell intrinsic factors, various breast cancer subclones emerge that can functionally cooperate by autocrine and paracrine signalling networks to drive breast cancer progression (**Figure 5, green**).

In Wnt-1-driven mammary tumours, for example, a heterogenous mixture of basal and luminal cancer cells supports their own tumour growth by interclonal cooperation since neither of the cell subpopulations alone is able to sustain tumour maintenance. Here, cancer-driving, Hras-mutant basal cells are deficient for Wnt signalling, but instead recruit heterologous Hras-wildtype, luminal cells to secrete Wnt-1 for bicalonal tumour growth^{98, 99}. A recent study by Malladi et al. has identified a stochastically occurring inhibition of Wnt-signalling caused by autocrine secretion of Dkk1 as an important determinant for the formation of slow-cycling breast cancer cells. These latent, stem-cell-like cells (expressing the transcription factors Sox2 and Sox9) decreased their cell surface innate immune sensor ULBP to evade a natural killer cell-mediated immune surveillance, thereby enabling their long-term persistence¹⁰⁰. Furthermore, metastasis-initiating breast cancer cells have been shown to secrete the extracellular matrix protein tenascin C in an autocrine manner, which stimulates Notch and Wnt signalling pathways and enhances their survival and metastatic fitness¹⁰¹. Along this line, in mammary epithelial cells autocrine and paracrine induced TGF β and (non)canonical Wnt signalling induce and maintain a mesenchymal/stem-cell like phenotype, increase tumorigenicity and enhance metastasis formation¹⁰².

However, breast cancer subclones cannot only cooperate with neighbouring clones within the primary tumour, but also with cells of metastatic lesions. Classically, it was believed that there is only unidirectional seeding from the primary breast tumour. However, a study by Kim et al. has pointed towards a bi-directional spread from distant, metastatic lesions back to the primary tumour. This so-called self-seeding did not require further adaptation and was conferred by an interleukin mediated (IL6 and IL8) CTC attraction by the primary tumour. Once CTCs arrived in the compatible soil of the

primary tumour, they were able to foster tumour growth, angiogenesis and stromal recruitment¹⁰³. Furthermore, whole-genome sequencing of matched primary and metastatic prostate cancer samples has indicated interclonal cooperation and polyclonal seeding from metastasis-to-metastasis¹⁰⁴.

Besides these homotypic, cancer-cell specific interactions that foster breast cancer progression, there is also heterotypic cooperation of breast cancer cells with various types of cells in the tumour microenvironment, e.g. fibroblasts, macrophages, neutrophils and T-cells, that dictate the pace of tumorigenesis, the occurrence of distant seeding and a patient's therapy response¹⁰⁵⁻¹⁰⁹ (**Figure 5, green**).

Experimental evidence from Marusyk et al. has verified – by transplanting a panel of sublines overexpressing single secreting factors implicated in tumour progression – such non-cell-autonomous driving forces, indicating that factors derived from the microenvironment are indispensable for subclonal heterogeneity¹¹⁰. Indeed, in breast cancer, fibroblast-derived caveolin-1 favours an increased contraction, matrix alignment and stiffening of the microenvironment. This, in turn, enhances breast cancer cell elongation, directional migration, and metastatic capacity¹¹¹. Another fibroblasts-breast cancer cell cooperativity has been exemplified by Malanchi et al., where cancer stem cells induced a fibroblast-mediated secretion of periostin, an extracellular matrix protein, to engage their Wnt-signalling and metastatic colonization capacity¹¹². Remarkably, stromal cells are also able to transfer exosomes harbouring noncoding transcripts and transposable elements to breast cancer cells, which induces an antiviral signalling pathway. This together with a juxtacrine JAG1-NOTCH3 mediated interaction is responsible for tumour-initiating cell expansion, resistance to radiation and chemotherapy and breast cancer re-initiation¹¹³. These and other studies highlight the emerging, causative role of stromal fibroblasts during breast cancer progression¹¹⁴.

Next to stromal fibroblasts, also macrophages are implicated in breast carcinogenesis. Multiphoton microscopy on primary murine mammary tumours has visualized a paracrine loop between breast cancer cells and macrophages close to vessels. A breast cancer-mediated secretion of the cytokine CSF-1 stimulated the perivascular macrophages to release the growth factor EGF; thereby enhancing breast cancer cell migration, invasion and metastatic potential^{115, 116}.

In addition to this breast cancer cell-macrophage cooperativity, also breast cancer-neutrophil cooperation has been illustrated recently by RNA-sequencing

analysis of polyclonal tumours. Polyclonal, metastatic breast tumours harbouring IL11 and VEGF-D-expressing subclones, systemically primed pulmonary stromal cells to secrete neutrophil-stimulating factors. Attracted, pro-tumorigenic neutrophils, in turn, stimulated the breast cancer cells' capacity to disseminate^{117, 118}.

Furthermore, immunogenomic analysis of one high-grade ovarian cancer patient treated with various chemotherapeutic reagents has illustrated the remarkable interplay of T-cells with the course of tumour progression: while regressing metastases harboured massive T-cell infiltrations, progressing secondary lesions were devoid of T-cells¹¹⁹.

Thus, these and other studies exemplify how interclonal cooperativity between various breast cancer subclones, but also between breast cancer and stromal cells (including fibroblasts, macrophages, neutrophils and T-cells), can have a great impact on breast cancer progression and therapy response.

1.2.3.8 Temporal and spatial heterogeneity

Cell-intrinsic and non-cell-autonomous traits do not only vary temporally throughout tumour progression, but also spatially within a primary tumour and its secondary lesions¹²⁰ (**Figure 5**). This intratumoural heterogeneity is particularly striking in a study, where multiregion sequencing was performed on 12 invasive primary breast cancers to representatively sample the whole tumour mass and to assess the geographical distribution of subclones. In order to do so, primary tumours were surgically removed, and six needle biopsies were resected from each half of a tumour and analysed by targeted gene sequencing. Most of the breast cancer biopsies exhibited one mutation confined to one to three neighbouring tissue regions, indicating a locally constrained expansion of subclones, while other breast cancer samples displayed clonal intermingling throughout the tissue. Remarkably, targetable mutations and landmarks of therapy resistance and invasive capability occurred in detectable subclones¹²¹. Similarly, multi-focal breast cancer lesions with a comparable grade and expression pattern of ER and HER2 status underwent whole-genome sequencing and targeted sequencing of 360 cancer-related genes. In two-thirds of the patients, all lesions shared specific substitutions/indels and were clonally related. Topographically, genetically similar lesions were located closer to each other than genetically distinct lesions, pointing towards an intra-mammary spread. Interestingly, despite pathological homology, in one-third of the patients, an enormous inter-lesion heterogeneity was

found for oncogenic mutations in PIK3CA, TP53, GATA3 and PTEN, the most frequently altered genes in breast cancer¹²². In a similar study performed by Miron et al., targeted sequencing of exon 9 and 20 of the PIK3CA gene of DCIS tissue, DCIS adjacent to invasive carcinoma lesions and IDCs, identified the same frequency (30 %) of PIK3CA mutations in all three histologically distinct groups. Nevertheless, in some cases of pre-invasive and invasive lesions, a discordance – thus intratumoural heterogeneity – in PIK3CA status was detected¹²³. Similarly, PIK3CA sequencing of single CTCs from metastatic breast cancer patients has revealed heterogeneity in the PIK3CA mutation status in one-third of PIK3CA-mutated CTCs¹²⁴. Furthermore, when morphological distinct lesions of six metaplastic breast carcinomas were laser capture microdissected and analysed by genomic hybridization, immunohistochemistry and *in situ* hybridization, they comprised a similar genetic pattern and clonal relationship with the exception of two lesions, that contained distinct genetic alterations¹²⁵. Thus, these and other studies highlight the presence of an enormous spatial heterogeneity in breast cancer during the various steps of malignant progression.

Waclaw et al. have developed a model for this spatial and temporal tumour evolution, in which virtually every single, tumorigenic cell can expand and dominate its respective lesions. They report that already a small, advantageous selection-driven cell turnover and a short-range spread result in a clonal intermixing, replacement of precursor cells and can account for chemotherapy resistance¹²⁶.

1.2.3.9 Breast Cancer Heterogeneity and its Clinical Consequences

The polyclonal nature of breast cancer in time and space increases a cell's adaptivity to constitutively changing constraints and is a significant hurdle for diagnosis and therapeutic response^{15, 120}. Thus, for a justified diagnosis and therapeutic decision, it is important to mitigate the risk of sampling bias and to evaluate the existing, substantial heterogeneity on a whole tumour-level. Comparison of multiple, spatially separated breast tumour sections, together with a non-invasive CTC analysis and the evaluation of matched, topographically dispersed metastatic lesions throughout breast oncogenesis is necessary during diagnosis and might enhance our understanding of intratumoural heterogeneity^{15, 32, 34, 62, 120, 127}. This idea needs to be approached from various angles (genetics, transcriptomics, proteomics, etc.) on bulk- and single tumour cells with the help of recent technologies to ultimately find tailored, suitable treatment combinations¹²⁸⁻¹³².

Based on Paget proposed “seed and soil” hypothesis, there are two treatment strategies: therapeutic regimens can target either the seed of a tumour, thus the cancer cells itself as exemplified by the HER2-targeting antibody trastuzumab in HER2-enriched breast cancers, or the soil, thus a tumour’s microenvironment^{25, 29, 107}. Examples for treating the tumour stroma in breast cancer is the anti-angiogenic antibody Bevacizumab (Avastin®), which binds to and inhibits the vascular endothelial growth factor VEGF-A¹³³.

However, “seed and soil” can also be simultaneously targeted by inhibiting focal adhesions, which are not only expressed by the cancer cells itself, but also by cells of the tumour microenvironment. Focal adhesions, the interacting link between the actin cytoskeleton and the extracellular matrix, are composed of integrin and signalling proteins, such as the cytoplasmic tyrosine kinase focal adhesion kinase (FAK). Once recruited by integrin- and growth factor-mediated signalling, FAK undergoes a conformational change, is autophosphorylated and serves as scaffold for Src family kinases¹³⁴. A fully active FAK-Src signalling complex regulates a plethora of downstream signalling pathways essential for cell migration, proliferation and death – cellular functions that are also exploited by cancer cells during tumour progression¹³⁵⁻¹³⁷. Hence, not surprisingly, previous studies have linked FAK expression levels with the initiation, progression and prognosis of a wide variety of malignancies, such as ovarian, head and neck, and breast cancer¹³⁷. Shibue et al. have discovered a FAK-dependent activation of the classical MAP kinase ERK, which in turn stimulated extravasated breast cancer cells in the lungs to survive and proliferate to form macrometastases^{71, 138, 139}. Remarkably, enhanced FAK signalling has also been correlated with drug- and chemoresistance in melanoma- and lymphoma^{36, 140, 141}. Here, endothelial cell-specific loss of FAK suppressed a DNA-damaged mediated NF- κ B activation and cytokine secretion from endothelial cells, thereby sensitizing cancer cells to chemotherapy and decreasing tumour growth¹⁴¹. Exploring this correlation of FAK expression and chemoresistance further, intravital imaging on BRAF-mutant melanomas emphasized a FAK-mediated reactivation of ERK/MAPK cascade, an increased survival signalling and the emergence of residual disease¹⁴⁰. These well-established tumorigenic activities together with its high expression and activity in different cancer types, suggests FAK as a promising target for cancer therapy, which has driven the screening and validation of small-molecule inhibitors¹⁴²⁻¹⁴⁵. The

assessment of the therapeutic potential of a novel, selective FAK-inhibitor will be delineated in the second part of my Ph.D. thesis.

While a targeted approach implicates the risk of an overtreatment of cells, which do not rely on the specified target or which circumvent the treatment by upregulating other relevant pathways, stromal cells are genetically more stable, which might attenuate the risk of resistance and recurrence^{108, 109, 146}. Just the presence, or the emergence of a single, drug-resistant cell (intrinsic vs. adaptive resistance) results in the expansion of a drug-resistant tumour within weeks or months after therapy initiation^{28, 50, 69}. Therefore, various combinatorial treatment options might be applied with(out) chemotherapy to eradicate the polyclonal behaviour of breast cancer to achieve the optimum therapeutic benefit^{33, 34}.

1.2.4 Models for Breast Cancer Heterogeneity

Until now, there are various model systems available to rebuild and understand breast cancer heterogeneity and to evaluate treatment regimens in a preclinical setting: 2D and 3D-cultures of murine and human breast cancer cell lines, syngeneic and xenograft transplantations, patient-derived xenografts (PDX), and autochthonous murine mammary tumour models. Each preclinical system implies model-specific advantages and disadvantages. While 2D and 3D-culture systems, such as tumour slices and organoid cultures, are robust, easy to manipulate and allow co-culturing of various cell types, the microenvironmental aspect of breast cancer heterogeneity cannot be implemented. However, syngeneic and xenograft transplantation models allow to explore the cooperation of breast cancer cells with its microenvironment. Nevertheless, transplanted murine and human breast cancer cell lines are originally homogenous cell lines, which are expanded *in vitro* and cannot recapitulate the polyclonal nature of breast cancer cells. PDX-models are able to incorporate these features of a heterogenous cell population of the original patient sample and the cooption with a host-derived microenvironment but lack the interaction with an intact immune system due to a cross-species incompatibility. Since in autochthonous murine mammary tumour models the tumour evolves over time in a native tissue context, this model incorporates the polyclonal nature of tumour cells and allows to study cooperations with immune- and microenvironmental cells^{15, 62, 69, 107, 109}.

1.2.4.1 *The MMTV-PyMT Mouse Model of Metastatic Breast Cancer*

Various genetically modified mouse models have been engineered to reconstruct breast cancer heterogeneity. Therefore, mammary gland-specific promoters (such as the mouse mammary tumour virus) have been exploited to drive the expression of oncogenes, e.g. polyomavirus middle T antigen (PyMT), HER2/neu and Wnt-1, amongst others. These mouse models in combination with a tamoxifen or tetracycline-inducible Cre/lox-P-mediated recombination or a Tet-On/Tet-Off system allow to induce a gene of interest specifically in time and space¹⁴⁷.

In the murine metastatic breast cancer model MMTV-PyMT, the mouse mammary tumour virus (MMTV) promoter is used to drive the mammary-gland specific expression of the PyMT oncogene. In female mice, this results in constitutive signalling of c-Src and PI3K, deregulated cell proliferation and widespread malignant transformation of the mammary gland epithelium from age 2-4 weeks on without further chromosomal aberrations. This ultimately culminates in the development of luminal-B type, multifocal mammary adenocarcinomas during a short latency and pulmonary metastases within 14 weeks in about 94 % of female mice. The MMTV-PyMT mouse model of metastatic breast cancer nicely resembles the human breast cancer pathogenesis by its stepwise progression to malignancy and the expression of specific biomarkers. During the various stages of mammary oncogenesis, the expression of ER, PR and integrin β 1 is gradually lost, while the expression of HER2/neu and the cell cycle-regulatory protein cyclin D1 increase over time^{11, 147-149}. During the pre-malignant stages more than 50 % of the cells are cyclin D1-positive, which are mainly located at the outer layer of luminal ducts but increase and redistribute as clusters through the entire tumour during tumorigenesis¹¹.

Maglione et al. have characterized various MMTV-PyMT-derived hyperplastic lesions from different topographical zones with regard to histopathological parameters and their transplantation potential. Various stable hyperplastic outgrowth lines possessed their own morphology, tumour latency and metastatic capacity; thus, representing the heterogenous nature of each lesion^{150, 151}. Microarray-analysis of the normal mammary fat pad and (pre)malignant lesions has revealed an upregulation of cell-cycle related genes and oncogenes during the premalignant transition¹⁵². Consistent with this, gene expression profiling of mammary tumours of different MMTV-PyMT mouse strains has shown a shared upregulation of genes mediating cell growth and a common downregulation of genes implicated in cell adhesion when

compared to normal tissue¹⁵³. MMTV-PyMT tumour piece transplantation into FVB hosts, followed by a combinatorial treatment with the chemotherapeutic reagents docetaxel, doxorubicin and cyclophosphamide and then microarray analysis, discovered biomarkers of dissemination, distant seeding and resistance¹⁵⁴. RNA-sequencing of bulk primary breast tumours of the MMTV-PyMT mouse model at various stages of mammary tumour progression, has indicated a high transcriptomic overlap between the various breast cancer stages with a high preservation in modules regulating oxidative phosphorylation, the Krebs-cycle, DNA replication and cell-cycle progression¹⁵⁵.

However, it is still questionable whether these genetically modified murine models are able to adequately rebuild human breast cancer pathogenesis. Yet, DNA microarray analysis, gene expression profiling and transcriptomic classification of diverse murine mammary tumour models and cell subpopulations have verified the conservation of human subtype-specific characteristics and shared genes and pathways in murine samples¹⁵⁶⁻¹⁵⁹. Intriguingly, these studies have validated the classification of the MMTV-PyMT mouse model as luminal B-type of breast cancer, with high expression of the human luminal-defining genes XBP1 and K8/18 and an enrichment in genes that predict secondary lesions^{156, 157, 159}.

Despite the similarities of murine mammary tumour models to human breast cancer, it is clear that species-mediated differences remain, such as variations in cell-intrinsic physiologies, growth kinetics, tumour size and site-specific metastatic seeding. Thus, various models are needed to start comprehending breast cancer heterogeneity that is present in the patient^{62, 107, 147}.

1.2.4.2 Lineage Tracing Studies

In order to permanently probe, track and dissect the functions of various cell types during organ development and malignant tumour progression in an unperturbed system *in vivo*, various *in situ* dual and multicolour lineage tracing systems have been generated¹⁶⁰.

To genetically label neurons to visualize synaptic circuits over time *in vivo*, Livet et al. have developed in 2007 the widely exploited “Brainbow” reporter system, in which different constructs with incompatible lox variants (Brainbow-1) and tandem, invertible DNA segments (Brainbow-2) allow a stochastic, combinatorial expression of several fluorescent proteins^{160, 161}. In Brainbow-1.0, incompatible lox variants can alternate,

enabling the switch from red fluorescent protein (RFP) expression to yellow fluorescent protein (YFP) or cyan fluorescent protein (CFP). Independent recombination of three transgene copies of the Brainbow-1.0 construct permits ten different colour combinations. In Brainbow-1.1, an additional, incompatible lox site has been inserted, which permits a fluorescent switch from orange to a stochastic recombination of CFP, YFP or RFP. Brainbow-2 makes use of tandem, invertible DNA segments. In Brainbow-2.0, a Cre-mediated recombination triggers the inversion of one DNA segment in opposite direction, switching gene expression from RFP to CFP. Brainbow-2.1 contains two tandem invertible DNA segments, which can be inverted and excised by Cre-activity, enabling the expression of four fluorescent proteins: CFP, green fluorescent protein (GFP), YFP and RFP¹⁶¹. In 2010, Snippert et al. have inserted the Brainbow-2.1 construct into the Rosa26 locus for the ubiquitous expression of the construct. The authors have utilized Lgr5-mediated Cre inversion and excision to fluorescently label stem-cell derived clones within the intestinal crypts in a random manner. Over time, these initially multicoloured clones progressed towards monoclonality, suggesting a life-long persistence of Lgr5⁺ stem cells by symmetric cell division¹⁶². Similarly, genetic lineage tracing and clonal analysis of K14-derived cells at distinct stages of squamous skin tumour progression unveils two proliferative cell compartments within papilloma's, a more persistent, stem cell-like population, and a slower cycling, differentiating compartment¹⁶³.

Over the years, several other *in situ* lineage tracing systems have been established to fate-track, visualize and functionally dissect different cell types during mammary gland development and tumour progression^{7, 164}. By using consecutive intravital imaging sessions on confetti-labelled mammary tumours (MMTV-PyMT; R26-Brainbow-2.1; R26-CreER^{T2}), Zomer et al. have visualized various clonal growth patterns indicative of cancer stem cell plasticity, such as disappearing clones, delayed growth onset, regression and continuous clonal growth¹⁶⁵. A comparable, thorough clonal restriction during malignant progression occurs in chemically induced or tumour-suppressor-inactivated confetti-labelled mammary tumours¹⁶⁶. This observation of clonal loss and dominance during carcinogenesis has also been confirmed by lineage tracing studies in other tumour models, such as pancreatic- and squamous skin carcinomas and sarcomas¹⁶⁷⁻¹⁶⁹. For example, by using a multicolour R26-Confetti-Cre reporter in an autochthonous mouse model of pancreatic cancer, Maddipati et al. have demonstrated a significant decrease in clonal heterogeneity during premalignant

progression¹⁶⁷. Furthermore, stage-specific lineage tracing of Confetti; K5-CreER^{T2} mice in a chemical-induced squamous carcinoma model has suggested the occurrence of a clonal sweep and clonal intermixing during the evolution of squamous skin carcinomas¹⁶⁸.

Besides these clonal analyses of primary tumour growth, the extent of tumour heterogeneity and clonal expansion can also be traced at the other steps of malignant progression, including invasion, dissemination into the circulation and seeding to secondary organs. In a study by Aceto et al., orthotopic transplantation experiments with GFP or mCherry-expressing MDA-MB231 cells has unveiled a plakoglobin-mediated polyclonal grouping of primary breast cancer cells as causative event for CTC cluster formation¹⁷⁰. In accordance with this, Cheung et al. have exploited multicolour lineage tracing combined with orthotopic transplantation of MMTV-PyMT; R26-Confetti organoids to prove the presence of a multicoloured, collective invasion and the existence of polyclonal, pulmonary breast cancer metastases^{171, 172}. Similar polyclonal metastases exist in pancreatic and squamous skin carcinomas^{167, 168}. Intriguingly, in the autochthonous mouse model of pancreatic cancer, the site-dependent selective pressures determine an either monoclonal or polyclonal outgrowth, with monoclonal metastases in liver and lung and polyclonal tumour cell expansion in peritoneum and diaphragm¹⁶⁷.

Throughout these exemplified, multifaceted studies, a universal pattern of clonal dynamics during tumour-evolution and growth becomes apparent, where clonal heterogeneity is followed by clonal dominance and subsequently polyclonal spread. This universality might – amongst other factors – be due to natural occurring tissue rearrangements, such as migration, proliferation and tissue deformation events, that foster fragmentation or merging of labelled cell clusters¹⁷³.

2. Aim of the Study

Breast cancer heterogeneity adversely affects a patient's prognosis, therapeutic response and clinical progression. In the past years, we have gained increasing knowledge about the extent of breast cancer heterogeneity and the treatment of the primary tumour. However, the functional contribution of tumour heterogeneity towards breast cancer progression is still obscure and standard of care treatments are rarely able to cure metastatic breast cancer patients. This demonstrates the urgent clinical need to functionally understand tumour heterogeneity during the progression of metastatic breast cancer to develop new targeted therapies.

This work aimed at identifying and characterizing clonal evolution and heterogeneity during metastatic breast cancer progression. To this end, we established a pipeline of multi-colour lineage tracing, in combination with laser capture microdissection and RNA-sequencing, to elucidate heterogeneity across and even within the stages of breast cancer progression. In collaboration with a pharmaceutical company, we have further assessed the therapeutic potential and biological effects of a novel, anti-proliferative inhibitor of focal adhesion kinase (FAK).

Specifically, I addressed the following aspects during my Ph.D. studies:

- The extent and functional contribution of clonal evolution and tumour heterogeneity to metastatic breast cancer progression.
- The therapeutic- and biological effects of the novel FAK-inhibitor BI 853520 *in vitro* and in preclinical mouse models of breast cancer.

3. Results

3.1 Multi-colour lineage tracing reveals intra-stage proliferative heterogeneity during mammary tumour progression

Stefanie Tiede^{1*}, Ravi Kiran Reddy Kalathur^{1,2}, Luca von Allmen¹, Barbara Maria Szczerba¹, Mathias Hess¹, Tatjana Vlajnic³, Benjamin Müller¹, James Canales Murillo¹, Nicola Aceto¹ and Gerhard Christofori^{1*}

¹ Department of Biomedicine, University of Basel, 4058 Basel, Switzerland

² Swiss Institute of Bioinformatics, 4053 Basel, Switzerland

³ Institute of Pathology, University Hospital Basel, 4056 Basel, Switzerland

* Co-corresponding authors

- in preparation -

3.1.1 Abstract

Despite major progress in breast cancer research, the process of tumour progression from a premalignant disease to metastatic outgrowth is still poorly understood and efficient therapeutic options to efficiently combat metastatic disease are still lacking. In particular, the functional contribution of distinct cancer cell clones within a tumour stage of a primary tumour to malignant tumour progression remains largely elusive. We have assessed clonal heterogeneity within individual primary tumours and metastasis and also during the distinct stages of malignant tumour progression using the Confetti lineage reporter in the background of the MMTV-PyMT mouse model of metastatic breast cancer. Comparative gene expression analysis of the clonal populations reveals a substantial level of heterogeneity across and also within the various stages of breast carcinogenesis. This intra-stage heterogeneity is manifested by differences in cell proliferation, where the fast-proliferating subclass is further enriched in oxidative phosphorylation and cell death. Notably, intra-stage proliferative heterogeneity is also found within lesions of ductal carcinoma *in situ* (DCIS) and in invasive carcinomas of breast cancer patients. However, cancer-related pathways underlying progression to malignancy seem to overrule the intrinsic proliferation differences; both slow and fast-proliferating subpopulations are found in all stages of malignant tumour progression and in metastasis. This new understanding of clonal variation and intra-stage tumour heterogeneity may have considerable impact on patient diagnosis and therapy response.

3.1.2 Introduction

In the past years, we have gained increasing knowledge about the extent of tumour heterogeneity. Tumour heterogeneity is known to be conferred by differences in various intrinsic parameters, e.g. the genetic profile and the transcriptomic landscape of a cell, the interplay between the genome, transcriptome and epigenome and the range of cell plasticity. Besides these cell-autonomous traits, also extrinsic factors in the tumour microenvironment affect tumour heterogeneity, including the tumour vasculature and tumour cell interactions with stromal and immune cells^{15, 32-36}. DNA- and gene expression profiling on bulk and single cells of primary tumours and metastatic lesions with and without chemotherapeutic regimens has revealed substantial differences in mutational profiles and large transcriptomic variations between clones of primary tumours and metastases^{18, 37-40, 44, 45, 51, 53-56, 64, 68, 121, 166, 168}.

This heterogeneity ultimately results in the formation of subclones with different biological capabilities, which will follow Darwin's principle of "survival of the fittest". Already in 1976, Nowell has described the idea of a clonal evolution, where subclones with a growth advantage in a certain ecosystem will persist, survive and eventually expand, whereas less fit subclones that are unable to compete against the others, will be extinct^{41, 42}. Multi-colour lineage tracing studies in breast, pancreatic, squamous skin carcinoma and sarcoma have identified the emergence of clonal dominance and loss during carcinogenesis and the existence of mono- or polyclonal metastasis^{165-169, 172}. Despite this increasing knowledge about the extent of breast cancer heterogeneity, the clonal diversification or, in other words, the topographical expansion of multiple cancer clones within a patient's primary tumour and secondary sites remains largely elusive. In particular, a comparative gene expression analysis of various clonal populations specifically within the stages of oncogenesis, such as hyperplasia, adenoma, carcinoma and metastasis, are still lacking. Moreover, the functional contributions of the various clonal subpopulations to malignant tumour progression and metastasis formation remain unknown, and defined gene signatures or biomarkers for diagnosis, prognosis and therapeutic outcome prediction are still to be identified.

Here, we have exploited the MMTV-PyMT mouse model of metastatic breast cancer expressing the Confetti lineage reporter and laser capture microscopy and next-generation RNA sequencing to interrogate clonal variation during breast cancer progression^{149, 161, 174}. Comparative analysis of gene expression profiles indicates a substantial level of heterogeneity across the stages and also within the specific stages of neoplastic progression. These intra-stage differences are mainly due to varying capacities in cell proliferation, survival and in oxidative phosphorylation and are dominated by a plethora of other cancer-related pathways to ensure malignant progression. The differences in the proliferation rate of clonal subpopulations within tumour stages has also been observed in patient samples, indicating the generality of the findings and their importance for the development of better diagnosis and therapy prediction.

3.1.3 Results

3.1.3.1 Confetti lineage tracing in murine mammary tumours

To assess clonal expansion and tumour heterogeneity during mammary gland carcinogenesis in the mouse, the MMTV-PyMT mouse model of metastatic breast

cancer was interbred with the Confetti lineage reporter mouse line (R26-CBW) and the K8-Cre-ER^{T2} driver mouse line^{149, 161, 174} (**Fig. 1a**). In the MMTV-PyMT mouse model, the MMTV promoter drives the mammary gland-specific expression of the PyMT oncogene, which results into the malignant transformation of the mammary gland epithelium, ultimately leading to the stepwise development of luminal mammary adenocarcinomas and pulmonary metastasis within 14 weeks of age^{149, 174}. The cytokeratin 8 promoter (K8, specific for luminal epithelial cells in the mammary gland) was used to drive the expression of a tamoxifen-inducible Cre-recombinase (K8-Cre-ER^{T2})^{3, 81}. To validate the choice of the K8 promoter as an adequate Cre driver, MMTV-PyMT primary tumours were first analysed for K8/K18 expression by immunofluorescence staining, demonstrating K8/18-expression in the different stages of murine mammary gland tumorigenesis, including hyperplasia, adenoma and carcinoma (**Supplementary Fig. 1a**).

To specifically trace K8-expressing luminal epithelial cells in the mammary gland, Cre recombinase was induced in five weeks old, female MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice by injecting one bolus of tamoxifen. Activation of Cre recombinase then led to recombination of the Confetti cassette and in the stable expression of one of the four Confetti reporter fluorescence proteins, membranous cyan fluorescent protein (mCFP), nuclear green fluorescent protein (nGFP), cytoplasmic yellow fluorescent protein (cYFP) and cytoplasmic red fluorescent protein (cRFP). The outgrowth of clonal cell populations was then analysed following 1, 4.5 and 8 weeks of tracing, or when the maximum primary tumour volumes had been reached, enabling the tracing of various breast cancer stages, such as hyperplasia, adenoma, carcinoma and pulmonary metastases (**Fig. 1b**). This lineage tracing strategy was used throughout this study.

Immunofluorescence staining of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} primary tumours for K8/18 expression verified the co-localization of recombined cells with K8/18 expression throughout tumour progression (**Fig. 1c**). Immunohistochemical and fluorescence analyses of vehicle or tamoxifen-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} primary tumours at indicated time points confirmed that the applied bolus of tamoxifen did not substantially affect primary tumour weight or malignant progression, tumour cell proliferation, apoptosis, immune cell infiltration, or tumour microvessel density (**Supplementary Fig. 1b**).

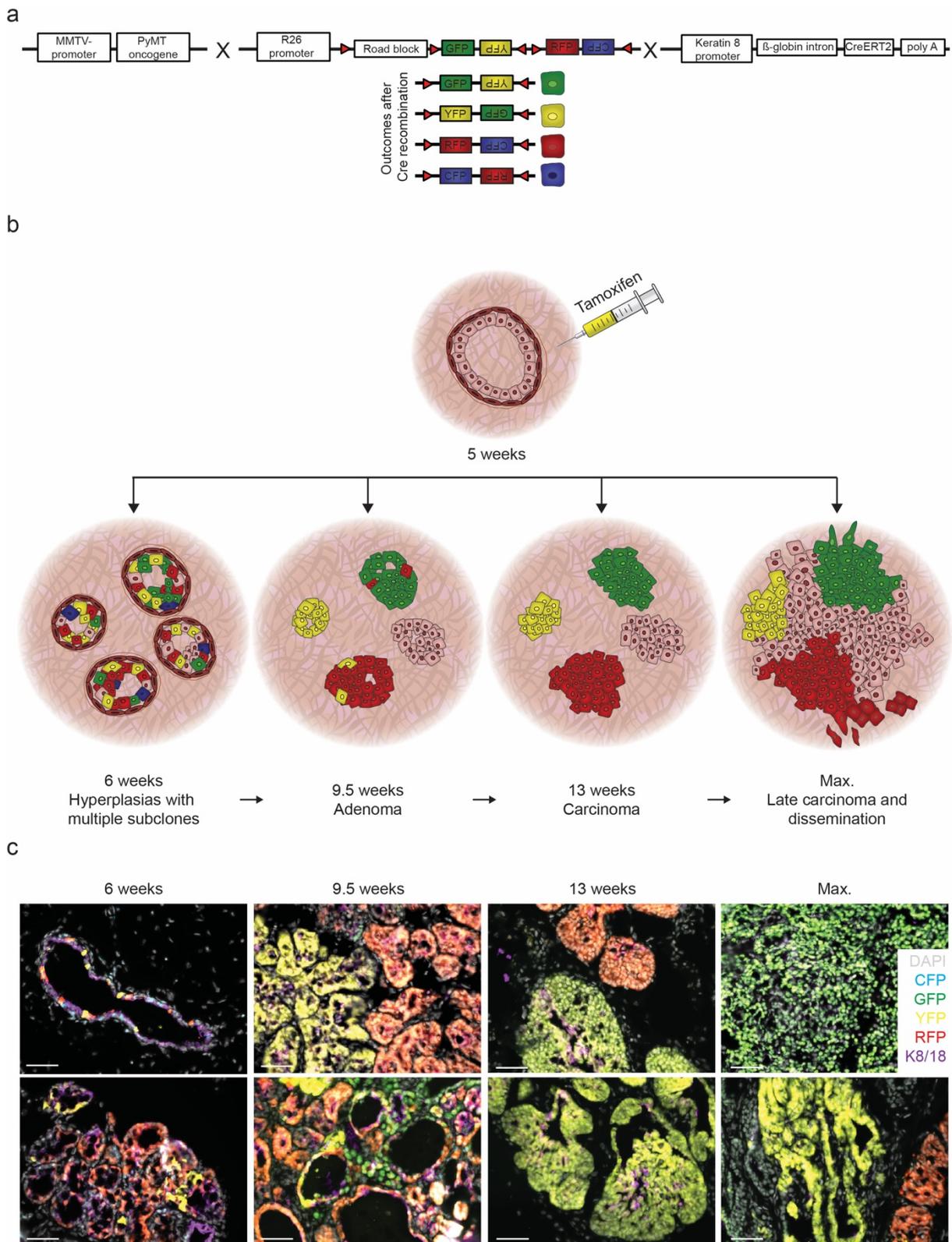


Figure 1. Tumour heterogeneity during the progression of metastatic breast cancer.

a, Lineage tracing strategy for a cyokeratin 8 (K8) promoter-driven colour randomization. In triple transgenic mice, the MMTV-promoter was used to drive the mammary gland-specific expression of the PyMT oncogene (MMTV-PyMT). The K8 promoter was used to drive the expression of tamoxifen-inducible Cre recombinase (K8-CreER^{T2}). Upon tamoxifen treatment, activation of Cre recombinase led to the stochastic expression of one of the four fluorescent proteins CFP, GFP, YFP and RFP (R26-CBW) in K8-expressing luminal epithelial cells and their progeny in the mammary gland.

b, Scheme depicting tamoxifen administration of five-weeks old MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice to allow lineage tracing for 1, 4.5, or 8 weeks or until the maximum tumour volume had been reached.

c, Microscopic images of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mammary gland tumours following 1, 4.5 and 8 weeks of tracing or until the maximum tumour volume had been reached, immunolabeled for K8 (magenta) and DAPI (grey) (n = 3 mice per time point). Scale bar, 50 μ m.

3.1.3.2 Tumour progression is associated with clonal dominance

To decipher tumour heterogeneity during the various stages of breast carcinogenesis, female MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice were colour-induced at five weeks of age, sacrificed at different time points (6, 9.5 and 13 weeks of age or once maximum tumour volume was reached), and clonal populations defined by the continuous and homogenous expression of one-colour fluorescence protein were analysed. FACS analysis and whole-mount imaging of mammary tumours of MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice one week after tamoxifen-induced colour randomization demonstrated single recombined luminal progenitor cells with a total labelling efficiency of around 3% in all four colours, but with a slight bias in frequency towards GFP and RFP (**Supplementary Fig. 2**).

Longitudinal imaging of MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice harbouring fluorescently labelled primary tumours throughout progression to malignancy was associated with the emergence of clonal populations of various stages in the primary tumour and culminated in a clonal dominance of large individual subclones in the later stages of tumour progression (**Fig. 2a**). This clonal restriction was manifested by a decrease in the number of clones and an increase in the area of clonal populations over time (**Fig. 2b**).

We next assessed the level of heterogeneity during tumour invasion, the initial step in the metastatic dissemination of tumour cells. Confocal microscopy imaging of thick cryosections of tamoxifen-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mammary gland tumours and subsequent 3D reconstructions revealed three modes of dissemination at the invasive fronts of primary tumours: disseminating single cells of all colours, monochromatic tumour cell clusters, and polychromatic tumour cell clusters (**Fig. 2c**).

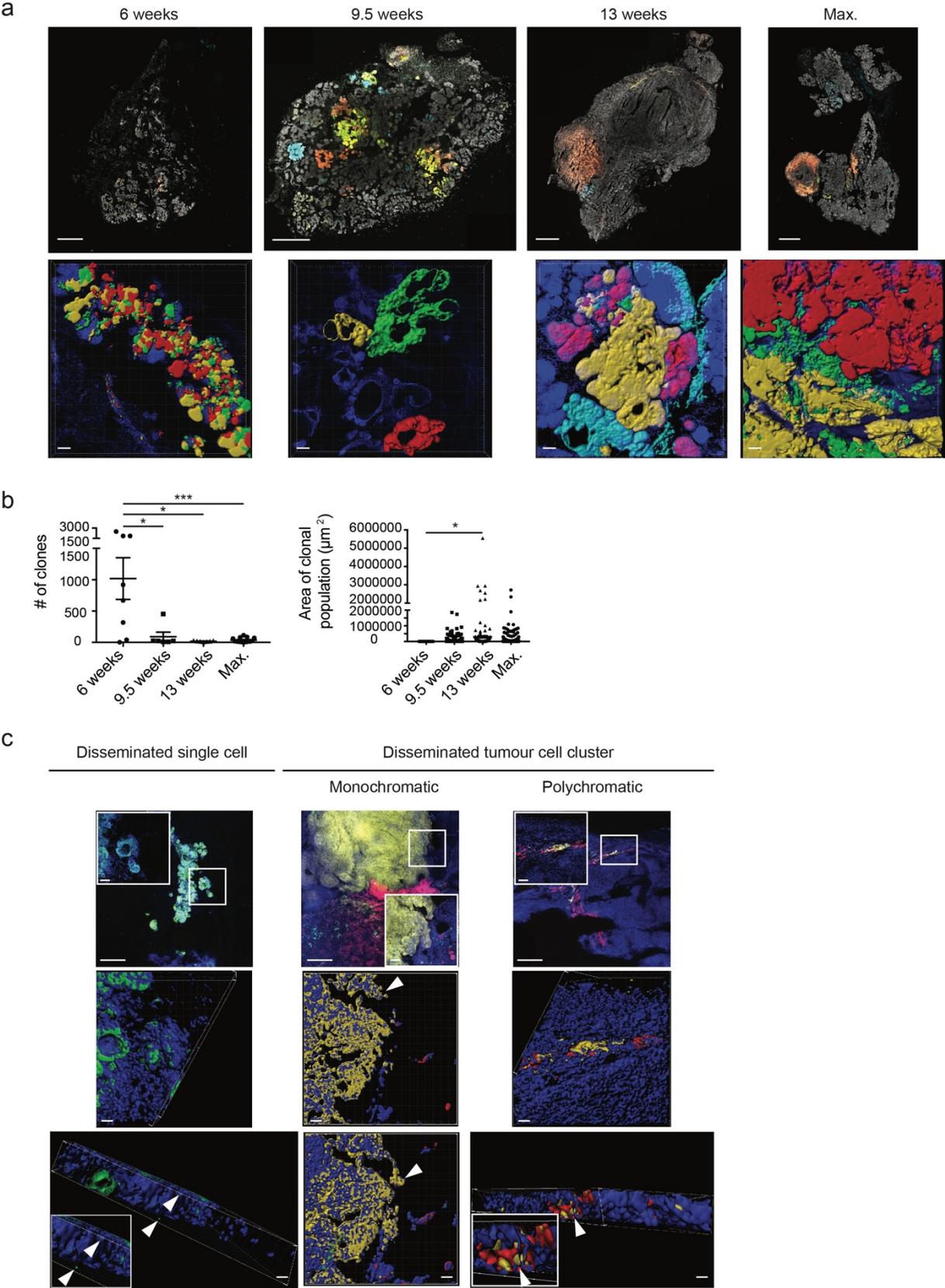


Figure 2. Temporal and spatial heterogeneity in primary breast tumours.
a, Representative fluorescence microscopy pictures (upper panels) and 3D-reconstructions (lower panels) of colour-induced mammary gland tumours of MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice at 6, 9.5, and 13 weeks of age and when the maximum tumour volume had been reached. DAPI, grey; n = 7-17 mice per time point. Scale bar, 1000 µm. 3D-reconstructions: DAPI, blue; n = 3-6 mice per time point. Scale bar, 100 µm.
b, Quantification of the numbers (left panel) and area sizes of clonal tumour cell populations (right panel) within tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice at 6, 9.5, and 13 weeks of age and when

the maximum tumour volume had been reached. n = 6-15 mice per time point. **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Student's t-test.

c, Representative fluorescence microscopy images (upper panels) and 3D-reconstructions (lower panels) of invasive and disseminating single cells and mono and polychromatic tumour cell clusters at the invasive front of late-stage tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice. DAPI, blue; n = 5. Scale bar, upper panels, 250 µm; rectangle, 50 µm. Scale bar, middle and lower panels, 30 µm.

3.1.3.3 Polyclonal metastatic spread

To address the functional effect of the observed tumour cell heterogeneity at the invasive front in the subsequent steps of the metastatic cascade, we next analysed circulating tumour cells (CTCs) in the bloodstream and pulmonary metastases in MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice bearing fluorescently labelled late-stage tumours. For CTC analysis, up to 1 ml of blood was isolated by cardiac puncture and CTCs were immediately captured with a Parsortix Cell Separation System. Microscopic analysis of cell separation cassettes identified CTCs as cells in the blood expressing one of the four fluorescent markers. We observed few single CTCs, yet more CTC clusters (harbouring on average 18 cells) (**Supplementary Fig. 3a**). The CTC clusters were either monochromatic or polychromatic and most of them were found to cluster with non-recombined tumour cells or other cell types, such as leukocytes (**Fig. 3a and Supplementary Fig. 3a**).

Once maximum tumour volume had been reached in colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice, serial sectioning of the whole lung followed by microscopic analysis detected approximately 30 metastatic lesions in the lungs of each mouse ranging from small microscopic to large macroscopic nodules and exhibiting mono and polychromatic cellular composition (**Fig. 3b, c**). Interestingly, polychromatic metastatic lesions were found larger but less frequent than their monochromatic counterparts in MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice and also in mice transplanted with a fragment of a triple-transgenic tumour (**Fig. 3d; Supplementary Fig. 3b**). The observation of larger polychromatic metastatic lesions might be the consequence of a survival benefit of polyclonal cell populations at the late stages of tumour progression, when metastatic outgrowth is maximal (**Fig. 3c,d; Supplementary Fig. 3c**). Analysis of the colour distribution of the clonal populations during progression to malignancy, including primary tumours, CTCs and pulmonary metastatic lesions, confirmed the extensive intra-tumoural heterogeneity with the concomitant preservation of colour, i.e. clonal dominance (**Fig. 3e**).

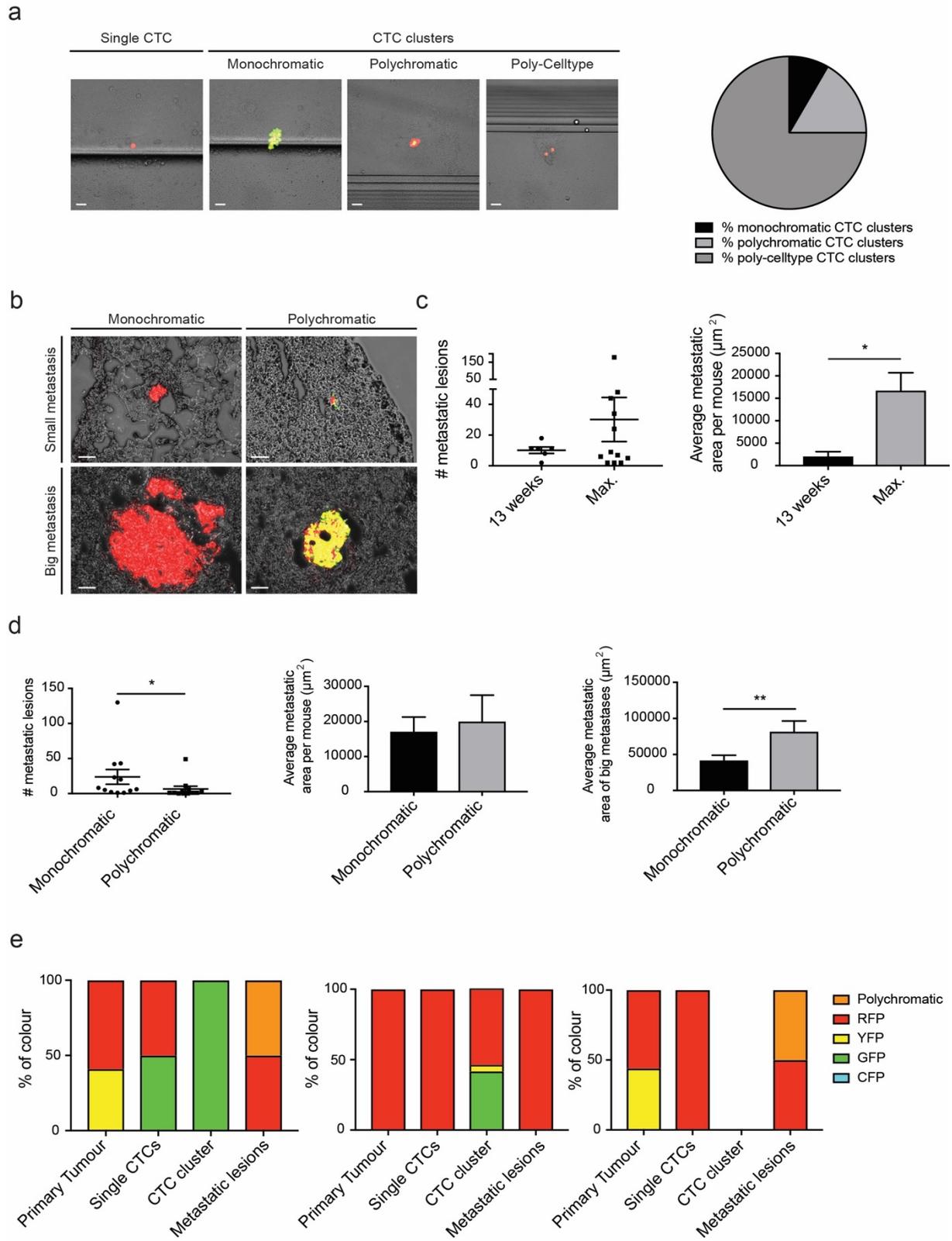


Figure 3. Polyclonal metastatic spread.

a, Representative fluorescence microscopy pictures of single circulating tumour cells (CTCs) and monochromatic, polychromatic and poly-celltype CTC clusters of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice at maximum tumour volume (left panels). A pie chart presents the percentage of CTC clusters being mono, polychromatic or poly-cell type (right panel). n = 3. Scale bar, 50 μm .

b, Representative fluorescence microscopy images of small or large and mono or polychromatic metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice. n = 31. Scale bar, 50 μm .

c, Quantification of the numbers and average metastatic areas of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice at 13 weeks of age and when the maximum tumour volume had been reached. n = 6-12 mice per time point. **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Student's t-test.

d, Quantification of the numbers and average area of monochromatic vs. polychromatic metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice when the maximum tumour volume had been reached (left and middle panel). Quantification of the average metastatic area of big monochromatic vs. polychromatic metastatic lesions is shown on the right. n = 12. **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, paired, two-tailed Student's t-test.

e, Quantification of colour dominance within primary tumours, CTCs, and metastatic lesions of three individual colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice when the maximum tumour volume had been reached. n = 3.

3.1.3.4 Substantial heterogeneity across and even within tumour stages

To investigate tumour heterogeneity during oncogenesis in colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice at the gene expression level, zonal patches of clonal populations of various stages within primary tumours as well as metastases were isolated by laser capture microdissection and subjected to RNA sequencing. Immunofluorescence microscopy analysis of histological sections of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice revealed the expression of GFP, YFP and RFP fluorescent markers in all stages of tumour progression, including metastatic lesions in the lungs (**Fig. 4a**). Since CFP expression was underrepresented at late-stage tumours, CFP-expressing clonal populations were excluded for further analysis. Clonal patches of GFP-, YFP- and RFP-expressing populations within primary tumours and metastatic lesions were then microdissected by laser capture microscopy and subjected to RNA sequencing (**Supplementary Fig. 4a**). Unsupervised hierarchical clustering of the top 500 most variable genes shared by hyperplasia, adenoma, carcinoma and metastatic lesions, suggested a substantial level of heterogeneity across the different stages of carcinogenesis (**Fig. 4b**). Interestingly, every stage with the exception of carcinoma harboured cell cycle-related terms in their top 10 reactome terms (**Fig. 4c, left panels, highlighted in bold**). Nevertheless, to adequately elucidate the extent and functional contribution of tumour heterogeneity to metastatic breast cancer progression, it is important to not only make a stage-wise comparison, but to specifically interrogate the level of heterogeneity between various subpopulations within the same tumour stages.

Strikingly, unsupervised hierarchical clustering of the top 500 most variable genes of each tumour stage revealed a tremendous heterogeneity in gene expression profiles and the identification of distinct subclasses within each tumour stage (**Fig. 4c, right panels, stage-specific subclasses are highlighted**). Intriguingly, when each tumour stage-specific subclass was compared to gene expression profiles of normal murine mammary glands, each tumour stage was found to harbour one subclass which was enriched for cell cycle-related reactome terms (**Fig. 4d**). Based on this

comparison, we classified each stage-specific subclass as fast or slow-proliferating for further analysis. A comparison of the pathways (top 10 KEGG terms) of all fast-proliferating subclasses of all tumour stages to all slow-proliferating subclasses validated their prominent vs. absent proliferative phenotype, respectively **(Supplementary Fig. 4b)**. Additionally, Integrated Motif Activity Response Analysis (ISMARA) on all subclasses pointed towards a stage-specific transcriptional regulation of the slow-proliferating vs. fast-proliferating phenotype, with the regulatory hubs highlighted in green and in red orchestrating the respective phenotypes **(Supplementary Fig. 4c)**.

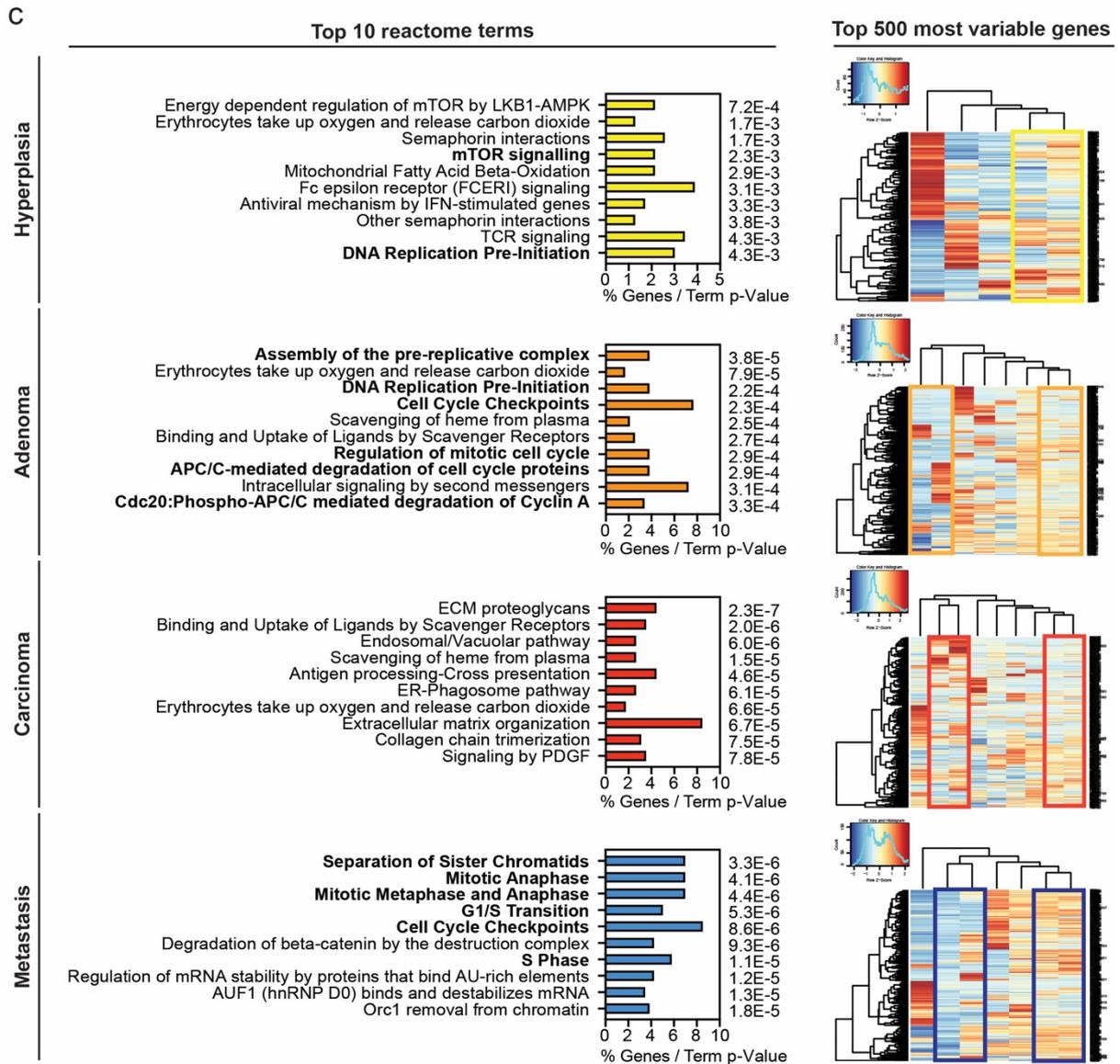
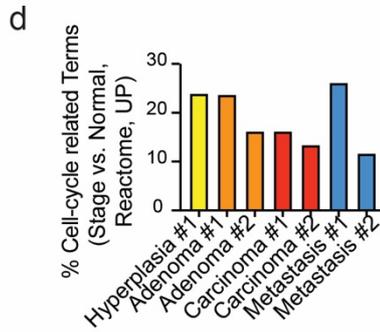
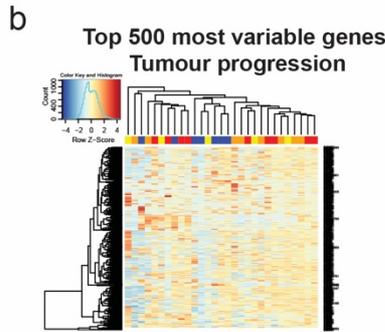
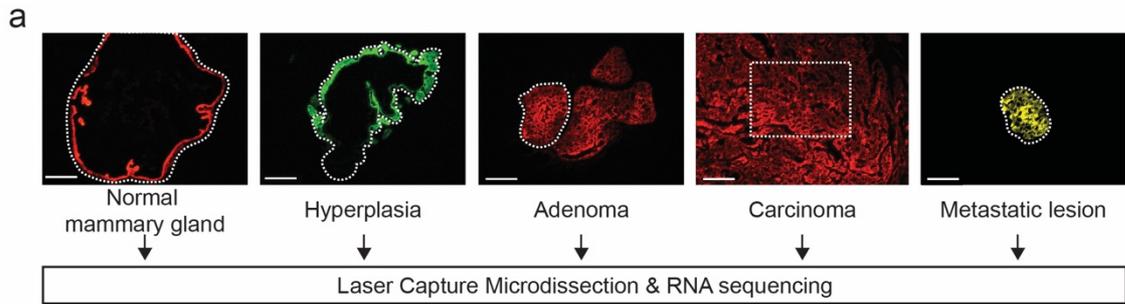


Figure 4. Intra-tumoural vs. intra-stage heterogeneity.

a, Outgrowth of clonal cell populations was analysed when the maximum tumour volume comprising all tumour stages (normal mammary gland, hyperplasia, adenoma, carcinoma, pulmonary metastases) had been reached in colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice. Zonal patches of clonal populations of the primary tumour as well as metastases were isolated by laser capture microdissection (encircled regions) and samples obtained were subjected to RNA sequencing. n = 3. Scale bar, 150 µm.

b, Unsupervised hierarchical clustering of the top 500 most variable genes across the various stages of mammary gland carcinoma progression. Hyperplasia, n = 5; adenoma, n = 8; carcinoma, n = 9; metastases, n = 7.

c, List of the top 10 reactome terms (percentage of expressed genes per term) of the top 500 most variable genes shared by all tumour stages are depicted in the left panels. Cell cycle-related terms are highlighted in bold. Unsupervised hierarchical clustering of the top 500 most variable genes within each stage (right panels) revealed stage-specific subclasses (highlighted in rectangles).

d, Bar graph representing the percentage of all upregulated cell-cycle related reactome terms in the comparisons of stage-specific subclasses vs. normal mammary gland. n = 2 for each tumour stage.

The following colour-code was used throughout the figure: hyperplasia, yellow; adenoma, orange; carcinoma, red; metastasis, blue.

3.1.3.5 Intra-tumour stage heterogeneity is due to differences in cell proliferation

In line with the slow-proliferating vs. fast-proliferating clonal populations in all tumour stages of MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice, gene set enrichment analysis of each tumour stage confirmed for the proliferative subclasses of each tumour stage a significant enrichment in genes implicated in the positive regulation of cell cycle (**Fig. 5a**). To assess whether the slow vs. fast-proliferating phenotypes can be observed in confetti-labelled tumours of MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice *in situ*, we performed immunofluorescence staining on primary tumours and lungs for phospho-histone H3 (pH3), a marker for proliferating cells. Within one tumour section, we detected several slow-proliferating clones in close proximity to fast-proliferating clones in hyperplasia, adenoma, carcinoma and metastases (**Fig. 5b and Supplementary Fig. 5a**). Quantification of pH3⁺-nuclei per clonal area verified the existence of slow vs. fast-proliferating clones within each tumour stage, which was accompanied by smaller vs. bigger clonal areas, respectively (**Fig. 5b and Supplementary Fig. 5b**). This correlation was not due to the selection of clone sizes for laser capture microdissection, since samples with identical size were isolated from one-colour clonal tumour patches for not introducing any confounder in the RNA sequencing analysis (**Supplementary Fig. 5c**). The finding of slow and fast-proliferating subclones throughout breast tumorigenesis could also be observed in monoclonal cell lines isolated from primary tumours and from pulmonary metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice and in mice transplanted with a fragment of a triple-transgenic tumour. Similar to primary tumours and metastatic lesions *in vivo*, the monoclonal cell lines displayed varying proliferation capacities, with some cell lines being slowly proliferating, whereas others harboured a high proliferative propensity (**Supplementary Fig. 5d**). The

differences in proliferation were found independent of the epithelial-to-mesenchymal transition state of the cell lines (data not shown). Notably, the intra-stage heterogeneity in tumour cell proliferation was also found in human breast cancer samples which had been stained for the proliferation marker Ki67 for clinical decision making. Also here, slow-proliferating and fast-proliferating lesions could be observed with high significance in ductal carcinomas *in situ* (DCIS) and invasive carcinomas, even though there was a considerable inter and intra-tumoural heterogeneity in the patient samples (**Fig. 5c and Supplementary Fig. 5e**).

Further gene-set enrichment analysis was performed to identify specific characteristics of the fast-proliferating subclasses. Interestingly, in addition to an induced cell cycle regulation the fast-proliferating subclasses were further enriched for oxidative phosphorylation and apoptosis. These computational results were validated by immunofluorescence staining of primary tumour sections of MMTV-PyMT mice for Ndufa2 and cleaved caspase 3 (ccaspase 3), markers for complex I of the oxidative phosphorylation system and for apoptosis, respectively (**Fig. 5d, e**). Here, fast-proliferative clonal populations also displayed an increase in Ndufa2 staining intensity. Furthermore, primary tumour sections of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice harboured areas, which were low or high in the apoptosis marker ccaspase3. However, the observed slow and fast-proliferating phenotypes were not attributable to differences in immune cell infiltrates and tumour microvessel density, as determined by immunofluorescence staining of primary tumour sections for CD45 and CD31, respectively (**Supplementary Fig. 5f, g**). Thus, throughout progression to malignancy, each tumour stage comprised not only inert but also highly dynamic clones, which were characterized by increased proliferation, oxidative phosphorylation, cell death and a larger clone size. In contrast, slow-proliferating populations of tumour cells did not apparently correlate with any specific gene expression profile or gene set enrichments.

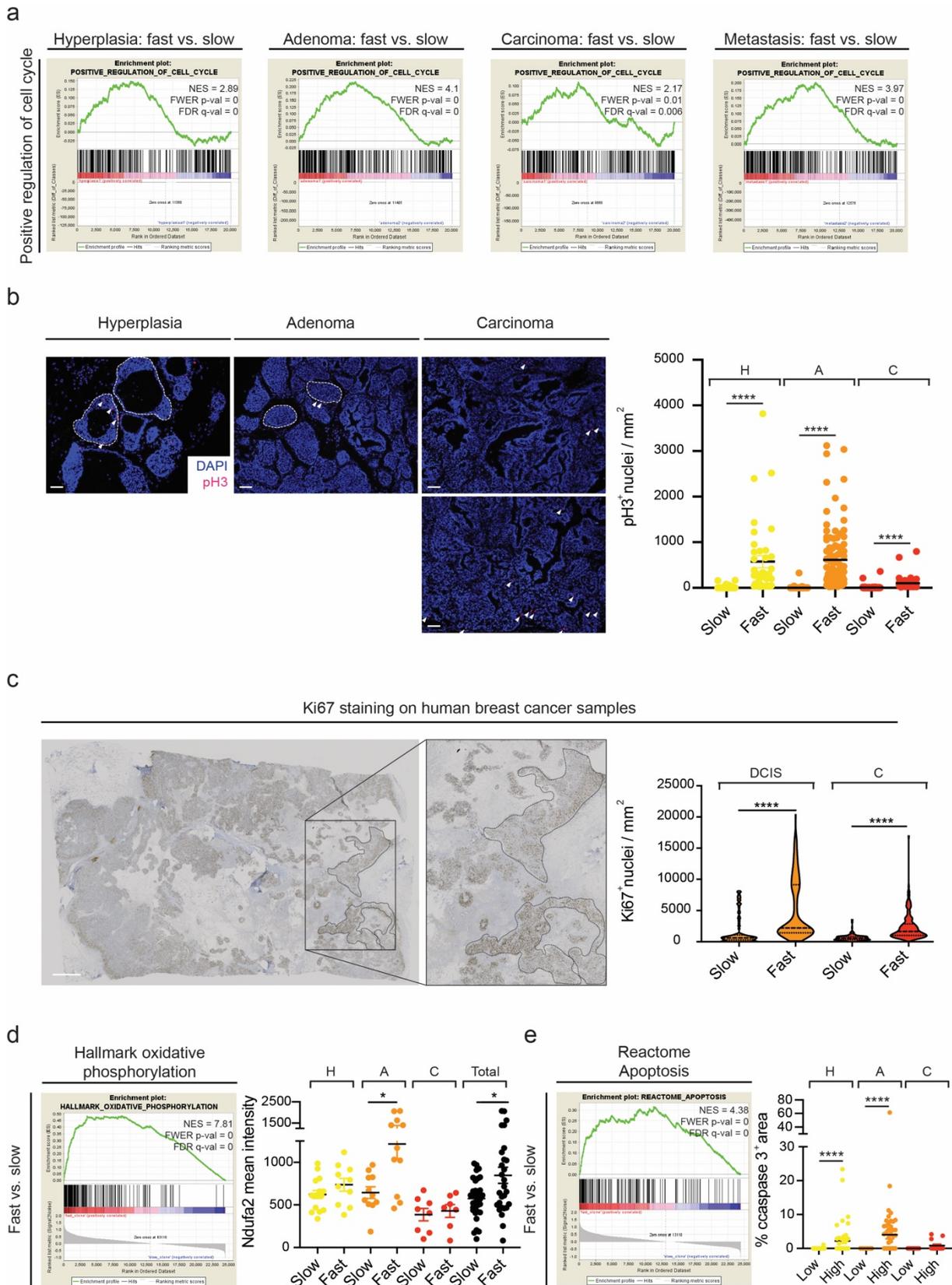


Figure 5. Intra-stage heterogeneity is due to differences in proliferation, oxidative phosphorylation and cell death.

a, Gene set enrichment analysis (GSEA) revealed an enrichment in gene sets involved in a positive regulation of cell cycle for the fast-proliferating subclasses of each tumour stage.

b, Representative immunofluorescence microscopy pictures of MMTV-PyMT; R26-CBW; K8-CreERT² mammary gland tumours, stained for phospho-histone H3 (pH3, magenta) and DAPI (blue) (left panels). Lesions below or above the average of pH3⁺ nuclei/mm² per tumour stage were classified as slow or fast-proliferating clonal populations, respectively. Differences in pH3⁺ nuclei per area of slow and fast-proliferating clonal populations within

each stage are depicted on the right. $n = 4$ mice. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, unpaired, two-tailed Mann-Whitney U test. Scale bar, 50 μm .

c, Representative light microscopy image of a human breast cancer biopsy immunolabeled for Ki67 indicated intra-stage heterogeneity due to proliferation differences (left panels). Lesions below or above the average of Ki67⁺ nuclei/ mm^2 per tumour stage per sample were classified as slow or fast-proliferating clonal populations, respectively. A quantification of Ki67⁺ nuclei per area of slow and fast-proliferating clonal populations within DCIS and invasive carcinoma lesions is shown on the right. $n = 9$ patients, with 1-4 sections each. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, unpaired, two-tailed Mann-Whitney U test. Scale bar, 2000 μm .

d, Gene-set enrichment analysis (GSEA) revealed for all fast-proliferating clonal lesions ($n = 8$) vs. all slow-proliferating lesions ones ($n = 9$) an enrichment in gene sets involved in oxidative phosphorylation (left panel). Lesions below or above the average of pH3⁺ nuclei/ mm^2 per tumour stage per MMTV-PyMT mouse were classified as slow or fast-proliferating clonal populations, respectively. A quantification of Ndufa2 mean intensity confirmed an increased capacity for oxidative phosphorylation in fast-proliferating clones (black labels). $n = 3$. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, unpaired, two-tailed Mann-Whitney U test.

e, Gene-set enrichment analysis (GSEA) for all fast-proliferating clonal populations ($n = 8$) vs. all slow-proliferating clones ($n = 9$) revealed an enrichment in gene sets involved in the regulation of apoptosis (left panel). A quantification of the percentage cleaved caspase 3⁺ area validates areas with low and high incidence of cell death in primary tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice. $n = 3$. **** $p < 0.0001$, *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, unpaired, two-tailed Mann-Whitney U test.

The following colour-code was used throughout the figure: hyperplasia, yellow; adenoma, orange; carcinoma, red; metastasis, blue.

3.1.3.6 Progression to malignancy is driven by unique pathways

To elucidate which unique pathways slow-proliferating and fast-proliferating tumour cell populations exploit to drive malignant progression, gene expression profiles of each stage-specific subclass were compared to that of normal mammary glands. The resulting differentially expressed genes and pathways (KEGG) of each subclass were then compared to each other to unravel unique up- and downregulated pathways for each stage-specific subclass (slow adenoma, slow carcinoma; slow carcinoma, slow metastasis; fast adenoma, fast carcinoma; fast carcinoma, fast metastasis) (**Fig. 6**). Interestingly, slow-proliferating clones progress from adenoma to carcinoma by upregulating oxidative phosphorylation and metabolic pathways in the adenoma stage and MAPK signalling pathways in the carcinoma stage and by downregulating ECM-receptor interactions and focal adhesions in the adenoma and carcinoma stage and cell cycle in the adenoma stage (upper left and middle columns). From invasive carcinoma to lung metastasis, slow-proliferating clonal populations further upregulate MAPK and Wnt signalling pathways in the carcinoma stage and oxidative phosphorylation in the metastases, which was accompanied by a downregulation of antigen processing and presentation in the carcinoma stage and protein digestion and absorption and ECM-receptor interactions in the metastases (lower left and middle columns). In contrast, fast-proliferating clones progress from adenoma to carcinoma by upregulating oxidative phosphorylation and pathways characteristic for Parkinson's, Huntington's and Alzheimer's disease in the adenoma and carcinoma stage, while cell adhesion molecules and focal adhesions were downregulated in the adenoma and carcinoma stage, respectively (upper left and right

columns). For metastatic spread, oxidative phosphorylation was further upregulated in the carcinoma stage and metastases with a concomitant upregulation of pathways implicated in Parkinson's (carcinoma and metastases), Huntington's and Alzheimer's disease (carcinoma stage). Conversely, complement and coagulation cascades and ABC transporters were downregulated in the carcinoma stage and focal adhesions were downregulated in the metastases (lower left and right columns). Thus, although each tumour stage harboured intrinsic differences in proliferation, this comparison exemplifies how both the slow and the fast-proliferating populations exploited various other pathways (including MAPK-and Wnt signalling and downregulation of cell adhesion molecules and focal adhesions) to progress from a hyperplastic lesion to distant metastasis. In other words, upregulated MAPK and Wnt signalling pathways were able to overrule the difference in proliferation rates within the various tumour stages in order to promote progression to malignancy.

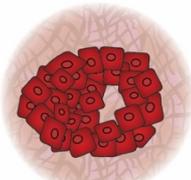
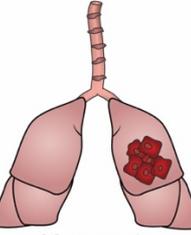
Stage	Slow	Fast
 Adenoma	UP Spliceosome Parkinson's disease Alzheimer's disease Oxidative phosphorylation Metabolic pathways	UP Oxidative phosphorylation Parkinson's disease Huntington's disease Alzheimer's disease Proteasome
	DOWN ECM-receptor interaction Pathways in cancer Cell cycle Focal adhesion ABC transporters	DOWN Staphylococcus aureus infection Chemokine signaling pathway Fc gamma R-mediated phagocytosis Cell adhesion molecules (CAMs) Focal adhesion
 Carcinoma	UP Autoimmune thyroid disease Histidine metabolism Thyroid cancer Viral myocarditis MAPK signaling pathway	UP Ribosome Parkinson's disease Oxidative phosphorylation Huntington's disease Alzheimer's disease
	DOWN Focal adhesion Staphylococcus aureus infection ECM-receptor interaction Complement and coagulation cascades Protein digestion and absorption	DOWN ABC transporters Long-term depression Progesterone-mediated oocyte maturation Focal adhesion Circadian rhythm
 Carcinoma	UP MAPK signaling pathway Taste transduction Autoimmune thyroid disease Viral myocarditis Wnt signaling pathway	UP Ribosome Oxidative phosphorylation Parkinson's disease Alzheimer's disease Huntington's disease
	DOWN Antigen processing and presentation Cytokine-cytokine receptor interaction Leukocyte transendothelial migration B cell receptor signaling pathway Rheumatoid arthritis	DOWN Complement and coagulation cascades Retinol metabolism Drug metabolism - cytochrome P450 ABC transporters Salivary secretion
 Metastasis	UP Huntington's disease Parkinson's disease Oxidative phosphorylation Alzheimer's disease Protein export	UP Protein export Oxidative phosphorylation RNA transport Protein processing in endoplasmic reticulum Parkinson's disease
	DOWN Protein digestion and absorption ECM-receptor interaction Dilated cardiomyopathy Staphylococcus aureus infection Osteoclast differentiation	DOWN Protein digestion and absorption Staphylococcus aureus infection ECM-receptor interaction Focal adhesion Osteoclast differentiation

Figure 6. Unique pathways drive tumour progression.

Gene expression profiles of stage-specific subclasses were compared to normal mammary glands. The resulting differentially expressed genes and KEGG terms of each subclass were compared to each other (left and middle column: slow adenoma vs. slow carcinoma, slow carcinoma vs. slow metastases; left and right column: fast adenoma vs. fast carcinoma, fast carcinoma vs. fast metastases) to find unique up- and downregulated pathways for each stage-specific subclass (slow adenoma, slow carcinoma; slow carcinoma, slow metastasis; fast adenoma, fast carcinoma; fast carcinoma, fast metastasis) that drives malignant progression. The top 5 KEGG terms are depicted (ranked by p-values < 0.05).

3.1.4 Discussion

Using MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice for multicolour lineage tracing in combination with laser capture microdissection and RNA sequencing^{3, 81, 128-132, 149, 161, 174}, we have assessed clonal evolution and tumour heterogeneity not only across the stages of neoplastic progression but specifically also within various tumour stages in a transgenic mouse model of metastatic breast cancer (**Supplementary Fig. 6**). Longitudinal imaging of Confetti-traced tumours identified clonal populations of various stages in primary tumours with the emergence of clonal dominance. Yet, metastatic spread appears associated with polyclonality at the invasive front, in circulating tumour cells (CTCs) in the blood stream, as well as in metastatic lesions in the lungs. Comparative analyses of gene expression profiles of various clonal primary tumour stages and metastasis indicate a substantial level of heterogeneity between the different oncogenic stages and, even more intriguing, within the same tumour stages. Unexpectedly, this intra-stage tumour heterogeneity is due to differences in tumour cell proliferation: each tumour stage harbours slow-proliferating, inert subpopulations as well as fast-proliferating, highly dynamic clones. These fast-proliferating clonal populations are further characterized by an enrichment in the expression of genes active in oxidative phosphorylation and programmed cell death and by increased clonal overgrowth. Interestingly, even though these intrinsic proliferation differences exist, slow and fast-proliferating subclasses across the various breast cancer stages are able to exploit various other cancer-related pathways when progressing to malignancy.

In line with our findings, others have previously used Confetti lineage tracing systems in breast, pancreatic, squamous skin carcinomas and sarcomas to report a similar replacement of a high polyclonality in early tumour lesions by few overgrowing clonal populations in later tumour stages, which is accompanied by a loss of clonal diversity and clonal restriction during tumorigenesis¹⁶⁵⁻¹⁶⁹. Furthermore, ours and other studies reveal the existence of polyclonal metastases in autochthonous and in tumour transplantation models^{167, 168, 172}. However, whether clonal expansion and polyclonal

spread is due to acquired mutations and/or to beneficial clonal cooperativity between various clonal subpopulations remains open¹⁰⁹. The presence of polychromatic and poly-cell type CTC clusters identified by us and others^{59, 169, 170, 172}, as well as our observation of larger polychromatic metastatic lesions, suggest a survival benefit for polyclonal lesions. However, the molecular contribution of tumour heterogeneity to malignant tumour progression and to metastasis formation is still not fully understood.

Here, we have combined laser capture microdissection, RNA sequencing and bioinformatics analyses to elucidate the transcriptomic profiles of a small number of cells of clonal populations within the various stages of tumour progression in the primary tumour and in pulmonary metastatic lesions. Consistent with previous findings, we observe a tremendous transcriptomic diversity between the different stages of carcinogenesis^{155, 166, 168}. As expected, subclonal diversification and heterogeneity is high between the longitudinal stages of tumour progression. However, to adequately assess the functional contribution of tumour heterogeneity to breast cancer progression, it is important to additionally interrogate the extent of heterogeneity between subpopulations within the same tumour stage. Indeed, we find an intriguing heterogeneity between the clonal subpopulations within individual stages of malignant tumour progression. This intra-stage heterogeneity is mainly based on differences in tumour cell proliferation, where the fast-proliferating subclass is further enriched in oxidative phosphorylation and programmed cell death. RNA sequencing of whole tumour lysates has previously also reported a role of cell cycle-related genes in the progression of MMTV-PyMT primary tumours¹⁵⁵. Other studies have found slow-cycling subpopulations in melanoma and other cancer types to be essential for continuous tumour growth, invasion, metastatic progression and drug resistance⁹¹⁻⁹⁴. Furthermore, the proliferative phenotype has been found predominantly accompanied by oxidative metabolism^{175, 176}. Similarly, our study reveals a tight association of the proliferative phenotype with a high oxidative capacity and larger clonal outgrowth, yet also with an enrichment in tumour cell death.

Noteworthy, the finding of slow and fast-proliferating subclones in the various stages of mammary gland tumorigenesis is translatable to human breast cancer patients. Here, Ki67 analysis confirms considerable inter and intra-tumour stage heterogeneity in human breast cancer punch biopsies. Intriguingly, low or high zonal expression of the proliferation marker Ki67 is found within lesions of both DCIS and of invasive carcinomas of patients. These data may be of great importance for patient

diagnosis. Histological parameters, including the cell-of-origin and invasion beyond the basement membrane, in combination with the expression of hormone receptors and further genomic and transcriptomic molecular profiling diagnostically determine distinct breast cancer subtypes. The temporal and spatial clonal variation and intra-stage proliferation differences specified here further add to the above-mentioned complexity, and we postulate that this new understanding of an intra-stage heterogeneity should be taken into consideration for diagnosis. More importantly, the observed proliferation differences within each tumour stage might enhance a tumour's robustness to anti-cancer therapies. For example, conventional medication for triple-negative breast cancers include chemotherapeutic reagents to specifically target fast-proliferating cells. Whether the slow-proliferating subpopulations of tumour cells are then responsible for the occurrence of therapy resistance and clinical progression, thus warrants further investigations.

3.1.5 Methods

Mice and Confetti labelling

MMTV-PyMT, R26-Confetti (here: R26-CBW) and K8-CreER^{T2} mice were previously described^{3, 81, 149, 161, 174}. To generate MMTV-PyMT; R26-CBW; K8-CreER^{T2} triple-transgenic mice, male MMTV-PyMT; R26-CBW mice were crossed to female K8-CreER^{T2} mice. To stochastically express one of the four Confetti reporter fluorescent proteins, female MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice of five weeks of age were induced intraperitoneally with five mg tamoxifen (Sigma-Aldrich) dissolved in 200 μ l sunflower oil (Sigma-Aldrich) containing 10% EtOH (Scharlau). These animal experiments have been performed under approval by the Swiss Federal Veterinary Office and the Cantonal Veterinary Office, Basel-Stadt, Switzerland (permit numbers 1878, 1907 and 1908).

Tumour-piece transplantation and primary tumour removal

Non-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} tumour pieces were subcutaneously transplanted into eight weeks old, female, isoflurane-anaesthetized NSG mice, the wounds were closed by metal clips (7mm; Alzet Wound Closure System) and the animals were placed under infrared light until walking. For the next two days, meloxicam was administered (5 mg/kg body weight per day) for pain relief. Once tumours became palpable, mice were intraperitoneally injected with vehicle or

five mg tamoxifen dissolved in 200 µl sunflower oil (containing 1:10 100% ethanol). Formed tumours were routinely monitored. Tumours with a tumour volume of 500 mm³ were surgically removed from isoflurane-anaesthetized mice, the wounds were closed by metal clips and the animals were kept warm upon recovery. For the next two days, meloxicam was administered (5 mg/kg body weight per day) for pain relief. At defined termination criteria, mice were sacrificed, and their lungs were further processed for histopathological analysis. All animal experiments were approved by and performed according to the guidelines and legislation of the Swiss Federal Veterinary Office and the Cantonal Veterinary Office, Basel-Stadt, Switzerland (permit number 1908).

Tissue harvesting and capture of circulating tumour cells (CTCs)

Primary tumours in the thoracic and abdominal mammary fat pad and lungs were isolated and tissue processing was performed as previously reported. Up to 1 ml of blood was collected by cardiac puncture and processed immediately on the Parsortix microfluidic device using Cell Separation Cassettes (GEN3D6.5, ANGLE). The number of single circulating tumour cells (CTCs) and CTC clusters was determined while cells were captured in the cassette.

Immunofluorescence staining on frozen tumour sections

For the preparation of cryostat sections, organs were fixed in 4% paraformaldehyde for two hours at 4°C, cryopreserved overnight at 4°C in 20% sucrose in PBS (Sigma-Aldrich) and embedded in OCT freezing matrix (Tissue Tek). For immunofluorescence staining, 7 µm thick cryosections of tumour- or lung samples were dried for 30 min and rehydrated in PBS. Following a 20 minutes' permeabilization in 0.1% Triton X-100 (Sigma-Aldrich) in PBS, slides were washed in PBS, blocked in 5% goat serum diluted in PBS (blocking buffer) for 1 h at room temperature and incubated overnight at 4 °C with primary antibodies diluted in blocking buffer. Following primary antibodies were used: guinea pig anti-mouse cytokeratin 8 + 18, Fitzgerald Industries, 20R-CP004; rabbit anti-mouse cytokeratin 14, Thermo Scientific, RB-9020-P0; rabbit anti-mouse phospho Histone H3 (pH3), Millipore, 06-570; rabbit anti-mouse cleaved caspase 3 (ccaspase 3), Cell Signaling, 9664, clone 5A1E; rat anti-mouse CD45, BD Pharmingen, 550539 and rat anti-mouse CD31, BD Pharmingen, 550274; rabbit anti-mouse Ndufa2, biorbyt, orb221658. On the next day, sections were incubated at room temperature for one hour with Alexa Fluor 488-labeled, 568-labeled or 647-labeled secondary

antibodies (Invitrogen, 1:400) diluted in blocking buffer and washed in PBS. Nuclei were counterstained with 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich, 1:5000) for 10 min and slides were washed in PBS and mounted in DAKO fluorescence mounting medium (DAKO). Images were analysed using the ImageJ software (ImageJ, Wayne Rasband, National Institutes of Health, USA) and the Zeiss microscopy software ZEN 2.6 pro. To determine whether the applied bolus of tamoxifen affects malignant progression, the number of pH3⁺-, ccaspase 3⁺-, CD45⁺-cells and CD31⁺ vessels were counted per primary tumour area (mm²) throughout tumorigenesis. For clonal analysis, the number of pH3⁺ nuclei per mm², the mean fluorescence intensity of Ndufa2, the percentage of ccaspase 3⁺ area and the number of CD45⁺-cells and CD31⁺ vessels were analysed per clonal lesion.

Quantification of Confetti labelling

To quantify Confetti labelling, cryosections (7 µm thick) of mammary gland tumours were rehydrated in PBS, incubated with DAPI (Sigma-Aldrich, 1:5000) at room temperature for 10 minutes and washed with PBS. The whole mammary gland section was imaged by using a 10X objective of the Zeiss Axio Imager Z1 microscope equipped with filters distinguishing CFP (474/23), GFP (510/20), YFP (535/22) and RFP (620/60). Total recombination efficiency (% fluorescence area/total tumour area) and the number and area of clonal populations (µm²) was determined by using the ZEN software ZEN 2.6 pro.

3 dimensional tissue analysis

80 µm thick cryosections were dried for two hours at room temperature, rehydrated in PBS and incubated with DAPI (Sigma-Aldrich, 1:1000) at room temperature for 25 minutes. Slides were then washed with PBS, mounted using DAKO fluorescence mounting medium (DAKO) and imaged using a Leica SP5 confocal microscope equipped with 405 nm diode-, 458 nm Argon- and 543 nm He-Ne lasers. Following emission signals were collected: DAPI (415-450), CFP (466-495), GFP (498-509), YFP (521-560), RFP (590-650). All images were acquired with a 10X and 40X objective using a Z-step size of (total Z-stack) and processed using ImageJ software. Representative 3D reconstructions for Confetti labelling of mammary gland tumours were created by Imaris software.

Flow Cytometry

For FACS analysis, mammary gland tumour pieces were enzymatically digested by collagenase D (Roche) and trypsin (Sigma-Aldrich) for 1.5 hours at 37 °C. Digested tumour pieces were washed, treated for 3-5 minutes with DNase (Sigma-Aldrich, 1:100) and washed twice. Cell pellets were resuspended in FACS buffer (1 mM EDTA in PBS, 2% FCS) and filtered through a 40 µm filter before analysis. Following FACS-settings were used: CFP (405; 450/50), GFP (488; 495 LP), YFP (488; 543/22), RFP (561; 582/15).

Whole mount imaging of mammary gland tumours

Mammary gland tumours were mounted on object glasses (Menzel) and imaged by using a Leica SP5 confocal microscope, equipped with 405 nm diode-, 458 nm Argon- and 543 nm He-Ne lasers. Following emission signals were collected: DAPI (415-450), CFP (466-495), GFP (498-509), YFP (521-560), RFP (590-650). All images were acquired with a 10X and 20X dry objective using a Z-step size of (total Z-stack) and processed using ImageJ software.

Serial sections lungs

For metastasis quantifications, OCT-embedded lungs were serially cut and pictures of mono- and polychromatic metastases were taken by a 10X (HC PL Fluotar) and 20X (HCX PL Fluotar L) objective of a Leica DMI 4000 microscope equipped with filters distinguishing CFP (460-500 nm), GFP (500-550 nm), YFP (520-550 nm) and RFP (604-644 nm).

Laser capture microdissection and transcriptome sequencing

The outgrowth of clonal cell populations has been analysed when the maximum allowed tumour volume comprising all stages (normal, hyperplasia, adenoma, carcinoma, pulmonary metastases) had been reached in colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice. By combining the recently published Geo-seq protocol using SmartSeq2 for laser capture microdissected samples with a recent report for parallel DNA and RNA sequencing from single cells, representative cDNA libraries of laser captured regions from fixed tumour sections were generated¹²⁸⁻¹³². In brief, 7 µm thick cryosections were mounted onto Menzel Superfrost Plus slides and placed in decreasing and increasing ethanol solutions. Zonal patches of clonal populations of

various stages in the primary tumour as well as monochromatic metastases were microdissected by using a Zeiss Laser Capture Microscope (energy: 49, focus: 67, delta-energy: 15, delta-focus: 5; CFP (474/23), GFP (510/20), YFP (535/22) and RFP (620/60)) in Auto LPC mode. Clonal populations of normal mammary glands were pooled to obtain a sufficient sample size. Collected laser capture microdissected samples were first lysed at 37 °C for 30 minutes by using a digestion buffer containing proteinase K (150 mM NaCl, 100 mM Tris pH 7.5, 0.5% Igepal, 0.5 µg/µl proteinase K) to reverse the cross-linking of formalin fixation and for further 30 minutes at 42 °C by using a 4M GuSCN solution (Thermo Fisher). Biotinylated oligo-dT30VN beads were used to discard gDNA in the supernatant and specifically start the reverse-transcriptase reaction and cDNA preamplification on poly-adenylated mRNA (exception to GeoSeq protocol: dNTP mix contained 10mM each of dATP, dCTP, dGTP, dTTP and SmartSeq2 ISPCR oligos were used, 5'-AAGCAGTGGTATCAACGCAGAGT-3'). cDNA quality was assessed by a Qubit Fluorometer with a dsDNA HS assay kit (Thermo Fisher), an Agilent high-sensitivity DNA chip for cDNA size distribution and quantitative real-time PCR to check for the expression of housekeeping genes and the absence of gDNA. Laser capture microdissected samples with best quality were subjected to RNA sequencing.

RNA sequencing

For library preparation, 500 pg cDNA was amplified using the Nextera XT DNA library preparation kit (Illumina). 12 PCR cycles were carried out, DNA was purified using AMPure XP beads in a 0.8:1 ratio and cDNA was sequenced on a NextSeq 500 platform (Illumina) with a 75-bp single-read mode (Index1:8, Index2: 8).

Analysis of RNA-sequencing data

Single-end RNA-seq reads (81-mers) were mapped to the mouse genome assembly, version mm10, with RNA-STAR¹⁷⁷, with default parameters except for allowing only unique hits to the genome (outFilterMultimapNmax=1) and filtering reads without evidence in spliced junction table (outFilterType="BySJout"). Expression levels per gene (counts over exons) for the RefSeq mRNA coordinates from UCSC (genome.ucsc.edu, downloaded in December 2015) were quantified using qCount function in QuasR package (version 1.12.0)¹⁷⁸. The samples with less than 10,000 genes expressed were removed from further analysis. The top 500 highly variable

genes across the samples were obtained by calculating variance in R on counts per million (CPM) values. The differentially expressed genes were identified using the edgeR package (version 3.14.0)¹⁷⁹. Genes with p-value \leq than 0.05 and minimum log2 fold change of +/-0.58 were considered as differentially regulated and were used for downstream functional and pathway enrichment analysis.

Functional enrichment analysis

Functional enrichment analysis of differentially expressed genes for biological processes or pathways were performed in R using several publicly available Bioconductor resources, including org.Ms.eg.db (version 3.6.0), GO.db (version 3.6.0), GOstats (version 2.48.0), KEGG.db (version 3.2.3) and ReactomePA (version 1.26.0)¹⁸⁰. Significance of each biological process or pathway identified was calculated using the hypergeometric test (equivalent to Fisher's exact test) and those processes and pathways with p-values \leq 0.05 were considered statistically significant.

Genome-wide predictions of regulatory sites

The Integrated System for Motif Activity Response Analysis (ISMARA) was used to identify binding sites recognized by transcription factors (TFs) and microRNAs (miRNAs)¹⁸¹. The fastq files of the samples obtained after sequencing were directly uploaded. Gained regulatory motifs of both TFs and miRNAs were sorted based on Z-score.

Gene-set enrichment analysis

The Gene Set Enrichment Analysis (GSEA) analysis was performed using JAVA application from the Broad Institute version 3.0¹⁸². The gene sets used for the analysis were derived from Molecular Signatures Database (MSigDB) version 6.2¹⁸³, gene ontology annotations and pathways were obtained from Gene Ontology¹⁸⁴ and Reactome databases¹⁸⁵.

Isolation of monoclonal cell lines

To establish monoclonal cell lines from mammary gland tumours and pulmonary metastases, isolated mammary gland tumours and metastatic lesions were minced into small pieces and were enzymatically digested by collagenase D (Roche) and trypsin (Sigma-Aldrich) for 1.5 hours at 37 °C. Digested pieces were washed, treated

for 3-5 minutes with DNase (Sigma-Aldrich, 1:100) and washed twice. Cell pellets were resuspended in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich), supplemented with fetal calf serum (FCS, Sigma-Aldrich, 10%), glutamine (Sigma-Aldrich, 2 mM), penicillin (Sigma-Aldrich, 100 U) and streptomycin (Sigma-Aldrich, 0.2 mg/l) and incubated in a humidified atmosphere of 37 °C and 5% CO₂. To ensure monoclonality, established cell lines were resuspended in FACS buffer (1 mM EDTA in PBS, 2% FCS), filtered and sorted by FACS for their respective colour.

Human breast cancer analysis

Paraffin-embedded, human breast cancer samples were routinely stained for Ki67 and cross-sections were imaged with a 20X objective of the Zeiss Axio Imager Z1 microscope. Lesions of ductal carcinoma *in situ* (DCIS) and invasive carcinoma were counted for the amount of Ki67⁺ nuclei as well as their respective area by using the ZEN software ZEN 2.6 pro.

Statistical evaluation

Graphs and statistical analyses were generated using GraphPad Prism software Version 8.2.0. To determine statistical significance, the (un)paired, two-tailed Student's t test or Mann-Whitney U test were applied (**** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05). Quantitative data were depicted as means ± SEM.

Competing interests

The authors declare no competing interests.

Data availability

The datasets generated and analyzed within the current study are deposited at Gene Expression Omnibus (GEO; currently in uploading process).

Author contributions

S. T. wrote the manuscript, designed the experiments and conducted most of the *in vivo* experiments, analyses, laser capture microdissections and validation studies, with support from L. v. A., M. H., B. M. and J. C. M.

R. K. R. K. performed all bioinformatics tools, including analysis of RNA-sequencing data, functional enrichment analysis, genome-wide predictions of regulatory sites and gene-set enrichment analysis.

B. S. and N. A. performed blood sample preparation on the Parsortix microfluidic device and gave technical advice.

T. V. provided human breast cancer samples and pathological input.

G. C. conceived and designed the experiments and wrote the manuscript with contributions from other co-authors.

Acknowledgements

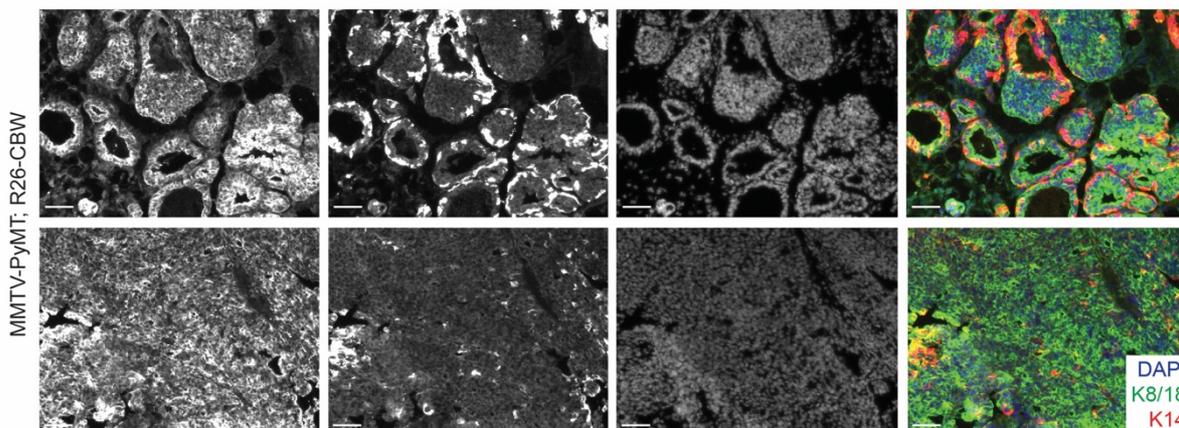
We are grateful to V. Taylor and M. Bentires-Alj (DBM, University of Basel) for providing R26-CBW-, and K8-CreER^{T2} mice, respectively. We thank E. Panoussis, I. Galm, U. Schmieder and T. Bürglin (DBM, University of Basel) for their technical assistance, P. Lorentz and the DBM microscopy core facility for expert support with microscopy (University of Basel), C. Beisel, E. Burcklen, K. Eschbach and the Genomics Facility Basel (University of Basel) for RNA sequencing and R. Ivanek for bioinformatic support. RNA sequencing calculations were conducted at sciCORE scientific computing core facility, University of Basel.

Funding

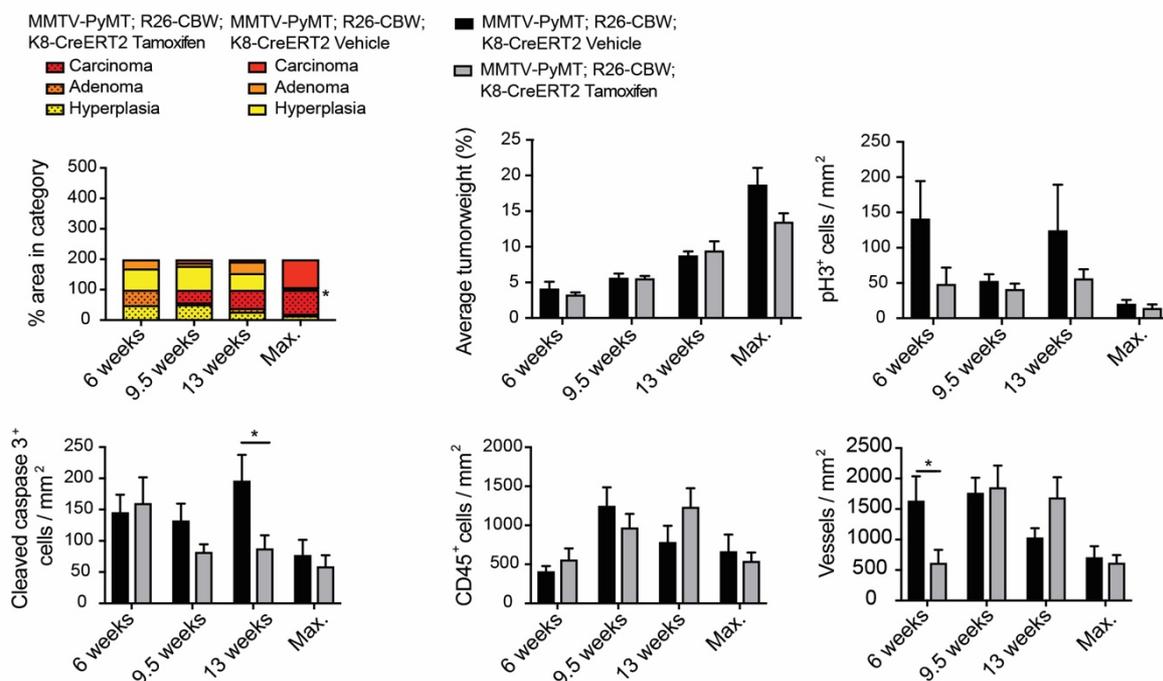
This work was supported by the SystemsX.ch MTD project MetastasiX, the Swiss National Science Foundation, the Swiss Cancer League, and the Krebsliga Beider Basel.

3.1.6 Supplemental Information

a



b

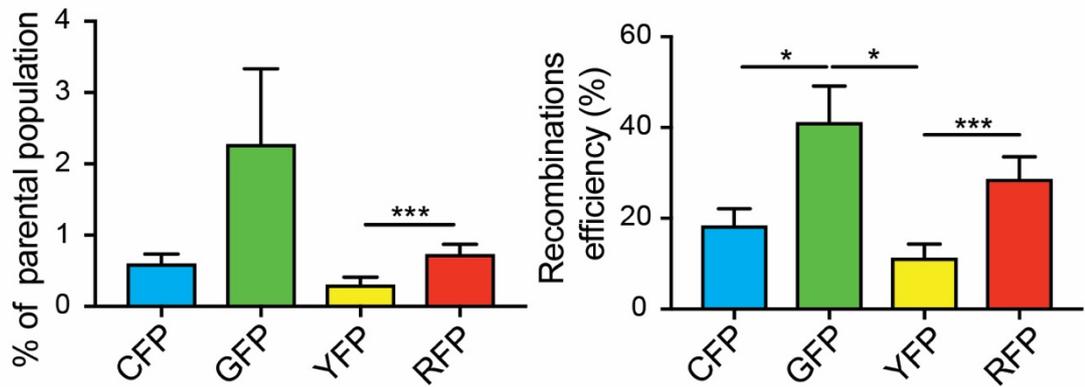


Supplementary Figure 1. Cytokeratin 8 expression in luminal epithelial cells in the mammary gland.

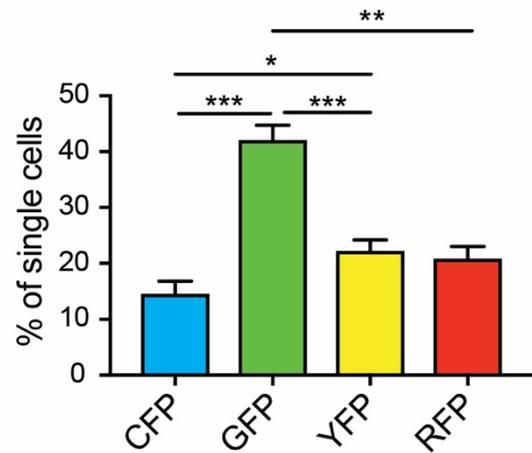
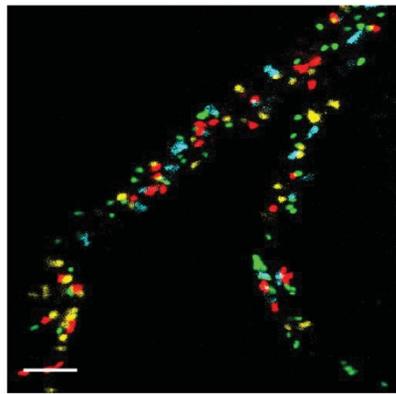
a, Representative pictures of MMTV-PyMT mammary gland tumours (14 weeks) immunolabeled for Keratin 8 (K8, green), Keratin 14 (K14, red) and DAPI (blue) (n = 3). Scale bar, 50 μ m.

b, Vehicle and tamoxifen-treated mammary gland tumours of MMTV-PyMT; R26-CBW; K8-CreERT² mice of 6, 9.5 and 13 weeks of age and when the maximum tumour volume had been reached, were H&E graded and quantified for the percentage average tumour weight, pH3⁺, ccaspase 3⁺ and CD45⁺ cells and tumour microvessel density (average tumour weight: n = 2-17 mice per time point and treatment group, for other analyses: n = 2-3 per time point and treatment group). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Student's t-test.

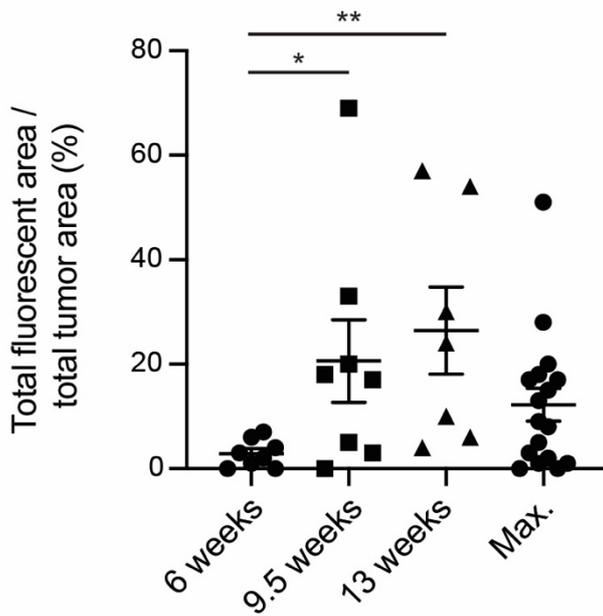
a



b



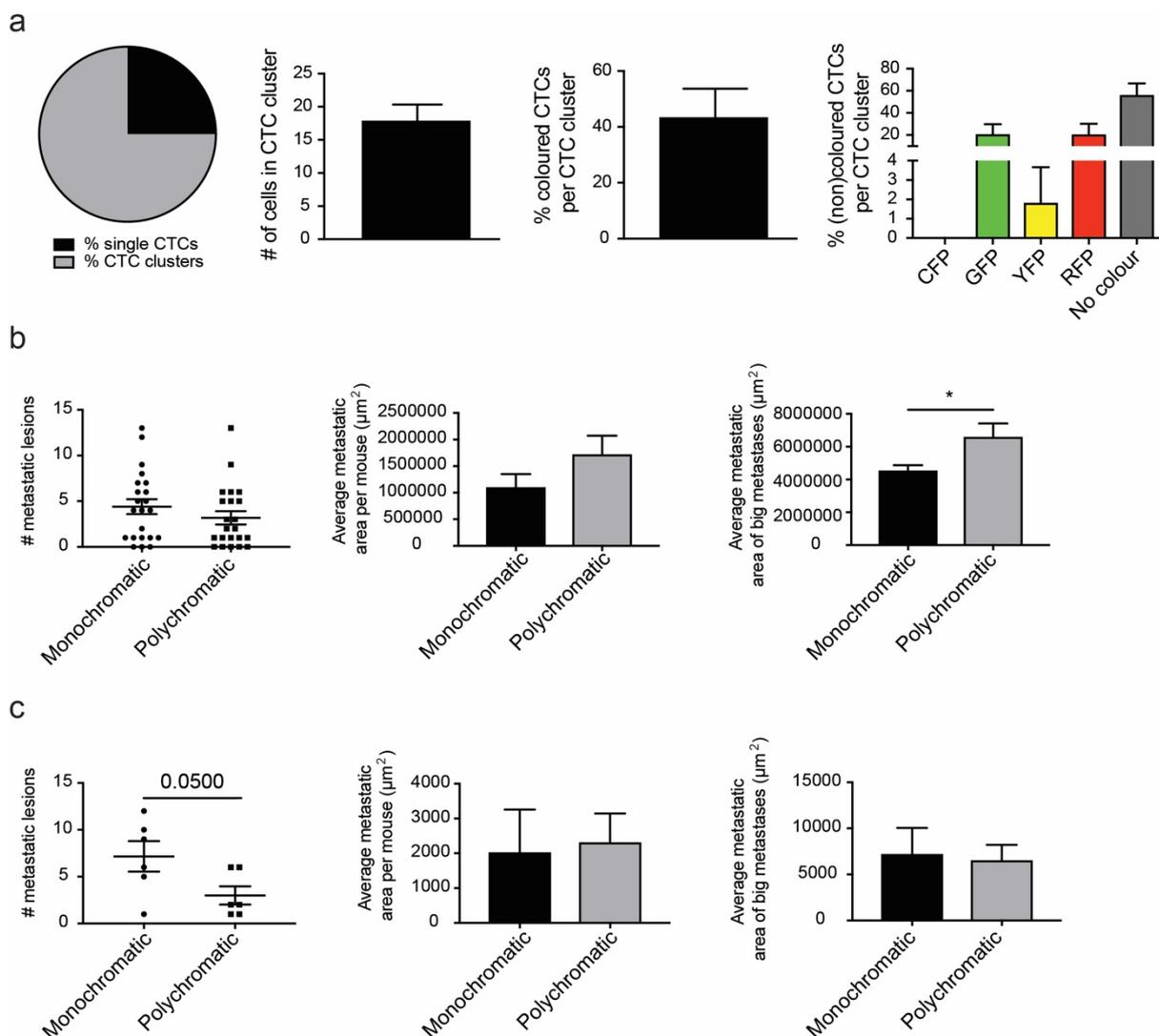
c



Supplementary Figure 2. A single tamoxifen administration results in an efficient K8-driven colour randomization. **a**, Digested, whole tumour lysates of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mammary gland tumours following one week of tracing were subjected to flow cytometry analysis. The percentage of CFP-, GFP-, YFP- and RFP-cells were quantified, and recombination's efficiencies were calculated (n = 6, two mammary glands each). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, paired, two-tailed Student's t-test.

b, Whole mount (left panel) and quantifications of the percentage single cells expressing CFP, GFP, YFP or RFP (right panel) of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mammary gland tumours following one week of tracing (n = 3). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, paired, two-tailed Student's t-test. Scale bar, 50 μm.

c, Total recombination efficiency (% fluorescence area/total tumour area) of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mammary gland tumours following one, 4.5- and eight weeks of tracing or when the maximum tumour volume had been reached (n = 7-17 mice per time point). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Student's t-test.



Supplementary Figure 3. Polychromatic metastatic spread.

a, Pie Chart indicating that primary mammary gland tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice shed few single CTCs and mainly CTC clusters (left panel). Quantification of the cell number in CTC cluster and percentage of (non) coloured CTCs per CTC cluster (right panels) (n = 3).

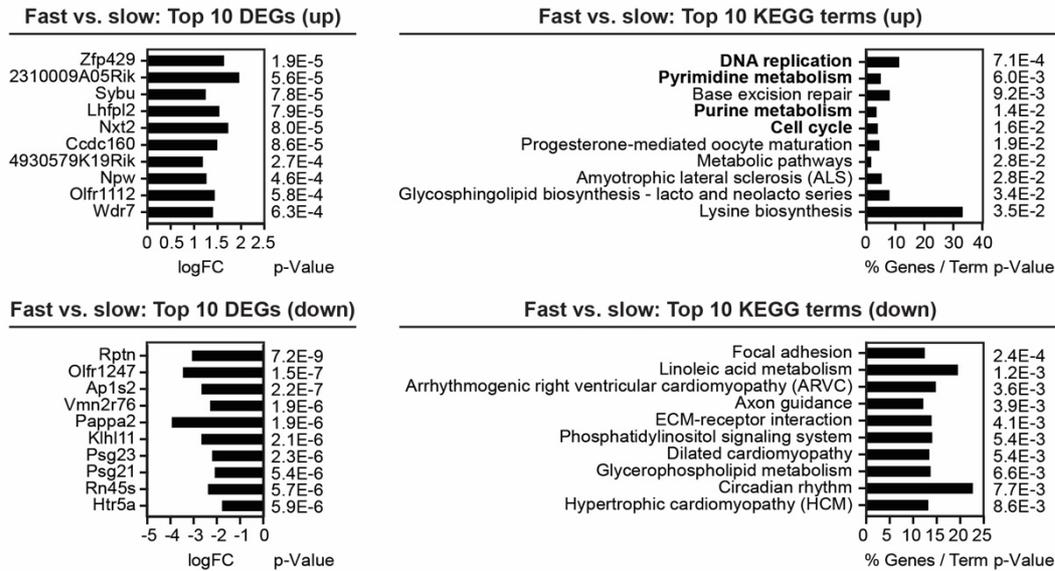
b, Quantification of the number and average area of monochromatic vs. polychromatic metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} tumour-piece transplanted NSG mice (left and middle panel). A quantification of the average metastatic area of big monochromatic vs. polychromatic metastatic lesions is shown on the right (n = 22). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, paired, two-tailed Student's t-test.

c, Quantification of the number and average area of monochromatic vs. polychromatic metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice of 13 weeks of age (left and middle panel). A quantification of the average metastatic area of big monochromatic vs. polychromatic metastatic lesions is shown on the right (n = 6). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, paired, two-tailed Student's t-test.

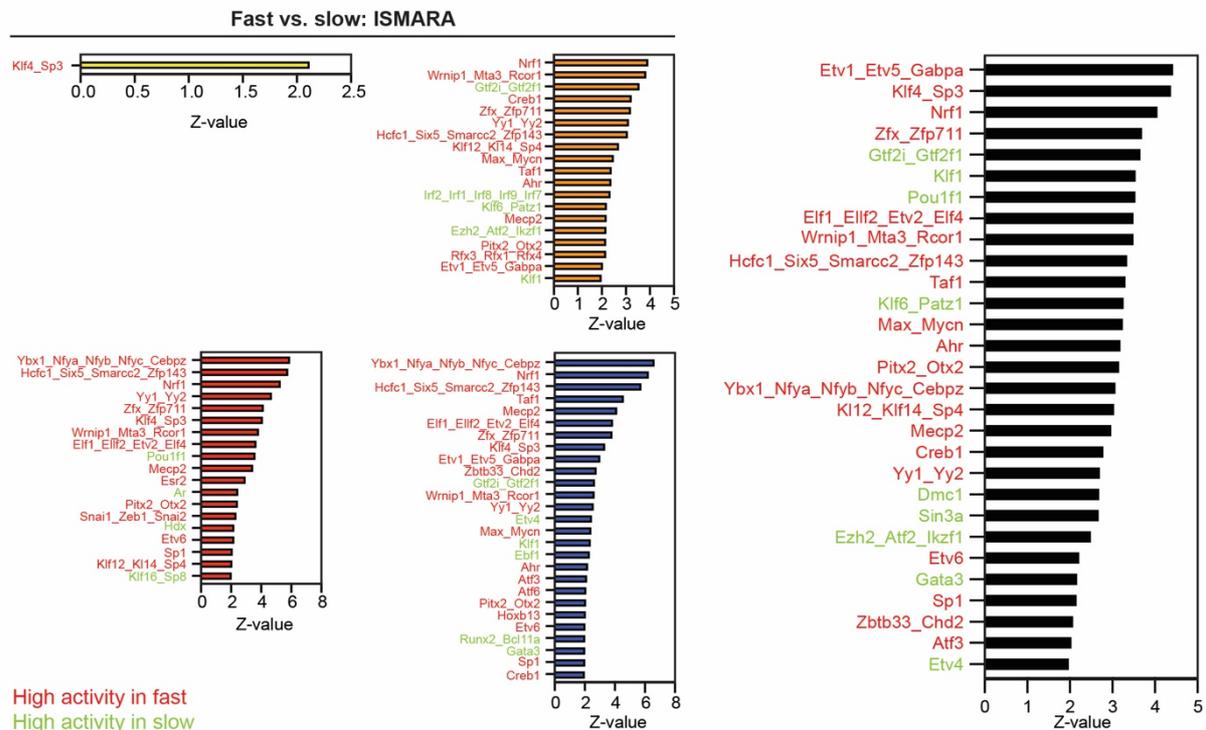
a

Colour	Normal	Hyperplasia	Adenoma	Carcinoma	Metastasis
GFP	Pool of clonal populations	2	1	1	1
YFP		0	3	3	1
RFP		3	4	5	5
Total	3	5	8	9	7

b



c



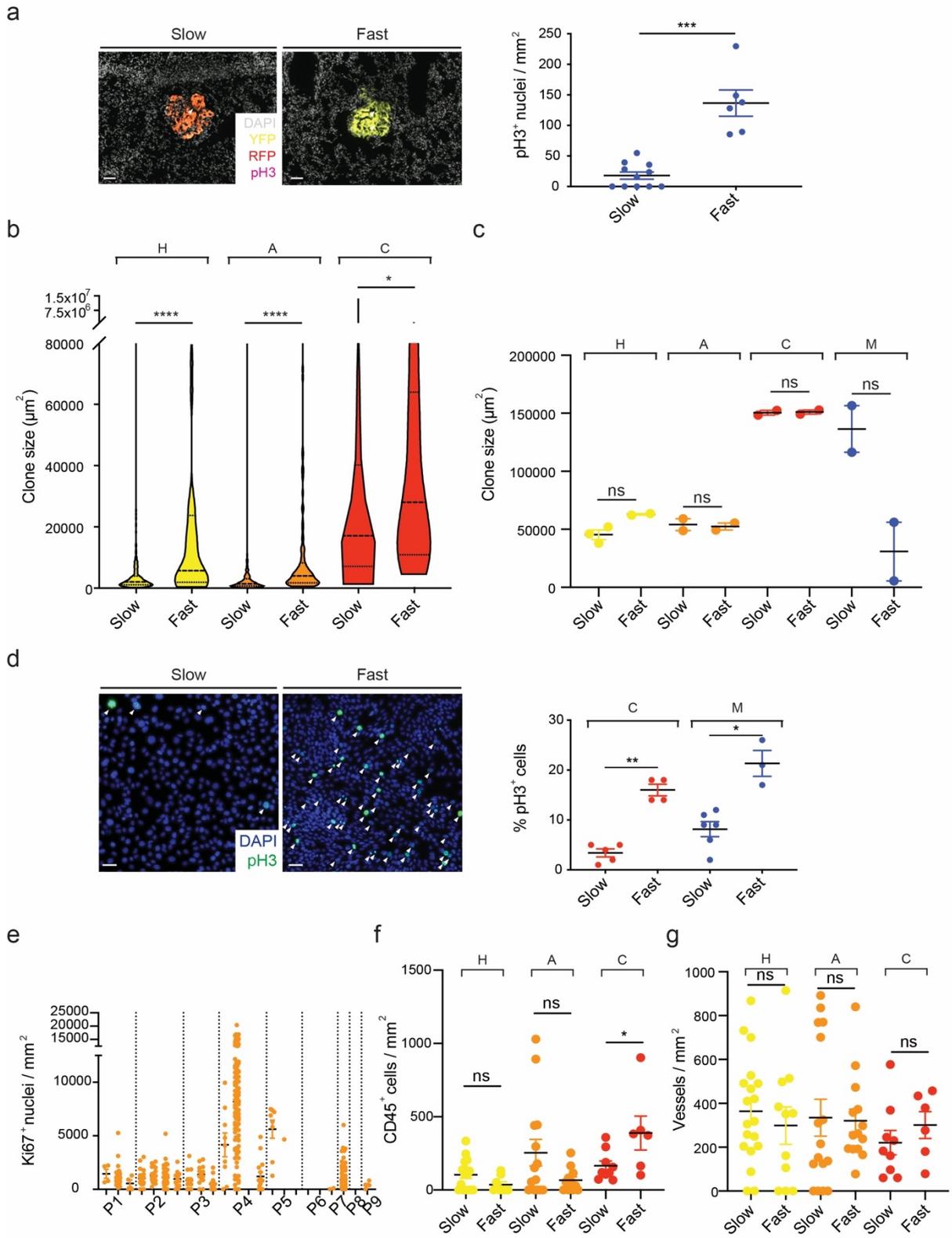
Supplementary Figure 4. Substantial level of intra-stage heterogeneity.

a, Overview of the number and fluorescent marker expression of monoclonal cell populations of various stages, e.g. normal, hyperplasia, adenoma, carcinoma and metastasis, which were laser capture microdissected from primary tumour sections of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice and subjected to RNA sequencing (normal: n = 3, hyperplasia: n = 5, adenoma: n = 8, carcinoma: n = 9, metastasis: n = 7).

b, Lists of the top 10 up (upper panels) and downregulated (lower panels) differentially expressed genes (DEGs) (logFC, left panels) and KEGG terms (percentage of expressed genes per term, right panels), comparing all fast

proliferative clonal populations (n = 8) to all slow proliferative ones (n = 9). All cell-cycle related terms are highlighted in bold.

c, Lists of all ISMARA motifs with $z > 2$, comparing per stage each fast-proliferating subclass (n = 2) to the slow-proliferating one (n = 2) (left panels). Right panel: a list of all ISMARA motifs with $z > 2$, comparing all fast-proliferating clonal populations (n = 8) to all slow-proliferating ones (n = 9). Motifs with a high activity in fast and slow-proliferating clones are highlighted in red and green, respectively. The following colour-code was used: hyperplasia (yellow), adenoma (orange), carcinoma (red) and metastasis (blue).



Supplementary Figure 5. Proliferative subclasses are further characterized by an increase in area.

a, Representative pictures of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} metastatic lesions, immunolabeled for pH3 (magenta) and DAPI (grey) (left panels). Differences in pH3⁺ nuclei per area of slow- and fast proliferating monochromatic metastatic lesions are depicted in the right panel (n = 3, total of 17 monochromatic metastatic lesions). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test. Scale bar, 50 μ m.

b, Violin plots of clonal areas in slow vs. fast-proliferating populations in each stage of primary tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice (n = 4). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test.

c, Clone size of all laser captured microdissected slow vs. fast-proliferating populations in each stage of primary tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice (n = 2-3 slow vs. fast-proliferating clonal populations per stage). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test.

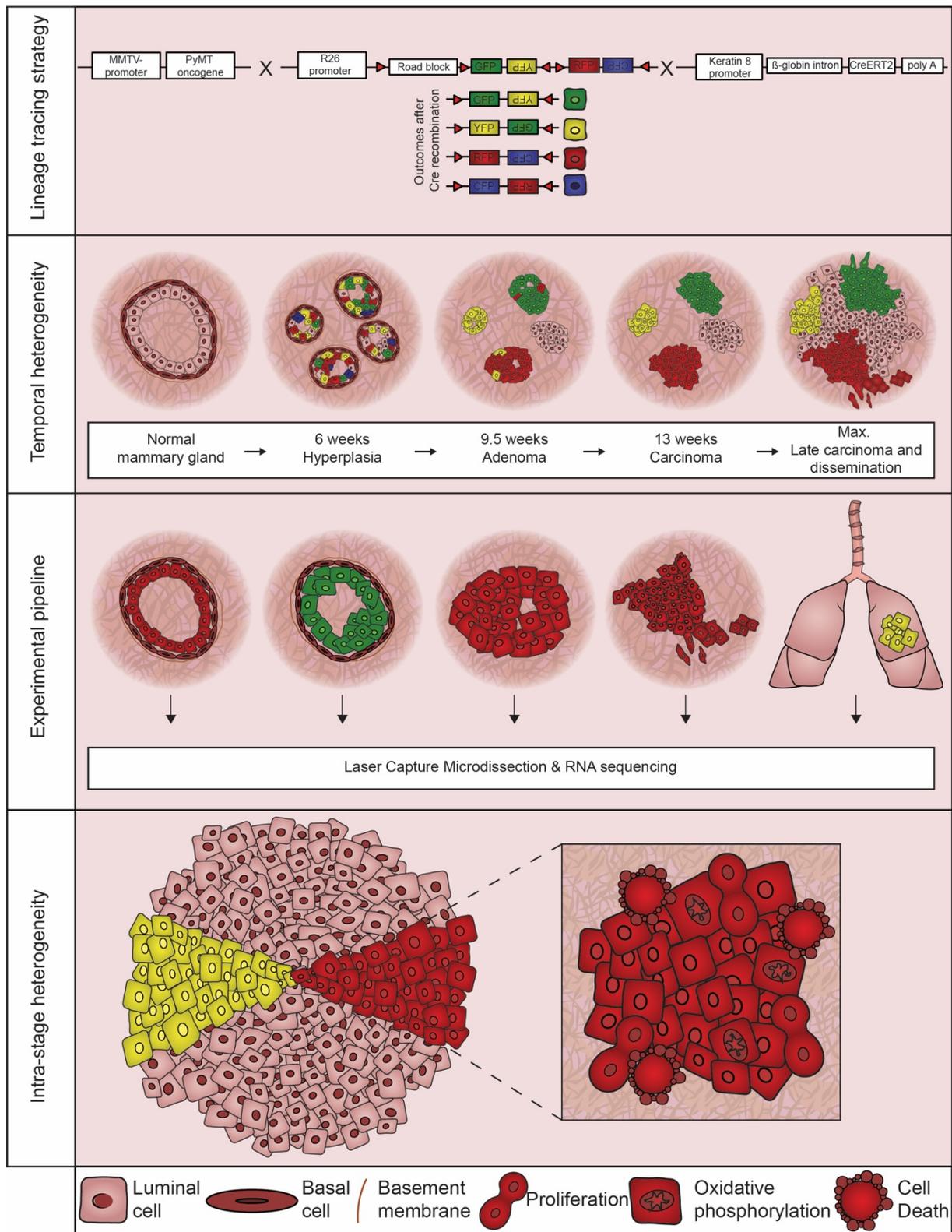
d, Microscopic images of monoclonal cell lines from primary tumours of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice and from metastatic lesions of colour-induced MMTV-PyMT; R26-CBW; K8-CreER^{T2} mice or mice transplanted with a triple-transgenic tumour piece, which were labelled for pH3 (green) and DAPI (blue) (left panels). Slow and fast-proliferating cell lines isolated from tumour (red, n = 9) and metastatic lesions (blue, n = 9) differed in their percentage of pH3⁺ cells (right panel). **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test. Scale bar, 50 μ m.

e, A quantification of Ki67⁺ nuclei per lesion area indicated a tremendous inter and intra-tumoural heterogeneity in human breast cancer patients (n = 9 patients, 23 specimens).

f, Lesions below or above the average of pH3⁺ nuclei/mm² per tumour stage per MMTV-PyMT mouse (13 weeks) were classified as slow or fast-proliferating clonal populations, respectively. A quantification of CD45⁺ cells per area (mm²) indicated that the proliferation phenotype is not attributable to variations in immune cell infiltrates. n = 3. **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test.

g, Lesions below or above the average of pH3⁺ nuclei/mm² per tumour stage per MMTV-PyMT mouse (13 weeks) were classified as slow or fast-proliferating clonal populations, respectively. A quantification of CD31⁺ vessels per area (mm²) revealed that the proliferation phenotype is not attributable to tumour microvessel density. n = 3. **** p < 0.0001, *** p < 0.001, ** p < 0.01, * p < 0.05, unpaired, two-tailed Mann-Whitney U test.

Following colour-code was used throughout the figure: hyperplasia (yellow), adenoma (orange), carcinoma (red) and metastasis (blue).



Supplementary Figure 6. Schematic overview to study heterogeneity during the progression of metastatic breast cancer.

Lineage tracing strategy: In this model, the Mouse Mammary Tumour Virus promoter (MMTV) is used to drive a mammary-gland specific expression of the Polyoma virus middle T antigen (PyMT) oncogene, which will result in malignant transformation of the mammary gland epithelium. The cytokeratin 8 promoter is used to drive the expression of a tamoxifen-inducible Cre recombinase. Upon tamoxifen treatment, activation of the Cre recombinase leads to stochastic expression of one of the four fluorescent proteins CFP, GFP, YFP and RFP in K8-expressing luminal epithelial cells and their progeny in the mammary gland.

Temporal heterogeneity: The confetti lineage tracing system allows to visualize clonal heterogeneity, culminates in clonal dominance and indicates polyclonal metastatic spread.

Experimental pipeline: Laser capture microdissection and RNA sequencing of clonal populations reveals transcriptomic variations across and particularly within various tumour stages.

Intra-stage heterogeneity: The observed intra-stage heterogeneity is conveyed by differences in proliferation, oxidative phosphorylation and cell death.

3.2 The FAK inhibitor BI 853520 exerts anti-tumor effects in breast cancer

Stefanie Tiede¹, Nathalie Meyer-Schaller¹, Ravi Kiran Reddy Kalathur^{1,2}, Robert Ivanek^{1,2}, Ernesta Fagiani¹, Philip Schmassmann¹, Patrick Stillhard¹, Simon Häfliger¹, Norbert Kraut³, Norbert Schweifer³, Irene C. Waizenegger³, Ruben Bill¹ and Gerhard Christofori¹

¹ Department of Biomedicine, University of Basel, 4058 Basel, Switzerland

² Swiss Institute of Bioinformatics, 4053 Basel, Switzerland

³ Boehringer Ingelheim RCV GmbH & Co KG, 1121 Vienna, Austria

These authors contributed equally: Ruben Bill, Gerhard Christofori

- published -

In September 2018 in *Oncogenesis*

<https://doi.org/10.1038/s41389-018-0083-1>

3.2.1 Abstract

Focal adhesion kinase (FAK) is a cytoplasmic tyrosine kinase that regulates a plethora of downstream signaling pathways essential for cell migration, proliferation and death, processes that are exploited by cancer cells during malignant progression. These well-established tumorigenic activities, together with its high expression and activity in different cancer types, highlight FAK as an attractive target for cancer therapy. We have assessed and characterized the therapeutic potential and the biological effects of BI 853520, a novel small chemical inhibitor of FAK, in several preclinical mouse models of breast cancer. Treatment with BI 853520 elicits a significant reduction in primary tumor growth caused by an anti-proliferative activity by BI 853520. In contrast, BI 853520 exerts effects with varying degrees of robustness on the different stages of the metastatic cascade. Together, the data demonstrate that the repression of FAK activity by the specific FAK inhibitor BI 853520 offers a promising anti-proliferative approach for cancer therapy.

3.2.2 Introduction

Every year, more than 1.4 million women worldwide are diagnosed with breast cancer, and over 450,000 women will lose their lives to this disease, mostly due to metastasis¹⁸⁶. Over the past decades, we have gained many important insights into breast cancer biology, which in turn have allowed the development of therapeutic approaches targeting molecules and signaling pathways specifically present in breast cancer cells^{187, 188}. Previous studies have linked the overexpression and activation of focal adhesion kinase (FAK) with the initiation and progression of a wide variety of malignancies, such as ovarian, head and neck, and breast carcinoma¹⁸⁷⁻¹⁹¹. FAK is a multifunctional cytoplasmic tyrosine kinase that forms an important component of focal adhesion sites^{135, 192-195}. Once recruited by signals initiated at integrin-mediated extracellular matrix attachment sites and by multiple growth factor receptors, such as epithelial growth factor receptor (EGFR), vascular endothelial growth factor receptor (VEGFR) and platelet derived growth factor receptor (PDGFR), FAK undergoes a conformational change, enabling autophosphorylation of the tyrosine residue (Y) 397 at its N-terminal domain^{188, 196, 197}. Subsequently, phosphorylated Y397 serves as a docking site for SRC homology 2 containing SRC family kinases, which results in a fully active FAK-SRC signaling complex that can trigger various downstream signaling

pathways known to control cell migration, invasion, proliferation and death – all activities pivotal for malignant tumor progression^{135, 188, 192, 195, 198-202}.

Previous studies have indicated that the forced expression of FAK in endothelial cells enhances angiogenesis and that the ectopic expression of a constitutive-active form of FAK in murine mammary cancer cells promotes their proliferation. Conversely, decreasing FAK expression impairs cancer cell proliferation *in vitro* and *in vivo*^{191, 195, 203-205} and inhibits endothelial cell proliferation *in vitro* and *in vivo*. These data together suggest a linear relationship between FAK activity and tumorigenesis^{193, 203, 204, 206, 207}. However, a recent study has reported that the heterozygous depletion of FAK in endothelial cells increases endothelial cell proliferation and tumor angiogenesis, indicating a non-linear effect of FAK activity in carcinogenesis^{188, 203, 204}. Supporting this notion, low dose treatment with the FAK inhibitor PF-573228 increases microvessel sprouting *ex vivo* and tumor growth *in vivo*²⁰³. These results indicate that the causal link between FAK activity and tumor progression still escapes a final conclusion, and further investigations are warranted to delineate the functional contribution of FAK to carcinogenesis.

We have evaluated the therapeutic and biological effects of BI 853520, a novel, potent and selective small chemical entity kinase inhibitor of FAK (FAK-I)¹⁴⁴, in cultured murine breast cancer cells *in vitro* and in various transplantation and transgenic mouse models of breast cancer *in vivo*. Gene expression profiling of primary tumors of mice treated with BI 853520 reveals a decrease in the expression of genes regulating cell proliferation. Indeed, treatment with BI 853520 provokes a significant reduction in cell proliferation *in vitro* and *in vivo*. In contrast, BI 853520 exerts heterogeneous effects on pulmonary metastasis at different levels of the metastatic cascade depending whether it is used in a neoadjuvant or adjuvant therapeutic setting. Thereby, the epithelial cell adhesion molecule E-cadherin may serve as a potential predictive marker for increased sensitivity of cancer cells to treatment with BI 853520.

3.2.3 Results

3.2.3.1 The FAK-I BI 853520 represses Y397-FAK autophosphorylation

To determine the *in vitro* efficacy of the FAK-I BI 853520 in repressing Y397-FAK phosphorylation in differentiated breast cancer cells and in breast cancer cells that have undergone an epithelial-mesenchymal transition (EMT) and also test the generality of the findings, we employed a combination of various murine mammary

cancer cell lines. 4T1 cells, which have been derived from a spontaneous tumor in a mammary gland of a Balb/c mouse, show hallmarks of a partial EMT and are highly metastatic upon transplantation into mice. The Py2T cell line has been established from a tumor of a MMTV-PyMT transgenic mouse. Py2T cells exhibit an epithelial cell morphology and undergo a reversible EMT upon long-term treatment with transforming growth factor β (TGF β) *in vitro* (Py2T-LT cells)²⁰⁸. 4T1, Py2T, and Py2T-LT cells were treated with different concentrations of BI 853520, and Y397-FAK phosphorylation was assessed by immunoblotting analysis and immunofluorescence staining using a phospho-Y397-FAK-specific antibody. BI 853520 significantly reduces Y397-FAK autophosphorylation in all cell types (**Fig. 1A and B**), while the phosphorylation levels of FAK's homologue PYK2 remained unaffected up to micromolar concentrations (**Fig. 1A**). Next, a time course experiment was performed on 4T1 cells treated for various time points with 0.1 μ M BI 853520 to analyze FAK-I efficacy in a time-dependent manner. Decreased Y397-FAK autophosphorylation following 0.1 μ M BI 853520 treatment occurred within 10 minutes and was substantially reduced at least for the following 48 hours (**Fig. 1C**), demonstrating a fast and potent inhibition of FAK by BI 853520 in this highly metastatic murine breast cancer cell line.

3.2.3.2 BI 853520 represses tumor cell proliferation and invasion only in 3D culture

We next assessed the repressive effect of the FAK-I BI 853520 on proliferation and invasion of 4T1 cells, Py2T cells and Py2T cells that have undergone an EMT (Py2T-LT). When cultured in 2D in plastic dishes, epithelial Py2T cells and mesenchymal Py2T-LT cells did not show any change in cell morphology or proliferation upon treatment with up to 3 μ M BI 853520. However, at a concentration of 10 μ M and above a cytotoxic effect and massive apoptosis could be observed with both cell types (**Fig. 1D**).

To further investigate the repressive effect of BI 853520 on breast cancer cell proliferation *in vitro*, we conducted a MTS-assay on 4T1, Py2T and Py2T-LT cells as previously reported (**Suppl. Fig. 1A**)²⁰⁹. Increasing concentrations of BI 853520 led to a modest yet significant reduction in cell viability of all three breast cancer cell types grown in 2D already at 10-100nM, with a massive reduction at the toxic concentrations above 3 μ M. Flow cytometry-based EdU/PI cell cycle analysis of 4T1 cells revealed a shift in cells leaving the S-phase and entering the G1-phase (**Suppl. Fig. 1B**). Consistent with these results, the percentage of pH3-positive nuclei was decreased in

4T1 cells with increasing concentrations of BI 853520, while increased rates of apoptosis were only observed at the toxic concentration of 10 μ M BI 853520 (**Suppl. Fig. 1C, D**).

In contrast, when cultured in a 3D environment, such as in soft agar, colony formation of mesenchymal Py2T-LT started to be compromised already at a concentration of 1nM BI 853520 (**Fig. 1E**). Epithelial Py2T cells did not form colonies in this assay even in the absence of BI 853520 (data not shown).

We have previously reported that mesenchymal Py2T-LT cells show a massive invasion into the surrounding matrix when cultured in 3D in Matrigel, while epithelial Py2T cells form highly differentiated spheres²⁰⁸ (**Figure1F; Suppl. Fig. 1E**). Comparable invasive growth was also observed with 4T1 cells when cultured in Matrigel (**Suppl. Fig. 1E; Suppl. Fig. 2A**). The invasion of mesenchymal Py2T-LT cells was already inhibited at 10nM BI 853520, while sphere formation of epithelial Py2T cells was only affected by increased cell death at the toxic concentrations of BI 853520 higher than 1 μ M (**Fig. 1F**). This effect was not as pronounced with 4T1 cells cultured in Matrigel, where BI 853520 at doses above 0.1 μ M repressed cell invasion, although without statistical significance, while high doses exerted a toxic effect (**Suppl. Fig. 1E**). The efficient repression of FAK autophosphorylation in the 3D Matrigel cultures was determined by immunofluorescence staining of the spheroids and by immunoblotting of spheroid lysates (**Suppl. Fig. 2A, B**). Despite a decrease in cell invasion, the immunofluorescence and immunoblotting analysis revealed that FAK inhibition did not substantially affect the expression of the mesenchymal markers fibronectin, vimentin or Zeb1 and the epithelial markers E-cadherin or ZO1 in the spheroid cultures.

These results indicate that the specific inhibition of cell proliferation and invasion at low doses of BI 853520 is functional only in three-dimensional cell culture conditions, whereas cells cultured on plastic only respond to BI 853520 at very high, toxic doses of BI 853520.

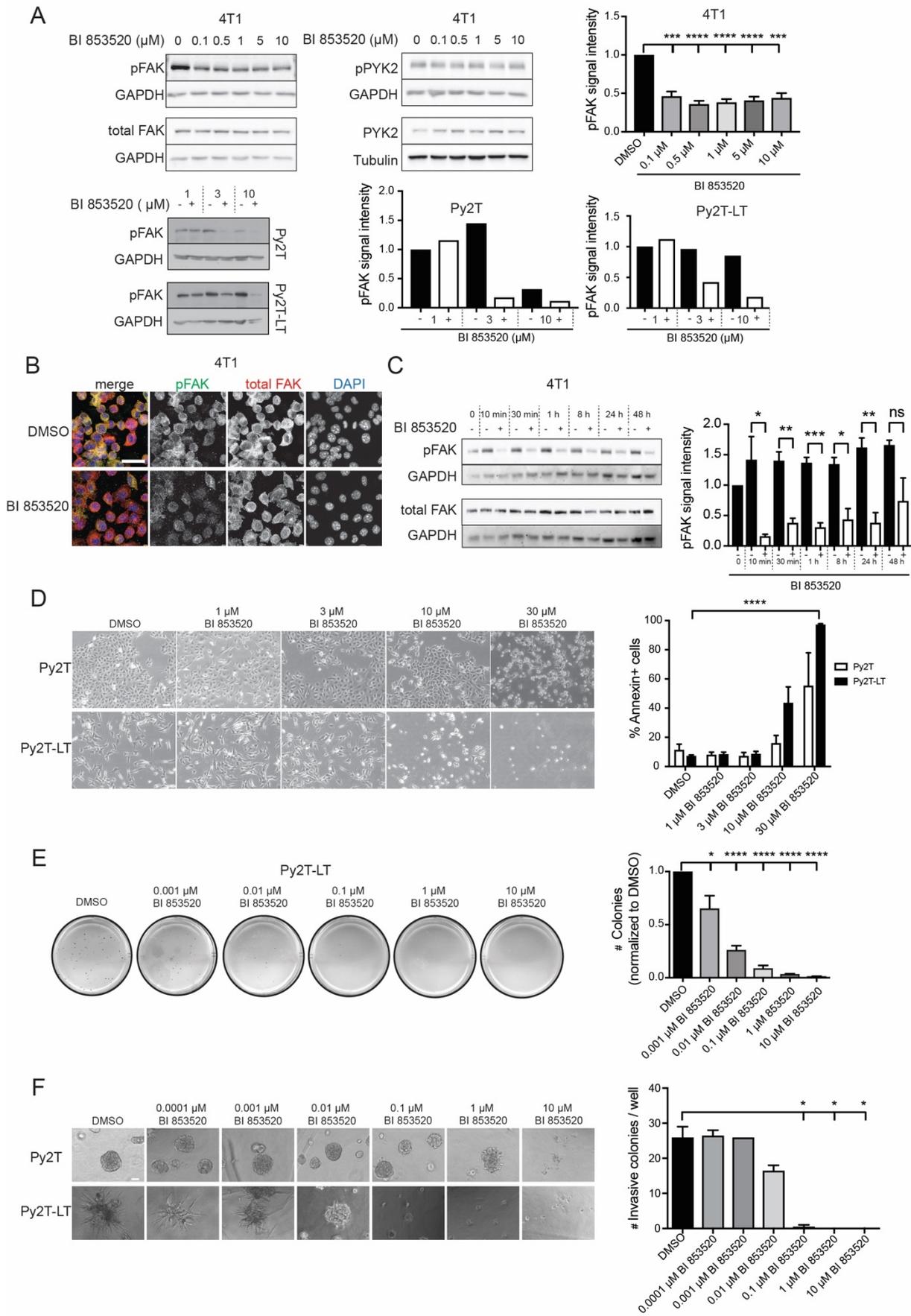


Figure 1: BI 853520 decreases Y397-FAK phosphorylation in a dose and time-dependent manner. (A) To determine the effective dose of BI 853520, 4T1, Py2T and Py2T-LT cells were exposed to increasing concentrations of BI 853520 (0μM, 0.1μM, 0.5μM, 1μM, 5μM and 10μM for 4T1 cells; 1μM, 3μM and 10μM for Py2T- and Py2T-LT cells) for 24 hours. Top left panel: 4T1 cell lysates were examined by immunoblotting analysis

for phospho-FAK-Y397 (pFAK) and total FAK (FAK). GAPDH served as loading control (n = 4). Immunoblotting for phospho-PYK2-Y402 (pPYK2), total PYK2 (PYK2), and GAPDH and tubulin as loading control is shown in the top middle panel (n = 2). Top right panel: Quantification of pFAK signal intensity normalized to GAPDH and total FAK/GAPDH ratios from the immunoblotting analysis shown on the left. Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05. Bottom left panel: Py2T and Py2T-LT cell lysates were examined by immunoblotting analysis for phospho-FAK-Y397 (pFAK). GAPDH served as loading control (n = 1). Right panel: Quantification of pFAK signal intensity normalized to GAPDH from the immunoblotting analysis shown on the left.

(B) 4T1 cells grown on coverslips were treated with the effective dose of BI 853520 (0.1 μM) for 96 hours, followed by immunofluorescence staining for pFAK (Y397, green), total FAK (red) and DNA (blue). Images were obtained using a laser-scanning confocal microscope Leica SP5 (n = 2). Scale bar, 30 μm.

(C) To determine the onset and duration of BI 853520-mediated pFAK downregulation, 4T1 cells were treated for the times indicated with 0.1 μM BI 853520. Cell lysates were analyzed by immunoblotting for pFAK (Y397), total FAK and GAPDH as loading control (n = 3). Right panel: Quantification of pFAK signal intensity normalized to GAPDH and total FAK/GAPDH ratios from the immunoblotting analysis shown on the left. Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.

(D) Cytotoxic effects of high doses of BI 853520 on murine breast cancer cells cultured on plastic surface. Epithelial Py2T and mesenchymal Py2T-LT murine breast cancer cells were treated with BI 853520 at the concentrations indicated and phase contrast micrographs were taken. Cells were then suspended, and staining for Annexin-V as marker of apoptosis was quantified by flow cytometry (n = 2). Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05. Scale bar: 120 μm.

(E) BI 853520 represses colony formation in three-dimensional growth conditions. Py2T-LT cells were seeded in soft agar gels and treated with increasing concentrations of BI 853520 as indicated. The numbers of colonies were quantified. (n = 3). Unpaired, two-tailed Student's t-test. ****, p < 0.0001; *, p < 0.05.

(F) BI 853520 represses mesenchymal murine breast cancer cell invasion cultured in Matrigel. Epithelial Py2T and mesenchymal Py2T-LT murine breast cancer cells were seeded in Matrigel and treated with increasing concentrations of BI 853520 as indicated. Phase contrast microscopy pictures were taken and the numbers of invasive colonies were quantified. Epithelial Py2T cells formed differentiated spheres which only responded to BI 853520 at high, toxic concentrations (n = 1). Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05. Scale bar: 60 μm.

3.2.3.3 BI 853520 attenuates primary tumor growth

To examine the repressive effect of the FAK-I BI 853520 on primary tumor growth *in vivo*, we have employed a variety of cellular and transgenic and transplantation mouse models of breast cancer. MMTV-PyMT transgenic mice develop mammary gland tumors and recapitulate the stepwise progression from differentiated mammary gland epithelia to hyperplasia, adenoma, carcinoma, and lung metastasis²¹⁰. Py2T murine epithelial breast cancer cells are highly tumorigenic and gain mesenchymal characteristics when implanted into mice *in vivo*, yet they are rarely metastatic²⁰⁸. 4T1 cells are highly metastatic breast cancer cells derived from a spontaneous mammary gland tumor of a Balb/c mouse²¹¹. Finally, E-cadherin-proficient MTfIECad cells have been isolated from a lymph node metastasis of a MMTV-Neu;*Cdh1fl/fl* tumor-bearing mouse and have been infected with an adenovirus expressing Cre recombinase to obtain the E-cadherin-deficient cell line MTΔECad²¹². Py2T, 4T1 or MTfIECad breast cancer cell lines were orthotopically transplanted into the mammary fat pad of syngeneic FVB/N, Balb/c or immunodeficient nude (nu/nu) mice, respectively, and mice were treated with 50mg/kg BI 853520 by oral gavage. BI 853520 treatment significantly suppressed primary tumor growth of all three cell lines *in vivo* (Fig. 2A-C; Suppl. Fig. 3A, B). Together, the data suggest a model-

independent suppression of primary tumor growth by BI 853520 which is consistent with data shown by Hirt et al.^{213, 214} Consistent with the results from experiments with cultured cell lines *in vitro* (**Fig. 1A and C**), immunoblotting analysis of primary Py2T tumors of mice treated with BI 853520 for three days displayed a decrease in the levels of Y397-FAK phosphorylation (**Suppl. Fig. 3C**). Similarly, Y397-FAK phosphorylation levels were reduced, although not with statistical significance, when mice bearing 4T1 tumors were treated with a single dose of BI 853520 and sacrificed 1, 4 or 8 hours post treatment (**Suppl. Fig. 3D**).

Since transplantation models of cancer cell lines lack the multi-step development of tumors as observed in patients, we next tested the effect of BI 853520 on primary tumor growth and progression in the MMTV-PyMT transgenic mouse model of metastatic breast cancer, a valuable tool to investigate metastatic breast cancer development and progression²¹⁰. Strikingly, both the administration of BI 853520 at an early stage of tumor development (5 weeks of age; “BI 853520 early”) and at 10 weeks of age (“BI 853520 late”), a time point when most MMTV-PyMT mice had palpable tumors, significantly reduced mammary fat pad weight as a surrogate endpoint for total tumor burden per mouse (**Fig. 2D**). Notably, administration of BI 853520 at 5 weeks of age significantly delayed the development of palpable tumors, i.e. tumor-free survival (**Fig. 2E**). In addition, BI 853520 inhibited tumor progression by reducing the formation of invasive carcinomas within the primary tumors of the transgenic mice (**Fig. 2F**). In summary, the FAK-I BI 853520 exerts a potent capability to repress primary tumor growth and tumor progression in multiple mouse models of breast cancer.

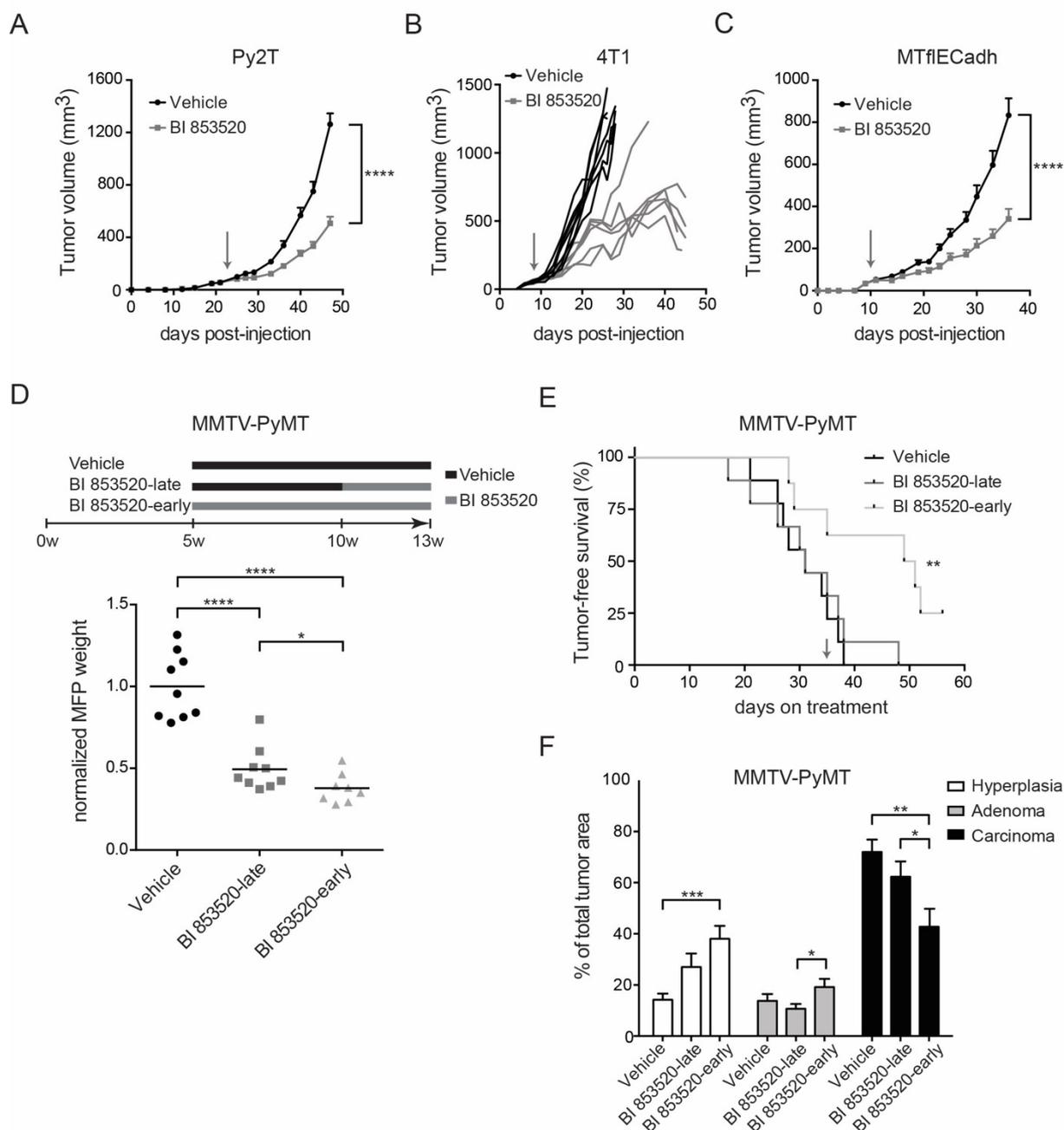


Figure 2: BI 853520 decreases primary tumor growth in various orthotopic pre-clinical mouse models of breast cancer.

(A) Oral treatment of mice bearing Py2T breast tumors with daily 50mg/kg BI 853520 dissolved in Natrosol or with Natrosol-vehicle alone from day 22 post injection (indicated by an arrow) for 25 consecutive days significantly decreased tumor volume over time. $n = 10$ mice per treatment cohort.

(B) Oral treatment of mice bearing 4T1 breast tumors with daily 50mg/kg BI 853520 dissolved in Natrosol or with Natrosol-vehicle alone from day eight post injection (indicated by an arrow) significantly decreased tumor volume over time. Tumor growth curves of individual mice are shown. Vehicle cohort, $n = 7$ mice; BI 853520 cohort, $n = 6$ mice.

(C) Oral treatment of mice bearing MTfIECad breast tumors with daily 50mg/kg BI 853520 dissolved in Natrosol or with Natrosol-vehicle alone from day 10 post injection significantly decreased tumor volume over time (indicated by an arrow). $n = 9$ mice per treatment cohort.

(D) Top: Oral treatment schedule for the treatment of MMTV-PyMT transgenic mice with vehicle, BI 853520-late (50mg/kg daily, from 10 to 13 weeks of age) or BI 853520-early (50mg/kg daily, from 5 to 13 weeks of age). Bottom: Normalized mammary fat pad (MFP) weights for each treatment cohort. All mice were sacrificed at 13 weeks of age, and primary tumors and lungs were used for further analysis. Vehicle cohort, $n = 9$; BI 853520-late, $n = 9$; BI 853520-early, $n = 8$.

(E) Early BI 853520 treatment significantly improves tumor-free survival of MMTV-PyMT mice. Days on treatment starting from week 5 of age are shown. The time reaching termination criteria was determined for the mice treated

as described in panel D. As shown here, at the time point of treatment initiation of the BI 853520-late group (arrow; day 35 on treatment), most mice in this group displayed palpable tumors.

(F) Grading of primary tumors of MMTV-PyMT transgenic mice treated as described in panel D. Data are shown as the percentage of the average area covered by each grade of two histological sections per mouse. Early BI 853520 treatment delays malignant tumor progression (increase in hyperplasia and adenoma stage, decrease in carcinoma stage).

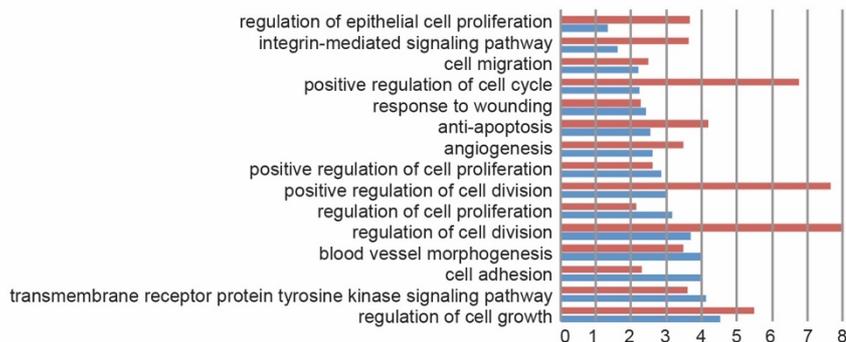
Statistical analysis was performed by unpaired, two-tailed Mann-Whitney U test (A, C, D), by Log-rank (Mantel-Cox) test (E), and unpaired, two-tailed Student's t-test (F). All data are depicted as means \pm SEM. ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

3.2.3.4 BI 853520 restrains tumor cell proliferation

In order to delineate the molecular mechanisms leading to the therapeutic effect of BI 853520 on primary tumor growth, mice harboring 4T1 primary tumors were treated for five days with BI 853520, and RNA extracted from total tumors was subjected to next generation sequencing. Only primary tumors with sufficient RNA quality were included into further analysis (**Suppl. Fig. 4A**). Gene expression correlation analysis displayed a clear separation of transcriptomic profiles derived from primary tumors of mice treated with BI 853520 or vehicle (**Suppl. Fig. 4B**). Comparative gene expression analysis of primary tumors of mice treated with BI 853520 versus vehicle control revealed 1293 upregulated and 475 downregulated genes (cutoffs: p -value ≤ 0.05 , fold change ± 1.5). Functional enrichment analysis for biological processes and signaling pathways indicated that the regulation of epithelial cell proliferation and positive regulation of cell cycle/cell proliferation/cell division were enriched in genes downregulated by BI 853520 treatment (**Fig. 3A**). In line with this finding, gene set enrichment analysis confirmed a significant reduction in the relative expression of genes important for cell cycle and positive regulation of mitotic cell cycle (including cyclin-dependent kinase 1 and 4; Cdk1: \log_2 fold-change = -0.4519, FDR = 4.24e-03; Cdk4: \log_2 fold-change = -0.3072, FDR = 0.003718), while the negative regulation of cell proliferation was increased following BI 853520 treatment (**Fig. 3C**; **Suppl. Fig. 4C**). Interestingly, biological processes linked to T cell differentiation and cell proliferation, (acute) inflammatory response, cytokine production and leukocyte activation were enriched in the group of genes upregulated by BI 853520 treatment (**Fig. 3B**).

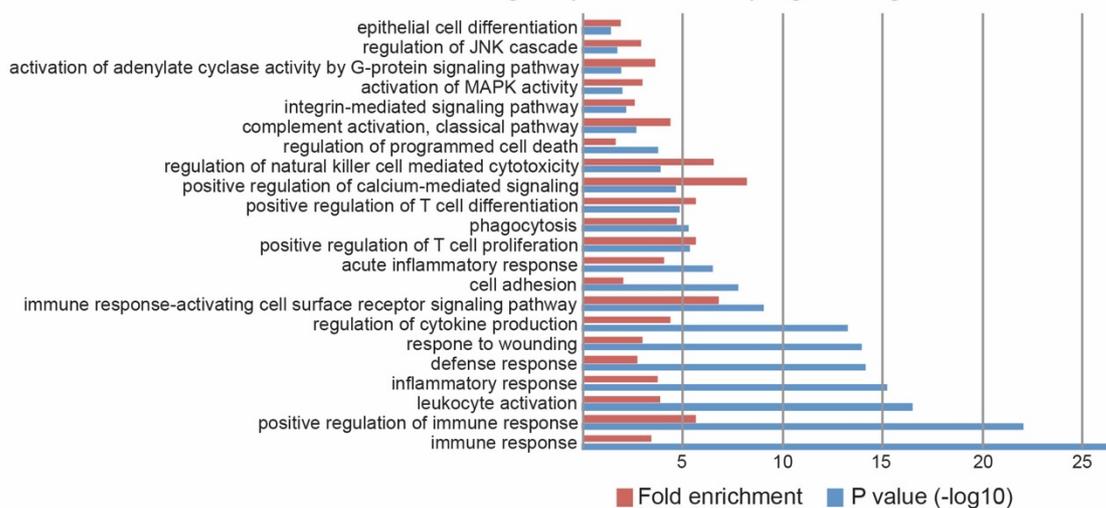
A

Enriched biological processes in downregulated genes



B

Enriched biological processes in upregulated genes



C

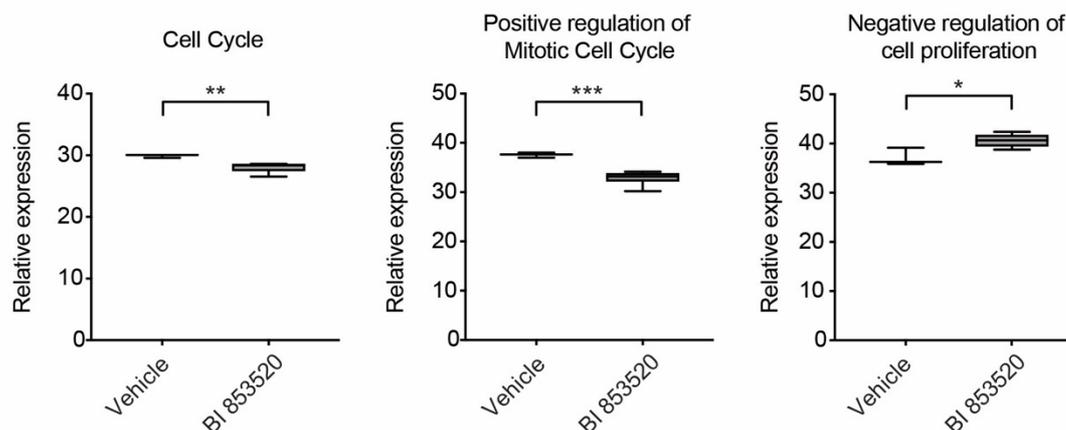


Figure 3: BI 853520 treatment downregulates genes implicated in cell proliferation and upregulates genes implicated in immune response.

(A, B) List of biological processes enriched in downregulated and upregulated genes, respectively, upon treatment of mice bearing primary 4T1 tumors with 50mg/kg BI 853520 dissolved in Natrosol-vehicle or vehicle alone for 5 days. All terms shown in the bar graphs are statistically significant (p -value ≤ 0.05). The blue and red bars represent p-values in $-\log_{10}$ format and the fold enrichment for each term, respectively. The p-values and enrichment were computed using Fisher's exact test.

(C) Box whisker plots showing the relative expression of genes involved in processes of cell cycle, positive regulation of mitotic cell cycle, and negative regulation of cell proliferation that were identified using gene set enrichment analysis. Relative gene expression was compared between vehicle-treated group ($n = 3$) and BI

853520-treated group (n = 6). The significance between the groups is shown as p-values using an unpaired, two-tailed Student's t-test. ***, p < 0.001; **, p < 0.01; *, p < 0.05.

In order to validate the anti-proliferative mechanism of BI 853520 on primary tumor growth, we analyzed the proliferation marker phospho-histone H3 (pH3) on histological sections of tumors from mice orthotopically transplanted with Py2T cells and short-term treated with BI 853520. Immunostaining for pH3 demonstrated a significant reduction of pH3-positive nuclei per area in mice treated with BI 853520 (**Fig. 4A**). Similarly, in both the 4T1 breast cancer model and the MMTV-PyMT transgenic mouse model, BI 853520 treatment led to a significant decrease in pH3-positive nuclei per mm² (**Fig. 4B and C**). In contrast, apoptosis as assessed by immunofluorescence staining for cleaved caspase 3 (cCasp3) was not significantly affected in the transgenic mouse model and in the short-term Py2T breast cancer model (**Fig. 4C and D**). In the 4T1 model, inter-experimental variability of cCasp3 staining and thus of the rate of apoptosis did not allow a reliable conclusion. Immunofluorescence staining of histological sections from the 4T1 tumors for blood vessel density (CD31), for vessel perfusion (FITC-Lectin perfusion), for vessel leakiness (FITC-Dextran) and for tumor hypoxia (Pimonidazole) did not reveal any significant changes (**Suppl. Fig. 5A-D**).

Consistent with the findings on murine breast cancer cells cultured *in vitro*, these results indicate that the repressive effect of BI 853520 on tumor growth is attributable to its anti-proliferative mode of action on tumor cells.

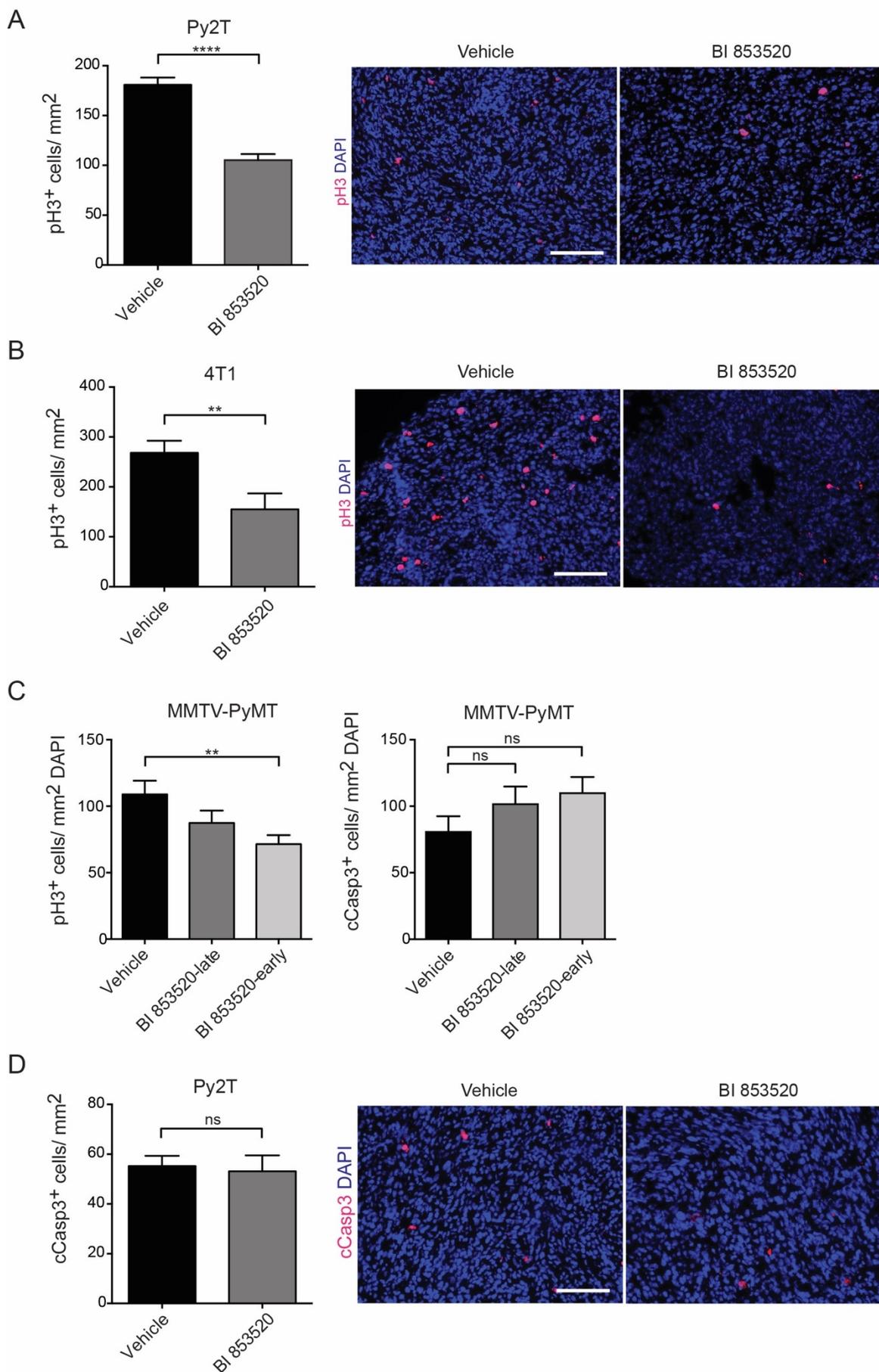


Figure 4: Reduced primary tumor growth is due to an anti-proliferative effect of BI 853520.

(A) Immunofluorescence microscopy analysis of primary tumor sections of Py2T breast tumors of mice treated for three days with vehicle control or with daily 50mg/kg BI 853520 for phospho-histone 3 (pH3; red) and DAPI (cell nuclei; blue). Quantification of pH3-positive nuclei per mm² is shown on the left, representative immunofluorescence images are shown on the right. n = 5 tumors in the vehicle group, n = 3 tumors in the BI 853520-treated group.

(B) Immunofluorescence microscopy analysis of primary tumor sections of 4T1 breast tumor bearing mice treated with vehicle control or daily 50mg/kg BI 853520 and analyzed at the experimental end-point as shown in Figure 2B for phospho-histone 3 (pH3; red) and DAPI (cell nuclei; blue). Quantification of pH3-positive nuclei per mm² is shown on the left, representative immunofluorescence images are shown on the right. n = 4 tumors in the vehicle group, n = 3 tumors in the BI 853520-treated group. Due to the widespread necrotic areas in this tumor model, only areas with intact cell nuclei were assessed.

(C) Left panel: number of pH3-positive nuclei per mm² on primary tumor sections from MMTV-PyMT transgenic mice following treatment with vehicle alone, daily 50mg/kg BI 853520-late (from 10 to 13 weeks of age) or daily 50 mg/kg BI 853520-early (from 5 to 13 weeks of age) as described in Figure 2D. n = 4 tumors in the vehicle group, n = 5 tumors in the BI 853520-late group and n = 5 tumors in the BI 853520-early group. Right panel: number of cleaved caspase-3 (cCasp3)-positive nuclei per mm² of primary tumor sections from MMTV-PyMT transgenic mice following BI 853520 treatment. n = 5 tumors in the vehicle group, n = 6 tumors in the BI 853520-late group and n = 7 tumors in the BI 853520-early group. Due to the massive difference in cellularity between different tumor areas in this particular tumor model, pH3 and cCasp3 counts were corrected to the area covered by DAPI-staining as a surrogate for cellularity.

(D) Immunofluorescence microscopy analysis of primary tumor sections of Py2T breast tumor-bearing mice treated for three days with vehicle control or 50mg/kg daily BI 853520 for cleaved caspase 3 (cCasp3; red) and DAPI (cell nuclei; blue). Quantification of cCasp3-positive nuclei per mm² is shown on the left, representative immunofluorescence images are shown on the right. n = 5 tumors in the control group, n = 3 tumors in the BI 853520-treated group.

Data indicate counts per field of view, shown as mean ± SEM. Statistical difference was determined by the unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.

3.2.3.5 Heterogeneous effects of BI 853520 at different stages of the metastatic cascade

Most studies have primarily focused on primary tumor growth when validating novel anti-cancer compounds, yet metastasis is the major cause of morbidity and mortality. To this end, we performed an in-depth analysis of BI 853520's effect on the different stages of the metastatic process, employing the 4T1 transplantation model of breast cancer that predominantly metastasizes to the lung. Assuming that the metastatic process involves the interplay between cancer cells and cell types of the tumor microenvironment, we orthotopically injected 4T1 cells in BALB/c mice and started BI 853520 treatment three days before (BI 853520 PRE) or seven days post 4T1 cell implantation (BI 853520 POST). Both treatment regimens efficiently reduced primary tumor volume (**Fig. 5A**). Notably, BI 853520 PRE-treated 4T1 tumors were significantly delayed in their initial growth, since they were palpable only at later time points as compared to the vehicle-PRE-treated control cohort (**Suppl. Fig. 6A**). Interestingly, even though BI 853520 POST-treatment led to a significant decline in 4T1 tumor volumes, pulmonary metastasis numbers per section and the metastasis area fraction were only decreased in the BI 853520 PRE-cohort (**Fig. 5A; Suppl. Figure 6B**). This result suggests that, once 4T1 tumor cells have entered the lung parenchyma and proliferation has been initiated, BI 853520 treatment does not longer reduce the number of metastatic nodules. To test this hypothesis, 4T1 tumor cells were

injected intravenously (i.v.) into the tail vein of mice to bypass the processes of primary tumor colonization and intravasation. BI 853520 treatment was either initiated 3 days before i.v. tumor cell injection (BI 853520 PRE) or 3 days after i.v. injection of tumor cells (BI 853520 POST) - a time point when metastatic clones have already started proliferating in the lung (**Suppl. Fig. 6C**). Indeed, pre-treating the mice before i.v. tumor cell injection reduced the number of metastasis and the metastatic area fraction (**Fig. 5B**). However, initiating BI 853520 treatment after metastasis had started to form did not significantly reduce the number of metastasis but decreased metastatic nodule size, as suggested by the reduced metastatic area fraction (**Fig. 5B**), consistent with BI 853520's anti-proliferative activity.

The latter result would argue that BI 853520 may not be able to suppress the metastatic seeding of breast cancer cells but rather inhibit outgrowth of already seeded metastatic modules. To test this possibility, we performed an experiment in which primary tumors were surgically resected 14 days after orthotopic implantation of 4T1 cells, and BI 853520 treatment was initiated 2 days later. Surprisingly, the numbers of metastatic nodules were even increased, however, the metastatic area fraction was slightly reduced, yet without statistical significance (**Fig. 5C; Suppl. Fig. 6D**). When 4T1-transplanted mice were treated with BI 853520 before surgical removal of the primary tumor and then therapy was stopped, primary tumor growth and also the number of metastasis were reduced, although without statistical significance (**Fig. 5D**). However, when reduced primary tumor growth was normalized to the number of metastasis (metastatic index), no significant change in the number of metastasis between BI 853520-treated and control-treated mice could be observed (**Suppl. Fig. 6E**).

A similar overall lack of significant suppression of metastasis by BI 853520 was also seen in the MMTV-PyMT mouse model of metastatic breast cancer. Although primary tumor growth and tumor progression was strongly reduced (**Fig. 2D**), the number of metastases was unchanged regardless of initiating BI 853520 treatment at an early or a late stage of multi-step carcinogenesis (**Suppl. Fig. 6F**). However, comparable to the 4T1 highly metastatic, syngeneic transplantation model of breast cancer, in the MMTV-PyMT transgenic mouse model of stepwise breast cancer progression and metastasis, the percentage of large metastases was significantly reduced by long-term BI 853520 treatment (**Suppl. Fig. 6F**), suggesting that BI 853520 repressed metastatic tumor cell proliferation. Together, these results

indicate that BI 853520 rather represses tumor cell proliferation than metastatic dissemination and that BI 853520 seems to exert heterogeneous effects on metastasis formation depending on the time and the stage of the metastatic cascade when BI 853520 is first administered.

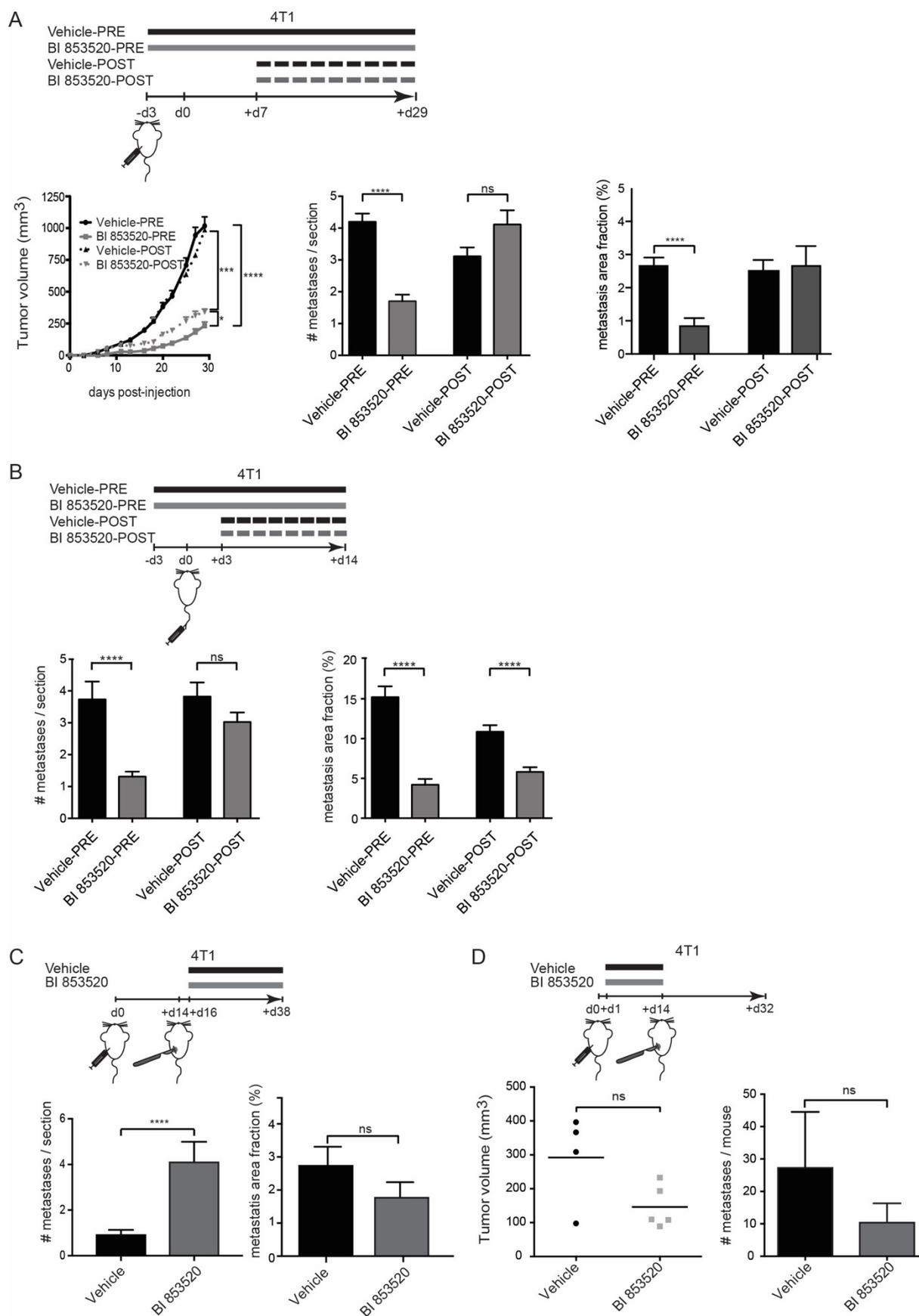


Figure 5: BI 853520 represses outgrowth of pulmonary metastases.

(A) Top panel: Schematic overview of the experimental setup. Mice were either pre-treated with BI 853520 three days prior to 4T1 orthotopic tumor cell inoculation (Vehicle-PRE or BI 853520-PRE groups) or treated seven days post 4T1 orthotopic tumor cell injection (Vehicle-POST or BI 853520-POST groups). Left bottom panel: Primary tumor growth in the indicated treatment cohorts was monitored over time. BI 853520 treatment (BI 853520-PRE

and BI 853520-POST) significantly reduced primary tumor growth. Middle bottom panel: Average numbers of individual pulmonary metastases per lung section. Only pre-treatment with BI 853520 decreased the numbers of pulmonary metastases. Right bottom panel: Average metastatic area fraction (%) in pulmonary cross sections. Only pre-treatment with BI 853520 decreased the outgrowth of pulmonary metastases. $n = 9$, Vehicle-PRE; $n = 9$, BI 853520-PRE; $n = 8$, Vehicle-POST; $n = 8$, BI 853520-POST. Statistical analysis was performed by unpaired, two-tailed Mann-Whitney U test (for tumor volumes) or unpaired, two-tailed Student's test (for number metastases/section and metastasis area fraction).

(B) Top panel: Schematic overview of the experimental setup. Mice were either pre-treated with BI 853520 three days prior intravenous injection of 4T1 cells (Vehicle-PRE or BI 853520-PRE groups) or treated three days post intravenous injection of 4T1 cells (Vehicle-POST or BI 853520-POST groups). Left bottom panel: Average numbers of pulmonary metastases per lung section. BI 853520 pre-treatment significantly decreased pulmonary metastases. Right bottom panel: Average metastatic area fraction (%) in pulmonary cross sections. Both BI 853520 treatments (BI 853520-PRE and BI 853520-POST) significantly reduced the metastatic area fraction. $n = 5$ for all treatment cohorts. Statistical analysis was performed using unpaired, two-tailed Student's t test.

(C) Schematic overview of the experimental setup. 4T1 cells were orthotopically transplanted into the mammary fat pad, followed by primary tumor removal 14 days post-injection and BI 853520 therapy in an adjuvant setting two days post-surgery. Mice were sacrificed 38 days post injection. Left bottom panel: Average numbers of pulmonary metastases per lung section in an adjuvant BI 853520 setting. BI 853520 significantly increased the number of pulmonary metastases. Right bottom panel: Average metastatic area fraction (%) in pulmonary cross sections, with a trend of decrease following adjuvant BI 853520 therapy. $n = 7$, Vehicle; $n = 5$, BI 853520. Statistical analysis was performed using unpaired, two-tailed Student's test.

(D) Schematic overview of the experimental setup. 4T1 cells were orthotopically transplanted into the mammary fat pad, followed by neoadjuvant BI 853520 therapy. Primary tumors were removed 14 days post-injection and BI 853520 therapy was discontinued. Mice were sacrificed 32 days post injection. Left bottom panel: Primary tumor volumes. Right bottom panel: number of metastatic lesions per mouse. $n = 5$, Vehicle; $n = 5$, BI 853520. Statistical analysis was performed by Mann-Whitney test for tumor volumes and Student's t-test for lung metastases.

3.2.3.6 E-cadherin status-dependent effects by BI 853520

FAK is seen as a key mediator in the crosstalk of integrin-mediated cell-matrix and E-cadherin-mediated cell-cell contacts – central mechanisms involved in cancer cell invasion and metastasis²¹⁵. In order to study the influence of the E-cadherin status and, with it, of an epithelial-mesenchymal transition (EMT) on the cells' susceptibility to BI 853520, we employed the E-cadherin-proficient and E-cadherin-deficient murine breast cancer cell lines MTflECad and MTΔECad, respectively. MTflECad cells exhibit a fully differentiated epithelial cell phenotype, while the genetic ablation of E-cadherin expression in MTΔECad resulted into a complete EMT, as manifested by changes in marker gene expression and a gain in the cells' migratory, invasive and metastatic capabilities²¹². MTflECad and MTΔECad cells were orthotopically injected into immunodeficient Rag2^{-/-};γc^{-/-} mice, and BI 853520 treatment was initiated when tumors reached the size of 100mm³ and continued until tumors measured ~1500mm³. Strikingly, BI 853520 treatment significantly delayed primary tumor growth of the highly differentiated, E-cadherin-proficient and non-metastatic MTflECad tumors (**Fig. 6A,B**), as already described above (**Fig. 2C**). In contrast, the growth of E-cadherin-deficient, invasive and metastatic MTΔECad tumors was not apparently affected by treatment with BI 853520 (**Fig. 6A, B**). Notably, BI 853520 reduced the number of metastasis in the lungs of mice transplanted with MTΔECad cells (**Fig. 6C**), whereas MTflECad did not form metastasis even in the absence of BI 853520 (data not shown).

In summary, the data suggest that BI 853520 is more potent in repressing tumor cell proliferation and primary tumor growth of differentiated, E-cadherin-expressing tumor cells (MTflECad, Py2T, in parts 4T1 cells and early stages of MMTV-PyMT tumor progression), while it represses the metastatic outgrowth of invasive, E-cadherin-deficient tumor cells (MTΔECad and late stages/metastasis of MMTV-PyMT tumor cells). The loss of E-cadherin expression is a hallmark of an epithelial-mesenchymal transition (EMT). While an EMT is mainly responsible for primary tumor cell invasion, its reversal, a mesenchymal-epithelial transition (MET), has been shown to contribute to the metastatic outgrowth of disseminated cancer cells in distant organs. Hence, the repressive effect of BI 853520 on metastatic outgrowth may be a result of the more potent anti-proliferative effect of BI 853520 on differentiated cancer cells.

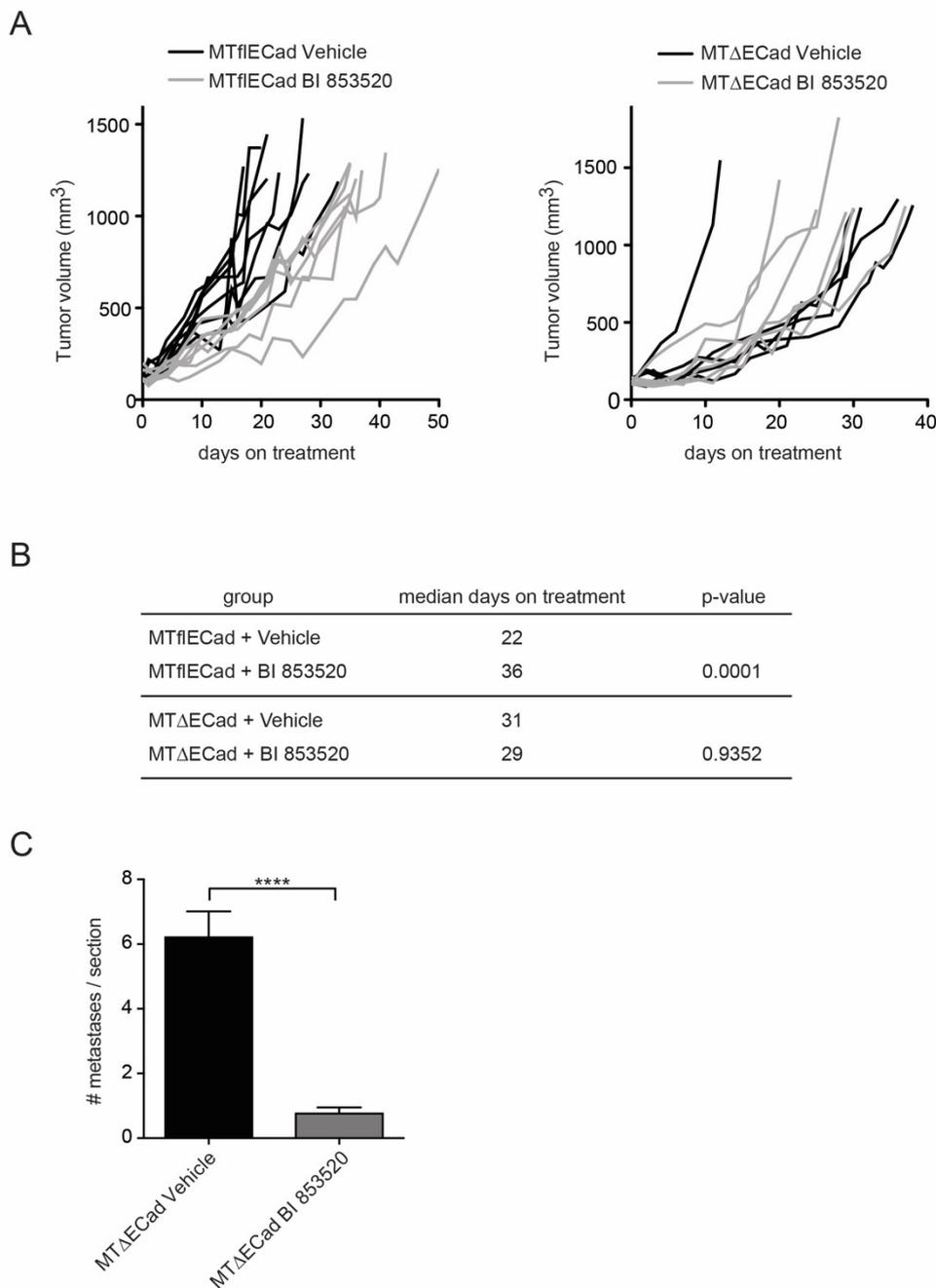


Figure 6: E-cadherin-dependent repression of metastasis formation by BI 853520.

(A) Mice were orthotopically transplanted with E-cadherin-expressing MTflECad or with E-cadherin-deficient MTΔECad murine breast cancer cells. BI 853520 treatment was started at a primary tumor size of 100 mm³, and animals were sacrificed before the primary tumor reached a size of ~1500 mm³. BI 853520 significantly increased survival of mice transplanted with MTflECad cells, but not in mice transplanted with MTΔECad cells.

(B) BI 853520 significantly increases the mean time to reaching termination criteria in mice transplanted with E-cadherin-proficient MTflECad cells, but not in mice transplanted with E-cadherin-deficient MTΔECad cells. Statistical analysis was performed using a log-rank (Mantel-Cox) test. MTflECad: vehicle cohort, n = 8; BI 853520 cohort, n = 7. MTΔECad: vehicle cohort, n = 5; BI 853520 cohort, n = 6.

(C) BI 853520 reduces the average numbers of pulmonary metastases per lung section in the cohort of mice bearing MTΔECad-tumors. Vehicle cohort, n = 5; BI 853520 cohort, n = 6. Statistical analysis was performed using unpaired, two-tailed Student's test. All data are depicted as mean ± SEM. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.

3.2.4 Discussion

A plethora of reports have highlighted that FAK's expression levels and its activation correlate with the initiation, progression and the prognosis of a wide variety of malignancies and, thus, FAK has been repeatedly proposed as an attractive target for cancer therapy^{187-189, 191, 192, 204}. Subsequently, a number of laboratories and pharmaceutical companies have invested into the development of specific FAK inhibitors which are now evaluated in pre-clinical animal models and in first clinical trials^{191, 204-206}. However, the actual mode of action of FAK in tumor progression and metastasis and the biological consequences of its pharmacological inhibition are still poorly understood.

Here, we have used the novel and highly specific small molecule FAK inhibitor (FAK-I) BI 853520¹⁴⁴ to repress FAK activity in cultured breast cancer cells *in vitro* and in various preclinical mouse models of breast cancer (Py2T, 4T1, MTfIECad/MTΔEcad, and MMTV-PyMT) *in vivo*. BI 853520 elicits a substantial repression of cancer cell proliferation and thus a repression of primary tumor growth and outgrowth of metastatic lesions. However, BI 853520 does not diminish the actual metastatic dissemination of murine breast cancer cells, in contrast we observe an increased number of pulmonary metastases upon treatment with BI 853520 in an adjuvant setting after surgical resection of 4T1 primary tumors.

On the other hand, BI 853520 most effectively hampers the establishment of pulmonary metastasis in mice harboring E-cadherin-deficient (MTΔEcad) tumors (**Fig. 6C**), the metastatic outgrowth of 4T1 breast cancer cells with downregulated E-cadherin expression (**Fig. 2A**), and the 3D Matrigel invasion of mesenchymal Py2T-LT cells (**Fig. 1F; Suppl. Fig. 2**). Shapiro and colleagues have linked FAK-I susceptibility of mesothelioma growth to merlin deficiency²¹⁶. However, the Py2T, MTfIECad and MTΔEcad cells employed here express similar levels of merlin, and the susceptibility of Py2T and MTfIECad tumors and the insensitivity of MTΔEcad tumors to BI 853520 appears to be independent of merlin expression status (data not shown). Rather, our results suggest that the lack or low expression of E-cadherin in 4T1 and MTΔEcad cells and in later stages of MMTV-PyMT tumor progression, and thus the invasive nature of the cancer cells, may serve as a molecular biomarker for an anti-metastatic activity of BI 853520. Intriguingly, this notion stands in stark contrast to previous studies, where FAK inhibition activates E-cadherin function and the reversal of an EMT, thereby restraining the initiation and establishment of proliferative programs

by micrometastatic cell clusters and not affecting macrometastatic expansion²¹⁷⁻²²². However, our histopathological analysis and gene expression profiling of primary tumors treated or not with BI 853520 have not revealed marked changes in the EMT status of the breast cancer cells. Hence, further investigations are warranted to decipher whether E-cadherin expression can be used as a potential predictive marker for the subset of patients that will respond to anti-metastatic FAK-I therapy^{188, 202, 204, 205}. Notably, Hirt *et al.* have recently reported that the expression of hsa-miR-200c-3p, an epithelial-specific miRNA efficiently repressing EMT transcription factors and thus promoting E-cadherin expression, strongly correlated with BI 853520 therapeutic efficacy in a number of transplantation mouse models of various cancer types¹⁴⁴. Confronted with these seemingly conflicting results, we hypothesize that intermediate stages of an EMT, such as a partial EMT with a combination of epithelial and mesenchymal marker expression, may represent a cell status of high cell plasticity and metastasis and thus may be preferentially targeted by FAK inhibition.

We have subsequently explored the direct biological effects of FAK inhibition on tumor cells and characterized potential mechanisms responsible for the significant decrease in tumor burden following BI 853520 administration. RNA sequencing of primary murine breast tumors treated with BI 853520 reveals a potent repression of tumor cell proliferation and thus primary tumor growth. Gene set enrichment analysis confirms a significant reduction in the relative expression of genes important for a positive regulation of mitotic cell cycle (including cyclin-dependent kinase 1 and 4) and an increase in the expression of genes relevant for a repression of the cell cycle. Previous studies have indicated that ectopic expression of a constitutive-active form of FAK leads to a rapid induction of proliferation in murine mammary tumor cells, whereas a decrease in FAK expression impairs their cell proliferation *in vitro* and *in vivo*^{191, 195, 201, 223}. In line with these findings, BI 853520 treatment elicits a significant reduction in murine breast cancer cell proliferation *in vitro* and *in vivo*. These findings are supported by flow cytometry-based cell cycle analysis where increasing concentrations of BI 853520 up to 1 μ M shift cells from the S-phase to the G1-phase. Together, the data suggest a linear mechanistic relationship between FAK activity and murine breast cancer cell proliferation.

Previously, it has been reported that tumor cells, that have been genetically ablated for FAK expression or treated with various FAK-I, undergo cell death^{191, 201, 204, 224}. However, using immunofluorescence analysis of primary tumor sections and

cellular assays we have not observed a significant induction of apoptosis following BI 853520 treatment. We conclude that the most prominent therapeutic effect observed by BI 853520 is based on its anti-proliferative activities. Along these lines, the results reported here also indicate that the kinase activity of FAK is required for 4T1 murine breast cancer cell pulmonary metastasis: the pharmacological inhibition of FAK kinase activity by BI 853520 is consistent with the previously reported findings that the kinase domain of FAK is required for metastatic outgrowth of 4T1 cells ²²⁵.

Although ongoing clinical trials with various FAK-I seem encouraging and methods for patient stratification are planned, it is questionable whether FAK inhibition will be used as a single-agent therapy as it is the case for most targeted therapies ^{197, 199, 204}. Previous studies have shown that repression of FAK signaling in combination with diverse chemotherapies enhanced chemo-cytotoxicity *in vitro* and *in vivo* ²²⁶⁻²³⁶. In line with FAK's role in an anti-tumor evasion mechanism and an upregulated immune response following BI 853520 treatment in the 4T1 breast cancer model (**Fig. 3B**), it has been shown in mouse models of squamous cell cancer and pancreatic ductal adenocarcinoma that FAK inhibition increases immune surveillance, enhances the sensitivity to T-cell immunotherapy and treatment with PD-1 antagonists, and ultimately delays tumor progression ²³⁷⁻²³⁹. Future efforts should be made to test BI 853520 and other FAK-I in combination with chemo- or immunotherapy reagents, thereby evaluating not only the anti-proliferative effects but also the anti-metastatic effects of a combination therapy that might be applicable for various cancer types ^{188, 206, 224, 240}.

3.2.5 Materials and Methods

Treatment of tumor transplantation and transgenic mouse models

The murine metastatic cell line 4T1 is described elsewhere ^{211, 241}. Py2T cells have been derived from a breast tumor of an MMTV-PyMT female mouse with an FVB/N background ²⁰⁸. E-cadherin-proficient MTflECad cells have been isolated from a lymph node metastasis of a MMTV-Neu;*Cdh1fl/fl* tumor-bearing mouse and have been infected with an adenovirus expressing Cre recombinase to obtain the E-cadherin-deficient cell line MTΔECad ²¹². All cell lines were maintained in Dulbecco's Modified Eagle Medium (DMEM), supplemented with fetal calf serum (FCS, 10%), glutamine (2mM), penicillin (100U; Sigma-Aldrich) and streptomycin (0.2mg/l; Sigma-Aldrich) in a humidified atmosphere of 37°C, 5% CO₂ and 95% humidity.

0.5 x 10⁶ Py2T, MTflECad or MTΔECad cells or 1 x 10⁶ 4T1 cells suspended in 100µl sterile PBS were injected into mammary fat pad number nine of anaesthetized, six to ten week-old female FVB/N (Py2T cells), BALB/c (4T1 cells), Rag2^{-/-};γc^{-/-} or nude mice (BALB/cAnN-*Foxn1nu/nu*/Rj; Janvier) mice (MTflECad and MTΔECad cells). The length (D) and width (d) of the tumors was measured by a digital caliper every two to three days and the tumor volume was calculated according the formula ($V = 0.543 \times D \times d^2$). Treatment protocols were initiated several days after tumor-cell inoculation, details are indicated in the respective figure legends. BI 853520¹⁴⁴ was dissolved in 0.5% Hydroxyethylcellulose Natrosol and administered at 50mg/kg body weight daily by oral gavage.

The MMTV-PyMT transgenic mouse model of metastatic breast cancer was used to determine the therapeutic and biological effect of BI 853520^{210, 242, 243}. BI 853520 was administered orally each day to mice of five weeks of age for eight consecutive weeks, referred to as “early BI 853520” treatment. In another cohort of mice, BI 853520 was administered orally each day to mice of ten weeks of age for three consecutive weeks, referred to as “late BI 853520” treatment. All mice of this experimental setup were sacrificed at 13 weeks of age and primary tumors and lungs were used for further processing.

For RNA sequencing, BI 853520 was administered from day 15 post injection for five consecutive days. To determine whether BI 853520 pre-treatment is able to prime the tumor microenvironment, mice were either pre-treated three days prior tumor cell inoculation (Vehicle PRE or BI 853520 PRE group) or BI 853520 was administered seven (subcutaneous tumor cell injection) or three days (intravenous tumor cell injection) post tumor cell injection (Vehicle POST or BI 853520 POST group).

For intravenous injection, 0.5 x 10⁶ mouse 4T1 breast cancer cells were injected into the tail vein of BALB/c mice.

To mimic the patient situation, BI 853520 was administered in an adjuvant setting to the preclinical 4T1 breast cancer mouse model. For this purpose, primary tumors were surgically removed 14 days after implantation of 4T1 cells in to the mammary fat pad, and BI 853520 therapy was started two days post-surgery for the consecutive 22 days.

For neoadjuvant BI 853520 treatment, BI 853520 administration was initiated one day after orthotopic tumor cell implantation (1 x 10⁶ 4T1 cells mixed with Matrigel)

until primary tumor removal on days 14 or 15 post tumor cell inoculation. Lung metastases were analyzed 32 days post tumor cell inoculation.

When not specified otherwise, animals were sacrificed when the maximum allowed primary tumor size of 1.5 cm³ was reached. For all experiments, primary tumors, lungs and peripheral blood were isolated and used for further processing as previously described^{208, 244}. Primary tumors in the abdominal mammary fat pad of MMTV-PyMT mice were graded by using a Zeiss Axio Imager.Z2 microscope and ZEN 2 Zeiss Microscopy software (version 2.0). Pulmonary cross sections were imaged with a 10 x objective of a Zeiss Axio Imager.Z2 or a Zeiss Axio Observer (Zeiss) microscope and the average amount of unique pulmonary metastases as well as the average metastatic area fraction (%) was quantified by using ImageJ software (ImageJ, Wayne Rasband, National Institutes of Health, USA), ZEN 2 Zeiss Microscopy software or a customized application from Visiopharm. All animal experiments were approved by and performed according to the guidelines and legislation of the Swiss Federal Veterinary Office and the Cantonal Veterinary Office, Basel-Stadt, Switzerland (permit numbers 1878, 1907 and 1908).

BI 853520 treatment of breast cancer cells in vitro

For immunoblotting analysis, immunofluorescence staining (for phospho-H3 and cleaved caspase 3) and EdU/PI cell cycle analysis, 1.3 x 10⁵ 4T1 cells were plated in a 6cm dish. For immunofluorescence staining (for phospho-Y397-FAK and total FAK), 3.3 x 10⁴ 4T1 cells were plated in a 6cm dish. For cell viability determination following 96 hours treatment with BI 853520, 0.5 x 10³ 4T1-, Py2T- and Py2T-LT cells were seeded in a 96-well format. 4T1, Py2T and Py2T-LT cells were treated for the indicated time points with varying concentrations of BI 853520.

Immunoblotting analysis

Tumor pieces were snap-frozen and lysates were used for immunoblotting analysis. Snap-frozen Py2T primary tumor pieces were homogenized by a Polytron PT1200 E (Kinematica AG) in lysis buffer (50mM HEPES pH7.5, 150mM NaCl, 1% Glycerol, 1% Triton X-100, 1.5mM MgCl₂, 5mM EGTA) supplemented with 0.1mM Sodium-Orthovanadate (Sigma-Aldrich), 20mM sodium fluoride (Sigma-Aldrich) and a protease inhibitor cocktail (Sigma-Aldrich, 1:200). The lysates were cleared by centrifugation and the protein content of supernatants was determined by using Protein Assay Dye

Reagent Concentrate (Bio-Rad). Snap-frozen 4T1 primary tumor pieces were homogenized by metal bead lysing matrix (MP Biomedicals) in RIPA buffer (Sigma-Aldrich, 150mM NaCl, 1.0% IGEPAL® CA-630, 0.5% sodium deoxycholate, 0.1% SDS, 50mM Tris, pH 8.0) supplemented with 20mM sodium fluoride (Sigma-Aldrich) and a protease inhibitor cocktail (Sigma-Aldrich, 1:100). 4T1, Py2T and Py2T-LT cells previously being treated with BI 853520 were lysed in sample solution (0.5M Tris-HCl pH 6.8, 10% SDS, glycerol) or NP40 buffer (20mM Tris pH 7.5, 100mM NaCl, 5mM MgCl₂, 0.2% NP40, 10% glycerol, 1mM NaF, 20mM beta-glycerophosphate and freshly added 0.5mM DTT /1x PIC (Sigma) /1mM Vanadate). The protein content was determined by Pierce BCA Protein Assay Kit (Thermo Scientific), loaded on 10% SDS-PAGE gels and subsequently used for immunoblotting analysis. Following blocking of the nitrocellulose blotting membranes (GE Healthcare) for one hour in 5% bovine serum albumin (BSA; Sigma-Aldrich) dissolved in 0.1% Tween (Sigma-Aldrich) in TBS (for Py2T tumors) or PBS (for 4T1 tumors), membranes were incubated with primary antibodies mouse anti-FAK (Millipore, 05-537, clone 4.47), rabbit anti-pFAK (Y397, Invitrogen, 44-624G), rabbit anti-PYK2 (Abcam, ab32571, clone YE353), rabbit anti-pPYK2 (Y402, Abcam, ab4800) overnight at 4°C. GAPDH (Abcam, ab9485 or Sigma-Aldrich, G8795) or α -tubulin (Sigma-Aldrich, T-9026, clone DM1A) were used as a loading control. Membranes were then incubated for one hour with HRP-conjugated secondary antibodies (Jackson) and subsequently visualized using Fusion Fx chemoluminescence reader (Vilber Lourmat) and Curix 60 (AGFA). For Py2T tumors, following visualization of the pFAK signal, antibodies were removed by using a stripping buffer (0.2M glycine pH 2.5, 0.1% SDS, 1% Tween) two times 30 minutes before initiating the staining procedure against total FAK protein.

Immunofluorescence staining of frozen sections

To prepare cryostat sections, organs were fixed for two hours at 4°C in 4% paraformaldehyde, incubated overnight in 20% sucrose in PBS (Sigma-Aldrich) at 4°C and then embedded and snap frozen in OCT freezing matrix (Tissue Tek). For immunofluorescence analysis of frozen sections, cryosections (7 μ m thick) of tumor samples were dried for 30 minutes and rehydrated three times for five minutes in PBS. After permeabilization for 20 minutes with 0.2% Triton X-100 (Sigma-Aldrich) in PBS, slides were washed three times for five minutes in PBS, blocked for one hour at room temperature in 5% goat serum diluted in PBS (blocking buffer) for the phospho-H3

staining and 20% goat serum diluted in PBS for the cleaved caspase-3 staining, and incubated overnight at 4°C with primary antibodies (rabbit anti-mouse phospho-H3, Millipore, 06-570; rabbit anti-mouse cleaved caspase-3, Cell Signaling, 9664, clone 5A1E) diluted in blocking buffer.

To detect perfused vessels, 150µg of fluorescein labeled lycopersicon esculentum lectin (Vector Laboratories, GL-1171) was injected i.v. and allowed to circulate for 20 minutes. Next, terminally anaesthetized animals were perfused with 4% PFA in PBS followed by PBS via the left cardiac ventricle.

For the detection of leaky blood vessels, 250 µg fluorescein-labeled Dextran (70kDa; Life Technologies, D-1822) was injected intravenously in a tail vein and allowed to circulate for five minutes. Then, terminally anaesthetized animals were perfused with PBS followed by a perfusion with 4% PFA in PBS via the left cardiac ventricle.

Hypoxic areas were detected by injecting pimonidazole-HCl (Hypoxyprobe Omni Kits, Hypoxyprobe, Inc.) at 60mg/kg intraperitoneally two hours prior to sacrifice of the animals.

Immunofluorescence stainings were revealed by incubating slides for one hour at room temperature with Alexa Fluor 488, Alexa Fluor 568, Alexa Fluor 633 or Alexa Fluor 647-labeled secondary antibodies (Invitrogen, 1:200) diluted in blocking buffer and washed three times for five minutes in PBS. Nuclei were counterstained for ten minutes with 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich, 1:5000) and slides were washed three times for five minutes in PBS. Finally, slides were mounted in DAKO fluorescence mounting medium and evaluated with a 20 x objective on a DMI 4000 microscope (Leica). Images were processed and analyzed using ImageJ software (ImageJ, Wayne Rasband, National Institutes of Health, USA).

Immunofluorescence staining of cultured cells

4T1 cells grown on coverslips were fixed for 20 minutes with 4% PFA in PBS at room temperature. Subsequently, cells were permeabilized for ten minutes with 0.1% Triton X-100 (Sigma-Aldrich) in PBS and blocked with 1% BSA (Sigma-Aldrich) in PBS for two hours. To visualize FAK and pFAK expression, cells were incubated with the primary antibodies mouse anti-FAK (Millipore, 05-537, clone 4.47), rabbit anti-pFAK (Y397, Invitrogen, 44-624G), rabbit anti-mouse phospho H3 (Millipore, 06-570), rabbit anti-mouse cleaved caspase-3 (Cell Signaling, 9664, clone 5A1E) diluted in 1% BSA

in PBS overnight at 4°C, followed by a 30 minutes incubation with DAPI (Sigma-Aldrich, 1:5000) and Alexa-Fluor 488 or 568-labeled secondary antibodies (Invitrogen, 1:200). The cells were sealed by mounting them with Dako fluorescence mounting medium. Images were obtained using a Confocal SP5 microscope with a 63 x objective and a DFC 360 FX Monochromatic Camera (Leica) or with a 10 x (cleaved caspase 3) or 20 x objective (pH3) on a DMI 4000 microscope (Leica).

In-gel immunofluorescence staining

Growth factor-reduced Matrigel (Corning) was diluted with ice-cold, serum-free growth medium (F12-HAM, Sigma-Aldrich) to 4 mg/ml protein concentration by gentle pipetting. Pellets of 1.5×10^3 4T1 cells and 2.5×10^3 Py2T and Py2T-LT cells, respectively, were resuspended in 10 μ l 4mg/ml Matrigel and seeded into each inner well of a μ -slide angiogenesis microscopy slide (ibidi, Martinsried, Germany). Following 20 minutes of gel solidification in a humidified atmosphere at 37°C and 5% CO₂, 30 μ l of normal growth medium (supplemented or not with 2ng/ml TGF β) was added to each well. Growth medium was refreshed every third day. After five days of culture, medium containing increasing concentrations of BI 853520 (0.01 μ M, 0.1 μ M, 1 μ M) was added. On the next day, structures were fixed in the matrix for 10 minutes with 4% PFA in PBS, washed for 5 minutes with 20mM glycine/PBS and permeabilized and blocked for two hours at room temperature in IF buffer (0.2% TritonX-100/0.1% BSA/0.05% Tween20/PBS) supplemented with 10% goat serum. Samples were stained with primary antibodies in IF buffer (rabbit anti-pFAK (Y397, Invitrogen, 44-624G); rabbit anti-ZO-1 (Zymed, 617300); mouse anti-vimentin (Sigma-Aldrich, V2258)) for two hours at room temperature, washed twice with IF buffer and incubated with Alexa-Fluor 488 or 568-labeled secondary antibodies diluted in IF buffer (Invitrogen, 1:200) for 45 minutes. Nuclei were stained using DAPI (Sigma-Aldrich, 1:5000) for 20 minutes at room temperature. Following two washes with IF buffer, samples were mounted with DAKO fluorescent mounting medium and imaged using a confocal microscope (SP5 confocal microscope with a 63 x objective, Leica).

Cell proliferation assay

The effects of varying BI 853520 concentrations on cell proliferation were determined by an MTS-assay²⁰⁹ (Promega) according to the manufacturer's protocol. Cells were seeded, and one day later varying concentrations of BI 853520 were added and cell

viability was measured on day five. Cell viability was measured using a SpectraMax 340 Microplate Reader (Molecular Devices) with a filter of 492 nm. For analysis, background absorbance from a “no-cell” control row was subtracted from all measured values and cell number of BI 853520-treated cells was normalized to the DMSO-treated controls.

Cell cycle analysis

For Edu/PI-incorporation cell cycle analysis, 4T1 cells were treated for 24 hours with varying concentrations of BI 853520, cells were then incubated for five minutes with 10 μ M EdU (EdU Flow Cytometry Kit, baseclick). Harvested cells were subsequently fixed for 15 minutes with 4% PFA in PBS at room temperature and permeabilized for 20 minutes using 0.5% Triton X-100 (Sigma-Aldrich) in PBS at room temperature. Following a 30 minutes incubation with the click reaction containing the catalyst solution, the Dye Azide, buffer additive and PBS, cells were stained with propidium iodide (Sigma-Aldrich) for two hours at a humidified atmosphere of 37°C and 5% CO₂. Cell cycle distribution was analyzed by using a BD FACSCANTO II analyzer and BD FACSDIVA software (BD Biosciences).

Soft agar colony formation assay

A bottom layer of 1.5ml 0.5% agarose/DMEM complete growth medium per 6-well plate with 2ng/ml TGF β and varying concentrations of BI 853520 was allowed to solidify for 30 minutes at room temperature. 3 ml of 0.3% agarose/DMEM complete growth medium with 2x10⁴ Py2T-LT cells, 2ng/ml TGF β and the target BI 853520 concentration were added on top and incubated for 1.5 hours in a tissue culture incubator. Finally, 1 ml DMEM complete growth medium with 2ng/ml TGF β and the target BI 853520 concentration was added on top. Every 4-6 days, 0.5 ml DMEM complete growth medium supplemented with 2ng/ml TGF β and the target BI 853520 concentration was refreshed. After 18-21 days, colonies were stained with MTT solution (Sigma-Aldrich) and quantified.

Apoptosis assay

2 x 10⁵ Py2T and Py2T-LT cells (constantly cultured in the presence of 2ng/ml TGF β) were plated per 60 mm petri dish and allowed to adhere overnight. The next day, BI 853520 at different concentrations was added. After 24 hours of incubation with BI

853520, cells were trypsinized, stained with Cy5 Annexin-V (BD Biosciences) and analyzed by flow cytometry on a BD FACSCANTO II analyzer and BD FACSDIVA software (BD Biosciences). Apoptotic cells were defined as Annexin-V positive.

Matrigel invasion assay

Growth factor-reduced Matrigel (Corning) was diluted with ice-cold serum-free DMEM (Sigma-Aldrich) by gentle pipetting using pre-chilled pipette tips to 4mg/ml protein (100%) concentration. 10 μ l Matrigel was plated into each inner well of a μ -slide angiogenesis microscopy slide (ibidi, Martinsried, Germany). The gel was allowed to solidify for 45 minutes in a humidified atmosphere at 37 °C and 5% CO₂. 0.3 x 10³ 4T1 cells and 1.25 x 10³ Py2T epithelial murine breast cancer cells and Py2T-LT mesenchymal cells, respectively, were seeded on the gel, and then a top layer of growth medium with 10% Matrigel, in the case for Py2T and Py2T-LT supplemented with twice the final concentration of BI 853520 and 4ng/ml TGF β (for Py2T-LT), was added. Following one day of incubation in a humidified atmosphere at 37 °C and 5% CO₂, twice the final concentration of BI 853520 was also added to the 4T1 cells. Medium containing 5% Matrigel and BI 853520, and in the case of Py2T-LT cells TGF β , was replaced after 3 days and after 6 days pictures were taken using a Leica DMIL microscope.

RNA isolation, RNA-sequencing (RNA-seq) library preparation, sequencing and data analysis

Snap-frozen mouse tumor pieces were homogenized in Trizol (Sigma-Aldrich) and total RNA was isolated with the RNeasy Lipid Tissue Mini Kit (Qiagen). Approximately 80ng of total RNA was subsequently used for TruSeq Stranded mRNA Library preparation on the NeoPrep System (Illumina, San Diego, CA, USA) and sequenced on the HiSeq1500 with the paired-end 50 cycle protocol and the fast-output mode (Illumina, San Diego, CA, USA). The RNA sequencing data is available under GEO accession number GSE116117.

Paired-end RNA-seq reads were mapped to the mouse genome assembly, version mm9, with RNA-STAR¹⁷⁷, with default parameters except that only unique hits were aligned to the genome (outFilterMultimapNmax=1) and that reads without evidence in spliced junction table were filtered (outFilterType="BySJout"). Using RefSeq mRNA coordinates from UCSC (genome.ucsc.edu, downloaded in July 2013)

and the qCount function from QuasR package (version 3.12.1)¹⁷⁸, gene expression was quantified as the number of reads that started within any annotated exon of a gene. The differentially expressed genes were identified using the edgeR package¹⁷⁹. Genes with a p-value smaller than 0.05 and a minimum log2 fold change of +/- 0.584 were considered statistically significant and these genes were subjected to downstream functional analysis.

Functional enrichment analysis

Functional enrichment analysis of differentially expressed genes for biological processes²⁴⁵ or pathways was performed in R using several publically available Bioconductor resources including GO.db (version 3.4.1), GOstats (version 2.42.0)²⁴⁶, KEGG.db (version 3.2.3) and ReactomePA (version 1.20.2)¹⁸⁰. The significance of each biological processes or pathways identified was calculated using the hypergeometric test (equivalent to Fisher's exact test). p values ≤ 0.05 were considered statistically significant. Clustering and heatmaps of the enriched biological processes were generated using hierarchical clustering and distances between the clusters were computed using the average linkage method.

Statistical analysis

Graphics and statistical analysis were performed using the GraphPad Prism Software Version 7.0. When not specified otherwise, the unpaired, two-tailed Student's t-test or Mann-Whitney U test were applied (****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$) to determine statistical significance. Quantitative data were depicted as means \pm standard error of the mean (SEM).

Conflict of interest

Part of this work has been supported by Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria. Irene C. Waizenegger, Norbert Schweifer and Norbert Kraut are employees of Boehringer Ingelheim Austria GmbH, Vienna, Austria. All other authors declare no conflict of interest.

Data availability

The RNA sequencing data is available under Gene Expression Omnibus (GEO) accession number GSE116117.

Authors' contributions

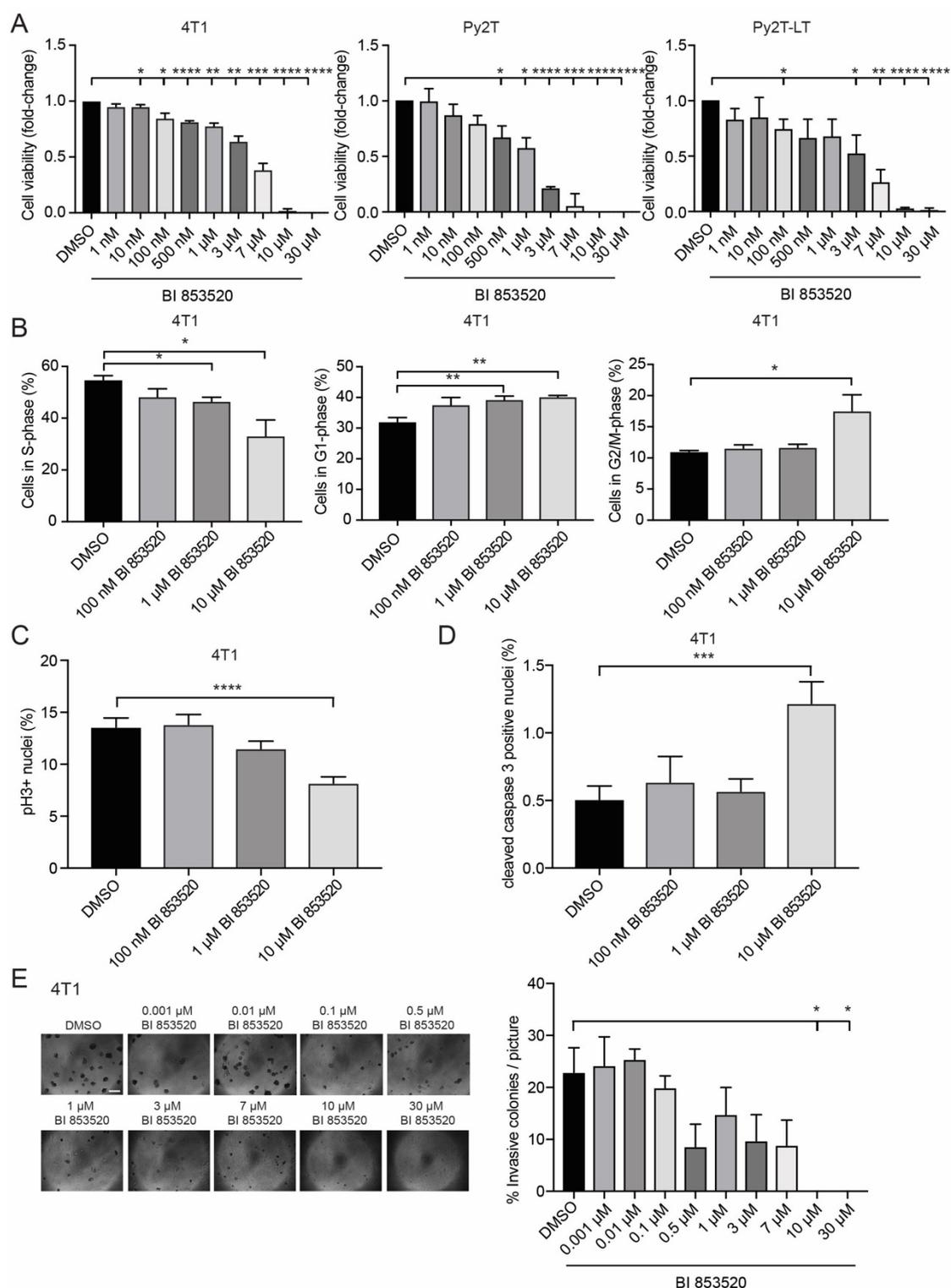
Conceived and designed the experiments: ST RB GC. Performed the experiments: ST PS RB. Analyzed the data: ST NM PS PST RK RI EF SH RB GC. Conducted RNA isolation of primary tumors and performed RNA sequencing: Boehringer Ingelheim, NS. Contributed reagents/materials/analysis tools: ST NM PS PST RK RI EF IW NK RB GC. Concept/design, review, discussion: IW NK. Wrote the paper: ST IW NK RB GC.

Acknowledgments

We are grateful to H. Antoniadis, I. Galm and U. Schmieder, T. Bürglin for their technical support and P. Lorentz and the DBM microscopy facility for expert help with microscopy. We thank F. Hilberg for critical comments on the manuscript. We greatly appreciate the support by the collaborative work with Boehringer Ingelheim.

This work was supported by the Swiss National Science Foundation and by Boehringer Ingelheim RCV GmbH & Co KG, Vienna, Austria. N.M-S was supported by a Marie-Heim Vögtlin grant from the Swiss National Foundation. R.B. was supported by a MD-PhD fellowship of the Swiss National Science Foundation and Swiss Cancer Research.

3.2.6 Supplemental information

**Supplementary Figure 1. Anti-proliferative effects of BI 853520 on 4T1 murine breast cancer cells *in vitro*.**

(A) Left panel: 4T1 breast cancer cells were exposed to increasing concentrations of BI 853520 one day after cell seeding and cell viability was measured at day five ($n = 3$). Middle panel: Py2T breast cancer cells were exposed to increasing concentrations of BI 853520 one day after cell seeding and cell viability was measured at day five ($n = 3$). Right panel: Py2T-LT breast cancer cells were exposed to increasing concentrations of BI 853520 one day after cell seeding and cell viability was measured at day five ($n = 3$).

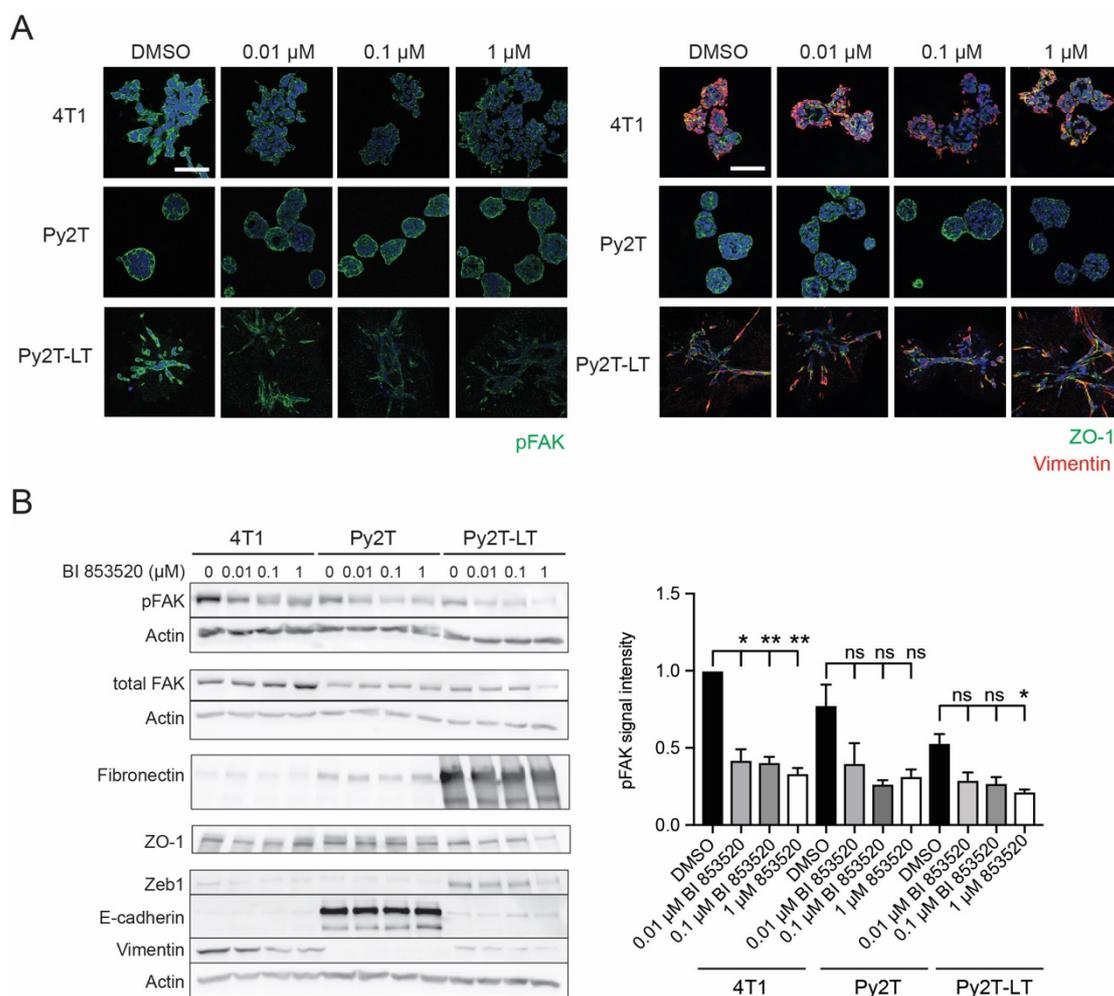
(B) 4T1 cells were treated for 24 hours with varying concentrations of BI 853520, incubated for five minutes with 10 μ M EdU and then subjected to EdU/Pi cell cycle flow cytometry analysis as described in Materials and Methods ($n = 4$). Cell numbers in S, G1 and G2/M-phases were quantified, the results are depicted in the three panels.

(C) 4T1 breast cancer cells were treated for 24 hours with varying concentrations of BI 853520 and then immunostained for phospho-histone 3 (pH3) and DAPI (nuclei). Quantification of the percentage of pH3-positive nuclei is shown, indicating a decrease in 4T1 cell proliferation (n = 4).

(D) 4T1 breast cancer cells were treated for 24 hours with varying concentrations of BI 853520 and then immunostained for cleaved caspase-3-positive nuclei. Quantification of the cleaved caspase-3-positive cells is shown, indicating no change in the rate of cell apoptosis except at the non-specific concentration of 10µM (n = 4). Statistical analyses were performed using an unpaired, two-tailed Student's t test.

All data are depicted as mean ± SEM. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.

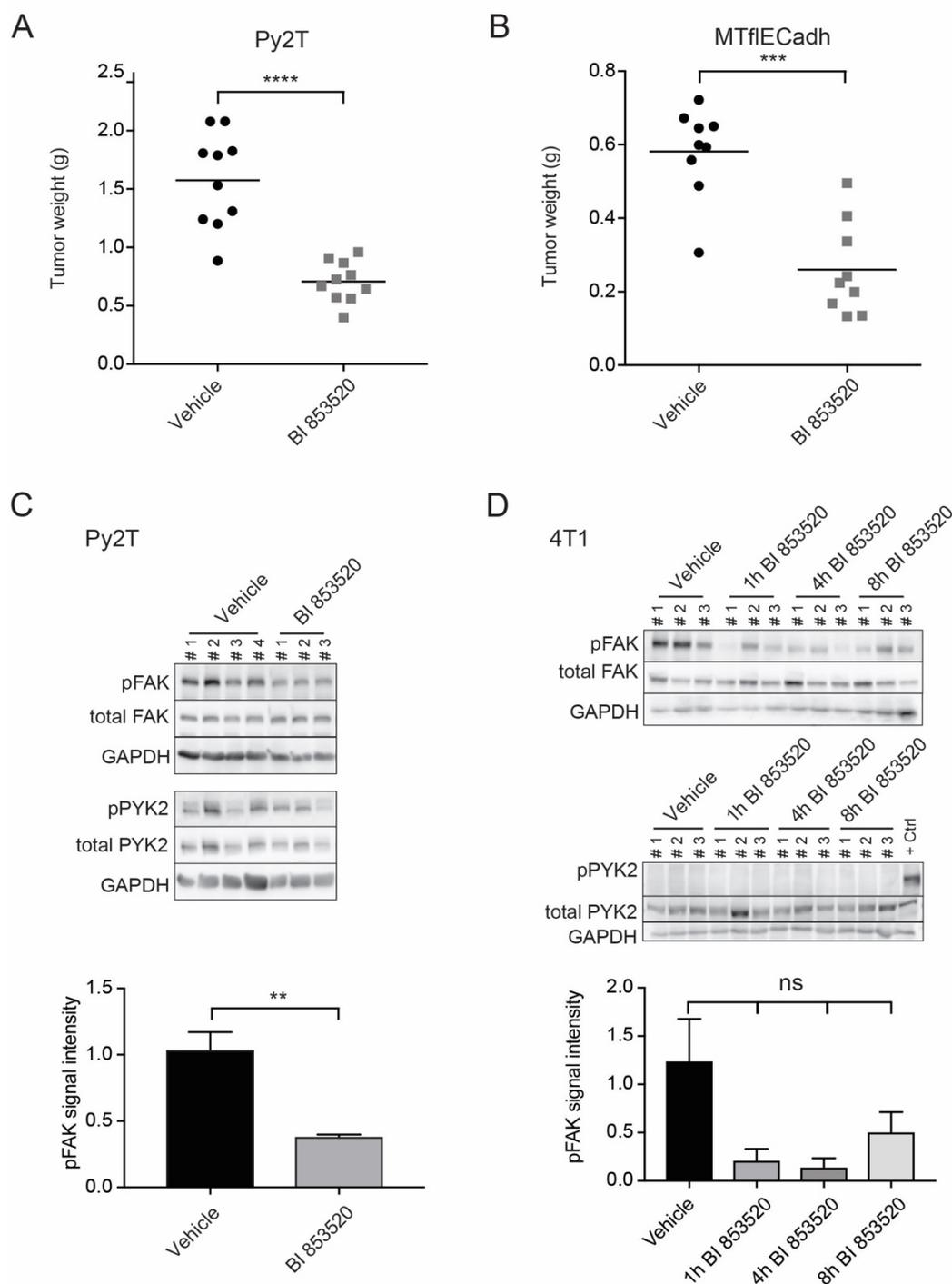
(E) BI 853520 represses 4T1 breast cancer cell invasion cultured in Matrigel. 4T1 breast cancer cells were seeded in Matrigel and treated with increasing concentrations of BI 853520 as indicated. Phase contrast microscopy pictures were taken and the numbers of invasive colonies per picture were quantified (n = 1). Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05. Scale bar, 250µm.



Supplementary Figure 2. BI 853520 decreases Y397-FAK phosphorylation in 3D Matrigel cultures in a dose-dependent manner.

(A) 4T1, Py2T and Py2T-LT cells grown in a 3D Matrigel culture (4mg/ml) were treated with increasing concentrations of BI 853520 (0.01µM, 0.1µM, 1µM) for 24 hours, followed by immunofluorescence staining for pFAK (Y397, green) and DNA (blue) (left panel). The 3D cultured cells were also stained for ZO-1 (green), vimentin (red) and DNA (blue) (right panel). Images were obtained using a SP5 laser-scanning confocal microscope (Leica) (n = 2). Scale bars, 30µm.

(B) 4T1, Py2T and Py2T-LT cells grown in a 5% 3D Matrigel culture were exposed to increasing concentrations of BI 853520 (0.01µM, 0.1µM, 1µM) for 24 hours (n = 2). Left panel: Cell lysates were examined by immunoblotting analysis for phospho-FAK-Y397 (pFAK) and total FAK (FAK), fibronectin, ZO-1, Zeb1, E-cadherin and vimentin. Actin was used as a loading control. Right panel: Quantification of the pFAK signal intensities normalized to actin and total FAK/actin ratios from the immunoblotting analysis shown in the left panel. Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.



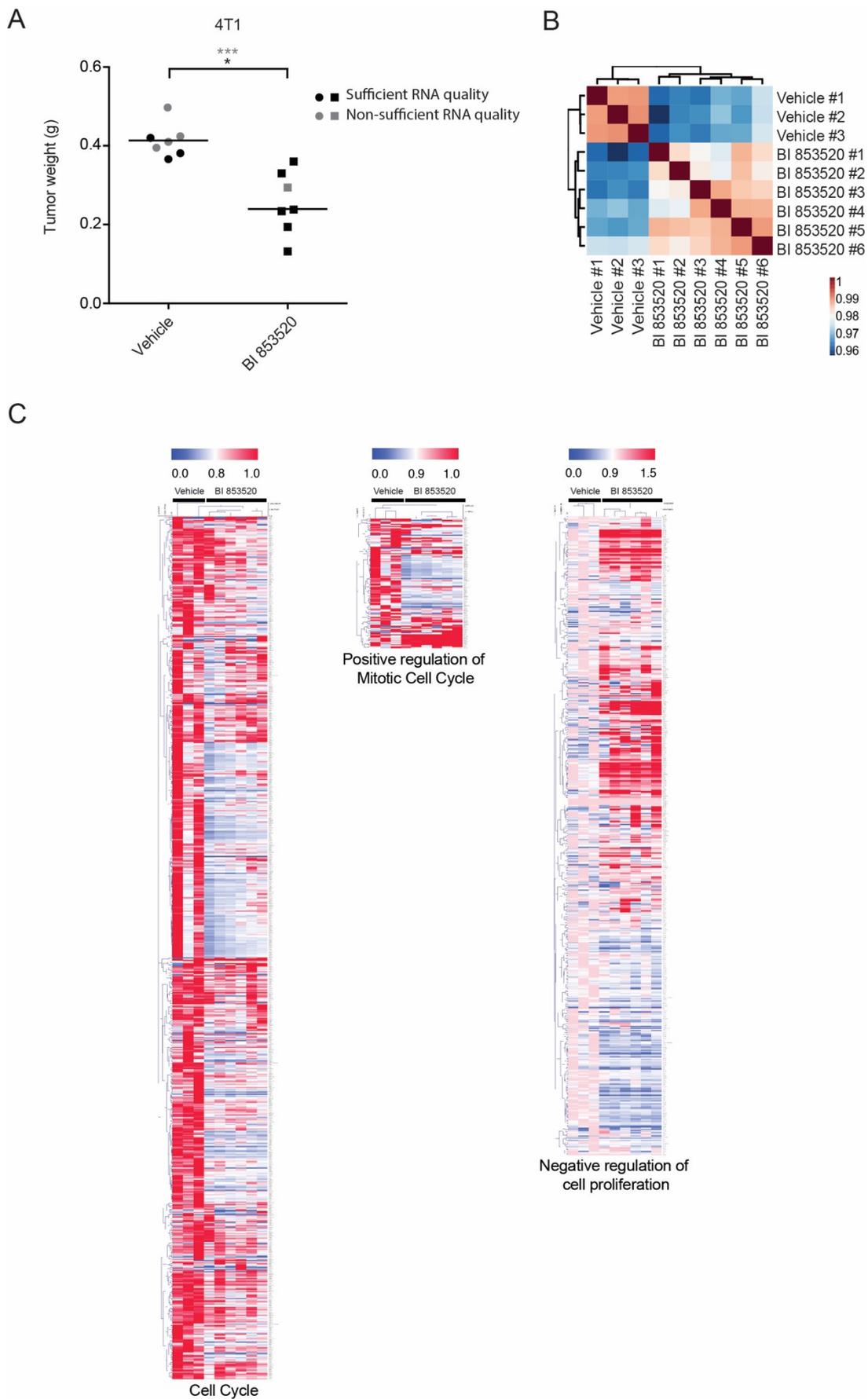
Supplementary Figure 3. BI 853520 reduces primary tumor in different preclinical breast cancer mouse models and delays malignant tumor progression.

(A) Treatment of mice bearing orthotopically transplanted Py2T breast tumors with 50mg/kg BI 853520 daily for 25 days significantly decreases tumor weight at the experimental end point. $n = 10$ mice per treatment cohort. Statistical analysis was performed by the unpaired, two-tailed Mann-Whitney U test.

(B) Treatment of mice bearing orthotopically transplanted MTfIECad breast tumors with 50 mg/kg BI 853520 daily significantly decreases tumor weight. $n = 9$ mice per treatment group. Statistical analysis was performed by the unpaired, two-tailed Mann-Whitney U test.

(C) Three-day daily 50mg/kg BI 853520-treated Py2T primary tumor pieces were snap-frozen 20h post last treatment. Top panel: Lysates were analyzed by immunoblotting analysis for phospho-FAK (pFAK, Y397), total FAK, phospho-PYK2 (pPYK2, Y402), total PYK2 and GAPDH as loading control. Bottom panel: Quantification of pFAK signal intensity normalized to GAPDH and total FAK/GAPDH ratios from the immunoblotting analysis shown above. Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

(D) Single dose 50mg/kg BI 853520-treated 4T1 primary tumor pieces were snap-frozen 1, 4 and 8 hours post treatment. Top panel: Lysates were analyzed by immunoblotting analysis for phospho-FAK (pFAK, Y397), total FAK, phospho-PYK2 (pPYK2, Y402), total PYK2 and GAPDH as loading control. Bottom panel: Quantification of pFAK signal intensity normalized to GAPDH and total FAK/GAPDH ratios from the immunoblotting analysis shown above. Statistical analysis was performed using an unpaired, two-tailed Student's t test. ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

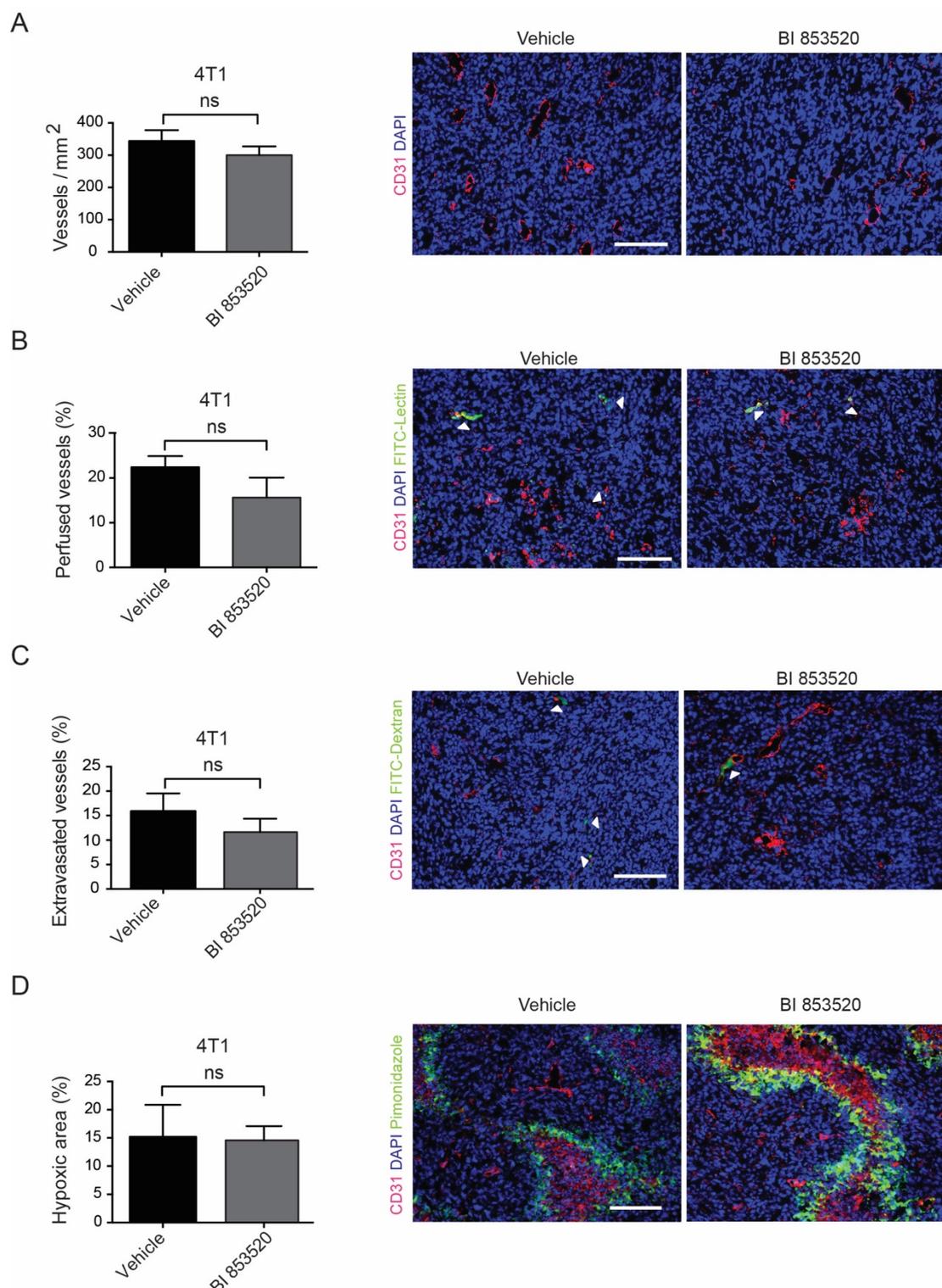


Supplementary Figure 4. Gene expression profiling reveals that BI 853520 represses murine breast cancer cell proliferation.

(A) BALB/c mice were implanted with 4T1 murine breast cancer cells and once tumors had formed treated for 5 days with vehicle control or with 50mg/kg BI 853520 daily. Tumor volumes are shown after treatment period. RNA was then extracted from primary tumors, and RNA samples with adequate quality were subjected to RNA sequencing and gene expression profiling analysis (indicated with black dots/rectangles). Statistical analysis was performed using an unpaired, two-tailed Mann-Whitney U test of all tumors (grey asterisks) and separately of only the tumors with sufficient RNA-quality (black asterisk; n = 7, tumors in the control group; n = 7, tumors in the BI 853520 -treated group). ****, $p < 0.0001$; ***, $p < 0.001$; **, $p < 0.01$; *, $p < 0.05$.

(B) Correlation analysis of gene expression data obtained from RNA-sequencing followed by hierarchical clustering showed a distinct grouping of vehicle and BI 853520-treated primary tumors. The color codes represent correlation values.

(C) Heatmaps of genes with changes in gene expression after BI 853520 treatment: the terms “general cell cycle”, “positive regulation of mitotic cell cycle” and “negative regulation of cell proliferation” were identified by gene set enrichment analysis.



Supplementary Figure 5. BI 853520 does not regulate angiogenesis in a relevant manner *in vivo*.

(A) Immunofluorescence microscopy analysis of primary tumor sections of 4T1 breast tumor bearing mice treated with vehicle control or 50mg/kg daily BI 853520 and analyzed at the experimental end-point for DAPI (cell nuclei; blue) and CD31 (red). Quantification of the numbers of vessels per mm² is shown on the left, representative immunofluorescence images are shown on the right. n = 7 tumors in the vehicle group, n = 8 tumors in the BI 853520 treated group.

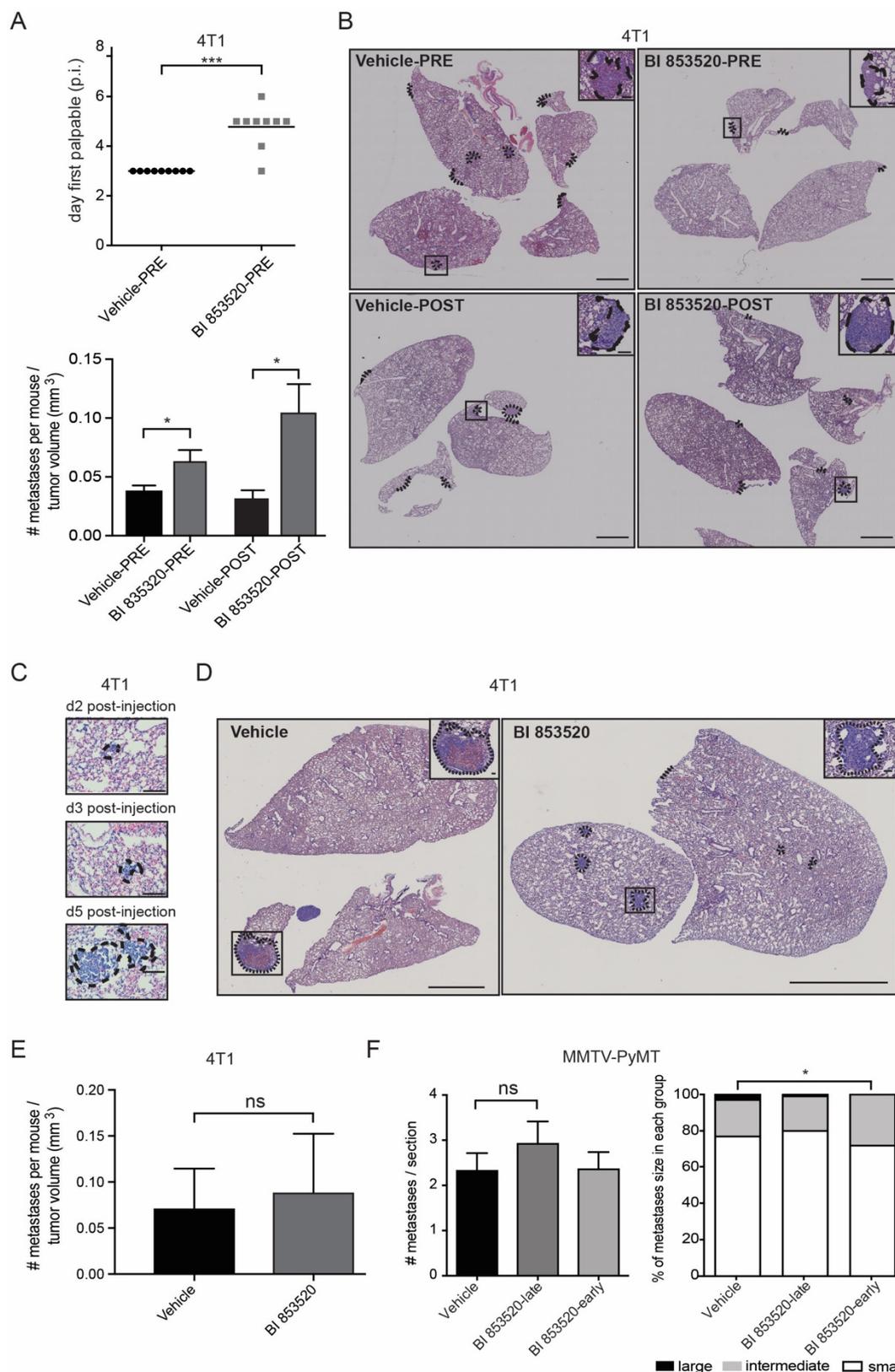
(B) Immunofluorescence microscopy analysis of primary tumor sections of 4T1 breast tumor bearing mice treated with vehicle control or 50mg/kg daily BI 853520 and analyzed at the experimental end-point for DAPI (cell nuclei; blue) and CD31 (red). Perfused vessels were FITC-lectin positive (green). Quantification of the percentage of lectin-positive vessels is shown in the left panel, representative immunofluorescence images are shown on the right. n = 2 tumors per group.

(C) Immunofluorescence staining of primary tumor sections of 4T1 breast tumor bearing mice treated with vehicle control or 50mg/kg daily BI 853520 and analyzed at the experimental end-point for DAPI (cell nuclei; blue) and

CD31 (red). Leaky vessels showed extravasated FITC-Dextran (green) in their vicinity. Quantification of the percentage of Dextran-positive vessels is shown in the left panel, representative immunofluorescence images are shown on the right. n = 2 tumors in the vehicle group, n = 3 tumors in the BI 853520 treated group.

(D) Immunofluorescence staining of primary tumor sections of the 4T1 breast tumor bearing mice treated with vehicle control or 50mg/kg daily BI 853520 and analyzed at the experimental end-point for DAPI (cell nuclei; blue), CD31 (red) and pimonidazole (green). Quantification of the percentage hypoxic area is shown in the left panel, representative immunofluorescence images are shown on the right. n = 4 tumors in the control group, n = 5 tumors in the BI 853520 treated group (two sections per tumor were analyzed).

Data indicate counts per field of view, shown as mean \pm SEM. Statistical difference was determined by the unpaired, two-tailed Student's t test. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05. Scale bars, 100 μ m.



Supplementary Figure 6. BI 853520 reduces pulmonary metastasis outgrowth.

(A) Tumor onset is significantly delayed in s.c. tumors pre-treated daily with 50mg/kg BI 853520 (palpable median five days post tumor cell injection) compared to vehicle pre-treated cohort (palpable three days post tumor cell injection). n = 9, Vehicle-PRE; n = 9, BI 853520-PRE. The ratio between the number of metastasis per mouse to the primary tumor volume represents the normalized metastatic index. Statistical analysis was performed using an unpaired, two-tailed Mann-Whitney U test.

(B) Mice were either pre-treated daily starting three days prior to 4T1 tumor cell inoculation (Vehicle-PRE or BI 853520-PRE; 50mg/kg BI 853520) or treated seven days post tumor cell injection (Vehicle-POST or BI 853520-POST).

POST groups) as described in Figure 5. Representative H&E staining of pulmonary histological sections are shown. Dashed lines indicate metastatic lesions. Scale bars, 2mm; scale bars in zoom-in, 100 μ m.

(C) Representative images of H&E-stained lungs following intravenous 4T1 injection. Cells appear to start proliferating three days post intravenous injection. Dashed lines indicate metastatic lesions. Scale bar, 100 μ m.

(D) Representative images of H&E stained pulmonary cross-sections of adjuvant BI 853520 therapy. Dashed lines indicate metastatic lesions. Scale bar, 2mm, scale bars in zoom-in, 100 μ m.

(E) Normalization of the number of metastasis to primary tumor growth (metastatic index) in mice transplanted with 4T1 cells and treated daily with 50mg/kg BI 853520 in a neoadjuvant setting until primary tumors were removed and therapy was discontinued. Numbers of metastasis were determined 18 days after primary tumor removal (**see also Fig. 5D**). The ratio between the number of metastasis per mouse to the primary tumor volume represents the normalized metastatic index. Statistical analysis by unpaired, two-tailed Student's t-test.

(F) BI 853520 therapy has no effect on the number of pulmonary metastases in the MMTV-PyMT transgenic mouse model treated with vehicle, late BI 853520 therapy or early BI 853520 therapy. n = 9, Vehicle; n = 9, BI 853520-late; n = 8, BI 853520-early (**see Fig. 2D-E**). Left panel: Numbers of metastases per section were quantified. Right panel: Quantification of the percentage of small, medium and large size individual lung metastatic lesions. Statistical analysis was performed using a Fisher's exact test. All data are depicted as mean \pm SEM. ****, p < 0.0001; ***, p < 0.001; **, p < 0.01; *, p < 0.05.

4. Concluding remarks and future perspectives

Despite recent advances in various sequencing techniques, we are unfortunately still far away from understanding and implementing the increasing extent of tumour heterogeneity to improve diagnosis and therapeutic options for breast cancer patients.

By establishing a pipeline of multicolour lineage tracing, laser capture microdissection and RNA-sequencing technology, we have shed light on the evolution of clonal populations in primary tumours and pulmonary secondary lesions in a mouse model of metastatic breast cancer. By longitudinal imaging, we could identify an initial clonal heterogeneity, which was followed by a clonal dominance and polyclonal seeding. However, the following question still remains: is this clonal sweep and polyclonal spread due to advantageous mutations and/or due to a co-option of various clonal populations? Paralleled genome and transcriptome sequencing could shed light on this question.

By comparative analysis of gene expression profiles, we have unravelled a substantial transcriptomic variation and clonal heterogeneity across the stages of breast cancer progression and particularly within each tumour stage. This intra-stage heterogeneity is marked by differences in proliferation, oxidative phosphorylation and cell death and could also be observed in human breast cancer biopsies. The imminent danger associated with intra-stage heterogeneity could be augmented robustness of the breast cancer lesions against chemotherapeutic reagents. Presumably, conventional chemotherapy will primarily eliminate the fast proliferating breast cancer clones. However, our new understanding of intra-stage heterogeneity points towards a survival benefit of the slow proliferating clones, which might foster therapeutic resistance and clinical progression. This is still hypothetical and might warrant further investigations in near future.

Though, one thing is certain: this polyclonal nature of breast cancer in “time and space” should be taken into consideration throughout clinical decision-making, which includes a patient’s diagnosis and potential treatment options. To mitigate the risk of sampling bias and to stratify patients for the right therapeutic regimen, heterogeneity should be evaluated on a whole tumour-level, by CTC analysis and by biopsies of accessible, secondary lesions.

Aside from deciphering tumour heterogeneity during breast cancer progression, we assessed the therapeutic potential and biological effects of the novel, small

Concluding remarks and future perspectives

molecule FAK-inhibitor BI 853520. Various *in vitro* studies and preclinical testing in several mouse models of breast cancer highlighted BI 853520 as a potent inhibitor with anti-tumour effects in breast cancer. Even though BI 853520 has accomplished clinical phase I studies in patients with progressed or metastatic nonhematologic cancers, it is uncertain whether BI 853520 will be efficient enough as a single-agent therapy. A combination of BI 853520 together with DNA-damaging therapies, such as chemotherapy, might enhance the breast cancer cell's chemosensitivity *in vitro* and *in vivo* and should be considered for future studies.

Last, taking these insights into account, the following question is sparking and needs to be addressed: would a combination of the anti-proliferative FAK-inhibitor BI 853520 together with chemotherapeutic regimens eradicate the whole tumour mass in our multi-colour mouse model? In other words: would a combination of a chemotherapy and the FAK-inhibitor BI 853520 target the fast proliferating clones as well as enhance the chemosensitivity of slow proliferating clones, respectively?

To adequately answer the above and other arising questions in the field of breast cancer heterogeneity, it is important to bundle all efforts in recent technologies (such as *in vivo* single cell barcoding, CRISPR/Cas9 screening and intravital imaging) to fully comprehend the functional contribution of tumour heterogeneity on progressing breast cancer. By going the extra mile, we might be able to improve diagnosis and to develop tailored, suitable treatment combinations to ultimately cure breast cancer.

5. Contribution to other projects

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

- Diepenbruck, M., **S. Tiede**, M. Saxena, R. Ivanek, R. K. R. Kalathur, F. Luond, N. Meyer-Schaller and G. Christofori (2017). "miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis." *Nat Commun* **8**(1): 1168. (<https://doi.org/10.1038/s41467-017-01197-w>)
- Meyer-Schaller, N., C. Heck, **S. Tiede**, M. Yilmaz and G. Christofori (2018). "Foxf2 plays a dual role during transforming growth factor beta-induced epithelial to mesenchymal transition by promoting apoptosis yet enabling cell junction dissolution and migration." *Breast Cancer Res* **20**(1): 118. (<https://doi.org/10.1186/s13058-018-1043-6>)
- Ishay-Ronen, D., M. Diepenbruck, R. K. R. Kalathur, N. Sugiyama, **S. Tiede**, R. Ivanek, G. Bantug, M. F. Morini, J. Wang, C. Hess and G. Christofori (2019). "Gain Fat-Lose Metastasis: Converting Invasive Breast Cancer Cells into Adipocytes Inhibits Cancer Metastasis." *Cancer Cell* **35**(1): 17-32. (<https://doi.org/10.1016/j.ccell.2018.12.002>)
- Meyer-Schaller, N., M. Cardner, M. Diepenbruck, M. Saxena, **S. Tiede**, F. Luond, R. Ivanek, N. Beerenwinkel and G. Christofori (2019). "A Hierarchical Regulatory Landscape during the Multiple Stages of EMT." *Dev Cell* **48**(4): 539-553. (<https://doi.org/10.1016/j.devcel.2018.12.023>)

By contributing to these projects, I have had the chance to learn more about EMT-regulatory mechanisms and how they can be exploited to suppress tumour growth and metastatic seeding *in vivo*.

6. References

1. Inman, J.L., Robertson, C., Mott, J.D. & Bissell, M.J. Mammary gland development: cell fate specification, stem cells and the microenvironment. *Development* **142**, 1028-1042 (2015).
2. Visvader, J.E. & Stingl, J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev* **28**, 1143-1158 (2014).
3. Van Keymeulen, A. *et al.* Distinct stem cells contribute to mammary gland development and maintenance. *Nature* **479**, 189-193 (2011).
4. Davis, F.M. *et al.* Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun* **7**, 13053 (2016).
5. Rios, A.C., Fu, N.Y., Lindeman, G.J. & Visvader, J.E. In situ identification of bipotent stem cells in the mammary gland. *Nature* **506**, 322-327 (2014).
6. Scheele, C.L. *et al.* Identity and dynamics of mammary stem cells during branching morphogenesis. *Nature* **542**, 313-317 (2017).
7. Rios, A.C., Fu, N.Y., Cursons, J., Lindeman, G.J. & Visvader, J.E. The complexities and caveats of lineage tracing in the mammary gland. *Breast Cancer Res* **18**, 116 (2016).
8. Organization, W.H., Vol. 2019 (2019).
9. Burstein, H.J., Polyak, K., Wong, J.S., Lester, S.C. & Kaelin, C.M. Ductal carcinoma in situ of the breast. *N Engl J Med* **350**, 1430-1441 (2004).
10. Lakhani, S.R. The transition from hyperplasia to invasive carcinoma of the breast. *J Pathol* **187**, 272-278 (1999).
11. Lin, E.Y. *et al.* Progression to Malignancy in the Polyoma Middle T Oncoprotein Mouse Breast Cancer Model Provides a Reliable Model for Human Diseases. *The American Journal of Pathology* **163**, 2113-2126 (2003).
12. Obenauf, A.C. & Massagué, J. Surviving at a Distance: Organ-Specific Metastasis. *Trends in Cancer* **1**, 76-91 (2015).
13. Weigelt, B., Peterse, J.L. & van 't Veer, L.J. Breast cancer metastasis: markers and models. *Nat Rev Cancer* **5**, 591-602 (2005).
14. Harbeck, N. *et al.* Breast cancer. *Nat Rev Dis Primers* **5**, 66 (2019).
15. Koren, S. & Bentires-Alj, M. Breast Tumor Heterogeneity: Source of Fitness, Hurdle for Therapy. *Mol Cell* **60**, 537-546 (2015).
16. Melchor, L. *et al.* Identification of cellular and genetic drivers of breast cancer heterogeneity in genetically engineered mouse tumour models. *J Pathol* **233**, 124-137 (2014).
17. Skibinski, A. & Kuperwasser, C. The origin of breast tumor heterogeneity. *Oncogene* **34**, 5309-5316 (2015).
18. Perou, C.M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747-752 (2000).
19. Prat, A. *et al.* Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. *Breast Cancer Res* **12**, R68 (2010).
20. Prat, A. & Perou, C.M. Deconstructing the molecular portraits of breast cancer. *Mol Oncol* **5**, 5-23 (2011).
21. Ellis, M.J. & Perou, C.M. The genomic landscape of breast cancer as a therapeutic roadmap. *Cancer Discov* **3**, 27-34 (2013).
22. Fidler, I.J. Tumor heterogeneity and the biology of cancer invasion and metastasis. *Cancer Res* **38**, 2651-2660 (1978).

23. Friedl, P. & Wolf, K. Tumour-cell invasion and migration: diversity and escape mechanisms. *Nat Rev Cancer* **3**, 362-374 (2003).
24. Ewing, J. Neoplastic Diseases. A Treatise on Tumors, 77-89 (1928).
25. Paget, S. The distribution of secondary growths in cancer of the breast. 1889. *Cancer Metastasis Rev* **8**, 98-101 (1889).
26. Al-Mehdi, A.B. *et al.* Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis. *Nat Med* **6**, 100-102 (2000).
27. Chambers, A.F., Groom, A.C. & MacDonald, I.C. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer* **2**, 563-572 (2002).
28. Massague, J. & Obenauf, A.C. Metastatic colonization by circulating tumour cells. *Nature* **529**, 298-306 (2016).
29. Valastyan, S. & Weinberg, R.A. Tumor metastasis: molecular insights and evolving paradigms. *Cell* **147**, 275-292 (2011).
30. Vanharanta, S. & Massague, J. Origins of metastatic traits. *Cancer Cell* **24**, 410-421 (2013).
31. Hanahan, D. & Weinberg, R.A. Hallmarks of cancer: the next generation. *Cell* **144**, 646-674 (2011).
32. Bedard, P.L., Hansen, A.R., Ratain, M.J. & Siu, L.L. Tumour heterogeneity in the clinic. *Nature* **501**, 355-364 (2013).
33. Brooks, M.D., Burness, M.L. & Wicha, M.S. Therapeutic Implications of Cellular Heterogeneity and Plasticity in Breast Cancer. *Cell Stem Cell* **17**, 260-271 (2015).
34. Lawson, D.A., Kessenbrock, K., Davis, R.T., Pervolarakis, N. & Werb, Z. Tumour heterogeneity and metastasis at single-cell resolution. *Nat Cell Biol* **20**, 1349-1360 (2018).
35. Marusyk, A., Almendro, V. & Polyak, K. Intra-tumour heterogeneity: a looking glass for cancer? *Nat Rev Cancer* **12**, 323-334 (2012).
36. McGranahan, N. & Swanton, C. Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future. *Cell* **168**, 613-628 (2017).
37. Banerji, S. *et al.* Sequence analysis of mutations and translocations across breast cancer subtypes. *Nature* **486**, 405-409 (2012).
38. Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61-70 (2012).
39. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346-352 (2012).
40. Shah, S.P. *et al.* The clonal and mutational evolution spectrum of primary triple-negative breast cancers. *Nature* **486**, 395-399 (2012).
41. Greaves, M. & Maley, C.C. Clonal evolution in cancer. *Nature* **481**, 306-313 (2012).
42. Nowell, P.C. The clonal evolution of tumor cell populations. *Science* **194**, 23-28 (1976).
43. Nguyen, L.V. *et al.* DNA barcoding reveals diverse growth kinetics of human breast tumour subclones in serially passaged xenografts. *Nat Commun* **5**, 5871 (2014).
44. Nguyen, L.V. *et al.* Barcoding reveals complex clonal dynamics of de novo transformed human mammary cells. *Nature* **528**, 267-271 (2015).
45. Eirew, P. *et al.* Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature* **518**, 422-426 (2015).
46. Campbell, P.J. *et al.* The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109-1113 (2010).

References

47. Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat Genet* **46**, 225-233 (2014).
48. Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *366*, 883-892 (2012).
49. Hunter, K.W., Amin, R., Deasy, S., Ha, N.H. & Wakefield, L. Genetic insights into the morass of metastatic heterogeneity. *Nat Rev Cancer* **18**, 211-223 (2018).
50. Klein, C.A. Selection and adaptation during metastatic cancer progression. *Nature* **501**, 365-372 (2013).
51. Casasent, A.K. *et al.* Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell* **172**, 205-217 e212 (2018).
52. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979-993 (2012).
53. Nik-Zainal, S. *et al.* The life history of 21 breast cancers. *Cell* **149**, 994-1007 (2012).
54. Yates, L.R. *et al.* Genomic Evolution of Breast Cancer Metastasis and Relapse. *Cancer Cell* **32**, 169-184 e167 (2017).
55. Wang, Y. *et al.* Clonal evolution in breast cancer revealed by single nucleus genome sequencing. *Nature* **512**, 155-160 (2014).
56. van 't Veer, L.J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530-536 (2002).
57. Wang, Y. *et al.* Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. *The Lancet* **365**, 671-679 (2005).
58. Allinen, M. *et al.* Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **6**, 17-32 (2004).
59. Szczerba, B.M. *et al.* Neutrophils escort circulating tumour cells to enable cell cycle progression. *Nature* **566**, 553-557 (2019).
60. Kang, Y. *et al.* A multigenic program mediating breast cancer metastasis to bone. *Cancer Cell* **3**, 537-549 (2003).
61. Minn, A.J. *et al.* Genes that mediate breast cancer metastasis to lung. *Nature* **436**, 518-524 (2005).
62. Alizadeh, A.A. *et al.* Toward understanding and exploiting tumor heterogeneity. *Nat Med* **21**, 846-853 (2015).
63. Navin, N. *et al.* Tumour evolution inferred by single-cell sequencing. *Nature* **472**, 90-94 (2011).
64. Shah, S.P. *et al.* Mutational evolution in a lobular breast tumour profiled at single nucleotide resolution. *Nature* **461**, 809-813 (2009).
65. Obradovic, M.M.S. *et al.* Glucocorticoids promote breast cancer metastasis. *Nature* **567**, 540-544 (2019).
66. Gkoutela, S. *et al.* Circulating Tumor Cell Clustering Shapes DNA Methylation to Enable Metastasis Seeding. *Cell* **176**, 98-112 e114 (2019).
67. Dravis, C. *et al.* Epigenetic and Transcriptomic Profiling of Mammary Gland Development and Tumor Models Disclose Regulators of Cell State Plasticity. *Cancer Cell* **34**, 466-482 e466 (2018).
68. Kim, C. *et al.* Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell* **173**, 879-893 e813 (2018).
69. Meacham, C.E. & Morrison, S.J. Tumour heterogeneity and cancer cell plasticity. *Nature* **501**, 328-337 (2013).
70. Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* **15**, 178-196 (2014).

71. Shibue, T. & Weinberg, R.A. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol* **14**, 611-629 (2017).
72. Fischer, K.R. *et al.* Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472-476 (2015).
73. Zheng, X. *et al.* Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525-530 (2015).
74. Meyer-Schaller, N. *et al.* A Hierarchical Regulatory Landscape during the Multiple Stages of EMT. *Dev Cell* **48**, 539-553 e536 (2019).
75. Dongre, A. & Weinberg, R.A. New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer. *Nat Rev Mol Cell Biol* **20**, 69-84 (2019).
76. Nieto, M.A., Huang, R.Y., Jackson, R.A. & Thiery, J.P. Emt: 2016. *Cell* **166**, 21-45 (2016).
77. Pastushenko, I. *et al.* Identification of the tumour transition states occurring during EMT. *Nature* **556**, 463-468 (2018).
78. Puram, S.V. *et al.* Single-Cell Transcriptomic Analysis of Primary and Metastatic Tumor Ecosystems in Head and Neck Cancer. *Cell* **171**, 1611-1624 e1624 (2017).
79. Aiello, N.M. *et al.* EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration. *Dev Cell* **45**, 681-695 e684 (2018).
80. Ishay-Ronen, D. *et al.* Gain Fat-Lose Metastasis: Converting Invasive Breast Cancer Cells into Adipocytes Inhibits Cancer Metastasis. *Cancer Cell* **35**, 17-32 e16 (2019).
81. Koren, S. *et al.* PIK3CA(H1047R) induces multipotency and multi-lineage mammary tumours. *Nature* **525**, 114-118 (2015).
82. Van Keymeulen, A. *et al.* Reactivation of multipotency by oncogenic PIK3CA induces breast tumour heterogeneity. *Nature* **525**, 119-123 (2015).
83. Gupta, P.B. *et al.* Stochastic state transitions give rise to phenotypic equilibrium in populations of cancer cells. *Cell* **146**, 633-644 (2011).
84. Klevebring, D. *et al.* Sequencing of breast cancer stem cell populations indicates a dynamic conversion between differentiation states in vivo. *Breast Cancer Res* **16**, R72 (2014).
85. Chaffer, C.L. *et al.* Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *Cell* **154**, 61-74 (2013).
86. Lawson, D.A. *et al.* Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131-135 (2015).
87. Granit, R.Z. *et al.* Regulation of Cellular Heterogeneity and Rates of Symmetric and Asymmetric Divisions in Triple-Negative Breast Cancer. *Cell Rep* **24**, 3237-3250 (2018).
88. Gupta, G.P. *et al.* Mediators of vascular remodelling co-opted for sequential steps in lung metastasis. *Nature* **446**, 765-770 (2007).
89. Wagenblast, E. *et al.* A model of breast cancer heterogeneity reveals vascular mimicry as a driver of metastasis. *Nature* **520**, 358-362 (2015).
90. Kelly, P.N., Dakic, A., Adams, J.M., Nutt, S.L. & Strasser, A. Tumor growth need not be driven by rare cancer stem cells. *Science* **317**, 337 (2007).
91. Roesch, A. *et al.* A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* **141**, 583-594 (2010).
92. Chen, K. *et al.* Differential Sensitivities of Fast- and Slow-Cycling Cancer Cells to Inosine Monophosphate Dehydrogenase 2 Inhibition by Mycophenolic Acid. *Mol Med* **21**, 792-802 (2016).

References

93. Perego, M. *et al.* A slow-cycling subpopulation of melanoma cells with highly invasive properties. *Oncogene* **37**, 302-312 (2018).
94. Wu, F.H. *et al.* Characterization and functional analysis of a slow-cycling subpopulation in colorectal cancer enriched by cell cycle inducer combined chemotherapy. *Oncotarget* **8**, 78466-78479 (2017).
95. Adams, J.M. & Strasser, A. Is tumor growth sustained by rare cancer stem cells or dominant clones? *Cancer Res* **68**, 4018-4021 (2008).
96. Dowsett, M. *et al.* Assessment of Ki67 in breast cancer: recommendations from the International Ki67 in Breast Cancer working group. *J Natl Cancer Inst* **103**, 1656-1664 (2011).
97. Wells, A., Griffith, L., Wells, J.Z. & Taylor, D.P. The dormancy dilemma: quiescence versus balanced proliferation. *Cancer Res* **73**, 3811-3816 (2013).
98. Cleary, A.S., Leonard, T.L., Gestl, S.A. & Gunther, E.J. Tumour cell heterogeneity maintained by cooperating subclones in Wnt-driven mammary cancers. *Nature* **508**, 113-117 (2014).
99. Polyak, K. & Marusyk, A. Cancer: Clonal cooperation. *Nature* **508**, 52-53 (2014).
100. Malladi, S. *et al.* Metastatic Latency and Immune Evasion through Autocrine Inhibition of WNT. *Cell* **165**, 45-60 (2016).
101. Oskarsson, T. *et al.* Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat Med* **17**, 867-874 (2011).
102. Scheel, C. *et al.* Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast. *Cell* **145**, 926-940 (2011).
103. Kim, M.Y. *et al.* Tumor self-seeding by circulating cancer cells. *Cell* **139**, 1315-1326 (2009).
104. Gudem, G. *et al.* The evolutionary history of lethal metastatic prostate cancer. *Nature* **520**, 353-357 (2015).
105. Archetti, M. & Pienta, K.J. Cooperation among cancer cells: applying game theory to cancer. *Nat Rev Cancer* **19**, 110-117 (2019).
106. Joyce, J.A. & Pollard, J.W. Microenvironmental regulation of metastasis. *Nat Rev Cancer* **9**, 239-252 (2009).
107. Junttila, M.R. & de Sauvage, F.J. Influence of tumour micro-environment heterogeneity on therapeutic response. *Nature* **501**, 346-354 (2013).
108. Quail, D.F. & Joyce, J.A. Microenvironmental regulation of tumor progression and metastasis. *Nat Med* **19**, 1423-1437 (2013).
109. Tabassum, D.P. & Polyak, K. Tumorigenesis: it takes a village. *Nat Rev Cancer* **15**, 473-483 (2015).
110. Marusyk, A. *et al.* Non-cell-autonomous driving of tumour growth supports sub-clonal heterogeneity. *Nature* **514**, 54-58 (2014).
111. Goetz, J.G. *et al.* Biomechanical remodeling of the microenvironment by stromal caveolin-1 favors tumor invasion and metastasis. *Cell* **146**, 148-163 (2011).
112. Malanchi, I. *et al.* Interactions between cancer stem cells and their niche govern metastatic colonization. *Nature* **481**, 85-89 (2011).
113. Boelens, M.C. *et al.* Exosome transfer from stromal to breast cancer cells regulates therapy resistance pathways. *Cell* **159**, 499-513 (2014).
114. Bhowmick, N.A., Neilson, E.G. & Moses, H.L. Stromal fibroblasts in cancer initiation and progression. *Nature* **432**, 332-337 (2004).
115. Wyckoff, J. *et al.* A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. *Cancer Res* **64**, 7022-7029 (2004).
116. Wyckoff, J.B. *et al.* Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. *Cancer Res* **67**, 2649-2656 (2007).

117. Janiszewska, M. *et al.* Subclonal cooperation drives metastasis by modulating local and systemic immune microenvironments. *Nat Cell Biol* **21**, 879-888 (2019).
118. Ombrato, L. & Malanchi, I. Subclonal cooperation rewrites metastasis. *Nat Cell Biol* **21**, 797-798 (2019).
119. Jimenez-Sanchez, A. *et al.* Heterogeneous Tumor-Immune Microenvironments among Differentially Growing Metastases in an Ovarian Cancer Patient. *Cell* **170**, 927-938 e920 (2017).
120. Swanton, C. Intratumor heterogeneity: evolution through space and time. *Cancer Res* **72**, 4875-4882 (2012).
121. Yates, L.R. *et al.* Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med* **21**, 751-759 (2015).
122. Desmedt, C. *et al.* Uncovering the genomic heterogeneity of multifocal breast cancer. *J Pathol* **236**, 457-466 (2015).
123. Miron, A. *et al.* PIK3CA mutations in in situ and invasive breast carcinomas. *Cancer Res* **70**, 5674-5678 (2010).
124. Pestrin, M. *et al.* Heterogeneity of PIK3CA mutational status at the single cell level in circulating tumor cells from metastatic breast cancer patients. *Mol Oncol* **9**, 749-757 (2015).
125. Geyer, F.C. *et al.* Molecular analysis reveals a genetic basis for the phenotypic diversity of metaplastic breast carcinomas. *J Pathol* **220**, 562-573 (2010).
126. Waclaw, B. *et al.* A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* **525**, 261-264 (2015).
127. Potts, S.J. *et al.* Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. *Lab Invest* **92**, 1342-1357 (2012).
128. Chen, J. *et al.* Spatial transcriptomic analysis of cryosectioned tissue samples with Geo-seq. *Nat Protoc* **12**, 566-580 (2017).
129. Macaulay, I.C. *et al.* G&T-seq: parallel sequencing of single-cell genomes and transcriptomes. *Nat Methods* **12**, 519-522 (2015).
130. Macaulay, I.C. *et al.* Separation and parallel sequencing of the genomes and transcriptomes of single cells using G&T-seq. *Nat Protoc* **11**, 2081-2103 (2016).
131. Picelli, S. *et al.* Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat Methods* **10**, 1096-1098 (2013).
132. Picelli, S. *et al.* Full-length RNA-seq from single cells using Smart-seq2. *Nat Protoc* **9**, 171-181 (2014).
133. Miller, K. *et al.* Paclitaxel plus bevacizumab versus paclitaxel alone for metastatic breast cancer. *N Engl J Med* **357**, 2666-2676 (2007).
134. Zhao, J. & Guan, J.L. Signal transduction by focal adhesion kinase in cancer. *Cancer Metastasis Rev* **28**, 35-49 (2009).
135. Kleinschmidt, E.G. & Schlaepfer, D.D. Focal adhesion kinase signaling in unexpected places. *Curr Opin Cell Biol* **45**, 24-30 (2017).
136. Mitra, S.K., Hanson, D.A. & Schlaepfer, D.D. Focal adhesion kinase: in command and control of cell motility. *Nat Rev Mol Cell Biol* **6**, 56-68 (2005).
137. Sulzmaier, F.J., Jean, C. & Schlaepfer, D.D. FAK in cancer: mechanistic findings and clinical applications. *Nat Rev Cancer* **14**, 598-610 (2014).
138. Shibue, T., Brooks, M.W., Inan, M.F., Reinhardt, F. & Weinberg, R.A. The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. *Cancer Discov* **2**, 706-721 (2012).
139. Shibue, T., Brooks, M.W. & Weinberg, R.A. An integrin-linked machinery of cytoskeletal regulation that enables experimental tumor initiation and metastatic colonization. *Cancer Cell* **24**, 481-498 (2013).

References

140. Hirata, E. *et al.* Intravital imaging reveals how BRAF inhibition generates drug-tolerant microenvironments with high integrin beta1/FAK signaling. *Cancer Cell* **27**, 574-588 (2015).
141. Tavora, B. *et al.* Endothelial-cell FAK targeting sensitizes tumours to DNA-damaging therapy. *Nature* **514**, 112-116 (2014).
142. Dunn, K.B., Heffler, M. & Golubovskaya, V.M. Evolving therapies and FAK inhibitors for the treatment of cancer. *Anticancer Agents Med Chem* **10**, 722-734 (2010).
143. Golubovskaya, V.M. Targeting focal adhesion kinase in cancer-part I. *Anticancer Agents Med Chem* **10**, 713 (2010).
144. Hirt, U.A. *et al.* Efficacy of the highly selective focal adhesion kinase inhibitor BI 853520 in adenocarcinoma xenograft models is linked to a mesenchymal tumor phenotype. *Oncogenesis* **7**, 21 (2018).
145. Tiede, S. *et al.* The FAK inhibitor BI 853520 exerts anti-tumor effects in breast cancer. *Oncogenesis* **7**, 73 (2018).
146. Polzer, B. & Klein, C.A. Metastasis awakening: the challenges of targeting minimal residual cancer. *Nat Med* **19**, 274-275 (2013).
147. Fantozzi, A. & Christofori, G. Mouse models of breast cancer metastasis. *Breast Cancer Res* **8**, 212 (2006).
148. Fluck, M.M. & Schaffhausen, B.S. Lessons in signaling and tumorigenesis from polyomavirus middle T antigen. *Microbiol Mol Biol Rev* **73**, 542-563, Table of Contents (2009).
149. Guy, C.T., Cardiff, R.D. & Muller, W.J. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. *Mol Cell Biol* **12**, 954-961 (1992).
150. Maglione, J.E. *et al.* Polyomavirus middle T-induced mammary intraepithelial neoplasia outgrowths: single origin, divergent evolution, and multiple outcomes. *Mol Cancer Ther* **3**, 941-953 (2004).
151. Maglione, J.E. *et al.* Transgenic Polyoma middle-T mice model premalignant mammary disease. *Cancer Res* **61**, 8298-8305 (2001).
152. Namba, R. *et al.* Molecular characterization of the transition to malignancy in a genetically engineered mouse-based model of ductal carcinoma in situ. *Mol Cancer Res* **2**, 453-463 (2004).
153. Qiu, T.H. *et al.* Global expression profiling identifies signatures of tumor virulence in MMTV-PyMT-transgenic mice: correlation to human disease. *Cancer Res* **64**, 5973-5981 (2004).
154. Franci, C. *et al.* Biomarkers of residual disease, disseminated tumor cells, and metastases in the MMTV-PyMT breast cancer model. *PLoS One* **8**, e58183 (2013).
155. Cai, Y. *et al.* Transcriptomic dynamics of breast cancer progression in the MMTV-PyMT mouse model. *BMC Genomics* **18**, 185 (2017).
156. Herschkowitz, J.I. *et al.* Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8**, R76 (2007).
157. Hollern, D.P. & Andrechek, E.R. A genomic analysis of mouse models of breast cancer reveals molecular features of mouse models and relationships to human breast cancer. *Breast Cancer Res* **16**, R59 (2014).
158. Lim, E. *et al.* Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. *Breast Cancer Res* **12**, R21 (2010).

159. Pfefferle, A.D. *et al.* Transcriptomic classification of genetically engineered mouse models of breast cancer identifies human subtype counterparts. *Genome Biol* **14**, R125 (2013).
160. Kretzschmar, K. & Watt, F.M. Lineage tracing. *Cell* **148**, 33-45 (2012).
161. Livet, J. *et al.* Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature* **450**, 56-62 (2007).
162. Snippert, H.J. *et al.* Intestinal crypt homeostasis results from neutral competition between symmetrically dividing Lgr5 stem cells. *Cell* **143**, 134-144 (2010).
163. Driessens, G., Beck, B., Caauwe, A., Simons, B.D. & Blanpain, C. Defining the mode of tumour growth by clonal analysis. *Nature* **488**, 527-530 (2012).
164. Sale, S. & Pavelic, K. Mammary lineage tracing: the coming of age. *Cell Mol Life Sci* **72**, 1577-1583 (2015).
165. Zomer, A. *et al.* Intravital imaging of cancer stem cell plasticity in mammary tumors. *Stem Cells* **31**, 602-606 (2013).
166. Rios, A.C. *et al.* Intraclonal Plasticity in Mammary Tumors Revealed through Large-Scale Single-Cell Resolution 3D Imaging. *Cancer Cell* **35**, 618-632 e616 (2019).
167. Maddipati, R. & Stanger, B.Z. Pancreatic Cancer Metastases Harbor Evidence of Polyclonality. *Cancer Discov* **5**, 1086-1097 (2015).
168. Reeves, M.Q., Kandyba, E., Harris, S., Del Rosario, R. & Balmain, A. Multicolour lineage tracing reveals clonal dynamics of squamous carcinoma evolution from initiation to metastasis. *Nat Cell Biol* **20**, 699-709 (2018).
169. Tang, Y.J. *et al.* Tracing Tumor Evolution in Sarcoma Reveals Clonal Origin of Advanced Metastasis. *Cell Rep* **28**, 2837-2850 e2835 (2019).
170. Aceto, N. *et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* **158**, 1110-1122 (2014).
171. Cheung, K.J., Gabrielson, E., Werb, Z. & Ewald, A.J. Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell* **155**, 1639-1651 (2013).
172. Cheung, K.J. *et al.* Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc Natl Acad Sci U S A* **113**, E854-863 (2016).
173. Rulands, S. *et al.* Universality of clone dynamics during tissue development. *Nat Phys* **14**, 469-474 (2018).
174. Lin, E.Y. *et al.* Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *Am J Pathol* **163**, 2113-2126 (2003).
175. Rodriguez-Enriquez, S. *et al.* Control of cellular proliferation by modulation of oxidative phosphorylation in human and rodent fast-growing tumor cells. *Toxicol Appl Pharmacol* **215**, 208-217 (2006).
176. Yao, C.H. *et al.* Mitochondrial fusion supports increased oxidative phosphorylation during cell proliferation. *Elife* **8** (2019).
177. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
178. Gaidatzis, D., Lerch, A., Hahne, F. & Stadler, M.B. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**, 1130-1132 (2015).
179. Robinson, M.D., McCarthy, D.J. & Smyth, G.K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139-140 (2010).
180. Yu, G. & He, Q.Y. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst* **12**, 477-479 (2016).

References

181. Balwierz, P.J. *et al.* ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. *Genome Res* **24**, 869-884 (2014).
182. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* **102**, 15545-15550 (2005).
183. Liberzon, A. *et al.* Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739-1740 (2011).
184. Carbon, S. *et al.* AmiGO: online access to ontology and annotation data. *Bioinformatics* **25**, 288-289 (2009).
185. Fabregat, A. *et al.* The Reactome Pathway Knowledgebase. *Nucleic Acids Res* **46**, D649-D655 (2018).
186. Acharyya, S. *et al.* A CXCL1 Paracrine Network Links Cancer Chemoresistance and Metastasis. *Cell* **150**, 165-178 (2012).
187. Dunn, K.B., Heffler, M. & Golubovskaya, V. Evolving Therapies and FAK Inhibitors for the Treatment of Cancer. *Anticancer Agents in Medicinal Chemistry* **10**, 722-734 (2010).
188. Lee, B.Y., Timpson, P., Horvath, L.G. & Daly, R.J. FAK signaling in human cancer as a target for therapeutics. *Pharmacology & Therapeutics* **146**, 132-149 (2015).
189. Golubovskaya, V. Focal Adhesion Kinase as a Cancer Therapy Target. *Anticancer Agents in Medicinal Chemistry* **10**, 735-741 (2010).
190. Schaller, M.D. Cellular functions of FAK kinases: insight into molecular mechanisms and novel functions. *Journal of Cell Science* **123**, 1007-1013 (2010).
191. Yoon, H., Dehart, J.P., Murphy, J.M. & Lim, S.T.S. Understanding the Roles of FAK in Cancer: Inhibitors, Genetic Models, and New Insights. *Journal of Histochemistry & Cytochemistry* **63**, 114-128 (2015).
192. Hao, H. *et al.* Focal adhesion kinase as potential target for cancer therapy. *Oncology Reports* **22**, 973-979 (2009).
193. Infusino, G.A. & Jacobson, J.R. Endothelial FAK as a therapeutic target in disease. *Microvascular Research* **83**, 89-96 (2012).
194. Lim, Y. *et al.* PyK2 and FAK connections to p190Rho guanine nucleotide exchange factor regulate RhoA activity, focal adhesion formation, and cell motility. *Journal of Cell Biology* **180**, 187-203 (2008).
195. Provenzano, P.P. & Keely, P.J. The role of focal adhesion kinase in tumor initiation and progression. *Cell Adhesion & Migration* **3**, 347-350 (2009).
196. Brami-Cherrier, K. *et al.* FAK dimerization controls its kinase-dependent functions at focal adhesions. *EMBO J* **33**, 356-370 (2014).
197. Zhao, X. & Guan, J.L. Focal adhesion kinase and its signaling pathways in cell migration and angiogenesis. *Advanced Drug Delivery Reviews* **63**, 610-615 (2011).
198. Cai, X. *et al.* Spatial and temporal regulation of focal adhesion kinase activity in living cells. *Mol Cell Biol* **28**, 201-214 (2008).
199. Lechertier, T. & Hoidalva-Dilke, K. Focal adhesion kinase and tumour angiogenesis. *Journal of Pathology* **226**, 404-412 (2012).
200. Mitra, S.K., Hanson, D.A. & Schlaepfer, D.D. Focal adhesion kinase: In command and control of cell motility. *Nature Reviews Molecular Cell Biology* **6**, 56-68 (2005).
201. Parsons, J.T. Focal adhesion kinase: the first ten years. *Journal of Cell Science* **116**, 1409-1416 (2003).

202. Zhao, J. & Guan, J.L. Signal transduction by focal adhesion kinase in cancer. *Cancer and Metastasis Reviews* **28**, 35-49 (2009).
203. Kostourou, V. *et al.* FAK-heterozygous mice display enhanced tumour angiogenesis. *Nature communications* **4**, 2020 (2013).
204. Golubovskaya, V. Targeting FAK in human cancer: from finding to first clinical trials. *Frontiers in Bioscience* **19**, 687–706 (2014).
205. Schultze, A. & Fiedler, W. Therapeutic potential and limitations of new FAK inhibitors in the treatment of cancer. *Expert Opinion on Investigational Drugs* **19**, 777-788 (2010).
206. Sulzmaier, F.J., Jean, C. & Schlaepfer, D.D. FAK in cancer: mechanistic findings and clinical applications. *Nature Reviews Cancer* **14**, 598-610 (2014).
207. Tavora, B. *et al.* Endothelial FAK is required for tumour angiogenesis. *Embo Mol Med* **2**, 516-528 (2010).
208. Waldmeier, L., Meyer-Schaller, N., Diepenbruck, M. & Christofori, G. Py2T Murine Breast Cancer Cells, a Versatile Model of TGF beta-Induced EMT In Vitro and In Vivo. *Plos One* **7** (2012).
209. Shapiro, I.M. *et al.* Merlin Deficiency Predicts FAK Inhibitor Sensitivity: A Synthetic Lethal Relationship. *Science Translational Medicine* **6** (2014).
210. Lin, E.Y. *et al.* Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. *American Journal of Pathology* **163**, 2113-2126 (2003).
211. Aslakson, C.J. & Miller, F.R. 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 (1992).
212. Lehenbre, F. *et al.* NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. *EMBO J* **27**, 2603-2615 (2008).
213. Hirt, U.A. *et al.* in Proceedings of the AACR-NCI-EORTC International Conference: Molecular Targets and Cancer Therapeutics, Vol. 10 (San Francisco, CA. Philadelphia; 2011).
214. Hirt, U.A., Haslinger, C., Schweifer, N., Garin-Chesa, P. & Adolf, G.R. 469: E-cadherin Expression Predicts Response of Carcinomas to Treatment with PTK2 Inhibitors. *European Journal of Cancer* **48** (2012).
215. Canel, M., Serrels, A., Frame, M.C. & Brunton, V.G. E-cadherin-integrin crosstalk in cancer invasion and metastasis. *J Cell Sci* **126**, 393-401 (2013).
216. Shapiro, I.M. *et al.* Merlin deficiency predicts FAK inhibitor sensitivity: a synthetic lethal relationship. *Sci Transl Med* **6**, 237ra268 (2014).
217. Chao, Y.L., Shepard, C.R. & Wells, A. Breast carcinoma cells re-express E-cadherin during mesenchymal to epithelial reverting transition. *Mol Cancer* **9**, 179 (2010).
218. Cicchini, C. *et al.* TGFbeta-induced EMT requires focal adhesion kinase (FAK) signaling. *Exp Cell Res* **314**, 143-152 (2008).
219. Hugo, H. *et al.* Epithelial--mesenchymal and mesenchymal--epithelial transitions in carcinoma progression. *J Cell Physiol* **213**, 374-383 (2007).
220. van Nimwegen, M.J., Verkoeijen, S., van Buren, L., Burg, D. & van de Water, B. Requirement for focal adhesion kinase in the early phase of mammary adenocarcinoma lung metastasis formation. *Cancer Res* **65**, 4698-4706 (2005).
221. Wendt, M.K. & Schiemann, W.P. Therapeutic targeting of the focal adhesion complex prevents oncogenic TGF-beta signaling and metastasis. *Breast Cancer Res* **11**, R68 (2009).

References

222. Wendt, M.K., Taylor, M.A., Schiemann, B.J. & Schiemann, W.P. Down-regulation of epithelial cadherin is required to initiate metastatic outgrowth of breast cancer. *Mol Biol Cell* **22**, 2423-2435 (2011).
223. Shibue, T., Brooks, M.W., Inan, M.F., Reinhardt, F. & Weinberg, R.A. The outgrowth of micrometastases is enabled by the formation of filopodium-like protrusions. *Cancer Discovery* **2**, 706–721 (2012).
224. Ward, K.K. *et al.* Inhibition of focal adhesion kinase (FAK) activity prevents anchorage-independent ovarian carcinoma cell growth and tumor progression. *Clinical and Experimental Metastasis* **30**, 579–594 (2013).
225. Mitra, S.K., Lim, S.T., Chi, A. & Schlaepfer, D.D. Intrinsic focal adhesion kinase activity controls orthotopic breast carcinoma metastasis via the regulation of urokinase plasminogen activator expression in a syngeneic tumor model. *Oncogene* **25**, 4429-4440 (2006).
226. Wu, Z.M., Yuan, X.H., Jiang, P.C., Li, Z.Q. & Wu, T. Antisense oligonucleotides targeting the focal adhesion kinase inhibit proliferation, induce apoptosis and cooperate with cytotoxic drugs in human glioma cells. *J Neurooncol* **77**, 117-123 (2006).
227. Smith, C.S. *et al.* Effect of focal adhesion kinase (FAK) downregulation with FAK antisense oligonucleotides and 5-fluorouracil on the viability of melanoma cell lines. *Melanoma Res* **15**, 357-362 (2005).
228. Chen, Y. *et al.* The effect of focal adhesion kinase gene silencing on 5-fluorouracil chemosensitivity involves an Akt/NF- κ B signaling pathway in colorectal carcinomas. *International Journal of Cancer* **127**, 195-206 (2010).
229. van Nimwegen, M.J., Huigsloot, M., Caier, A., Tijdens, I.B. & van de Water, B. Focal Adhesion Kinase and Protein Kinase B Cooperate to Suppress Doxorubicin-Induced Apoptosis of Breast Tumor Cells. *Molecular Pharmacology* **70**, 1330-1339 (2006).
230. Halder, J. *et al.* Focal Adhesion Kinase Targeting Using In vivo Short Interfering RNA Delivery in Neutral Liposomes for Ovarian Carcinoma Therapy. *Clinical Cancer Research* **12**, 4916-4924 (2006).
231. Zhang, H.M. *et al.* Induced focal adhesion kinase expression suppresses apoptosis by activating NF- κ B signaling in intestinal epithelial cells. *American Journal of Physiology* **290**, C1310-C1320 (2006).
232. Huanwen, W. *et al.* Intrinsic chemoresistance to gemcitabine is associated with constitutive and laminin-induced phosphorylation of FAK in pancreatic cancer cell lines. *Molecular Cancer* **8**, 1-18 (2009).
233. Duxbury, M.S. *et al.* RNA interference targeting focal adhesion kinase enhances pancreatic adenocarcinoma gemcitabine chemosensitivity. *Biochem Biophys Res Commun* **311**, 786-792 (2003).
234. Halder, J. *et al.* Therapeutic efficacy of a novel focal adhesion kinase inhibitor TAE226 in ovarian carcinoma. *Cancer Res* **67**, 10976-10983 (2007).
235. Hochwald, S.N. *et al.* A novel small molecule inhibitor of FAK decreases growth of human pancreatic cancer. *Cell Cycle* **8**, 2435-2443 (2009).
236. Kang, Y. *et al.* Role of Focal Adhesion Kinase in Regulating YB-1-Mediated Paclitaxel Resistance in Ovarian Cancer. *Journal of the National Cancer Institute* **105**, 1485–1495 (2013).
237. Symeonides, S.N., Anderton, S.M. & Serrels, A. FAK-inhibition opens the door to checkpoint immunotherapy in Pancreatic Cancer. *J Immunother Cancer* **5**, 17 (2017).
238. Jiang, H. *et al.* Targeting focal adhesion kinase renders pancreatic cancers responsive to checkpoint immunotherapy. *Nat Med* **22**, 851-860 (2016).

239. Serrels, A. *et al.* Nuclear FAK controls chemokine transcription, Tregs, and evasion of anti-tumor immunity. *Cell* **163**, 160-173 (2015).
240. Tavora, B. *et al.* Endothelial-cell FAK targeting sensitizes tumours to DNA-damaging therapy. *Nature Letter* **514**, 112-116 (2014).
241. Cooke, V.G. *et al.* Pericyte depletion results in hypoxia-associated epithelial-to-mesenchymal transition and metastasis mediated by met signaling pathway. *Cancer Cell* **21**, 66-81 (2012).
242. Clarke, R.B., Stingl, J., Vivanco, M. & Bentires-Alj, M. 'The charmingest place': non-coding RNA, lineage tracing, tumor heterogeneity, metastasis and metabolism - new methods in mammary gland development and cancer: the fifth ENBDC Workshop. *Breast Cancer Research* **15** (2013).
243. Fantozzi, A. & Christofori, G. Mouse models of breast cancer metastasis. *Breast Cancer Research* **8** (2006).
244. Diepenbruck, M. *et al.* miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis. *Nat Commun* **8**, 1168 (2017).
245. Gene Ontology, C. Gene Ontology Consortium: going forward. *Nucleic Acids Res* **43**, D1049-1056 (2015).
246. Falcon, S. & Gentleman, R. Using GOstats to test gene lists for GO term association. *Bioinformatics* **23**, 257-258 (2007).

7. Acknowledgements

First of all, I would like to express my greatest gratitude to Prof. Dr. Gerhard Christofori for offering me an interesting internship position and – most importantly – also the possibility to perform my Ph.D. in his laboratory. Under your guidance, I had the freedom to develop myself as a researcher and to strengthen my personality. Our fruitful discussions, your never-ending support, together with an exciting lab retreat and several lab excursions made the time unforgettable to me.

Furthermore, I am thankful to my committee members Prof. Dr. Mohamed Bentires-Alj and Prof. Dr. Sven Rottenberg for helpful comments and inspiring input.

Special thanks also goes to Ruben Bill, my master thesis supervisor and BI 853520 project partner. Your methodological input, together with your confidence in me and your cheerful nature, have contributed to successfully accomplish the BI 853520 project. I would also like to thank all members of the “small-lab”, especially Fabiana Lüönd, Maren Diepenbruck and Marco Morini, for your unlimited assistance, critical questions and jokes and sometimes – in dire straits – encouraging words. Of course, a big “thank you” also goes to all other current and previous members of the Christofori laboratory for answering my questions, sharing knowledge and advices, helping out, cheering-up and having a great time together. I am also grateful to all contributing authors of my (un)published stories, especially to Prof. Dr. Nicola Aceto and Barbara Szczerba for microfluidic CTC isolation, to the technical support provided by E. Panoussis, I. Galm, U. Schmieder and T. Bürglin, to microscopic advice by P. Lorentz, to the animal care takers and to all the other people dedicated to ensuring a functioning infrastructure. You made my life much easier.

Most importantly, I would not have reached so far without the support of my family. Even though most of you are distant-wise far away, you have always been in my heart and you are my closest. I am so glad having you.

Last but not least, I would like to express my deepest love to you, Leonhard. My home is where you are and I am nothing without you.

ARTICLE

DOI: 10.1038/s41467-017-01197-w

OPEN

miR-1199-5p and Zeb1 function in a double-negative feedback loop potentially coordinating EMT and tumour metastasis

Maren Diepenbruck¹, Stefanie Tiede¹, Meera Saxena¹, Robert Ivanek ¹, Ravi Kiran Reddy Kalathur¹, Fabiana Lüönd¹, Nathalie Meyer-Schaller¹ & Gerhard Christofori¹

Epithelial tumour cells can gain invasive and metastatic capabilities by undergoing an epithelial-mesenchymal transition. Transcriptional regulators and post-transcriptional effectors like microRNAs orchestrate this process of high cellular plasticity and its malignant consequences. Here, using microRNA sequencing in a time-resolved manner and functional validation, we have identified microRNAs that are critical for the regulation of an epithelial-mesenchymal transition and of mesenchymal tumour cell migration. We report that miR-1199-5p is downregulated in its expression during an epithelial-mesenchymal transition, while its forced expression prevents an epithelial-mesenchymal transition, tumour cell migration and invasion in vitro, and lung metastasis in vivo. Mechanistically, miR-1199-5p acts in a reciprocal double-negative feedback loop with the epithelial-mesenchymal transition transcription factor Zeb1. This function resembles the activities of miR-200 family members, guardians of an epithelial cell phenotype. However, miR-1199-5p and miR-200 family members share only six target genes, indicating that, besides regulating Zeb1 expression, they exert distinct functions during an epithelial-mesenchymal transition.

¹Department of Biomedicine, University of Basel, 4058 Basel, Switzerland. Nathalie Meyer-Schaller and Gerhard Christofori contributed equally to this work. Correspondence and requests for materials should be addressed to N.M-S. (email: nathalie.meyer-schaller@unibas.ch) or to G.C. (email: gerhard.christofori@unibas.ch)

An epithelial-to-mesenchymal transition (EMT) as well as its reversal, a mesenchymal-to-epithelial transition (MET), reflect two gradual, well-controlled processes during embryogenesis and wound healing in adults to promote tissue and organ formation and homeostasis. Both processes induce a global reorganization of a cell's constitution and allow switching back and forth between two different cell phenotypes to endow the necessity of tissue plasticity¹. In the context of malignant tumour progression, epithelial tumour cells can undergo an EMT upon diverse extracellular stimuli and consequently gain metastatic capabilities. Among many growth factors and environmental cues, such as tissue hypoxia, transforming growth factor β (TGF β) strongly activates the dedifferentiation process in epithelial tumour cells and, thus, induces global changes in a cell's transcriptional and post-transcriptional networks^{2–4}. This allows tumour cells to disseminate from the primary tumour and to intravasate and survive in the blood circulation. At the distant organ, cells extravasate into the organ parenchyma and eventually grow out as lethal metastases, possibly promoted by a MET^{5–8}. An EMT also provides cells with increased chemoresistance, thus impeding efficient therapy of malignant mesenchymal cancer cells^{9, 10}.

MicroRNAs (miRNAs) represent a class of ~22 nucleotide-long non-coding RNAs that can regulate gene expression at the post-transcriptional level by either inducing target messenger RNA (mRNA) degradation or preventing mRNA translation^{11–13}. In the context of an EMT, miRNA-200 family members (miR-200a/b/c, miR-141 and miR-429) take a central stage: they are required to maintain an epithelial cell morphology by degrading the transcripts of the EMT-inducing transcription factors (TFs) Zeb1 and Zeb2^{14–17}. Members of the miR-200 family bind to specific seed sequences in the 3' untranslated region (3' UTR) of Zeb1 and 2 mRNAs and destabilize them. During an EMT, members of the miR-200 family are downregulated in their expression, which results in the increased expression of Zeb1 and 2. Conversely, Zeb1 and Zeb2 directly suppress the transcription of miR-200 family members^{14–17}. Such a double-negative feedback loop is a major example for a reciprocal TF-miRNA regulation during an EMT. Similar other molecular switches also regulate EMT/MET plasticity and malignant tumour progression^{18–22}.

Here, we report the identification of miR-1199-5p, as a repressor of EMT, tumour cell invasion and metastasis, which comparable to miR-200 family members targets Zeb1 mRNA for degradation. Conversely, Zeb1 represses the expression of miR-1199-5p and of the miR-200 family. However, miR-200 family members and miR-1199-5p seem to exert distinct functions; they share only six of their many target mRNAs, among them Zeb1.

Results

Identification of EMT-associated miRNAs. To identify regulatory miRNAs involved in the gradual process of an EMT, we performed miRNA sequencing on a detailed time course of a TGF β -induced EMT in normal murine mammary gland cells (NMuMG subclone E9; NMuMG/E9). Analysis of the kinetics of miRNA transcript regulation during an EMT in a time-resolved manner identified 32 differentially expressed miRNAs. Unsupervised hierarchical clustering illustrated that approximately half of the differentially expressed miRNAs showed a continuous increase in their expression during an EMT, whereas the other half exhibited decreased expression (Fig. 1a). In order to identify the miRNAs functionally impacting on a TGF β -induced EMT, we performed a microscopy-based screen in which NMuMG/E9 cells were transfected with miRNA mimics and cultured in the absence or presence of TGF β for 4 days (Fig. 1b). Subsequently, mesenchymal cell characteristics were monitored by high-content

fluorescence microscopy analysis and quantified, including the deposition of the extracellular matrix protein fibronectin and the formation of focal adhesions and of actin stress fibres²³. Nine out of 32 differentially expressed miRNAs were able to block or at least delay the EMT process: miR-125b-5p, miR-181b-2-3p, miR-1247-3p, miR-200a-3p, miR-200b-3p, miR-429-3p, miR-1199-5p, miR-145a-3p and miR-504-5p. Conversely, the forced expression of miR-145a-5p and miR-6944-3p promoted an EMT in epithelial NMuMG/E9 cells. Changes in cell morphology, the localization of the epithelial adhesion junction protein E-cadherin as well as the mRNA levels of E-cadherin, N-cadherin, fibronectin and Zeb1 further confirmed the impact of the different miRNAs on the EMT process (Supplementary Fig. 1a–c).

Increased cell migration is a functional output of EMT, which allows tumour cell invasion and dissemination into the surrounding tissue¹. In a second screening step, we identified those miRNAs that were able to affect mesenchymal tumour cell migration (Fig. 1c). Mesenchymal, highly migratory and tumorigenic Py2T cells that have been treated >20 days with TGF β ²⁴ were transiently transfected with mimics of EMT-affecting miRNAs. Four out of the 11 miRNAs tested significantly reduced cell migration in a trans-well Boyden chamber assay: miR-200b-3p, miR-429-3p, miR-1199-5p, and to a lesser extent miR-125b-5p (Supplementary Fig. 2a, b). As a result, our experimental strategy identified miRNAs whose transcriptional regulation and function affected both a TGF β -induced EMT and mesenchymal breast cancer cell migration (Fig. 1d). Four miRNAs significantly fulfilled these criteria: the miRNA-200 family members miR-200b-3p and miR-429-3p, and miR-125b-5p and miR-1199-5p. Since ectopic expression of miR-1199-5p, an EMT-regulatory miRNA, seemed to affect EMT with similar potency as the well-studied miR-200 family members^{15, 17, 18}, we further investigated the functional role and the mechanisms of action of miR-1199-5p in the regulation of an EMT.

miRNA-1199-5p inhibits EMT and tumour cell invasion. The continuous downregulation of miR-1199-5p expression during a TGF β -induced EMT was due to transcriptional repression as determined by a miR1199-promoter/luciferase-reporter assay during a TGF β -induced EMT in NMuMG/E9 cells and Py2T murine breast cancer cells, and in mesenchymal Py2T cells and metastatic 4T1 murine breast cancer cells treated with TGF β for >20 days (Fig. 2a, Supplementary Fig. 3a, b). Analysis of miR-1199 expression in different human breast cancer cell lines²⁵ further confirmed a higher expression in epithelial than in mesenchymal breast cancer cells (Supplementary Fig. 3c).

The forced expression of miR-1199-5p by the transfection of a construct encoding a 1199-5p miRNA mimic caused sustained epithelial cell morphology in NMuMG/E9 and in human untransformed mammary gland MCF10A cells induced to undergo an EMT by TGF β treatment (Fig. 2b–e, Supplementary Fig. 3d–f). The cells maintained their epithelial morphology and the cell surface localization of the adherens junction protein E-cadherin and of the tight junction protein ZO-1 at the plasma membrane, while miR-Ctr-transfected cells lost these epithelial characteristics and progressed with an EMT (Fig. 2b, Supplementary Fig. 3d). The reorganization of cortical actin to stress fibres as well as the formation of focal adhesions were also prevented by miR-1199-5p mimics (Fig. 2c). Mesenchymal markers, such as fibronectin, vimentin or N-cadherin were repressed at the protein (Fig. 2d, Supplementary Fig. 3e) and mRNA (Fig. 2e, Supplementary Fig. 3f) expression levels in miR-1199-5p-transfected cells. Furthermore, miR-1199-5p induced a significant decrease in Zeb1 mRNA, however, transcript levels of other key EMT TFs, such as Zeb2²⁶ or Sox4²⁷ remained unchanged (Fig. 2e).

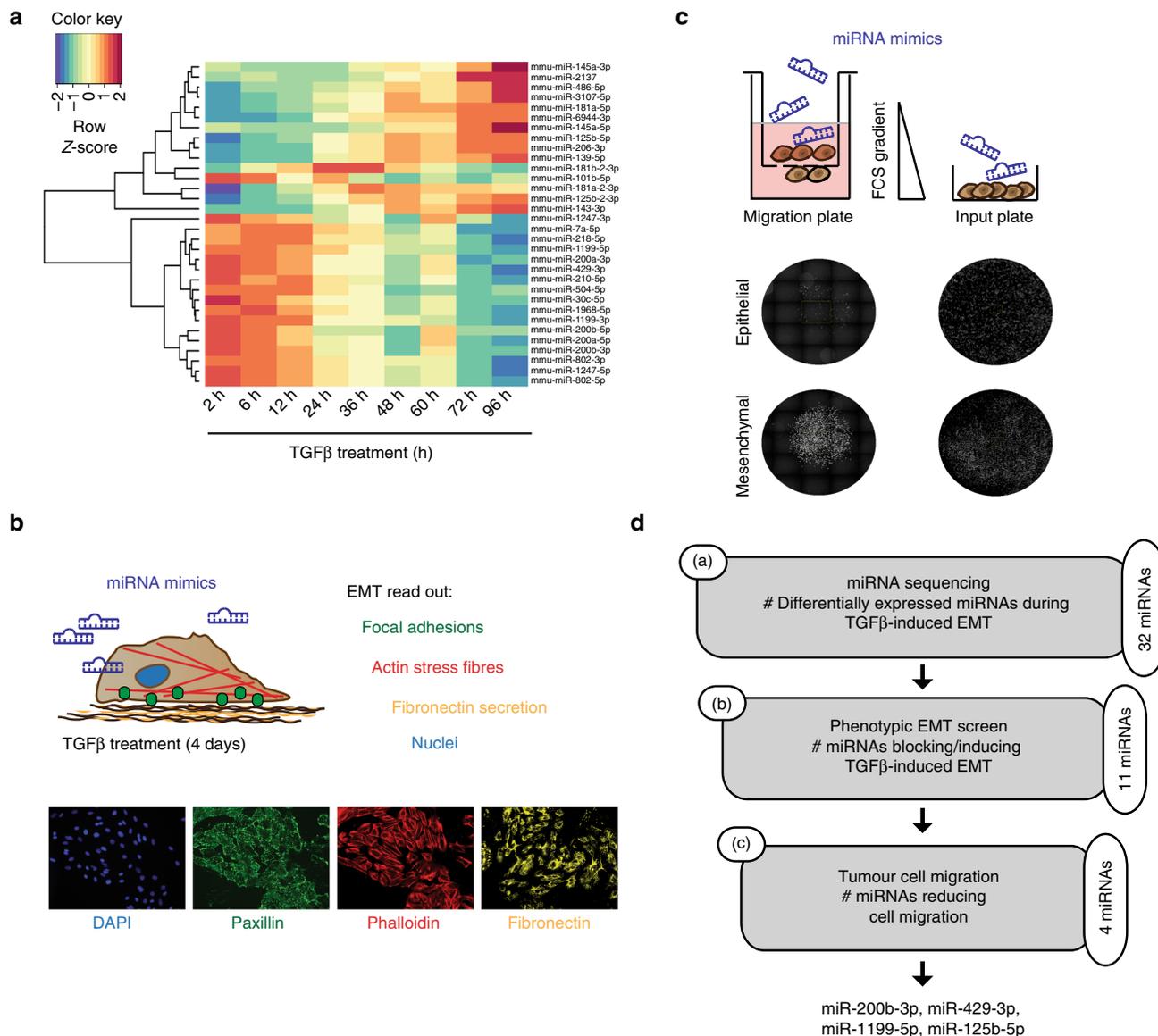


Fig. 1 Identification of miRNAs critical for an EMT and cell migration. **a** Expression profiling of miRNAs during a TGFβ-induced EMT in NMuMG/E9 cells. NMuMG/E9 cells were treated with TGFβ for the time points indicated, and total RNA was isolated for miRNA sequencing. The heat map summarizes the hierarchical clustering of differentially expressed miRNAs ($\log_2FC(\pm 2)$; $FDR < 0.05$) compared to epithelial, untreated cells during an EMT according to the indicated colour scale. **b** Identification of miRNAs controlling EMT. NMuMG/E9 cells were transfected with miRNA mimics for individual miRNAs and analysed for mesenchymal characteristics in the absence and presence of TGFβ for 4 days. Formation of focal adhesions (green), actin cytoskeleton reorganization to actin stress fibres (red), fibronectin deposition (yellow) and nuclei (blue) were visualized by high-content fluorescence screening microscopy. **c** Identification of miRNAs regulating mesenchymal tumour cell migration. Mesenchymal, migratory Py2T cells (>20 days TGFβ) were transfected with different miRNA mimics, plated in 96-well Boyden chamber migration inserts within a FCS gradient and in parallel on a 96-well input plate. Cell nuclei were imaged using a fluorescence screening microscope and quantified. Migrated cells were normalized to the total cell number on the input plate. **d** Scheme depicting the workflow and the number of **(a)** differentially expressed miRNAs during an EMT, **(b)** the number of identified miRNAs functionally contributing to a TGFβ-induced EMT and **(c)** the number and names of the miRNAs functionally contributing to mesenchymal tumour cell migration

A gain in cell migration and invasion can be one of the consequences of a TGFβ-induced EMT and allows tumour cells to leave the primary tumour and intravasate into the blood circulation and to extravasate at a distant organ^{1, 5}. Ectopic expression of miR-1199-5p displayed only minor effects on the early stages of an EMT in Py2T cells (Supplementary Fig. 3g-i), yet it induced a significant reduction in migration and invasion of mesenchymal (>20 days TGFβ) Py2T cells, as analysed by transwell Boyden chamber assays (Fig. 2f). The inhibitory effects of miR-1199-5p on in vitro cell migration and invasion were also

observed in 4T1 cells treated for >20 days with TGFβ (Fig. 2g). In summary, miR-1199-5p is sufficient to sustain an epithelial cell phenotype.

The direct target genes of miR-1199-5p during an EMT. To elucidate the genome-wide function of miR-1199-5p during an EMT, we performed RNA-sequencing analysis on NMuMG/E9 cells transiently transfected with either a miR-1199-5p mimic or a miR-Ctr mimic and cultured for 4 days in the presence of TGFβ.

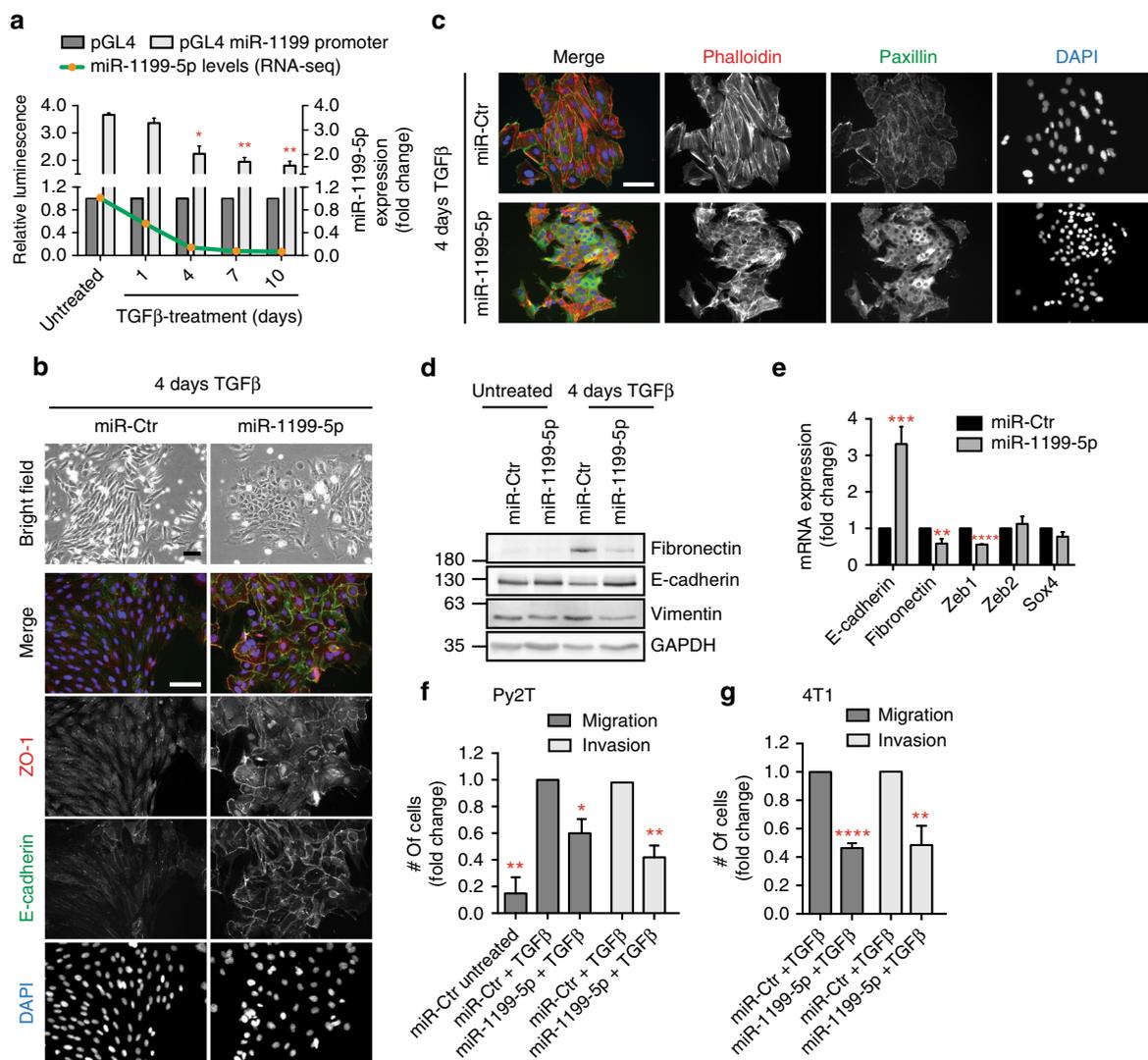


Fig. 2 miRNA-1199-5p inhibits an EMT and tumour cell migration and invasion. **a** miR-1199-5p transcript and miR-1199 promoter activity levels decrease during a TGFβ-induced EMT. Green line: expression profile of miR-1199-5p during EMT in NMuMG/E9 cells as determined by RNA sequencing. Grey bars: miR-1199 promoter activity during an EMT. NMuMG/E9 cells were treated with TGFβ for the time points indicated and transfected with a Renilla Luciferase reporter along with either a miR-1199 promoter Firefly Luciferase reporter (pGL4 miR-1199 promoter) or a control reporter (pGL4; Supplementary Fig. 3b). Relative luminescence (Firefly/Renilla) was calculated and normalized to the control reporter (mean fold changes ± s.e.m.; $n = 3$; significance determined by an unpaired, two-sided t test; * $P < 0.05$, ** $P < 0.01$). **b, c** NMuMG/E9 cells were transiently transfected with the miRNA mimics indicated and cultured in presence of TGFβ for 4 days (4-day TGFβ). Bright-field images of NMuMG/E9 cells illustrate the differences in cell morphology. Immunofluorescence images of NMuMG/E9 cells visualize different epithelial and mesenchymal cell structures as indicated. Scale bars: 100 μm. **d, e** Immunoblot (**d**) and quantitative RT-PCR mRNA (**e**) expression analysis of NMuMG/E9 cells transiently transfected with miRNA mimics for the EMT markers indicated (mean fold changes ± s.e.m.; $n = 6$; significance determined by an unpaired, two-sided t test, Student's t test; ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$). **f, g** Mesenchymal (>20 days TGFβ) Py2T (**f**) and 4T1 (**g**) cells were transiently transfected with the miRNA mimics indicated and re-plated within a FCS gradient of a Boyden chamber migration or invasion insert. Nuclei of transmigrated and invaded cells were quantified after 18 h (mean fold changes ± s.e.m.; migration: $n = 3$; invasion Py2T: $n = 3$, 4T1: $n = 4$; significance determined by an unpaired, two-sided t test; * $P < 0.05$, ** $P < 0.01$, **** $P < 0.0001$)

As expected, forced expression of miR-1199-5p induced a block in a TGFβ-induced EMT in these cells and led to an overall anti-correlative ($r = -0.324$) transcriptomic profile compared to genes regulated in their expression during an EMT (miR-Ctr 4d vs. miR-Ctr 0d) (Fig. 3a). We found 787 genes differentially expressed (log2 fold change of ± 1 (log2FC(± 1)); False Discovery Rate (FDR) < 0.05) by the ectopic expression of miR-1199-5p during an EMT. Subsequent functional annotation analysis by DAVID (Database for Annotation, Visualization and Integrated Discovery)^{28, 29} for biological processes (GO (gene ontology)) and cellular pathways (Kyoto Encyclopedia of Genes and Genomes; KEGG) revealed their involvement predominantly in cell

adhesion processes, extracellular matrix (ECM)-receptor interactions and focal adhesions (Fig. 3b).

In order to delineate the mechanism by which miR-1199-5p maintains an epithelial cell morphology, we set out to determine its direct mRNA targets during an EMT. Computational analysis by miRWalk³⁰ revealed 1789 potential target mRNAs of miR-1199-5p, which display a seed sequence in their 3' UTR, 5' UTR or coding sequence (CDS). Overlaying these mRNA targets with genes regulated by miR-1199-5p during an EMT (Fig. 3a; 787 genes) uncovered 90 target genes of which 66 displayed a significant reduction in their transcript levels upon the forced expression of miR-1199-5p (Fig. 3c, Supplementary Table 1).

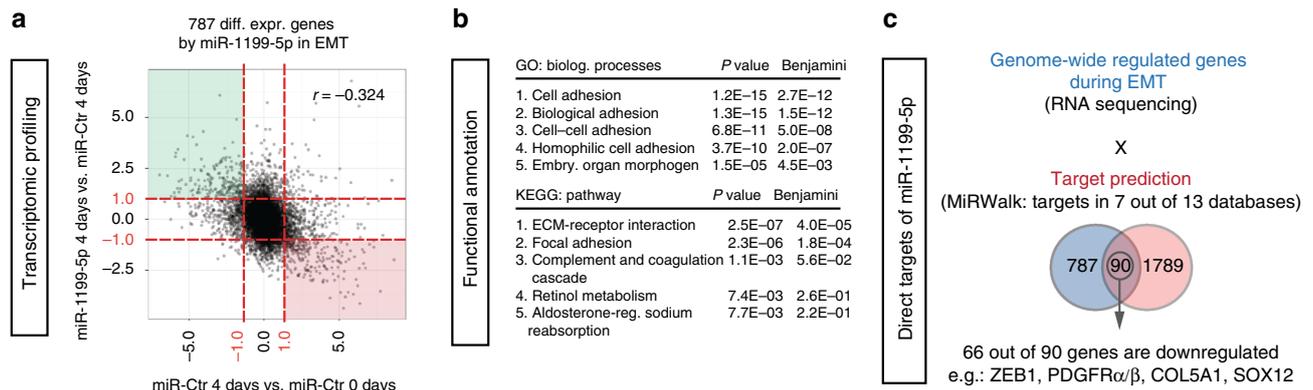


Fig. 3 Targets of miR-1199-5p function during an EMT. **a** Overall gene expression regulation by miR-1199-5p during a TGF β -induced EMT. RNA-sequencing analysis was performed on NMuMG/E9 cells transiently transfected with a miR-1199-5p or a negative control (miR-Ctr) mimic. Cells were cultured for 4 days in the presence (4 days; miR-Ctr- and miR-1199-5p-transfected cells) or absence of TGF β (only miR-Ctr-transfected cells). The scatterplot depicts overall gene expression (log₂FC) anti-correlation between miR-1199-5p 4 days vs. miR-Ctr 4 days over miR-Ctr 4 days vs. miR-Ctr 0 days. Differential expression analysis (red dashed line: log₂FC(\pm 1); False Discovery Rate (FDR) <0.05) identified 787 genes regulated by miR-1199-5p during an EMT. Green quadrant: increased gene expression; red quadrant: decreased gene expression. **b** Functional annotation clustering analysis of differentially expressed genes from **a**. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analyses by DAVID for the top five biological processes (top), pathways (bottom) and their associated *P* and Benjamini-Hochberg values. **c** Identification of miR-1199-5p direct targets in EMT. The Venn diagram depicts the number of genes differentially expressed upon miR-1199-5p expression during an EMT (blue), predicted direct targets of miR-1199-5p by miRWalk2.0 (red) and the number of overlapping genes. Sixty-six out of 90 genes demonstrate a downregulation in their expression upon forced expression of miR-1199-5p in EMT, including the EMT TF Zeb1

Among these genes, we identified the key EMT TF Zeb1³¹ with a potential conserved 8-mer seed sequence for miR-1199-5p in its 3' UTR (Fig. 4a).

miR-1199-5p directly targets Zeb1 mRNA. The zinc-finger E-box-binding homeobox TF Zeb1 has a well-established role as transcriptional repressor of epithelial genes and, hence, as an activator of an EMT^{3, 31}. Zeb1 function is associated with increased stemness, cell survival and metastasis^{32, 33}, and high levels of Zeb1 have been linked to aggressive breast cancer subtypes, therapy resistance, high risk for distant metastasis and poor survival^{34, 35}. As expected, Zeb1 transcript levels were increased during a TGF β -induced EMT of different murine and human cellular systems (Supplementary Fig. 4a). Furthermore, small-interfering RNA-mediated knockdown of Zeb1 maintained an epithelial cell morphology in NMuMG/E9 and MCF10A cells induced to undergo an EMT and significantly reduced the migratory properties of mesenchymal (>20 days TGF β) Py2T and 4T1 cells in trans-well Boyden chamber assays (Supplementary Fig. 4b–e).

We next examined whether miR-1199-5p indeed regulated Zeb1 levels during an EMT. Ectopic expression of miR-1199-5p in NMuMG/E9 and Py2T cells induced to undergo an EMT resulted in the stabilization of cell junction protein E-cadherin, while nuclear levels of Zeb1 were reduced (Fig. 4b, Supplementary Fig. 4f). Immunoblotting analyses for Zeb1 confirmed the repression of Zeb1 expression by miR-1199-5p in these murine cells (Fig. 4c, Supplementary Fig. 4g). Even though the seed match of human miR-1199-5p is not perfectly complementary to the seed sequence in the Zeb1 3' UTR (mouse and human miR-1199-5p seed matches differ in one base), forced expression of hsa-miR-1199-5p still significantly reduced Zeb1 mRNA (Supplementary Fig. 3f) and protein (Supplementary Fig. 4h) levels in human MCF10A cells. These results further imply that even imperfect binding of a miRNA to its target RNA can efficiently downregulate its expression.

To validate the direct regulation of Zeb1 by miR-1199-5p on the post-transcriptional level during an EMT we made use of two

Zeb1 3' UTR luciferase reporter constructs, one containing the species conserved, wild-type miR-1199-5p seed sequence and the other one carrying a mutated version of the seed sequence with five nucleotides exchanged (Fig. 4a, d). Transient transfection of a miR-1199-5p mimic and the Zeb1 3' UTR wild-type reporters in NMuMG/E9 cells revealed a significant decrease in luminescence, which was not observed with the mutant version of the reporter (Fig. 4d). Together, these data identify the key EMT TF Zeb1 as a direct target of miR-1199-5p, which thus represses Zeb1 expression at the post-transcriptional level and prevents an EMT.

Zeb1 directly controls the expression of miR-1199-5p. An important mechanism that appears to be responsible for EMT cell plasticity are double-negative feedback loops between miRNAs and key EMT TFs, functioning as molecular switches for various cell differentiation states^{18–20, 22, 36, 37}. Because miR-1199-5p regulates the expression of Zeb1, we assessed whether Zeb1 would in turn regulate the expression of miR-1199-5p during an EMT. siRNA-mediated ablation of Zeb1 expression in NMuMG/E9 cells cultured in the presence of TGF β for 4 days and transfected with either a miR-1199-promoter luciferase-reporter (pGL4 miR-1199 promoter) or a control promoter luciferase-reporter (pGL4; Supplementary Fig. 3b) revealed a significant increase in miR-1199 promoter activity upon loss of Zeb1 (Fig. 5a, left). Conversely, transient overexpression of Myc-tagged Zeb1 in epithelial NMuMG/E9 cells significantly decreased miR-1199 promoter activity (Fig. 5a, right). Furthermore, the regulation of miR-1199 promoter activity by loss or gain of function experiments of Zeb1 also correlated with an increase or decrease in endogenous miR-1199-5p transcript levels, respectively (Fig. 5b).

The gene encoding for miR-1199-5p is located within the first CDS of an unknown, protein-coding gene (2210011C24Rik) on chromosome 8 in the mouse genome. Its localization is conserved in the human genome, where it is located in the first CDS of LOC113230, an orthologue of 2210011C24Rik, on chromosome 19 (Supplementary Fig. 5a). Notably, the transcript levels of the murine host gene are significantly downregulated during a TGF β -induced EMT in different cellular models, as it was observed for

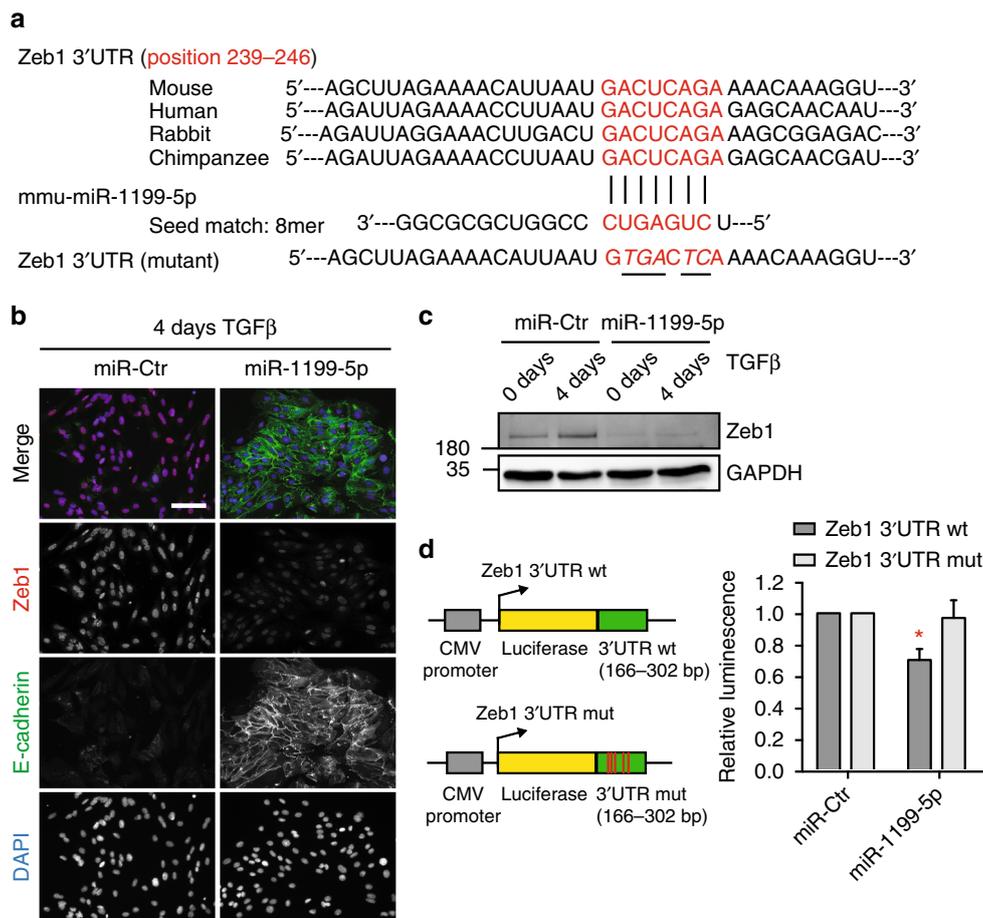


Fig. 4 miR-1199-5p directly controls Zeb1 expression. **a** Species-conserved miR-1199-5p-binding site in the 3' UTR of Zeb1. The scheme represents: top: sequence alignments of predicted binding sites of miR-1199-5p (red; mouse: position 239–246) in Zeb1 3' UTRs of different species. Middle: alignment of the 8mer seed match in mouse miR-1199-5p. Bottom: mutated miR-1199-5p seed sequence in the 3' UTR of Zeb1 mRNA utilized for the control reporter construct (Zeb1 3' UTR mut) in **d**. Exchanged nucleotides are underlined. **b, c** Forced expression of miR-1199-5p reduces Zeb1 nuclear localization and protein levels during an EMT. NMuMG/E9 cells were transiently transfected with miRNA mimics as indicated and cultured in the absence (0 day) or presence of TGFβ (4 days). Immunofluorescence (**b**) and immunoblotting (**c**) analyses illustrate the differences in Zeb1 protein levels. Scale bar: 100 μm. **d** Post-transcriptional regulation of Zeb1 by miR-1199-5p. NMuMG/E9 cells were transfected with the miRNA mimics indicated, with a Renilla luciferase reporter construct and with either a Zeb1 3' UTR wild-type (wt) or a Zeb1 3' UTR mutant (mut) Firefly luciferase reporter construct. Relative luminescence (Firefly/Renilla) was calculated and normalized to miR-Ctr-transfected cells (mean fold changes ± s.e.m.; $n = 3$; significance determined by an unpaired, two-sided t test; * $P < 0.05$)

miR-1199-5p (Supplementary Fig. 5b; Fig. 2a). Furthermore, gain and loss of function studies by siRNA-mediated ablation or transient overexpression of Zeb1 in NMuMG/E9 cells led to an increase or decrease in 2210011C24Rik expression, respectively (Supplementary Fig. 5c, d), indicating a co-regulation of miR-1199-5p and its host gene.

The murine miR-1199/2210011C24Rik promoter region encompasses several E-box motifs (CANNTG; $N = G$ or C) as potential Zeb1-binding sites (Fig. 5c). We next performed chromatin immunoprecipitation (ChIP) for endogenous Zeb1 in epithelial Py2T cells and in 4 days TGFβ-treated Py2T cells, a time point displaying a robust increase in Zeb1 expression (Supplementary Fig. 4a), followed by quantitative RT-PCR analysis with various primer pairs covering different regions of the miR-1199/2210011C24Rik promoter. These experiments confirmed a direct binding of Zeb1 to a region with the closest proximity (+51/–52 bps) to the transcriptional start site (TSS) (Fig. 5d). Notably, Zeb1 binding was only observed in cells undergoing an EMT and not in their epithelial counterparts (Fig. 5d). Using the miR-1199/2210011C24Rik promoter luciferase reporter (Supplementary Fig. 3b), we individually mutated

each of four Zeb1 E-box-binding sites (CAGGTG to CATTG). Only mutation of E-box 1 (–18 bp from TSS) significantly ablated Zeb1-mediated repression of luciferase activity in NMuMG/E9 cells (Fig. 5e).

Together, the results show that Zeb1 is a direct transcriptional repressor of the *miR-1199* gene. Hence, Zeb1 and miR-1199-5p act in a reciprocal fashion to repress each other.

Comparing miR-1199-5p and miR-200s function during an EMT. MiR-1199-5p's function during an EMT as well as its regulation by the EMT TF Zeb1 resembles the activities of members of the miR-200 family, two of which, miR-200b-3p and miR-429-3p, have been identified by our functional EMT screen (Fig. 1)^{14–17}.

Comparable to miR-1199-5p, miR-200b-3p and miR-429-3p transcript levels are also strongly reduced during a TGFβ-induced EMT in NMuMG/E9 cells (Supplementary Fig. 6a, Fig. 2a). The forced expression of miR-200b-3p or miR-429-3p in human MCF10A (Supplementary Fig. 6b, top) and murine NMuMG/E9 cells (Supplementary Fig. 1a–c) also prevented an EMT and

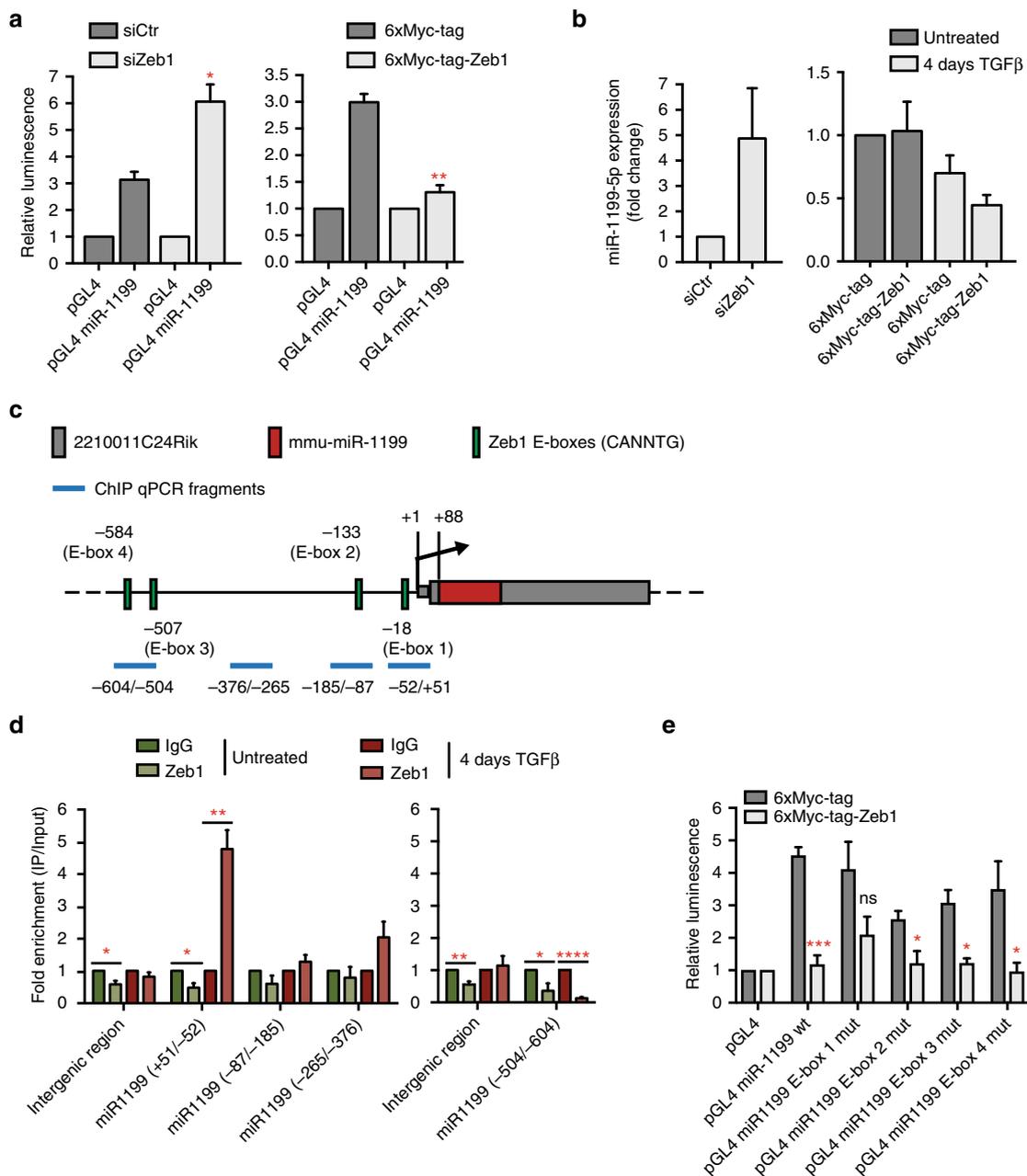


Fig. 5 Zeb1 directly regulates miR-1199-5p expression. **a** Zeb1 controls the promoter activity of the miR-1199 gene. NMuMG/E9 cells were transfected with a miR-1199 promoter Firefly luciferase reporter construct, a Renilla luciferase reporter construct as well as with (left) a siRNA against Zeb1 (siZeb1) or a negative control siRNA (siCtr) or (right) a Zeb1 expression construct (6xMyc-tag-Zeb1) or a negative control construct (6xMyc-tag). Cells transfected with siRNAs were cultured for 4 days with TGFβ, while cells transfected with 6xMyc or 6xMyc-tagged Zeb1 were cultured in the absence of TGFβ. Relative luminescence (Firefly/Renilla) was calculated and normalized to the control Firefly luciferase reporter (pGL4; mean fold changes±s.e.m.; n = 3; significance determined by an unpaired, two-sided t test; *P < 0.05, **P < 0.01). **b** Zeb1 controls miR-1199-5p transcript levels. NMuMG/E9 cells were transfected with siRNAs (left) and Zeb1 constructs (right) as described in **a** and further cultured in the absence or presence of TGFβ. MiR-1199-5p transcript levels were examined by RT-PCR analysis (mean fold changes±s.e.m.; left: n = 5; right: n = 3). **c** Schematic presentation of the genomic localization of the murine miR-1199 gene and its promoter region. Red: mmu-miR-1199; grey: 2210011C24Rik gene; green: E-boxes (CANNTG, N = G or C); blue: promoter fragments examined by ChIP-qPCR analysis. **d** Zeb1 directly binds to the miR-1199 promoter. Chromatin of Py2T cells cultured in the absence (green) and presence of TGFβ (red) was subjected to chromatin immunoprecipitation with antibodies against Zeb1 followed by RT-PCR analysis using primers amplifying different regions of the miR-1199 promoter illustrated in **c**. An intergenic region was used as negative control. Data were normalized to control IgG and are presented as mean fold enrichment above background±s.e.m. (n = 3; significance determined by an unpaired, two-sided t test; *P < 0.05, **P < 0.01, ****P < 0.0001). **e** Analysis of E-box-binding motifs in the miR-1199 promoter as potential Zeb1-binding sites. NMuMG/E9 cells were transfected and cultured as described in **a** (right). Relative luminescence (Firefly/Renilla) was calculated as described in **a** (mean fold changes±s.e.m.; n = 3; significance determined by an unpaired, two-sided t test; *P < 0.05, ****P < 0.001)

maintained an epithelial morphology in the presence of TGF β . However, similar to miR-1199-5p, expression of miR-200b-3p and miR-429-3p mimics in tumorigenic Py2T cells failed to completely block mesenchymal cell morphology, even though the mRNA levels of E-cadherin and Zeb1 were significantly increased or decreased, respectively (Supplementary Fig. 6b, c). Also comparable to miR-1199-5p, ectopic expression of miR-200b-3p or miR-429-3p significantly reduced cell migration and invasion of mesenchymal Py2T and 4T1 cells (Supplementary Fig. 2, Supplementary Fig. 6d). Notably, all three miRNAs are mechanistically embedded in a double-negative feedback regulation with Zeb1^{18, 19} (Fig. 4, Fig. 5, Supplementary Fig. 6e).

Since miR-200b-3p, miR-429-3p and miR-1199-5p induced comparable functional outputs with regard to an EMT process, we assessed their effects on tumour progression *in vivo*. Stable expression of miR-1199, miR-200b (which also induced the expression of miR-429 and slightly miR-1199; Supplementary Fig. 6f) and miR-429 in metastatic 4T1 cells tagged by ZsGreen led to a reduction in Boyden chamber trans-well cell migration compared to an empty vector control (Supplementary Fig. 6g) and confirmed our results for their transient overexpression (Fig. 2g, Supplementary Fig. 6d).

Upon orthotopic transplantation of 4T1 cells into the mammary fat pads of immunodeficient NSG and NMRI mice, the forced expression of miR-1199, miR-429 or miR-200b induced a significant reduction in primary tumour growth over time (Fig. 6a). Notably, the high potential of 4T1 cells to metastasize to the lungs was reduced by the expression of miR-1199 and miR-429, however not by the expression of miR-200b (Fig. 6b–d). The failure of miR-200b to repress 4T1 metastasis was surprising, yet has been observed by others as well (G.J. Goodall, personal communication of unpublished data). Examining the metastatic outgrowth of 4T1 cells in the lung, expression of each of the three miRNAs significantly increased the size of the few metastatic nodules compared to the many nodules observed with empty vector-transduced cells (Fig. 6e). Therefore, the repressive growth effect observed for miR-1199, miR-200b and miR-429 in the primary tumour was not maintained in the outgrowth of tumour cells at the metastatic site. Of note, the number of circulating tumour cells isolated from the blood of tumour-bearing mice was decreased by the forced expression of miR-1199 and miR-429, but not by the expression of miR-200b (Fig. 6f, g).

In summary, miR-1199 and miR-429, but not miR-200b, are sufficient to reduce tumour cell intravasation into the blood circulation and the seeding of lung metastases. However, all three miRNAs seem to promote metastatic outgrowth once tumour cells have seeded in the lung parenchyma.

Common and distinct targets of miR-200s and miR-1199-5p.

The functional similarities between miR-1199-5p and miR-200 family members observed during an EMT *in vitro* as well as during tumour progression *in vivo* begs the question about the shared and the distinct functions of miR-200 family members and miR-1199-5p in regulating an EMT. To reveal the transcriptomic effects of the various miRNAs during a TGF β -induced EMT, we performed RNA sequencing of NMuMG/E9 cells transfected with miRNA mimics for miR-200b-3p, miR-429-3p and miR-1199-5p. As expected, miR-200b-3p or miR-429-3p mimics induced a block in EMT of NMuMG/E9 cells and led to an overall anti-correlative gene expression profile compared to mesenchymal, miR-Ctr-transfected cells ($r = -0.505$ and $r = -0.51$, respectively; Fig. 7a), similar to the profile observed with miR-1199-5p mimic ($r = -0.324$; Fig. 3a). Differential gene expression analysis (log₂FC (± 1); FDR < 0.05) revealed 1097 and 1058 genes affected by miR-

200b-3p and miR-429-3p, respectively, during an EMT, of which 982 genes were shared by the two miRNAs (Fig. 7a, b). Since they belong to the same miRNA family, such a high number of co-regulated genes can be explained by an identical seed sequence on their target mRNAs. However, the difference in the effects of the two miRNAs on lung metastasis is surprising (Fig. 6b–d), and we can only speculate whether the small pool of individual target genes for each of the miRNAs is the reason for the different outcomes *in vivo*.

Directly compared to each other, miR-200b-3p, miR-429-3p and miR-1199-5p share 465 regulated genes during an EMT, which is more than half of the 787 genes regulated by miR-1199-5p alone (Fig. 7b, Fig. 3a). Functional annotation analysis revealed that these genes control the same biological processes and pathways as miR-1199-5p alone, that is, cell adhesion, ECM-receptor interactions and focal adhesions (Fig. 7c, Fig. 3b). A total of 267 genes are exclusively regulated by miR-1199-5p during an EMT.

Target prediction analysis by miRWalk2.0³⁰ for miR-200b-3p and miR-429-3p together with the differential expression analysis presented in Fig. 7a revealed 54 direct target genes commonly regulated by both miR-200 family members during an EMT (Fig. 7d, Supplementary Table 2), among them the key EMT TFs Snail1, Zeb1 and Zeb2³. Six out of the 54 genes were also directly controlled by miR-1199-5p (Fig. 7d). Besides the common target Zeb1, mRNA levels for Ncs1, Cdon, Sox12, Zfp9, Col5a1 were also decreased by all three miRNAs during an EMT (Fig. 7e). Of note, 58 target genes were uniquely regulated by miR-1199-5p.

The results show that miR-1199-5p, miR-200b-3p and miR-429-3p tightly control TGF β -induced EMT plasticity with similar potency, however, they only share the regulation of six gene transcripts, one of which is the critical EMT TF Zeb1. Yet, each of the miRNAs seems to have distinct functions by repressing a larger number of unique target mRNAs.

Discussion

A cancer-associated EMT can drive early steps of the metastatic cascade by converting epithelial tumour cells into invasive metastatic cancer cells^{1, 5}. MiRNAs are powerful, post-transcriptional regulators of a cell's transcriptome and, therefore, they are able to coordinate changes of a cell's morphology and functional capabilities necessary for an EMT^{11–13}. In the present study, we have set out to identify and characterize such EMT-regulatory miRNAs in normal and tumorigenic breast cancer cells. Using miRNA-sequencing analysis of the kinetics of EMT, we identified 32 strongly regulated miRNAs by a global, time-resolved miRNA profile of TGF β -induced EMT. Subsequently, we have tested the miRNAs' functional contribution to an EMT and to mesenchymal tumour cell migration. Among a number of miRNAs, we identify miR-1199-5p as a strong regulator of an EMT. Its transcript levels as well as its promoter activity are continuously decreased during an EMT. Furthermore, we demonstrate that miR-1199-5p downregulation is required for cell dedifferentiation, for mesenchymal tumour cell migration and invasion, and for lung metastasis formation *in vivo*. Finally, we report that during an EMT, this exonic miRNA is mechanistically embedded in a reciprocal, direct negative feedback loop with the TF Zeb1. The double-negative feedback loop with Zeb1 during an EMT is reminiscent of the miR-200 family members^{14–16, 18, 19}. Indeed, consistent with previous reports, the miR-200 family members miR-200b-3p and miR-429-3p have been identified by our EMT screen as critical mediators of an epithelial cell phenotype^{14–17}.

While the gain of function approaches of miR-1199-5p and miRNA family members has clearly repressed an EMT, we have

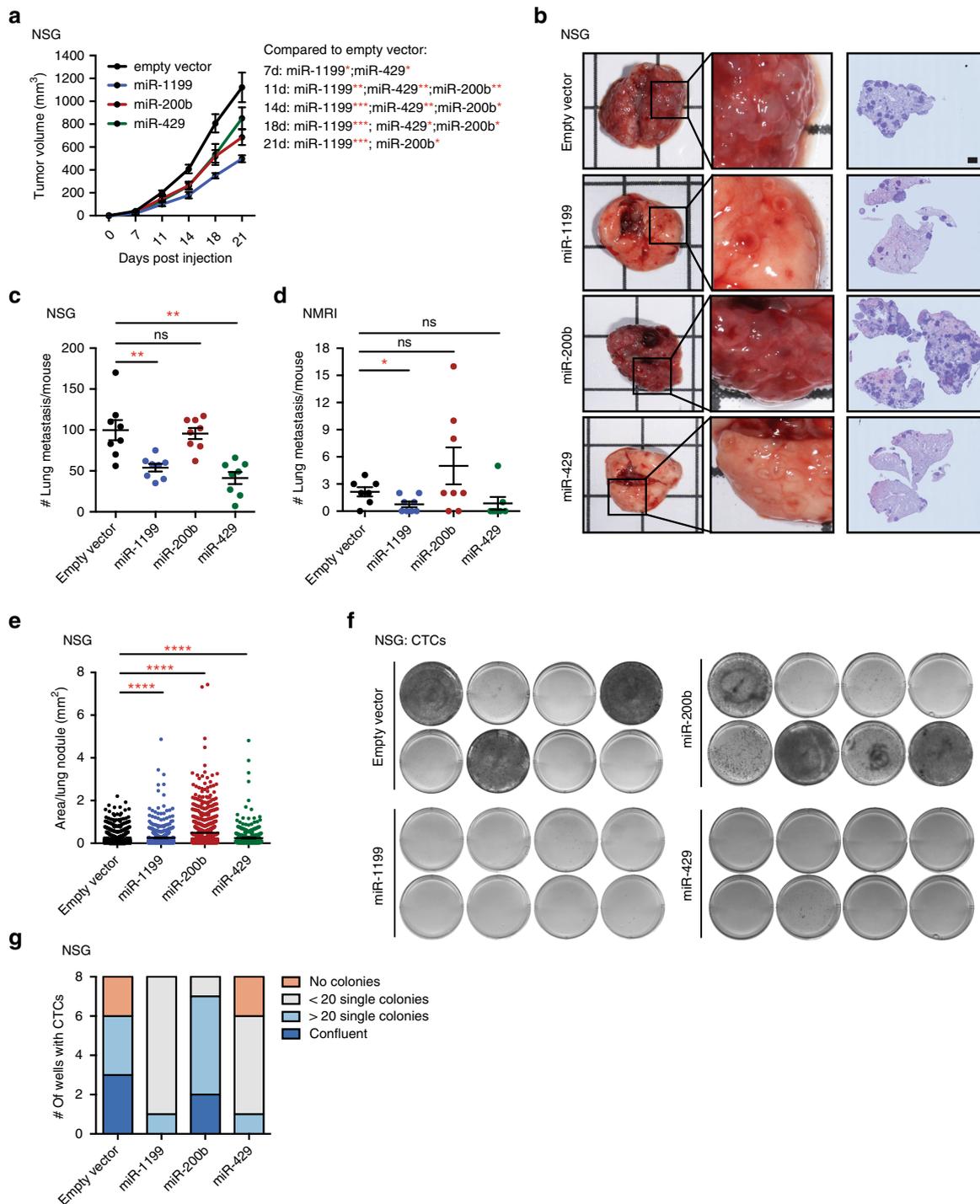


Fig. 6 miR-1199 and miR-200s repress 4T1 murine breast cancer primary tumour growth and metastasis. **a** ZsGreen-labelled 4T1 cells stably expressing miR-1199, miR-200b or miR-429 were injected into the mammary fat pad of female NSG or NMRI mice as indicated, and primary tumour growth was analysed over time. Statistical significance compared to empty vector control: 7 days: miR-1199*, miR-429*; 11 days: miR-1199**, miR-429**, miR-200b**; 14 days: miR-1199***, miR-429***, miR-200b*; 18 days: miR-1199***, miR-429***, miR-200b*; 21 days: miR-1199***, miR-200b*. **b** Representative bright-field images of whole lungs (left), enlargement of indicated lung areas (middle) and H&E staining of lung sections (right) isolated from mice killed 21 days post cell injection. Scale bar, 1 mm. **c, d** The number of lung metastases per mouse was quantified microscopically by H&E staining. **e** The tissue area of all lung nodules was analysed and quantified microscopically by H&E staining. **f, g** CTCs isolated from the blood of tumour-bearing mice were isolated 21 days post cell injection and cultured for 5 days ex vivo. Cell colonies were visualized by MTT staining, imaged (**f**) and quantified (**g**) as indicated. NSG: 8 mice per group ± s.e.m.; NMRI: 7–8 mice per group ± s.e.m.; statistical analyses: Mann-Whitney *U* test; **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001

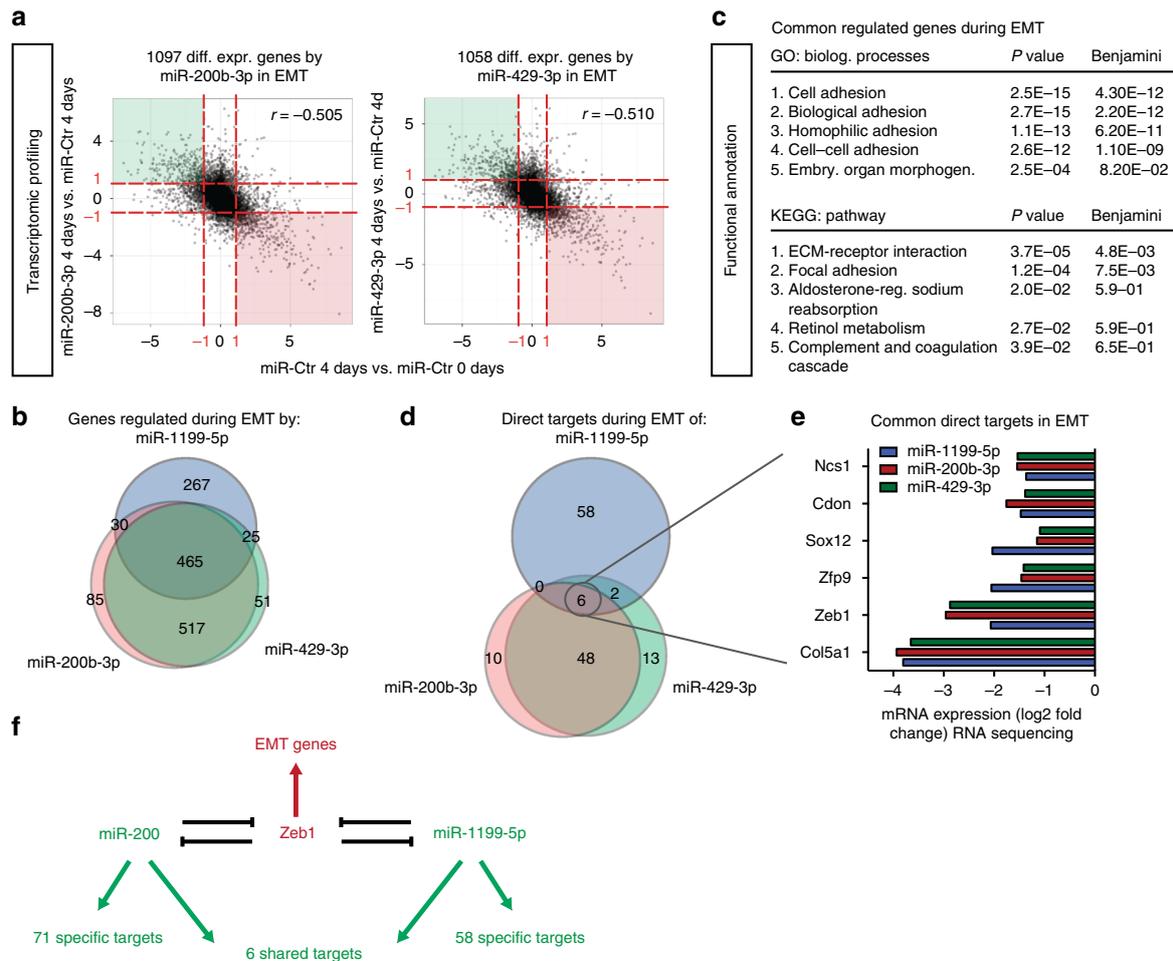


Fig. 7 Shared and distinct target genes of miR-1199-5p, miR-200b-3p and miR-429-3p. **a** Overall gene expression regulation by miR-200b-3p and miR-429-3p during a TGFβ-induced EMT. RNA-sequencing analyses were performed on NMuMG/E9 cells transiently transfected with a miR-200b-3p, miR-429-3p or a negative control (miR-Ctr) mimic. Cells were then cultured in the presence or absence of TGFβ (0 day vs. 4 days). RNA sequencing and data analysis was performed as described for miR-1199-5p in Fig. 3a. Scatterplots depict the overall gene expression (log2FC) (anti-)correlation between miR-200b-3p 4 days vs. miR-Ctr 4 days (left) or miR-429-3p 4 days vs. miR-Ctr 4 days (right) over miR-Ctr 4 days vs. miR-Ctr 0 day. Differential expression analysis (red dashed line: log2FC(±1); FDR <0.05) identified 1097 genes and 1058 genes regulated during an EMT by miR-200b-3p and miR-429-3p, respectively. **b** The Venn diagram summarizes the number of genes commonly and individually regulated by miR-1199-5p (blue), miR-200b-3p (red) and miR-429-3p (green) during a TGFβ-induced EMT (data from Fig. 7a, Fig. 3a). **c** Functional annotation clustering analysis by DAVID of genes commonly regulated by miR-1199-5p, miR-200b-3p and miR-429-3p during an EMT. Presented are the top five biological processes, pathways (KEGG) and their associated *P* and Benjamini-Hochberg values. **d** The Venn diagram presents the number of commonly and individually regulated, predicted target genes of miR-1199-5p (blue), miR-200b-3p (red) and miR-429-3p (green) during an EMT. MirWalk2.0 was used to predict miRNA target genes that show reduced transcript levels upon forced expression of miR-1199-5p, miR-200b-3p and miR-429-3p during an EMT. **e** Shown are the mRNA levels of genes commonly repressed by miR-1199-5p, miR-200b-3p and miR-429-3p during an EMT (RNA-sequencing data: log2FC compared to miR-Ctr-transfected cells). **f** Schematic representation of the double-negative feedback loops between miR-200 family members and Zeb1 and between miR-1199-5p and Zeb1. Only six target genes are shared between miR-200 and miR-1199-5p among which is Zeb1. Each of them has individual target genes as indicated and thus may affect distinct biological functions

repeatedly failed to demonstrate an effect of loss function approaches on EMT by interfering with miR-1199-5p and also with miR-200 family member activities. Interestingly, transient transfection of a miR-1199-5p sponge construct or a miR-1199-5p inhibitor did not lead to a loss of miR-1199-5p function in epithelial NMuMG/E9 cells, as determined by 3' UTR reporter assays, yet was not sufficient to increase Zeb1 expression or to induce an EMT. Of note, sponge constructs or miRNA inhibitors against miR-200b-3p and miR-200c-3p also did not alter Zeb1 expression or the induction of an EMT (Supplementary Fig. 7). From these data, we conclude that the regulation of Zeb1 expression on the transcriptional and posttranscriptional level is complex and we suppose that the loss of one Zeb1 regulator

(miR-1199-5p or miR-200b/c-3p) is not sufficient to increase Zeb1 levels and to jumpstart an EMT. The overlapping double-negative feedback loops between miRNAs and TFs may act as a 'buffered system' to compensate for the loss of one regulator and thus ensure a system's functionality.

Our genome-wide analyses of the three miRNAs revealed a large number of regulated genes, which underscores their critical functions during an EMT. However, only six genes were identified as common direct targets of the three miRNAs, including Zeb1, Col5a1, Zfp9, Sox12, Cdon and Ncs1. The latter five are so far 'unknowns' in the context of an EMT and malignant tumour progression, most of which might be of interest considering their potential function. For instance, Sox12, together with Sox11 and

Sox4, is a member of the SoxC TF family, of which Sox4 has an established role in EMT and breast cancer progression²⁷. The cell surface receptor encoded by Cdon, similar to the EMT inducer neuronal cell adhesion molecule, consists of immunoglobulin/fibronectin type III domains known to mediate cell signalling³⁸. Finally, collagen type V alpha 1 (Col5a1) is mainly found in organ tissue together with collagen type I, a major ECM component of breast cancer. Furthermore, Col5a1 facilitates breast cancer formation, invasion and metastasis^{39, 40}.

In summary, we report the identification and characterization of miR-1199-5p as a miRNA acting in a negative reciprocal regulation with Zeb1 during an EMT. Our findings add this negative feedback loop to other known reciprocal regulations between a miRNA and a key EMT TF^{18, 20–22}. These regulatory units may work as functional epithelial/mesenchymal switches supporting the reversibility, robustness and effectiveness of EMT/MET, endowing cells with a high degree of cell plasticity necessary for embryonic development as well as for malignant tumour progression.

Methods

Reagents and antibodies. E-cadherin (BD Transduction Labs, 610182), N-cadherin (Takara, M142), ZO-1 (Zymed, 617300), paxillin (BD, 610052), fibronectin (Sigma-Aldrich, F3648), vimentin (Sigma-Aldrich, V2258), GAPDH (Sigma-Aldrich, G8795), Zeb1 (Cell Signaling, 3396 and Santa Cruz Biotechnology, sc-25388), Alexa-Fluor 488 and 568 (Molecular Probes), secondary horse radish peroxidase (HRP)-conjugated antibodies against mouse and rabbit (Jackson ImmunoResearch), recombinant human TGFβ1 (R&D Systems, 240-B), 4',6-diamidino-2-phenylindole (DAPI, Sigma-Aldrich) and phalloidin Alexa-Fluor 568 (Molecular Probes, A12380).

Cell lines and cell culture. A subclone of normal murine mammary gland cells (NMuMG/E9)⁴¹, the MCF10A human epithelial mammary gland cell line⁴², the Py2T murine breast cancer cell line²⁴ and the murine metastatic cell line 4T1⁴³ have been described previously. NMuMG/E9, Py2T and 4T1 cells were grown in Dulbecco's modified eagle medium (DMEM; Sigma-Aldrich) supplemented with foetal calf serum (FCS, 10%; Sigma-Aldrich), glutamine (2 mM; Sigma-Aldrich), penicillin (100 U; Sigma-Aldrich) and streptomycin (0.2 mg/l; Sigma-Aldrich). MCF10A cells were cultured in DMEM/F12 medium supplemented with horse serum (5%; Biocconcept Amimed), insulin (10 µg/ml; Sigma-Aldrich), hydrocortisone (0.5 µg/ml; Sigma-Aldrich), human EGF (0.02 µg/ml; Invitrogen) and cholera toxin (0.01 µg/ml; Sigma-Aldrich). All cell lines were grown at 37 °C, 5% CO₂, 95% humidity.

Plasmids. pCS3-6xMyc-tag-Zeb1 and pCS3-6xMyc-tag-Zeb2 were a kind gift from T. Brabletz (Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany). The control plasmid pCS3-6xMyc-tag was generated by digesting a pCS3-6xMyc-tag-Zeb2 vector with EcoRI/XbaI, thereby removing the Zeb2 gene but retaining the 6xMyc tag.

miR-1199 promoter reporter. The genomic promoter region of miR-1199/2210011C24Rik (808 nucleotides upstream of 2210011C24Rik TSS) of NMuMG/E9 cells was PCR-amplified with the following primers: 5'-tcgactcgaggaccgggaaacactctgta-3' and 5'-tcgaaagcttgcgtctccactgcaattccgc-3', which exhibit restriction enzyme sites for XhoI and HindIII, respectively (underlined). The PCR amplicon was digested with XhoI and HindIII and subcloned into the pGL4.10[luc2] Firefly luciferase reporter vector purchased from Promega.

Using the QuikChange II site-directed mutagenesis Kit (Agilent), four different miR-1199/2210011C24Rik promoter luciferase mutants exhibiting point mutations in the Zeb1 E-box-binding motifs (wt: CAGGTG; mut: CATTG) at positions -18, -133, -507 and -584 bp from the TSS were generated from the original miR-1199/2210011C24Rik promoter reporter. The following primers were used to introduce the different point mutations:

E-box 1 (-18 bp):
5'-ctcggctcagaggctccattgtgtttctaccgaag-3',
5'-cttcggtagaaaacacaaatggagcctctgagccgag-3';
E-box 2 (-133 bp):
5'-cagagagccaggtcccagcttggcaggtgagc-3',
5'-gctcaactggcaaatgcccagctgctctctg-3';
E-box 3 (-507 bp):
5'-ccactgaaatccagctggcattggacggggcg-3',
5'-cgccccgtcaaatgcccagctgattcagtg-3';
E-box 4 (-584 bp):
5'-ccttcagtcacccgtgtgtttctggcggaatccg-3',
5'-cggattcccgaatgaacacacgggtgactgaagg-3'.

Zeb1 3'UTR reporter. A mouse Zeb1 3' UTR luciferase reporter for miR-1199-5p was generated by subcloning the murine Zeb1 3' UTR sequence (166–302 bps) containing a wild-type (wt: GACTCAGA) or mutated miR-1199-5p-binding site (mut: GTGACTCA) into the pMIR-REPORT Firefly luciferase vector (Applied Biosystems) via SpeI/HindIII sites. The wild-type and mutant Zeb1 3' UTR luciferase reporter for miR-200b-3p and miR-429-3p contain two common miRNA-binding sites (wt: CAGTATTA, CAGTATT; mut: CTC AATAA, GTCTAA) were generated by subcloning the murine Zeb1 3' UTR sequence (276–431 bps) into the pMIR-REPORT vector as described for Zeb1 3' UTR/miR-1199-5p. The different DNA fragments were synthesized by Integrated DNA Technologies.

Plasmids expressing ZsGreen1 and mmu-miR-1199 or mmu-miR-200b or mmu-miR-429. Stem loop sequences of murine miR-1199, miR-200b and miR-429 plus additional 100 bps up and downstream of the miRNAs were synthesized by Integrated DNA Technologies (miR-1199) or PCR-amplified from genomic DNA of NMuMG/E9 cells with the primers:

miR-200b:
5'-tcgaacgcgctctcctctgcaatgctctg-3',
5'-tcgagctagcctaactcttgcctccatagcc-3';
miR-429:
5'-tcgaacgcgctggaagggtgaacccaagaat-3' and
5'-tcgagctagcagcggggcctgtatattt-3'.

The miRNA constructs were flanked with MluI and NheI sites (underlined) and subsequently digested and subcloned into the 3' UTR of ZsGreen1 of the pmRi-ZsGreen1 vector (Clontech) via the restriction enzyme sites mentioned.

miRNA sponge constructs: constructs with 5× miR-1199-5p, miR-200b/c-3p or scrambled bulged binding sites separated by a spacer sequence (GAATAT) were synthesized and inserted in the 3' UTR of an enhanced green fluorescent protein reporter gene driven by the CMV promoter (pEGFP-C3, Clontech) by Invitrogen.

RNA interference. *miRNA mimic transfection:* NMuMG/E9, MCF10A, Py2T and 4T1 cells were transfected with 20 nM or 50 nM of mouse or human pre-miR miRNA precursors (Ambion) by using Lipofectamine RNAiMax (Invitrogen) according to the manufacturer's instructions. A random sequence has been used as negative control (Ambion, AM17110 control #1). To maintain a miRNA's over-expression, the transfection was repeated every third day. All pre-miR miRNA precursors used in this study are listed in Supplementary Table 3.

miRNA inhibitor transfection: NMuMG/E9 cells were transfected with 50 nM of miRNA inhibitors against miR-1199-5p from Ambion (miRVana; MH13577) and Exiqon (miRCURY LNA; 4108998).

siRNA transfection: 20 nM of siRNA against Zeb1 (Ambion: mouse: s74841, 74843 10 nM each; human: s229971) or a negative control (Ambion: 4390846) was used for transient gene knockdown experiments. Lipofectamine RNAiMax (Invitrogen) was used for the transfection according to the manufacturer's instructions and was repeated every third day.

Stable miRNA overexpression. 4T1 cells were serially transfected with pTet-On Advanced and plasmids expressing ZsGreen1 and mmu-miR-1199 or mmu-miR-200b or mmu-miR-429 or ZsGreen1 only using Lipofectamine 2000 (Invitrogen) and selected with Neomycin and Hygromycin (Clontech). Additionally, cells stably expressing ZsGreen1 were sorted by flow cytometry, and ectopic miRNA expression was validated by quantitative RT-PCR.

Luciferase reporter assays. A dual-luciferase reporter assay (Promega) was used to measure miR-1199/2210011C24Rik promoter activity and posttranscriptional repression of Zeb1 3' UTRs by miR-1199-5p, miR-200b-3p and miR-429-3p according to the manufacturer's instructions. Firefly/Renilla luciferase activity was measured with a luminometer (Berthold Technologies; Centro LB 960).

miR-1199/2210011C24Rik promoter reporter: Cells were plated in triplicates in a 24-well plate and treated with TGFβ. Three days before measuring luciferase activity, cells were transfected with 0.5 µg of pGL4-miR-1199/2210011C24Rik or pGL4 Firefly luciferase promoter reporter by using Lipofectamine 3000 (Invitrogen). Additionally, cells were transfected with 10 ng of pRL-CMV (Promega), which encodes a Renilla luciferase and was used for subsequent cell number normalization.

To test miR-1199/2210011C24Rik promoter activity upon Zeb1 gain and loss of function studies, NMuMG/E9 cells were plated as described before and reverse transfected with 20 nM of siRNAs against Zeb1 or a negative control (Ambion). On the next day, cells were transfected with the different reporter constructs mentioned above and treated with TGFβ for 3 days. For Zeb1 gain of function studies, NMuMG/E9 cells were forward transfected with 50 ng of pCS3-6xMyc-tag or pCS3-6xMyc-tag-Zeb1 along with the different luciferase reporters.

Zeb1 3'UTR reporter: NMuMG/E9 cells were plated and reverse transfected with 20 nM of miR-1199-5p, miR-200b-3p, miR-429-3p pre-miR miRNA precursors or a negative control (Ambion) as described above. On the next day, cells were transfected with 0.2 µg of pMIR-Report Zeb1-3' UTR/miR-1199/miR-200b/miR429 wild-type or mutant vector along with 10 ng of pRL-CMV by using Lipofectamine 3000 (Invitrogen).

RNA isolation and quantitative RT-PCR. In order to quantify gene transcripts, total RNA was isolated using Tri Reagent (Sigma-Aldrich) following the manufacturer's instructions. Complementary DNA (cDNA) was generated by reverse transcription of RNA using M-MLV reverse transcriptase (Promega) and was quantified by real-time PCR using Mesa Green qPCR MasterMix plus (Eurogentec). Riboprotein L19 was used as internal normalization control. FCs were calculated using the comparative Ct method ($\Delta\Delta C_t$). Primers used for quantitative RT-PCR are listed in Supplementary Table 4.

For quantitative miRNA analysis, total RNA was isolated using miRNeasy kit (Qiagen). Mature miRNAs were reverse-transcribed by using the miRCURY LNA Universal RT miRNA PCR kit (Exiqon). MiR-1199-5p (Exiqon; 206004) expression was measured by RT-PCR and normalized to U6 small nuclear RNA (Exiqon; 203907) expression as internal control. FCs in miRNA expression were calculated using the comparative Ct method ($\Delta\Delta C_t$).

Immunoblotting. Cells were lysed on ice in RIPA buffer (50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 10% glycerol, 1% NP40, 0.5% NaDOC, 0.1% SDS, 2 mM MgCl₂, 2 mM CaCl₂, 1 mM DTT, 1 mM NaF, 2 mM Na₃VO₄ and 1× protease inhibitor cocktail (Sigma-Aldrich)) for 20 min, centrifuged and protein concentrations were determined via Bio-Rad Bradford solution according to the manufacturer's instructions. Proteins were mixed with SDS-PAGE loading buffer (10% glycerol, 2% SDS, 65 mM Tris, 1 mg/100 ml bromophenol blue, 1% β -mercaptoethanol) and equally loaded on a SDS polyacrylamide gel. After size fractionation, proteins were transferred on an Immobilon-P PVDF membrane (Millipore) using an electrophoretic transfer cell (Bio-Rad) and blocked with 5% skim milk powder dissolved in TBS/0.05% Tween for 1 h. The membrane was incubated with primary antibodies at 4 °C overnight or at room temperature for 1 h. HRP-conjugated secondary antibodies were used to visualize specific proteins on a Fusion Fx7 chemoluminescence reader (Vilber Lourmat). Uncropped immunoblot scans from main blots are displayed in Supplementary Fig. 8.

Immunofluorescence microscopy. Cells were grown on uncovered glass cover slips (Menzel-Glaser), fixed with 4% paraformaldehyde/PBS for 20 min, permeabilized with 0.5% NP40 for 5 min and blocked with 3% BSA/0.01% Triton X-100/PBS for 30 min. Afterwards, cells were incubated at room temperature with the primary antibodies (listed above) diluted in 3% BSA/0.01% Triton X-100/PBS for 1.5 h, followed by an incubation with a fluorophore-coupled secondary antibody (Alexa Fluor, Invitrogen) for 1 h. Cell nuclei were visualized with DAPI (Sigma-Aldrich). After staining, the cells were mounted in fluorescence mounting medium (Dako) on microscope slides and imaged using fluorescence microscopy (Leica DMI 4000).

EMT phenotypic microscopy-based screen. The general setup of the screen has been described recently²³. In brief, NMuMG/E9 cells were plated as duplicates in a 96-well plate and reverse transfected with 20 nM pre-miR miRNA precursors (Ambion; Supplementary Table 4) using Lipofectamine RNAiMax (Invitrogen) 2 days prior to the addition of TGF β for 4 days. Epithelial control cells were transfected for 3 days in the absence of TGF β . For immunofluorescence analysis, cells were processed as described above and mesenchymal characteristics, such as formation of focal adhesions (paxillin), actin stress fibres (phalloidin) and fibronectin deposition, were stained as described in ref. ²³. Images of cells in a 96-well plate were taken with an Operetta HCS microscope (Perkin Elmers). EMT features were quantified (Columbus software, version 2.5.0) and the median of two replicates was calculated and compared to the median value of miR-Ctr-transfected cells. Standard deviations (s.d.) were estimated. MiRNAs showing at least 3× s.d. for all three EMT read-outs or 4× s.d. in two EMT read-outs aside from the median value of miR-Ctr-transfected cells were considered as hits to block or induce EMT. Some miRNAs, which turned out to be no hits in the EMT phenotypic microscopy-based screen, but maintained an epithelial cell morphology of NMuMG/E9 cells cultured in the presence of TGF β (judged by eye), were also considered hits.

Trans-well migration and invasion assays. Boyden chamber invasion assay (24-well plate format): mesenchymal Py2T and 4T1 cells were reverse transfected with 50 nM of pre-miR miRNA precursors in a six-well plate. Two days later, the cells were trypsinized, washed once with PBS and counted. A total of 25,000 cells were plated as duplicates in a 24-trans-well migration or invasion insert (membrane with 8 μ m pores, covered with a layer of growth factor reduced Matrigel; BD Biosciences). After 18 h, the cells were fixed with 4% paraformaldehyde/PBS, nuclei were stained with DAPI (Sigma-Aldrich) and non-migrated/invaded cells on the upper surface of the membrane were removed with a cotton swab. Migrated/invaded cells on the bottom of the membrane were imaged with a fluorescence microscope (Leica DMI 4000) and quantified. As chemo-attractant, a gradient of 0.2–20% FCS was used.

Boyden chamber migration screen (96-well plate format): mesenchymal Py2T cells were reverse transfected with 50 nM of pre-miR miRNA precursors in a 96-well plate. After 2 days, cells were trypsinized and washed with PBS in a round-bottom 96-well plate (Corning). MiR-Ctr-transfected cells were counted and 6000 cells were plated as duplicates in the inserts of a 96-well FluoroBlok fluorescent-blocking high-density positron emission tomography insert (BD Biosciences) and,

in parallel, in a 96-well reference plate (Cell Carrier 96, Perkin Elmers). The same cell suspension volume was used to plate cells transfected with other pre-miR miRNA precursors. As chemo-attractant, a gradient of 0.2–20% FCS was used. Migrated cells (bottom of membrane of migration insert) as well as cells in the reference plate were stained with DAPI and imaged with an Operetta HCS microscope (Perkin Elmers), quantified (Columbus software, version 2.5.0) and normalized to each other.

Chromatin immunoprecipitation. ChIP experiments were performed as previously described⁴⁴. Briefly, crosslinked protein-bound DNA of Py2T cells was sonicated (Bioruptor, Diagenode) to achieve chromatin fragments of an average size of 300 bps. For ChIP of endogenous Zeb1, 150 μ g of chromatin was incubated with 10 μ g of Zeb1 antibody (sc-25388) and immunocomplexes were precipitated with 40 μ l of pre-blocked magnetic Protein G beads (Invitrogen). Immunocomplexes were eluted from the beads, de-crosslinked, and genomic DNA was purified by phenol/chloroform extraction and precipitated with sodium acetate. One out of forty of the ChIP sample and 1% of input DNA were used for quantitative RT-PCR. Fold enrichments for specific miR-1199 promoter regions were calculated by IP over input samples and normalized to isotype-specific IgG as negative control. Primers targeting different genomic regions of the miR-1199 promoter are listed below:

–52/+51 bp from TSS: fwd: 5'-AGTTGTGCCCTGTCTCG-3'; rev: 5'-GATGGCGTCTCCATCTG-3'.

–185/–87 bp from TSS: fwd: 5'-CCTCTGGAGAGGAGCACTTG-3'; rev: 5'-CCCCAGTACCTCCGTTATACT-3'.

–265/–376 bp from TSS: fwd: 5'-TATTTGGGCATCTCAATTTCG-3'; rev: 5'-GACCCAGGACTCCACTCTC-3'.

–604/+504 bp from TSS: fwd: 5'-CAGTCAACCCGTGTGTTTCAG-3'; rev: 5'-GTCACCTGCCAGCTGGATT-3'.

Microarray data and analysis. In order to compare the expression of the miR-1199-5p in human cells, we downloaded raw data (CEL files) of the experiment GSE32474²⁵ from the Gene Expression Omnibus database. Data was analysed in R (<https://www.r-project.org/>)/Rstudio (<https://www.rstudio.com>) using Bioconductor add-on packages. The background signal correction, normalization and summarization were performed by robust multiarray averaging⁴⁵ from the affy package⁴⁶. To identify differences in gene expression between epithelial and mesenchymal samples, the linear models for microarray data (limma)⁴⁷ package was used.

miRNA sequencing and analysis. Total RNA was isolated from NMuMG/E9 cells undergoing TGF β -induced EMT by using the miRNeasy Mini Kit (Qiagen) according to the manufacturer's instructions. Quality and quantity of total RNA was analysed by using the RNA 6000 Pico kit from Agilent. MiRNA-sequencing libraries of two biological replicates were prepared with the Illumina TruSeq Small RNA Sample Prep as described by the manufacturer. cDNA libraries were size fractionated on a 6% Tris-Borate-EDTA (TBE) gel and fragments of 140–160 nt were extracted. Libraries were loaded and sequenced on an Illumina HiSeq 2500 using protocols defined by the manufacturer.

For bioinformatics analysis, sequencing adapters were removed from obtained single-end miRNA-seq reads (51-mers) using the preprocessReads function from R/Bioconductor package QuasR, version 1.0.9 43: 'preprocessReads(filename = tab\$FileName, outputPath = tab\$FileNameCleaned, Rpartner = "TGGAATTCTC GGGTGCCAAGGAACTCCAGTCA")'. Preprocessed reads were mapped to mouse genome assembly, version mm10, with Bowtie allowing up to 50 hits in the genome (44, included in the QuasR package) using the command: 'qAlign ("samples_preprocessed.txt", "BSgenome.Hsapiens.UCSC.mm10", maxHits = 50)'. The coordinates of mature miRNAs downloaded from miRBase (<http://www.mirbase.org/>, v21) were extended by 3 bp on each side and passed as an argument to the qCount function, by which we quantified expression as the number of reads that started within any mature miRNA. The differentially expressed miRNAs were identified using the edgeR package (version 3.2.4). MiRNAs with FDR <0.05 and minimum log₂ FC of ± 2 were used for further analysis.

Whole-transcriptome RNA sequencing and analysis. NMuMG/E9 cells were transfected with pre-miR miRNA precursors (Ambion) as described before prior to 4 days TGF β treatment or not. Total RNA was isolated from cells of two independent experiments using the RNeasy Mini Kit (Qiagen) according to the manufacturer's instruction. RNA quality control was performed with a fragment analyser using the standard or high-sensitivity RNA analysis kit (DNF-471-0500 or DNF-472-0500) from Labgene and RNA concentration was measured by using the Quanti-iTTM RiboGreen RNA assay Kit (Life Technologies/Thermo Fisher Scientific). A total of 200 ng of RNA was utilized for library preparation with the TruSeq stranded total RNA LT sample prep Kit (Illumina). Poly-A + RNA was sequenced with HiSeq SBS Kit v4 (Illumina) on an Illumina HiSeq 2500 using protocols defined by the manufacturer.

Obtained single-end RNA-seq reads (51-mers) were mapped to the mouse genome assembly, version mm10, with default RNA-STAR⁴⁸ parameters except for allowing only unique hits to genome (outFilterMultimapNmax = 1) and filtering reads without evidence in spliced junction table (outFilterType = "BySJout").

Using RefSeq mRNA coordinates from UCSC (<http://www.genome.ucsc.edu>, downloaded in December 2015) and the qCount function from QuasR package (version 3.12.1), we quantified gene expression as the number of reads that started within any annotated exon of a gene. The differentially expressed genes were identified using the edgeR package (version 1.10.1). Genes with FDR <0.05 and minimum log₂ FC of ±1 were used for downstream analysis.

Tumour transplantation. A total of 0.5×10^6 4T1 cells stably expressing ZsGreen-miR-1199/-200b/-429 or ZsGreen only were injected into the mammary fat pad of 7–9-week-old female NMRI nude or NSG mice. Mice were killed after 18 or 21 days post injection, respectively, and lungs were isolated. Tissue processing and haematoxylin and eosin staining were performed as previously described²⁴. Lung sections were imaged using an Axio Imager scanning microscope (Zeiss), and the number of lung metastasis as well as the area per lung nodule were quantified. Mann–Whitney *U* test was used for statistical analysis. All studies involving mice were approved by the Swiss Federal Veterinary Office (SFVO) and the regulations of the Cantonal Veterinary Office of Basel Stadt (licence 1907).

Isolation of circulating tumour cells. Eighty microlitres of blood was drawn by heart puncture of tumour-bearing mice killed 21 days post orthotopic cell injection, mixed with normal growth medium and cultured in a six-well plate for 5 days. Adherent cells were washed with PBS daily. A MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) staining was used to visualize tumour cell colonies. In brief, cells were incubated with 0.4 mg/ml MTT solution diluted in normal growth medium for 3 h. Afterwards, cell colonies were imaged and quantified.

Statistical analysis. Statistical analyses and graphs were generated using GraphPad Prism software (version 6). Statistical analyses were performed using an unpaired, two-sided *t* test or Mann–Whitney *U* test with **P* < 0.05, ***P* < 0.01, ****P* < 0.001, *****P* < 0.0001.

Data availability. The miRNA expression data of a TGFβ time course experiment in NMuMG/E9 cells as well as RNA expression data of NMuMG/E9 cells ectopically transfected with miR-Ctr, miR-1199-5p, miR-200b-3p or miR-429-3p are deposited at Gene Expression Omnibus (GEO, accession number: GSE86026). All other remaining data are available within the article and supplementary files, or available from the authors upon request.

Received: 10 January 2017 Accepted: 25 August 2017

Published online: 27 October 2017

References

- Nieto, M. A., Huang, R. Y., Jackson, R. A. & Thiery, J. P. EMT: 2016. *Cell* **166**, 21–45 (2016).
- Massague, J. & Obenauf, A. C. Metastatic colonization by circulating tumour cells. *Nature* **529**, 298–306 (2016).
- Lamouille, S., Xu, J. & Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat. Rev. Mol. Cell Biol.* **15**, 178–196 (2014).
- Diepenbruck, M. & Christofori, G. Epithelial-mesenchymal transition (EMT) and metastasis: yes, no, maybe? *Curr. Opin. Cell Biol.* **43**, 7–13 (2016).
- Chaffer, C. L. & Weinberg, R. A. A perspective on cancer cell metastasis. *Science* **331**, 1559–1564 (2011).
- Brabletz, T. To differentiate or not--routes towards metastasis. *Nat. Rev. Cancer* **12**, 425–436 (2012).
- Ocana, O. H. et al. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell* **22**, 709–724 (2012).
- Tsai, J. H., Donaher, J. L., Murphy, D. A., Chau, S. & Yang, J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell* **22**, 725–736 (2012).
- Fischer, K. R. et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature* **527**, 472–476 (2015).
- Zheng, X. et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* **527**, 525–530 (2015).
- Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
- Zaravinos, A. The regulatory role of microRNAs in EMT and cancer. *J. Oncol.* **2015**, 865816 (2015).
- Pencheva, N. & Tavazoie, S. F. Control of metastatic progression by microRNA regulatory networks. *Nat. Cell Biol.* **15**, 546–554 (2013).
- Burk, U. et al. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep.* **9**, 582–589 (2008).
- Gregory, P. A. et al. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10**, 593–601 (2008).
- Park, S. M., Gaur, A. B., Lengyel, E. & Peter, M. E. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev.* **22**, 894–907 (2008).
- Korpala, M., Lee, E. S., Hu, G. & Kang, Y. The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2. *J. Biol. Chem.* **283**, 14910–14914 (2008).
- Brabletz, S. & Brabletz, T. The ZEB/miR-200 feedback loop--a motor of cellular plasticity in development and cancer? *EMBO Rep.* **11**, 670–677 (2010).
- Bracken, C. P. et al. A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition. *Cancer Res.* **68**, 7846–7854 (2008).
- Siemens, H. et al. miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. *Cell Cycle* **10**, 4256–4271 (2011).
- Shi, L. et al. p53-Induced miR-15a/16-1 and AP4 form a double-negative feedback loop to regulate epithelial-mesenchymal transition and metastasis in colorectal cancer. *Cancer Res.* **74**, 532–542 (2014).
- Kundu, S. T. et al. The miR-200 family and the miR-183 similar to 96 similar to 182 cluster target Foxf2 to inhibit invasion and metastasis in lung cancers. *Oncogene* **35**, 173–186 (2016).
- Lotz-Jenne, C. et al. A high-content EMT screen identifies multiple receptor tyrosine kinase inhibitors with activity on TGFβ receptor. *Oncotarget* **7**, 25983–26002 (2016).
- Waldmeier, L., Meyer-Schaller, N., Diepenbruck, M. & Christofori, G. Py2T murine breast cancer cells, a versatile model of TGFβ-induced EMT in vitro and in vivo. *PLoS ONE* **7**, e48651 (2012).
- Kohn, K. W., Zeeberg, B. M., Reinhold, W. C. & Pommier, Y. Gene expression correlations in human cancer cell lines define molecular interaction networks for epithelial phenotype. *PLoS ONE* **9**, e9269 (2014).
- Taube, J. H. et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc. Natl Acad. Sci. USA* **107**, 15449–15454 (2010).
- Tiwari, N. et al. Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. *Cancer Cell* **23**, 768–783 (2013).
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).
- Dweep, H. & Gretz, N. miRWalk2.0: a comprehensive atlas of microRNA-target interactions. *Nat. Methods* **12**, 697 (2015).
- Peinado, H., Olmeda, D. & Cano, A. Snail, ZEB and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* **7**, 415–428 (2007).
- Wellner, U. et al. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nat. Cell Biol.* **11**, 1487–U1236 (2009).
- Krebs, A. M. et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat. Cell Biol.* **19**, 518–529 (2017).
- Karhitala, P. et al. Vimentin, zeb1 and Sip1 are up-regulated in triple-negative and basal-like breast cancers: association with an aggressive tumour phenotype. *Breast Cancer Res. Treat.* **138**, 81–90 (2013).
- Lehmann, W. et al. ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types. *Nat. Commun.* **7**, 10498 (2016).
- Li, X. L. et al. A p21-ZEB1 complex inhibits epithelial-mesenchymal transition through the microRNA 183-96-182 cluster. *Mol. Cell. Biol.* **34**, 533–550 (2014).
- Liu, Y. N. et al. MiR-1 and miR-200 inhibit EMT via Slug-dependent and tumorigenesis via Slug-independent mechanisms. *Oncogene* **32**, 296–306 (2013).
- Lehembre, F. et al. NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. *EMBO J.* **27**, 2603–2615 (2008).
- Provenzano, P. P. et al. Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med.* **4**, 38 (2006).
- Provenzano, P. P. et al. Collagen density promotes mammary tumor initiation and progression. *BMC Med.* **6**, 11 (2008).
- Maeda, M., Johnson, K. R. & Wheelock, M. J. Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. *J. Cell. Sci.* **118**, 873–887 (2005).
- Soule, H. D. et al. Isolation and characterization of a spontaneously immortalized human breast epithelial cell line, MCF-10. *Cancer Res.* **50**, 6075–6086 (1990).
- Aslakson, C. J. & Miller, F. R. 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 (1992).

44. Cortazar, D. et al. Embryonic lethal phenotype reveals a function of TDG in maintaining epigenetic stability. *Nature* **470**, 419–423 (2011).
45. Irizarry, R. A. et al. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res.* **31**, e15 (2003).
46. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy-analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315 (2004).
47. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
48. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

Acknowledgements

We thank T. Brabletz for sharing vector constructs with us. We are grateful to R. Goosen for bioinformatics support. Furthermore, we thank A. Vettiger, P. Schmidt, H. Antoniadis and I. Galm for technical assistance. We are grateful to T. Roloff, S. Dessus-Babus, C. Beisel and the Genomics Facility Basel for next-generation RNA sequencing. Calculations were performed at sciCORE (<http://scicore.unibas.ch/>) scientific computing core facility at University of Basel. This work was supported by the SystemsX.ch RTD project Cellplasticity, the SystemsX.ch MTD project MetastasiX, the Swiss National Science Foundation and the Swiss Cancer League. N.M.-S. was supported by a Marie-Heim Vögtlin grant from the Swiss National Foundation.

Author contributions

M.D. designed and performed the experiments, analysed the data and wrote the paper. S.T., F.L., N.M.-S. and M.S. designed and performed experiments and analysed the data. R.I. and R.K.R.K. performed bioinformatics analyses. G.C. oversaw the project, designed experiments, analysed data and wrote the paper.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01197-w.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017

RESEARCH ARTICLE

Open Access



Foxf2 plays a dual role during transforming growth factor beta-induced epithelial to mesenchymal transition by promoting apoptosis yet enabling cell junction dissolution and migration

Nathalie Meyer-Schaller^{1,2†}, Chantal Heck^{1,3†}, Stefanie Tiede¹, Mahmut Yilmaz^{1,4} and Gerhard Christofori^{1*} 

Abstract

Background: The most life-threatening step during malignant tumor progression is reached when cancer cells leave the primary tumor mass and seed metastasis in distant organs. To infiltrate the surrounding tissue and disseminate throughout the body, single motile tumor cells leave the tumor mass by breaking down cell-cell contacts in a process called epithelial to mesenchymal transition (EMT). An EMT is a complex molecular and cellular program enabling epithelial cells to abandon their differentiated phenotype, including cell-cell adhesion and cell polarity, and to acquire mesenchymal features and invasive properties.

Methods: We employed gene expression profiling and functional experiments to study transcriptional control of transforming growth factor (TGF) β -induced EMT in normal murine mammary gland epithelial (NMuMG) cells.

Results: We identified that expression of the transcription factor forkhead box protein F2 (Foxf2) is upregulated during the EMT process. Although it is not required to gain mesenchymal markers, Foxf2 is essential for the disruption of cell junctions and the downregulation of epithelial markers in NMuMG cells treated with TGF β . Foxf2 is critical for the downregulation of E-cadherin by promoting the expression of the transcriptional repressors of E-cadherin, Zeb1 and Zeb2, while repressing expression of the epithelial maintenance factor Id2 and miRNA 200 family members. Moreover, Foxf2 is required for TGF β -mediated apoptosis during EMT by the transcriptional activation of the proapoptotic BH3-only protein Noxa and by the negative regulation of epidermal growth factor receptor (EGFR)-mediated survival signaling through direct repression of its ligands betacellulin and amphiregulin. The dual function of Foxf2 during EMT is underscored by the finding that high Foxf2 expression correlates with good prognosis in patients with early noninvasive stages of breast cancer, but with poor prognosis in advanced breast cancer.

Conclusions: Our data identify the transcription factor Foxf2 as one of the important regulators of EMT, displaying a dual function in promoting tumor cell apoptosis as well as tumor cell migration.

Keywords: Apoptosis, Breast cancer, Cell migration, E-cadherin, EGFR, EMT, Foxf2, Noxa

* Correspondence: gerhard.christofori@unibas.ch

[†]Nathalie Meyer-Schaller and Chantal Heck contributed equally to this work.

¹Department of Biomedicine, University of Basel, Mattenstrasse 28, 4058 Basel, Switzerland

Full list of author information is available at the end of the article



Background

The process of epithelial to mesenchymal transition (EMT) describes a complex molecular and cellular program by which epithelial cells abandon their differentiated features and acquire mesenchymal characteristics, including motility, invasiveness, and increased resistance to apoptosis [1–5]. EMT has been implicated in several physiological as well as pathological processes. While it is a critical mechanism for embryonic development, EMT is re-engaged in adults during wound healing, tissue regeneration, organ fibrosis, and cancer progression and metastasis [2]. Recent studies implicate that primary tumors displaying an EMT-like gene expression profile are more likely to be associated with distant metastasis formation and a worse prognosis for overall survival [6–8]. In contradiction to these findings are the observations that distant metastases frequently exhibit an epithelial phenotype highly similar to the primary tumor [9, 10]. Explaining this observation, it has been shown that disseminated mesenchymal cancer cells undergo the reverse process (mesenchymal to epithelial transition (MET)) after metastatic spread and colonization and revert to a differentiated, epithelial cell state enabling them to establish in the distant location [11–13]. However, the contribution of EMT to the metastatic process is debated. Recent lineage tracing experiments have suggested that EMT is required for the development of drug resistance but not for metastasis [14, 15]. However, these reports have been met with great skepticism and data questioning these results [16–18]. In summary, overwhelming evidence supports the conclusion that EMT and its reverse process MET are pivotal regulators of cell plasticity in malignant tumor progression and play important roles in drug resistance, relapse, and metastatic progression [19].

Recently, a number of transcription factors have been identified that play critical roles in the initiation and execution of an EMT and in the metastatic process, including Snai1 (Snail), Snai2 (Slug), Zeb1 (δ EF1), Zeb2 (Sip1), E47, Twist, goosecoid, Foxc2, Dlx2, RBPjk, Yap/Taz, Sox4 and 9, Klf4, and NFkB [3, 19–22]. We have previously established a list of genes that change in their expression during the consecutive morphological states of transforming growth factor (TGF) β -induced EMT in normal murine mammary gland (NMuMG) epithelial cells [20, 23]. This analysis identified forkhead box protein F2 (Foxf2) as a transcription factor that is upregulated in its expression during EMT in NMuMG cells and in several other experimental EMT systems. The family of Forkhead box (Fox) genes are defined by a conserved DNA binding domain of a winged helix structure acting as transcription factors, which have been found to serve as key regulators in embryogenesis, signal transduction, maintenance of differentiated cell states, and tumorigenesis [24]. There are three families of Fox

genes, Foxc, Foxf, and Foxl1, that form paralogous clusters in the genome and that are extensively expressed in mesodermal tissue [25, 26]. One of the best characterized members of this family is Foxc2, which has been implicated in the regulation of EMT by interacting with Smad proteins and to be a key player in metastasis [27, 28]. Moreover, Foxc1 and Foxc2 are highly expressed in the claudin-low metaplastic breast cancer subtype, which is associated with EMT and cancer stemness [29]. Furthermore, the overexpression of Foxm1 in pancreatic cancer cells leads to the acquisition of an EMT phenotype via upregulation of Zeb1 and Zeb2 as well as stem cell-like characteristics [30].

Foxf2 (also known as Freac-2 or Fkhl6) is a widely expressed protein in various mesenchymal tissues and was first identified as a transcriptional activator containing a forkhead domain for nuclear localization and two independent C-terminal activation domains [31]. Foxf2 interacts with TBP and TFIIB, two components of the general transcriptional activator complex binding a specific DNA motif [32, 33]. Expression of several Fox family genes, including Foxf1 and Foxf2, is specifically regulated by sonic hedgehog (Shh) signaling in a crosstalk with Notch, epidermal growth factor (EGF)/fibroblast growth factor (FGF), and TGF β signaling [34, 35]. Indeed, TGF β -induced EMT is one of the mechanisms strongly involved in regulating fusion of the palatal cleft, and Foxf2 levels are high in the mesenchyme of the secondary palate and in the mesenchyme of the lung and gut [36, 37]. Accordingly, Foxf2-deficient mice die shortly after birth due to cleft palate and abnormal tongue and gut development, indicating an essential role of Foxf2 in this EMT-associated developmental process [38, 39]. Epithelial cells of Foxf2-deficient mice show typical signs of depolarization, and the subcellular localization of adherens junctions, normally confined to lateral membranes, expands into the basal and apical membranes.

In many cancer types, the expression of Foxf2 is repressed by promoter hypermethylation or by oncogenic microRNAs (miRNAs), such as miR-301, which promotes breast cancer cell proliferation, invasion, and tumor growth [40, 41], indicating that Foxf2 may act as a tumor suppressor. However, other studies have reported a protumorigenic role of Foxf2 in other cancer types (reviewed in [41]), for example by repressing intestinal stem cells and preventing adenoma formation by inhibiting Wnt signaling [42]. Furthermore, low Foxf2 expression has been reported to correlate with early-onset metastasis and poor prognosis in breast cancer patients [43], and loss of Foxf2 expression promotes an EMT and metastasis of experimental cancer [44, 45]. These conflicting results seem to mirror a double-sided role for Foxf2 in maintaining tissue homeostasis, in regulating an EMT, and in breast cancer progression [46].

Here, we have employed a TGF β -induced EMT in NMuMG cells and in murine and human breast cancer cells to demonstrate a critical role of Foxf2 during an EMT by concomitantly regulating an EMT, cell survival, and apoptosis. The mechanistic insights into Foxf2 functions also support a dual role of Foxf2 in breast cancer progression and metastasis, on one hand by affecting cell junction homeostasis, and by regulating cell proliferation and survival on the other hand.

Methods

Reagents and antibodies

See Additional file 1.

Cell culture and cell lines

All reagents used for cell culture were obtained from Sigma/Fluka (Basel, Switzerland) if not otherwise mentioned. All cells were cultured at 37 °C with 5% CO₂ in Dulbecco's modified Eagle's medium (DMEM) supplemented with glutamine (2 mM), penicillin (100 U), streptomycin (0.2 mg/l) and 10% fetal bovine serum (FBS). The subclone NMuMG/E9 (hereafter called NMuMG) is expressing E-cadherin and has previously been described [47]. MT Δ Ecad and MCF7-shEcad have been described previously [23]. NMuMG-shSmad4 and NMuMG-shCtr were obtained from P. ten Dijke (Leiden University Medical Center, The Netherlands) [48]. Py2T breast cancer cells were established from a tumor of the MMTV-PyMT mouse model of breast cancer as previously described [49]. NMuMG cells were treated with TGF β (2 ng/ml) without serum deprivation, and TGF β was replenished every 2 days. siRNA transfections with lipofectamine RNAiMAX (Invitrogen) were performed according to the manufacturer's protocol 24 h before treatment with TGF β .

Generation of lentivirus

A cDNA encoding Foxf2 (kindly provided by Leif Lundh, Goteborg University, Sweden) [50] was tagged N-terminally with HA-tag and cloned into the lentiviral expression vector pLenti-CMV-Puro (kindly provided by Matthias Kaeser, Bern). Lentiviral particles were produced by transfecting HEK293T cells with the lentiviral expression vectors in combination with the packaging vector pR8.91 and the envelope encoding vector pVSV using Eugene HD (Roche). After 2 days, the virus-containing HEK293T supernatant was harvested, filtered (0.45 μ m), supplemented with polybrene (8 ng/ml), and used for target cell infection. Infections were performed twice a day on 2 consecutive days.

Growth curves

One day before t₀, 1.6 \times 10⁴ NMuMG cells were seeded in triplicate into 24-well plates and transfected with the

indicated siRNA. After 24 h the cells were treated with TGF β and cell numbers were determined using a Neubauer counting chamber.

Migration assay

NMuMG cells (2 \times 10⁴/well) pretreated for 18 days with TGF β were seeded in DMEM, 2% FBS, and TGF β into the upper chamber of a cell culture insert (pore size 8 μ m; Falcon BD, Franklin Lakes, NJ). The lower chamber was filled with DMEM, 20% FBS, and TGF β . After 16 h incubation at 37 °C and 5% CO₂, the cells that had traversed the membrane were fixed in 4% paraformaldehyde/phosphate-buffered saline (PBS) (15 min at room temperature), stained with DAPI (0.5 μ g/ml), and counted using a fluorescence microscope.

Quantitative real-time polymerase chain reaction (RT-PCR)

Total RNA was prepared using Tri Reagent (Sigma-Aldrich), reverse transcribed with ImProm-II Reverse Transcriptase (Promega) and transcription levels were quantified using SYBR-green PCR Mastermix (Eurogentec) in a real-time PCR system (Step One Plus, Applied Biosystems). Human or mouse riboprotein L19 primers were used for normalization. PCR assays were performed in duplicate and the fold induction was calculated against control-treated cells using the comparative Ct method ($\Delta\Delta$ C_t). To quantify miRNA levels, RNA was isolated with the miRNeasy kit (Qiagen) followed by polyadenylation and reverse transcription using QuantiMir RT kit (BioCat). Primers are listed in Additional file 1.

Immunoblotting and immunofluorescence staining

See Additional file 1.

Apoptosis assay

Cells were washed twice in ice-cold PBS and suspended in 1 \times Annexin-V binding buffer (0.01 M HEPES, pH 7.4, 0.14 M NaCl, 2.5 mM CaCl₂) at a concentration of 1 \times 10⁶ cells/ml; 5 μ l of Cy5 Annexin-V was added to 1 \times 10⁵ cells and incubated for 15 min on ice in the dark. Stained cells were filtered through a 40- μ m mesh and analyzed on a FACSCanto II using DIVA Software (Becton Dickinson). Cell debris and duplets were excluded by a combination of light scatter and forward scatter plus width.

Cell cycle analysis

Cells were incubated with 10 μ M BrdU for 2 h at 37 °C and 5% CO₂. The cells were then fixed in 70% ice-cold ethanol and lysed by incubating first with 2 N HCl and 0.5% Triton X-100 for 30 min and then in 0.1 M Na₂B₄O₇, pH 8.5, for 2 min at room temperature. Nuclei were washed with 0.5% Tween-20, 1% bovine serum albumin (BSA)/PBS and incubated with FITC-labeled anti-BrdU antibody (#347583, Beckton Dickinson) for

30 min at room temperature. Nuclei were stained for DNA content by incubating with 5 µg/ml propidium iodide (PI) for a minimum of 1 h at room temperature. Stained cells were filtered through a 40-µm mesh and analyzed on a FACSCanto II using DIVA Software (Becton Dickinson).

Chromatin immunoprecipitation

Chromatin immunoprecipitation (ChIP) experiments were performed as previously described with some modifications [51]. In brief, cells were crosslinked either with 1% formaldehyde or in combination with 2 mM EGS (ethylene glycol bis(succinimidyl succinate); Thermo-Fisher, 21,565). Crosslinked chromatin was sonicated to receive an average fragment size of 500 bp. ChIP was performed with 100 µg of chromatin and 2.5–5 µg HA-tag antibody per IP, and 1% of ChIP material or input material was used for quantitative RT-PCR using specific primers covering Foxf2 binding sites in promoter regions of *Btc* (–450 to –253 from TSS), of *Ereg* (–851 to –654 from TSS), of *Areg* exon2 (+1086 to 1210 from TSS), and of *Noxa* (–696 to –499 from TSS). Primers covering an intergenic region were used as control, and the amplification efficiencies were normalized between the primer pairs. Enrichment of IP/input over IgG background control was calculated and the specificity measured as fold change to an unspecific intergenic region.

Transcriptome, survival, and metastasis correlation analysis

See Additional file 1.

Statistical analysis

Statistical analysis and graphs were generated using the GraphPad Prism software (GraphPad Software Inc., San Diego CA). All statistical analyses were performed as indicated by paired or unpaired two-sided *t* test.

Results

Foxf2 expression is induced during EMT

We screened for changes in gene expression by DNA oligonucleotide microarray analysis during an EMT in three independent in vitro model systems. First, MTfE-cad cells have been derived from a breast tumor of MMTV-Neu transgenic mice [52] in which both E-cadherin alleles were flanked by LoxP recombination sites [53]. Genetic ablation of E-cadherin was achieved by the transient expression of Cre-recombinase (MTΔE-cad) [23]. Second, EMT was induced in the human breast cancer cell line MCF7 by downregulation of E-cadherin using stable expression of shRNA [23] and, thirdly, EMT was induced in normal murine mammary epithelial (NMuMG) cells by treatment with TGFβ [54] (Additional file 1: Figure S1A). The forkhead transcription factor Foxf2 was identified as a commonly

upregulated gene during EMT in all three experimental systems (Additional file 1: Figure S1B, C). To assess whether Foxf2 is a target of canonical or noncanonical TGFβ signaling, we monitored Foxf2 expression in NMuMG cells stably depleted of Smad4 expression (NMuMG-shSmad4) [48]. Foxf2 mRNA expression levels were significantly reduced in TGFβ-treated NMuMG-shSmad4 cells compared with control cells, indicating that Foxf2 is regulated via canonical Smad4-dependent TGFβ signaling (Additional file 1: Figure S1D).

Foxf2 is partially required for EMT

We first assessed whether the expression of Foxf2 is able to induce an EMT by infecting NMuMG cells with lentiviral particles encoding HA-tagged human Foxf2. Although the cells expressed Foxf2 in their nuclei, the cells did not gain an EMT-like phenotype (data not shown). Conversely, to investigate whether the upregulation of Foxf2 expression is required for an EMT, we stably infected NMuMG cells with lentiviral particles expressing two different shRNAs against murine Foxf2 (shFoxf2 #703, shFoxf2 #704). NMuMG-shFoxf2 cells treated with TGFβ apparently changed to a mesenchymal cell morphology, comparable to TGFβ-treated NMuMG-shCtrl cells. However, NMuMG-shFoxf2 cells did not completely lose their tight cell-cell contacts, a key step during an EMT (Fig. 1a). Indeed, quantitative RT-PCR (Fig. 1b, c) and immunoblotting (Fig. 1d) analysis revealed that the shRNA-mediated ablation of Foxf2 expression resulted in a sustained expression of the epithelial adherens and tight junction molecules E-cadherin and ZO-1, whereas the increased expression of the mesenchymal markers fibronectin, Ncam1, and N-cadherin remained unaffected.

To investigate whether the shRNA-mediated depletion of Foxf2 expression affects EMT-associated changes in cell adhesion, cell junctions, and/or cytoskeletal composition, we performed immunofluorescence microscopy analysis for the cell adhesion proteins E-cadherin, N-cadherin, and Ncam1, the tight junction protein ZO-1, the focal adhesion protein paxillin, and actin stress fibers (phalloidin). NMuMG-shFoxf2 cells did not show a classical cadherin switch when treated with TGFβ. In Foxf2-ablated cells, a normal upregulation of the mesenchymal marker N-cadherin was observed, but the expression of the epithelial markers E-cadherin and ZO-1 was partially maintained at the cell membrane, in contrast to shCtrl-expressing cells which showed a bona-fide EMT (Additional file 1: Figure S2A, B). Upregulated expression of Foxf2 during an EMT was also not required for the EMT-associated cytoskeletal reorganization of cortical actin into actin stress fibers, for the

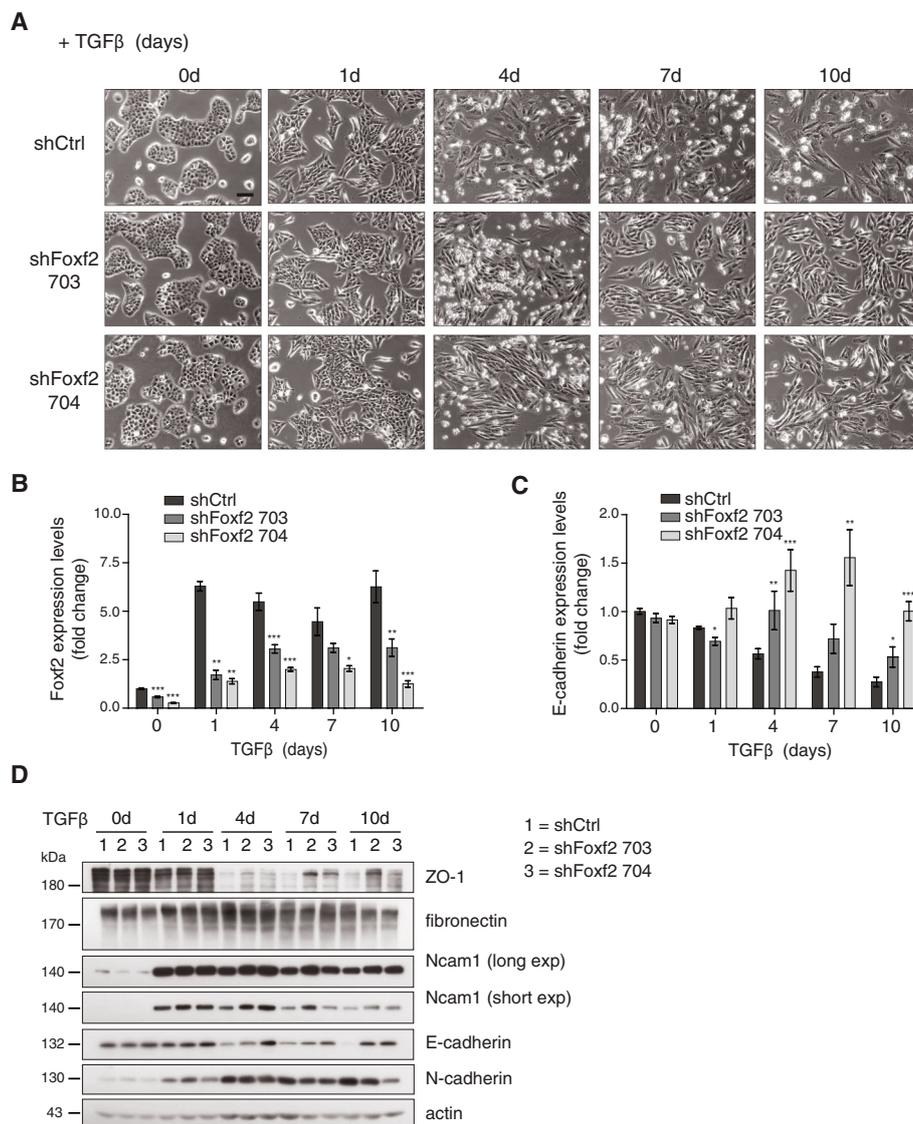


Fig. 1 Downregulation of Foxf2 attenuates TGFβ-induced EMT. **a** Phase-contrast micrographs of NMuMG cells stably expressing a control shRNA (shCtrl) or shRNAs against Foxf2 (shFoxf2 703, shFoxf2 704) treated with transforming growth factor (TGFβ) for the times indicated. Scale bar = 100μm. **b** Foxf2 knockdown efficiency was determined by quantitative RT-PCR in NMuMG cells stably infected with shCtrl, shFoxf2 703, or shFoxf2 704 and treated with TGFβ for the times indicated. Values were normalized to RPL19 and presented as fold changes compared with untreated shCtrl NMuMG cells. **c** Loss of E-cadherin expression during TGFβ-induced EMT depends on Foxf2. E-cadherin mRNA levels in shFoxf2 and shCtrl-transfected NMuMG cells were determined by quantitative RT-PCR. Values were normalized to RPL19 and reported as fold changes compared with untreated shCtrl NMuMG cells. **d** Knocking down Foxf2 leads to a sustained expression of cell junction components. Immunoblotting analysis for the epithelial markers E-cadherin and ZO-1 as well as the mesenchymal markers Ncam1, N-cadherin, and fibronectin in shFoxf2 knockdown and shCtrl NMuMG cells treated with TGFβ for the times indicated. Actin was used as a loading control. Data are shown as mean ± SEM of three independent experiments. Statistical values were calculated using a paired two-tailed *t* test. **p* ≤ 0.05; ***p* ≤ 0.01; ****p* ≤ 0.001

upregulation of the mesenchymal marker Ncam1, or for the formation of focal adhesions shown by paxillin staining (Additional file 1: Figure S2B, C). Together, these results indicate that Foxf2 is required for the disruption of cell-cell junctions but not for the induction of a mesenchymal cellular phenotype and mesenchymal marker expression.

Foxf2 regulates EMT transcriptional regulators and cell migration

The maintenance of an epithelial morphology and E-cadherin expression in Foxf2-depleted NMuMG cells became even more apparent when treated for 19 to 20 days with TGFβ, a time frame necessary for control NMuMG cells to acquire an EMT stage associated with

cellular migration (Fig. 2a–d). Since *Foxf2* appeared essential for the loss of E-cadherin expression and the disruption of cell-cell junctions, we assessed whether the loss of *Foxf2* expression affected the migratory

capabilities of cells. Transwell migration assays revealed a decrease in motility for cells stably expressing shRNAs against murine *Foxf2* compared with NMuMG cells expressing control shRNA (Fig. 2e).

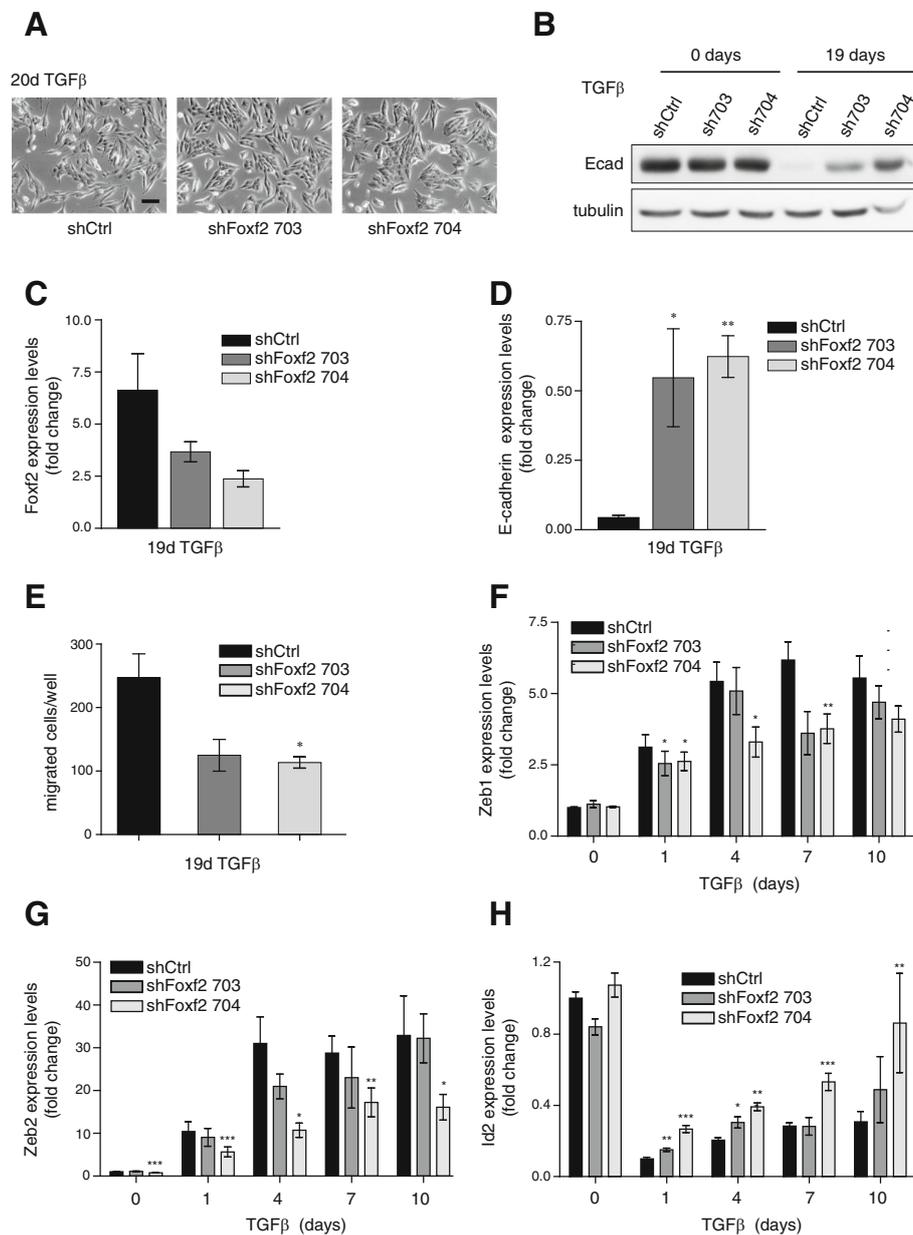


Fig. 2 *Foxf2* regulates cell migration and the expression of E-cadherin transcriptional repressors. **a** Phase-contrast micrographs of NMuMG cells stably expressing a control shRNA (shCtrl) or a *Foxf2*-specific shRNA (shFoxf2 703, 704) treated with transforming growth factor (TGF)β for 20 days. Scale bar = 100μm. **b** shRNA-mediated ablation of *Foxf2* leads to sustained E-cadherin expression after 19 days of TGFβ treatment shown by immunoblotting analysis in shFoxf2 and shCtrl NMuMG cells. Tubulin was used as a loading control. **c** *Foxf2* and **d** E-cadherin mRNA levels were determined by quantitative RT-PCR in shFoxf2 and shCtrl NMuMG cells treated with TGFβ for 19 days. Values were normalized to RPL19 and reported as fold changes compared with untreated shCtrl NMuMG cells. **e** Depletion of *Foxf2* leads to reduced cell migration of NMuMG cells treated for 19 days with TGFβ through transwell filters compared with control cells (shCtrl). mRNA levels of **f** Zeb1, **g** Zeb2, and **h** Id2 were determined by quantitative RT-PCR in shFoxf2- and shCtrl-transfected NMuMG cells. Values were normalized to RPL19 and reported as fold changes compared with untreated shCtrl NMuMG cells. Data are shown as mean ± SEM of three independent experiments. Statistical values were calculated using a paired two-tailed *t* test between shCtrl and shFoxf2 cells. **p* ≤ 0.05; ***p* ≤ 0.01; ****p* ≤ 0.001

The loss of E-cadherin is often attributed to transcriptional dysregulation. Several transcription factors have been identified that are able to repress E-cadherin gene expression, among them the zinc-finger-homeodomain transcription factors Zeb1 and Zeb2 [55–57]. Quantitative RT-PCR analysis revealed that shRNA-mediated ablation of Foxf2 expression during a TGF β -induced EMT in NMuMG cells attenuated the upregulation of Zeb1 and Zeb2 expression observed in shCtrl-transfected cells (Fig. 2f, g). Inhibitors of differentiation (Ids) act as positive regulators of proliferation and as negative regulators of differentiation. The Id proteins lack a DNA-binding motif and inhibit, for example, E2A-dependent suppression of the E-cadherin promoter [58]. Consistent with the sustained expression of E-cadherin, ablation of Foxf2 in TGF β -treated NMuMG cells interfered with the downregulation of Id2 expression during an EMT (Fig. 2h).

The loss of Foxf2 maintained E-cadherin expression during a TGF β -induced EMT in NMuMG cells in a similar manner as the loss of the major transcriptional repressor of E-cadherin expression, Zeb1 (Additional file 1: Figure S3A, B). Also comparable to the siRNA-mediated ablation of Zeb1 expression, loss of Foxf2 expression resulted in the upregulated expression of the miR-200 family members miR-200a-3p, miR-200b-3p, and miR-429-3p (Additional file 1: Figure S3C). On the other hand, both Foxf2 and Zeb1 are predicted targets of miR-200 family members, and ectopic expression of the miR-200 family members miR-200b-3p, miR-200c-3p, and miR-429-3p downregulated both transcription factors (Additional file 1: Figure S3D, E). Consistent with a regulatory role of Foxf2 on the expression of Zeb family proteins, Foxf2 upregulation after 1 day of TGF β treatment of NMuMG cells was followed by the induction of Zeb1 and Zeb2 expression, leading to a continuous downregulation of E-cadherin (Additional file 1: Figure S3F). These results indicate tight control of E-cadherin expression by a double-negative feedback loop between Foxf2 and miR-200 family members, as well as regulation of the expression of known targets of miR200 family members and transcriptional repressors of E-cadherin expression, such as Zeb1, Zeb2, and Id2. Similarly, Foxf2 is essential for the proper downregulation of E-cadherin and the regulation of Zeb2 and Id2 during a TGF β -induced EMT of Py2T murine breast cancer cells that have been derived from a tumor of the MMTV-PyMT mouse model of breast cancer (Additional file 1: Figure S3G).

In conclusion, the upregulation of Foxf2 in NMuMG cells undergoing an EMT is essential for the transcriptional repression of E-cadherin, for the disruption of cell-cell adhesions, and for EMT-associated cell migration, yet has only minor effects on the induction of mesenchymal marker expression.

Foxf2 regulates cell death and survival pathways

To identify the actual genes and signaling pathways that are regulated by Foxf2 during an EMT, we performed gene expression profiling by RNA sequencing of NMuMG cells that were transfected with control siRNA (siCtrl; epithelial state) or with siCtrl or siRNA targeting Foxf2 (siFoxf2) in the presence of TGF β for 4 days (mesenchymal state). We found 1789 genes to be differentially expressed at least twofold upon Foxf2 knockdown compared with siRNA control at 4 days of TGF β treatment, and 2689 genes were significantly changed upon induction of EMT comparing siCtrl in the absence or presence of TGF β (siCtrl 0 days vs 4 days TGF β ; EMT). In total, 792 genes were commonly regulated by the loss of Foxf2 expression and by the induction of an EMT with TGF β (Fig. 3a; Additional file 1: Table S1). Unsupervised hierarchical clustering revealed that Foxf2-deficient cells treated with TGF β more closely resembled the mesenchymal state of TGF β -treated control cells, however they formed a separate clustering arm (Fig. 3b). To study in more detail which transcripts were specifically altered compared with the mesenchymal and epithelial control states, gene expression signatures were generated using weighted gene coexpression network analysis (WGCNA). Six different gene expression signatures were extracted, with the yellow and the brown signatures summarizing genes from an intermediate mesenchymal state (Fig. 3b; Additional file 1: Table S2). The EMT-induced expression of the genes in the yellow signature was strongly reduced by the ablation of Foxf2 knockdown. Conversely, the EMT-repressed expression of the genes in the brown signature was blocked by the loss of Foxf2 (Fig. 3c).

Pathway enrichment analysis using ingenuity pathway analysis (IPA) revealed major functions of the yellow signature-associated genes in cellular movement and cell-cell signaling and interaction, functions that can be attributed to the loss of E-cadherin expression as described above (Fig. 3d). On the other hand, the brown signature was found to be associated with molecular transport and metabolism (lipid, vitamin, and mineral metabolism) pathways (Fig. 3d). Interestingly, both up- and downregulated EMT signatures that are also affected by Foxf2 knockdown (brown and yellow signatures) were enriched in pathways describing cell death and survival, indicating a regulatory role of Foxf2 in these processes (Fig. 3d). Indeed, NMuMG cells stably expressing shRNA against Foxf2 showed significantly less TGF β -mediated growth inhibition compared with shCtrl-transfected cells, and the cell number increased significantly compared with shCtrl-expressing cells (Fig. 4a).

Foxf2 affects apoptosis by regulating the expression of Noxa

To assess whether the increase in cell numbers was due to increased proliferation or decreased cell death, we

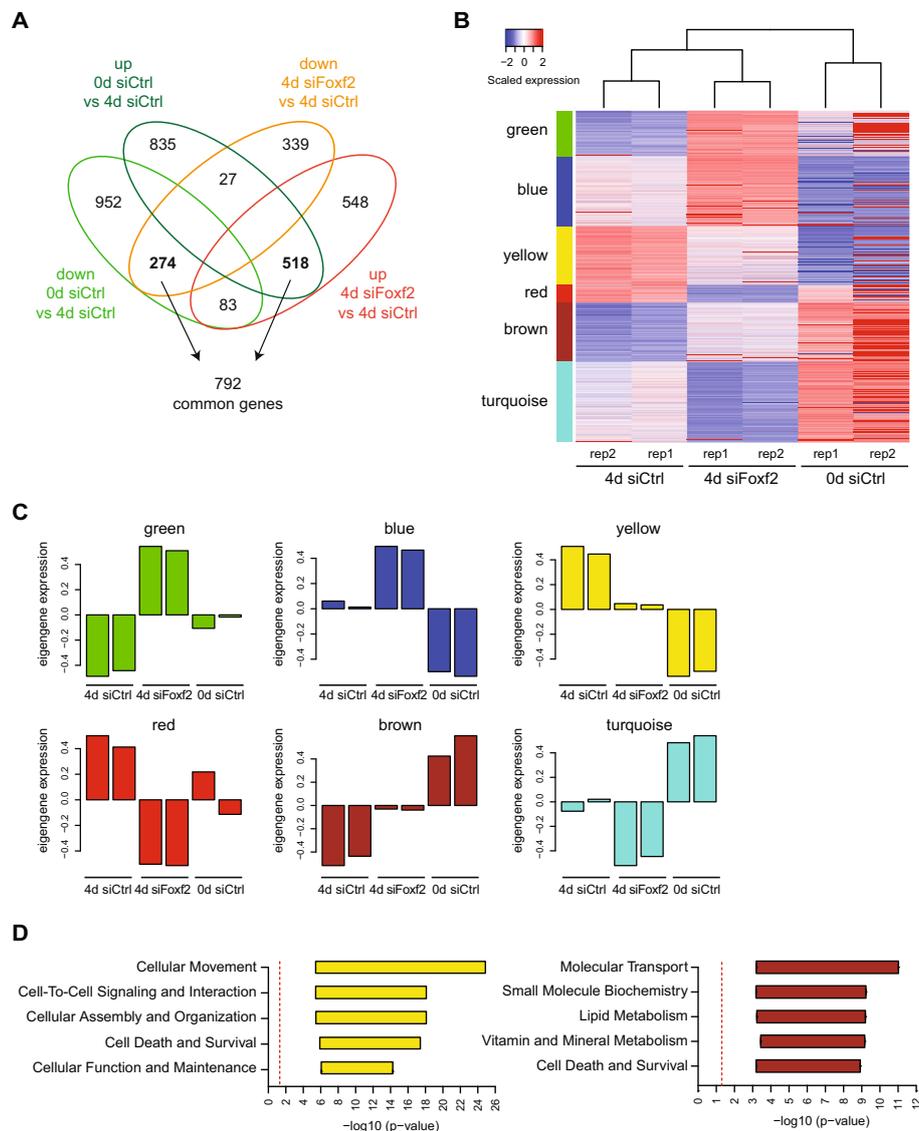


Fig. 3 RNA sequencing analysis of Foxf2-dependent gene expression. **a** Venn diagram illustrating the overlap of significantly differentially expressed genes (p adjusted < 0.05 , $\text{abs}(\log_2 \text{fold change}) \geq 1$) between the epithelial and mesenchymal state (0 days vs 4 days siRNA control (siCtrl)) or for the Foxf2 perturbation (4 days siFoxf2 vs 4 days siCtrl). Highlighted in bold are genes whose change during EMT is reversed by Foxf2 knockdown. **b** Heatmap of gene expression from RNA-sequencing of epithelial control samples (untreated siCtrl NMuMG cells) and of siCtrl- or siFoxf2-treated NMuMG cells in the presence of TGF β for 4 days. The sample order in the heatmap was obtained from an unsupervised hierarchical clustering, while rows (genes) were arranged according to the gene signatures derived from WGCNA. **c** The eigengene expression of the different gene signatures derived from WGCNA (see **a**) illustrates a general expression trend of all genes belonging to a gene signature. **d** IPA analysis of the brown and yellow gene signature. Shown in the bar plot are the significance ranges of pathways belonging to the most significantly enriched categories. The dotted red line indicates a p value of 0.05. Differential gene expression and gene signature memberships are reported in Additional file 1 (Table S1 and S2, respectively)

compared the rates of apoptosis and proliferation by Annexin-V staining and BrdU incorporation and PI staining, respectively, in NMuMG cells treated with TGF β and depleted or not for Foxf2 expression. The loss of Foxf2 significantly reduced the levels of apoptosis when compared with control cells (Fig. 4b), while only a moderate difference in the number of cycling cells was observed (Additional file 1: Figure S4A). The lack of Foxf2 reduced caspase-dependent programmed cell

death, as the levels of cleaved caspase-3 and its downstream cleavage target poly-(ADP-ribose) polymerase (PARP) were diminished upon knockdown of Foxf2 (Fig. 4c). In summary, the results show that the increased expression of Foxf2 during EMT is critical for promoting TGF β -induced cell death.

TGF β has been shown to act as a tumor suppressor in the early stages of tumorigenesis by inducing the expression of cell cycle inhibitors and proapoptotic factors

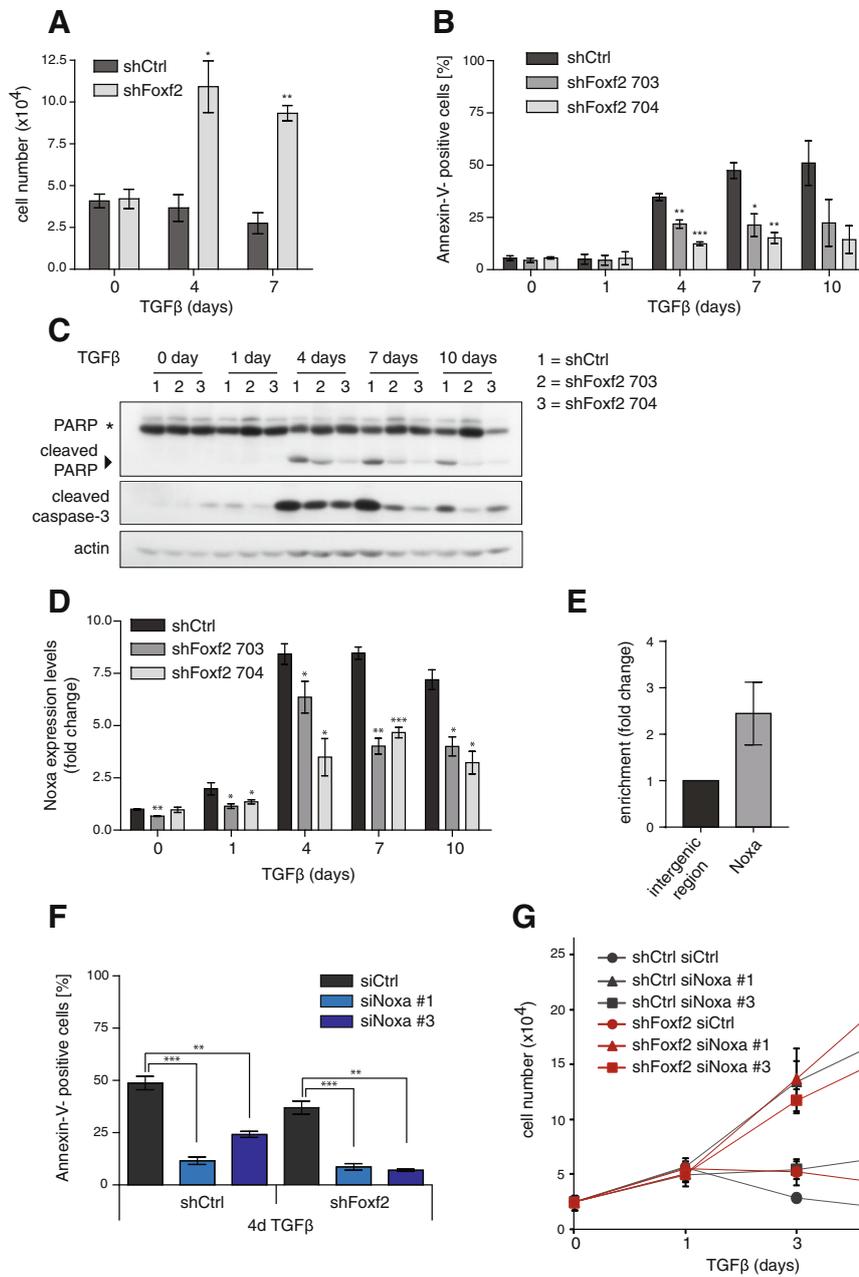


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 Depletion of Foxf2 attenuates TGFβ-induced apoptosis and the expression of the proapoptotic protein Noxa. **a** Downregulation of Foxf2 promotes cell proliferation. shFoxf2- and shCtrl-expressing NMuMG cells were treated with transforming growth factor (TGF)β for the times indicated and counted using a Neubauer chamber. **b** Foxf2 depletion decreases apoptosis during TGFβ-induced EMT. shFoxf2- and shCtrl-expressing NMuMG cells were treated with TGFβ for the times indicated, and apoptosis was detected by Annexin-V staining and flow cytometry analysis. **c** TGFβ induces classical caspase-mediated apoptosis dependent on the upregulation of Foxf2. Immunoblotting analyses of the same experiment as shown in Fig. 1d for cleaved caspase-3 and PARP in shFoxf2- and shCtrl-expressing NMuMG cells treated with TGFβ for the times indicated. Actin was used as a loading control. **d** Knockdown of Foxf2 attenuates the upregulation of Noxa expression. Noxa mRNA levels in shFoxf2- and shCtrl-expressing NMuMG cells treated with TGFβ for the times indicated were determined by quantitative RT-PCR. Values were normalized to RPL19 and presented as fold changes compared with untreated shCtrl NMuMG cells. **e** Foxf2 regulates Noxa expression by direct transcriptional activation. Chromatin immunoprecipitation of HA-tagged Foxf2 was performed either on Foxf2-expressing or control NMuMG cells treated for 2 days with TGFβ. Immunoprecipitated DNA fragments were quantified by quantitative PCR using primers covering base pairs -696 to -499 of the *noxa* promoter region. Enrichment (IP/input) for specific primers was calculated relative to primers covering an intergenic region. **f** Noxa depletion significantly decreases TGFβ-induced apoptosis. shFoxf2- and shCtrl-expressing NMuMG cells were transfected with control siRNA (siCtrl) and two different siRNAs specific for murine Noxa (siNoxa #1, siNoxa #3) and incubated with TGFβ for 4 days. The extent of apoptosis was measured by Annexin-V staining and flow cytometry. **g** The impairment of Noxa expression leads to increased cell proliferation. shFoxf2- and shCtrl-expressing NMuMG cells were transfected with control siRNA (siCtrl) and two different siRNAs specific for murine Noxa (siNoxa #1, siNoxa #3) and incubated with TGFβ for the times indicated. Cell numbers were determined using a Neubauer chamber. Results show the mean ± SEM of three independent experiments. Statistical values were calculated using paired/unpaired two-tailed *t* test. **p* ≤ 0.05; ***p* ≤ 0.01; ****p* ≤ 0.001

[59]. Differential gene expression analysis between TGFβ-treated control and Foxf2-deficient NMuMG cells revealed a substantial regulation of the BH3-only factor Noxa by Foxf2, which was confirmed by quantitative RT-PCR analysis (Fig. 4d). ChIP experiments with NMuMG cells expressing HA-tagged Foxf2 treated for 2 days with TGFβ demonstrated a weak direct binding of Foxf2 to the *noxa* gene promoter, suggesting additional indirect regulatory mechanisms (Fig. 4e). As the upregulation of Foxf2 is necessary for TGFβ-induced Noxa expression, we next assessed whether loss of Noxa is sufficient to prevent apoptosis in TGFβ-treated NMuMG cells. Noxa expression in NMuMG cells was ablated by transient transfection of two different siRNAs (siNoxa #1, siNoxa #3) or control siRNA (siCtrl) in shFoxf2- and shCtrl-expressing cells. Following the reduction of Noxa mRNA levels in siNoxa #1- and #3-treated cells (Additional file 1: Figure S4B), apoptosis was attenuated and cell growth inhibition was compensated in TGFβ-treated cells (Fig. 4f, g; Additional file 1: Figure S4C). These results demonstrate that depletion of Noxa is sufficient to protect NMuMG cells from TGFβ-induced apoptosis. Together, these data indicate that Foxf2 mediates TGFβ-induced apoptosis by the transcriptional activation of the proapoptotic protein Noxa.

Foxf2 promotes TGFβ-induced growth arrest by repressing EGF receptor signaling

As well as activating transcription of the *noxa* gene and inducing apoptosis, we investigated whether Foxf2 regulates any prosurvival signaling pathway. EGF receptor (EGFR) family members are known to provide protection from TGFβ-induced cell cycle arrest and apoptosis by activating the PI3K pathway [60, 61]. Indeed, immunoblotting analyses revealed that the levels of activated (tyrosine 1173-phosphorylated) forms of EGFR were higher in

Foxf2-depleted NMuMG cells compared with control NMuMG cells when treated with TGFβ (Fig. 5a).

Previously, we have reported that survival of NMuMG cells undergoing a TGFβ-induced EMT depends on activated EGFR signaling [20]. We thus investigated whether inhibition of EGFR signaling impaired the antiapoptotic effect of Foxf2 depletion. Towards this aim, shFoxf2- and shCtrl-transfected NMuMG cells were treated with the EGFR inhibitor (EGFRi) AG1478 during TGFβ treatment, and cell growth and rates of apoptosis were determined. Combined treatment with EGFRi and TGFβ for 4 days led to a significantly reduced growth in shFoxf2-expressing NMuMG cells compared with the solvent (dimethyl sulfoxide (DMSO))-treated shFoxf2-expressing NMuMG cells (Fig. 5b; Additional file 1: Figure S5A). In addition, treatment of shFoxf2-expressing NMuMG cells with AG1478 increased apoptosis to a similar extent as observed in shCtrl-expressing cells (Fig. 5c). The extent of apoptosis thereby correlated with the levels of EGFR inhibition (Fig. 5d). This result indicates that TGFβ-resistant growth of Foxf2 knockdown cells relies on the activation of EGFR survival signaling.

To investigate how Foxf2 influences EGFR activation, we assessed whether the expression of EGFR ligands was affected by the modulation of Foxf2 expression. Gene expression profiling and validation by quantitative RT-PCR revealed that the EGFR-ligands betacellulin (*Btc*), amphiregulin (*Areg*), and (moderately) epiregulin (*Ereg*) showed sustained expression upon knockdown of Foxf2 during an EMT in both NMuMG and Py2T cells (Additional file 1: Figure S5B, C). Promoter binding prediction programs indicated a potential direct binding of Foxf2 to the *Btc* promoter (data not shown). ChIP followed by quantitative PCR of HA-tagged Foxf2 in NMuMG cells during TGFβ-induced EMT revealed a direct binding of Foxf2 to the *Btc* promoter region, to a

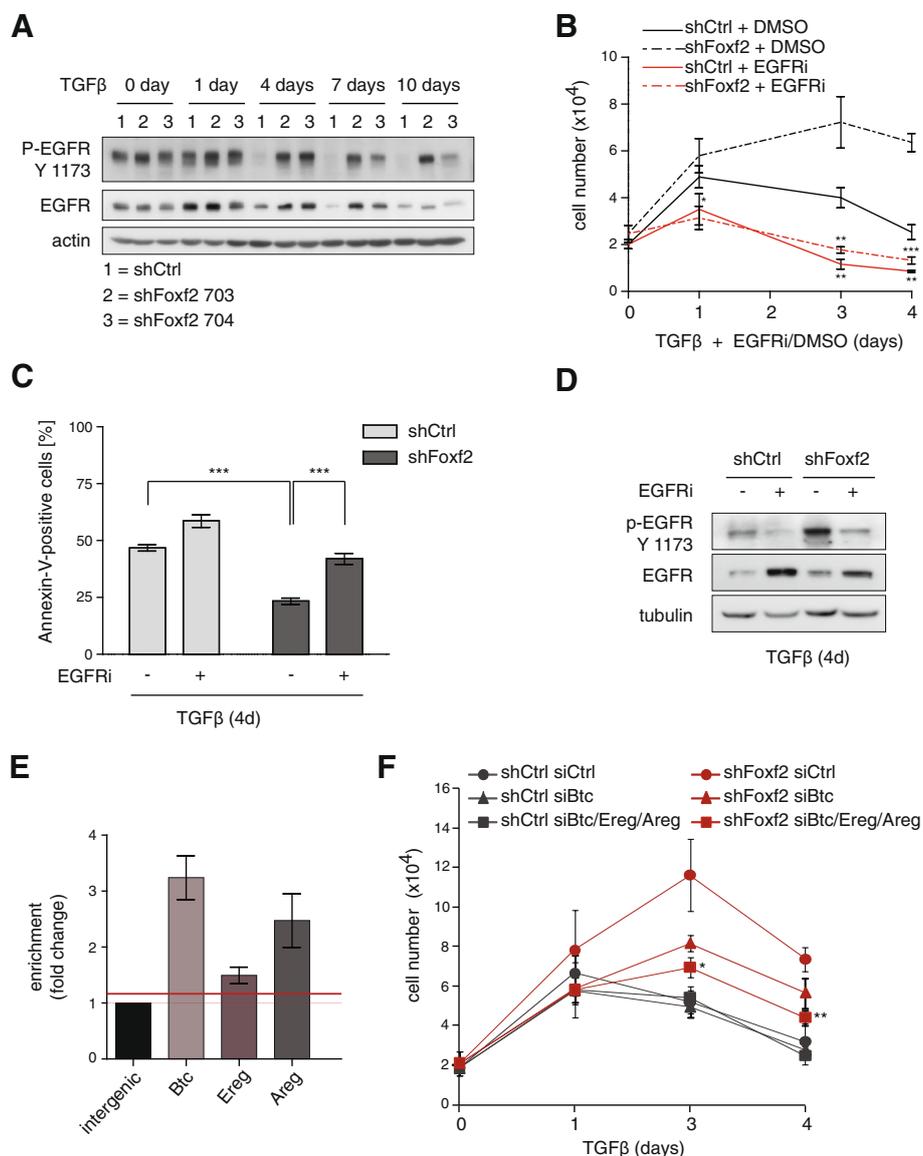


Fig. 5 Inhibition of EGFR signaling increases apoptosis in Foxf2-depleted cells. **a** Depletion of Foxf2 leads to sustained epidermal growth factor receptor (EGFR) activation. Immunoblotting analysis of the phosphorylation status of EGFR and total EGFR protein levels in shFoxf2- and shCtrl-expressing NMuMG cells treated with transforming growth factor (TGF)β for the times indicated. Actin was used as a loading control. **b–d** NMuMG cells expressing shRNA specific for Foxf2 or control shRNA were treated with TGFβ and AG1478 (EGFR inhibitor (EGFRi)) or control solvent (dimethyl sulfoxide (DMSO)) for the indicated times. **b** Cell numbers were determined using a Neubauer chamber. **c** EGFR inhibition significantly increases apoptosis in Foxf2 knockdown cells. Apoptosis was detected by Annexin-V staining and flow cytometry. **d** Treatment with AG1478 decreases EGFR activation in shFoxf2 cells to a similar extent as seen in TGFβ-treated NMuMG cells expressing control shRNA. Immunoblotting analysis for EGFR phosphorylation and total EGFR levels is shown. Tubulin was used as a loading control. **e** Foxf2 regulates the expression of EGFR ligands by direct transcriptional repression. Chromatin immunoprecipitation of Foxf2 was performed either on HA-Foxf2 expressing or control NMuMG cells treated for 2 days with TGFβ. Immunoprecipitated DNA fragments were quantified by quantitative PCR using primers covering base pairs –450 to –253 of the *Btc* promoter region, base pairs –851 to –654 of the *Ereg* promoter, and base pairs +1086 to 1210 of the *Areg* exon 2. Enrichment (IP/input) for specific primers was calculated relative to primers covering an intergenic region. **f** Individual depletion of Btc or combined depletion of betacellulin (Btc), epiregulin (Ereg), and amphiregulin (Areg) reduces cell numbers in shFoxf2- but not in shCtrl-expressing NMuMG cells in the presence of TGFβ for 4 days. Cell numbers were determined using a Neubauer chamber. Data are shown as mean ± SEM of at least three independent experiments. Statistical values were calculated using a paired/unpaired two-tailed *t* test. **p* ≤ 0.05; ***p* ≤ 0.01; ****p* ≤ 0.001

regulatory region in exon 2 of the *Areg* gene, and with less efficiency to the *Ereg* promoter region (Fig. 5e).

To assess whether *Btc*, *Areg*, or *Ereg* were responsible for the stimulation of EGFR and increased cell survival of *Foxf2*-depleted cells, NMuMG cells stably expressing shRNA against *Foxf2* or a control shRNA were transiently transfected with siRNAs against *Btc* or with a mix of siRNAs against *Btc*, *Areg*, and *Ereg* and treated with TGF β . The efficiency of *Btc* or combined *Btc/Areg/Ereg* ablation was determined by quantitative RT-PCR (Additional file 1: Figure S5D, E). Knockdown of *Btc* alone or in combination with the other two family members resulted in reduced cell growth in sh*Foxf2*-expressing cells when treated with TGF β (Fig. 5f). These results indicate that attenuation of EGFR activation by siRNA-mediated depletion of its ligands abrogates the survival benefit of *Foxf2* depletion. We conclude that TGF β -induced *Foxf2* expression represses the transcriptional activation of the *Btc* and *Areg* genes, resulting into a reduced expression of these EGFR ligands and a repression of EGFR survival signaling.

Foxf2 expression correlates with poor prognosis in patients

Cancer-associated gene expression profiling has emerged as an appropriate tool to predict the relapse risk and to identify genes that mediate disease progression. To investigate whether *Foxf2* expression is predictive for tumor progression or metastasis formation, we analyzed a breast cancer database of the Memorial Sloan-Kettering Cancer Center (MSKCC), published by Minn et al. [62]. This “Minn” database consists of microarray gene expression analysis of tumor samples from 82 patients with advanced breast cancer (T2–T4). The tumors were divided into two groups according to the log expression levels relative to the median expression of the investigated gene. The low and high *Foxf2*-expressing groups were further stratified for lymph node (LN) metastasis status. Interestingly, low *Foxf2* expression significantly correlated with early distant metastasis formation in lymph node-negative (LN⁻) tumors, whereas the opposite tendency was found in tumors of patients that were positive for lymph node metastasis (LN⁺) (Fig. 6a, b).

To further substantiate a potential correlation between *Foxf2* expression and patient survival, we analyzed the Netherlands Cancer Institute (NKI295) breast cancer database for *Foxf2* expression [63]. The NKI295 database consists of microarray gene expression analysis of tumor samples from 295 patients with early-stage breast cancer (stage I or stage II primary breast carcinomas). Although *Foxf2* expression was not predictive for metastasis formation or survival in the total patient pool (Fig. 6c), high expression of *Foxf2* correlated with poor overall survival in patients with luminal subtype B breast cancer (Fig. 6d). High expression of *Foxf2* in tumors with negative estrogen receptor (ER) status correlated

with high significance of early metastasis onset as well as poor overall survival (Fig. 6e, f). Similarly, in a large tumor collection from the Metabric consortium [64, 65], high *Foxf2* expression predicted worse survival in the luminal B breast cancer subtype (Fig. 6g). Interestingly, *Foxf2* expression was significantly higher in more aggressive tumor subtypes, such as ER⁻ compared with ER⁺, triple-negative compared with all other subtypes, and in claudin-low tumors (the breast cancer subtype associated with an EMT signature), compared with all others (Fig. 6h).

Together, the expression of *Foxf2* in human patient samples and its prediction for clinical outcome reflect the dual function of *Foxf2* observed in our in-vitro studies. *Foxf2* may function as a tumor suppressor in early cancer development by promoting apoptosis, hence showing a poor prognosis in LN⁻ patients with low *Foxf2* expression. More advanced tumors with high *Foxf2* expression correlate with shorter metastasis-free survival, supporting a role of *Foxf2* in cancer cell invasion and metastasis formation.

Discussion

Overcoming the growth inhibitory effect of TGF β during the early stages of tumorigenesis as well as the conversion of TGF β -mediated growth inhibition into TGF β -induced tumor progression are fundamental processes during primary tumor growth and metastasis formation [1, 66]. Thus, understanding the mechanisms underlying this dual role of TGF β in cancer progression and the strategies of cancer cells to circumvent TGF β -induced apoptosis may offer new opportunities for novel cancer therapies.

Here, we have employed nontransformed NMuMG cells and Py2T murine breast cancer cells to delineate the molecular mechanisms underlying a TGF β -induced EMT, including overcoming TGF β -induced resistance to apoptosis and the acquisition of invasive properties. We report a dual function of the transcription factor *Foxf2*, acting as a tumor suppressor by promoting apoptosis and by repressing survival, while exerting protumorigenic activity at later stages of tumor progression, such as a promigratory function by inducing the disruption of cell-cell adhesion. *Foxf2* is upregulated via the canonical TGF β pathway, and gain of function studies in NMuMG cells reveal that its expression is not sufficient to induce an EMT. However, loss of function studies demonstrate that *Foxf2* is essential for the disruption of cell junctions and the repression of epithelial marker expression but not for the gain of mesenchymal marker expression. Notably, *Foxf2* is crucial for TGF β -mediated cell death by the transcriptional activation of the *Noxa* gene, encoding for a BH3-only proapoptotic factor, and the subsequent induction of caspase-dependent apoptosis. Moreover, *Foxf2* directly represses transcription of the *Btc*

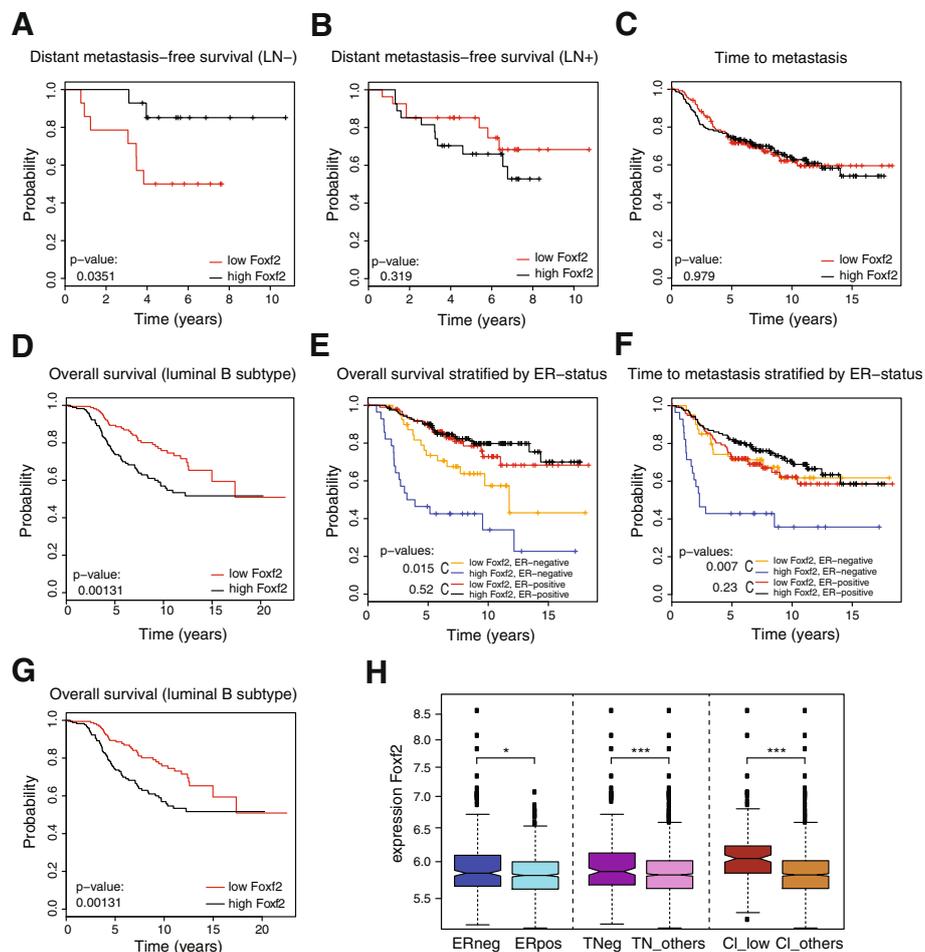
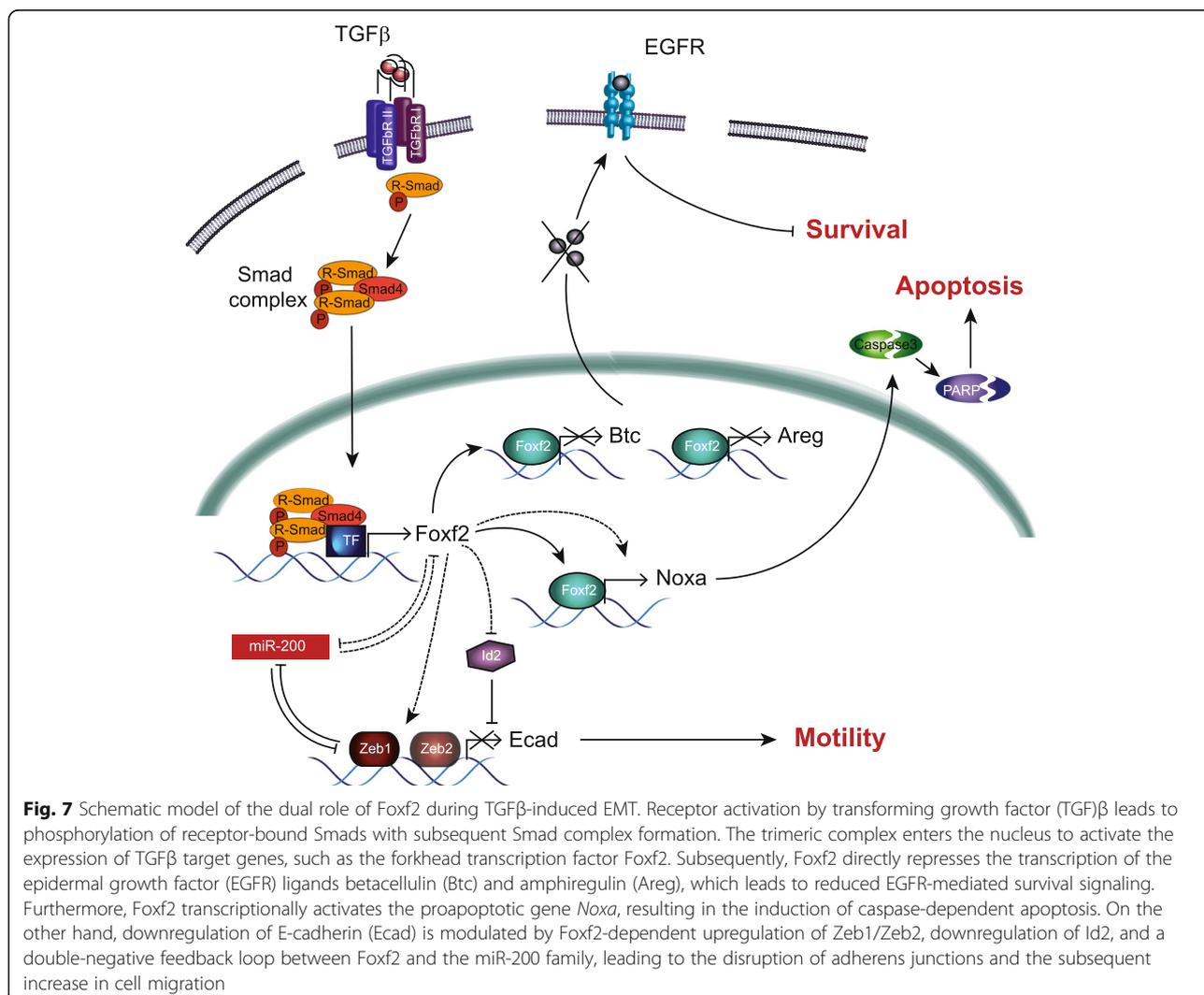


Fig. 6 High Foxf2 expression correlates with good prognosis in early-onset breast cancer patients but with poor prognosis in late-stage estrogen receptor-negative and luminal B breast cancer patients. **a,b** Statistical analysis of the Memorial Sloan-Kettering Cancer Center (Minn) database. Tumors were divided into high and low Foxf2-expressing groups based on the median expression of Foxf2 mRNA. Expression of Foxf2 was correlated with distant metastasis free survival in lymph node-negative (LN⁻) tumors ($n = 28$) or LN⁺ tumors ($n = 54$). Foxf2 expression is predictive for metastasis incidence in LN⁻ tumors (**a**), where low Foxf2 expression correlates with early metastases onset (p value = 0.0351), but not for LN⁺ tumors (**b**). **c-f** Statistical analysis of the Netherlands Cancer Institute (NKI295) database. Tumors were divided into high and low Foxf2-expressing groups based on the relative expression of Foxf2 compared with the tumor pool (log fold change = 0). Expression of Foxf2 was correlated with time to metastasis (**c, f**) or overall survival (**d, e**) either in all tumors analyzed (**c**; $n = 288$) or stratified for estrogen receptor (ER) expression (**e, f**; ER-negative: $n = 68$; ER-positive: $n = 220$). Foxf2 expression does not correlate with tumor metastasis in the total tumor pool (**c**), but high Foxf2 expression is predictive for overall patient survival in luminal subtype B (**d**; $p = 0.00131$) and ER-negative tumors (**e**; $p = 0.0154$), respectively. High Foxf2 expression levels also correlate with an early onset of metastasis in ER-negative patients (**f**; $p = 0.007$). **g, h** Statistical analysis of the Metabric database ($n = 1298$ tumors). **g** Tumors were divided into high and low Foxf2-expressing groups based on the median expression of Foxf2 mRNA. In the luminal subtype B tumors ($n = 352$), Foxf2 expression is predictive for overall survival (p value of the likelihood-ratio test). **h** Expression of Foxf2 across multiple different tumor subtypes shows significantly increased levels in ER-negative (ERneg) vs ER-positive (ERpos) tumors, in triple negative (TNeg), and in claudin-low (CI_low) compared with all other tumors (p value of the Kruskal-Wallis test; * $p \leq 0.05$; *** $p \leq 0.001$)

and *Areg* genes, encoding for ligands of EGFR, and thus attenuates EGFR-mediated survival signaling (Fig. 7).

The failure of Foxf2-depleted cells to disrupt tight and adherens junctions exemplifies its indispensable role in the acquisition of an invasive cell morphology. Loss of E-cadherin is an early event during EMT resulting in the disruption of the polarity complex, a prerequisite for the dissociation and invasion of cancer cells [67–69]. Direct

transcriptional repression has emerged as one common regulatory mechanism of E-cadherin expression in various cancer types [70, 71]. Here we demonstrate that, by mediating the TGF β -induced upregulation of the transcriptional repressors Zeb1 and Zeb2 as well as the repression of Id2, Foxf2 mediates the transcriptional downregulation of E-cadherin and consequently the disruption of cell-cell adhesion. In addition, we show that



Foxf2 also affects the expression of the miR-200 family which are potent EMT-repressing noncoding RNAs that target Zeb1 and Zeb2 transcripts [72–74]. Interestingly, the expression of Foxf2 and miR-200 are controlled in a double-negative feedback loop, similar to the well-studied Zeb1-miR200 loop. The mechanism by which Foxf2 regulates the expression of these transcription (co)factors remains elusive, but the presence of putative Foxf2 binding sites in the promoter region of Zeb1 and Zeb2 (data not shown) and data from Kundu et al. [75] in nonsmall-cell lung cancer (NSCLC) cells suggest a direct regulatory mechanism, thus ensuring the downregulation of cell-cell junctions at multiple levels (Fig. 7). Moreover, consistent with our finding that Foxf2 predicts poor survival in a subset of breast cancer patients, our results identify Foxf2 as a promigratory and prometastatic factor (Fig. 7).

Our results also show that Foxf2 is essential for TGFβ-mediated apoptosis. The reduction of TGFβ-induced apoptosis in Foxf2-deficient cells is a consequence of the loss of the transcriptional activation of *Noxa* gene expression. *Noxa* has been shown to be critical in fine-tuning cell death decisions via degradation of the prosurvival molecule Mcl1 [76]. *Noxa* has been identified as a primary p53 target gene; however, oncogenic stress, such as irradiation and hypoxia, results in efficient induction of *Noxa* also in the absence of p53 [77, 78]. Our findings identify Foxf2 as a novel transcriptional activator of the tumor suppressor *Noxa* (Fig. 7).

Besides triggering apoptosis via regulation of *Noxa* expression, Foxf2 is also involved in the negative regulation of survival signals by the transcriptional repression of the EGFR ligands betacellulin (Btc), amphiregulin (Areg), and, to a lesser degree, epiregulin (Ereg). Although the regulatory effect of Foxf2 on the transcription level of

the individual EGFR ligands is moderate (Additional file 1: Figure S5B, C), a cumulative effect by the simultaneous modulation of Btc, Areg, and Ereg expression enables a significant shift towards less EGFR signaling (Fig. 5a). Reduced EGFR signaling leads to a reduction in EGFR phosphorylation and, hence, reduced PKB activation. Blocking EGFR signaling has been shown to amplify the apoptotic response to TGF β [61]. Here, we show that both pharmacological receptor inhibition as well as the combined reduction of the expression of the EGFR ligands Btc, Areg, and Ereg increased TGF β -induced apoptosis in Foxf2-depleted NMuMG cells. These results indicate that, in addition to Noxa regulation, Foxf2 mediates its apoptotic effect by blocking EGFR-mediated survival signaling (Fig. 7).

To support the importance of a Foxf2 function during tumor development and progression, we have performed correlation studies for Foxf2 on three different breast cancer databases [62–64]. Analysis of a lymph node-negative stratified patient subset demonstrates that low Foxf2 expression significantly correlates with early metastasis formation. Comparably, low Foxf2 expression has been recently reported to correlate with early-onset metastasis and poor prognosis in breast cancer patients [43]. Interestingly, the opposite is found in more progressive breast cancers, where high Foxf2 expression correlates with poor prognosis. Stratification for luminal subtype B or for ER status reveals a highly significant correlation of high Foxf2 expression and early metastasis onset, concomitant with reduced overall survival. ER⁻ tumors represent highly aggressive breast cancer subtypes. In addition, Foxf2 transcript levels are increased in more aggressive ER⁻, in triple negative, and in the EMT-like, claudin-low breast cancer subtypes. Together, our findings identify high levels of Foxf2 as a marker for good prognosis in early noninvasive stages of tumor development, but with poor prognosis in malignant stages [79]. These findings substantiate the dual role of Foxf2 in cancer patients.

Conclusions

In our study, we demonstrate a dual role for the transcription factor Foxf2. It induces proapoptotic and represses antiapoptotic genes and, thus, acts as a tumor suppressor, likely with the help of specific cofactors. On the other hand, it induces the expression of EMT-inducing genes and thus exerts prometastatic functions to cells that have overcome the apoptotic crisis and undergone EMT. The role of Foxf2 in pre- and post-EMT cells reflects the well-studied dual role of TGF β in cancer progression. Our results also substantiate findings in knockout mouse models where Foxf2 was found to play an important role in EMT-associated developmental processes and maintenance of the epithelial-mesenchymal structure in lung and gut tissues

[38, 39]. Hence, fine-tuning of the expression of Foxf2 and its cofactors could be pivotal during carcinogenesis, and insights into its regulation and molecular function are critical for the design of novel therapeutic strategies.

Additional file

Additional file 1: Figure S1. Foxf2 expression is upregulated during EMT via the canonical Smad pathway. The increased expression of Foxf2 during an EMT was assessed in normal murine mammary gland epithelial cells (NMuMG) and in murine and human breast cancer cells. **Figure S2.** Foxf2 is required for TGF β -induced disruption of adherens junctions. RNAi-mediated ablation prevents an EMT as visualized by immunofluorescence staining for epithelial and mesenchymal markers. **Figure S3.** Foxf2 regulates the expression of Zeb1, Zeb2, Id2, and members of the miR-200 family as determined by quantitative RT-PCR of their expression during an EMT in the absence or presence of Foxf2. **Figure S4.** Foxf2 regulates Noxa expression and thus affects cell proliferation and apoptosis. Foxf2 regulated the expression of Noxa, and siRNA-mediated depletion of Noxa prevented an increase in cell death induced by the loss of Foxf2 expression as assessed by quantitative RT-PCR. **Figure S5.** EGF ligand-mediated EGF receptor signaling overcomes Foxf2-controlled cell survival. Foxf2 represses the expression of EGF receptor ligands as assessed by quantitative RT-PCR. **Supplementary material and methods.** Detailed information is given on the antibodies and reagents, on biochemical and cell biological methods, and on RNA sequencing and bioinformatics analysis used in the study. **Table S1.** Excel file summarizing the differential expression analysis (siFoxf2 to siCtrl after 4 days TGF β treatment or siCtrl with vs without TGF β for 4 days) of all transcripts detected with RNA-sequencing. **Table S2.** Excel file showing the list of genes belonging to the different gene signatures (modules) and the strength of their modular membership (kME values). (ZIP 14675 kb)

Abbreviations

Areg: Amphiregulin; Btc: Betacellulin; ChIP: Chromatin immunoprecipitation; EGFR: Epidermal growth factor receptor; EGFRi: EGFR inhibitor; EMT: Epithelial to mesenchymal transition; Ereg: Epiregulin; Foxf2: Forkhead box protein F2; Ids: Inhibitors of differentiation; IPA: Ingenuity pathway analysis; MET: Mesenchymal to epithelial transition; NMuMG: Normal murine mammary gland; PARP: Poly-(ADP-ribose) polymerase; TGF: Transforming growth factor; WGCNA: Weighted gene coexpression network analysis

Acknowledgments

We thank Drs. L. Lundh, M. Kaeser, and P. ten Dijke for sharing important reagents. We are grateful to P. Schmidt, H. Antoniadis, I. Galm, U. Schmieler, and R. Jost for excellent technical support. We thank P. Lorentz and the DBM microscopy facility (DBM, University of Basel) for excellent support with microscopy, C. Beisel and the Genomics Facility Basel (D-BSE) for next generation RNA sequencing, and R. Ivanek and the Bioinformatics Core Facility (DBM, University of Basel) for help with RNA sequencing analysis.

Funding

This research has been supported by the Swiss National Science Foundation, the Swiss Initiative for Systems Biology (RTD Cellplasticity), and the Swiss Cancer League. NM-S was supported by a Marie-Heim Vögtlin grant from the Swiss National Foundation.

Availability of data and materials

The RNA expression data from the RNA-sequencing are deposited at Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>; GEO accession number: GSE112796).

Authors' contributions

NM-S, CH, ST, and MY designed and performed the experiments, analyzed the data, and wrote the manuscript. GC designed the experiments, analyzed the data, and wrote the manuscript. All authors read and approved the final manuscript.

Ethics approval and consent to participate

In this work, no experiments have been performed with patients or patient material, and hence no approval or consent was required. Human cell lines have been purchased from ATCC. Mouse cell lines have been derived under the Animal Experiment Permissions 1887, 1907, and 1908 by the Swiss Authorities (Kantonale Veterinärämter Basel-Stadt).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

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

Author details

¹Department of Biomedicine, University of Basel, Mattenstrasse 28, 4058 Basel, Switzerland. ²Present address: Institute of Pathology, University Hospital of Basel, Basel, Switzerland. ³Present address: Integra Biosciences AG, Zizers, Switzerland. ⁴Present address: Roche Pharma, Basel, Switzerland.

Received: 23 March 2018 Accepted: 20 August 2018

Published online: 01 October 2018

References

- Yang J, Weinberg RA. Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis. *Dev Cell*. 2008;14(6):818–29.
- Kalluri R. EMT: when epithelial cells decide to become mesenchymal-like cells. *J Clin Invest*. 2009;119(6):1417–9.
- Thiery JP, Acloque H, Huang RY, Nieto MA. Epithelial-mesenchymal transitions in development and disease. *Cell*. 2009;139(5):871–90.
- Chaffer CL, San Juan BP, Lim E, Weinberg RA. EMT, cell plasticity and metastasis. *Cancer Metastasis Rev*. 2016;35(4):645–54.
- Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science*. 2013;342(6159):1234850.
- Bloushtain-Qimron N, Yao J, Snyder EL, Shipitsin M, Campbell LL, Mani SA, Hu M, Chen H, Ustyansky V, Antosiewicz JE, et al. Cell type-specific DNA methylation patterns in the human breast. *Proc Natl Acad Sci U S A*. 2008;105(37):14076–81.
- Shipitsin M, Campbell L, Argani P, Wieremowicz S, Bloushtain-Qimron N, Yao J, Nikolskaya T, Serebryskaya T, Beroukhim R, Hu M, et al. Molecular definition of breast tumor heterogeneity. *Cancer Cell*. 2007;11(3):259–73.
- Graff JR, Gabrielson E, Fujii H, Baylín SB, Herman JG. Methylation patterns of the E-cadherin 5' CpG island are unstable and reflect the dynamic, heterogeneous loss of E-cadherin expression during metastatic progression. *J Biol Chem*. 2000;275(4):2727–32.
- Shipitsin M, Polyak K. The cancer stem cell hypothesis: in search of definitions, markers, and relevance. *Lab Invest*. 2008;88(5):459–63.
- Brabletz T, Hlubek F, Spaderna S, Schmalhofer O, Hiendlmeyer E, Jung A, Kirchner T. Invasion and metastasis in colorectal cancer: epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. *Cells Tissues Organs*. 2005;179(1–2):56–65.
- Chaffer C, Thompson E, Williams E. Mesenchymal to epithelial transition in development and disease. *Cells Tissues Organs*. 2007;185(1–3):7–19.
- Ocana OH, Corcoles R, Fabra A, Moreno-Bueno G, Acloque H, Vega S, Barrallo-Gimeno A, Cano A, Nieto MA. Metastatic colonization requires the repression of the epithelial-mesenchymal transition inducer Prrx1. *Cancer Cell*. 2012;22(6):709–24.
- Tsai JH, Donaher JL, Murphy DA, Chau S, Yang J. Spatiotemporal regulation of epithelial-mesenchymal transition is essential for squamous cell carcinoma metastasis. *Cancer Cell*. 2012;22(6):725–36.
- Fischer KR, Durrans A, Lee S, Sheng J, Li F, Wong ST, Choi H, El Rayes T, Ryu S, Troeger J, et al. Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance. *Nature*. 2015;527(7579):472–6.
- Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H, Wu CC, LeBleu VS, Kalluri R. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature*. 2015;527(7579):525–30.
- Aiello NM, Brabletz T, Kang Y, Nieto MA, Weinberg RA, Stanger BZ. Upholding a role for EMT in pancreatic cancer metastasis. *Nature*. 2017;547(7661):E7–8.
- Ye X, Brabletz T, Kang Y, Longmore GD, Nieto MA, Stanger BZ, Yang J, Weinberg RA. Upholding a role for EMT in breast cancer metastasis. *Nature*. 2017;547(7661):E1–3.
- Krebs AM, Mitschke J, Lasierra Losada M, Schmalhofer O, Boerries M, Busch H, Boettcher M, Mougiakakos D, Reichardt W, Bronsert P, et al. The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nat Cell Biol*. 2017;19(5):518–29.
- Polyak K, Weinberg RA. Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. *Nat Rev Cancer*. 2009;9(4):265–73.
- Yilmaz M, Maass D, Tiwari N, Waldmeier L, Schmidt P, Lehembre F, Christofori G. Transcription factor Dlx2 protects from TGFbeta-induced cell-cycle arrest and apoptosis. *EMBO J*. 2011;30(21):4489–99.
- Tiwari N, Meyer-Schaller N, Arnold P, Antoniadis H, Pachkov M, van Nimwegen E, Christofori G. Klf4 is a transcriptional regulator of genes critical for EMT, including Jnk1 (Mapk8). *PLoS One*. 2013;8(2):e57329.
- Tiwari N, Tiwari VK, Waldmeier L, Balwierz PJ, Arnold P, Pachkov M, Meyer-Schaller N, Schubeler D, van Nimwegen E, Christofori G. Sox4 is a master regulator of epithelial-mesenchymal transition by controlling Ezh2 expression and epigenetic reprogramming. *Cancer Cell*. 2013;23(6):768–83.
- Lehembre F, Yilmaz M, Wicki A, Schomber T, Strittmatter K, Ziegler D, Kren A, Went P, Derksen PW, Berns A, et al. NCAM-induced focal adhesion assembly: a functional switch upon loss of E-cadherin. *EMBO J*. 2008;27(19):2603–15.
- Kaufmann E, Knochel W. Five years on the wings of fork head. *Mech Dev*. 1996;57(1):3–20.
- Kaestner KH, Knochel W, Martinez DE. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev*. 2000;14(2):142–6.
- Wotton KR, Shimeld SM. Analysis of lamprey clustered fox genes: insight into fox gene evolution and expression in vertebrates. *Gene*. 2011;489(1):30–40.
- Fuxe J, Vincent T, Garcia de Herreros A. Mesenchymal cross-talk between TGF-beta and stem cell pathways in tumor cell invasion: role of EMT promoting Smad complexes. *Cell Cycle*. 2010;9(12):2363–74.
- Mani SA, Yang J, Brooks M, Schwanninger G, Zhou A, Miura N, Kutok JL, Hartwell K, Richardson AL, Weinberg RA. Mesenchyme forkhead 1 (FOXC2) plays a key role in metastasis and is associated with aggressive basal-like breast cancers. *Proc Natl Acad Sci U S A*. 2007;104(24):10069–74.
- Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, Yang J, Hartwell K, Onder TT, Gupta PB, Evans KW, et al. Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. *Proc Natl Acad Sci U S A*. 2010;107(35):15449–54.
- Bao B, Wang Z, Ali S, Kong D, Banerjee S, Ahmad A, Li Y, Azmi AS, Miele L, Sarkar FH. Over-expression of FoxM1 leads to epithelial-mesenchymal transition and cancer stem cell phenotype in pancreatic cancer cells. *J Cell Biochem*. 2011;112(9):2296–306.
- Aitola M, Carlsson P, Mahlapuu M, Enerback S, Pelto-Huikko M. Forkhead transcription factor FoxF2 is expressed in mesodermal tissues involved in epithelial-mesenchymal interactions. *Dev Dyn*. 2000;218(1):136–49.
- Blixt A, Mahlapuu M, Bjursell C, Darnfors C, Johannesson T, Enerback S, Carlsson P. The two-exon gene of the human forkhead transcription factor FREAC-2 (FKHL6) is located at 6p25.3. *Genomics*. 1998;53(3):387–90.
- Hellqvist M, Mahlapuu M, Blixt A, Enerback S, Carlsson P. The human forkhead protein FREAC-2 contains two functionally redundant activation domains and interacts with TBP and TFIIIB. *J Biol Chem*. 1998;273(36):23335–43.
- Lan Y, Jiang R. Sonic hedgehog signaling regulates reciprocal epithelial-mesenchymal interactions controlling palatal outgrowth. *Development*. 2009;136(8):1387–96.
- Katoh Y, Katoh M. Hedgehog signaling, epithelial-to-mesenchymal transition and miRNA (review). *Int J Mol Med*. 2008;22(3):271–5.
- Yu W, Ruest LB, Svoboda KK. Regulation of epithelial-mesenchymal transition in palatal fusion. *Exp Biol Med (Maywood)*. 2009;234(5):483–91.
- Wilkie AO, Morriss-Kay GM. Genetics of craniofacial development and malformation. *Nat Rev Genet*. 2001;2(6):458–68.
- Wang T, Tamakoshi T, Uezato T, Shu F, Kanzaki-Kato N, Fu Y, Koseki H, Yoshida N, Sugiyama T, Miura N. Forkhead transcription factor Foxf2 (LUN)-deficient mice exhibit abnormal development of secondary palate. *Dev Biol*. 2003;259(1):83–94.

39. Ormestad M, Astorga J, Landgren H, Wang T, Johansson BR, Miura N, Carlsson P. Foxf1 and Foxf2 control murine gut development by limiting mesenchymal Wnt signaling and promoting extracellular matrix production. *Development*. 2006;133(5):833–43.
40. Shi W, Gerster K, Alajez NM, Tsang J, Waldron L, Pintilie M, Hui AB, Sykes J, P'ng C, Miller N, et al. MicroRNA-301 mediates proliferation and invasion in human breast cancer. *Cancer Res*. 2011;71(8):2926–37.
41. Lo HW, Hsu SC, Xia W, Cao X, Shih JY, Wei Y, Abbruzzese JL, Hortobagyi GN, Hung MC. Epidermal growth factor receptor cooperates with signal transducer and activator of transcription 3 to induce epithelial-mesenchymal transition in cancer cells via up-regulation of TWIST gene expression. *Cancer Res*. 2007;67(19):9066–76.
42. Nik AM, Reyahi A, Ponten F, Carlsson P. Foxf2 in intestinal fibroblasts reduces numbers of Lgr5(+) stem cells and adenoma formation by inhibiting Wnt signaling. *Gastroenterology*. 2013;144(5):1001–11.
43. Kong PZ, Yang F, Li L, Li XQ, Feng YM. Decreased FOXF2 mRNA expression indicates early-onset metastasis and poor prognosis for breast cancer patients with histological grade II tumor. *PLoS One*. 2013;8(4):e61591.
44. Cai J, Tian AX, Wang QS, Kong PZ, Du X, Li XQ, Feng YM. FOXF2 suppresses the FOXC2-mediated epithelial-mesenchymal transition and multidrug resistance of basal-like breast cancer. *Cancer Lett*. 2015;367(2):129–37.
45. Wang QS, Kong PZ, Li XQ, Yang F, Feng YM. FOXF2 deficiency promotes epithelial-mesenchymal transition and metastasis of basal-like breast cancer. *Breast Cancer Res*. 2015;17:30.
46. Lo PK. The controversial role of forkhead box F2 (FOXF2) transcription factor in breast cancer. *PRAS Open*. 2017;1.
47. Maeda M, Johnson KR, Wheelock MJ. Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition. *J Cell Sci*. 2005;118(Pt 5):873–87.
48. Deckers M, van Dinther M, Buijs J, Que I, Lowik C, van der Pluijm G, ten Dijke P. The tumor suppressor Smad4 is required for transforming growth factor beta-induced epithelial to mesenchymal transition and bone metastasis of breast cancer cells. *Cancer Res*. 2006;66(4):2202–9.
49. Waldmeier L, Meyer-Schaller N, Diepenbruck M, Christofori G. Py2T murine breast cancer cells, a versatile model of TGFbeta-induced EMT in vitro and in vivo. *PLoS One*. 2012;7(11):e48651.
50. Hellqvist M, Mahlapuu M, Samuelsson L, Enerback S, Carlsson P. Differential activation of lung-specific genes by two forkhead proteins, FREAC-1 and FREAC-2. *J Biol Chem*. 1996;271(8):4482–90.
51. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*. 2007;39(4):457–66.
52. Muller WJ, Sinn E, Pattengale PK, Wallace R, Leder P. Single-step induction of mammary adenocarcinoma in transgenic mice bearing the activated c-neu oncogene. *Cell*. 1988;54(1):105–15.
53. Derksen PW, Liu X, Saridin F, van der Gulden H, Zevenhoven J, Evers B, van Beijnum JR, Griffioen AW, Vink J, Krimpenfort P, et al. Somatic inactivation of E-cadherin and p53 in mice leads to metastatic lobular mammary carcinoma through induction of anoikis resistance and angiogenesis. *Cancer Cell*. 2006;10(5):437–49.
54. Piek E, Moustakas A, Kurisaki A, Heldin CH, ten Dijke P. TGF-(beta) type I receptor/ALK-5 and Smad proteins mediate epithelial to mesenchymal transdifferentiation in NMuMG breast epithelial cells. *J Cell Sci*. 1999;112(Pt 24):4557–68.
55. Brabletz S, Brabletz T. The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO Rep*. 2010;11(9):670–7.
56. Caramel J, Ligier M, Puisieux A. Pleiotropic roles for ZEB1 in cancer. *Cancer Res*. 2018;78(1):30–5.
57. Gheldof A, Hulpiau P, van Roy F, De Craene B, Bex G. Evolutionary functional analysis and molecular regulation of the ZEB transcription factors. *Cell Mol Life Sci*. 2012;69(15):2527–41.
58. Kondo M, Cubillo E, Tobiume K, Shirakihara T, Fukuda N, Suzuki H, Shimizu K, Takehara K, Cano A, Saitoh M, et al. A role for Id in the regulation of TGF-beta-induced epithelial-mesenchymal transdifferentiation. *Cell Death Differ*. 2004;11(10):1092–101.
59. Derynck R, Akhurst RJ, Balmain A. TGF-beta signaling in tumor suppression and cancer progression. *Nat Genet*. 2001;29(2):117–29.
60. Fabregat I, Herrera B, Fernandez M, Alvarez AM, Sanchez A, Roncero C, Ventura JJ, Valverde AM, Benito M. Epidermal growth factor impairs the cytochrome C/caspase-3 apoptotic pathway induced by transforming growth factor beta in rat fetal hepatocytes via a phosphoinositide 3-kinase-dependent pathway. *Hepatology*. 2000;32(3):528–35.
61. Murillo MM, del Castillo G, Sanchez A, Fernandez M, Fabregat I. Involvement of EGF receptor and c-Src in the survival signals induced by TGF-beta1 in hepatocytes. *Oncogene*. 2005;24(28):4580–7.
62. Minn A, Gupta G, Siegel P, Bos P, Shu W, Giri D, Viale A, Olshen A, Gerald W, Massague J. Genes that mediate breast cancer metastasis to lung. *Nature*. 2005;436(7050):518–24.
63. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature*. 2002;415(6871):530–6.
64. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, Speed D, Lynch AG, Samarajiwa S, Yuan Y, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*. 2012;486(7403):346–52.
65. Dvinge H, Git A, Graf S, Salmon-Divon M, Curtis C, Sottoriva A, Zhao Y, Hirst M, Armitage J, Miska EA, et al. The shaping and functional consequences of the microRNA landscape in breast cancer. *Nature*. 2013;497(7449):378–82.
66. Thiery JP, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. *Nat Rev Mol Cell Biol*. 2006;7(2):131–42.
67. Birchmeier W, Behrens J. Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness. *Biochim Biophys Acta*. 1994;1198(1):11–26.
68. Christofori G. New signals from the invasive front. *Nature*. 2006;441(7092):444–50.
69. Kalluri R, Weinberg R. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009;119(6):1420–8.
70. Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer*. 2007;7(6):415–28.
71. Perk J, Iavarone A, Benezra R. Id family of helix-loop-helix proteins in cancer. *Nat Rev Cancer*. 2005;5(8):603–14.
72. Park SM, Gaur AB, Lengyel E, Peter ME. The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2. *Genes Dev*. 2008;22(7):894–907.
73. Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol*. 2008;10(5):593–601.
74. Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, Brabletz T. A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. *EMBO Rep*. 2008;9(6):582–9.
75. Kundu ST, Byers LA, Peng DH, Roybal JD, Diao L, Wang J, Tong P, Creighton CJ, Gibbons DL. The miR-200 family and the miR-183~96~182 cluster target Foxf2 to inhibit invasion and metastasis in lung cancers. *Oncogene*. 2016;35(2):173–86.
76. Willis SN, Chen L, Dewson G, Wei A, Naik E, Fletcher JL, Adams JM, Huang DC. Proapoptotic Bak is sequestered by Mcl-1 and Bcl-xL, but not Bcl-2, until displaced by BH3-only proteins. *Genes Dev*. 2005;19(11):1294–305.
77. Kim JY, Ahn HJ, Ryu JH, Suk K, Park JH. BH3-only protein Noxa is a mediator of hypoxic cell death induced by hypoxia-inducible factor 1alpha. *J Exp Med*. 2004;199(1):113–24.
78. Ploner C, Kofler R, Villunger A. Noxa: at the tip of the balance between life and death. *Oncogene*. 2009;27:S84–92.
79. Dairkee SH, Ljung BM, Smith H, Hackett A. Immunolocalization of a human basal epithelium specific keratin in benign and malignant breast disease. *Breast Cancer Res Treat*. 1987;10(1):11–20.