Technische Universität München

Fakultät für Medizin

Development and characterization of an *in vitro* assay that elicits invasive potential from human mammary luminal progenitor cells

Hilary Marlene Ganz

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades einer **Doktorin der Naturwissenschaften (Dr. rer. nat.)** genehmigten Dissertation.

Vorsitz: Prof. Dr. Susanne Kossatz

Prüfende/-r der Dissertation:

1. Prof. Dr. Heiko Lickert

2. Prof. Dr. Wolfgang Wurst

Die Dissertation wurde am 15.11.2021 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 12.04.2022 angenommen.

Within the scope of the PhD thesis presented here, the following original article has been published in an international peer-reviewed journal:

Ganz, H. M., Buchmann, B., Engelbrecht, L. K., Jesinghaus, M., Eichelberger, L., Gabka, C. J., Schmidt, G. P., Muckenhuber, A., Weichert, W., Bausch, A. R., & Scheel, C. H. (2021). Generation of ductal organoids from normal mammary luminal cells reveals invasive potential. The Journal of Pathology. https://doi.org/10.1002/path.5790

Summary

The histopathological analysis of invasive carcinomas unequivocally shows that during tumorigenesis, proliferating luminal cells have the potential to invade into the surrounding extracellular matrix (ECM). Particularly in carcinomas of no special type (NST), the most common subtype of breast cancer, invading luminal cells undergo a morphogenetic process. In detail, a network of ducts is formed that especially in low-grade carcinomas, resembles the branched, ductal morphology of the normal mammary gland. However, to this day, no *in vitro* model has been able to recapitulate the invasive branching capacity displayed by luminal cells *in vivo*. Interestingly, while genetic aberrations have been connected to disease morphology, it has so far not been possible to identify universal genetic aberrations that discriminate invasively growing luminal cells from those that remain noninvasive during tumorigenesis. Based on these findings, it is possible that cellular context, rather than genetic aberrations, determines whether luminal breast cancer cells acquire invasiveness.

In this study, suitable *in vitro* conditions were created that elicit invasive potential from normal luminal progenitor cells. To this end, viability of the luminal subset was ensured by optimizing culture conditions. Moreover, collagen type I was employed as matrix based on the invasion-promoting impact of this ECM component as well as its abundance in invasive breast cancers. Employing these specific culture conditions enabled FACS-sorted primary human luminal progenitor cells to display an innate potential to invade and generate multi-cellular, branched, ductal structures while maintaining luminal characteristics such as expression of known lineage markers and cell-cell adhesion components. Consequently, the phenotype of the generated organoids was reminiscent of low-grade NST carcinomas. Moreover, reduced actomyosin contractility was identified as a prerequisite for collective luminal cell invasion via matrix-remodeling luminal leader cells at the tips of the formed branches. Knock-out of *CDH1* encoding for adherens junction component E-cadherin using the CRISPR-Cas9 system caused a morphological switch from a ductal phenotype to diffuse invasion. The latter morphology was reminiscent of lobular carcinoma of the breast, a rare subtype of breast cancer characterized by loss-of-function mutations or promoter methylation of *CDH1*. These observations revealed that genetically modified organoids showed an aberrant morphology reflecting the gene-specific impact of certain aberrations on *in vivo* disease morphology.

Zusammenfassung

Die histopathologische Analyse invasiver Karzinome zeigt auf, dass proliferierende Luminalzellen während der Tumorentstehung das Potenzial haben, in die angrenzende extrazelluläre Matrix (EZM) einzudringen. Insbesondere bei Karzinomen des nicht-spezifischen Typs (NST), dem häufigsten Subtyp von Brustkrebs, durchlaufen die eindringenden Luminalzellen dabei einen morphogenetischen Prozess. Genauer gesagt bildet sich ein Netzwerk von Gängen, das insbesondere bei niedriggradigen Karzinomen Ähnlichkeiten mit der verzweigten, duktalen Morphologie der normalen Brustdrüse aufweist. Bis heute ist jedoch kein *in vitro*-Modell in der Lage, die Fähigkeit der Luminalzellen, während des invasiven Wachstumes verzweigte Gänge zu bilden, zu rekapitulieren. Verschiedene genetische Aberrationen konnten mit der Krankheitsmorphologie in Verbindung gebracht werden. Jedoch war es bisher nicht möglich, universelle genetische Aberrationen zu identifizieren, die invasiv wachsende Luminalzellen von solchen unterscheiden, die während der Tumorentstehung nicht-invasiv bleiben. Diese Erkenntnisse weisen womöglich darauf hin, dass vielmehr der zelluläre Kontext und nicht genetische Aberrationen darüber entscheiden, ob luminale Brustkrebszellen invasiv werden.

In dieser Studie wurden geeignete *in vitro* Bedingungen geschaffen, um invasives Potenzial in normalen, luminalen Vorläuferzellen freizusetzen. Zu diesem Zweck wurde zunächst durch Optimierung der Kulturbedingungen die Lebensfähigkeit dieses luminalen Zelltyps sichergestellt. Darüber hinaus wurde Kollagen Typ I als Matrix verwendet, da diese EZM-Komponente eine invasionsfördernde Wirkung hat und bei invasivem Brustkrebs verstärkt angereichert ist. Tatsächlich ermöglichte die Anwendung dieser spezifischen Kulturbedingungen Invasion und Bildung multizellulärer, verzweigter, duktaler Strukturen aus FACS-sortierten, humanen, luminalen Vorläuferzellen. Gleichzeitig behielten die Zellen ihre luminalen Charakteristika, wie die Expression bekannter abstammungsspezifischer Marker und Zell-Zell-Adhäsionskomponenten. Folglich erinnerte der Phänotyp der erzeugten Organoide an niedriggradige NST-Karzinome. Darüber hinaus wurde im Rahmen dieser Arbeit verminderte Aktomyosin-Kontraktilität als Voraussetzung für kollektive luminale Zellinvasion ausgemacht. Des Weiteren konnte gezeigt werden, dass Matrix-Remodellierung durch luminale Leitzellen an den Spitzen der gebildeten Gänge, maßgeblich zur Invasion beitrugen. Schließlich wurde herausgefunden, dass die Deletion des Gens *CDH1*, welches für die Adherens Junction-Komponente E-Cadherin kodiert, zu einem Wechsel des duktalen Phänotyps hin zu diffuser Invasion führte. Die letztgenannte Morphologie erinnerte dabei an das lobuläre Karzinom der Brust, eine seltene Unterart von Brustkrebs, die durch Funktionsverlustmutationen oder Promotor-Methylierung von *CDH1* gekennzeichnet ist. Diese

Beobachtungen zeigten, dass genetisch veränderte Organoide eine abweichende Morphologie aufweisen, welche die genspezifischen Auswirkungen bestimmter Aberrationen auf die Morphologie der Krankheit *in vivo* widerspiegelt.

Table of Contents

1 Abbreviations

%	Percent
*	Multiplication symbol
°C	Degree Celsius
µ	Micro
2D	Two-dimensional
3D	Three-dimensional
7-AAD	7-Aminoactinomycin
ATAC	Assay for transposase-accessible chromatin
B	Basal
BCOM	Breast cancer organoid medium
BLOM	Branched luminal organoid medium
bp	Base pair
BRCA1	Breast cancer 1, early onset
BSA	Bovine serum albumin
B-SFU	Branched structure-forming unit
Cas	CRISPR associated
CD	Cluster of differentiation
CDH1	Cadherin-1
cDNA	Complementary DNA
CI	Confidence interval
CK	Cytokeratin
CO_2	Carbon dioxide
CRISPR	Clustered regularly interspaced short palindromic repeats
DAPI	4′,6-diamidino-2-phenylindole
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl-sulfoxide
DNA	Deoxyribonucleic acid
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
ELDA	Extreme limiting dilution analysis

EpCAM	Epithelial cell adhesion molecule
ER	Estrogen receptor
FACS	Fluorescence activated cell sorting
F-actin	Actin filaments
FCS	Fetal calf serum
FGF	Fibroblast growth factor
FSC	Forward scatter
fwd	Forward
g	Gram(s)
GATA3	Trans-acting T-cell-specific transcription factor
gDNA	Genomic DNA
GFP	Green fluorescent protein
gRNA	Guide RNA
hr	hours
H&E	Hematoxylin and eosin
HBSS	Hank’s balanced salt solution
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HER2	Human epidermal growth factor receptor 2
ILC	Invasive lobular carcinoma
Ko	Knock-out
KRT	Keratin
LP	Luminal progenitor
m	Milli
M	Molar
MECGM	Mammary epithelial cell growth medium
ML	Mature luminal
MMP	Matrix-metalloproteinase
mRNA	Messenger RNA
Muc1	Mucin-1
n	Nano
n	Cell count
n.s.	Not significant
NRG1	Neuregulin 1

NST	Carcinomas of no special type
O_2	Oxygen
p	Pico
p63	Tumor protein 63
PAB	Para-amino blebbistatin
PB	Pacific blue
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PE	Phytoerythrin
PR	Progesterone receptor
rev	Reverse
RhoA	Ras homolog family member A
RNA	Ribonucleic acid
ROCK	Rho-associated protein kinase
s.d.	Standard deviation
SSC	Sideward scatter
t	Time
TNS	Trypsin neutralizing solution
TP53	Tumor protein 53
U	Enzyme unit
V	Volume
V	Velocity
wt	Wild-type
x	Mean of counted squares
ZO-1	Zonula occludens-1
α	Alpha
αSMA	Alpha smooth muscle actin

2 Introduction

2.1 The normal adult mammary gland

2.1.1 Composition of the normal adult mammary gland

The adult human mammary gland is built up by a bilayered, branched, ductal network. This network consists of two major mammary epithelial cell types, an inner layer of luminal cells and an outer layer of basal/myoepithelial cells. The two cell types can be distinguished by their characteristic, mutually exclusive, marker expression. Thereby, it has been reported that basal cells have a characteristic expression of the microfilament alpha smooth muscle actin (αSMA), the membrane metallo-endopeptidase CD10 as well as the transcription factor tumor protein 63 (p63) (Santagata et al., 2014) (Fig. 1A,B). Luminal cells can be identified unequivocally by certain intermediate filaments, the cytokeratins (CKs) 7, 8, and 18, certain tight junction proteins such as claudin-4 and zonula occludens-1 (ZO-1) (Santagata et al., 2014) as well as the trans-acting T-cell-specific transcription factor GATA3 (GATA3) (Naylor & Ormandy, 2007).

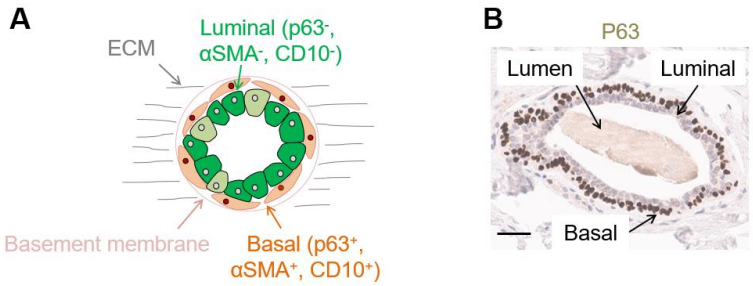

Figure 1. Organization of mammary gland ducts in the normal breast. (A) Schematic illustration of bilayered normal mammary gland duct indicating expression of basal cell markers: inner layer of luminal cells depicted in green ($p63^-$, $αSMA^-$, $CD10^-$) and outer layer of basal/myoepithelial cells depicted in orange ($p63^+$, $αSMA^+$, $CD10^+$). The mammary gland epithelium is surrounded by the basement membrane (pink) and embedded into the ECM (grey). **(B)** Paraffin section of normal mammary duct with immunohistochemical staining for p63 (brown). Scale bar: 50 µm. Immunohistochemistry and associated imaging were performed by Moritz Jesinghaus (Institute of Pathology, Technical University of Munich).

Moreover, the luminal compartment can further be divided into two cellular subsets: mature luminal cells and luminal progenitor cells. Mature luminal cells are described as estrogen receptor (ER) positive and rarely proliferative. Luminal progenitor cells can be ER-positive (ER^+) or ER-negative (ER^-) and comprise progenitors for ductal and alveolar luminal cells (Bach et al., 2017; Eirew et al., 2008; Lim et al., 2009; Shehata et al., 2012).

The epithelium is surrounded by a thin basement membrane composed of polymerized laminins and type IV collagens tethered to glycoproteins and proteoglycans (Yurchenco & Patton, 2009). Beyond the basement membrane, the epithelium is embedded into an extracellular matrix (ECM) rich in collagens (type I and III), proteoglycans, tenascins, hyaluronic acid, and fibronectin (Muschler & Streuli, 2010) (Fig. 1A). Furthermore, the ECM contains a variety of stromal cells such as adipocytes and fibroblasts (Muschler & Streuli, 2010).

2.1.2 Bipotency in epithelial cells of the normal adult mammary gland

The female adult mammary gland is a remarkably dynamic organ that goes through a cyclic evolution process during pregnancy, lactation, and involution. Upon onset of pregnancy, mammary epithelial cells undergo intensive mitosis. Thereby, alveolar buds are generated that further differentiate into distinct alveoli and finally become milk-secreting lobules during lactation (Macias & Hinck, 2012). Here, the epithelial lineages execute distinct functions. Thereby, the luminal cells produce milk and seal the ducts via tight junctions (D.-A. D. Nguyen & Neville, 1998). At the same time, the basal cells are contractile, which is a crucial competence for transportation of milk produced by luminal cells along the ductal network and for its ejection (Deugnier et al., 1995; Haaksma et al., 2011). Upon weaning, the process of involution is initiated during which the epithelial tree is remodeled back into its nearly original state. In detail, alveoli collapse and milk-producing cells are removed (Jena et al., 2019; Macias & Hinck, 2012).

The occurrence of such major developmental processes during adulthood requires the presence of uniquely potent cells to allow extensive cycles of expansion and renewal. Consequently, the differentiation hierarchy of adult human mammary epithelial cells has been an active field of investigation. Thereby, one major point of discussion is whether during mammary gland remodeling in the adult, the luminal and basal lineage are recreated from one bipotent progenitor cell type or from two separate, lineage-restricted basal and luminal progenitor cells, respectively. To resolve this matter, researchers have tried to uncover the existence of adult stem cells with bipotent potential within the two epithelial subsets.

One option to approach this question is the isolation of epithelial lineages from the human mammary gland and subsequent monitoring of one single basal or luminal cell’s ability to recreate multilineage tissues. Thereby, cellular development can either be observed *in vitro* or *in vivo* upon transplantation of human epithelial cells into mouse mammary fat pads that have been cleared of the mouse glandular tissue. Doing so, several *in vitro* studies found

that the basal subset harbors cells with bipotent potential as identified by luminal and basal marker expression in cultures inoculated purely with basal cells (Eirew et al., 2008; Linnemann et al., 2015; Stingl et al., 2001). Notably, multilineage potential has also been found in the luminal subset when culture conditions geared towards maintenance of basal cells were used (Pechoux et al., 1999) as well as in reconstitution assays (Gudjonsson, Villadsen, et al., 2002; Rosenbluth et al., 2020; Shehata et al., 2012). However, several studies deemed the regenerative capacity of luminal cells significantly lower than the one of the basal subset (Linnemann et al., 2015; Shehata et al., 2012). Similarly, it was found that upon transplantation of human epithelial cells into cleared mouse mammary fat pads, both lineages had the capacity to repopulate the mouse mammary gland with a multilineage ductal tree (Shehata et al., 2012). However, several of these studies reinforced the notion that the basal subpopulation contains a vast majority of cells with repopulating capacity (Eirew et al., 2008; Shehata et al., 2012; Stingl et al., 2006).

While the above-described *in vitro* and transplantation assays are valuable for determination of cellular regenerative potential, their transferability to actual cellular behavior during normal development *in vivo* is limited by the fact that within these assays, epithelial cells are extracted from their normal tissue architecture and reintroduced into a different environment. In search for a method that does not include drastic micro-environmental changes, lineage tracing has emerged as a tool for studying stem cell properties *in vivo*. Thereby, a combination of lineage-specific promoters with inducible recombinases, multicolor receptor constructs and live-cell imaging allows identification of all progenies of a single cell *in vivo*. However, lineage tracing studies in the adult mouse mammary gland yielded contradicting results. Several groups could trace cells of basal origin that grew into cells of the basal as well as the luminal lineage, indicating the existence of bipotent stem cells (Rios et al., 2014; Song et al., 2019; Van Amerongen et al., 2012; Wang et al., 2015). In contrast, others found that both lineages were exclusively generated by lineage-restricted basal or luminal progenitor cells (Davis et al., 2016; Elias et al., 2017; Prater et al., 2014; Van Keymeulen et al., 2011; Wuidart et al., 2016, 2018). The inconsistencies between lineage tracing studies might be a consequence of study design. Particularly, high labelling frequency and the choice of not completely lineage-restricted promoters for reporter gene expression can result in inaccurate conclusions. For example the supposedly basal-specific promoter for CK14 frequently used to monitor basal cell expansion by for instance Rios et al. (Rios et al., 2014), has recently been found to be active in a subset of luminal cells (Boras-Granic et al., 2014; P. Sun et al., 2010). Analogous to this, it was found that the presumably luminal-specific CK8 can also be expressed in a subset of basal cells at low levels (Koren et al., 2015). In combination with the often high

proportion of labelled cells (Rios et al., 2014; Wang et al., 2015), there is a certain likelihood that more than one clone arises from the same region, and thus erroneously indicates the presence of a bipotent progenitor cell.

Overall, a majority of lineage tracing studies gave reason to question the existence of bipotent stem cells previously observed in high frequencies in *in vitro* and transplantation studies. Interestingly, Van Amerogen et al. found that a subset of cells within the basal population displayed extensive, artificially introduced bipotency upon transplantation that was beyond their capabilities observed under homeostatic conditions (Van Amerongen et al., 2012). This finding could at least partly explain the contradicting results acquired with different methods. However, besides the fact that lineage tracing studies have not delivered an unequivocal answer to the potency discussion, there has been criticism that these studies might not be capable to properly characterize rare and extremely heterogeneous cell populations. Furthermore, lineage tracing studies are restricted to animal models, and thus limited in their relevance for understanding human mammary gland biology.

An additional approach is the analysis of mammary epithelial cells via the assay for transposase-accessible chromatin (ATAC)-sequencing, whereby the chromatin accessibility is investigated genome-wide. Dravis et al. performed ATAC-sequencing in combination with RNA-sequencing on the separate epithelial lineages (Dravis et al., 2018). Doing so, they found that basal mammary cells of adult mice bear high chromatin accessibility at basal as well as luminal genes similar to multipotent fetal mammary stem cells (Dravis et al., 2018). However, RNA-sequencing revealed that in cells of the basal subset, luminal genes are, in contrast to basal genes, only expressed at low levels during tissue homeostasis. Of note, Chung et al. validated these results with ATAC-sequencing on the single cell level (Chung et al., 2019). These findings help to explain how basal cells could be lineage-restricted during homeostasis and yet show multilineage potential when taken out of their original context during transplantation.

Based on the controversial findings in this area, the recent advances of single cell-specific sequencing methods have encountered great interest in the mammary gland field. Here, several single cell RNA-sequencing studies have underpinned that no bipotent stem cells exist in the healthy adult mammary gland. Rather, they found that basal and luminal lineages are clearly transcriptionally segregated (Bach et al., 2017; Giraddi et al., 2018; Martin Carli et al., 2020). In contrast, in a variety of other single cell RNA-sequencing studies, luminal and myoepithelial cells were traced back to a common cellular origin that was transcriptionally related to the basal subset (W. Chen et al., 2019; Q. H. Nguyen et al., 2018; Pal et al., 2017; H. Sun et al., 2018; Thong et al., 2020), once again not fully resolving

this controversial topic. However, a unifying notion that can be made from recent single cell RNA-sequencing studies is that the vast diversity of observed transcriptional profiles has uncovered that the classical models of human mammary gland maturation over-simplify the complex *in vivo* differentiation processes. Instead, these studies point towards a high plasticity throughout epithelial development in the mammary gland whereby cells gradually differentiate through a differentiation trajectory rather than maturating along restricted paths (Anstine & Keri, 2019).

In summary, contradicting results throughout all different techniques have prevented a clear resolution of the question whether bipotent mammary progenitors participate in tissue homeostasis. However, there is irrevocable evidence that bipotency whether existing during tissue homeostasis or not, can artificially be induced within both, luminal and basal epithelial cells under certain circumstances such as transplantation. Moreover, the studies agree that a majority of bipotent potential resides within cells of the basal subpopulation.

2.2 Classification of breast cancer

2.2.1 Epidemiology

Organs like the breast with a high dynamic evolution during adulthood, are prone to erroneous development. With approximately 2.3 million new cases, female breast cancer was, by total numbers, the most frequently diagnosed cancer in 2020, worldwide. In women, it accounted for roughly 25% of all cancers and was the number one cause of cancer-related death (Sung et al., 2021).

Breast cancer is a genetically and clinically heterogeneous disease (Stingl & Caldas, 2007). Thus, classification systems have been developed to organize this heterogeneity into a standardized scheme. The first major subdivision is made between sarcomas and carcinomas. Sarcomas arise from cells of the connective tissue and only make up around 1% of primary breast cancers. Carcinomas arise from cells of the epithelial lineages and make up the other 99% of cases (Al-Benna et al., 2010). As the great majority of cases are carcinomas, further classification systems described here will focus solely on epithelial cancers.

2.2.2 Histopathology

Historically, breast cancers have been classified based on their histology. Here, a first distinction is made between invasive and *in situ* carcinomas. *In situ* cancer cells grow inside of the pre-existing normal lobules or ducts and do not invade the surrounding tissue. Conversely, invasive cancers infiltrate outside of the normal breast lobules and ducts and grow into the breast connective tissue. Thereby, *in situ* cancers are viewed as a nonobligatory precursor lesion for invasive cancers (Cowell et al., 2013). Of all cases diagnosed in the United States in the year 2019, an estimated 15% were *in situ* carcinomas while 85% were categorized as invasive (DeSantis et al., 2019). Among invasive carcinomas, the most common types are invasive ductal carcinoma, also known as invasive carcinoma of no special type (NST), followed by invasive lobular carcinoma (ILC). Within the group of invasive cancers, NST carcinomas comprise over 70% of all cases(Li et al., 2003) while ILCs make up around 10-15% (Ciriello et al., 2015; Desmedt et al., 2016). Of note, the classification as ductal and lobular does not provide information on the spatial origin of the cancer. Instead, the classification of carcinomas as 'ductal' is based on the appearance of ducts that lack the basal cell layer but otherwise resemble the branched ductal network of the normal mammary gland (Fig. 2A,B).

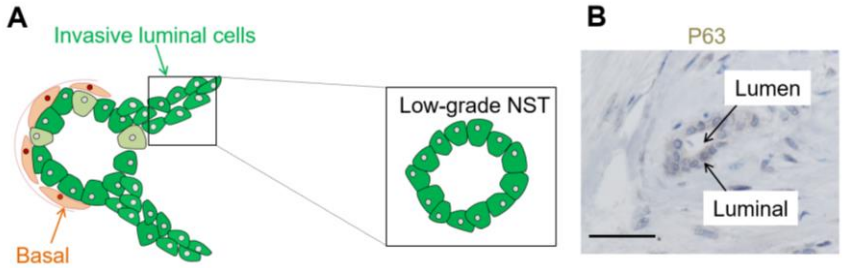

Figure 2. Invasive outgrowth of luminal cells during carcinoma development. (A) Schematic illustration of invasively growing luminal cells (green) growing beyond the borders formed by basal cells (orange). Outgrowing ducts consist of luminal cells and are correctly polarized with a distinct lumen formation (low-grade NST carcinoma) **(B)** Paraffin section of polarized duct arising in low-grade NST carcinoma with immunohistochemical staining for p63 (brown). Scale bar: 50 µm. Immunohistochemistry and associated imaging were performed by Moritz Jesinghaus (Institute of Pathology, Technical University of Munich).

2.2.3 Tumor staging and grading

Solid tumors such as breast cancer are further categorized via staging and grading. Thereby, the breast cancer stage describes the tumor's size and spread. The three main parameters considered for the typically used TNM classification are the size of the tumor (T), the spread to nearby lymph nodes (N) as well as metastases at distant sites (M). The

lowest stage is 0, which is used to describe *in situ* carcinomas while the highest stage is IV, which is used to describe cancer in a highly advanced stage (Sobin et al., 2011). In contrast, grading describes the tumors malignancy level by comparing it to normal mammary gland tissues. Grades from 1 to 3 are given, whereby grade 1 means the tumor cells are well differentiated while grade 3 represents poor differentiation. Three aspects are taken into consideration for grading of breast cancers. The first aspect is the extent of tubulogenesis, meaning the capacity of tumor cells to form ductal structures like in the normal mammary gland. The second aspect is the occurrence of atypical nuclear pleomorphisms, which means it categorizes how much the size and shape of cellular nuclei differ from healthy ones. Finally, the third aspect concerns the mitotic count that gives insight into the growth rate of a tumor (Bloom & Richardson, 1957; Patey & Scarff, 1928).

2.2.4 Molecular subtypes of breast cancer

Within the past two decades, the above-described morphology-based approach for breast cancer classification has been supplemented by approaches that take molecular tumor biology into account. In detail, global gene expression profiles of human breast cancers have unraveled four repeatedly identified intrinsic subtypes of the disease: basal-like, human epidermal growth factor receptor 2 (HER2)-enriched, luminal A and luminal B breast cancers (Perou et al., 2000; Sorlie et al., 2001, 2003). Thereby, the two luminal subtypes display high expression of genes associated with the luminal lineage such as *ESR1* encoding ER. In contrast, HER2-enriched as well as basal-like tumors are characterized by a low expression of genes associated with ER. Besides that, HER2-enriched tumors show amplification of *HER2* and *HER2*-associated genes (Perou et al., 2000; Sorlie et al., 2001). Moreover, as the name suggests, basal-like breast cancers are characterized by the expression of genes associated with basal mammary epithelial cells such as *Keratin 5 (KRT5), Keratin 6 (KRT6) and Keratin 17 (KRT17)* (Perou et al., 2000; Sorlie et al., 2001). Of note, the classification into these breast cancer subtypes bears prognostic value and directs treatment decisions (Goldhirsch et al., 2013; Sorlie et al., 2003).

Clinically, an approximation to these subtypes is made by immunohistochemical assessment of ER, progesterone receptor (PR), HER2 and Ki-67 expression within tumor sections. Thereby, luminal A and luminal B tumors are typically positive for ER and/or PR and mostly negative for HER2 (Goldhirsch et al., 2011). However, luminal A and luminal B display distinct Ki-67 expression, which is typically low in luminal A tumors and high in luminal B tumors (Goldhirsch et al., 2011). In contrast, HER2-enriched and basal-like tumors are typically negative for ER and PR (Goldhirsch et al., 2011). Moreover, HER2-

enriched tumors show strong expression of HER2 (Goldhirsch et al., 2011) while a subset of basal-like tumors expresses basal markers such as p63 and αSMA on the protein level(Livasy et al., 2006).

Taken together, tumor grading and staging as well as classification according to receptor status allow for an assessment of breast cancer with prognostic value. Despite the development of more sophisticated technologies, those classifications are still routinely used for individual treatment decision-making in clinical practice today (Johansson et al., 2021).

2.3 What drives breast tumorigenesis?

2.3.1 Genetic aberrations in breast cancer

A variety of parameters are well-known risk factor for the development of breast cancer. Those factors include old age, late menopause, a family history of breast cancer, high mammographic density and many lifestyle-associated factors (Y. S. Sun et al., 2017). Several of those risk factors are related to a common cause of breast cancer development: the accumulation of genetic lesions. Thereby, two types of lesions can be distinguished, which drive tumorigenesis: somatic alterations that occur de-novo in an individual's body cells and germline mutations, which affect gametes and can be inherited. Such genetic alterations foster tumor development if they either decrease the function of a gene with tumor-suppressing role or reinforce the function of a tumor-promoting oncogene (Hanahan & Weinberg, 2011).

The specific impact of affected genes is many-faceted. The gene most frequently affected in breast cancer is *TP53*, encoding for the tumor suppressor p53, which is involved in key cellular processes such as control of proliferation as well as stability and integrity of the whole genome (Zilfou & Lowe, 2009). Aberrations in *TP53* can be caused by germline as well as somatic variants (Malkin et al., 1990; Packwood et al., 2019) and are found in around 20-30% of breast tumors (Pharoah et al., 1999; Silwal-Pandit et al., 2014).

Of note, while *TP53* mutations can be found in all subtypes of breast cancer, other frequently occurring genetic aberrations are strongly associated with a certain tumor subtype and appear to determine tumor phenotype. One prominent example is the cell adhesion molecule E-cadherin which is frequently lost in breast cancers, mostly due to somatic alterations (Masciari et al., 2007; Rahman et al., 2000; Schrader et al., 2011; Xie et al., 2011). These alterations result in carcinomas of the lobular type, either in *in situ*

lobular carcinomas or ILCs that are histologically characterized by a distinctive growth pattern (Gamallo et al., 1993; Moll et al., 1993).

Another well-known example for a mutant tumor suppressor that affects breast cancer subtype is *BRCA1*, which encodes for proteins contributing to repair of damaged DNA. *BRCA1* mutations are often germline mutations, and thus account for a considerable share of hereditary breast cancers. In detail, long term studies found that the cumulative risk for developing breast cancer by age 70 lies between 46% and 84% in *BRCA1* mutation carriers (S. Chen et al., 2006; Ford, 1994; Mavaddat et al., 2013). While the correlation here is not as distinctive as the one of E-cadherin and lobular carcinoma, the susceptibility to basal-like breast cancers clearly correlates with the presence of dysfunctional *BRCA1* (Turner & Reis-Filho, 2006).

In summary, genetic alterations are a major risk factor for breast cancer development. Thereby, certain alterations encourage the development of specific subtypes of breast cancer.

2.3.2 Cancer cell of origin

Following the histological and functional heterogeneity of breast carcinomas, the possibility has been discussed that the different malignancies are not only determined by genetic events but arise from different epithelial 'cells of origin'.

The cell of origin for luminal A and luminal B malignancies has, largely due to the luminal characteristics of those breast cancer subtypes, mostly been suspected in the luminal compartment. However, based on the absence of ER and PR expression in most HER2-enriched and basal-like malignancies, the cellular origin of those cancers is less obvious. In particular, based on the expression of basal markers in the basal-like cancer subtype, it has been proposed that basal-like breast cancers originate from the basal compartment (Foulkes, 2004). This notion was further corroborated by the theory that highly plastic, long-lived stem cells reside within the basal compartment as described above (Rios et al., 2014; Van Amerongen et al., 2012; Wang et al., 2015). Consequently, those cells are supposedly susceptible to oncogenic transformation over time while transformation of such plastic cells via for instance aberrations in *BRCA1* functionality, would explain the co-expression of basal and luminal characteristics by basal-like carcinomas (Dontu et al., 2003; Foulkes, 2004; Melchor & Benitez, 2008). However, this hypothesis was challenged by the observation that pre-neoplastic *BRCA1*-mutant breast tissue contained an enlarged luminal progenitor population as compared to tissue from non-carriers (Lim et al., 2009).

Furthermore, gene expression prolife analysis revealed that basal-like tumors were transcriptionally most closely related to the luminal progenitor subpopulation of the normal mammary gland (Molyneux et al., 2010). Finally, *Brca1* deletion within the luminal progenitors of the mouse mammary gland was shown to generate tumors phenocopying basal-like breast cancers. In contrast, tumors generated from *Brca1* deficient basal cells did not resemble the histology of basal-like breast cancers (Molyneux et al., 2010). Moreover, while the hormone receptors ER and PR, whose presence or absence strongly determines subtype categorization, are markers of hormone sensing luminal cells in the healthy breast (Tarulli et al., 2015), they are not general luminal lineage markers. In fact, assessment of luminal lineage markers revealed that most invasive carcinomas of all subtypes express luminal-specific markers such as GATA3, CK8/18, Mucin-1 (MUC-1) and ZO-1 in a majority of their cells (Bell et al., 2003; Livasy et al., 2006; Matsukita et al., 2003; Rattan et al., 2012; Santagata et al., 2014; Shaoxian et al., 2017). Notably, while basal-like breast cancers partly display low expression of certain luminal markers such as GATA3 (Shaoxian et al., 2017), other luminal-specific markers such as CK8/18 are expressed reliably also in a majority of basal-like cancers (Livasy et al., 2006).

Building from these findings, the field by now largely agrees that the cell of origin for almost all breast cancers lies within the luminal progenitor population and that breast cancer subtype is largely determined by genetic events (Visvader & Stingl, 2014). However, with respect to the recently more and more apparent heterogeneity also within the luminal compartment as uncovered by the above-described single cell RNA-sequencing studies, it is conceivable that the differentiation stage of a luminal cell at the point of oncogenic transformation has influence on the resulting cancer subtype (Anstine & Keri, 2019). Of note, a possible explanation for the bipotent potential of luminal progenitor cells unraveled during basal-like breast cancer development has been provided by Dravis et al. (Dravis et al., 2018). Using ATAC-sequencing, they showed that the chromatin accessibility of adult luminal progenitors most closely resembles the profile of multipotent fetal mammary stem cells. This relation could explain how luminal cells can easily acquire basal-like features during malignant transformation while still being lineage-restricted during tissue homeostasis (Dravis et al., 2018).

In summary, the evidence strongly suggests that all major subtypes of breast cancer including basal-like tumors are generated by epithelial cells of the luminal progenitor subset.

2.3.3 To invade or not to invade

While certain tumors remain non-invasive and grow in the confined spaces of pre-exiting ducts (*in situ* carcinomas), others infiltrate the surrounding tissue, which is an initial step for metastasis formation and overall associated with a drastically worse prognosis for the patient (Van Seijen et al., 2019). Thereby, the comparison of *in situ* and invasive carcinomas is the ideal platform to identify potential specific determinants of invasion. Different evolutionary models have been proposed in the attempt to explain why certain tumors transition to the invasive state while others do not. The 'independent lineage' model proposes that invasive and *in situ* carcinomas stem from distinct cell lineages, which have no common genetic modifications (Sontag & Axelrod, 2005). Opposing this view, the 'evolutionary bottleneck' theory states that *in situ* and invasive carcinoma cells are directly related and that invasion is the result of one dominant clone (Cowell et al., 2013). Finally, a third theory has emerged recently. The 'multiclonal invasion' model postulates that invasive cells and pre-invasive cells are directly related, similar to the evolutionary bottleneck model. However, this model proposes that multiple clones invade concomitantly. The evolutionary bottleneck as well as the multiclonal invasion model are supported by gene expression analysis and genotyping studies that find largely overlapping profiles of invasive and *in situ* carcinomas (Hernandez et al., 2012; Ma et al., 2003; Petridis et al., 2016; Porter et al., 2003). Notably, invasive carcinomas are typically representative for a later stage of the disease, and thus additional genetic lesions have been found in invasive cells, including some characteristic mutations and copy-number alterations (Kim et al., 2015; Krøigård et al., 2015; Newburger et al., 2013; Yates et al., 2015). Nevertheless, the fact that those lesions are not universally detected in invasive carcinomas, indicates that they are not necessarily required for invasion. Recently, findings from topographic single cell-sequencing confirmed the high genetic similarity of *in situ* and invasive carcinomas as they showed that most aberrations found in invasive regions were already existent in cells of the pre-invasive state (Casasent et al., 2018). In addition, this study showed that *in situ*, a common aberrant progenitor results in a variety of aberrant sub-clones. At some stage those clones co-migrate across the basement membrane and invade, supporting the multiclonal invasion model (Casasent et al., 2018).

The inability to identify significant genetic drivers of the invasive transition has encouraged the theory that tumor cells invade for reasons largely unrelated to their genetic profile. Thereby, the role of the ECM during the invasion process has gained a lot of attention. Of note, one independent breast cancer risk factor that has been established by radiological studies, is breast tissue density. Thereby, high tissue density is representative for a

pronounced collagen deposition within the parenchyma whereas low density is characterized by a high proportion of fat (Boyd et al., 1992). In the respective mammographic studies, it was found that a high breast density, and therefore the abundance of collagen, correlates with increased overall breast cancer risk (Boyd et al., 1998; Byrne et al., 1995; Wolfe, 1976). Of note, also the risk for developing invasive breast cancer has been described to be directly affected by breast tissue density. In detail, women with high breast tissue density have a 3.5-fold greater risk of developing invasive breast cancer than women with low breast tissue density (Gill et al., 2006). Similarly, women with ductal carcinoma *in situ* had a 3.2-fold higher risk of the disease progressing to invasive carcinoma when they had a high breast tissue density as compared to women with low breast tissue density (Habel et al., 2004). These epidemiological observations imply that the collagen accumulations within the ECM are a driver of invasive breast cancer development. Of note, collagen type I is the most abundant collagen type in the central regions of invasive breast carcinomas (Lagace et al., 1985). Thereby, it is widely accepted that collagen type I drives invasion of established tumor cells (Egeblad et al., 2010). For instance, the overexpression of collagen type I in the mouse mammary gland *in vivo*, results in increased tumor formation and a more invasive phenotype with elevated metastatic capacity (Provenzano, Inman, Eliceiri, Knittel, et al., 2008).

Notably, during tissue homeostasis as well as in *in situ* carcinomas, neither luminal nor basal epithelial cells are in direct contact with the ECM. Instead, they are completely separated from the ECM by the surrounding basement membrane (Lopez-Garcia et al., 2010). Consequently, this membrane needs to be degraded or overcome by the cells to allow contact to the ECM during invasion. Notably, in *in vitro* culture models, the direct contact between epithelial cells and collagen type I can easily be recreated. Doing so, invasive behavior has been elicited from normal mammary epithelial cell lines and primary epithelial cells of the basal epithelial subset in collagen type I based *in vitro* culture systems, further corroborating the pro-invasive impact of collagen type I (Buchmann et al., 2021; Carey et al., 2017; Linnemann et al., 2015; Nguyen-Ngoc et al., 2012). Within the system as described by Linnemann et al. as well as Buchmann et al., the outgrowth of primary basal cells in collagen type I gels is thereby not merely an invasive process but accompanied by invasive branching morphogenesis (Buchmann et al., 2021; Linnemann et al., 2015). If a comparable invasive branching morphogenesis process could be triggered within luminal cells, this would explain how the network of luminal ducts as observed in NST carcinomas can be generated from luminal cells. However, in previous studies, when healthy luminal cells were put into collagen type I-containing matrices, they were not able to undergo an invasive process. Rather, when cultured in collagen type I gels, human luminal progenitor

cells grew out into non-invasive sphere- or budding-like structures (Linnemann et al., 2015) with partly reversed apical-basal polarity (Gudjonsson, Rønnov-Jessen, et al., 2002). In contrast, collagen-induced invasive behavior has been observed in luminal tumor explants (Cheung et al., 2013). However, in order to invade, the cells required the acquisition of basal characteristics (Cheung et al., 2013). Furthermore, while rudimental invasion could be recreated from these tumor explants, no invasive branching morphogenesis or duct formation occurred, which means that the arising structures did not resemble the complex *in vivo* morphology of most invasive carcinomas (Cheung et al., 2013).

In summary, it is known that during invasive cancer development, luminal cells invade the matrix while largely maintaining luminal marker expression. Furthermore, it appears that luminal cells invading the collagen type I-rich ECM are not necessarily genetically different from those luminal cells that remain within the ducts. Based on these considerations, it is thinkable that there is invasive potential also within non-invasive luminal cells which is unraveled by microenvironmental cues. Nevertheless, it has never been possible to recreate invasive behavior from normal luminal cells when exposed to collagen type I as potential driver of invasion.

2.4 Breast cancer *in vitro* models

For modelling various aspects of breast cancer, several *in vitro* systems have been established. Breast cancer cell lines are tumor-derived, immortal cells that have been widely used as a tool to study breast cancer biology and perform drug screenings(Iorio et al., 2016). Those cell lines bear striking advantages such as easy handling, high accessibility, and good modifiability wherefore a variety of 2-dimensional (2D) and 3-dimensional (3D) culture systems has been established using epithelial cell lines such as MCF-7, MDA-MB-231 or HeLa (Burdall et al., 2003; Debnath & Brugge, 2005; Ravid-Hermesh et al., 2018). Using those cell lines, certain assays have been developed that put a particular focus on recapitulating dynamic processes of tumorigenesis such as cellular invasion. Thereby, a high *in vivo* relevance is particularly promised by methods that allow investigation of invasion into a 3D matrix. However, in the so far established models, invasion is purely viewed as a migratory dissemination process (Debnath & Brugge, 2005; Moon et al., 2020; Ravid-Hermesh et al., 2018). Thereby, important features of cancer cell invasion are neglected such as the fact that invasion *in vivo* is often accompanied by a morphogenesis process for instance in the development of the ductal morphology of NST carcinomas. Therefore, 'invasive' morphologies resulting from cell line-based models are not fully comparable to the *in vivo* morphology of the disease, which calls the *in vivo* relevance of

the observed invasion process into question. Furthermore, an overall emerging problem of breast cancer cell lines is that the high selective pressure in culture makes cell lines prone to genotypic and phenotypic drift (Burdall et al., 2003). As a result, they often show low correlation to normal or cancerous mammary gland tissue (Ertel et al., 2006; Gillet et al., 2011; Santagata et al., 2014). Furthermore, cell lines such as MDA-MB-231 that show *in vitro* invasiveness are derived from patients in an extremely late stage of the disease, which calls into question whether they are suitable models in search for causes of invasive growth in the early stage of breast cancer or for screening of therapeutics that are supposed to prevent invasion early on. The insufficient transferability from cell line-based preclinical research to actual tumors is one of the suspected reasons for the vast number of drugs failing in clinical studies while showing promising results *in vitro* (Wilding & Bodmer, 2014). In detail, a recent study analyzing the success rate of clinical trials for more than 21,000 compounds over a period of 15 years found that only 13.8% of all compounds were successful in these trials. Moreover, merely 3.4% of compounds tested for cancer treatment, were successful throughout clinical studies, emphasizing a need for improved preclinical cancer models (Wong et al., 2019).

The inherent lack of transferability from breast cancer cell lines to *in vivo* human cancers has fueled the development of patient-derived organoid models. Organoids are miniature models of tissues that are derived from stem or progenitor cells with the ability to self-organize in 3D environments. Thereby, they bridge the modeling gap between *in vivo* organ function and 2D culture systems (Lancaster & Knoblich, 2014). While mammary organoids can be derived from immortalized cell lines or mouse mammary gland tissue, they can also be generated from freshly isolated human mammary epithelial cells, which promises the most accurate depiction of the human mammary gland (Lancaster & Knoblich, 2014).

Within the breast cancer research field, the Clevers group has pioneered with their work on human-derived breast cancer organoid cultures. In their recent work, Sachs et al. describe the long-term maintenance of mammary epithelium within fragments (Sachs et al., 2018). Thereby, the fragments were isolated from breast cancer patient tissue and cultured in Matrigel (Sachs et al., 2018). Notably, within these organoids, cells of the luminal subset were maintained and showed correct polarization over extended period of time while retaining their characteristic marker expression (Rosenbluth et al., 2020; Sachs et al., 2018). This is of interest as due to the role of luminal progenitor cells as cell of origin for most breast cancers, there is an increasing interest in meaningful *in vitro* systems for this subset. However, within the model presented by Sachs et al., the organoids were maintained as round shaped, cyst-like, partly discohesive structures (Sachs et al., 2018). Consequently, the organoid morphologies did not resemble the ductal phenotype of most

breast cancers. Moreover, the focus of the system described by Sachs et al. was to maintain cancerous cells *in vitro* rather than modelling of dynamic processes (Sachs et al., 2018). Therefore, dynamic processes such as invasion are not recapitulated in this model.

An additional challenge in studies that are based on cancerous material is the availability of the tissue. Neoadjuvant chemotherapy increasingly limits access to untreated tumor tissue that could be used to establish *in vitro* models. Moreover, especially in cancers that are diagnosed at an early stage, there is simply only a small quantity of cancerous tissue in the diseased breast (Balani et al., 2017).

In summary, in none of the established *in vitro* models, a tumor morphology was generated that resembled the ductal morphology of most invasive carcinomas. However, invasive processes can be monitored using breast cancer cell line-based approaches. Nevertheless, when it comes to advanced and presumably superior primary cell-based culture systems, there is, to this day, no system that allows monitoring the dynamics of cellular invasion.

2.5 Aim of this project

A variety of highly dynamic, complex processes contribute to tumorigenesis in the human mammary gland. As it is not possible to live-track tumor development within the human breast, our understanding of these processes is largely limited to histopathological sections that represent a 'frozen in time' picture of malignant development. For NST carcinomas as the most common type of invasive carcinomas, histologically, we can determine that luminal cells form ducts growing outside of the boundaries of the otherwise bilayered mammary gland epithelium.

Three-dimensional modelling systems have been developed to recapitulate dynamic processes of tumorigenesis such as the invasion of the ECM. However, those systems are based on cell lines which do not reflect the cellular state of mammary cells during early tumorigenesis and generated structures do not resemble the ductal morphology of NST carcinomas. Within the past years, using primary human tissue as a base for *in vitro* organoid systems has been established as the closest possible approximation to *in vivo* processes. However, it has so far merely been possible to model cellular dynamics of tumorigenesis with human material-based organoids. Therefore, in this work, I set out to develop a model based on primary human material that recreates the formation of invasively growing breast cancer. Thereby, I wished to recapitulate the actual ductal morphology of most invasive carcinomas as known from histopathological assessment.

Building on the fact that there are no known universal genetic determinants of invasively growing luminal cells that discriminate them from normal cells, I suspected that genetically normal luminal cells already possess invasive potential. I aimed to develop culture conditions that allow to uncover this invasive potential *in vitro*. Moreover, I set out to implement genetic modifications with the CRISPR-Cas9 system to model the impact of specific genetic aberrations on the morphology of invasively growing luminal cells.

3 Materials

3.1 Reduction mammoplasty donors

Donor	Age (years)	Parity
M16	17	0
M28	38	1
M35	48	0
M36	63	3
M42	41	0
M44	19	0
M45	54	1
M46	51	1
M47	55	2
M48	49	1
M50	58	1
M51	37	2

3.2 Reagents and chemicals

Reagent/chemical	Supplier (Catalog number)
2-Propanol	Roth (9866)
Acetic acid	Th.Geyer (2234)
Agarose powder	VWR (35-1020)
Albumin, fraction V (BSA)	Roth (CP84.2)
Aluminum potassium sulfate	Sigma (A-7167)
Ampicillin sodium salt	Roth (K029)
Aqua-Poly/Mount mounting medium	Polysciences (18606)
Carmine	Sigma (C-1022)
Collagen type I (rat-tail)	Corning (354236)
Deoxynucleotide (dNTP) solution mix	NEB (N0447)

DNAzol	Life Technologies (10503027)
Dimethyl-sulfoxide (DMSO)	Sigma (D8418)
DMEM/F12 without phenol red	Life technologies (21041025)
Donkey serum	GeneTex (GTX73205)
Ethanol	Merck (100983)
Ethylenediaminetetraacetic acid (EDTA)	Roth (X986)
Gel loading dye, orange (6X)	NEB (B7022S)
Glutamine	Sigma (G7513)
Hank's balanced salt solution (HBSS)	Life technologies (14025092)
LB Agar	Sigma (L2897)
LB medium	Roth (X964)
Marimastat	Sigma (444289)
Paraformaldehyde (PFA) 16% (v/v)	VWR International (43368.9M)
Phosphate buffered saline (PBS)	Thermo Fisher Scientific (10010-015)
S.O.C. medium	Thermo Fisher Scientific (15544034)
Sodium hydroxide solution (1M)	Sigma (28-3010)
SYBR™ safe DNA stain	Thermo Fisher Scientific (S33102)
TransIT-X2® dynamic delivery system	Mirus Bio LLC (MIR 6003)
TriDye™ 1 kb plus DNA ladder	NEB (N3270S)
Tris	Sigma (T1503)
Triton X-100	Sigma (T8787)
Trypsin-EDTA 0.05%	Life technologies (25300)
Trypsin-EDTA 0.25%	Life technologies (25200)
Trypsin neutralizing solution (TNS)	PromoCell (C-41120)
UltraPure™ BSA	Sigma (T8787)
UltraPure™ DNase/RNase-free distilled water (Milli-Q water)	Thermo Fisher Scientific (10977035)

3.3 Enzymes and growth factors

Enzyme/growth factor	**Supplier (Catalog number)**
Collagenase	Sigma (C9407)
Dispase	Stem Cell Technologies (07913)

DNase I	Sigma (11284932001)
DpnI, recombinant	NEB (R0176)
Gibson assembly® master mix	NEB (E2611)
Hyaluronidase	Sigma (H3506)
Insulin	Sigma (I6634)
Phusion® high-fidelity DNA polymerase	NEB (M0530)
Q5® high-fidelity DNA polymerase	NEB (M0491)
Taq DNA polymerase with thermopol® buffer	NEB (M0267)

3.4 Cell culture media components

Component	**Supplier (Catalog number)**
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Sigma (H4034)
A-83-01	Tocris (2939)
Advanced DMEM/F12	Life technologies (12634010)
B27 50x	Life technologies (17504044)
Epidermal growth factor (EGF), recombinant	PeproTech (AF-100-15)
Fetal calf serum (FCS)	Anprotec (AC-SM-0160)
Fibroblast growth factor 10 (FGF10), recombinant	PeproTech (100-26)
Fibroblast growth factor 7 (FGF7), recombinant	PeproTech (100-19)
Forskolin	Biomol (AG-CN2-0089)
GlutaMax 100x	Life technologies (35050061)
Mammary epithelial cell growth medium (MECGM)	PromoCell (C-21010)
MECGM supplement mix	PromoCell (C-39115)
N-acetylcysteine	Sigma (A9165)
Neuregulin 1 (NRG1), recombinant	PeproTech (100-03)
Nicotinamide	Sigma (N0636)
Noggin, recombinant	PeproTech (120-10C)
Penicillin/streptomycin	Invitrogen (5000956)
R-Spondin-3, recombinant	PeproTech (120-44)

SB202190	Sigma (S7067)
Y-27632	Biomol (Y0110)

3.5 Cell culture media compositions

All media shown here are depicted in their 1x concentration. Media were sterile filtered, if not all components used for medium preparation, were sterile.

3.5.1 BCOM, BLOM and Basic medium

Component	BCOM	BLOM	Basic medium	Concentration
A-83-01	+	+	-	500 nM
Advanced DMEM/F12	+	+	+	-
B27 50x	+	+	+	1x
EGF, recombinant	+	+	+	5 ng/ml
FCS	-	+	+	0.5%
FGF10, recombinant	+	+	-	20 ng/ml
FGF7, recombinant	+	+	-	5 ng/ml
GlutaMax 100x	+	+	+	1x
HEPES	+	+	+	10 mM
N-acetylcysteine	+	+	+	1.25 mM
Nicotinamide	+	+	+	5 mM
Noggin, recombinant	+	+	-	100 ng/ml
NRG1, recombinant	+	+	+	5 nM
Penicillin/streptomycin	+	+	+	100 µg/ml
R-Spondin-3	+	+	-	250 ng/ml
SB202190	+	+	+	500 nM
Y-27632	+	+	-	5 µM

3.5.2 Linnemann et al. seeding and maintenance medium

Component	Seeding	Maintenance	Concentration

FCS	+	-	0.5%
Forskolin	+	+	10 µM
MECGM	+	+	-
MECGM supplement mix	+	+	10 mM
Penicillin/streptomycin	+	+	100 µg/ml
Y-27632	+	-	3 µM

3.6 Composition of buffers, solutions and media (not for cell culture)

All components and their respective suppliers are listed in 3.2 and 3.4.

Solution/ buffer/ medium	**Component and concentration**
Blocking solution	BSA 10% (m/m)
	Donkey serum 10% (v/v)
	PBS
Carmine staining solution	Carmine 2.0% (m/m)
	Aluminum potassium sulfate 5.0% (m/m)
	1 Thymol crystal (per 500 ml)
	Milli-Q water
DAPI staining solution	DAPI 167 ng/ml
	PBS
FACS buffer	BSA 10% (m/m)
	PBS
Freezing medium	FCS 50% (v/v)
	WB 40% (v/v)
	DMSO 10% (v/v)
HEPES solution	HEPES 0.1 M
	Milli-Q water
HF buffer	HEPES 10 mM
	Penicillin/ streptomycin 100 U/ml
	FCS 10% (v/v)

	HBSS
Neutralizing solution	HEPES 500 mM
	11x PBS
Permeabilization solution	Triton X-100 0.2% (v/v)
	PBS
Tris-acetate-EDTA (TAE) buffer	Tris 0.04 M
	Acetic acid 0.02 M
	EDTA 1 mM
	Milli-Q water
Tissue digestion buffer	Insulin 1 µg/ml
	Collagenase 300 U/ml
	Hyaluronidase 100 U/ml
	WB
Trypsin 0.15%	Trypsin-EDTA 0.05% 50% (v/v)
	Trypsin-EDTA 0.25% 50% (v/v)
Washing buffer for breast tissue digestion (WB)	HEPES 10 mM
	Penicillin/ streptomycin 100 U/ml
	Glutamine 2 mM
	DMEM/F12 without phenol red

3.7 Consumables

Compound	**Supplier**
2-and 8-well dishes, uncoated	Ibidi
Cell culture dishes (6-cm)	Nunc
Cell strainer (40 µm)	Corning
Conical centrifugation tube (15 ml and 50 ml)	Corning
CryoTubes®	Nunc
Easy grip polystyrene storage bottle (500 ml)	Corning
Filter system for sterile filtration	Sigma
Microscopy slides (cut edges, matt strip)	Thermo fisher scientific
Microscopy cover glass 22 mm x 40 mm	VWR international

Neubauer counting chambers	Herenz
Nitril® nextgen® gloves	Meditrade
Optically clear bottom dishes (96-well)	Perkin Elmer
Pipette tips (filtered)	Starlab
Pipette tips (unfiltered)	Starlab
Polystyrene cell culture plates (6-, 24- and 48-well)	Becton Dickinson
Polystyrene petri dish (for agar plates)	Greiner Bio-One
Safe-lock tubes 1.5 ml PCR clean	Eppendorf
Safe-lock tubes 2.0 ml	Eppendorf
Scalpels	VWR international
Sterile serological pipets (2, 5, 10, 25, 50 ml)	Greiner Bio-One
Test tube with cell strainer cap (FACS tube)	Corning
XL10-Gold bacteria	Agilent technologies

3.8 Plasmids

Plasmid	Description	Source
STAgR_Neo	Vector Backbone for STAgR cloning	(Addgene, 102992)* Provided by Christopher Breunig and Stefan Stricker
STAgR_gRNA Scaffold_hU6	STAgR cloning template for generation of STAgR fragments	(Addgene, 102840)* Provided by Christopher Breunig and Stefan Stricker
pSpCas9(BB)-2A-GFP (PX458)	Cas9 expressing plasmid with connected 2A-GFP for selection at the FACS/ suitable for transfection	(Addgene, 48138)* Provided by Christopher Breunig and Stefan Stricker

As described by Breunig et al. 2018 (Breunig et al., 2018)

3.9 Antibodies and fluorescent dyes

3.9.1 Antibodies and cellular stains for FACS

Epitope (Clone)	Conjugation	Host	Volume (µl)*	Supplier (Catalog number)
7-AAD	-	-	4	BD (559925)

CD10 (HI10A)	APC	Mouse	5	Biozol (BLD-312210)
CD31 (WM59)	V450	Mouse	1	BD (561653)
CD31 (WM59)	V450	Mouse	1	BD (561653)
CD326/EpCAM (VU-D19)	FITC	Mouse	10	Biozol (GTX79849)
CD45 (HI30)	V450	Mouse	1	BD (560368)
CD45 (HI30)	V450	Mouse	1	BD (560368)
CD49f (GOH3)	PE	Rat	5	BD (555736)
SYTOX™ blue	-	-	1/1,000**	Life technologies (S34857)

*For staining 1x 10^6 cells **Dilution based on sample volume

3.9.2 Primary antibodies and fluorescent dyes for immunofluorescence and blocking

Epitope (Clone)	Conjugation	Host	Dilution	Supplier (Catalog number)
4′,6-Diamidin-2-phenylindol (DAPI)	-	-	167 ng/ml	Sigma (D9542)
αSMA	-	Rabbit	1/100	Abcam (ab5694)
Atto 488 nm	-	-	-	Sigma (41051)
CK8/18 (5D3)	-	Mouse	1/250	Dianova (DLN-10750)
CD49f (GOH3)	-	Rat	1/100	Santa cruz (sc-19622)
E-cadherin (24E10)	Alexa 488	Rabbit	1/50	CST (3199)
E-cadherin (HECD1)	-	Mouse	1/25	Abcam (ab1416)
GATA3 (L50-823)	-	Mouse	1/250	Biocare medical (CM405)
Ki-67	-	Rabbit	1/300	Abcam (ab15580)
Laminin	-	Rabbit	1/100	Sigma (L9393)
MUC-1	-	Rabbit	1/100	Abcam (ab15481)
P63 (EPR5701)	-	Rabbit	1/300	Abcam (ab124762)
Phalloidin	Atto 647N	-	1/250	Sigma (65906)
SirDNA	SPY650	-	1/1,000	Spirochrome (SC501)
Vimentin (V9)	-	Mouse	1/100	Abnova (MAB3578)
ZO-1 (1A12)	-	Mouse	1/100	Life technologies (339100)

3.9.3 Primary antibodies used for immunohistochemistry

Epitope (Clone)	Conjugation	Host	Dilution	Supplier (Catalog number)
E-cadherin (EP700Y)	-	Rabbit	1:100	Cell marque (246R-16)
GATA3 (L50-823)	-	Mouse	1:200	Biocare (CM405B)
P63 (SFI-6)	-	Mouse	1:50	DCS (PI006C01)

3.9.4 Secondary antibodies used for immunofluorescence

Host/Isotope	Conjugation	Species reactivity	Dilution	Supplier (Catalog number)
Donkey/IgG	Alexa 488	Mouse	1/250	Life technologies (A-21202)
Donkey/IgG	Alexa 594	Rabbit	1/250	Life technologies (A-21207)
Donkey/IgG	Alexa 647	Mouse	1/250	Life technologies (A-31571)
Donkey/IgG	Alexa 488	Rat	1/250	Life technologies (A-21208)

3.10 Kits

Kit	Supplier (Catalog number)
Monarch gel extraction kit	NEB (T1020)
Mix2Seq kit	Eurofins (3094-000MSK)
Qiagen plasmid midi kit	Qiagen (12143)

3.11 Instruments

Instrument	Manufacturer
ChemiDoc™ MP imaging system	Bio-rad laboratories
Benchmark XT system	Ventana medical systems
Blue light table, LED transilluminator	Serva electrophoresis
FACSAria IIIu	BD
FLUOVIEW FV1200 inverted confocal laser scanning microscope	Olympus

HERAcell® 240i incubator	Thermo fisher scientific
Heraeus™ Fresco™ 21	Thermo fisher scientific
Intelli-mixer RM-2L LFT	Labortechnik GmbH
Leica DM IL LED microscope	Leica microsystems
Leica SP8 lighting confocal microscope	Leica microsystems
Mastercycler® nexus gradient	Eppendorf
NanoDrop® ND 1000 spectrophotometer	Thermo fisher scientific
On-stage incubation system	Ibidi
Safe 2020 Class II biological safety cabinet	Thermo fisher scientific
ThermoMixer® C 1.5 ml	Eppendorf
Zeiss axio imager.M2 with colibri 7	Carl zeiss microscopy

3.12 Software

Software	**Manufacturer**
Benchling	Benchling
FACS Diva™ v6.1.3	BD
FlowJo® 10.5.0	FlowJo LLC
FV10-ASW software v4.2.b	Olympus
Gimp 2.10.24	GIMP development team
GraphPad prism 8	GraphPad software inc.
ImageJ 1.52	U.S. national institute of health
Leica application suite 4.2	Leica microsystems
Matlab R2020b	MathWorks
MS office 365	Microsoft
ZEN 2.3 pro	Carl zeiss microscopy

4 Methods

4.1 Cell biological methods

The cells were cultivated under sterile conditions and whenever possible, work was performed under a hood with laminar air flow. Cells were maintained in an incubator at 37°C with 95% humidity and CO_2 levels at 5%. Moreover, O_2 levels were at 3%. All centrifugation steps were performed at 300 × g, at 4 °C for 5 minutes.

4.1.1 Isolation of human mammary epithelial cells

Breast tissue was obtained from healthy women that underwent reduction mammoplasty for cosmetic reasons. The tissue was provided by the Nymphenburg Clinic for Plastic and Aesthetic surgery (Böcklinstraße 1, 80638 München). Thereby, tissue collection was performed in accordance with the regulations of the ethics committee of the Ludwig-Maximilian University, Munich, Germany (proposal 19-989). Information on age and parity of all donors used in this study can be found in 3.1.

Fresh tissue was stored on ice for transportation and processed immediately. As first step for epithelial cell isolation, a large fraction of the adipose tissue was removed by carefully scraping it off with scalpels. Next, the remaining tissue was minced into pieces of approximately 1 mm^3 using scalpels.

Once all tissue was processed accordingly, approximately 5 ml of tissue pieces were filled into 50 ml falcons together with 20 ml of tissue digestion buffer. Next, insulin (1 µg/ml), hyaluronidase (100 U/ml) and collagenase (300 U/ml) were added, and the falcons were rotated at 6-10 rpm on an intelli-mixer at 37°C for 16-18 hrs to facilitate enzymatic digest.

After digestion, the falcons were filled with washing buffer and centrifuged to pellet the tissue. Next, the fatty supernatant was aspirated and discarded. The pellet was resuspended in 25 ml washing buffer, transferred into a fresh tube and centrifugation was repeated.

Finally, the cells were either frozen down as fragments or single cells. For freezing as single cells, fragments were dissociated into single cells in 0.15% pre-warmed trypsin-EDTA while mixing by slowly pipetting for 5 minutes. Next, cold HF buffer was added, cells were pelleted by centrifugation and 5 mg/ml pre-warmed dispase was added for 3 minutes for further dissociation. Cells were counted and spun down. Next, cells were resuspended in

appropriate volume of freezing medium at a density of approximately 5x10^6 cells/ml, aliquoted, frozen down at -80°C and transferred to liquid nitrogen for long-term storage. For freezing fragments, a small aliquot of cells was dissociated to single cells as described above to estimate total cell count. Finally, the remaining fragments were resuspended in freezing medium and frozen down in aliquots as described above.

4.1.2 Thawing of primary human mammary epithelial cells

The required number of vials with primary human mammary epithelial cells in freezing was placed in a pre-heated water bath (37°C). After thawing, the cells were immediately transferred into 50 ml falcons containing PBS. Cells frozen as single cells were filtered through a 40 µm strainer to remove residual aggregates. Cells frozen as fragments were dissociated into single cells in 0.15% pre-warmed trypsin-EDTA while mixing by slowly pipetting for 5 minutes. Cold HF buffer or trypsin neutralizing solution (TNS) was added to stop the reaction. If visible DNA was released from dead cells, 1 mg/ml DNAse I was added for 1-2 minutes before the cells were filtered through a 40 µm cell strainer. Next, cells were centrifuged, and pelleted cells were used for FACS as described in 4.2.1.

4.1.3 Cell counting

For counting, pelleted cells were resuspended in a defined volume. Three 10 µl aliquot of single cell suspension were taken and injected into a Neubauer chamber. Counting of three squares in each chamber was performed using a light microscope (Leica DM IL LED, 10x objective). Thereby, living cells were identified based on their overall morphology and granularity. The total number of living cells was calculated according to the following formula:

$$n(total\ cell\ count) = x\ (mean\ of\ counted\ squares) * V\ (sample\ volume) * 10\text{^}4$$

4.1.4 Collagen coating

Polystyrene culture plates were coated with rat-tail collagen type I in a concentration of 5 µg/cm². Collagen type I batches differing in concentration were diluted in PBS to create the coating solution with a concentration of 38 µg/ml. Coating solution was added to the respective wells and dishes were incubated for 2-3 hrs at 37°C or overnight at 4°C. Subsequently, the dishes were washed twice with PBS and either used directly or stored for up to 2 weeks at 4°C.

4.1.5 2D culture of human mammary epithelial cells

Single cell suspensions containing the desired amount of fresh human mammary epithelial cells in BCOM medium were seeded into collagen type I-coated polystyrene cell culture plates. Upon reaching a confluency of 80-90%, cells were split at a ratio of 1:2 to 1:12, depending on the experiment. In detail, medium was removed, and cells were washed with PBS once. Next, 0.15% trypsin-EDTA was added, and cells were incubated at 37°C for detachment. For neutralization, TNS was added with three times the volume of trypsin. In addition, cell aggregates were dissociated by pipetting up and down. Optionally, 10 µl of cell suspension was used to determine cell count (as described in 4.1.3). Next, the required number of cells was transferred to a 50 ml tube and centrifuged. The pellet was resuspended in BCOM medium and seeded into a new collagen-coated 2D culture plate or into a collagen gel for 3D culture as described in section 4.1.6.

4.1.6 3D culture of human mammary epithelial cells

For preparation of 3D collagen type I gels, FACS-sorted single cells were resuspended in a small volume of medium. In case of BLOM gels, the cells were resuspended in BLOM foundation. For culture conditions according to Linnemann et al., cells were resuspended in seeding medium that lacked Y-27632, Forskolin and FCS. Next, cells were counted. The desired number of cells was mixed with medium (BLOM foundation or seeding medium lacking Y-27632, Forskolin and FCS), rat-tail collagen type I and neutralizing solution. Thereby, the final collagen concentration was 1.3 mg/ml. Neutralizing solution was added in a ratio of 1:10 relative to the volume of collagen. The pH of the polymerization solution was probed using indicator strips. If the pH was too acidic (pH < 7.4), 1M sodium hydroxide solution was added up to a volume of 10 µl per 1 ml polymerization solution. The gels were put to 37°C for 1 hr for polymerization. Upon polymerization, medium was added carefully. Thereby, the volume of medium was equivalent to the volume of the gel. To compensate for the volume of the gel, all components of the medium that had not been added to the collagen gel, were added in a 2x concentration. Finally, gels were encircled with a pipet tip for detachment. For Matrigel experiments, luminal progenitor cells in a BLOM suspension were mixed with growth factor reduced Matrigel, all other steps were performed analogously to culture in collagen type I gels.

The first medium exchange was performed 5 days after seeding with 1x concentrated medium. In case of gels cultured according to Linnemann et al., from now on maintenance

medium was used that lacked FCS and Y-27632. All consecutive medium exchanges were performed every 2-3 days with 1x concentrated medium.

Outgrowing structures were categorized as branched structures if they had more than two branching points. Structures were categorized as sticks if they had less than two branching points and were at least twice as long as they were wide. The remaining structures were categorized as spheres. For determination of branched structure sizes and of maximum structure width, pictures were taken on a Leica DM IL LED microscope and ImageJ was used for image analysis. For assessment of branched structure complexity, the longest duct was determined as structure body and the hierarchy of all remaining ducts relative to the structure body was classified.

4.1.7 Extreme limiting dilution analysis

Extreme limiting dilution analysis (ELDA) was used to determine branched structure-forming units (B-SFUs). To this end, gels were prepared that contained freshly sorted luminal progenitor cells in degrading dilutions. The gels were prepared in 24-Well plates and had a volume of 400 µl. Thereby, six different cell concentrations were tested, and six gels were prepared per dose. Gels were counted as positive if they contained at least one branched structure. Evaluation of B-SFU and calculation of 95% confidence intervals was performed using ELDA software as described previously (Y. Hu & Smyth, 2009).

4.1.8 Inhibitor treatment

For investigation of collagen degradation via matrix-metalloproteinases (MMPs), the MMP inhibitor Marimastat was added to the cell culture medium in a 10 µM concentration at day 0 or day 7 as indicated in the respective figure. In live-cell imaging experiments, 10 µM Marimastat was added during elongation phase prior to imaging. For blocking of E-cadherin function, HECD1, an E-cadherin blocking-antibody, was added in a dilution of 1:25 on day 5 of 3D culture. The myosin II inhibitor para-amino blebbistatin (PAB) was used in analogy to the medium component Y-27632 and therefore added to 3D cultures on day 0. The final concentration of PAB was 10 µM.

4.1.9 CRISPR-Cas9

To generate a sufficient number of cells for gene editing, freshly sorted luminal progenitor cells were expanded in 2D culture as described in chapter 4.1.5. Once the cells had grown

confluent, they were split and seeded in a density of 1.5x10^4 cells/cm^2. For gene editing, the cells were transiently transfected 6 hrs after splitting with the TransIT-X2 transfection reagent according to the manufacture's recommendations. Thereby, a GFP-expressing Cas9 (pSpCas9(BB)-2A-GFP) and a gRNA-containing plasmid (STAgR_Neo with inserted gRNAs) were co-transfected in a molar ratio of 1:3. For controls, cells were transfected with the Cas9-GFP plasmid and a gRNA scaffold plasmid (STAgR_gRNAScaffold_hU6) that was void of the respective gRNAs. After 12 to 16 hrs, the medium was exchanged. Finally, 48 hrs after splitting the cells, they were harvested for FACS. Based on the temporary expression of GFP within Cas9-GFP positive cells, successfully transfected cells were collected (see section 4.2.2). To validate the successful induction of deletions at the targeted locus, DNA of GFP positive cells was collected and polymerase chain reaction (PCR) on the target locus was performed (see section 4.5.11). To analyze implications of the respective deletion on branched structure formation, freshly sorted GFP-positive cells were seeded into 3D culture in BLOM medium as described in section 4.1.6. To assess editing efficiency on the protein level and to differentiate between knock-out and wild-type structures, immunofluorescence was performed on the outgrowing structures (see section 4.3.3). For design of gRNAs, cloning and PCR see section 4.5.2.

4.2 Flow cytometry

4.2.1 Sorting of uncultured human mammary epithelial cells

For analysis and sorting of HMEC populations, surface staining with fluorescently conjugated antibodies was performed. In detail, frozen human mammary epithelial cells were thawed as described in section 4.1.2. Next, cells were taken up in FACS buffer at a concentration of 1x10^7 cells/ml. The sample for the actual FACS sort was stained with the antibodies CD45-V450, CD31-PB, EpCAM-FITC, CD49f-PE and CD10-APC (see section 3.9.1 for more information on the antibodies). Additionally, for each antibody, a small volume of cells (usually 50 µl) was stained with this particular antibody as single stain controls for setting the gates. The cells were incubated on ice for 40 minutes while being protected from light. Next, the cells were washed with PBS and resuspended in FACS buffer. Thereby, the sample was resuspended in appropriate volume to achieve a final concentration of 1x10^7 cells/ml while single stain controls were resuspended in 250 µl to ensure sufficient volume for analysis. Next, 7-AAD was added to one of the controls and to the sample for dead-cell exclusion. After an additional incubation for 15 minutes on ice (protected from light), the cells were ready for analysis/sorting. Forward scatter (FSC) and side scatter (SSC) were used to exclude cell debris, cell clumps and doublets. Three gates were used for this: SSC-

A/FSC-A followed by FSC-A/FSC-W and SSC-A/SSC-W. Next, dead cells were excluded using 7-AAD. Moreover, endothelial, and hematopoietic cells were excluded based on their expression of CD31 and CD45, respectively. 7-AAD^{-}/CD31^{-}/CD45^{-} cells were divided into four populations based on their expression of CD49f and EpCAM. These four populations were stroma (CD49f^{low}/EpCAM^{low}), mature luminal (CD49f^{low}/EpCAM^{high}), luminal progenitor (CD49f^{high}/EpCAM^{high}), and basal cells (CD49f^{high}/EpCAM^{low}). The basal population was further characterized by expression of CD10. The sorted cells were collected in FACS tubes containing FACS buffer using the purity mode. After sorting, the cells were subjected to a re-analysis to ensure the purity of the sorted populations. Only sorts with a purity of at least 99.5% were used for experiments. Tubes were filled with PBS and cells were spun down and resuspended. In case of BLOM gels, the cells were resuspended in BLOM foundation, for culture conditions according to Linnemann et al., cells were resuspended in seeding medium that lacked Y-27632, Forskolin and FCS. To ensure proper cell count with respect to varying recovery rates after the sort, a small aliquot of cells was used for counting as described in section 4.1.3.

4.2.2 Sorting of GFP positive cells

For analysis or sorting of previously 2D-cultured cells that were transfected with Cas9-GFP plasmid, cells were harvested and brought to a concentration of 1x10^6/ml in FACS buffer. Next, cells were stained for 5 minutes with the SYTOX^{TM} Blue in a 1 µM concentration to distinguish between viable and dead cells. During analysis or sorting, cell debris, cell clumps and doublets were excluded using SSC and FSC as described above.

For sorting of GFP-positive cells, respective populations were guided into tubes filled with 1 ml of FACS buffer using the purity mode. No re-analysis was performed as sorted populations were typically small. Next, the tubes were filled up with PBS, spun down and resuspended in a small volume of medium or FACS buffer for counting. The cells were brought to the desired concentration and seeded for further experiments.

4.3 Staining and imaging

4.3.1 Immunohistochemistry on tissue sections

Immunohistochemistry (IHC) on tissue sections was performed by Moritz Jesinghaus (Institute of Pathology, Technical University of Munich). Thereby, paraffin section with a thickness of 3 µM were used. For IHC, the Benchmark XT system, an automated slide

preparation system and associated reagents and buffers supplied from Ventana Medical Systems were used. All antibodies used as well as respective dilutions are listed in 3.9.3.

4.3.2 2D immunofluorescence

Cells for 2D immunofluorescence were grown on collagen-coated 96-well dishes with optically clear bottom. For fixation, medium was removed from cells with the desired confluency and wells were rinsed with PBS three times. Next, cells were fixated with 4% PFA for 15 minutes. Of note, plates were slowly shaking on an orbital shaker at room temperature during all incubation steps unless indicated otherwise. Subsequently, the wells were again rinsed three times with PBS and cells were permeabilized with permeabilization solution for 2 minutes to allow intracellular stainings. The wells were again rinsed with PBS before blocking solution was added to block potential non-specific binding sites. Blocking was performed for 1 hr while shaking. After washing with PBS three times, 75 µl of primary antibody solution in 0.1% BSA was added to the wells. Of note, all antibodies and their respective dilutions are listed in 3.9.2. Primary antibody was incubated overnight at 4°C while shaking. The next day, wells were washed three times with PBS, and 50 µl of secondary antibody solution in 0.1% BSA, was added and incubated for 2-3 hrs while shaking. Afterwards, wells were washed three times with PBS and DAPI staining solution was added for 2 minutes to stain nuclei. Next, the wells were washed three times with PBS and three times with deionized water. Wells were mounted by addition of one drop Aqua-Poly/Mount mounting medium and subsequently dried for 2-3 days until the mounting medium had solidified completely. Plates were either imaged immediately or stored under light protection at 4°C for up to 4 weeks.

4.3.3 3D immunofluorescence

During staining of 3D structures grown in collagen gels, all incubation and washing steps were performed on an orbital shaker at room temperature unless stated otherwise. All PBS washing steps were always performed for 10 minutes on an orbital shaker. Prior to fixation, medium was removed from the gels and the gels were washed with PBS once. Next, cells were fixated with 4% PFA for 15 minutes and subsequently washed four times with PBS. Cells were permeabilized with permeabilization solution for 10 minutes and washed with PBS, followed by blocking with blocking solution overnight at 4°C on an orbital shaker. After another washing step with PBS, primary antibodies in 0.1% BSA were added and incubated overnight at 4°C. Of note, all antibodies and their respective dilutions are listed in 3.9.2. The next day, gels were washed three times with PBS before secondary antibody in 0.1% BSA

was added for 2-3 hrs. From this point on, the experiment was proceeded in the dark (wrapped in aluminum foil). Two PBS washing steps were performed and DAPI staining solution was added for 2 minutes to stain nuclei. Next, gels were washed with PBS three times and subsequently two times with deionized water for 5 minutes. Finally, the collagen gels were transferred to a microscopy slide. Here, remaining water at the edges of the gels was removed carefully using a dust-free tissue. Next, 2-3 drops of Aqua-Poly/Mount mounting medium were added onto the gels. A microscopy cover glass was placed on top of the gels while avoiding the inclusion of bubbles. Mounted gels were dried at room temperature for 2-3 days and stored at 4°C for up to 4 weeks or at -20°C for up to 6 months.

4.3.4 Carmine staining

For carmine staining of 3D structures, gels were washed with PBS once, fixated with 4% PFA for 15 minutes and subsequently washed four times with PBS. Next, carmine staining solution was added and incubated overnight at 4°C. Prior to imaging, three PBS washing steps were performed.

4.3.5 Collagen labelling and imaging

Collagen labelling and imaging were performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich). For visualization of the collagen structure, collagen was fluorescently labelled using Atto 488 as described previously (Geraldo et al., 2013). Thereby, collagen was dialyzed to a pH of 7.0 at 4°C. In order to facilitate conjugation, collagen was incubated with Atto 488 overnight at 4°C. Next, non-bound Atto 488 was removed by further dialysis for 8 hrs. Finally, another dialysis step was performed with the addition of acid in order to prevent unwanted polymerization. A SP8 lightning confocal microscope was used for fluorescence microscopy and confocal reflection. Thereby, the collagen network was visualized through a HC PL APO 40x/1.10 water immersion objective upon illumination with light at 488 nm. Next, emitted fluorescent light was collected at 510 nm and 550 nm while reflected light was collected at 488 nm. Thereby, the distance between single slides was set to 1 µm. The degree of collagen fiber alignment was calculated as described previously by Buchmann et al. (Buchmann et al., 2021).

4.3.6 Imaging of fixated cells

Immunofluorescent images of 2D cells were acquired on a wide-field fluorescence microscope, the axio imager.M2 using 10x, 20x and 40x objectives and the Zeiss ZEN 2.3

pro software. For imaging of fluorescently stained 3D structures, the inverted FluoView FV1200 confocal laser scanning microscope was used. This confocal microscope allowed to selectively image single focal planes of the gels, thereby avoiding out of focus fluorescence. The microscope contained four laser lines (405, 488, 543, and 633 nm) as well as several objectives (20x, 40x, 60x, UPLSAPO). Carmine-stained structures were imaged on a Leica DM IL LED microscope using 10x, 20x and 40x objectives.

2D and 3D immunofluorescence as well as carmine images were processed with Gimp 2.8.22 and ImageJ software to adjust brightness across the entire image field or to retrospectively alter the color of a certain staining.

4.3.7 Live-cell imaging

Live-cell imaging was performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich). A Leica lightning confocal microscope was used for imaging. Maintaining the cells in a viable state was facilitated with an on-stage incubation system was used for regulating CO_2, O_2, humidity and temperature for maintaining the cells in a viable state during imaging. Imaging was performed during organoid elongation phase (mainly day 7 to day 9). Thereby, images were taken every 10 minutes. If required, organoids were incubated with sirDNA (10 µM) for 3 hrs prior to the measurement to facilitate nuclear labeling. The stain was excited at 633 nm and the fluorescent signal was collected at 674 nm with a HCX PL APO 10x/0.40 CS dry objective. For visualization of bead displacements and deformation fields, ImageJ 1.48v was used. In detail, the nuclear fluorescent signal at the first-time step was summed up along the z-axis allowing for visualization of the organoid's shape. Next, the fluorescent signal of the fluorescent beads was summed up along the z-axis and consecutively summed up along all subsequent steps. Finally, the two calculated pictures were merged into one image. Bead movement was calculated using Matlab R2020b. To do so, background was reduced by covering the fluorescent signal of the beads with an intensity threshold. Next, the center of an interpolated intensity grid was used to define bead positions. Finally, bead coordinates of each individual bead were matched in three consecutive images to calculate the tracks of individual beads. To guarantee correct tracking, beads touching the boundaries as well as beads in direct proximity to other beads, were excluded.

4.4 Methods working with bacteria

For all methods working with chemically competent bacteria, non-autoclaved tips as well as nuclease free tubes were used. All steps were performed at room temperature, unless stated otherwise.

4.4.1 Bacterial transformation

The bacterial strain XL10-Gold was used or transformation. The bacteria were stored at -80°C and were thawed on ice prior to transformation. Plasmid DNA (1 ml, typically less than 1 µg) was put into a 1.5 ml Eppendorf tube. A small aliquot of thawed bacteria (25 µl) was added to the Eppendorf tube and incubated on ice for 10 minutes. Next, the solution was subjected to a heat shock by placing the tube at 42°C (water bath) for 30 seconds. Next, the tube was quickly placed on ice for 2 minutes. 200 µl of S.O.C. medium was added, and the tube was incubated in a ThermoMixer at 37°C for 1 hr while shaking at 1,000 rpm. Afterwards, an aliquot of bacteria was placed on LB agar plates containing Ampicillin (100 µg/ml) for selection of successfully transformed bacteria. Plates were incubated overnight at 37°C or over the weekend at room temperature until colonies had formed. Single colonies were multiplied, and plasmid DNA was isolated as descried below in section 4.5.1.

4.4.2 Bacterial plasmid DNA isolation

Single colonies were picked from the agar plates using a pipette tip. For mini preparation, a single colony was transferred into 3 ml of LB medium containing Ampicillin (100 µg/ml) for selection. In case of midi preparations, 100 ml of LB medium with Ampicillin were inoculated with a single colony. The inoculated cultures were expanded overnight by vigorous shaking (120 rpm) at 37°C. After expansion of the bacteria for midi culture, plasmid DNA was isolated using the Qiagen plasmid midi kit as proposed by the manufacturers. Afterwards, the DNA was resuspended in 150-200 µl Milli-Q water. In case of mini preparation, plasmid DNA was isolated using the buffers supplied by the Qiagen midi kit. In brief, 2 ml of the expanded bacterial culture was transferred into an Eppendorf tube. Next, the tube was spun down at 13,000 rpm for 3 minutes. Supernatant was removed and the bacteria were resuspended in 250 µl buffer P1 from the Qiagen midi kit. Next, 250 µl of buffer P2 was added and the tube was inverted 10x for lysis of bacteria. Finally, 300 µl of buffer P3 was added and the sample was inverted 10x for neutralization. The sample was spun down for 20 minutes at 13,000 rpm at 4°C. After this step, the plasmid was contained in the supernatant. Therefore, the supernatant was transferred to a fresh 1.5 ml Eppendorf tube.

For precipitation, 600 µl of isopropanol (100%) was added and the tube was inverted 5x. Finally, the tube was spun down for 25 minutes at 13,000 rpm at 4°C to pellet DNA. The supernatant was removed, and the plasmid DNA was washed 2x with 70% ethanol (500 µl) for 5 min at 13,000 rpm at 4°C. After the second washing step, as much as possible of the supernatant was removed and the tubes were placed in the pre-heated ThermoMixer at 85°C (no shaking) in order to evaporate remaining ethanol. After mini preparation, plasmid DNA was resuspended in 30 µl Milli-Q water. The resulting DNA concentration of midi- and mini preparation was measured on a NanoDrop spectrophotometer by measuring the absorption at λ=230 nm, λ=260 nm and λ=280 nm. For further experiments, only DNA with purity greater than 1.8 (based on 260/280 nm ratio) was used.

4.5 Methods working with DNA

All methods working with DNA were performed at room temperature with non-autoclaved tips and nuclease free Eppendorf tubes.

4.5.1 Extraction of genomic DNA from epithelial cells

For extraction of genomic DNA (gDNA), cells were grown in a 12-well plate until they were at least 75% confluent. Next, cells were washed once with PBS before 250 µl DNAzol reagent were added per well. DNA isolation was performed either immediately or plates were sealed with parafilm and stored overnight at 4°C. DNA was precipitated by addition of 125 µl of ultra-pure ethanol (100%) and subsequent, careful shaking. After a short incubation period, DNA became visible as a white precipitate. A 1.5 ml Eppendorf tube was prepared by filling it with 500 µl of 70% ethanol. The DNA was taken up with a pipet tip and transferred into the ethanol-containing Eppendorf tube. After centrifugation for 1 minute at 13,000 rpm, the supernatant was removed, and DNA was washed twice with 70% ethanol. Following the final centrifugation step, the remaining ethanol was removed, and the tube was placed in the pre-heated ThermoMixer at 85°C (no shaking) in order to evaporate remaining ethanol. Finally, the DNA was dissolved in 8 mM NaOH (typically 30 µl) and the pH was adjusted to 7.5 with a 0.1 M HEPES solution.

4.5.2 STAgR cloning

STAgR cloning as described by Breunig et al. (Breunig et al., 2018) was performed to generate plasmids which contained two sgRNAs framing exons of interest for gene knock-out. Cas9-GFP expressing plasmid pSpCas9(BB)-2A-GFP, STAgR cloning template

STAgR_gRNAScaffold_hU6 and the STAgR_Neo plasmid for gRNA insertion were kindly provided by the group of Stefan Stricker. In brief, the required gRNAs were designed as described in 4.5.3. Plasmids containing the two gRNAs were generated by first, synthesizing the required building blocks via PCR as described in 4.5.5. Next, the vector reaction was digested using restriction enzyme DpnI as described in 4.5.6 to get rid of template plasmid. PCR reactions were loaded onto 1% agarose gels as described in 4.5.7 and fragments of interest were extracted as described in 4.5.8. Next, Gibson assembly was performed as described in 4.5.9 and after transformation into chemically competent bacteria (see section 4.4.1), colony PCR was performed as described in 4.5.10. DNA was isolated from samples with supposedly correct modification and sequenced as described in 4.5.11.

4.5.3 gRNA design

sgRNAs targeting exons of interest were designed using benchling software. Thereby, pairs of gRNAs were designed to span exons that are present in all transcript variants of the gene. Two gRNAs typically spanned regions of between 100 and 200 base pairs (bps). The following gRNAs were used:

Table 4.1: Sequences of sgRNAs used for CRISPR.

Target Gene	gRNA identifier	Sequence (5' to 3')
BRCA1	BRCA1_1	ACACAACAAAGAGCATACAT
BRCA1	BRCA1_2	GTTTCTATCATCCAAAGTAT
BRCA1	BRCA1_3	TGTGCTTTTCAGCTTGACAC
BRCA1	BRCA1_4	AAATAAGGTGTGAGACCAGT
BRCA1	BRCA1_5	TAAATCTCGTACTTTCTTGT
BRCA1	BRCA1_6	CACAATTACTTTCTATGACG
TP53	TP53_1	GGAGACAGAGTTGAAAGTCA
TP53	TP53_2	GCAGTCACAGCACATGACGG
TP53	TP53_3	GTTGGCAAAACATCTTGTTG
TP53	TP53_4	GAGCGCTGCTCAGATAGCGA
TP53	TP53_5	GGAGTACTGTAGGAAGAGGA
TP53	TP53_6	GCTTGTAGATGGCCATGGCG
CDH1	CDH1_1	TAAAGAAGGATCCCAACACT
CDH1	CDH1_2	GCAATGCGTTCTCTATCCAG
CDH1	CDH1_3	ATAATAAAGACACCAACAGG

CDH1	CDH1_4	TTAGAAGCTTGTTGACACCG
CDH1	CDH1_4	TTAGAAGCTTGTTGACACCG

4.5.4 Primer design

Primers were designed with benchling and the NCBI primer designing tool. Primers were designed to have an optimal annealing temperature between 55°C and 70°C. However, the actual optimal annealing temperature for each primer pair was determined experimentally. Primers were stored as stocks with a concentration of 100 pmol/µl and further diluted for experiments to a working concentration of 100 nM.

For STAgR cloning (see also section 4.5.2), primers with overhangs for cloning were ordered. Thereby, the forward primer consisted of the sequence of the desired gRNA (sense) and overlapping overhangs matching the sequence of scaffold. The reverse primer consisted of the reverse complementary sequence of the gRNA and was designed to have matching overhangs with the hU6 promotor.

Forward primer: 5‘ $N_{20(\text{sgRNA sequence sense})}$GTTTTAGAGCTAGAAATAGCAAGTT 3‘
Reverse primer: 5‘ $N_{20(\text{sgRNA sequence reverse complement})}$CGGTGTTTCGTCCTTT 3‘

All primers were ordered from Eurofins.

4.5.5 Polymerase chain reaction (PCR)

DNA regions of interest were multiplied with PCR using Phusion or Q5 polymerase. Thereby, PCR with Q5 polymerase was used to investigate potential gene knock-out on short regions of DNA. PCR using Phusion polymerase was used to amplify gRNA pieces and vector backbone for STAgR cloning.

For PCR using Q5 polymerase, reaction tubes with a total volume of 30 µl were prepared. In detail, one tube contained 100 ng of gDNA, Q5 buffer, 1 µl of the respective forward and reverse primer (10 µM), 20 units/ml of Q5 polymerase, 1 µl of dNTPs (10 mM) and Milli-Q water. Denaturation, annealing, and extension cycles were performed as recommended by the supplier of Q5 polymerase. The gradient function of the Eppendorf Mastercycler® nexus gradient was used to test different annealing temperatures.

PCR using Phusion polymerase during STAgR cloning was used to generate plasmids containing two gRNAs. For amplification of STAgR pieces, the forward primer containing

the gRNA sequence of the first gRNA and the reverse primer, containing the gRNA sequence of the second gRNA, were used. Each STAgR piece amplification was prepared as follows: An aliquot of STAgR_gRNAScaffold_hU6 (10 ng) was placed at the bottom of a PCR tube. 1 µl of dNTPs (10 µM) was added together with 2.5 µl of the respective primers (10 µM). Phusion buffer (5x) and 20 units/ml Phusion polymerase were added together with 1.5 µl Dimethyl-sulfoxide (DMSO). Finally, the volume was filled up to 50 µl using Milli-Q water.

The following conditions were used for PCR:

Table 4.2: Cycling program for amplification of gRNA STAgR pieces.

Step	Temperature (in °C)	Duration (in seconds)	Cycles
Initialization	98	90	1x
Denaturation	98	10	
Annealing	59	30	38x
Extension	72	30	
Final Extension	72	10	1x

For amplification of the vector backbone, forward primer containing the gRNA sequence of the second gRNA and the reverse primer, containing the gRNA sequence of the first gRNA, were used. The PCR mixture was equivalent to the one used for STAgR pieces with the alteration that 10 ng of gRNA_Neo vector were used instead of STAgR_gRNAScaffold_hU6 vector. The following conditions were used for PCR:

Table 4.3: Cycling program for amplification of vector backbone.

Step	Temperature (in °C)	Duration (in seconds)	Cycles
Initialization	98	90	1x
Denaturation	98	10	
Annealing	59	30	38x
Extension	72	90	
Final Extension	72	10	1x

PCR products were analyzed on agarose gels and, if required, extracted from those for further use (see section 4.5.8).

4.5.6 Restriction enzyme digest

Restriction digestions were performed using restriction endonucleases from NEB. Restriction digestions were performed to prepare DNA (such as PCR products) for cloning as well as to control correctness of cloning. In detail, 20-50 µl reactions were prepared to digest DNA. Thereby, reactions contained DNA, Milli-Q water, restriction enzyme(s) and a buffer as recommended by the manufacturer. Digest was performed on a ThermoMixer according to the manufacturer's recommendation.

4.5.7 Agarose gel electrophoresis

Agarose gels were prepared with agarose powder. In detail, agarose powder was mixed with 1x TAE buffer to a final concentration between 1% and 2% (m/m). The solution was heated in the microwave to start the gelling. Afterwards, SYBR Safe DNA stain was added at a ratio of 1:10,000. Next, the gel was poured into gel carriers with appropriate gel combs while avoiding bubbles. After the gel had solidified, the gel carrier was filled with 1x TAE buffer, and the gel comb was removed. Before loading DNA samples onto the gel pockets, samples were mixed with orange loading dye (6x). To determine the size of the appearing DNA lanes, at least one pocket was filled with DNA weight marker. DNA fragments were separated at 120 V for 45-60 minutes. Following the separation step, either pictures of the gel were taken on the ChemiDoc UV transilluminator for analysis or DNA lanes were visualized using a blue light table.

4.5.8 Gel extraction of DNA fragments

For isolation of DNA fragments from agarose gels, the separated DNA lanes were visualized on a blue light table. The fragments of desired size were carefully cut out of the gel using a scalpel. Gel pieces were transferred into 1.5 ml Eppendorf tubes and purified using the Monarch gel extraction kit. Thereby, the manufacturer's instructions were followed. The required heating step was performed on a ThermoMixer. DNA was eluted using 15-20 µl of pre-heated elution buffer (provided within the kit) and DNA concentration was measured on a NanoDrop spectrophotometer.

4.5.9 Ligation of DNA fragments

DNA fragments were ligated using Gibson assembly, a method to connect DNA fragments with overlapping ends. Here, I used Gibson assembly to assemble PCR products after gel

extraction in order to generate sgRNA-containing plasmids for gene editing. For generating plasmids containing two sgRNAS, two fragments had to be connected: the vector backbone and one gRNA STAgR piece. The 2x Gibson assembly master mix was used according to the manufacturer's recommendations. Thereby, a Gibson assembly reaction was set up using 100 ng of purified vector backbone DNA and 29.64 ng gRNA STAgR piece (equals 3x molar excess of insert). After incubation at 50°C or 60 minutes in a ThermoMixer (no shaking), the reaction was diluted 1:4 in Milli-Q water and then 2 µl of the reaction were transformed into competent bacteria (see section 4.4.1).

4.5.10 Colony PCR and sequencing

After transformation into competent bacteria and enrichment on a LB plate, single bacterial colonies were picked. For colony PCR, the bacterial colony was picked using a pipet tip and smeared onto the bottom of a PCR reaction tube. Next, the remaining bacteria on the pipet tip were used to inoculate 2 ml of LB medium with Ampicillin (100 µg/ml). For PCR, the Taq polymerase was used according to the manufacturer's recommendations with modifications. In detail, per reaction, only a total volume of 10 µl was prepared. The primers used here were designed to bind within the vector backbone while flanking the region in which the sgRNAs were supposedly inserted.

For screening assembled plasmids with the desired amount of sgRNAs, the following primers were used (5' to 3'):

STAgR_seq_fwd2: ACTGGATCCGGTACCAAGG

STAgR_seq_rev: TTACGGTTCCTGGCCTTTTG

The following conditions were used for PCR:

Table 4.5: Cycling program for colony PCR.

Step	Temperature (in °C)	Duration (in seconds)	Cycles
Initialization	94	360	1x
Denaturation	94	30	
Annealing	55	30	38x
Extension	72	120	
Final Extension	72	600	1x

The PCR products were analyzed on 1% agarose gels (see section 4.5.7). Thereby, the size of the PCR product after successful introduction of both sgRNAs, was approximately 740 bps. For samples that showed the correct size in colony PCR, plasmid DNA was isolated from previously inoculated bacterial colonies. To validate the correctness of plasmid sequences, DNA of selected bacteria colonies was sequenced. For sequencing, the Mix2Seq kit was used according to the manufacturer's instructions. The same primers as used for colony PCR were used. Upon validation of the sequence, plasmid DNA was amplified using midi preparation (see section 4.4.2).

4.5.11 Validation of genetic modification in luminal progenitor cells

After transfection and enrichment of successfully transfected luminal progenitor cells at the FACS, PCR was used to validate genetic modification at the target region. To this end, primers flanking the target exon of interest were designed as described in 4.5.4. First, optimal annealing temperatures were determined experimentally using wild-type gDNA. Then, gDNA was isolated from transfected bulk cells (see section 4.5.1). PCR was conducted using Q5 polymerase as described in 4.5.5.

Table 4.6: Sequences of all primers used for checking gene editing after transfection.

Gene	Primers (fwd/rev) (5' to 3')	Product length (wt)	T (Annealing)
BRCA1	CAGGAGCCTACAAGAAAGTACGA	825 bp	65°C
BRCA1	TCGGGTTCACTCTGTAGAAGTC		
TP53	AGACGCCAACTCTCTCTAGC	500 bp	65°C
TP53	GCACCACCACACTATGTCGAAAA		
CDH1	GCTGTCTGGCTAGGTTGGAC	583 bp	67°C
CDH1	GATCCAGCATGGGTTGACCA		

After PCR, DNA was run on 1.5% agarose gels for analysis as described in 4.5.7.

4.6 Data analysis and statistical evaluation

Structure prevalence, structure type, structure size and structure complexity were analyzed using GraphPad prism 8. Calculation of nuclei velocity within organoids was facilitated via optical flow using the Farneback algorithm implemented in Matlab. Thereby, a branch axis was defined, and the velocities were separated into a parallel and an orthogonal component relative to the branch axis. Therefore, if the directionality equals 1, this means that the

parallel component equals the orthogonal component. A value higher or lower than 1, indicates that either the parallel or the orthogonal component prevail. Visualization of the migration patterns of individual cells was facilitated by manually tracking individual cells. To do so, a single nucleus was tracked three times and the average path was calculated using these values.

For conveying statistical information, data was several times presented as box-and-whisker plots. Those plots show data as median ± 25% and were generated with GraphPad prism software using the 'min to max' method. The remaining data is depicted as mean ± s.d.. P values were calculated using an unpaired two tailed *t*-test that allows to compare two groups. Thereby, $P > 0.05$ was considered not significant (n.s.) while $P \leq 0.05$ was considered significant. The level of significance is indicated as follows: $^{*}P \leq 0.05$, $^{**}P \leq 0.01$, $^{***}P \leq 0.001$, $^{****}P \leq 0.0001$.

5 Results

5.1 Generation of complex branched ductal organoids from human mammary epithelial cells in collagen type I gels

To allow cell-type-specific observations under defined conditions, I isolated the epithelial lineages of the mammary gland from women undergoing reduction mammoplasties for aesthetic reasons via fluorescence activated cell sorting (FACS). In detail, after exclusion of cellular debris and doublets using side scatter (SSC) and forward scatter (FSC) (Fig. 3A), dead cells were removed based on 7-AAD uptake (Fig. 3B) (Ganz et al., 2021). Next, endothelial and hematopoietic cells were excluded based on their expression of CD31 (DeLisser et al., 1994) and CD45 (Hermiston et al., 2003; Shivtiel et al., 2008), respectively (Fig. 3C) (Ganz et al., 2021). Finally, the epithelial lineages of the breast were separated from each other and from the remaining stroma via the established markers CD49f and EpCAM (Eirew et al., 2008) (Fig. 3D) (Ganz et al., 2021). For exclusion of a formerly reported stromal contamination within the basal subset, I additionally used the previously described marker CD10 (Bachelard-Cascales et al., 2010; Linnemann et al., 2015) (Fig. 3D) (Ganz et al., 2021).

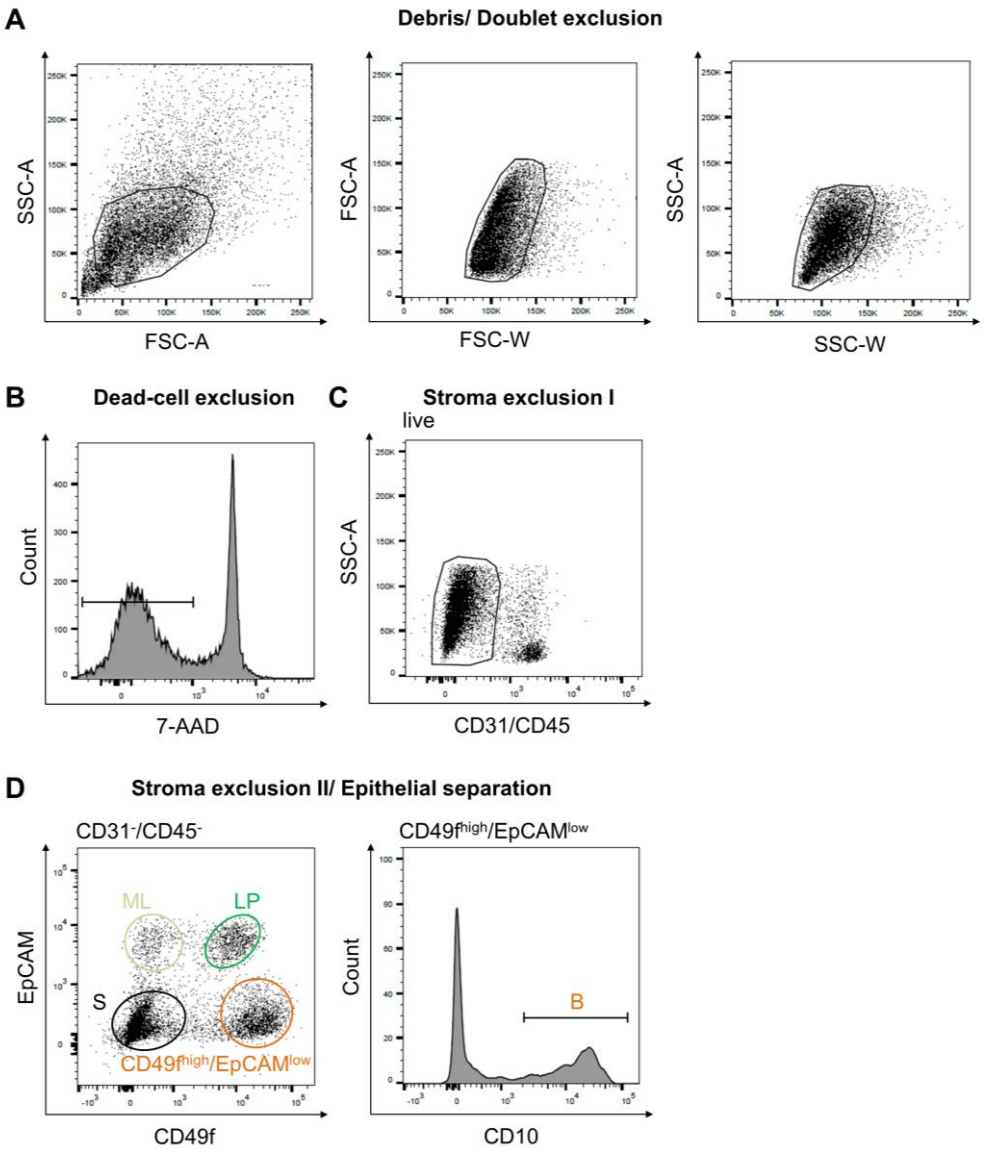

Figure 3. FACS sorting layout of human mammary tissue. (A) Exclusion of debris and cell doublets based on side scatter (SSC) and forward scatter (FSC). Cells outside of the indicated gates were excluded. **(B)** Exclusion of dead cells via 7-AAD uptake. Cells that were positive for 7-AAD (outside indicated gate) were excluded. **(C)** First exclusion of stromal cells based on CD45 and CD31. Cells outside of the indicated gate were excluded. **(D)** Second exclusion of stromal cells and separation into epithelial subsets based on CD49f and EpCAM expression. The four populations arising are S=stroma, ML=mature luminal, LP=luminal progenitor and a CD49f^{high}/EpCAM^{low} population. Further sorting of the CD49f^{high}/EpCAM^{low} population via CD10 results in separation of remaining stroma from the B=basal population.

In summary, the three epithelial lineages isolated were luminal progenitor (CD49f^{high}/EpCAM^{high}), mature luminal (CD49f^{low}/EpCAM^{high}) and basal cells (CD49f^{high}/EpCAM^{low}/CD10^{+}). After every sort, I validated population purity by re-analyzing the sample by FACS (Fig. 4A,B) (Ganz et al., 2021). Sorted cells were only used for subsequent experimental procedures when sorting purity was at least 99.5%.

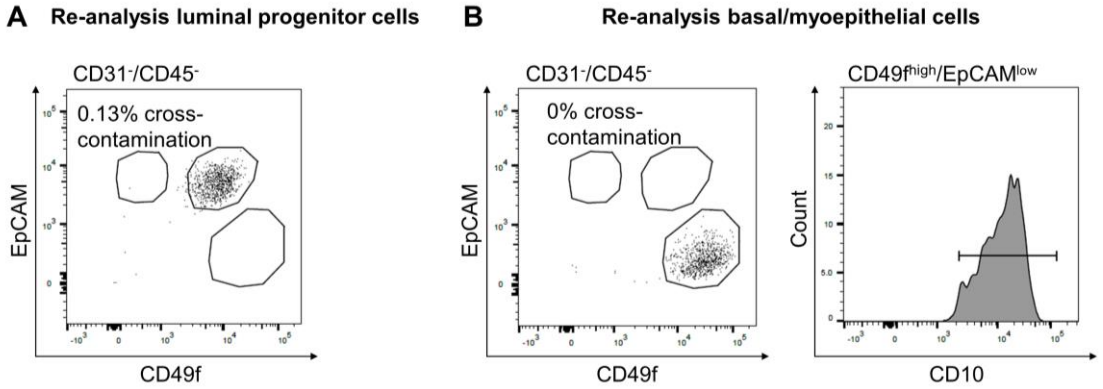

Figure 4. Re-analysis of FACS-sorted epithelial populations. (A) Re-analysis of FACS-sorted luminal progenitor population based on EpCAM/CD49f expression using gating strategy applied during previous sort. Cross contamination was calculated in % of all $CD31^-/CD45^-$ cells. **(B)** Re-analysis of FACS-sorted basal population based on EpCAM/CD49f as well as CD10 expression using gating strategy applied during previous sort. Cross contamination was calculated in % of all $CD31^-/CD45^-$ cells. Depicted is one representative re-analysis.

5.1.1 Branched structure formation from luminal progenitor cells

Due to the role of luminal progenitors ($CD49f^{high}/EpCAM^{high}$) as cell of origin in breast cancer as outlined above, the focus of the work described here was unravelling the capacity of the luminal progenitor subset to undergo invasion and form branched ductal organoids. To ensure robust rates of proliferation and maintenance of luminal cell characteristics in culture, I chose the breast cancer organoid medium (BCOM) (Sachs et al., 2018) as liquid compound of the culture system. This medium was originally used for culturing mammary gland fragments in a Matrigel matrix. Thereby, it supported the maintenance of basal as well as luminal lineage characteristics (Rosenbluth et al., 2020; Sachs et al., 2018). As matrix, I chose collagen type I, which is a prevailing component of the mammary gland ECM. When I seeded single freshly sorted luminal progenitor cells into these conditions, I found that a subset of cells formed small cell clusters that gave rise to simple branched organoids with short ducts (Fig. 5A) (Ganz et al., 2021). Thereby, I defined a structure as branched organoid if it had at least two branching points. However, only 8% of all arising structures were branched organoids and the majority of cells grew into sticks or spheres instead (Fig. 5B) (Ganz et al., 2021). In contrast to the original application of BCOM for support of cells within fragments, here, I was aiming for promotion of single cells in culture and set out to optimize the medium for this purpose.

Previously, it had been reported that small amounts of fetal calf serum (FCS) can support initial survival and outgrowth of single primary mammary epithelial cells in

culture(Linnemann et al., 2015). And indeed, upon addition of 0.5% FCS, I observed a 12-fold increase in branched structure formation (Fig. 5A-C) (Ganz et al., 2021).

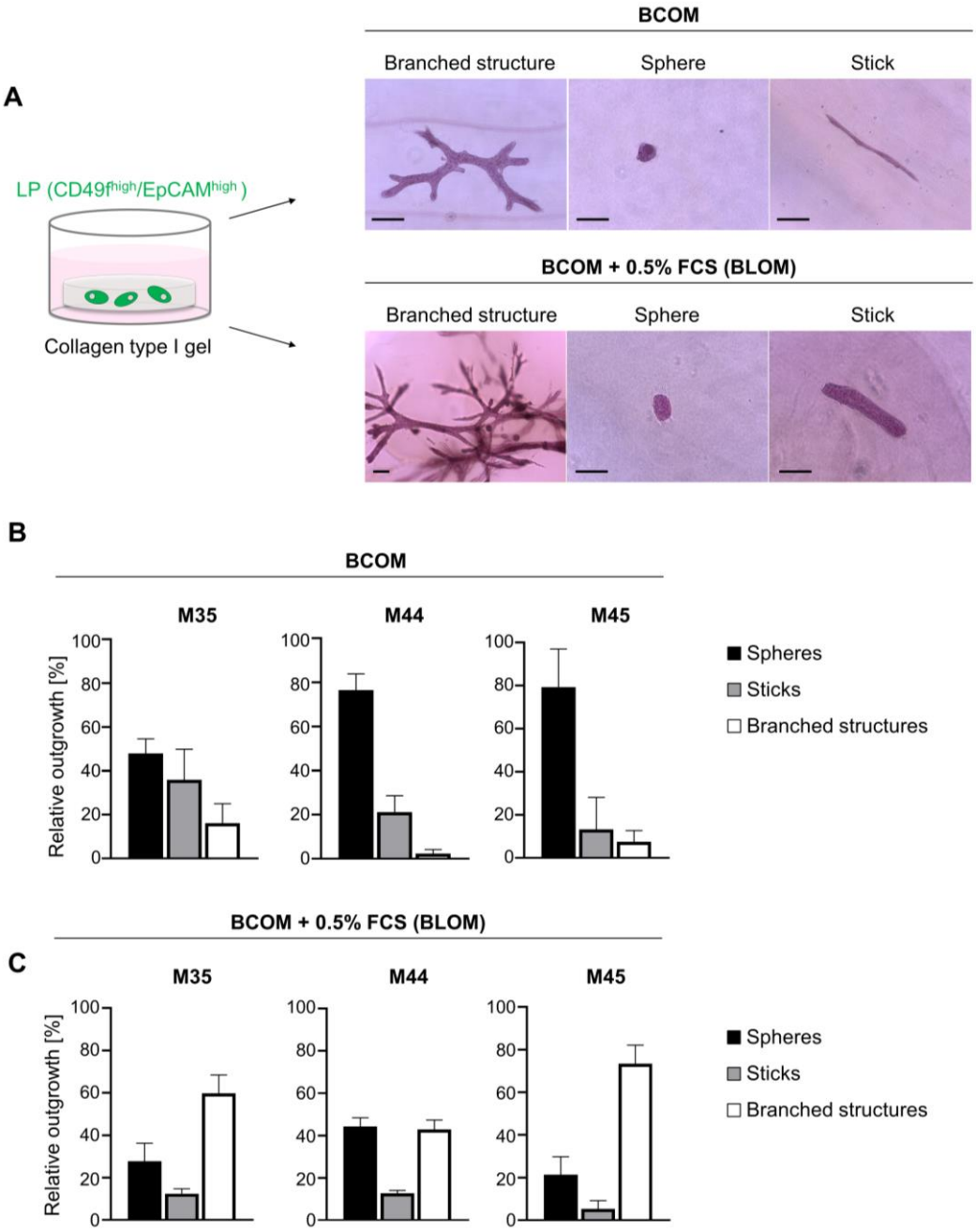

Figure 5. Formation of structures from human mammary luminal progenitor cells in BCOM vs. BLOM in collagen type I gels. (A) Illustration of experimental setup with single, FACS-sorted luminal progenitor cells seeded into collagen type I gels. Carmine staining of arising structure types in BCOM and BLOM medium. Scale bars: 100 µm. **(B)** Quantification of outgrowing structure type in BCOM medium. Depicted are donors M35, M44 and M45. Structure type was normalized to total structures arising per gel. Data are mean ± s.d. *n*=4 gels/condition. **(C)** Quantification of outgrowing structure type in BLOM medium. Depicted are donors M35, M44 and M45. Structure type was normalized to total structures arising per gel. Data are mean ± s.d. *n*=4 gels/condition.

In addition, the emerging branched structures were in average 2.6-fold larger and significantly more complex (Fig. 6A,B) (Ganz et al., 2021). In detail, upon FCS addition, the percentage of structures with at least tertiary branching increased from 4% to 65% (Fig. 6B) (Ganz et al., 2021). Based on the advances upon FCS addition, I hereafter named the FCS-containing medium 'branched luminal organoid medium' (BLOM).

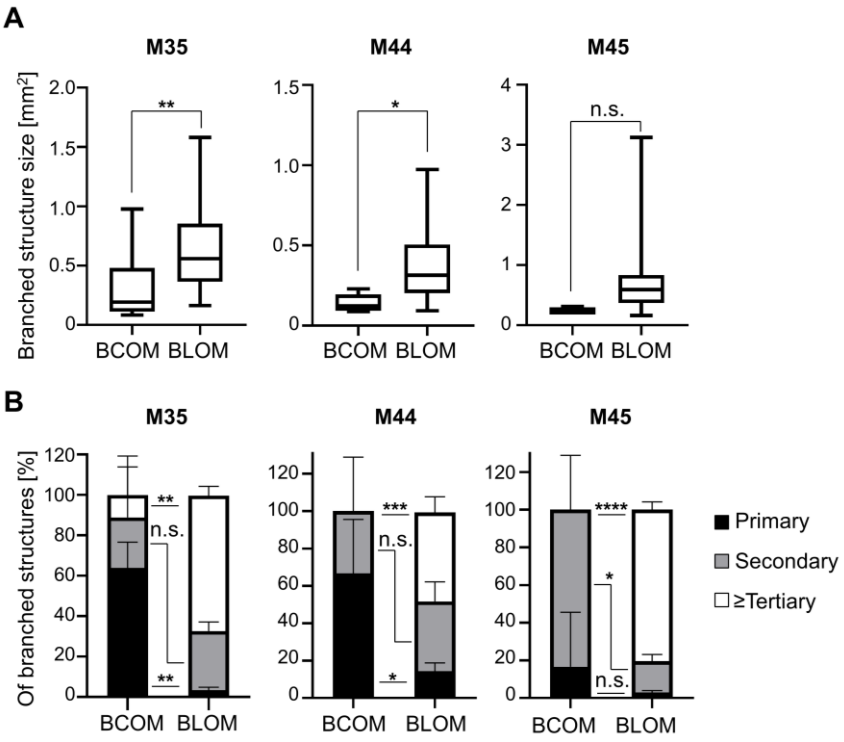

Figure 6. Size and complexity of branched ductal structures from human mammary luminal progenitor cells in BCOM vs. BLOM in collagen type I gels. (A) Quantification of branched structure sizes in BCOM vs. BLOM. Depicted are donors M35, M44 and M45. Size was defined as the area covered by the organoid when all tips were connected virtually using ImageJ. Data are shown as median ± 25%. n=4 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, *P ≤ 0.05, **P ≤ 0.01. **(B)** Quantification of branched structure complexity in BCOM vs. BLOM medium. Depicted are donors M35, M44 and M45. Data are presented as stacked bars. Data are mean ± s.d. *n*=3 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, *P ≤ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001.

Importantly, I found that the ability to form branched luminal organoids in BLOM was conserved over several donors. In total, branched organoid forming capacity was observed in 12 out of 12 donors (Fig. 7A) (Ganz et al., 2021). Importantly, the ductal phenotype of branched structures was reminiscent of the morphology of low-grade NST carcinomas with a high degree of duct formation (Fig. 7B) (Ganz et al., 2021). Moreover, branched organoid formation only occurred within collagen type I gels. In contrast, when Matrigel was used as ECM surrogate, only sphere-like or budding-like structures formed (Fig. 7C) (Ganz et al., 2021).

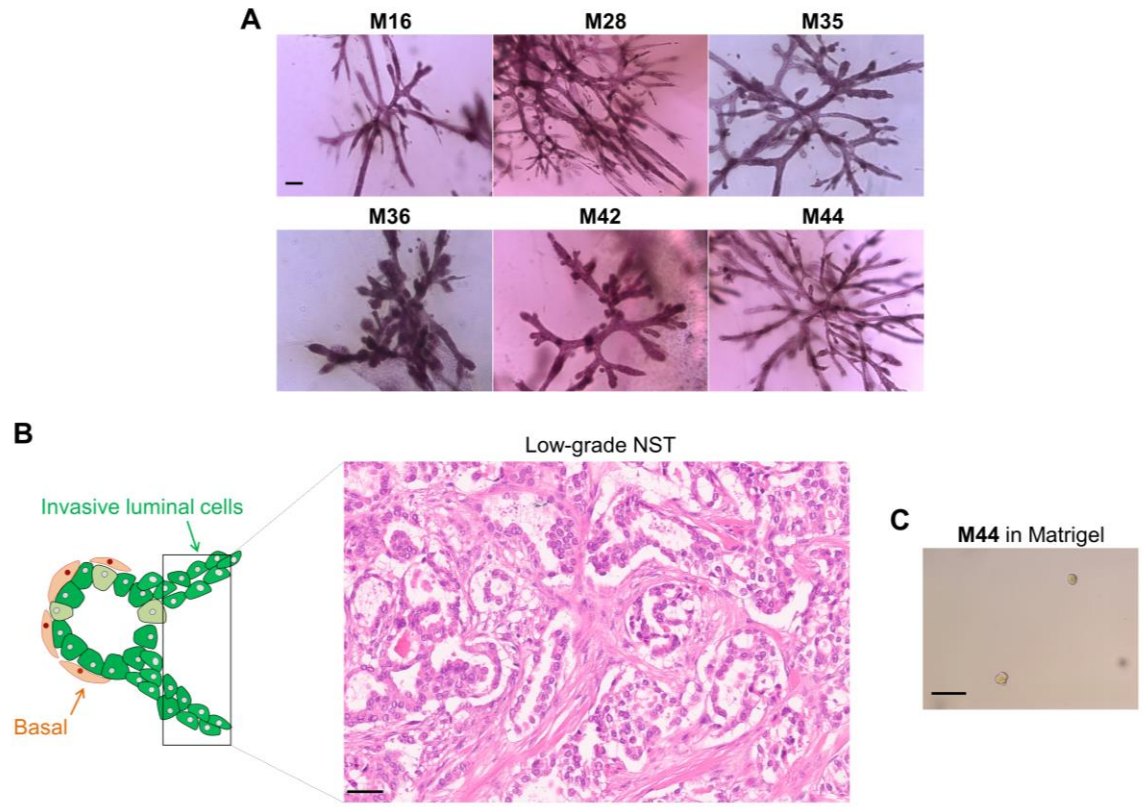

Figure 7. Formation of ductal structures from various luminal progenitor donors in collagen type I and comparison to Matrigel as matrix. (A) Carmine stainings of branched structures arising from luminal progenitors in collagen type I gels from donors M16, M28, M35, M36, M42 and M44 in BLOM medium after 8 to 12 days in culture. Scale bar: 100 µm**. (B)** Schematic illustration of invasively growing luminal cells (green) growing beyond the borders formed by basal cells (orange) and H&E-stained section of low-grade NST carcinoma. Scale bar: 50 µm. **(C)** Light microscopy picture of luminal progenitor cells from donor M44 in BLOM in Matrigel matrix after 7 days in culture. Scale bar: 100 µm.

These data show that healthy primary luminal progenitor cells in collagen type I matrix have the capacity to give rise to complex branched ductal structures. Thereby, the resulting ductal organoids morphologically resemble low-grade NST carcinomas. The cell of origin role of luminal cells and the morphological resemblance of the organoids to invasive cells *in vivo*, suggests that formation of luminal cell-derived ductal organoids could be a process closely related to cancerous invasion events.

5.1.2 Lack of branched structure forming potential in mature luminal cells

Next, I set out to investigate whether under the same conditions the potential to generate branched ductal organoids could also be elicited from the more differentiated luminal subset, the mature luminal cells (CD49f^{low}/EpCAM^{high}). Doing so, I found that freshly sorted mature luminal cells in collagen type I gels barely attached and did not show development beyond the single cell state (Fig. 8A) (Ganz et al., 2021). This is in line with previous reports

stating difficulties in maintaining and expanding this subset in culture(Lim et al., 2009; Linnemann et al., 2015).

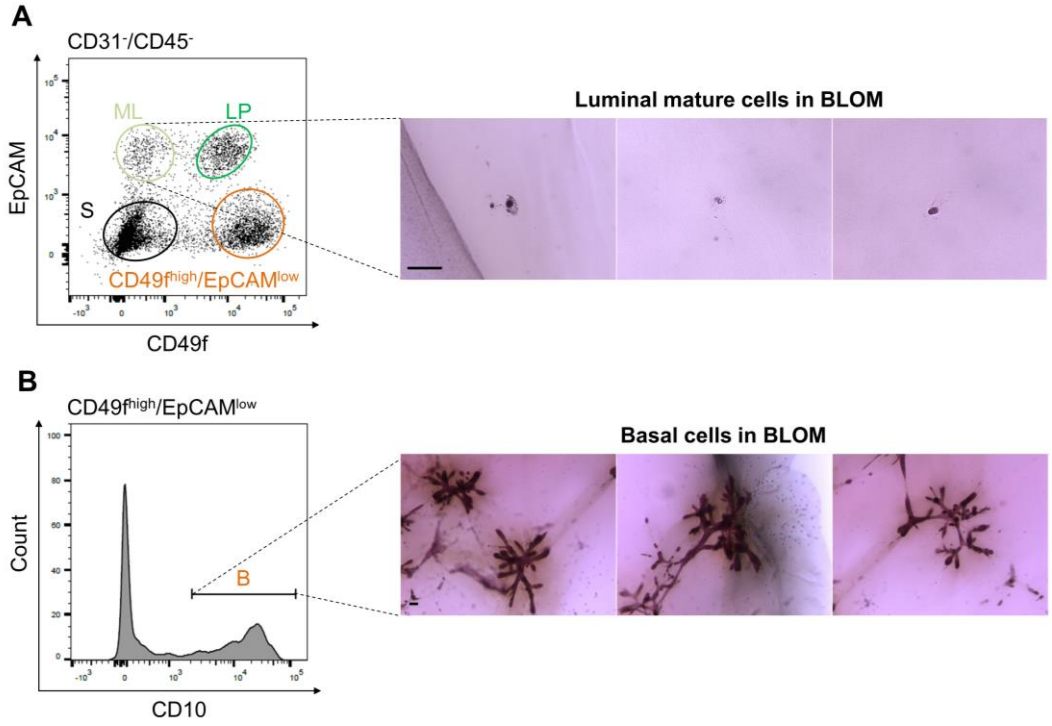

Figure 8. Branched structure formation from basal and mature luminal cells in collagen type I gels. (A) Sorting of epithelial cells from reduction mammoplasties via FACS based on EpCAM and CD49f expression. The four populations arising are S=stroma, ML=mature luminal, LP=luminal progenitor and a CD49f^{high}/EpCAM^{low} population. Carmine stainings of outgrowth from sorted mature luminal cells in collagen type I gels in BLOM medium after 10 days. Depicted is donor M16. Scale bar: 100 µm. **(B)** Further sorting of the CD49f^{high}/EpCAM^{low} population via CD10 results in separation of remaining stroma from the B=basal population. Carmine stainings of outgrowth from sorted B cells in collagen type I gels in BLOM medium after 10 days. Scale bar: 100 µm. Experiments were repeated in *n*=3 donors.

5.1.3 Basal cell-derived branched structures in different media

In contrast to luminal mature cells, freshly FACS-sorted basal cells (CD49f^{high}/EpCAM^{low}/CD10^{+}) generated branched structures in the assay, an ability that has been described for this subset before (Linnemann et al., 2015) (Fig. 8B) (Ganz et al., 2021). Different to branched ductal organoid formation from luminal progenitor cells, there is no indication that formation of such structures from basal cells reflects cancerous development. Rather, the formation of branched structures from single basal cells has previously been described as a reflection of regenerative capacity within this subset (Linnemann et al., 2015). This is mainly based on the fact that basal cells have shown bipotent potential in collagen type I gels where they differentiated into bilayered branched structures with distinct alveoli resembling terminal ductal lobular units, the functional units of the normal mammary

gland (Linnemann et al., 2015). I set out to investigate whether the same regenerative potential was unraveled in the basal subset within collagen type I gels in combination with the BLOM medium. Direct comparison of branched structures arising from basal cells in BLOM to those derived in culture conditions as described by Linnemann et al. revealed minor differences in overall morphology. In detail, basal cell-derived structures in BLOM had a rather uniform width with minor alveoli-like thickening at the tip of each duct (Fig. 9A).

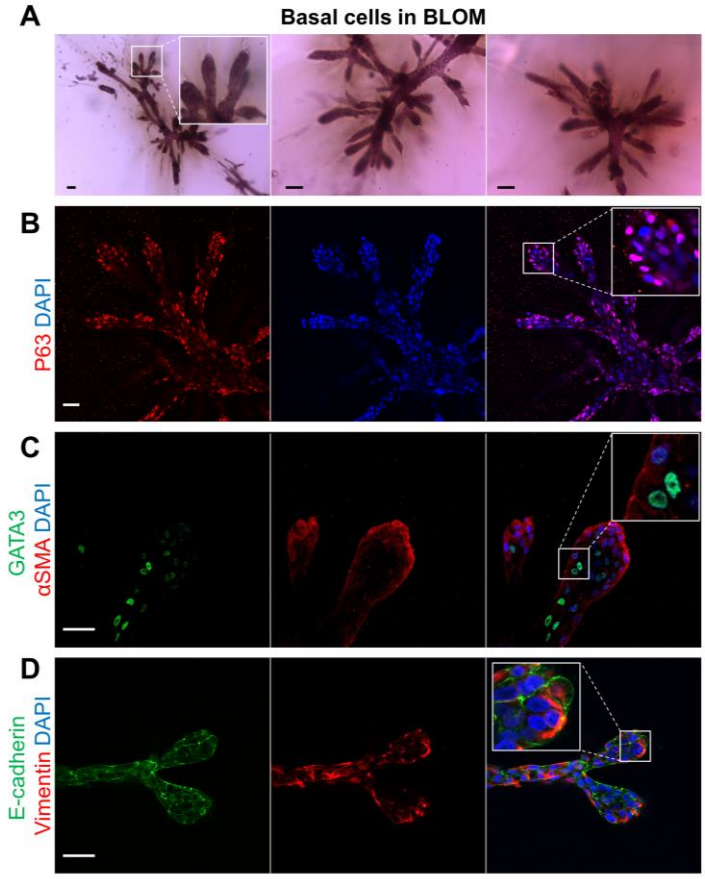

Figure 9. Basal cell-derived branched structures in BLOM in collagen type I gels. (A) Carmine stainings of outgrowth from sorted basal cells in collagen type I gels in BLOM medium after 10 days. Scale bars: 100 µm. **(B)** Confocal microscopy on basal cell-derived branched structures in BLOM medium after 10 days. Representative images of basal lineage marker p63 (red) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(C)** Confocal microscopy on basal cell-derived branched structures in BLOM medium after 10 days. Representative images of basal lineage marker αSMA (red) and luminal lineage marker GATA3 (green) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(D)** Confocal microscopy on basal cell-derived branched structures in BLOM medium after 10 days. Representative images of basal marker vimentin (red) and luminal marker E-cadherin (green) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. All imaging experiments were repeated in *n*=30 structures/condition; 3 donors; 10 structures/donor.

In contrast, basal cell-derived structures grown in conditions according to Linnemann et al. (Linnemann et al., 2015) had a more distinct alveologenesis whereas their ducts were comparatively thin (Fig. 10A,E).

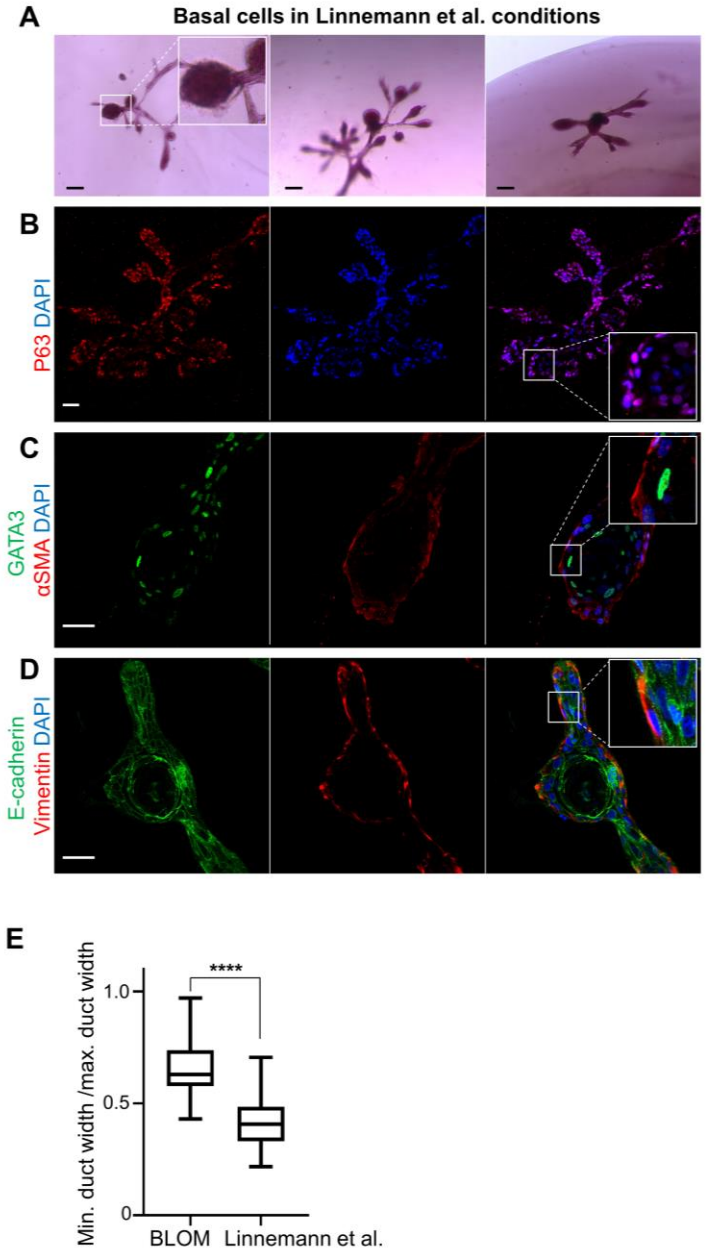

Figure 10. Basal cell-derived branched structures in Linnemann et al. medium in collagen type I gels. (A) Carmine stainings of outgrowth from sorted basal cells in collagen type I gels in medium according to Linnemann et al. after 10 days. Scale bars: 100 µm. **(B)** Confocal microscopy on basal cell-derived branched structures in medium according to Linnemann et al. after 10 days. Representative images of basal lineage marker p63 (red) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(C)** Confocal microscopy on basal cell-derived branched structures in medium according to Linnemann et al. Representative images of basal lineage marker αSMA (red) and luminal lineage marker GATA3 (green) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(D)** Confocal microscopy on basal cell-derived branched structures in medium according to Linnemann et al. Representative images of basal marker vimentin (red) and luminal marker E-

cadherin (green) expression. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. All imaging experiments were repeated in *n*=30 structures/condition; 3 donors; 10 structures/donor. **(E)** Quotient of minimum duct width and maximum duct width (alveolus). Data are shown as median ± 25%. *n*=25 ducts. P values were calculated using an unpaired two-tailed *t*-test, ****P ≤ 0.0001.

I used immunofluorescence in combination with confocal microscopy to further characterize and compare basal cell-derived branched structures. Thereby, staining for basal lineage marker p63 confirmed the overall basal identity of the structures (Fig. 9B, 10B) (Ganz et al., 2021). However, I noticed that while in both conditions outer layers of basal cells were generally positive for p63, basal cells at the inner, luminal positions were often negative for p63 (Fig. 9B, 10B). This finding indicated that in both, BLOM conditions as well as conditions according to Linnemann et al., dedifferentiation occurred within the inner layer of organoid cells. A co-staining of basal lineage marker αSMA and the luminal transcription factor GATA3 confirmed this notion and further unraveled bipotent potential in both conditions. In detail, I found that the outer layer of basal cells reliably expressed αSMA while cells at luminal positions did not (Fig. 9C, 10C). At the same time, the cells at luminal locations had upregulated GATA3 indicating a shift towards luminal cell identity (Fig. 9C, 10C). Furthermore, I examined expression of E-cadherin and vimentin. Cell-cell adhesion molecule E-cadherin is in the normal mammary gland mainly expressed within the generally more epithelial luminal subset(Andrews et al., 2012). However, E-cadherin is not an exclusive luminal lineage marker and in line with previous reports on basal cells in culture (Linnemann et al., 2015), I found E-cadherin localized at cell-cell junctions of basal cell-derived structures (Fig. 9D, 10D). In contrast, vimentin is *in situ* a marker that is typically associated with the basal subset (Peuhu et al., 2017). In line with that notion, I found vimentin expression in both conditions examined. Strikingly, vimentin was mostly localized towards the basolateral side of the organoids (Fig. 9D, 10D), which again supported the notion that at this side characteristic basal marker expression is preferably maintained.

In summary, the experiments presented here reveal that basal cells undergo branched structure formation within BLOM conditions. Moreover, the inner layer of organoid cells shows bipotent behavior and differentiates towards a luminal phenotype. Thereby, the basal cells behave similar to what has been described for this population in collagen type I gels within a different medium.

5.2 Branched luminal organoids arise clonally and express luminal lineage and polarization markers

5.2.1 Branched organoids arise at high frequency from single luminal progenitor cells

Based on the aim of this work to model cancer relevant invasion processes of the mammary epithelium rather than to characterize regenerative capacity, I focused all subsequent experiments on the luminal progenitor subset. To this end, I set out to further characterize organoid formation from luminal progenitor cells, particularly focusing on frequency and morphogenetic steps. For determination of frequency, I seeded luminal progenitor cells into collagen type I gels at low seeding density in limiting dilutions and confirmed their single cell state by light microscopy to verify that generated structures stem from one cell. Furthermore, by daily microscopic inspections during the growth process, I could rule out that branched organoids are the result of merging cells during the growth process. To determine the proportion of cells that have structure forming capacity, I used extreme limiting dilution analysis (ELDA). In this assay, cells are seeded in degrading solutions. By counting of cultures in which activity (in this case branched structure formation) can be detected, the frequency of cells of interested is determined (Y. Hu & Smyth, 2009). I used this method to determined branched structure forming units (B-SFU) within the luminal progenitor population. In detail, it was found that the frequency of B-SFUs varied between luminal progenitor populations of different donors. Overall, 1 out of 15 freshly sorted luminal progenitor cells was a B-SFU (Table 5.1) (Ganz et al., 2021). Moreover, it was found that B-SFU frequency varied from donor to donor with the lowest frequency in donor M50 (1 out of 21 cells) and the highest frequency in M44 (1 out of 10 cells) (Table 5.1) (Ganz et al., 2021).

Table 5.1: Extreme limiting dilution analysis (ELDA): determination of branched structure-forming units (B-SFU)

Number of seeded cells	Positive gels/total gels		
	M36	**M44**	**M50**
5	1/6	3/6	n/A
10	2/6	4/6	2/6
20	5/6	5/6	4/6
40	6/6	6/6	5/6

80	6/6	6/6	6/6
160	6/6	6/6	6/6
320	n/A	n/A	6/6
B-SFU	1/15.1	1/9.52	1/20.7
95% CI	1/8.85 – 1/26	1/5.53 – 1/17.7	1/11.8 – 1/36.5
Mean B-SFU		**1/15.11**	

Data are mean with 95% confidence intervals (CI).

In addition, tracking and monitoring organoid formation over time revealed distinct growth phases. In detail, single luminal progenitor cells attached slowly, started proliferating and formed small cell clusters (establishment phase, day 1-5) (Fig. 11) (Ganz et al., 2021). After day 5 of culture, those small clusters started to spread though the matrix increasingly and became more and more complex due to the continuous formation of side branches (elongation/branching phase, day 5-10) (Fig. 11) (Ganz et al., 2021). Lastly, branch tips started rounding-up and thickening while duct elongation decreased (rounding-up phase, day 11-13) (Fig. 11) (Ganz et al., 2021).

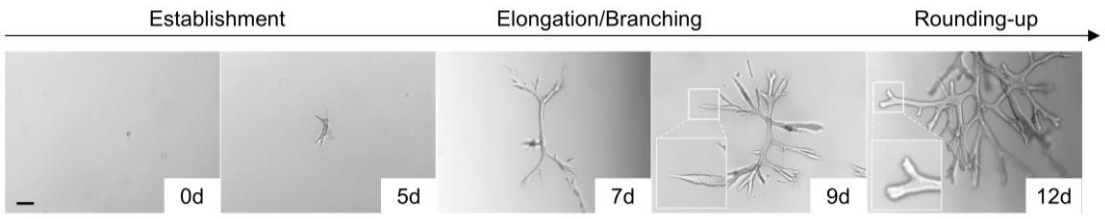

Figure 11. Branched structures arise in less than two weeks from single luminal progenitor cells. Light microscopy pictures of a branched luminal organoid developing from a single cell over the course of 12 days. Scale bar: 100 µm.

Taken together, these experiments show that branched organoids are generated at high frequency from single luminal progenitor cells. Moreover, the fact that branched organoids arise even at very low seeding density, indicates that they can be generated from one single cell and are therefore clonal.

5.2.2 Luminal progenitor-derived branched ductal organoids express luminal markers characteristic for NST carcinomas

Based on the morphological resemblance to NST carcinomas, I set out to investigate the expression of luminal markers that are expressed by a majority of NST carcinomas. These luminal markers include CK8/18 (Rattan et al., 2012), the transcription factor GATA3 (Guo et al., 2017; Shaoxian et al., 2017) (Fig. 12A,B) (Ganz et al., 2021), tight junction protein

ZO-1 (Bell et al., 2003) as well as the mucin MUC-1 (Matsukita et al., 2003). I performed immunostaining and confocal microscopy on fully grown organoids (day 11-13). This revealed that all organoids examined expressed CK8/18 at cytoplasmic locations (Fig. 12C) as well as nuclear GATA3 (Fig. 12D) (Ganz et al., 2021). Confocal microscopy further revealed the existence of a non-continuous lumen usually framed by a single layer of cells throughout the organoids (Fig. 12D) (Ganz et al., 2021).

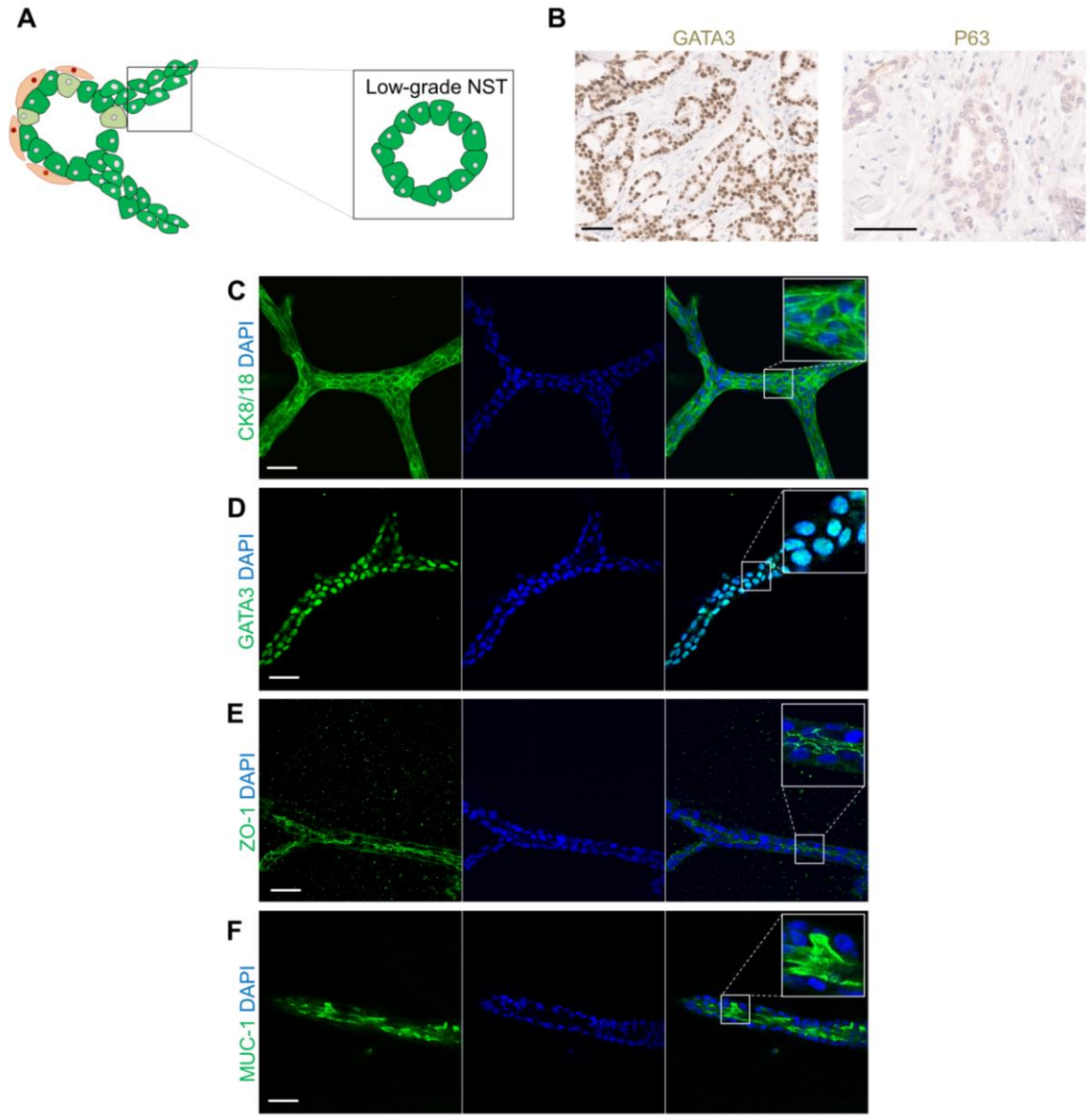

Figure 12. Luminal progenitor-derived branched ductal structures express luminal markers characteristic for low-grade NST. (A) Schematic illustration of invasively growing luminal progenitor cells with cross-section of an invasive and polarized duct composed of luminal cells (green) as characteristic for low-grade NST **(B)** GATA3 (brown) or p63 (brown) stained section of ducts arising in low-grade NST carcinomas. Scale bars: 50 µm. Immunohistochemistry and associated imaging were performed by Moritz Jesinghaus (Institute of Pathology, Technical University of Munich). **(C-F)** Confocal microscopy on fully grown branched luminal progenitor-derived organoids: representative images of luminal markers CK8/18 (green), GATA3 (green), ZO-1 (green),

MUC-1 (green), Laminin (green). Nuclei are stained with DAPI (blue). *n*=30 structures/condition; 3 donors; 10 structures/donor. Scale bar: 50 µm.

Of note, it has been shown previously that in order to achieve lumen formation during collective invasion of epithelial cells, correct apical-basal polarity is required (Friedl & Gilmour, 2009). Accordingly, organoids expressed ZO-1 at the apical side of most ducts indicating correct apical-basal polarization (Fig. 12E, Fig. 13C) (Ganz et al., 2021). In addition, I found luminal differentiation marker MUC-1 at the apical side in 89% of ducts (Fig. 12F, Fig. 13C) (Ganz et al., 2021) further corroborating the notion that organoids are correctly polarized. Notably, the formation of correctly polarized, lumen-containing ducts is a hallmark for the pathology of low-grade NST carcinomas (Fig. 12A,B) (Ganz et al., 2021).

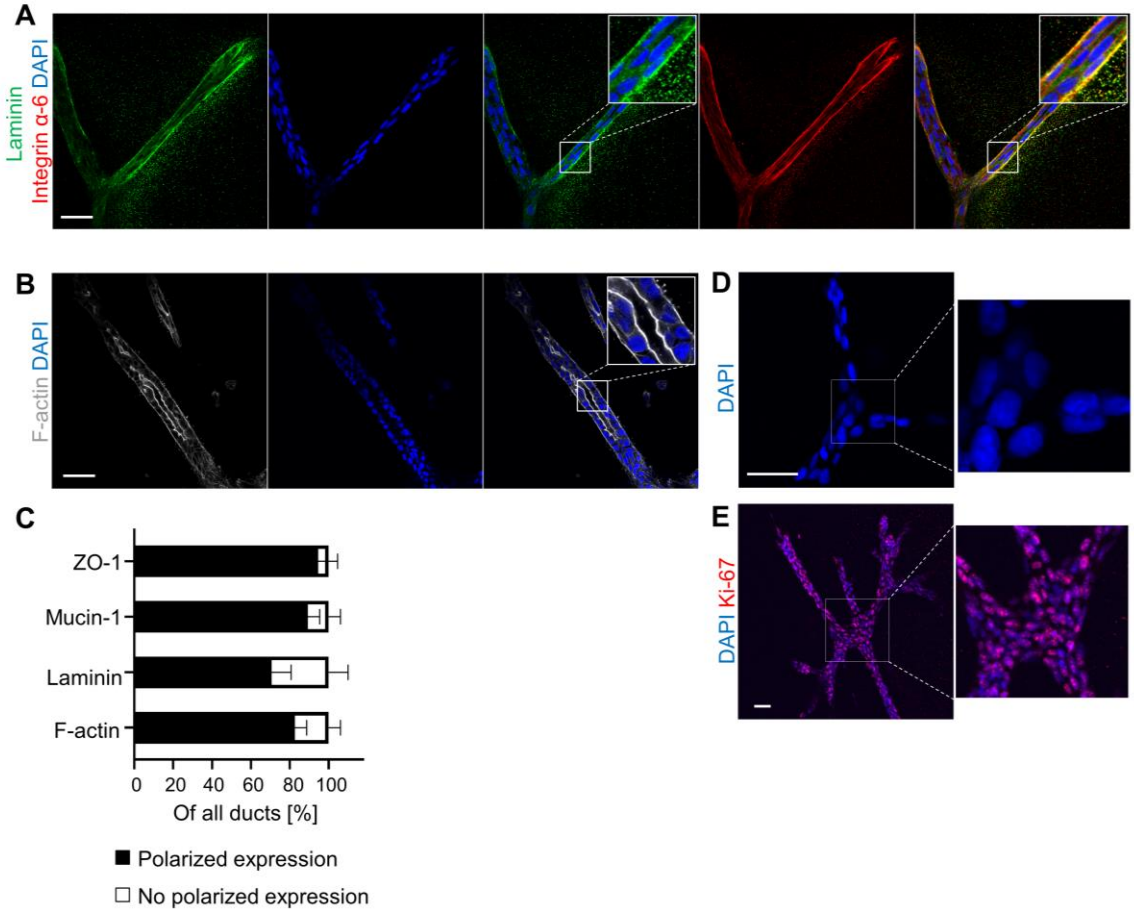

Figure 13. Characteristics of luminal progenitor-derived branched ductal structures with respect to grading system. (A) Confocal microscopy on fully grown branched luminal progenitor-derived organoid: representative images of polarization markers Laminin (green) and Integrin α-6 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(B)** Confocal microscopy on fully grown branched luminal progenitor-derived organoid: representative image of polarization marker F-actin (white). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(C)** Quantification of ducts with and without correctly polarized expression of markers. Data are presented as stacked bars with mean ± s.d. *n*=25 ducts. **(D)** Confocal microscopy on fully grown branched luminal progenitor-derived organoid: representative high-power image of nuclei stained with DAPI (blue). Scale bar: 50 µm. **(E)** Confocal microscopy on branched organoid during elongation phase on day 9. Shown is a

representative picture of Ki-67 expression in nuclei (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. *n*=30 structures/condition; 3 donors; 10 structures/donor.

5.2.3 Luminal progenitor-derived branched ductal organoids exhibit characteristics of low-grade carcinomas

I set out to further characterize polarization of the luminal progenitor-derived organoids. To this end, I examined the expression of the polarization markers laminin, which is a component of the basement membrane, and filamentous actin (F-actin). Thereby, I found that the basolateral sides of organoid ducts were partially covered by laminin (71% of all ducts) (Fig. 13A,C) (Ganz et al., 2021). In those cases, laminin deposition was co-localized with laminin receptor integrin α-6 (Fig. 13A) (Ganz et al., 2021). Furthermore, at apical locations, I observed an accumulation of F-actin framing the arising lumen (83% of all ducts) (Fig. 13B,C) (Ganz et al., 2021). In combination with the correct localization of luminal-specific markers (Fig. 10E,F) (Ganz et al., 2021), this demonstrates that the high degree of duct formation is accompanied by a pronounced apical-basal polarity. As mentioned above, the high degree of duct formation observed in the luminal progenitor-derived organoids was reminiscent of the morphology of well differentiated, low-grade NST carcinomas. Notably, the degree of differentiation and therefore formation of correctly polarized ducts, is just one hallmark of low-grade carcinomas. In addition, the tumor grade is determined by the presence or absence of nuclear pleomorphisms and the mitotic count.

The nuclei within branched organoids had a small, overall uniform size and appeared homogenous with indistinct nucleoli, representing the lowest degree of nuclear pleomorphism, which is in line with the classification as low-grade carcinoma based on differentiation (Fig. 13D). For the determination of the mitotic count for grading of patient samples, the number of actively dividing cells within a specified area of the tissue section is determined. Within the elongating organoids, Ki-67 staining revealed that a vast majority of the cells were in an actively dividing state (Fig. 13E). Of note, due to the strikingly different total number of cells within one tissue section as compared to an organoid, these observations cannot be compared quantitatively. However, the high proliferative activity with almost every single cell actively dividing in the organoids would most likely fit to a higher-grade carcinoma characterized by high mitotic count.

Taken together, these observations reveal that organoids generated from luminal progenitor cells upon transplantation into collagen type I gels, maintain luminal lineage markers. Moreover, the mature organoids exhibit apical-basal polarity and are characterized by

normal-like nuclei morphology. Consequently, the organoids arising in this assay largely resemble low-grade NST carcinomas with respect to morphology.

5.2.4 Lack of basal marker expression in luminal progenitor-derived branched ductal organoids

Invasive carcinomas are typically negative for a majority of basal lineage markers such as p63 (Zaha, 2014) (Fig. 12B) (Ganz et al., 2021). Therefore, I hypothesized that single luminal progenitor cells could form NST-like structures without the acquisition of basal markers. However, invasive behavior *in vitro* has previously never been observed within purely luminal mammary epithelial cells. Instead, primary human mammary cells that showed invasive capacity *in vitro* were either of basal origin (Buchmann et al., 2021) or were luminal cells that had converted towards a basal phenotype (Cheung et al., 2013). Therefore, I wanted to investigate whether bipotency was triggered within luminal progenitors upon transplantation of single cells and consecutive branched organoid formation in collagen type I gels as described above for the basal mammary epithelial cells. To this end, I used immunofluorescence and confocal microscopy to assess expression of basal lineage markers αSMA and p63. As a control I used basal cell-derived branched structures obtained within the same assay that reliably express basal lineages markers as outlined above (Fig. 9B,C) (Ganz et al., 2021). Doing so, I found that neither αSMA nor p63 were expressed within luminal progenitor-derived branched organoids, supporting the previous notion that luminal cell identity is maintained within these organoids (Fig. 14A,B) (Ganz et al., 2021). Finally, I examined the expression of vimentin and E-cadherin. As mentioned above, while E-cadherin is an epithelial marker, vimentin is associated with myoepithelial cell state. Thereby, besides being expressed within the basal subset of the breast, vimentin is also a marker for the epithelial-mesenchymal transition, a cellular program employed by epithelial cells during tumorigenesis (Kalluri & Weinberg, 2009). Thereby, recent advances in the field have put a spotlight on the acquisition of a hybrid phenotype by epithelial cells. In detail, rather than a full transition towards a mesenchymal phenotype, invasive cells have been described to exhibit both, epithelial and mesenchymal characteristics (Dongre & Weinberg, 2019). In line with this notion, I found epithelial marker E-cadherin as well as mesenchymal marker vimentin to be expressed within luminal progenitor-derived organoids (Fig. 14C) (Ganz et al., 2021). Thereby, the expression of vimentin was mainly localized towards the tips of extending ducts (Fig. 14C). Notably, vimentin as well as E-cadherin are both frequently expressed in NST carcinomas. However, while expression of E-cadherin is found in almost all cases (Gamallo et al., 1993), vimentin

is only expressed in a subset of NST carcinomas (Domagala et al., 1990; Korsching et al., 2005).

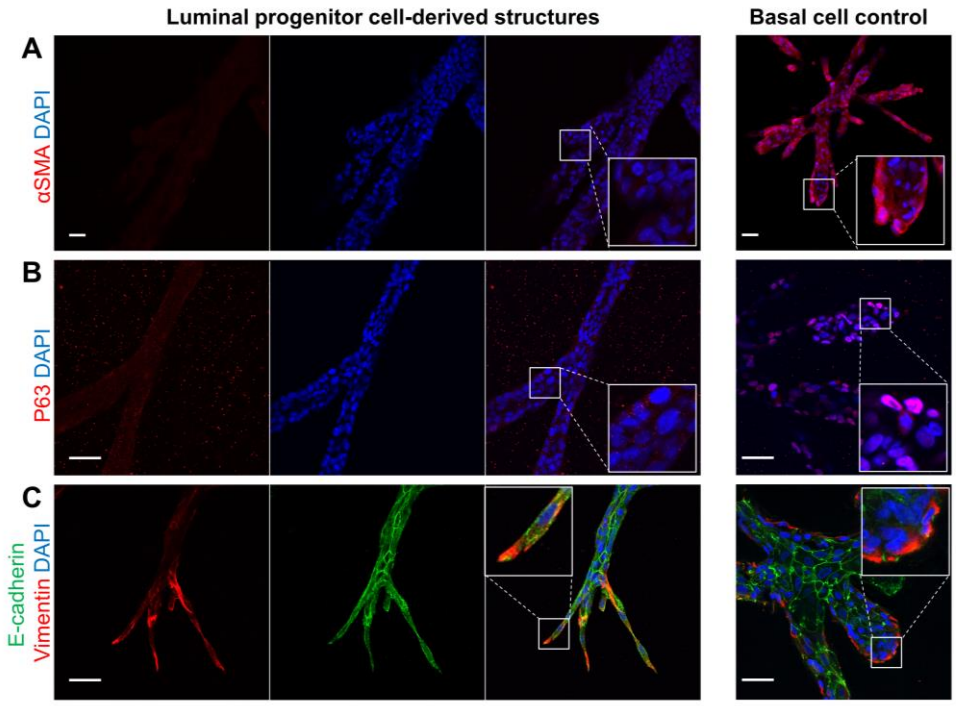

Figure 14. Basal marker expression in luminal progenitor-derived branched ductal organoids. (A-C) Confocal microscopy on fully grown branched luminal progenitor- or basal cell-derived organoids. Organoids were all derived within BLOM medium. Representative images of basal markers αSMA (red), p63 (red), Vimentin (red) and luminal marker E-cadherin (green) are shown. Nuclei are stained with DAPI (blue). *n*=30 structures/condition; 3 donors; 10 structures/donor. Scale bar: 50 µm.

In summary, the overall absence of basal lineage markers within luminal progenitor-derived organoids further corroborates the similarity to NST carcinomas. However, the expression of vimentin might indicate a certain ambivalence of epithelial cell identity.

5.3 Luminal progenitor-derived branched ductal organoid formation requires inhibition of ROCK-myosin II signaling

5.3.1 Contribution of single BLOM compounds to branched organoid formation from luminal progenitor cells

The observation that luminal progenitors have branched organoid-forming capacity is in drastic contrast to the properties previously attributed to this subset. In former reports, luminal progenitor cells merely formed spheres and partly showed reverted polarity in collagen type I gels (Gudjonsson, Rønnov-Jessen, et al., 2002; Linnemann et al., 2015).

While differences in overall proliferative capacity due to culture conditions might be one factor explaining this extensive divergence, I aimed to disclose whether and how the single components of the BLOM medium are connected to branched organoid forming capacity.

To this end, I reduced the composition of the BLOM medium to a minimum of components relevant for cell survival and proliferation. Within this medium, that I termed 'Basic medium', branched organoid formation was almost completely absent. In detail, while in control conditions, 53% of cells grew into branched structures, only around 1% did so in Basic medium (Fig. 15A). Instead, within Basic medium 88% of cells grew out into spheres and 11% into sticks (Fig. 15A). As a next step, I added each remaining BLOM component individually to the Basic medium. Strikingly, I found that the addition of Y-27632 to the Basic medium alone restored 76% of the branched structure formation capacity (Fig. 15B,C). Notably, none of the other components had comparable impact (Fig. 15B).

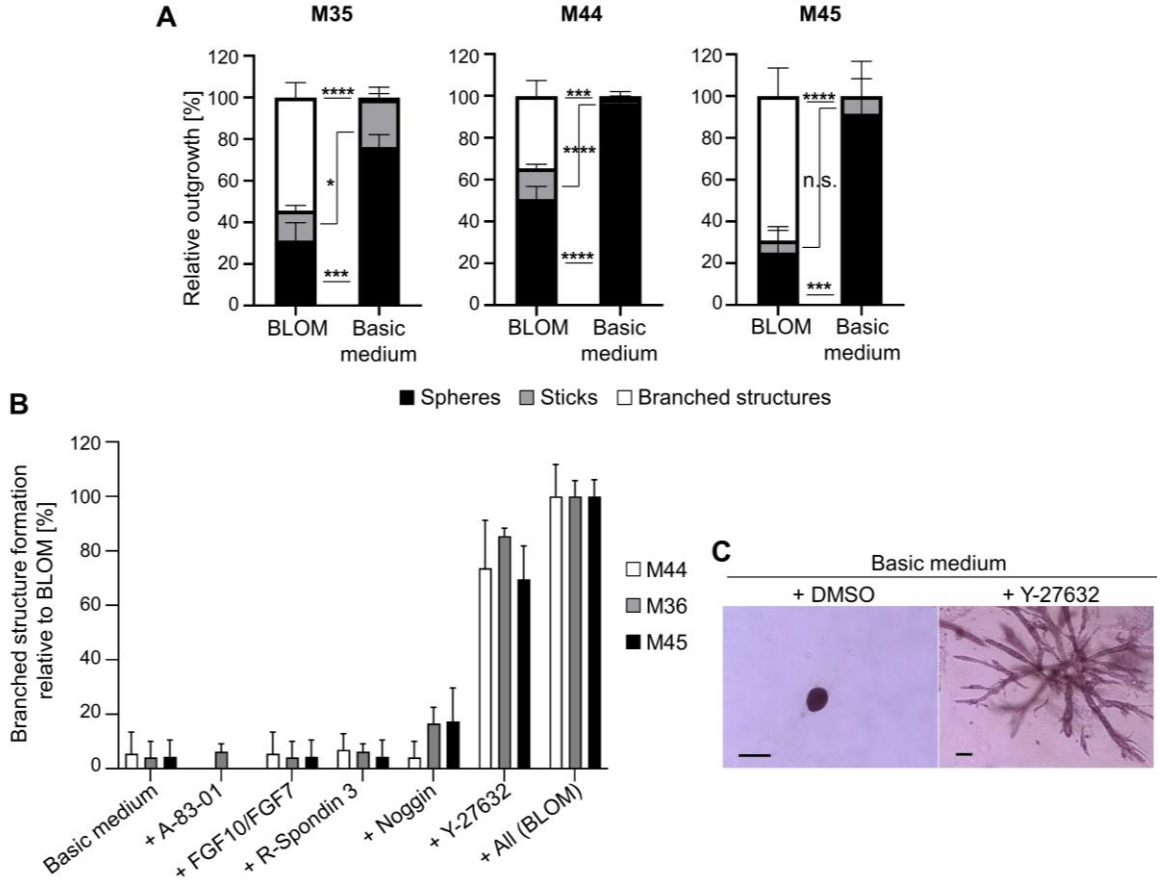

Figure 15. Contribution of single BLOM compounds to branched structure formation from luminal progenitors. (A) Quantification of outgrowing structure type in BLOM compared to Basic medium. Depicted are donors M35, M44 and M45. Structure type was normalized to total structures arising per gel and data are presented as stacked bars. Data are mean ± s.d. *n*=4 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, *P ≤ 0.05, ***P ≤ 0.001, ****P ≤ 0.0001. **(B)** Quantification of branched structure arising in Basic medium with various BLOM additives in comparison to full medium (BLOM). The number of branched structures was

normalized to branched structures arising in BLOM medium. Depicted are donors M36, M44 and M45. Data are mean ± s.d. *n*=4 gels/condition. **(C)** Carmine stainings of outgrowth from sorted luminal progenitor cells in Basic medium with DMSO control vs. Basic medium with addition of Y-27632 after 10 days. Scale bars: 100 µm. Experiments were repeated in *n*=3 donors.

5.3.2 Loss of branched organoid forming capacity in the absence of ROCK inhibitor

Y-27632 is an inhibitor of the Rho-associated protein kinase (ROCK). ROCK is a mediator of the small GTPase RhoA, which regulates cellular functions related to the cytoskeleton such as contractility. Consequently, Y-27632 is experimentally frequently employed to modulate the impact of loss of cellular contractility. Interestingly, previous publications have demonstrated that reduction of RhoA-ROCK-myosin II signaling via inhibition of ROCK can foster epithelial cell invasion into collagen type I gels (Zhou & Kramer, 2005). Notably, in previously described setups where luminal progenitor cells within collagen type I matrices did not show branched structure forming or invasive capacity, ROCK inhibitor was never a constant component of the culture medium. I set out to confirm the function of Y-27632 by monitoring branched organoid formation in BLOM in the absence of this inhibitor. As suspected, I found that depriving the medium of Y-27632 strongly reduced branched organoid formation. In detail, the formation of branched organoids decreased by 79% compared to BLOM conditions (Fig. 16A) (Ganz et al., 2021). Instead, cells mostly grew out into spheres (80% of all structures) (Fig. 16A) (Ganz et al., 2021). Furthermore, the morphology of the rarely emerging branched organoids resembled budding-like structures with a non-invasive appearance and were therefore strikingly different to the ones observed in BLOM containing Y-27632 (Fig. 16B) (Ganz et al., 2021). Quantitatively, branched organoids derived in the absence of Y-27632 were in average 66% smaller in size (Fig. 16C) and had reduced branching complexity (Fig. 16D) (Ganz et al., 2021). In detail, while 65% of control organoids showed tertiary or even more complex branching, only 6% of organoids derived in the absence of Y-27632 were this complex (Fig. 16D) (Ganz et al., 2021). I further confirmed the impact of ROCK inhibition on luminal progenitor branched organoid forming capacity by withdrawing Y-27632 during elongation phase. Thereby, I observed that withdrawal of ROCK inhibitor resulted in premature rounding-up of previously extending branches (Fig. 16E).

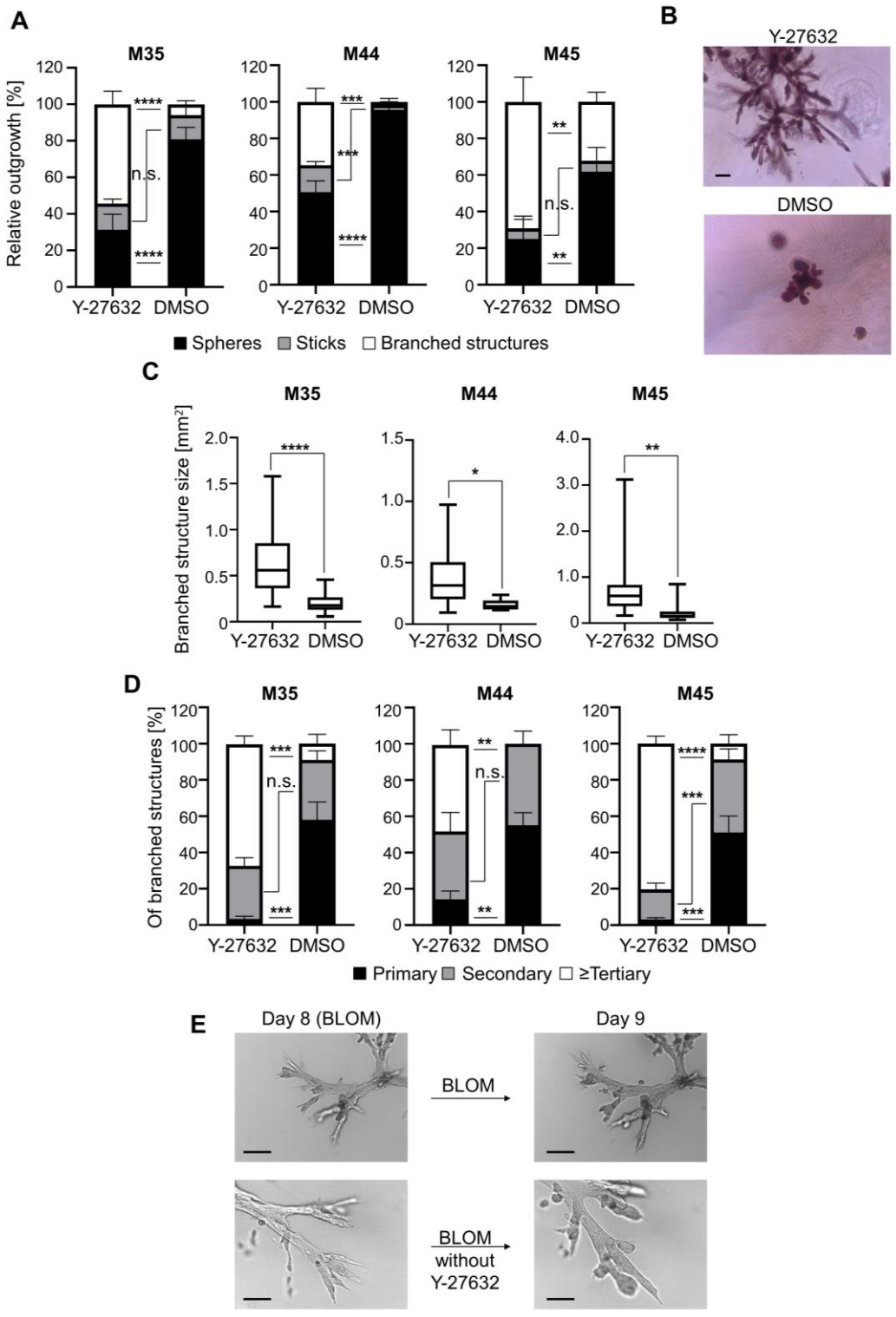

Figure 16. Loss of branched structure formation capacity in BLOM without ROCK inhibitor. **(A)** Quantification of outgrowing structure type in BLOM with Y-27632 or without (DMSO control). Depicted are donors M35, M44 and M45. Structure type was normalized to total structures arising per gel and data are presented as stacked bars. Data are mean ± s.d. *n*=4 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. **(B)** Carmine stainings of branched structures arising in BLOM with Y-27632 or without (DMSO control). Depicted is donor M45. Scale bar: 100 µm. **(C)** Quantification of branched structure sizes in BLOM with Y-27632 or without (DMSO control). Depicted are donors M35, M44 and M45.

Size was defined as the area covered by the organoid when all tips were connected virtually using ImageJ. Data are shown as median ± 25%. *n*=4 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, *P ≤ 0.05, **P ≤ 0.01, ****P ≤ 0.0001. **(D)** Quantification of branched structure complexity in BLOM with Y-27632 or without (DMSO control). Depicted are donors M35, M44 and M45. Branched structure types were normalized to total branched structures arising per gel and data are presented as stacked bars. Data are mean ± s.d. *n*=3 gels/condition. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, **P ≤ 0.01, ***P ≤ 0.001, ****P ≤ 0.0001. **(E)** Light microscopy pictures of structure during elongation phase in BLOM conditions compared to conditions where Y-27632 was withdrawn. Shown is the development from day 8 to day 9. Depicted is donor M51. Scale bars: 100 µm. Experiments were repeated in *n*=3 donors.

Taken together, these results show that the ROCK inhibitor Y-27632 is essential for the formation of large and complex branched, ductal organoids from luminal progenitor cells in this assay. Building on the assumption that branched organoid formation is the result of luminal cell invasion, these observations are in line with the previously described invasion-promoting role of ROCK on epithelial cells.

5.3.3 ROCK inhibition is required for correct polarization of branched ductal organoids

In addition to its role for invasion, inhibition of ROCK has been described as a prerequisite for correct apical-basal polarization of epithelial structures in collagen type I gels (Yu et al., 2008). In line with this notion, I found a loss of apical-basal polarization in luminal progenitor structures derived in the absence of Y-27632. Specifically, immunostaining and confocal microscopy revealed a lack of lumen formation in the respective organoids (Fig. 17 A,B) (Ganz et al., 2021). Furthermore, I found neither F-actin accumulation at apical positions nor laminin deposition at the basolateral sides as characteristic for organoids derived in the presence of Y-27632 (Fig. 17 A,B) (Ganz et al., 2021). Consequently, it can be assumed that the compound Y-27632 is an indispensable factor to facilitate apical-basal polarization within luminal progenitor-derived organoids in this assay.

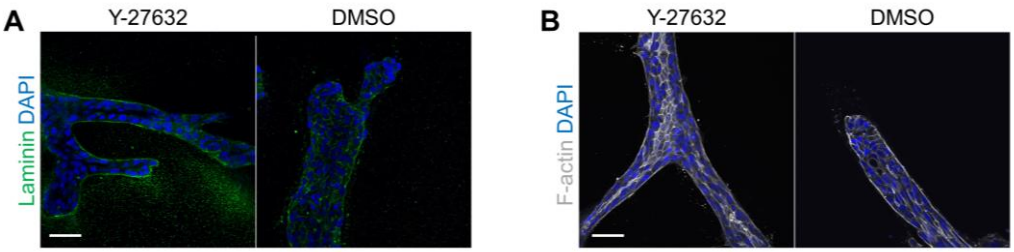

Figure 17. ROCK inhibition is required for correct polarization of branched ductal structures. **(A/B)** Confocal microscopy on fully grown branched organoids arising in BLOM with Y-27632 or without (DMSO control). Shown are representative images of polarization markers Laminin (green) and F-actin (white). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. Experiments were repeated in *n*=3 donors.

5.3.4 Cellular dynamics depend on ROCK inhibition

Next, I set out to assess the dynamics of branched organoid formation from luminal progenitor cells to unravel the dependency on ROCK inhibition. Specifically, I wanted to understand the role of Y-27632 in inter-cellular dynamics as well as cell-matrix interactions. To this end, live-cell confocal microscopy was performed during the organoid elongation phase. Thereby, cellular movements within luminal progenitor-derived organoids were monitored via nuclear labelling in the presence and absence of ROCK inhibitor. These measurements revealed that under the influence of ROCK inhibitor, the cellular velocity and therefore migration was preferably directed outwards along the axis of the extending duct (V_{II}) (Fig. 18A). However, without ROCK inhibition, this directionality was significantly decreased. Here, the parallel (V_{II}) and orthogonal (V_I) velocity, which is directed perpendicular to the duct axis, were more similar, resulting in a less directed movement (Fig. 18A) (Ganz et al., 2021). Next, live-cell imaging of organoids was performed in collagen type I gels that were spiked with fluorescent tracer particles/beads. Thereby, interactions with the surrounding ECM could be visualized. By that, it appeared that organoids derived in the presence of ROCK inhibitor, exhibited bead displacement towards the organoids in front of elongating branches (Fig. 18B) (Ganz et al., 2021). As this bead placement is indicative for contractility, this means that despite ROCK inhibition, weak contractility is existent. In contrast, organoids derived in the absence of ROCK inhibitor, showed a pronounced and generalized bead displacement towards the organoid which suggests high contractility (Fig. 18C) (Ganz et al., 2021).

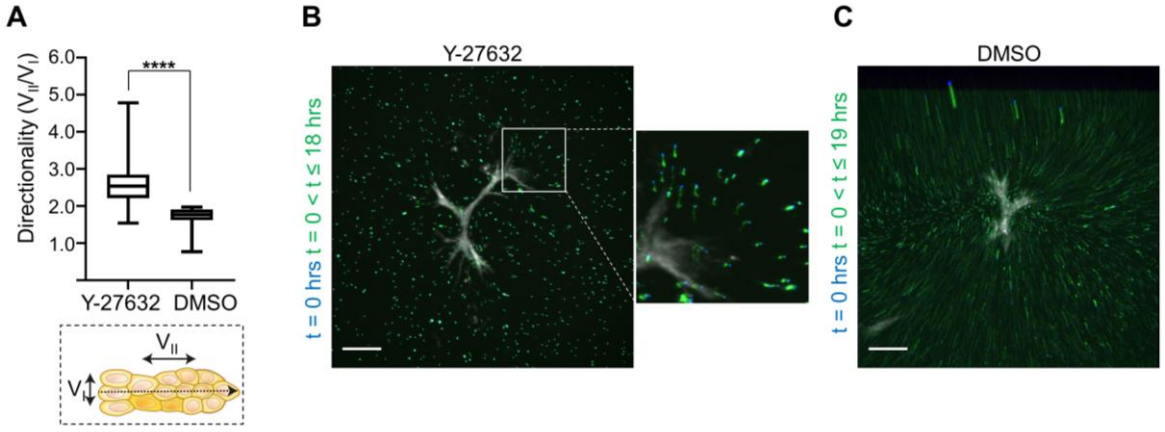

Figure 18. Cellular dynamics depend on ROCK inhibition. (A) Top: Directionality of cellular movement calculated as quotient of velocity directed in parallel to extending duct (V_{II}) and orthogonal to it (V_I). Data are shown as median ± 25%. P value was calculated using an unpaired two-tailed *t*-test, ****P ≤ 0.0001. Bottom: Schematic depiction of elongating duct. **(B/C)** Displacement of fluorescent beads relative to t=0 measured by live cell imaging with Y-27632 or without (DMSO control). T=0 hrs (blue) vs. t=0 < t ≤ 18/19 hrs (green). Scale bar: 200 µm. Experiments were repeated

in *n*=3 donors with a total of individual organoids *n*≥30. Live-cell imaging and associated measurements were performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich).

In summary, these observations reveal that reduction of the ROCK signaling via addition of Y-27632 ensures reduced and directed contractility, thereby enabling migration of normal luminal progenitor cells into collagen type I gels.

5.3.5 ROCK inhibition can be replaced by myosin II inhibition

Besides the impact of the ROCK inhibitor on actomyosin-related processes, this inhibitor has previously been attributed a role in increasing survival and proliferation of cells in 3D culture (Linnemann et al., 2015; Sato et al., 2009; Zhang et al., 2011). Accordingly, the above-described reduction of branched organoid sizes and complexity within ROCK inhibitor-deficient conditions could be indicative of a survival and/or proliferation increasing effect of Y-27632 in this system. To exclude the possibility that branched organoid formation observed upon addition of Y-27632 is merely a result of increased proliferative capacity, I set out to target a different segment of the ROCK-myosin II pathway. In detail, I replaced Y-27632 by para-amino blebbistatin (PAB), an inhibitor of the ROCK downstream target myosin II (Fig. 19A). In fact, upon addition of PAB, branched ductal organoids appeared that were morphologically similar to the BLOM phenotype providing further support for the involvement of the ROCK-myosin II pathway in branched organoid formation (Fig. 19B) (Ganz et al., 2021). However, the quantities of branched organoids derived in the presence of PAB were reduced as compared to conditions employing Y-27632 (Fig. 19C) (Ganz et al., 2021).

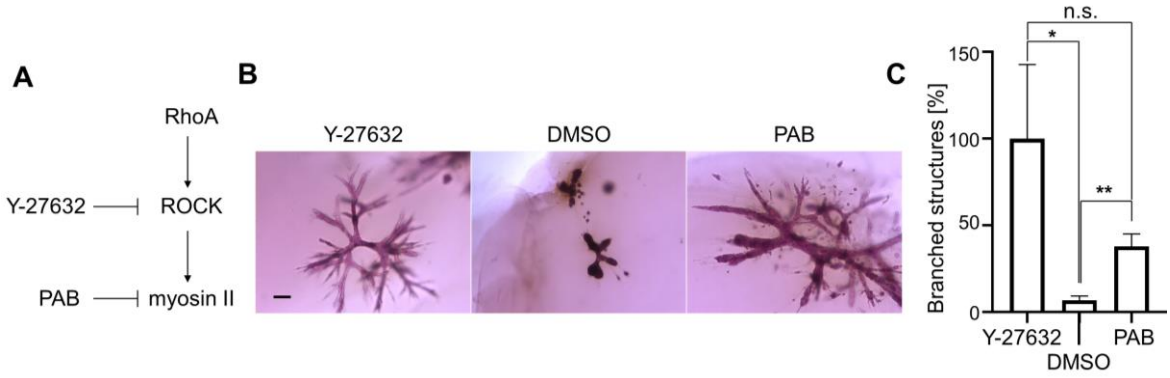

Figure 19. ROCK inhibition can be replaced by myosin II inhibition. (A) Schematic depiction of RhoA-ROCK-myosin II pathway and inhibitors. **(B)** Carmine stainings of branched structures arising in BLOM with Y-27632 or without Y-27632. In conditions without Y-27632, either DMSO or PAB was added. Scale bar: 100 µm. **(C)** Quantification of branched structures in BLOM with Y-27632 or without Y-27632. In conditions without Y-27632, either DMSO or PAB was added. Branched structure formation was normalized to BLOM with Y-27632. Data are mean ± s.d. *n*=3 independent

experiments with donors M16, M28 and M44. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05, *P ≤ 0.05, **P ≤ 0.01.

All in all, these data confirm that inhibition of the ROCK-myosin II pathway is required for branched structure formation from luminal progenitor cells and that the crucial role of Y-27632 is not merely an effect of increased survival or proliferative capacity.

5.4 Luminal leader cells guide collective invasion via matrix remodeling

5.4.1 ECM remodeling reveals active cellular invasion

I hypothesized that to undergo branching morphogenesis resulting in the observed low-grade carcinoma morphology, luminal progenitor cells in BLOM conditions need to invade the ECM actively and collectively. Of note, one important factor of active cellular invasion is the remodeling of the surrounding ECM. During cancer progression in the mammary gland, active epithelial cell invasion has been reported as a result of the interplay between the ECM with stromal and epithelial cells (Winkler et al., 2020). As the assay described here is free of stromal cells, all invasion relevant processes must be performed by luminal progenitor cells in cooperation with the collagen type I ECM surrogate. Based on the engagement with the ECM during branched organoid formation visualized by bead displacement as described above (Fig. 18B), I suspected that normal luminal progenitor cells possess matrix remodeling capacity in this assay. Consequently, the focus was first put on matrix topography to unravel invasive mechanisms exploited by luminal progenitor cells. In the normal mammary gland, the epithelium is typically surrounded by anisotropic, curl-like collagen fibers (Provenzano et al., 2006). However, those collagen fibers are linearized upon tumor initiation. In return, these aligned fibers have been shown to guide and promote migration and invasion of cancerous cells (Provenzano et al., 2006). Previous reports have shown that such alignment can be facilitated by fibroblasts or invading epithelial cells themselves (Levental et al., 2009; Provenzano et al., 2006; Provenzano, Inman, Eliceiri, Trier, et al., 2008). In this work, fluorescently labelled collagen type I was used to visualize fiber orientation. Thereby, it appeared that in the cell-free periphery of the ducts, collagen fibers exhibited no preferred orientation (Fig. 20A) (Ganz et al., 2021). In contrast, collagen fibers were aligned right in from of invading ducts (Fig. 20A) (Ganz et al., 2021). In detail, the degree of alignment was significantly increased in front of invading ducts by 132% as compared to the periphery (Fig. 20B) (Ganz et al., 2021). Of note, it is well described that cellular contractility via ROCK is required for such fiber alignment

(Brownfield et al., 2013; Provenzano, Inman, Eliceiri, Trier, et al., 2008). Along with the above-described bead displacement in front of invading ducts, this supported the concept that even under influence of ROCK inhibitor, contractility was not fully abolished. Rather, it appeared that ROCK-mediated contractions were reduced to a level that allowed rearrangement of fibers in a controlled manner, and thus a directed movement of invading cells.

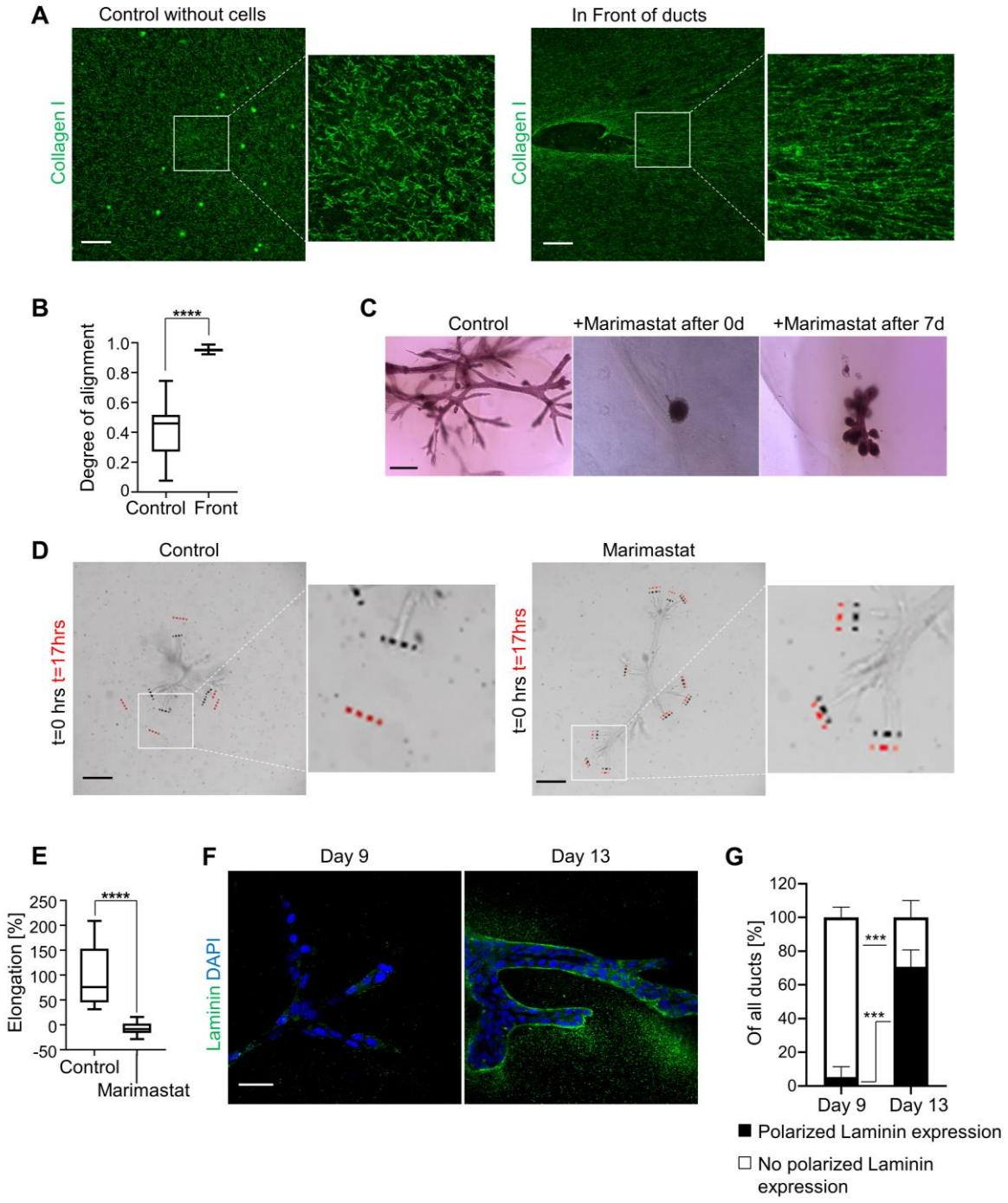

Figure 20. Active collective invasion is facilitated by matrix remodeling. (A) Alignment of collagen type I fibers in control sections of the gel without cells vs. in front of invading ducts. Scale

bar: 50 µm. *n*=22 individual organoids. **(B)** Quantification of fiber alignment degree in control sections of the gel without cells vs. in front of invading ducts. Data are shown as median ± 25%. P value was calculated using an unpaired two-tailed *t*-test, ****P ≤ 0.0001. *n*=22 individual organoids. **(C)** Carmine stainings of branched structures arising in BLOM upon treatment with Marimastat at day 0 or day 7 as compared to untreated control. Pictures were taken on day 10. Scale bar: 100 µm. **(D)** Duct elongation upon treatment with Marimastat vs. control organoids measured by live-cell imaging with t=0 hrs (black) vs. t=17 hrs (red). Scale bar: 200 µm. *n*=18 individual organoids. **(E)** Quantification of duct elongation upon treatment with Marimastat vs. control organoids tracked by live-cell imaging. Data are shown as median ± 25%. P value was calculated using an unpaired two-tailed *t*-test, ****P ≤ 0.0001. *n*=18 individual organoids. All experiments were repeated in at least *n*=3 donors. **(F)** Confocal microscopy on branched luminal progenitor-derived organoids: representative images of Laminin (green), and DAPI (blue) on day 9 and day 13. *n*=30 structures/condition; 3 donors; 10 structures/donor. Scale bar: 50 µm. **(G)** Quantification of ducts with and without expression of correctly polarized Laminin on day 9 as compared to day 13. Data are presented as stacked bars with mean ± s.d. *n*=3 donors, 25 ducts per donor. **(A)** and live-cell imaging in **(B)** as well as associated measurements were performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich).

In parallel to fiber alignment, during tumorigenesis, actively invading cells can clear a path for invasion by degrading the matrix. This degradation is driven by matrix-metalloproteinases (MMPs), which can be secreted by stromal cells such as fibroblasts as well as epithelial cells (Friedl & Wolf, 2008; Kessenbrock et al., 2010). To unravel the contribution of MMPs to branched organoid formation in this assay, I treated freshly seeded luminal progenitor cells with the MMP inhibitor Marimastat. Upon MMP inhibition during the establishment phase, branched organoid formation was completely abolished, and luminal progenitor cells grew into spheres (Fig. 20C). Furthermore, Marimastat treatment at a later timepoint of culture, during the elongation/branching phase, resulted in the formation of small, branched organoids with pronounced round structure tips (Fig. 20C). Both observations were indicative of a halt of invasive growth upon MMP inhibition while proliferative capacity appeared to be largely maintained. In line with these static observations, live-cell imaging revealed that Marimastat treatment during elongation phase prevented the invasion of ducts into the collagen type I gels (Fig. 20D,E) (Ganz et al., 2021). Of note, I had previously observed that the ducts were partly covered with laminin which called into question how invading cells could interact with the collagen during elongation phase. To address this question, I performed immunofluorescence staining for laminin during elongation (day 9) as compared to rounding-up phase (day 13). Doing so, I found that elongating ducts were, in contrast to fully grown ducts, not covered with laminin (Fig. 20F,G) (Ganz et al., 2021). This observation helps to explain the direct interaction of invading cells with the collagenous ECM during elongation.

Taken together, these observations show that during the generation of branched, ductal organoids in the presence of ROCK inhibitor, the collagenous matrix is remodeled. As the only cellular component of the system are the luminal progenitors, this demonstrates that

luminal progenitor cells with balanced contractile properties remodel the ECM to invade actively.

5.4.2 Luminal leader cells guide invasive branching morphogenesis

In this model, cellular invasion is accompanied by a branching morphogenesis process whereby the organoids increase in complexity continuously during invasion. Notably, it has been described that invasive branching morphogenesis is typically guided by a single leader cell (Varner & Nelson, 2014). Since fiber alignment and matrix degradation had been observed exclusively in front of the tips of invading ducts, I suspected that matrix remodeling in this assay might be facilitated by luminal leader cells guiding the ducts. Matching this assumption, a single cell was observed at the tip of each duct during elongation phase. This cell was characterized by an extended morphology with filopodia-like membrane protrusions (Fig. 21A) (Ganz et al., 2021). In line with that, live-cell imaging during elongation phase showed that there was typically one cell at the tip of each invading duct, which remained at the front position over several hours and continuously grew outwards while staying connected to the stalk cells of the remaining duct (Fig. 21B) (Ganz et al., 2021). However, from time to time leader cells were exchanged with one of the stalk cells that had previously followed the leader cell. In that case, one of the stalk cells took over the front position and the former leader cell fell back into the stalk cell pool (Fig. 21C) (Ganz et al., 2021). As described above, the elongation phase observed here was characterized by steadily rising complexity of the branched organoid. In line with the proposed role of leader cells for guiding the extending branches, live-cell imaging showed that new ducts developed mainly via bifurcation of a tip (21 out of 26 events, organoids from 4 different donors) (Ganz et al., 2021). Thereby, two neighboring leader cells with disparately oriented membrane protrusions were present at the duct's tip and gave rise to two distinct sub-branches, which highlights the role of leader cells in directing invasive branching morphogenesis (Fig. 21D) (Ganz et al., 2021). Rarely, the formation of new ducts from the structure body was observed, which supports the notion that stalk cells can convert to leader cells spontaneously (5 out of 26 events, 4 donors) (Ganz et al., 2021).

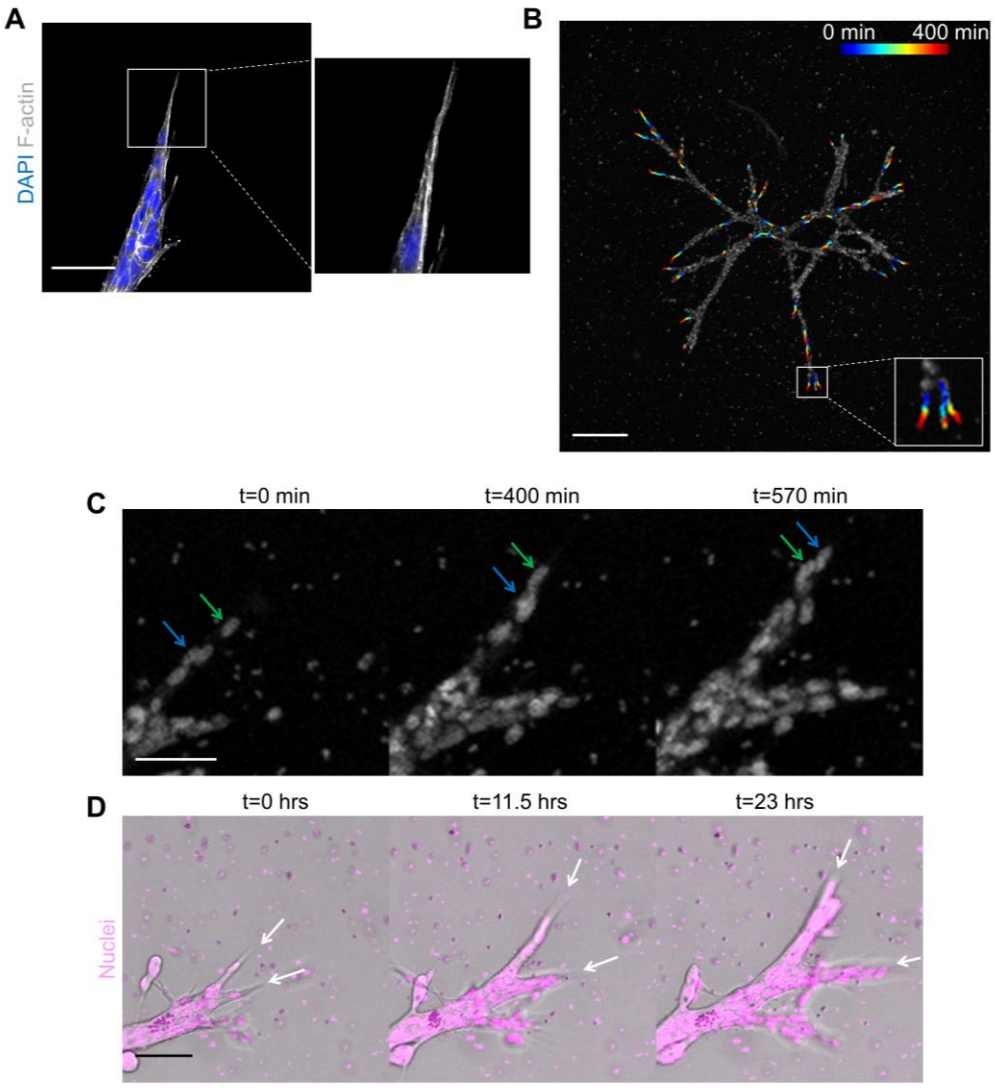

Figure 21. Leader cells guide collective invasion. (A) Confocal microscopy on leader cell. Shown is a representative picture of a F-actin (white) staining. Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(B)** Movement of leader cells over the course of 0-400 minutes measured by live-cell imaging via nuclear labelling with sirDNA (white). Scale bar: 200 µm. **(C)** Leader cell exchange. Cells tracked by live-cell imaging. Leader cell 1 (green arrow), leader cell 2 (blue arrow). Organoids measured by live-cell imaging via nuclear labelling with sirDNA (white). Scale bar: 50 µm. **(D)** Bifurcation at tip. Organoids measured by live-cell imaging via nuclear labelling with sirDNA (pink). Scale bar: 80 µm. All experiments were repeated in at least *n*=3 donors. Live-cell imaging as well as associated measurements in **(B**-**D)** were performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich).

In summary, it appears that certain luminal cells take over leader cell function and guide invasive branching morphogenesis in the organoids. Of note, while leader cells execute a specialized function during their time at the leading edge, they are not per se specialized as they can be replaced by other luminal cells.

5.4.3 E-cadherin ensures cell-cell connectivity during collective invasion

It has previously been described that leader cells generate forces that are transferred onto the epithelial stalk cells to allow collective invasion of whole ducts (Friedl & Gilmour, 2009). Thereby, strong cadherin-mediated cell-cell junctions play a critical role (Friedl & Gilmour, 2009). In line with this notion, I found that the luminal progenitor-derived organoids displayed strong E-cadherin expression between leader and stalk cells (Fig. 22A) as well as among stalk cells in the cellular body (Fig. 22A) (Ganz et al., 2021). The role of E-cadherin for mediating collective invasion was unraveled by inhibiting E-cadherin in structures during elongation phase with the function blocking antibody HECD1. Thereby, consecutive live-cell imaging revealed dissolution of invasive branching morphogenesis upon E-cadherin blockage (Fig. 22B) (Ganz et al., 2021). In detail, upon HECD1 inhibition, single cells and even full ducts lost contact to the structure body. As a result, cells were scattered around the structure whereby duct coherence was subsequently lost (Fig. 22C) (Ganz et al., 2021).

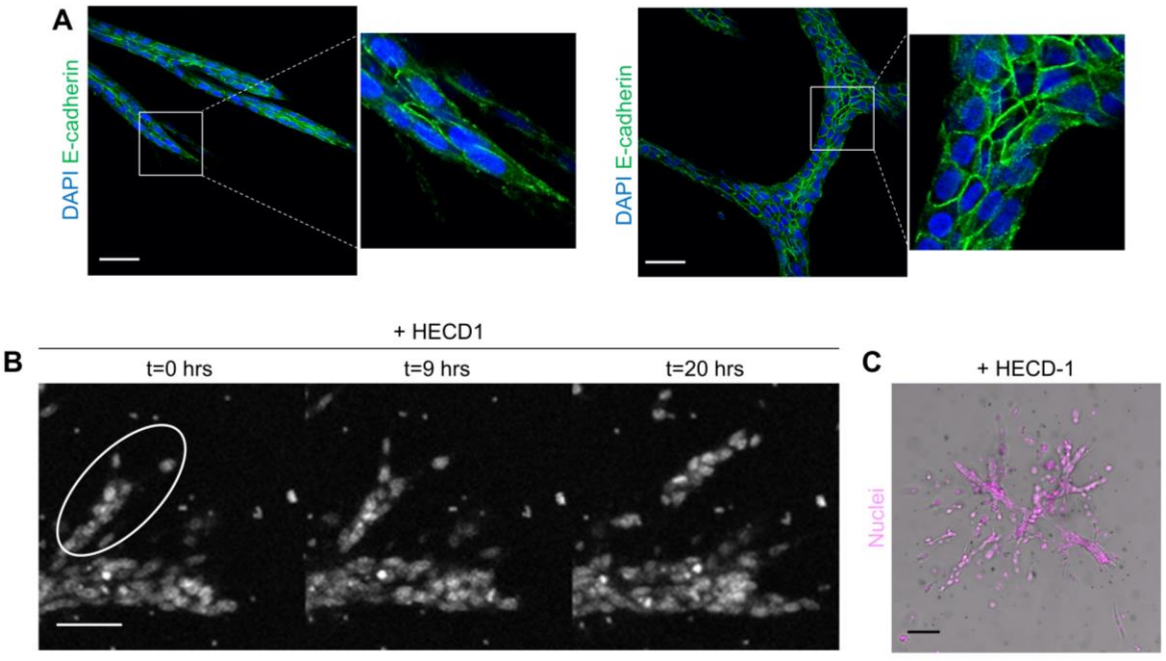

Figure 22. E-cadherin ensures cell-cell connectivity during collective invasion. (A) Confocal microscopy on invading organoid. Shown are representative images of cell junction marker E-cadherin (green). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(B)** Duct elongation of organoid (day 8) upon addition of HECD1 antibody on day 5. Organoids measured by live-cell imaging over the course of 20 hrs via nuclear labelling with sirDNA (white). Scale bar: 50 µm. *n*=30 individual organoids**. (C)** Image of organoid on day 9 after HECD1 treatment from day 5 on. Nuclear labelling with sirDNA (pink). Scale bar: 100 µm. All experiments were repeated in *n*=3 donors. Live-cell imaging in **(B)** was performed by Benedikt Buchmann (Department of Biophysics, Technical University of Munich).

Taken together, these observations reveal a critical role of E-cadherin-mediated cell-cell junctions for maintaining duct integrity during collective invasion.

5.5 Editing of luminal progenitor cells with CRISPR-Cas9

5.5.1 Expansion of luminal progenitor cells in 2D and subsequent organoid formation

Research with primary human cells, particularly when it comes to the application of elaborate methods such as CRISPR-Cas9 gene editing, is often hampered by small amounts of starting material. In addition, it is a challenge to maintain and expand luminal progenitor cells in culture while maintaining their cellular identity as reflected in their original lineage marker expression. To investigate the suitability of the model described here for a wider range of applications that require maintaining and expanding luminal progenitor cells in culture, I investigated the potential to expand this subset in 2D culture focusing on proliferative capacity and correct lineage marker expression. Interestingly, I found that luminal progenitor cells did not attach on tissue culture treated plastic dishes (Fig. 23A). To enable luminal progenitor attachment in 2D, dishes required collagen type I coating (Fig. 23A). Expanding on this, I found that cells could be passaged at least five times without losing expression of luminal marker CK8/18 (Fig. 23B). A large share of luminal progenitor cells also maintained Ki-67 expression indicating proliferative activity, however, the share of Ki-67 positive cells started decreasing around passage 5 (Fig. 23B). Furthermore, luminal progenitor cells did not acquire basal lineage characteristics in culture as shown in absence of p63 expression (Fig. 23C). In contrast, basal control cells reliably expressed p63 while showing an absence of CK8/18 expression (Fig. 23D). Analysis of population doublings revealed that after an initial attachment phase, the luminal progenitor populations grew exponentially between passage 1 and passage 5 to 7 (Fig. 23E). After that, proliferation slowed down, which is in line with the decreasing expression of Ki-67 around this time. The cells finally became senescent after a total of 13 to 17 population doublings (Fig. 23E).

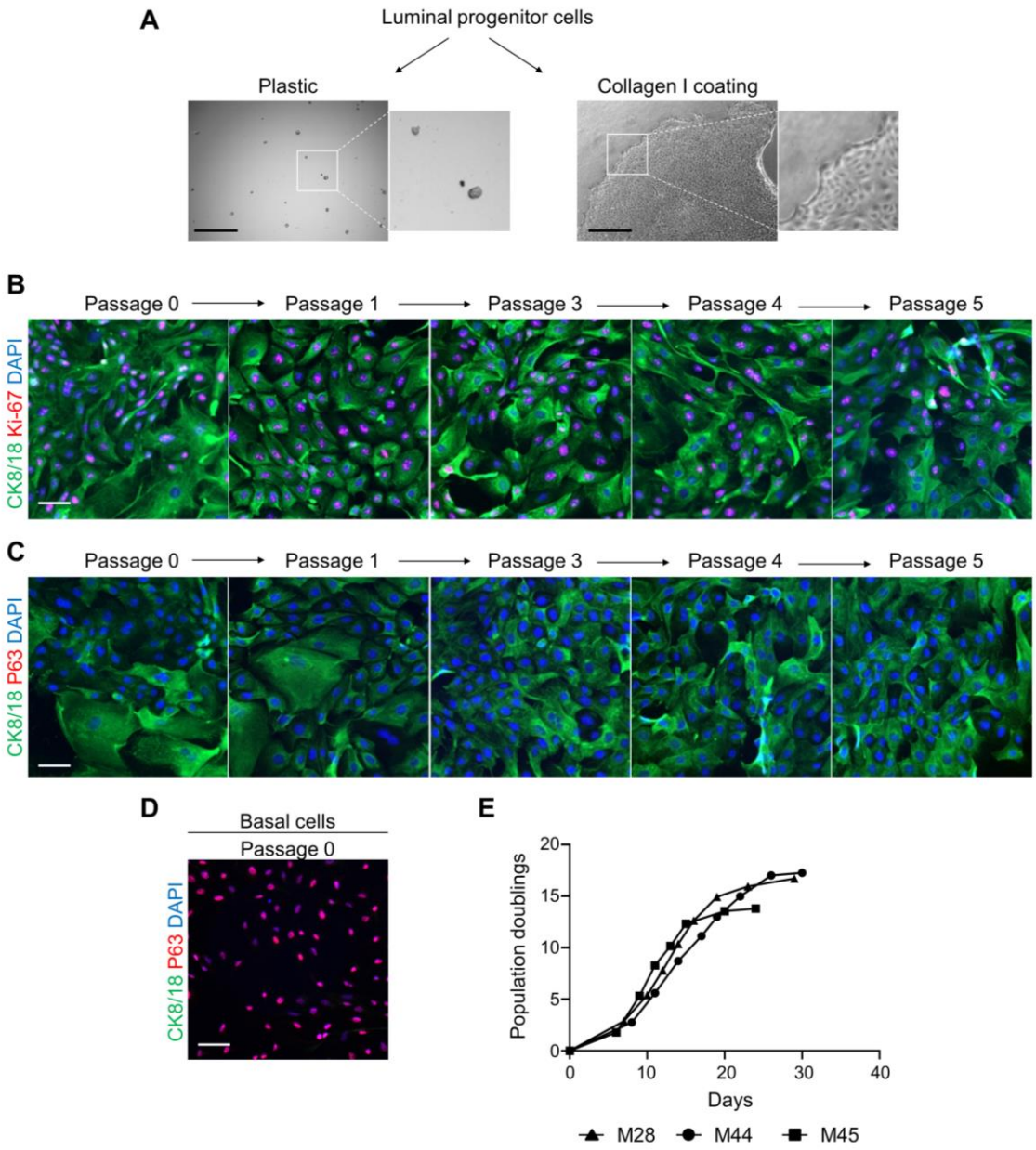

Figure 23. Expansion of luminal progenitor cells in 2D culture. (A) Light microscopy pictures of luminal progenitors at passage 0 in 2D culture grown in BCOM medium 7 days after FACS sorting. Cells were seeded either on cell culture-treated plastic or collagen type I-coated plastic. Scale bar: 500 µm. **(B)** Immunofluorescence of luminal progenitor cells grown in BCOM medium on collagen type I-coated plastic. Depicted are cells between passage 0 and passage 5. Cells are stained for CK8/18 (green) and Ki-67 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm **(C)** Immunofluorescence of luminal progenitor cells grown in BCOM medium on collagen type I-coated plastic. Depicted are cells between passage 0 and passage 5. Cells are stained for CK8/18 (green) and p63 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(D)** Immunofluorescence of basal cells grown in BCOM medium on collagen type I-coated plastic. Depicted are cells at passage 0. Cells are stained for CK8/18 (green) and p63 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. **(E)** Population doublings in 2D culture: Every dot represents a splitting event. Depicted are growth curves for M28, M44 and M45. Experiments were repeated in *n*=3 donors.

As the goal of the expansion was to generate a sufficient number of cells for the 3D organoid assay, I set out to examine whether branched organoid forming capacity was maintained

upon 2D expansion. To this end, luminal progenitor cells were propagated in 2D culture and subsequently transplanted into collagen type I gels. This revealed that branched ductal organoids could be generated from luminal progenitor cells even after three passages in 2D culture (Fig. 24) (Ganz et al., 2021).

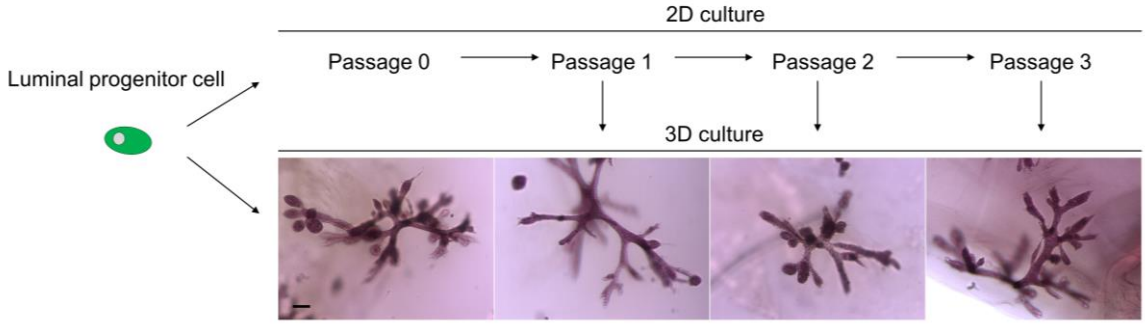

Figure 24. Organoid formation following 2D expansion. Carmine stainings of luminal progenitor cell-derived branched structures after 2D expansion for up to 3 passages. Scale bar: 100 µm. The experiment was repeated in *n*=3 donors.

In conclusion, the conditions presented here are suitable to expand luminal progenitor cells in 2D culture. Thereby, the luminal identity of the cells is maintained. Moreover, luminal progenitor cells maintain branched organoid forming capacity upon 2D culture. These observations increase the number of possible applications of the system presented here and therefore add to the value of the 3D model.

5.5.2 Tumor suppressor knock-out via CRISPR-Cas9

One application that requires maintenance of cells over time is genetic editing with the CRISPR-Cas9 system. As outlined above, the branched ductal organoids arising from normal luminal progenitor cells resemble low-grade NST carcinomas. However, genetic aberrations can have a strong impact on invasive morphology *in vivo*. Therefore, I set out to facilitate CRISPR-Cas9 mediated gene editing on the luminal progenitor cells. Moreover, I wanted to investigate whether the genetic impact on invasive morphology *in vivo* is recapitulated within the organoids. To achieve genetic modifications, I attempted to transiently transfect luminal progenitor cells in 2D with a gRNA-containing vector in parallel to a GFP-positive Cas9 plasmid. Next, I enriched for successfully transfected cells via FACS sorting for GFP positive cells (Fig. 25A) (Ganz et al., 2021). The assumption hereby was that if cells had taken up the larger Cas9-containing plasmid, they also had taken up the required gRNA vector. After the sort, luminal progenitor cells were either expanded for validating successful knock-out on the genetic level or seeded into 3D culture for subsequent knock-out organoid formation (Fig. 25A) (Ganz et al., 2021). In this approach

double-paired gRNAs were designed as described previously (Breunig et al., 2018) in order to ensure impactful modification of the genome.

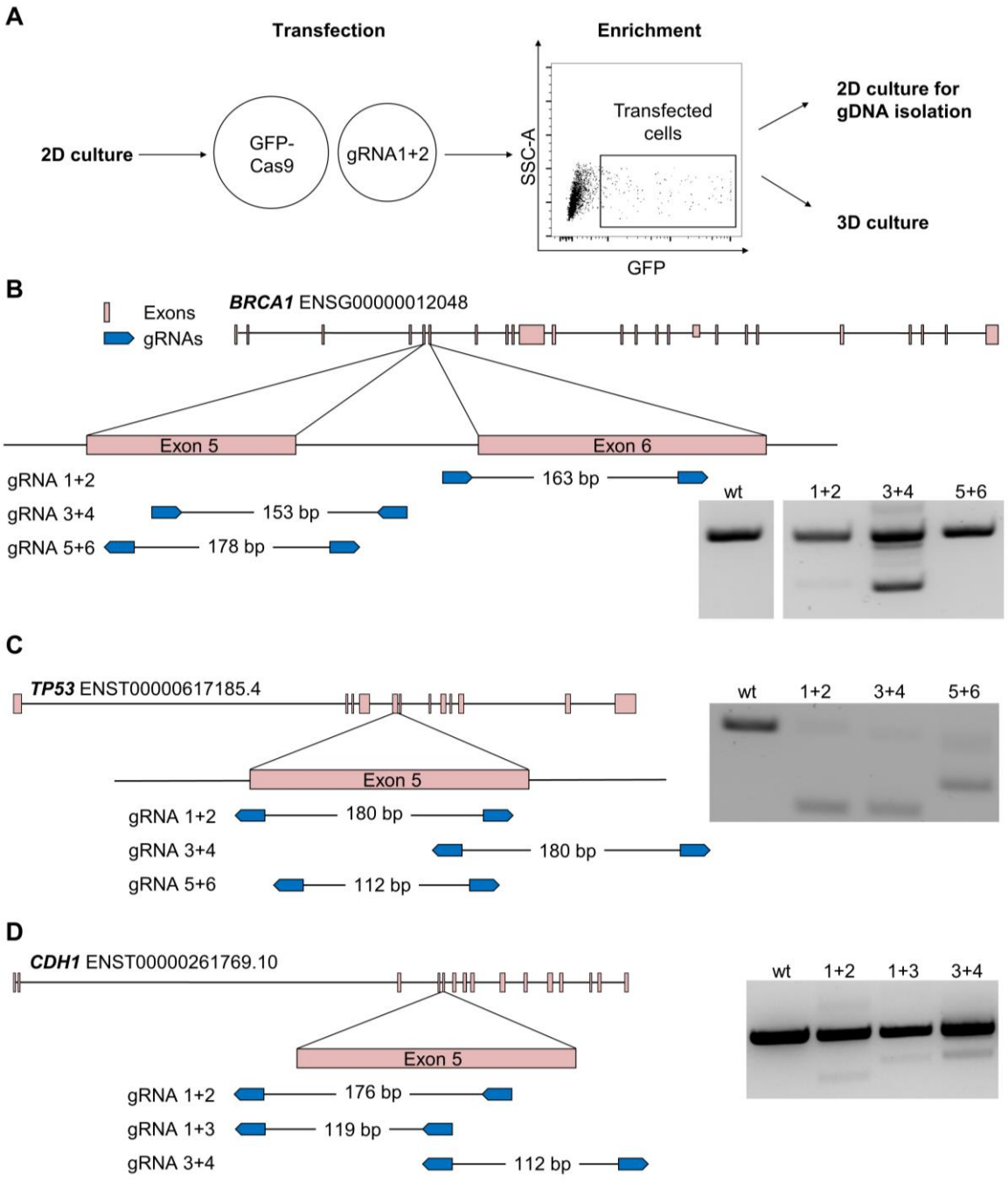

Figure 25. Tumor suppressor knock-out via CRISPR-Cas9. (A) General scheme for CRISPR-Cas9 modification after 2D expansion. **(B)** Sequence of BRCA1 encoding gene *BRCA1* and modification efficiency in the bulk population of transfected luminal progenitor cells on the genetic level vs. wild-type (wt) cells. Exons are depicted in pink. gRNAs are depicted in blue. Schematic target sites of the tested gRNAs on the gene are shown. **(C)** Sequence of p53 encoding gene *TP53* and modification efficiency in the bulk population of transfected luminal progenitor cells on the genetic level vs. wt cells. Exons are depicted in pink. gRNAs are depicted in blue. Schematic target sites of the tested gRNAs on the gene are shown. **(D)** Sequence of E-cadherin encoding gene *CDH1* and modification efficiency in the bulk population of transfected luminal progenitor cells on the genetic

level vs. wt cells. Exons are depicted in pink. gRNAs are depicted in blue. Schematic target sites of the tested gRNAs on the gene are shown. The experiment was repeated in *n*=3 donors.

Initially, I chose three different known tumor suppressor genes as potentially interesting targets: *BRCA1*, *TP53* and *CDH1*. For each gene, I cloned three constructs containing two gRNAs each. Thereby, gRNA pairs were designed to span a relevant region of an exon of the respective gene (Fig. 25B-D) (Ganz et al., 2021). Upon separate transfection with each construct, enrichment and 2D expansion, I performed PCR on the loci of interest. Doing so, I found that most gRNA pairs caused a full genetic deletion in a subset of the bulk population (Fig. 25B-D) (Ganz et al., 2021). The most efficient gRNA pair for each gene knock-out was chosen to perform 3D culture for branched organoid formation subsequent to the knock-out procedure. The gRNA pairs of choice were gRNA 3/4 for *BRCA1*, 5/6 for *TP53* and 3/4 for *CDH1*. As controls, I used luminal progenitors that were transfected with empty gRNA vector and GFP-positive Cas9 plasmid.

I was aiming for a read-out based on organoid morphology. Therefore, I first analyzed morphological differences using light microscopy. This revealed that for *TP53* as well as *BRCA1* knock-out, no aberrant morphologies appeared compared to empty-vector control gels (data not shown). However, in case of *CDH1*, I noticed obvious differences between empty-vector transfected and knock-out gels, therefore I focused the further analysis on the knock-out of this gene.

5.5.3 Deletion of E-cadherin results in ILC-like morphology

Due to the comparably low knock-out efficiency observed on the genetic level for *CDH1*, I assumed that gels contained a mixture of wild-type and knock-out cells on the protein level (Fig. 26A) (Ganz et al., 2021). I set out to investigate whether morphological aberrations observed with light microscopy were direct consequences of E-cadherin status. To this end, I used immunofluorescence in combination with confocal microscopy for investigating E-cadherin expression of the outgrowing structures. Thereby, I took advantage of the fact that cells are seeded as single cells in this assay and grow out clonally. Consequently, I could classify each structure (including spheres, sticks and branched organoids) separately based on E-cadherin status. Doing so, I found that knock-out gels contained a mixture of wild-type (E-cadherin positive) and knock-out (E-cadherin negative) structures, whereby around 25% were bearing the knock-out (Fig. 26B) (Ganz et al., 2021). Of note, in empty-vector control gels, no E-cadherin negative structures were observed.

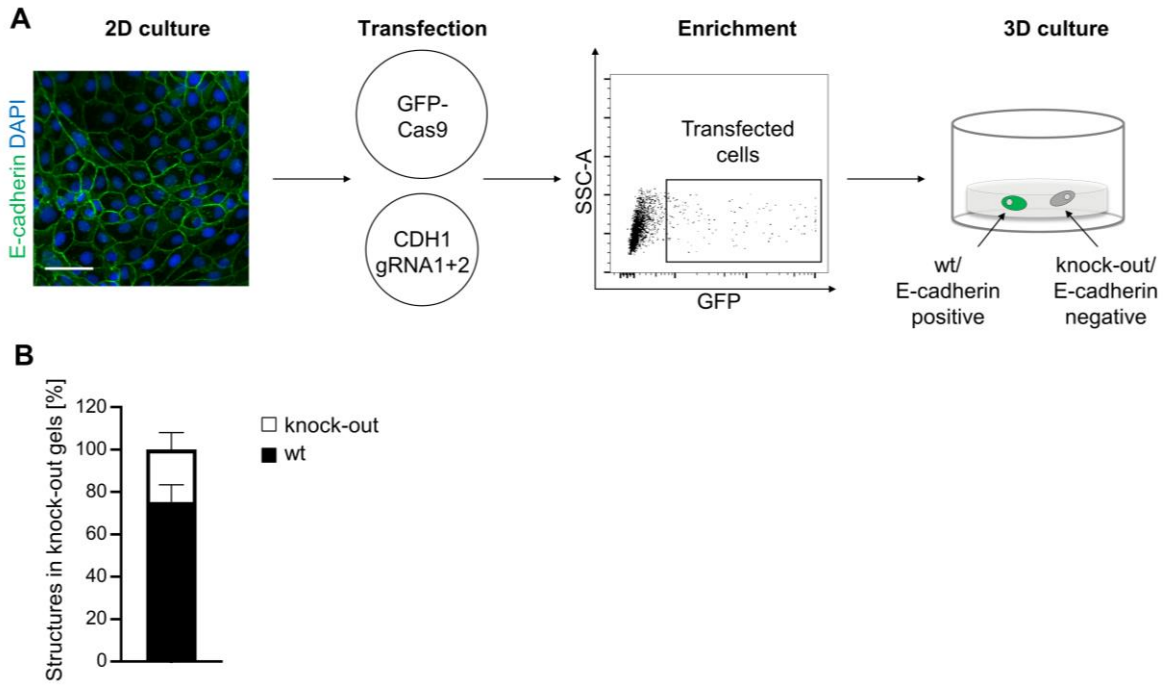

Figure 26. E-cadherin deletion. (A) Immunofluorescence of E-cadherin expression in 2D culture and schematic illustration of clonal outgrowth in 3D culture upon transfection and enrichment. E-cadherin (green), DAPI (blue). Scale bar: 50 µm. **(B)** Modification efficiency on the protein level. All structures found in gels after transfection and enrichment were classified via confocal microcopy by E-cadherin status into wild-type (wt) vs. knock-out structures. Data are presented as stacked bars and values were normalized to the total number of structures per condition. The experiment was repeated in *n*=3 donors.

Classifying structures according to their subtype revealed that as suspected from light microscopic analysis, knock-out structures indeed showed aberrant morphologies. In detail, wild-type cells grew into spheres, sticks and branched organoids of the previously described morphology (Fig. 27A,C) (Ganz et al., 2021). In contrast, knock-out cells rather grew out into sticks instead of spheres or complex branched organoids and strikingly, several scattered structures appeared (Fig. 27A,C) (Ganz et al., 2021). Scattered structures were thereby characterized by loss of cell-cell coherence resulting in singulation of cells or fragmentation of ducts. Furthermore, knock-out structures were also significantly thinner, especially within sticks and branched organoids, cells often formed single cell files (Fig. 27B,C) (Ganz et al., 2021).

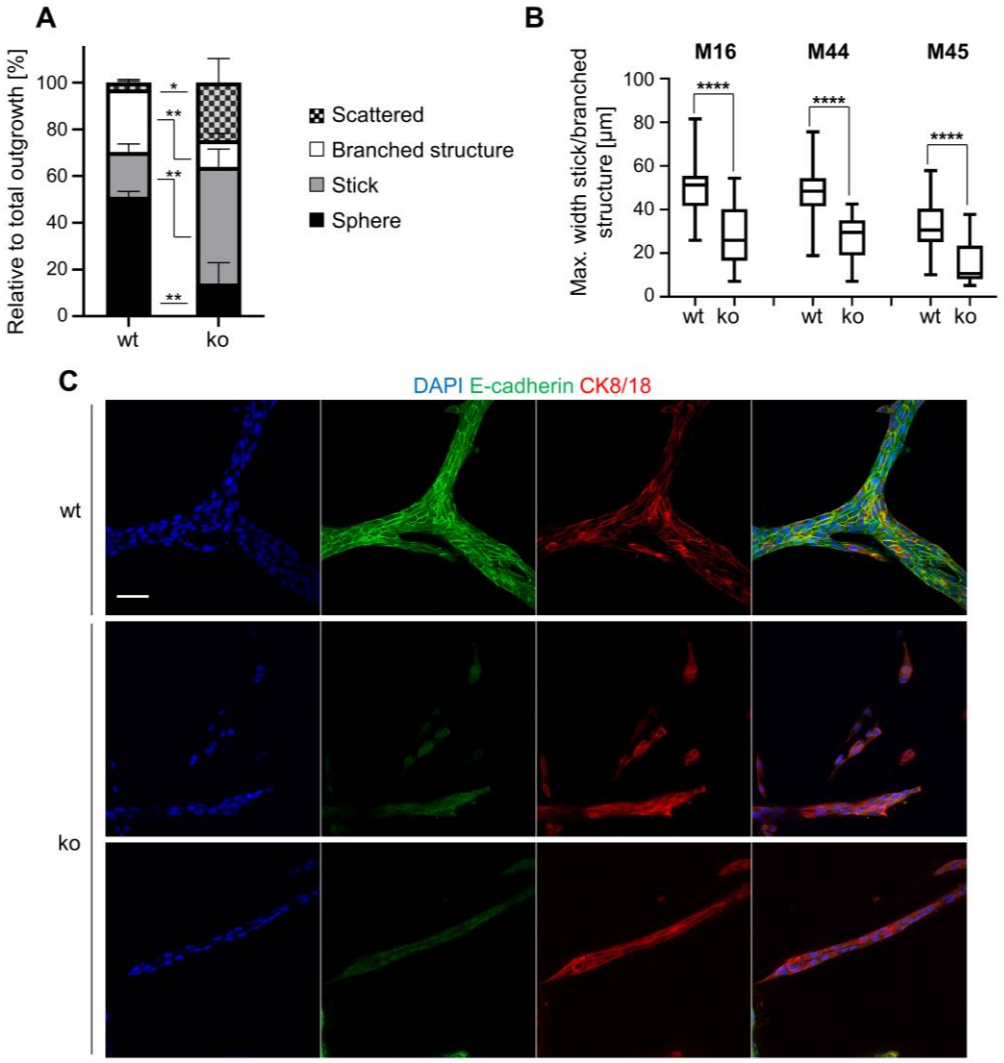

Figure 27. Analysis of E-cadherin knock-out phenotype. (A) Quantification of structure type arising in collagen type I gels after transfection and enrichment divided by E-cadherin status of the wild-type (wt) compared to knock-out (ko) structures. Data are presented as stacked bars and each structure category is normalized to the total number of structures per condition. **(B)** Quantification of maximum width of arising branched structures and sticks. Structures are divided by E-cadherin status into wt vs. ko. Data are shown as median ± 25%. P values were calculated using an unpaired two-tailed *t*-test, ****P ≤ 0.0001. *n*=3 donors. **(C)** Confocal microscopy of structures grown out in collagen type I gels after transfection and enrichment that either still maintained E-cadherin (wt) or had lost it (ko). Structures are stained for E-cadherin (green) and CK8/18 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm.

To exclude the possibility that the strong impact on organoid morphology was not a consequence of loss of proliferative capacity in knock-out structures, I performed Ki-67 staining. Thereby, I found that Ki-67 expression was similar between wild-type and knock-out structures (Fig. 28A,B) (Ganz et al., 2021). Therefore, I concluded that E-cadherin deletion directly impacts organoid morphology.

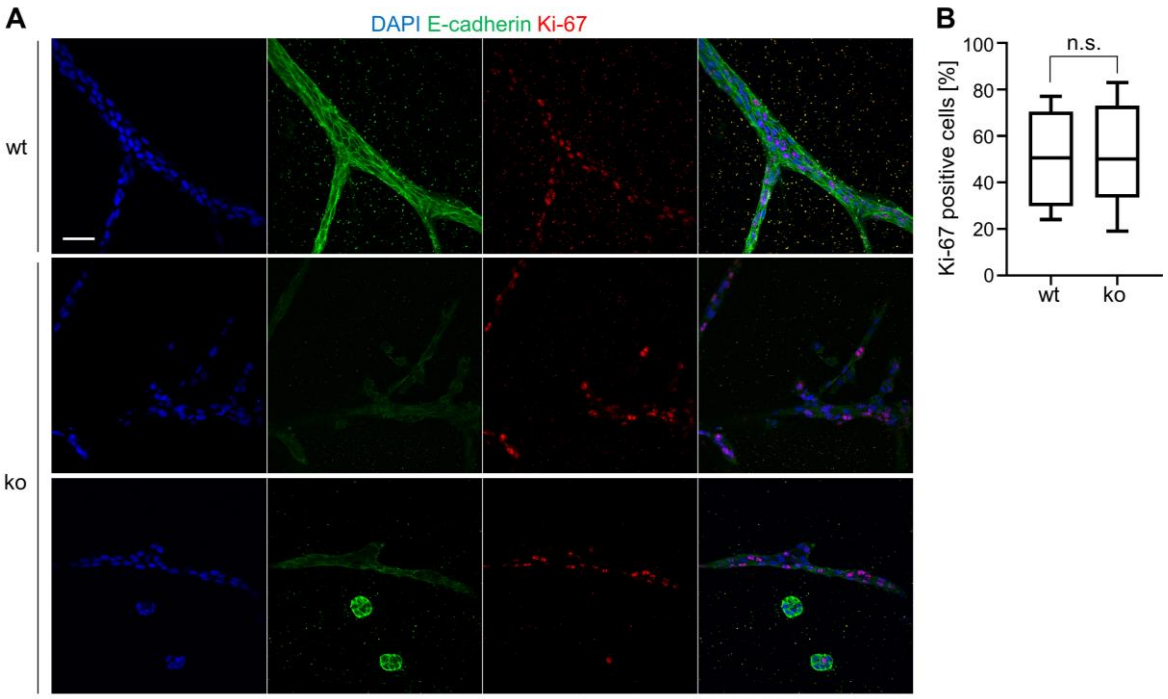

Figure 28. Proliferative activity in wt and E-cadherin knock-out phenotype. (A) Confocal microscopy of structures growing out in collagen type I gels after transfection and enrichment that either still maintained E-cadherin (wt) or had lost it (ko). Structures are stained for E-cadherin (green) and Ki-67 (red). Nuclei are stained with DAPI (blue). Scale bar: 50 µm. The experiment was repeated in *n*=3 donors. **(B)** Quantification of percentage of Ki-67 positive cells within a structure in wt as compared to ko structures. Data are shown as median ± 25%. *n*=10 structures. P values were calculated using an unpaired two-tailed *t*-test, n.s.=P ≥ 0.05.

The strong impact on organoid morphology is in line with the critical role of E-cadherin status in invasive cancer formation. Here, E-cadherin status is the main discriminator between NST and invasive lobular carcinoma (ILC), a specific, morphologically distinct subtype of invasive cancer. While NST carcinomas typically express high levels of E-cadherin, full loss of protein expression is typically observed in ILCs (Gamallo et al., 1993; Moll et al., 1993) (Fig. 29) (Ganz et al., 2021). Morphologically, ILCs are, in contrast to NSTs, characterized by an absence of duct formation and presence of cells penetrating the matrix as single cells or thin files of cells (McCart Reed et al., 2015) (Fig. 29) (Ganz et al., 2021).

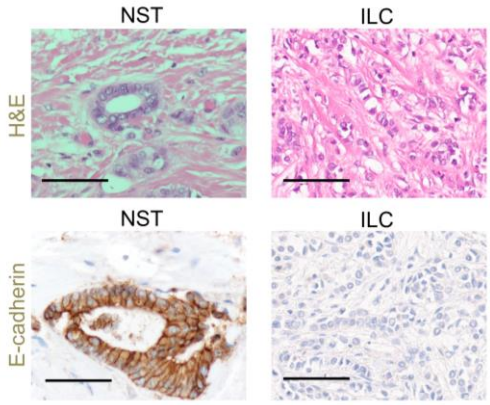

Figure 29. *In vivo* relevance of E-cadherin knock-out. H&E and E-cadherin (brown) stained section of NST carcinoma and ILC. Scale bar: 50 µm. Immunohistochemistry and associated imaging were performed by Moritz Jesinghaus (Institute of Pathology, Technical University of Munich).

In summary, the morphology observed upon E-cadherin knock-out was reminiscent of the growth pattern that ILCs exhibit *in vivo.* Therefore, I concluded that knock-out of a gene relevant for invasive branching morphogenesis has an impact on the resulting phenotype in this model that is comparable to its effect on cancerous outgrowth *in vivo*.

6 Discussion

6.1 Characteristics of luminal progenitor-derived organoids with respect to tumor grading

In this work, I have generated organoids from healthy luminal cells that histologically resemble invasive carcinomas. The possibility to create carcinoma-like structures from normal human cells, is a nonintuitive finding. Our manifested knowledge on breast carcinomas includes the notion that carcinoma cells are strikingly different from normal epithelial cells. However, there is only a vague definition of these differences. One characteristic of breast cancer cells is genetic instability (Hanahan & Weinberg, 2011). This instability causes a sheer endless number of possible modifications that is reflected by the extensive inter- and intra-tumor heterogeneity (Polyak, 2011). Despite this diversity, approximately 70% of all invasive carcinomas are histologically classified as the same subtype, the NST carcinoma. Thereby, all NST carcinomas share one common feature: the appearance of luminal ducts without a surrounding myoepithelium. Consequently, the appearance of such purely luminal ducts histologically classifies invasive breast carcinomas. Since the luminal cells forming these ducts are genetically extremely heterogenous in breast cancer, it is conceivable that the potential to invade and form ducts without a surrounding myoepithelium, already exists within healthy luminal cells as this work strongly suggests.

Of note, the luminal organoids generated in the assay described here, histologically resemble the low-grade version of NST carcinomas. The most important observation that led to this conclusion is the high degree of differentiation exhibited by luminal cell-derived organoids. In detail, the organoids formed correctly polarized ducts containing a lumen, and thus strongly resembled normal mammary gland morphology. In line with the characterization as low-grade based on degree of differentiation, luminal organoid cells showed no nuclear pleomorphisms, which is the second hallmark of low-grade carcinomas (Bloom & Richardson, 1957). An approximation to the third grading parameter, the mitotic count, can best be made by examining Ki-67 expression within luminal organoids. Notably, the elongation of luminal organoids in this assay is a rapid process and consequently, most cells were Ki-67 positive during this phase. With respect to grading, this would be the expected profile of a high-grade carcinoma. However, as cells in culture are subject to strong proliferative stimuli, the observed process can be interpreted as an artificially

accelerated version of carcinoma development. Therefore, the correlation of Ki-67 expression *in vitro* and histological grade *in vivo* might have its limitations.

Notably, the grading of breast carcinomas has prognostic value for the patients. In detail, low-grade carcinomas are associated with a good prognosis and overall high survival rate whereas high-grade carcinomas are accompanied by a less favorable prognosis (Bloom & Richardson, 1957; Elston & Ellis, 1991; Patey & Scarff, 1928). Thereby, the tumor grade is strongly associated with the number of mutated genes whereby a high mutational load correlates with high-grade carcinomas (Budczies et al., 2015). Considering that the system presented here employs genetically normal cells, it is reasonable that induction of invasive growth in these cells results in organoids that resemble low-grade carcinomas. Here, future studies could address whether organoids generated from tumor-derived luminal progenitor cells, exhibit characteristics of higher-grade carcinomas.

In summary, while the conditions here are not fully transferrable to the classification system applied on breast cancer tissue samples, the morphology generated from single luminal progenitor cells strongly resembles low-grade NST carcinomas. Thereby, this model is the only human material based *in vitro* model known to me that reflects histopathological features of invasive breast cancers.

6.2 Parameters of luminal progenitor-derived branched ductal organoid formation

Within this work I found three parameters whose combination is essential for the formation of invasively growing branched ductal organoids.

6.2.1 Ensuring sufficient cellular proliferation

Excessive proliferation is one of the hallmarks of breast cancer (Hanahan & Weinberg, 2011). Thereby, proliferation can be drastically increased due to alterations in specific genes. Genes that frequently result in deregulated proliferation include oncogenes such as *MYC* as well as tumor suppressor genes such as *PTEN* and *TP53* (Hanahan & Weinberg, 2011; Kinzler & Vogelstein, 1997; Vogelstein & Kinzler, 2004). In the model presented here, the luminal cells are of healthy origin. Consequently, for modelling invasive outgrowth, the cells need to be put into a highly proliferative state artificially by the culture conditions.

Maintenance of the luminal subset is a known difficulty in culture of mammary gland tissue. Therefore, I built my work on the BCOM medium (Sachs et al., 2018), that has previously

been described as suitable to maintain the luminal subset in 3D culture over time. Within BCOM medium, ductal, branched organoids appeared. However, those organoids were relatively small and non-complex. Here, supplementation with small amounts of FCS drastically increased outgrowth of branched ductal organoids. FCS is a mixture of several hormones, lipids and growth factors and its main function is fostering proliferation and growth in cell culture (Puck & Marcus, 1955). Nevertheless, as FCS is a complex, organic product, it might theoretically have unknown functions in this assay apart from encouraging proliferation. However, due to the appearance of invasively shaped branched ductal organoids already within the FCS-free conditions, it can be excluded that FCS unlocks invasive capacity within luminal progenitor cells. Rather, it reinforces the invasive outgrowth by ensuring proliferative activity.

6.2.2 Direct contact to collagen type I

As a component of the ECM, collagen type I is in the normal mammary gland usually locally separated from the luminal subset by the basal cell layer as well as the basement membrane. Invasive carcinomas are characterized by an absence of the basal cell layer and basement membrane and therefore, invading carcinoma cells are in direct contact with the ECM (Lopez-Garcia et al., 2010). Of note, the presence of an intact basal cell and basement membrane barrier has been described as suppressor of luminal cell invasion *in vivo* (M. Hu et al., 2008; Polyak & Kalluri, 2010). Or in other words, the lack of separation between luminal cells and the ECM drives luminal cell invasion. Moreover, in breast cancer, the accumulation of collagen type I within the ECM has been described as a major risk factor for invasive cancer development (Boyd et al., 1998; Byrne et al., 1995; Wolfe, 1976). Based on these considerations, I suspected that exposing luminal cells to collagen type I could foster luminal cell invasion. Supporting this hypothesis, an invasion-promoting effect of collagen type I on other epithelial cells *in vitro* has previously been described (Buchmann et al., 2021; Carey et al., 2017; Linnemann et al., 2017).

In this work, seeding the freshly sorted luminal progenitor cells into collagen type I gels, resulted in invasive branching morphogenesis. Furthermore, in line with the non-invasive behavior of luminal cells under homeostatic conditions, I found that culturing the cells within the basement membrane surrogate Matrigel, resulted in proliferation in the form of sphere formation but not in the induction of invasive growth. This confirms my initial assumption that under certain circumstances, collagen type I can have invasion-promoting effect on normal luminal cells. Furthermore, it supports the notion that the basement membrane and

its components, in this work recapitulated in Matrigel culture, have an invasion-suppressing function.

Of note, it was observed that luminal progenitor-derived organoids secrete the basement membrane component laminin at a late stage of their development. The secretion of this component likely hinters direct contact of luminal progenitor cells and collagen type I matrix. Therefore, collagen-driven invasion might be hindered at this stage resulting in the above-described rounding-up of extending ducts. This would indicate that invasion by luminal cells is self-limiting.

Based on these observations, I conclude that the collagen type I-containing 3D matrix is the second determinant of luminal progenitor-derived organoid formation.

6.2.3 Balanced contractility

The fact that previous studies failed to recapitulate luminal cell invasion besides employing various culture media and collagen type I matrices (Gudjonsson, Rønnov-Jessen, et al., 2002; Linnemann et al., 2015), led me to reason that a third factor controls luminal cell invasion that discriminates the conditions created in the system presented here from earlier approaches.

Previous *in vitro* studies with immortalized normal-like mammary epithelial cells (Carey et al., 2017) as well as epithelial kidney cells (Zhou & Kramer, 2005) showed that inhibition of cellular contractility via ROCK enabled cellular invasion into a collagen type I matrix. Notably, cellular contractility is largely regulated via the Rho-ROCK myosin II pathway (Amano et al., 2010). In the assay described here, the ROCK inhibitor Y-27632 is a component of the BLOM medium. Interestingly, it was observed that if the BLOM medium was deprived of ROCK inhibitor, strong cellular contractility appeared, and invasion and consequently ductal organoid formation was prohibited. In contrast, in full BLOM medium, the presence of ROCK inhibitor had a strongly reductive effect on contractile behavior of the luminal cells, which enabled directed invasive outgrowth. However, residual bead displacement and fiber alignment monitored despite the presence of ROCK inhibitor in BLOM medium hinted towards a role of the ROCK inhibitor in balancing contractility of luminal cells to a certain level rather than fully obstructing it.

Of note, effects besides the reduction of contractility have been attributed to the ROCK inhibitor Y-27632. In detail, it has been described that ROCK inhibitor Y-27632 increases outgrowth of murine (Makarem et al., 2013) and human (Linnemann et al., 2015) mammary epithelial cells in 3D culture. Consequently, I had to exclude the possibility that the observed

dependency of invasion on ROCK inhibition was merely an effect of increased cellular viability. I did so by inhibiting a different segment of the contractility cascade via the myosin II inhibitor PAB. As expected, inhibition of contractility via PAB also resulted in the formation of branched organoids from luminal progenitor cells. However, the lower efficiency of branched organoid formation in the presence of PAB compared to conditions with ROCK inhibitor Y-27632, indicates that Y-27632 additionally supports luminal cell outgrowth.

Overall, this work reveals that three parameters are required to allow generation of luminal organoids. First, proliferation within the luminal cells needs to be stimulated based on an optimized culture medium to guarantee sufficient cellular material for generation of complex organoids from single cells. Second, direct contact to a collagen type I-containing ECM surrogate needs to be available. And third, contractility within luminal cells needs to be reduced via inhibition of the Rho-ROCK myosin II pathway. Importantly, I found that if only one of these parameters is not fulfilled, the generation of complex branched ductal organoids is impaired in this assay.

6.2.4 Potential *in vivo* relevance of parameters for luminal organoid formation as cause for invasive cancer formation

The need for proliferative capacity and the invasion-promoting effect of collagen type I are renowned elements of cancerous development *in vivo*. In contrast, the reduction of cellular contractility, which is the third parameter I found to be essential within this *in vitro* system, is a less prominent criterion for luminal cell invasion.

The need for inhibiting contractility in luminal progenitor cells was surprising with respect to the fact that *in vivo*, contractile abilities have so far mainly been attributed to the αSMA-rich basal subset of the human mammary gland (Deugnier et al., 1995; Haaksma et al., 2011). A possible explanation for increased contractile ability of luminal progenitor cells in collagen type I gels could be a raise in activity of the small GTPase RhoA, which has been described as a consequence of epithelial cell adhesion to collagen type I *in vitro* (Zhou & Kramer, 2005). In line with that, it has been shown that RhoA upregulation prevents cellular invasion and correct polarization of epithelial cells (Yu et al., 2008; Zhou & Kramer, 2005). Within the respective studies, both downstream effects of RhoA upregulation could be reverted by addition of ROCK inhibitor Y-27632 (Yu et al., 2008; Zhou & Kramer, 2005). Therefore, RhoA-mediated contractility in the luminal progenitor subset might represent an invasion-preventing cellular safety mechanism triggered in the event of direct contact to collagen type I.

A recent mouse-based study by Schipper et al. supports the relevance of the above-described mechanism with *in vivo* data (Schipper et al., 2019). In detail, Schipper et al. report that *in vivo* the mere contact of luminal cells to the ECM did not result in cellular invasion. However, once functionality of actomyosin contractility regulator MYPT1 was reduced within the luminal cells, invasion was enabled (Schipper et al., 2019). Consequently, the contractile reaction of luminal cells to collagen type I likely is an innate invasion-suppressing mechanism triggered by the ECM contact. Moreover, Schipper et al. show that full obstruction of MYPT1 and therefore complete blockage of contractile ability, results in loss of invasive capacity *in vivo* (Schipper et al., 2019). This is in accordance with the above-described observations that even upon ROCK inhibition, residual contractility remains within the luminal subset and likely is required for invasion as it supports ECM remodeling. However, the question remains how inhibition of contractility might be an enabler of luminal cell invasion in the human mammary gland.

One possible explanation is the acquisition of genetic aberrations by luminal cells. For example, as shown by Schipper et al., the truncation of MYPT1 via mutations in its encoding gene *Ppp1r12a* can reduce cellular contractility in luminal cells (Schipper et al., 2019). However, loss of MYPT1 function is largely restricted to the ILC subtype of invasive breast cancer (Kas et al., 2017) and can therefore not explain how contractility reduction could be enabled in NST carcinoma. Moreover, the inability to identify universal genetic aberrations between invasive and *in situ* carcinomas (Hernandez et al., 2012; Ma et al., 2003; Petridis et al., 2016; Porter et al., 2003) makes it unlikely that contractility-reducing genetic aberrations are a determinant of luminal cell invasion. Nevertheless, this does not exclude that non-genetic aberrations within the epithelial cells could alter contractile behavior and thereby encourage cellular invasion.

A contractility-associated pathway frequently disturbed in all types of invasive breast cancers is non-canonical Wnt signaling via Wnt5a. In this pathway, Wnt5a is upstream of ROCK and mechanistically, loss of Wnt5a results in diminished contractility via reduced ROCK signaling (Asem et al., 2016). Thereby, loss or reduction of Wnt5a protein expression is associated with relapse of invasive breast cancer and increased metastasis (Serra et al., 2011). This aberration is found frequently, with 45% to 75% of human mammary breast cancers having lost or strongly reduced Wnt5a protein expression (Borcherding et al., 2015; Dejmek et al., 2005; Prasad et al., 2018). Wnt5a is of particular interest in this context as lack of Wnt5a protein expression has been described mainly as a feature of invasive carcinomas. In contrast, pre-invasive lesions (*in situ* carcinomas) typically maintain and partly even overexpress Wnt5a (Jönsson et al., 2002). This notion seemingly opposes studies that do not find differences in genotype or gene expression between pre-

transinvasive and invasive cells. Interestingly, it has previously been reported that in human breast epithelial cells, Wnt5a mRNA levels do not correlate to the Wnt5a protein levels in breast cancer tissue due to post-transcriptional modification, which would explain how this factor could remain undetected in gene expression studies (Dejmek et al., 2005; Leandersson et al., 2006).

Overall, there are hints that reduction of contractility is a mechanism that might contribute to invasion of luminal cells *in vivo*. However, invasion is generally viewed as a multifactorial process, a notion that is also supported by the fact that I found three parameters that are required for invasion *in vitro*. Based on that, it is thinkable that if invasion *in vivo* also depends on contractility-reducing alterations in the epithelium, it is further regulated by the surrounding matrix and proliferative stimuli.

6.3 Interpretation of the results in light of invasion theories

The theories on what causes a carcinoma to grow invasively *in vivo* can be roughly divided into two categories based on whether it identifies genetic aberrations as determinants of invasion. Two of the evolutionary models, as described in the introduction, rely on genetic differences between invading clones and cells that remain non-invasive within the ducts in order to explain why only certain cells undergo invasion. One of these models is the *independent lineage* model that proposes completely different lineages of cells are responsible for *in situ* and invasive cancers (Sontag & Axelrod, 2005). The second model is the *evolutionary bottleneck* model, which suggests that genetic accumulations occur within the ducts prior to invasion and only a subpopulation of cells with a certain genetic profile overcomes the evolutionary bottleneck and invades into the ECM (Cowell et al., 2013). In contrast, the *multiclonal invasion* model proposes that genomic heterogeneity is largely maintained during progression to the invasive state, and thus emphasizes the importance of non-genetic factors such as changes in the stroma or ECM (Casasent et al., 2018). The last model is thereby supported by the above-mentioned notion that it has not yet been possible to identify universal invasion-specific genetic aberrations in invasive carcinoma cells (Hernandez et al., 2012; Ma et al., 2003; Petridis et al., 2016; Porter et al., 2003).

Within this work, I have defined three parameters that are required for luminal cell invasion *in vitro*. Thereby, the system outlined in this work indicates that a single alteration is not sufficient to make the difference between invasive and non-invasive growth in luminal cells. For instance, cells can have reduced contractility and high proliferative potential within BLOM medium but will not invade unless contact to collagen type I is given. This situation

is reminiscent of *in situ* carcinomas in which cells have partly acquired numerous alterations that can largely overlap with those in invasive carcinomas, yet are separated from the ECM and do not invade (Petridis et al., 2016). Therefore, this work suggests that it is likely the same cells growing non-invasively in *in situ* carcinomas that also grow out into the surrounding tissue in invasive carcinomas, thus opposing the idea of generally different lineages in *in situ* and invasive carcinomas as proposed in the *independent lineage* model.

Moreover, the finding that contact to the ECM component collagen type I enables the induction of invasion within genetically normal cells emphasizes the role of this non-genetic factor. However, within the model described here, contact to collagen type I alone is not sufficient to induce invasion unless combined with proliferative capacity and contractility inhibition. As outlined above, it is conceivable that genetic aberrations are involved in the regulation of proliferation and cellular contractility *in vivo*. Therefore, this work does not exclude a crucial role of genetic aberrations during invasion. However, this work does indicate that contact to the ECM is one important non-genetic factor during the complex invasion process.

Moreover, the findings described in this work give a possible explanation for the emergence of the partly opposing theories of the *evolutionary bottleneck* and *multiclonal invasion* model. Tumor development is not a standardized process. Thus, in every cancer, events required for invasion might occur in a variable timeline. For instance, in the *in situ* stage, several clones could possess alterations that increase their proliferative activity and decrease their contractile ability. If then the barrier to the collagen type I-containing ECM is lost, multiple clones will likely invade concomitantly. However, in other carcinomas, this physical barrier could be lost at an earlier timepoint, making it a matter of time until an evolutionary bottleneck occurs due to a specific clone acquiring high proliferative capacity and reduced contractility. This theory also fits with the fact that *in situ* carcinomas can be a precursor to the invasive disease but are not necessarily required (Cowell et al., 2013). In detail, the accumulation of cells within the ducts, and thus the development of an *in situ* carcinoma will according to this logic only occur if an increase in proliferative activity precedes other steps required for invasion.

In summary, the crucial role of contact to collagen type I indicates that the ECM is a non-genetic factor that plays an important role during luminal cell invasion. However, the other two factors, proliferative capacity and reduced contractility, might be a result of genetic aberrations or could be mediated by stromal factors. Therefore, the observations described here do not allow a definite conclusion about how much genetic factors are involved in the

invasion process and it is thinkable that invasion is in the end generated by a combination of genetic and non-genetic factors.

6.4 Unknown capabilities of luminal progenitor cells unraveled in this work

Here, an assay is described in which single luminal progenitor cells undergo an invasive branching morphogenesis process resulting in ductal structures with correct polarization. These observations reveal several largely unknown capabilities of luminal progenitor cells *in vitro*.

In previous culture systems, neither comparable invasive nor branched structure formation capacity has been described for the luminal progenitor subset. Instead, in similar approaches employing collagen type I matrices, luminal progenitor cells mainly grew out as non-invasive spheres (Gudjonsson, Rønnov-Jessen, et al., 2002; Linnemann et al., 2015). However, while Gudjonsson et al. exclusively observed formation of spheres (Gudjonsson, Rønnov-Jessen, et al., 2002), Linnemann et al. described the appearance of branched structures in rare occasions (Linnemann et al., 2015). The final morphology of these structures was characterized by a low complexity and grape-like, non-invasive shape. Of note, while Gudjonsson et al. did not use the ROCK inhibitor Y-27632 in their medium, Linnemann et al. added Y-27632 for the first five days of culture in order to foster initial attachment. Based on the key role of the ROCK inhibitor for the formation of branched structures from luminal progenitor cells as unraveled in this work, it is likely that this component caused the observed differences between the two protocols.

Another novel aspect of the ductal structures described in this work, is their pronounced polarity. *In vivo*, invasive luminal cells can maintain strong polarization similar to the ductal network of the normal mammary gland. Histologically, this is reflected in the formation of luminal ducts that preserve hallmarks of polarization such as lumen formation as typically observed in low-grade NSTs. Gudjonsson et al. described luminal progenitor-derived spheres with reversed polarity in their model (Gudjonsson, Rønnov-Jessen, et al., 2002). In contrast, while Linnemann at al. did not specify the expression of polarization markers in luminal progenitor-derived branched structures or spheres, those exhibited lumen formation to some extent, indicating correct polarization (Linnemann et al., 2015). Again, the addition of Y-27632 for the initial attachment phase by Linnemann et al. could explain these differences and explain why in the assay described here, such a distinct polarization can be observed under constant ROCK inhibition.

In this work, it was shown that duct elongation and branching are directed by luminal leader cells. As mentioned above, in previous approaches, healthy luminal cells never showed comparable invasive capacity. Consequently, it has never been described that luminal cells of healthy origin could take over leader cell function during invasive processes. However, the concept that mammary epithelial cell invasion can be guided by epithelial leader cells is well established. In culture systems in which branched structures are derived from normal basal mammary epithelial cells, it has been shown that basal leader cells guide branch elongation (Buchmann et al., 2021). Similar results were obtained in experiments conducted by Sokol et al., who embedded fragments of healthy mammary gland tissue into hydrogels containing collagen type I (Sokol et al., 2016). Here, in line with the capabilities of basal cells as described by Buchmann et al. (Buchmann et al., 2021), duct elongation was guided by basal leader cells (Sokol et al., 2016). In both set-ups, bilayered branched structures formed, which had an inner layer of luminal cells and an outer layer of basal cells, and thus resembled the build-up of the normal human mammary gland. In contrast, invasive behavior of tumor-derived tissue has rarely been recreated *in vitro*. The only somewhat successful approach for recreating tumor cell invasion from human mammary gland tissue has been presented by Cheung et al. (Cheung et al., 2013). In detail, they found that when fragments of tumor tissue were cultured in a collagen type I matrix, groups of cells invaded collectively, guided by epithelial leader cells. However, these leader cells and practically all invasively growing stalk cells, exhibited basal characteristics (Cheung et al., 2013). Notably, it has previously been reported that luminal cells can acquire basal cell characteristics in culture (Gudjonsson, Villadsen, et al., 2002; Pechoux et al., 1999; Rosenbluth et al., 2020; Shehata et al., 2012). However, strong expression of basal markers in all invasively growing cells is not reminiscent of the *in vivo* morphology of most NSTs and would only be expected in the relatively rare subset of basal-like cancers (Livasy et al., 2006). Furthermore, within the set-up by Cheung et al., the cells did not undergo invasive branching morphogenesis, neither did the resulting structures show polarized ductal morphology(Cheung et al., 2013). Therefore, it remains questionable whether the basal leader mechanism described previously, is of universal relevance in the development of NST carcinomas. Once again, it can be noted that in none of the described approaches, contractility of the cells was balanced via for instance addition of ROCK inhibitor for extended periods of time. With respect to the significance of balanced contractility as described in this work, this detail might be the crucial difference between previous *in vitro* culture systems and the approach described here. Of note, these findings do not exclude that *in vivo* luminal cell invasion might be guided by basal or stromal leader cells or that initially luminal leader cells can acquire basal-like phenotypes during tumorigenesis. However, this work clearly shows, that

theoretically the leader cell role can be taken over by luminal progenitor cells that maintain luminal characteristics.

Importantly, in the conditions defined in this work, up to 1 out of 10 freshly isolated normal luminal progenitor cells generated complex multicellular organoids through invasive branching morphogenesis. Prior to organoid formation in 3D culture, primary luminal progenitor cells undergo the tissue isolation procedure and FACS, which most likely results in reduced viability of this subset. Considering this, rather than being the ability of a rare luminal subset, it is likely that almost every single luminal progenitor cell bears matrix invasion potential. Thus, my work suggests that invasive potential merely needs to be triggered by suitable conditions and can result in extensive invasion as long as the proliferative capacity of the cells allows for it.

In summary, this work is the first one to show that luminal cells have the potential to self-sufficiently invade and organize into structures resembling the morphology of low-grade NST carcinomas. Thereby, all cells, including the leader cells which guide invasion, retain luminal characteristics. Furthermore, the large share of normal luminal cells that show invasive capacity in this assay, supports the theory that invasive potential is already existent within pre-invasive malignancies and merely needs to be unlocked in the luminal compartment.

6.5 Parameters of basal cell-derived organoid formation

While the focus of this work was on unravelling the potential of luminal cells, I also showed that it is possible to derive branched structures from basal cells under the same conditions that are required for organoid formation from luminal cells.

Thereby, the basal cells behaved similar to what has been described previously within alternate culture conditions (Linnemann et al., 2015). In detail, besides undergoing a branching morphogenesis process, the potency to partly covert towards a luminal phenotype, was unraveled within the inner layer of organoid cells. However, the morphology of branched structures arising in BLOM medium was slightly different to structures formed in the medium as employed by Linnemann et al.. Specifically, it appeared that within the conditions by Linnemann et al., alveologenesis was more distinct. A possible explanation is that the medium used by Linnemann et al. contains the diterpene Forskolin, which they described as an essential factor of the medium composition, fostering alveologenesis (Linnemann et al., 2015). Consequently, the lower levels of alveologenesis observed in organoids grown in BLOM medium might be based on the fact that this medium does not

contain Forskolin. However, experiments employing BLOM with addition of Forskolin would be required to validate this hypothesis.

Another noteworthy variation between the two media is the variant application of ROCK inhibitor as culture supplement. While the ROCK inhibitor Y-27632 is a constant supplement of the BLOM medium, in conditions described by Linnemann et al., Y-27632 is merely added for the initial attachment phase until day 5 (Linnemann et al., 2015). Therefore, it appears that in contrast to luminal progenitor cells, basal cells do not require extended inhibition of contractility for undergoing branching morphogenesis. Moreover, in the conditions described by Linnemann et al., extended treatment with Y-27632 after day 5 of culture has reportedly resulted in loss of branched structure formation capacity (Buchmann et al., 2021; Linnemann et al., 2015). This is based on the fact that basal cell-derived organoids in conditions developed by Linnemann et al. relied on strong collective contractility for successful elongation, which is drastically reduced by addition of Y-27632 (Buchmann et al., 2021). Interestingly, it appears that basal cell-derived organoids in BLOM are not affected in the same way upon extended exposure to the ROCK inhibitor as they grow into coherent organoids, which calls into question whether they are generated via the same morphogenesis mechanism.

One possible explanation for these contradicting observations concerning culture conditions that allow for branched structure formation from basal cells, could be that the branched structures within BLOM and PC medium are generated by distinct subsets of basal cells. Due to the advancement of single cell techniques, the idea has been put forward that the previously identified three subsets of mammary epithelial cells with only a single basal population, is a severe over-simplification of mammary epithelial cell heterogeneity. Specifically, single cell RNA-sequencing of normal human mammary gland tissue has identified two potential subsets of basal cells (Q. H. Nguyen et al., 2018). Therefore, it is thinkable that even though in both set-ups, basal cells are isolated via the same FACS sorting scheme and express the same lineage characteristics in culture, they stem from transcriptionally distinct basal populations.

In summary, morphological differences in branched structure formation from basal cells could be based on certain components of the different media or could be the result of heterogeneity within the FACS-sorted basal population. However, additional studies are required to address this question.

6.6 Branched ductal organoids as a new tool for breast cancer research

Besides the biological implications of this work, the herein presented organoid model is a promising tool for invasive breast cancer research.

Organoids are a popular and relatively new tool in cancer research. They allow a more accurate depiction of the *in vivo* situation and are therefore emerging systems for screening purposes in preclinical drug discovery (Lancaster & Knoblich, 2014). However, established breast cancer organoid models do not recapitulate the dynamic processes of invasion but are focused merely on maintaining the cellular heterogeneity. As a result, in studies employing patient-derived cancer organoids, candidate drugs for solid tumors are evaluated predominantly by their ability to induce tumor organoid shrinkage or reduce cell viability in culture (Driehuis et al., 2020). Furthermore, these models rely on sparsely available tumor tissue samples.

The assay developed here bypasses several drawbacks of established breast cancer models. First, this assay employs luminal progenitor cells and therefore the actual cells of origin for breast cancer, which allows the observation of events specific to this cellular subset. Thereby, invasion is modelled from initially healthy luminal cells, which means that the focus lies on early processes occurring in luminal cell invasion. Furthermore, as healthy mammary material is abundantly available due to mammoplasties performed for aesthetic reasons, the issue of tissue shortage can be bypassed. Importantly, the invasion process recapitulated here actually results in a morphology similar to the one observed *in vivo*, which suggests a mechanistic resemblance. This feature is especially relevant for screening approaches in drug discovery that aim to reduce invasive and migrative behavior of cancerous cells, so called migrastatics (Gandalovičová et al., 2017). Here, a specific example for the impact that an invasion-suppressing compound can have in this model is given by the treatment with MMP inhibitor Marimastat, which obstructs luminal cell invasion into the matrix.

Moreover, while the standard assay is based on genetically normal luminal progenitor cells, gene-specific impact can be recapitulated by editing the cells with the CRISPR-Cas9 system prior to organoid formation. This possibility allows recapitulating the influence of individual genetic aberrations on tumor cell invasion in a highly specific manner. A further advantage here is the clonality of outgrowing organoids, which allows differentiating between successfully edited organoids in which every single cell bears the same aberration as opposed to unmodified wild-type organoids. In contrast, CRISPR approaches in

established organoid models with bulk-tumor tissues rely on a close to 100% editing efficiency or the option to select for positively edited cells in order to be able to present meaningful and defined knock-out organoids (Sachs et al., 2018).

Currently, the read-out of this model is purely morphological. However, while all molecular subtypes of breast cancer can belong to the histopathological group as NST carcinomas and can therefore be morphologically similar, they exhibit a variety of transcriptomic profiles. Future studies could therefore address the analysis of the organoid transcriptome and the potential influence that specific genetic modifications have on the molecular subtype. An interesting example here could be the deletion of *BRCA1* which is *in vivo* associated with the basal-like breast cancer subtype (Turner & Reis-Filho, 2006) and might result in a shift towards a basal-like expression profile.

Taken together, the model presented here, offers a new approach for investigating certain aspects of tumorigenesis. The main advantage of this model is that it employs the actual cells of origin for breast cancer. Thereby, this model will be extremely helpful as a human material-based system for validating results acquired in animal models or with breast cancer cell lines. Moreover, this model has great potential in drug screening approaches in search for compounds that specifically tackle luminal cell invasion early on.

7 Bibliography

Al-Benna, S., Poggemann, K., Steinau, H.-U., & Steinstraesser, L. (2010). Diagnosis and management of primary breast sarcoma. *Breast Cancer Research and Treatment*, *122*(3), 619–626. https://doi.org/10.1007/s10549-010-0915-y

Amano, M., Nakayama, M., & Kaibuchi, K. (2010). Rho-kinase/ROCK: A key regulator of the cytoskeleton and cell polarity. *Cytoskeleton*, *67*(9), 545–554. https://doi.org/10.1002/cm.20472

Andrews, J. L., Kim, A. C., & Hens, J. R. (2012). The role and function of cadherins in the mammary gland. *Breast Cancer Research*, *14*(1), 203. https://doi.org/10.1186/bcr3065

Anstine, L. J., & Keri, R. (2019). A new view of the mammary epithelial hierarchy and its implications for breast cancer initiation and metastasis. *Journal of Cancer Metastasis and Treatment*, *2019*. https://doi.org/10.20517/2394-4722.2019.24

Asem, M., Buechler, S., Wates, R., Miller, D., & Stack, M. (2016). Wnt5a Signaling in Cancer. *Cancers*, *8*(9), 79. https://doi.org/10.3390/cancers8090079

Bach, K., Pensa, S., Grzelak, M., Hadfield, J., Adams, D. J., Marioni, J. C., & Khaled, W. T. (2017). Differentiation dynamics of mammary epithelial cells revealed by single-cell RNA sequencing. *Nat Commun*, *8*(1), 2128. https://doi.org/10.1038/s41467-017-02001-5

Bachelard-Cascales, E., Chapellier, M., Delay, E., Pochon, G., Voeltzel, T., Puisieux, A., Caron de Fromentel, C., & Maguer-Satta, V. (2010). The CD10 enzyme is a key player to identify and regulate human mammary stem cells. *Stem Cells*, *28*(6), 1081–1088. https://doi.org/10.1002/stem.435

Balani, S., Nguyen, L. V., & Eaves, C. J. (2017). Modeling the process of human tumorigenesis. *Nature Communications*, *8*(1), 15422. https://doi.org/10.1038/ncomms15422

Bell, J., Walsh, S., Nusrat, A., & Cohen, C. (2003). Zonula occludens-1 and Her-2/neu expression in invasive breast carcinoma. *Appl Immunohistochem Mol Morphol*, *11*(2), 125–129. https://doi.org/10.1097/00129039-200306000-00006

Bloom, H. J. G., & Richardson, W. W. (1957). Histological Grading and Prognosis in Breast Cancer. *British Journal of Cancer*, *11*(3), 359–377. https://doi.org/10.1038/bjc.1957.43

Boras-Granic, K., Dann, P., & Wysolmerski, J. J. (2014). Embryonic cells contribute directly to the quiescent stem cell population in the adult mouse mammary gland. *Breast Cancer Research*, *16*(6), 487. https://doi.org/10.1186/s13058-014-0487-6

Borcherding, N., Kusner, D., Kolb, R., Xie, Q., Li, W., Yuan, F., Velez, G., Askeland, R., Weigel, R. J., & Zhang, W. (2015). Paracrine WNT5A Signaling Inhibits Expansion of Tumor-Initiating Cells. *Cancer Research*, *75*(10), 1972–1982. https://doi.org/10.1158/0008-5472.CAN-14-2761

Boyd, N. F., Jensen, H. M., Cooke, G., & Han, H. L. (1992). Relationship Between Mammographic and Histological Risk Factors for Breast Cancer. *JNCI: Journal of the National Cancer Institute*, *84*(15), 1170–1179. https://doi.org/10.1093/jnci/84.15.1170

Boyd, N. F., Lockwood, G. A., Byng, J. W., Tritchler, D. L., & Yaffe, M. J. (1998). Mammographic densities and breast cancer risk. *Cancer Epidemiology, Biomarkers & Prevention : A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*, *7*(12), 1133–1144. http://www.ncbi.nlm.nih.gov/pubmed/9865433

Breunig, C. T., Durovic, T., Neuner, A. M., Baumann, V., Wiesbeck, M. F., Koferle, A., Gotz, M., Ninkovic, J., & Stricker, S. H. (2018). One step generation of customizable gRNA vectors for multiplex CRISPR approaches through string assembly gRNA cloning (STAgR). *Plos One*, *13*(4), e0196015. https://doi.org/10.1371/journal.pone.0196015

Brownfield, D. G., Venugopalan, G., Lo, A., Mori, H., Tanner, K., Fletcher, D. A., & Bissell, M. J. (2013). Patterned collagen fibers orient branching mammary epithelium through distinct signaling modules. *Curr Biol*, *23*(8), 703–709. https://doi.org/10.1016/j.cub.2013.03.032

Buchmann, B., Engelbrecht, L. K., Fernandez, P., Hutterer, F. P., Raich, M. K., Scheel, C. H., & Bausch, A. R. (2021). Mechanical plasticity of collagen directs branch elongation in human mammary gland organoids. *Nature Communications*, *12*(1), 2759. https://doi.org/10.1038/s41467-021-22988-2

Budczies, J., Bockmayr, M., Denkert, C., Klauschen, F., Lennerz, J. K., Györffy, B., Dietel, M., Loibl, S., Weichert, W., & Stenzinger, A. (2015). Classical pathology and mutational load of breast cancer - integration of two worlds. *The Journal of Pathology. Clinical Research*, *1*(4), 225–238. https://doi.org/10.1002/cjp2.25

Burdall, S. E., Hanby, A. M., Lansdown, M. R., & Speirs, V. (2003). Breast cancer cell lines: friend or foe? *Breast Cancer Research*, *5*(2), 89. https://doi.org/10.1186/bcr577

Byrne, C., Schairer, C., Wolfe, J., Parekh, N., Salane, M., Brinton, L. A., Hoover, R., & Haile, R. (1995). Mammographic Features and Breast Cancer Risk: Effects With Time, Age, and Menopause Status. *JNCI Journal of the National Cancer Institute*, *87*(21), 1622–1629. https://doi.org/10.1093/jnci/87.21.1622

Carey, S. P., Martin, K. E., & Reinhart-King, C. A. (2017). Three-dimensional collagen matrix induces a mechanosensitive invasive epithelial phenotype. *Scientific Reports*, *7*(1), 42088. https://doi.org/10.1038/srep42088

Casasent, A. K., Schalck, A., Gao, R., Sei, E., Long, A., Pangburn, W., Casasent, T., Meric-Bernstam, F., Edgerton, M. E., & Navin, N. E. (2018). Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing. *Cell*, *172*(1–2), 205-217 e12. https://doi.org/10.1016/j.cell.2017.12.007

Chen, S., Iversen, E. S., Friebel, T., Finkelstein, D., Weber, B. L., Eisen, A., Peterson, L. E., Schildkraut, J. M., Isaacs, C., Peshkin, B. N., Corio, C., Leondaridis, L., Tomlinson, G., Dutson, D., Kerber, R., Amos, C. I., Strong, L. C., Berry, D. A., Euhus, D. M., & Parmigiani, G. (2006). Characterization of BRCA1 and BRCA2 Mutations in a Large United States Sample. *Journal of Clinical Oncology*, *24*(6), 863–871. https://doi.org/10.1200/JCO.2005.03.6772

Chen, W., Morabito, S. J., Kessenbrock, K., Enver, T., Meyer, K. B., & Teschendorff, A. E. (2019). Single-cell landscape in mammary epithelium reveals bipotent-like cells associated with breast cancer risk and outcome. *Communications Biology*, *2*(1), 306. https://doi.org/10.1038/s42003-019-0554-8

Cheung, K. J., Gabrielson, E., Werb, Z., & Ewald, A. J. (2013). Collective invasion in breast cancer requires a conserved basal epithelial program. *Cell*, *155*(7), 1639–1651. https://doi.org/10.1016/j.cell.2013.11.029

Chung, C.-Y., Ma, Z., Dravis, C., Preissl, S., Poirion, O., Luna, G., Hou, X., Giraddi, R. R., Ren, B., & Wahl, G. M. (2019). Single-Cell Chromatin Analysis of Mammary Gland Development Reveals Cell-State Transcriptional Regulators and Lineage Relationships. *Cell Reports*, *29*(2), 495-510.e6. https://doi.org/10.1016/j.celrep.2019.08.089

Ciriello, G., Gatza, M. L., Beck, A. H., Wilkerson, M. D., Rhie, S. K., Pastore, A., Zhang, H., McLellan, M., Yau, C., Kandoth, C., Bowlby, R., Shen, H., Hayat, S., Fieldhouse, R., Lester, S. C., Tse, G. M. K., Factor, R. E., Collins, L. C., Allison, K. H., … Zmuda, E. (2015). Comprehensive Molecular Portraits of Invasive Lobular Breast Cancer. *Cell*,

163(2), 506–519. https://doi.org/10.1016/j.cell.2015.09.033

Cowell, C. F., Weigelt, B., Sakr, R. A., Ng, C. K., Hicks, J., King, T. A., & Reis-Filho, J. S. (2013). Progression from ductal carcinoma in situ to invasive breast cancer: revisited. *Mol Oncol*, *7*(5), 859–869. https://doi.org/10.1016/j.molonc.2013.07.005

Davis, F. M., Lloyd-Lewis, B., Harris, O. B., Kozar, S., Winton, D. J., Muresan, L., & Watson, C. J. (2016). Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun*, *7*, 13053. https://doi.org/10.1038/ncomms13053

Debnath, J., & Brugge, J. S. (2005). Modelling glandular epithelial cancers in three-dimensional cultures. *Nature Reviews Cancer*, *5*(9), 675–688. https://doi.org/10.1038/nrc1695

Dejmek, J., Leandersson, K., Manjer, J., Bjartell, A., Emdin, S. O., Vogel, W. F., Landberg, G., & Andersson, T. (2005). Expression and signaling activity of Wnt-5a/discoidin domain receptor-1 and Syk plays distinct but decisive roles in breast cancer patient survival. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, *11*(2 Pt 1), 520–528. http://www.ncbi.nlm.nih.gov/pubmed/15701836

DeLisser, H. M., Newman, P. J., & Albelda, S. M. (1994). Molecular and functional aspects of PECAM-1/CD31. *Immunology Today*, *15*(10), 490–495. https://doi.org/10.1016/0167-5699(94)90195-3

DeSantis, C. E., Ma, J., Gaudet, M. M., Newman, L. A., Miller, K. D., Goding Sauer, A., Jemal, A., & Siegel, R. L. (2019). Breast cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*, *69*(6), 438–451. https://doi.org/10.3322/caac.21583

Desmedt, C., Zoppoli, G., Gundem, G., Pruneri, G., Larsimont, D., Fornili, M., Fumagalli, D., Brown, D., Rothé, F., Vincent, D., Kheddoumi, N., Rouas, G., Majjaj, S., Brohée, S., Van Loo, P., Maisonneuve, P., Salgado, R., Van Brussel, T., Lambrechts, D., … Sotiriou, C. (2016). Genomic Characterization of Primary Invasive Lobular Breast Cancer. *Journal of Clinical Oncology*, *34*(16), 1872–1881. https://doi.org/10.1200/JCO.2015.64.0334

Deugnier, M.-A., Moiseyeva, E. P., Thiery, J. P., & Glukhova, M. (1995). Myoepithelial cell diffeentiation in the developing mammary gland: Progressive acquisition of smooth muscle phenotype. *Developmental Dynamics*, *204*(2), 107–117. https://doi.org/10.1002/aja.1002040202

Domagala, W., Lasota, J., Bartkowiak, J., Weber, K., & Osborn, M. (1990). Vimentin is preferentially expressed in human breast carcinomas with low estrogen receptor and high Ki-67 growth fraction. *The American Journal of Pathology*, *136*(1), 219–227. http://www.ncbi.nlm.nih.gov/pubmed/2153347

Dongre, A., & Weinberg, R. A. (2019). New insights into the mechanisms of epithelial–mesenchymal transition and implications for cancer. *Nature Reviews Molecular Cell Biology*, *20*(2), 69–84. https://doi.org/10.1038/s41580-018-0080-4

Dontu, G., Al-Hajj, M., Abdallah, W. M., Clarke, M. F., & Wicha, M. S. (2003). Stem cells in normal breast development and breast cancer. *Cell Proliferation*, *36*, 59–72. https://doi.org/10.1046/j.1365-2184.36.s.1.6.x

Dravis, C., Chung, C. Y., Lytle, N. K., Herrera-Valdez, J., Luna, G., Trejo, C. L., Reya, T., & Wahl, G. M. (2018). Epigenetic and Transcriptomic Profiling of Mammary Gland Development and Tumor Models Disclose Regulators of Cell State Plasticity. *Cancer Cell*, *34*(3), 466-482 e6. https://doi.org/10.1016/j.ccell.2018.08.001

Driehuis, E., Kretzschmar, K., & Clevers, H. (2020). Establishment of patient-derived cancer organoids for drug-screening applications. *Nature Protocols*, *15*(10), 3380–3409. https://doi.org/10.1038/s41596-020-0379-4

Egeblad, M., Rasch, M. G., & Weaver, V. M. (2010). Dynamic interplay between the collagen scaffold and tumor evolution. *Curr Opin Cell Biol*, *22*(5), 697–706. https://doi.org/10.1016/j.ceb.2010.08.015

Eirew, P., Stingl, J., Raouf, A., Turashvili, G., Aparicio, S., Emerman, J. T., & Eaves, C. J. (2008). A method for quantifying normal human mammary epithelial stem cells with in vivo regenerative ability. *Nat Med*, *14*(12), 1384–1389. https://doi.org/10.1038/nm.1791

Elias, S., Morgan, M. A., Bikoff, E. K., & Robertson, E. J. (2017). Long-lived unipotent Blimp1-positive luminal stem cells drive mammary gland organogenesis throughout adult life. *Nature Communications*, *8*(1), 1714. https://doi.org/10.1038/s41467-017-01971-w

Elston, C. W., & Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology*, *19*(5), 403–410. https://doi.org/10.1111/j.1365-2559.1991.tb00229.x

Ertel, A., Verghese, A., Byers, S. W., Ochs, M., & Tozeren, A. (2006). Pathway-specific differences between tumor cell lines and normal and tumor tissue cells. *Molecular Cancer*, *5*(1), 55. https://doi.org/10.1186/1476-4598-5-55

Ford, D. (1994). Risks of cancer in BRCA1-mutation carriers. *The Lancet*, *343*(8899), 692–695. https://doi.org/10.1016/S0140-6736(94)91578-4

Foulkes, W. D. (2004). BRCA1 functions as a breast stem cell regulator. *Journal of Medical Genetics*, *41*(1), 1–5. https://doi.org/10.1136/jmg.2003.013805

Friedl, P., & Gilmour, D. (2009). Collective cell migration in morphogenesis, regeneration and cancer. *Nat Rev Mol Cell Biol*, *10*(7), 445–457. https://doi.org/10.1038/nrm2720

Friedl, P., & Wolf, K. (2008). Tube travel: the role of proteases in individual and collective cancer cell invasion. *Cancer Res*, *68*(18), 7247–7249. https://doi.org/10.1158/0008-5472.CAN-08-0784

Gamallo, C., Palacios, J., Suarez, A., Pizarro, A., Navarro, P., Quintanilla, M., & Cano, A. (1993). Correlation of E-cadherin expression with differentiation grade and histological type in breast carcinoma. *Am J Pathol*, *142*(4), 987–993. https://www.ncbi.nlm.nih.gov/pubmed/7682767

Gandalovičová, A., Rosel, D., Fernandes, M., Veselý, P., Heneberg, P., Čermák, V., Petruželka, L., Kumar, S., Sanz-Moreno, V., & Brábek, J. (2017). Migrastatics—Anti-metastatic and Anti-invasion Drugs: Promises and Challenges. *Trends in Cancer*, *3*(6), 391–406. https://doi.org/10.1016/j.trecan.2017.04.008

Ganz, H. M., Buchmann, B., Engelbrecht, L. K., Jesinghaus, M., Eichelberger, L., Gabka, C. J., Schmidt, G. P., Muckenhuber, A., Weichert, W., Bausch, A. R., & Scheel, C. H. (2021). Generation of ductal organoids from normal mammary luminal cells reveals invasive potential. *The Journal of Pathology*. https://doi.org/10.1002/path.5790

Geraldo, S., Simon, A., & Vignjevic, D. M. (2013). Revealing the Cytoskeletal Organization of Invasive Cancer Cells in 3D. *Journal of Visualized Experiments*, *80*. https://doi.org/10.3791/50763

Gill, J. K., Maskarinec, G., Pagano, I., & Kolonel, L. N. (2006). The association of mammographic density with ductal carcinoma in situof the breast: the Multiethnic Cohort. *Breast Cancer Research*, *8*(3), R30. https://doi.org/10.1186/bcr1507

Gillet, J.-P., Calcagno, A. M., Varma, S., Marino, M., Green, L. J., Vora, M. I., Patel, C.,

Orina, J. N., Eliseeva, T. A., Singal, V., Padmanabhan, R., Davidson, B., Ganapathi, R., Sood, A. K., Rueda, B. R., Ambudkar, S. V., & Gottesman, M. M. (2011). Redefining the relevance of established cancer cell lines to the study of mechanisms of clinical anti-cancer drug resistance. *Proceedings of the National Academy of Sciences*, *108*(46), 18708–18713. https://doi.org/10.1073/pnas.1111840108

Giraddi, R. R., Chung, C. Y., Heinz, R. E., Balcioglu, O., Novotny, M., Trejo, C. L., Dravis, C., Hagos, B. M., Mehrabad, E. M., Rodewald, L. W., Hwang, J. Y., Fan, C., Lasken, R., Varley, K. E., Perou, C. M., Wahl, G. M., & Spike, B. T. (2018). Single-Cell Transcriptomes Distinguish Stem Cell State Changes and Lineage Specification Programs in Early Mammary Gland Development. *Cell Rep*, *24*(6), 1653-1666 e7. https://doi.org/10.1016/j.celrep.2018.07.025

Goldhirsch, A., Winer, E. P., Coates, A. S., Gelber, R. D., Piccart-Gebhart, M., Thürlimann, B., Senn, H.-J., Albain, K. S., André, F., Bergh, J., Bonnefoi, H., Bretel-Morales, D., Burstein, H., Cardoso, F., Castiglione-Gertsch, M., Coates, A. S., Colleoni, M., Costa, A., Curigliano, G., … Wood, W. C. (2013). Personalizing the treatment of women with early breast cancer: highlights of the St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2013. *Annals of Oncology*, *24*(9), 2206–2223. https://doi.org/10.1093/annonc/mdt303

Goldhirsch, A., Wood, W. C., Coates, A. S., Gelber, R. D., Thurlimann, B., Senn, H. J., & Panel, members. (2011). Strategies for subtypes--dealing with the diversity of breast cancer: highlights of the St. Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2011. *Ann Oncol*, *22*(8), 1736–1747. https://doi.org/10.1093/annonc/mdr304

Gudjonsson, T., Rønnov-Jessen, L., Villadsen, R., Rank, F., Bissell, M. J., & Petersen, O. W. (2002). Normal and tumor-derived myoepithelial cells differ in their ability to interact with luminal breast epithelial cells for polarity and basement membrane deposition. *J Cell Sci*, *115: 39-50*.

Gudjonsson, T., Villadsen, R., Nielsen, H. L., Ronnov-Jessen, L., Bissell, M. J., & Petersen, O. W. (2002). Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. *Genes Dev*, *16*(6), 693–706. https://doi.org/10.1101/gad.952602

Guo, Y., Yu, P., Liu, Z., Maimaiti, Y., Chen, C., Zhang, Y., Yin, X., Wang, S., Liu, C., & Huang, T. (2017). Prognostic and clinicopathological value of GATA binding protein 3 in breast cancer: A systematic review and meta-analysis. *PloS One*, *12*(4), e0174843.

https://doi.org/10.1371/journal.pone.0174843

Haaksma, C. J., Schwartz, R. J., & Tomasek, J. J. (2011). Myoepithelial Cell Contraction and Milk Ejection Are Impaired in Mammary Glands of Mice Lacking Smooth Muscle Alpha-Actin1. *Biology of Reproduction*, *85*(1), 13–21. https://doi.org/10.1095/biolreprod.110.090639

Habel, L. A., Dignam, J. J., Land, S. R., Salane, M., Capra, A. M., & Julian, T. B. (2004). Mammographic Density and Breast Cancer After Ductal Carcinoma In Situ. *JNCI Journal of the National Cancer Institute*, *96*(19), 1467–1472. https://doi.org/10.1093/jnci/djh260

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, *144*(5), 646–674. https://doi.org/10.1016/j.cell.2011.02.013

Hermiston, M. L., Xu, Z., & Weiss, A. (2003). CD45: A Critical Regulator of Signaling Thresholds in Immune Cells. *Annual Review of Immunology*, *21*(1), 107–137. https://doi.org/10.1146/annurev.immunol.21.120601.140946

Hernandez, L., Wilkerson, P. M., Lambros, M. B., Campion-Flora, A., Rodrigues, D. N., Gauthier, A., Cabral, C., Pawar, V., Mackay, A., A'Hern, R., Marchiò, C., Palacios, J., Natrajan, R., Weigelt, B., & Reis-Filho, J. S. (2012). Genomic and mutational profiling of ductal carcinomas in situ and matched adjacent invasive breast cancers reveals intra-tumour genetic heterogeneity and clonal selection. *The Journal of Pathology*, *227*(1), 42–52. https://doi.org/10.1002/path.3990

Hu, M., Yao, J., Carroll, D. K., Weremowicz, S., Chen, H., Carrasco, D., Richardson, A., Violette, S., Nikolskaya, T., Nikolsky, Y., Bauerlein, E. L., Hahn, W. C., Gelman, R. S., Allred, C., Bissell, M. J., Schnitt, S., & Polyak, K. (2008). Regulation of in situ to invasive breast carcinoma transition. *Cancer Cell*, *13*(5), 394–406. https://doi.org/10.1016/j.ccr.2008.03.007

Hu, Y., & Smyth, G. K. (2009). ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. *J Immunol Methods*, *347*(1–2), 70–78. https://doi.org/10.1016/j.jim.2009.06.008

Iorio, F., Knijnenburg, T. A., Vis, D. J., Bignell, G. R., Menden, M. P., Schubert, M., Aben, N., Gonçalves, E., Barthorpe, S., Lightfoot, H., Cokelaer, T., Greninger, P., van Dyk, E., Chang, H., de Silva, H., Heyn, H., Deng, X., Egan, R. K., Liu, Q., … Garnett, M. J. (2016). A Landscape of Pharmacogenomic Interactions in Cancer. *Cell*, *166*(3), 740–754. https://doi.org/10.1016/j.cell.2016.06.017

Jena, M. K., Jaswal, S., Kumar, S., & Mohanty, A. K. (2019). Molecular mechanism of mammary gland involution: An update. *Developmental Biology*, *445*(2), 145–155. https://doi.org/10.1016/j.ydbio.2018.11.002

Johansson, A. L. V., Trewin, C. B., Fredriksson, I., Reinertsen, K. V., Russnes, H., & Ursin, G. (2021). In modern times, how important are breast cancer stage, grade and receptor subtype for survival: a population-based cohort study. *Breast Cancer Research*, *23*(1), 17. https://doi.org/10.1186/s13058-021-01393-z

Jönsson, M., Dejmek, J., Bendahl, P.-O., & Andersson, T. (2002). Loss of Wnt-5a protein is associated with early relapse in invasive ductal breast carcinomas. *Cancer Research*, *62*(2), 409–416. http://www.ncbi.nlm.nih.gov/pubmed/11809689

Kalluri, R., & Weinberg, R. A. (2009). The basics of epithelial-mesenchymal transition. *Journal of Clinical Investigation*, *119*(6), 1420–1428. https://doi.org/10.1172/JCI39104

Kas, S. M., de Ruiter, J. R., Schipper, K., Annunziato, S., Schut, E., Klarenbeek, S., Drenth, A. P., van der Burg, E., Klijn, C., ten Hoeve, J. J., Adams, D. J., Koudijs, M. J., Wesseling, J., Nethe, M., Wessels, L. F. A., & Jonkers, J. (2017). Insertional mutagenesis identifies drivers of a novel oncogenic pathway in invasive lobular breast carcinoma. *Nature Genetics*, *49*(8), 1219–1230. https://doi.org/10.1038/ng.3905

Kessenbrock, K., Plaks, V., & Werb, Z. (2010). Matrix metalloproteinases: regulators of the tumor microenvironment. *Cell*, *141*(1), 52–67. https://doi.org/10.1016/j.cell.2010.03.015

Kim, S. Y., Jung, S.-H., Kim, M. S., Baek, I.-P., Lee, S. H., Kim, T.-M., Chung, Y.-J., & Lee, S. H. (2015). Genomic differences between pure ductal carcinoma in situ and synchronous ductal carcinoma in situ with invasive breast cancer. *Oncotarget*, *6*(10), 7597–7607. https://doi.org/10.18632/oncotarget.3162

Kinzler, K. W., & Vogelstein, B. (1997). Gatekeepers and caretakers. *Nature*, *386*(6627), 761–763. https://doi.org/10.1038/386761a0

Koren, S., Reavie, L., Couto, J. P., De Silva, D., Stadler, M. B., Roloff, T., Britschgi, A., Eichlisberger, T., Kohler, H., Aina, O., Cardiff, R. D., & Bentires-Alj, M. (2015). PIK3CAH1047R induces multipotency and multi-lineage mammary tumours. *Nature*, *525*(7567), 114–118. https://doi.org/10.1038/nature14669

Korsching, E., Packeisen, J., Liedtke, C., Hungermann, D., Wülfing, P., van Diest, P. J., Brandt, B., Boecker, W., & Buerger, H. (2005). The origin of vimentin expression in

invasive breast cancer: epithelial–mesenchymal transition, myoepithelial histogenesis or histogenesis from progenitor cells with bilinear differentiation potential? *The Journal of Pathology*, *206*(4), 451–457. https://doi.org/10.1002/path.1797

Krøigård, A. B., Larsen, M. J., Lænkholm, A.-V., Knoop, A. S., Jensen, J. D., Bak, M., Mollenhauer, J., Kruse, T. A., & Thomassen, M. (2015). Clonal expansion and linear genome evolution through breast cancer progression from pre-invasive stages to asynchronous metastasis. *Oncotarget*, *6*(8), 5634–5649. https://doi.org/10.18632/oncotarget.3111

Lagace, R., Grimaud, J. A., Schurch, W., & Seemayer, T. A. (1985). Myofibroblastic stromal reaction in carcinoma of the breast: variations of collagenous matrix and structural glycoproteins. *Virchows Arch A Pathol Anat Histopathol*, *408*(1), 49–59. https://doi.org/10.1007/BF00739962

Lancaster, M. A., & Knoblich, J. A. (2014). Organogenesis in a dish: Modeling development and disease using organoid technologies. *Science*, *345*(6194), 1247125–1247125. https://doi.org/10.1126/science.1247125

Leandersson, K., Riesbeck, K., & Andersson, T. (2006). Wnt-5a mRNA translation is suppressed by the Elav-like protein HuR in human breast epithelial cells. *Nucleic Acids Research*, *34*(14), 3988–3999. https://doi.org/10.1093/nar/gkl571

Levental, K. R., Yu, H., Kass, L., Lakins, J. N., Egeblad, M., Erler, J. T., Fong, S. F., Csiszar, K., Giaccia, A., Weninger, W., Yamauchi, M., Gasser, D. L., & Weaver, V. M. (2009). Matrix crosslinking forces tumor progression by enhancing integrin signaling. *Cell*, *139*(5), 891–906. https://doi.org/10.1016/j.cell.2009.10.027

Li, C. I., Anderson, B. O., Daling, J. R., & Moe, R. E. (2003). Trends in incidence rates of invasive lobular and ductal breast carcinoma. *JAMA*, *289*(11), 1421–1424. https://doi.org/10.1001/jama.289.11.1421

Lim, E., Vaillant, F., Wu, D., Forrest, N. C., Pal, B., Hart, A. H., Asselin-Labat, M. L., Gyorki, D. E., Ward, T., Partanen, A., Feleppa, F., Huschtscha, L. I., Thorne, H. J., kConFab, Fox, S. B., Yan, M., French, J. D., Brown, M. A., Smyth, G. K., … Lindeman, G. J. (2009). Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. *Nat Med*, *15*(8), 907–913. https://doi.org/10.1038/nm.2000

Linnemann, J. R., Meixner, L. K., Miura, H., & Scheel, C. H. (2017). *An Organotypic 3D Assay for Primary Human Mammary Epithelial Cells that Recapitulates Branching*

Morphogenesis (pp. 125–137). https://doi.org/10.1007/978-1-4939-7021-6_9

Linnemann, J. R., Miura, H., Meixner, L. K., Irmler, M., Kloos, U. J., Hirschi, B., Bartsch, H. S., Sass, S., Beckers, J., Theis, F. J., Gabka, C., Sotlar, K., & Scheel, C. . H. (2015). Quantification of regenerative potential in primary human mammary epithelial cells. *Development*, *142*(18), 3239–3251. https://doi.org/10.1242/dev.123554

Livasy, C. A., Karaca, G., Nanda, R., Tretiakova, M. S., Olopade, O. I., Moore, D. T., & Perou, C. M. (2006). Phenotypic evaluation of the basal-like subtype of invasive breast carcinoma. *Modern Pathology*, *19*(2), 264–271. https://doi.org/10.1038/modpathol.3800528

Lopez-Garcia, M. A., Geyer, F. C., Lacroix-Triki, M., Marchió, C., & Reis-Filho, J. S. (2010). Breast cancer precursors revisited: molecular features and progression pathways. *Histopathology*, *57*(2), 171–192. https://doi.org/10.1111/j.1365-2559.2010.03568.x

Ma, X.-J., Salunga, R., Tuggle, J. T., Gaudet, J., Enright, E., McQuary, P., Payette, T., Pistone, M., Stecker, K., Zhang, B. M., Zhou, Y.-X., Varnholt, H., Smith, B., Gadd, M., Chatfield, E., Kessler, J., Baer, T. M., Erlander, M. G., & Sgroi, D. C. (2003). Gene expression profiles of human breast cancer progression. *Proceedings of the National Academy of Sciences*, *100*(10), 5974–5979. https://doi.org/10.1073/pnas.0931261100

Macias, H., & Hinck, L. (2012). Mammary gland development. *Wiley Interdisciplinary Reviews. Developmental Biology*, *1*(4), 533–557. https://doi.org/10.1002/wdev.35

Makarem, M., Kannan, N., Nguyen, L. V., Knapp, D. J. H. F., Balani, S., Prater, M. D., Stingl, J., Raouf, A., Nemirovsky, O., Eirew, P., & Eaves, C. J. (2013). Developmental Changes in the in Vitro Activated Regenerative Activity of Primitive Mammary Epithelial Cells. *PLoS Biology*, *11*(8), e1001630. https://doi.org/10.1371/journal.pbio.1001630

Malkin, D., Li, F., Strong, L., Fraumeni, J., Nelson, C., Kim, D., Kassel, J., Gryka, M., Bischoff, F., Tainsky, M., & Et, A. (1990). Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science*, *250*(4985), 1233–1238. https://doi.org/10.1126/science.1978757

Martin Carli, J. F., Trahan, G. D., Jones, K. L., Hirsch, N., Rolloff, K. P., Dunn, E. Z., Friedman, J. E., Barbour, L. A., Hernandez, T. L., MacLean, P. S., Monks, J., McManaman, J. L., & Rudolph, M. C. (2020). Single Cell RNA Sequencing of Human Milk-Derived Cells Reveals Sub-Populations of Mammary Epithelial Cells with Molecular Signatures of Progenitor and Mature States: a Novel, Non-invasive Framework for Investigating Human Lactation Physiology. *Journal of Mammary Gland*

Biology and Neoplasia, *25*(4), 367–387. https://doi.org/10.1007/s10911-020-09466-z

Masciari, S., Larsson, N., Senz, J., Boyd, N., Kaurah, P., Kandel, M. J., Harris, L. N., Pinheiro, H. C., Troussard, A., Miron, P., Tung, N., Oliveira, C., Collins, L., Schnitt, S., Garber, J. E., & Huntsman, D. (2007). Germline E-cadherin mutations in familial lobular breast cancer. *Journal of Medical Genetics*, *44*(11), 726–731. https://doi.org/10.1136/jmg.2007.051268

Matsukita, S., Nomoto, M., Kitajima, S., Tanaka, S., Goto, M., Irimura, T., Kim, Y. S., Sato, E., & Yonezawa, S. (2003). Expression of mucins (MUC1, MUC2, MUC5AC and MUC6) in mucinous carcinoma of the breast: comparison with invasive ductal carcinoma. *Histopathology*, *42*(1), 26–36. https://doi.org/10.1046/j.1365-2559.2003.01530.x

Mavaddat, N., Peock, S., Frost, D., Ellis, S., Platte, R., Fineberg, E., Evans, D. G., Izatt, L., Eeles, R. A., Adlard, J., Davidson, R., Eccles, D., Cole, T., Cook, J., Brewer, C., Tischkowitz, M., Douglas, F., Hodgson, S., Walker, L., … Easton, D. F. (2013). Cancer Risks for BRCA1 and BRCA2 Mutation Carriers: Results From Prospective Analysis of EMBRACE. *JNCI: Journal of the National Cancer Institute*, *105*(11), 812–822. https://doi.org/10.1093/jnci/djt095

McCart Reed, A. E., Kutasovic, J. R., Lakhani, S. R., & Simpson, P. T. (2015). Invasive lobular carcinoma of the breast: morphology, biomarkers and 'omics. *Breast Cancer Res*, *17*, 12. https://doi.org/10.1186/s13058-015-0519-x

Melchor, L., & Benitez, J. (2008). An integrative hypothesis about the origin and development of sporadic and familial breast cancer subtypes. *Carcinogenesis*, *29*(8), 1475–1482. https://doi.org/10.1093/carcin/bgn157

Moll, R., Mitze, M., Frixen, U. H., & Birchmeier, W. (1993). Differential loss of E-cadherin expression in infiltrating ductal and lobular breast carcinomas. *Am J Pathol*, *143*(6), 1731–1742. https://www.ncbi.nlm.nih.gov/pubmed/8256859

Molyneux, G., Geyer, F. C., Magnay, F.-A., McCarthy, A., Kendrick, H., Natrajan, R., MacKay, A., Grigoriadis, A., Tutt, A., Ashworth, A., Reis-Filho, J. S., & Smalley, M. J. (2010). BRCA1 Basal-like Breast Cancers Originate from Luminal Epithelial Progenitors and Not from Basal Stem Cells. *Cell Stem Cell*, *7*(3), 403–417. https://doi.org/10.1016/j.stem.2010.07.010

Moon, H. R., Ospina-Munoz, N., Noe-Kim, V., Yang, Y., Elzey, B. D., Konieczny, S. F., & Han, B. (2020). Subtype-specific characterization of breast cancer invasion using a

microfluidic tumor platform. *Plos One*, *15*(6), e0234012. https://doi.org/10.1371/journal.pone.0234012

Muschler, J., & Streuli, C. H. (2010). Cell-matrix interactions in mammary gland development and breast cancer. *Cold Spring Harb Perspect Biol*, *2*(10), a003202. https://doi.org/10.1101/cshperspect.a003202

Naylor, M. J., & Ormandy, C. J. (2007). Gata-3 and mammary cell fate. *Breast Cancer Research : BCR*, *9*(2), 302. https://doi.org/10.1186/bcr1661

Newburger, D. E., Kashef-Haghighi, D., Weng, Z., Salari, R., Sweeney, R. T., Brunner, A. L., Zhu, S. X., Guo, X., Varma, S., Troxell, M. L., West, R. B., Batzoglou, S., & Sidow, A. (2013). Genome evolution during progression to breast cancer. *Genome Research*, *23*(7), 1097–1108. https://doi.org/10.1101/gr.151670.112

Nguyen-Ngoc, K.-V., Cheung, K. J., Brenot, A., Shamir, E. R., Gray, R. S., Hines, W. C., Yaswen, P., Werb, Z., & Ewald, A. J. (2012). ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proceedings of the National Academy of Sciences*, *109*(39), E2595–E2604. https://doi.org/10.1073/pnas.1212834109

Nguyen, D.-A. D., & Neville, M. C. (1998). Tight junction regulation in the mammary gland. *Journal of Mammary Gland Biology and Neoplasia*, *3*(3), 233–246. https://doi.org/10.1023/a:1018707309361

Nguyen, Q. H., Pervolarakis, N., Blake, K., Ma, D., Davis, R. T., James, N., Phung, A. T., Willey, E., Kumar, R., Jabart, E., Driver, I., Rock, J., Goga, A., Khan, S. A., Lawson, D. A., Werb, Z., & Kessenbrock, K. (2018). Profiling human breast epithelial cells using single cell RNA sequencing identifies cell diversity. *Nature Communications*, *9*(1), 2028. https://doi.org/10.1038/s41467-018-04334-1

Packwood, K., Martland, G., Sommerlad, M., Shaw, E., Moutasim, K., Thomas, G., Bateman, A. C., Jones, L., Haywood, L., Evans, D. G., Birch, J. M., Alsalmi, O. A., Henderson, A., Poplawski, N., & Eccles, D. M. (2019). Breast cancer in patients with germline TP53 pathogenic variants have typical tumour characteristics: the Cohort study of TP53 carrier early onset breast cancer (COPE study). *The Journal of Pathology: Clinical Research*, *5*(3), 189–198. https://doi.org/10.1002/cjp2.133

Pal, B., Chen, Y., Vaillant, F., Jamieson, P., Gordon, L., Rios, A. C., Wilcox, S., Fu, N., Liu, K. H., Jackling, F. C., Davis, M. J., Lindeman, G. J., Smyth, G. K., & Visvader, J. E. (2017). Construction of developmental lineage relationships in the mouse mammary

gland by single-cell RNA profiling. *Nature Communications*, *8*(1), 1627. https://doi.org/10.1038/s41467-017-01560-x

Patey, D. H., & Scarff, R. W. (1928). Histology in the prognosis of carcinoma of the breast. *Lancet*, *1*, pp-801-804.

Pechoux, C., Gudjonsson, T., Ronnov-Jessen, L., Bissell, M. J., & Petersen, O. W. (1999). Human mammary luminal epithelial cells contain progenitors to myoepithelial cells. *Dev Biol*, *206*(1), 88–99. https://doi.org/10.1006/dbio.1998.9133

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A., Fluge, O., Pergamenschikov, A., Williams, C., Zhu, S. X., Lonning, P. E., Borresen-Dale, A. L., Brown, P. O., & Botstein, D. (2000). Molecular portraits of human breast tumours. *Nature*, *406*(6797), 747–752. https://doi.org/10.1038/35021093

Petridis, C., Brook, M. N., Shah, V., Kohut, K., Gorman, P., Caneppele, M., Levi, D., Papouli, E., Orr, N., Cox, A., Cross, S. S., Dos-Santos-Silva, I., Peto, J., Swerdlow, A., Schoemaker, M. J., Bolla, M. K., Wang, Q., Dennis, J., Michailidou, K., … Sawyer, E. J. (2016). Genetic predisposition to ductal carcinoma in situ of the breast. *Breast Cancer Res*, *18*(1), 22. https://doi.org/10.1186/s13058-016-0675-7

Peuhu, E., Virtakoivu, R., Mai, A., Wärri, A., & Ivaska, J. (2017). Epithelial vimentin plays a functional role in mammary gland development. *Development*, *144*(22), 4103–4113. https://doi.org/10.1242/dev.154229

Pharoah, P. D. P., Day, N. E., & Caldas, C. (1999). Somatic mutations in the p53 gene and prognosis in breast cancer: a meta-analysis. *British Journal of Cancer*, *80*(12), 1968–1973. https://doi.org/10.1038/sj.bjc.6690628

Polyak, K. (2011). Heterogeneity in breast cancer. *Journal of Clinical Investigation*, *121*(10), 3786–3788. https://doi.org/10.1172/JCI60534

Polyak, K., & Kalluri, R. (2010). The Role of the Microenvironment in Mammary Gland Development and Cancer. *Cold Spring Harbor Perspectives in Biology*, *2*(11), a003244–a003244. https://doi.org/10.1101/cshperspect.a003244

Porter, D., Lahti-Domenici, J., Keshaviah, A., Bae, Y. K., Argani, P., Marks, J., Richardson, A., Cooper, A., Strausberg, R., Riggins, G. J., Schnitt, S., Gabrielson, E., Gelman, R., & Polyak, K. (2003). Molecular markers in ductal carcinoma in situ of the breast. *Molecular Cancer Research : MCR*, *1*(5), 362–375.

http://www.ncbi.nlm.nih.gov/pubmed/12651909

Prasad, C. P., Manchanda, M., Mohapatra, P., & Andersson, T. (2018). WNT5A as a therapeutic target in breast cancer. *Cancer and Metastasis Reviews*, *37*(4), 767–778. https://doi.org/10.1007/s10555-018-9760-y

Prater, M. D., Petit, V., Alasdair Russell, I., Giraddi, R. R., Shehata, M., Menon, S., Schulte, R., Kalajzic, I., Rath, N., Olson, M. F., Metzger, D., Faraldo, M. M., Deugnier, M. A., Glukhova, M. A., & Stingl, J. (2014). Mammary stem cells have myoepithelial cell properties. *Nat Cell Biol*, *16*(10), 1-7,942-950. https://doi.org/10.1038/ncb3025

Provenzano, P. P., Eliceiri, K. W., Campbell, J. M., Inman, D. R., White, J. G., & Keely, P. J. (2006). Collagen reorganization at the tumor-stromal interface facilitates local invasion. *BMC Med*, *4*(1), 38. https://doi.org/10.1186/1741-7015-4-38

Provenzano, P. P., Inman, D. R., Eliceiri, K. W., Knittel, J. G., Yan, L., Rueden, C. T., White, J. G., & Keely, P. J. (2008). Collagen density promotes mammary tumor initiation and progression. *BMC Medicine*, *6*(1), 11. https://doi.org/10.1186/1741-7015-6-11

Provenzano, P. P., Inman, D. R., Eliceiri, K. W., Trier, S. M., & Keely, P. J. (2008). Contact guidance mediated three-dimensional cell migration is regulated by Rho/ROCK-dependent matrix reorganization. *Biophys J*, *95*(11), 5374–5384. https://doi.org/10.1529/biophysj.108.133116

Puck, T. T., & Marcus, P. I. (1955). A rapid method for viable cell titration and clone production with HeLa cells in tissue culture: the use of X-irradiated cells to supply conditioning factors. *Proc Natl Acad Sci USA*.

Rahman, N., Stone, J. G., Coleman, G., Gusterson, B., Seal, S., Marossy, A., Lakhani, S. R., Ward, A., Nash, A., McKinna, A., A'Hern, R., Stratton, M. R., & Houlston, R. S. (2000). Lobular carcinoma in situ of the breast is not caused by constitutional mutations in the E-cadherin gene. *British Journal of Cancer*, *82*(3), 568–570. https://doi.org/10.1054/bjoc.1999.0965

Rattan, B., Manjari, M., Kahlon, S. K., Kalra, N., Bhalla, A., & Paul, S. (2012). The Immunohistochemical Expression of the Oestrogen Receptor (ER), HER-2/NEU and Cytokeratin 8/18 and 5/6 in Invasive Breast Carcinoma. *J Clin Diagn Res*, *6*(9), 1495–1498. https://doi.org/10.7860/JCDR/2012/4086.2542

Ravid-Hermesh, O., Zurgil, N., Shafran, Y., Afrimzon, E., Sobolev, M., Hakuk, Y., Bar-On Eizig, Z., & Deutsch, M. (2018). Analysis of Cancer Cell Invasion and Anti-metastatic

Drug Screening Using Hydrogel Micro-chamber Array (HMCA)-based Plates. *Journal of Visualized Experiments : JoVE*, *140*. https://doi.org/10.3791/58359

Rios, A. C., Fu, N. Y., Lindeman, G. J., & Visvader, J. E. (2014). In situ identification of bipotent stem cells in the mammary gland. *Nature*, *506*(7488), 322–327. https://doi.org/10.1038/nature12948

Rosenbluth, J. M., Schackmann, R. C. J., Gray, G. K., Selfors, L. M., Li, C. M., Boedicker, M., Kuiken, H. J., Richardson, A., Brock, J., Garber, J., Dillon, D., Sachs, N., Clevers, H., & Brugge, J. S. (2020). Organoid cultures from normal and cancer-prone human breast tissues preserve complex epithelial lineages. *Nat Commun*, *11*(1), 1711. https://doi.org/10.1038/s41467-020-15548-7

Sachs, N., de Ligt, J., Kopper, O., Gogola, E., Bounova, G., Weeber, F., Balgobind, A. V, Wind, K., Gracanin, A., Begthel, H., Korving, J., van Boxtel, R., Duarte, A. A., Lelieveld, D., van Hoeck, A., Ernst, R. F., Blokzijl, F., Nijman, I. J., Hoogstraat, M., … Clevers, H. (2018). A Living Biobank of Breast Cancer Organoids Captures Disease Heterogeneity. *Cell*, *172*(1–2), 373-386 e10. https://doi.org/10.1016/j.cell.2017.11.010

Santagata, S., Thakkar, A., Ergonul, A., Wang, B., Woo, T., Hu, R., Harrell, J. C., McNamara, G., Schwede, M., Culhane, A. C., Kindelberger, D., Rodig, S., Richardson, A., Schnitt, S. J., Tamimi, R. M., & Ince, T. A. (2014). Taxonomy of breast cancer based on normal cell phenotype predicts outcome. *J Clin Invest*, *124*(2), 859–870. https://doi.org/10.1172/JCI70941

Sato, T., Vries, R. G., Snippert, H. J., van de Wetering, M., Barker, N., Stange, D. E., van Es, J. H., Abo, A., Kujala, P., Peters, P. J., & Clevers, H. (2009). Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche. *Nature*, *459*(7244), 262–265. https://doi.org/10.1038/nature07935

Schipper, K., Seinstra, D., Paulien Drenth, A., van der Burg, E., Ramovs, V., Sonnenberg, A., van Rheenen, J., Nethe, M., & Jonkers, J. (2019). Rebalancing of actomyosin contractility enables mammary tumor formation upon loss of E-cadherin. *Nature Communications*, *10*(1), 3800. https://doi.org/10.1038/s41467-019-11716-6

Schrader, K. A., Masciari, S., Boyd, N., Salamanca, C., Senz, J., Saunders, D. N., Yorida, E., Maines-Bandiera, S., Kaurah, P., Tung, N., Robson, M. E., Ryan, P. D., Olopade, O. I., Domchek, S. M., Ford, J., Isaacs, C., Brown, P., Balmana, J., Razzak, A. R., … Huntsman, D. G. (2011). Germline mutations in CDH1 are infrequent in women with early-onset or familial lobular breast cancers. *Journal of Medical Genetics*, *48*(1), 64–

68. https://doi.org/10.1136/jmg.2010.079814

Serra, R., Easter, S. L., Jiang, W., & Baxley, S. E. (2011). Wnt5a as an Effector of TGFβ in Mammary Development and Cancer. *Journal of Mammary Gland Biology and Neoplasia*, *16*(2), 157–167. https://doi.org/10.1007/s10911-011-9205-5

Shaoxian, T., Baohua, Y., Xiaoli, X., Yufan, C., Xiaoyu, T., Hongfen, L., Rui, B., Xiangjie, S., Ruohong, S., & Wentao, Y. (2017). Characterisation of GATA3 expression in invasive breast cancer: differences in histological subtypes and immunohistochemically defined molecular subtypes. *Journal of Clinical Pathology*, *70*(11), 926–934. https://doi.org/10.1136/jclinpath-2016-204137

Shehata, M., Teschendorff, A., Sharp, G., Novcic, N., Russell, I. A., Avril, S., Prater, M., Eirew, P., Caldas, C., Watson, C. J., & Stingl, J. (2012). Phenotypic and functional characterisation of the luminal cell hierarchy of the mammary gland. *Breast Cancer Res*, *14*(5), R134. https://doi.org/10.1186/bcr3334

Shivtiel, S., Kollet, O., Lapid, K., Schajnovitz, A., Goichberg, P., Kalinkovich, A., Shezen, E., Tesio, M., Netzer, N., Petit, I., Sharir, A., & Lapidot, T. (2008). CD45 regulates retention, motility, and numbers of hematopoietic progenitors, and affects osteoclast remodeling of metaphyseal trabecules. *Journal of Experimental Medicine*, *205*(10), 2381–2395. https://doi.org/10.1084/jem.20080072

Silwal-Pandit, L., Vollan, H. K. M., Chin, S.-F., Rueda, O. M., McKinney, S., Osako, T., Quigley, D. A., Kristensen, V. N., Aparicio, S., Børresen-Dale, A.-L., Caldas, C., & Langerød, A. (2014). TP53 Mutation Spectrum in Breast Cancer Is Subtype Specific and Has Distinct Prognostic Relevance. *Clinical Cancer Research*, *20*(13), 3569–3580. https://doi.org/10.1158/1078-0432.CCR-13-2943

Sobin, L., Gospodarowicz, M., & C, W. (2011). *TNM Classification of Malignant Tumours* (7th ed.). Wiley-Blackwell.

Sokol, E. S., Miller, D. H., Breggia, A., Spencer, K. C., Arendt, L. M., & Gupta, P. B. (2016). Growth of human breast tissues from patient cells in 3D hydrogel scaffolds. *Breast Cancer Res*, *18*(1), 19. https://doi.org/10.1186/s13058-016-0677-5

Song, W., Wang, R., Jiang, W., Yin, Q., Peng, G., Yang, R., Yu, Q. C., Chen, J., Li, J., Cheung, T. H., Jing, N., & Zeng, Y. A. (2019). Hormones induce the formation of luminal-derived basal cells in the mammary gland. *Cell Res*, *29*(3), 206–220. https://doi.org/10.1038/s41422-018-0137-0

Sontag, L., & Axelrod, D. E. (2005). Evaluation of pathways for progression of heterogeneous breast tumors. *Journal of Theoretical Biology*, *232*(2), 179–189. https://doi.org/10.1016/j.jtbi.2004.08.002

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Thorsen, T., Quist, H., Matese, J. C., Brown, P. O., Botstein, D., Lonning, P. E., & Borresen-Dale, A. L. (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A*, *98*(19), 10869–10874. https://doi.org/10.1073/pnas.191367098

Sorlie, T., Tibshirani, R., Parker, J., Hastie, T., Marron, J. S., Nobel, A., Deng, S., Johnsen, H., Pesich, R., Geisler, S., Demeter, J., Perou, C. M., Lonning, P. E., Brown, P. O., Borresen-Dale, A. L., & Botstein, D. (2003). Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc Natl Acad Sci U S A*, *100*(14), 8418–8423. https://doi.org/10.1073/pnas.0932692100

Stingl, J., & Caldas, C. (2007). Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. *Nat Rev Cancer*, *7*(10), 791–799. https://doi.org/10.1038/nrc2212

Stingl, J., Eaves, C. J., Zandieh, I., & Emerman, J. T. (2001). Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue. *Breast Cancer Research and Treatment*, *67*(2), 93–109. https://doi.org/10.1023/A:1010615124301

Stingl, J., Eirew, P., Ricketson, I., Shackleton, M., Vaillant, F., Choi, D., Li, H. I., & Eaves, C. J. (2006). Purification and unique properties of mammary epithelial stem cells. *Nature*, *439*(7079), 993–997. https://doi.org/10.1038/nature04496

Sun, H., Miao, Z., Zhang, X., Chan, U. I., Su, S. M., Guo, S., Wong, C. K. H., Xu, X., & Deng, C.-X. (2018). Single-cell RNA-Seq reveals cell heterogeneity and hierarchy within mouse mammary epithelia. *Journal of Biological Chemistry*, *293*(22), 8315–8329. https://doi.org/10.1074/jbc.RA118.002297

Sun, P., Yuan, Y., Li, A., Li, B., & Dai, X. (2010). Cytokeratin expression during mouse embryonic and early postnatal mammary gland development. *Histochemistry and Cell Biology*, *133*(2), 213–221. https://doi.org/10.1007/s00418-009-0662-5

Sun, Y. S., Zhao, Z., Yang, Z.-N., Xu, F., Lu, H.-J., Zhu, Z.-Y., Shi, W., Jiang, J., Yao, P.-P., & Zhu, H.-P. (2017). Risk Factors and Preventions of Breast Cancer. *International*

Journal of Biological Sciences, *13*(11), 1387–1397. https://doi.org/10.7150/ijbs.21635

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, caac.21660. https://doi.org/10.3322/caac.21660

Tarulli, G. A., Laven-Law, G., Shakya, R., Tilley, W. D., & Hickey, T. E. (2015). Hormone-Sensing Mammary Epithelial Progenitors: Emerging Identity and Hormonal Regulation. *Journal of Mammary Gland Biology and Neoplasia*, *20*(1–2), 75–91. https://doi.org/10.1007/s10911-015-9344-1

Thong, T., Wang, Y., Brooks, M. D., Lee, C. T., Scott, C., Balzano, L., Wicha, M. S., & Colacino, J. A. (2020). Hybrid Stem Cell States: Insights Into the Relationship Between Mammary Development and Breast Cancer Using Single-Cell Transcriptomics. *Frontiers in Cell and Developmental Biology*, *8*. https://doi.org/10.3389/fcell.2020.00288

Turner, N. C., & Reis-Filho, J. S. (2006). Basal-like breast cancer and the BRCA1 phenotype. *Oncogene*, *25*(43), 5846–5853. https://doi.org/10.1038/sj.onc.1209876

Van Amerongen, R., Bowman, A. N., & Nusse, R. (2012). Developmental stage and time dictate the fate of Wnt/beta-catenin-responsive stem cells in the mammary gland. *Cell Stem Cell*, *11*(3), 387–400. https://doi.org/10.1016/j.stem.2012.05.023

Van Keymeulen, A., Rocha, A. S., Ousset, M., Beck, B., Bouvencourt, G., Rock, J., Sharma, N., Dekoninck, S., & Blanpain, C. (2011). Distinct stem cells contribute to mammary gland development and maintenance. *Nature*, *479*(7372), 189–193. https://doi.org/10.1038/nature10573

Van Seijen, M., Lips, E. H., Thompson, A. M., Nik-Zainal, S., Futreal, A., Hwang, E. S., Verschuur, E., Lane, J., Jonkers, J., Rea, D. W., & Wesseling, J. (2019). Ductal carcinoma in situ: to treat or not to treat, that is the question. *British Journal of Cancer*, *121*(4), 285–292. https://doi.org/10.1038/s41416-019-0478-6

Varner, V. D., & Nelson, C. M. (2014). Cellular and physical mechanisms of branching morphogenesis. *Development*, *141*(14), 2750–2759. https://doi.org/10.1242/dev.104794

Visvader, J. E., & Stingl, J. (2014). Mammary stem cells and the differentiation hierarchy: current status and perspectives. *Genes Dev*, *28*(11), 1143–1158.

https://doi.org/10.1101/gad.242511.114

Vogelstein, B., & Kinzler, K. W. (2004). Cancer genes and the pathways they control. *Nature Medicine*, *10*(8), 789–799. https://doi.org/10.1038/nm1087

Wang, D., Cai, C., Dong, X., Yu, Q. C., Zhang, X.-O., Yang, L., & Zeng, Y. A. (2015). Identification of multipotent mammary stem cells by protein C receptor expression. *Nature*, *517*(7532), 81–84. https://doi.org/10.1038/nature13851

Wilding, J. L., & Bodmer, W. F. (2014). Cancer Cell Lines for Drug Discovery and Development. *Cancer Research*, *74*(9), 2377–2384. https://doi.org/10.1158/0008-5472.CAN-13-2971

Winkler, J., Abisoye-Ogunniyan, A., Metcalf, K. J., & Werb, Z. (2020). Concepts of extracellular matrix remodelling in tumour progression and metastasis. *Nature Communications*, *11*(1), 5120. https://doi.org/10.1038/s41467-020-18794-x

Wolfe, J. (1976). Breast patterns as an index of risk for developing breast cancer. *American Journal of Roentgenology*, *126*(6), 1130–1137. https://doi.org/10.2214/ajr.126.6.1130

Wong, C. H., Siah, K. W., & Lo, A. W. (2019). Estimation of clinical trial success rates and related parameters. *Biostatistics*, *20*(2), 273–286. https://doi.org/10.1093/biostatistics/kxx069

Wuidart, A., Ousset, M., Rulands, S., Simons, B. D., Van Keymeulen, A., & Blanpain, C. (2016). Quantitative lineage tracing strategies to resolve multipotency in tissue-specific stem cells. *Genes & Development*, *30*(11), 1261–1277. https://doi.org/10.1101/gad.280057.116

Wuidart, A., Sifrim, A., Fioramonti, M., Matsumura, S., Brisebarre, A., Brown, D., Centonze, A., Dannau, A., Dubois, C., Van Keymeulen, A., Voet, T., & Blanpain, C. (2018). Early lineage segregation of multipotent embryonic mammary gland progenitors. *Nature Cell Biology*, *20*(6), 666–676. https://doi.org/10.1038/s41556-018-0095-2

Xie, Z. M., Li, L. S., Laquet, C., Penault-Llorca, F., Uhrhammer, N., Xie, X. M., & Bignon, Y. J. (2011). Germline mutations of the E-cadherin gene in families with inherited invasive lobular breast carcinoma but no diffuse gastric cancer. *Cancer*, *117*(14), 3112–3117. https://doi.org/10.1002/cncr.25876

Yates, L. R., Gerstung, M., Knappskog, S., Desmedt, C., Gundem, G., Van Loo, P., Aas, T., Alexandrov, L. B., Larsimont, D., Davies, H., Li, Y., Ju, Y. S., Ramakrishna, M.,

Haugland, H. K., Lilleng, P. K., Nik-Zainal, S., McLaren, S., Butler, A., Martin, S., … Campbell, P. J. (2015). Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nature Medicine*, *21*(7), 751–759. https://doi.org/10.1038/nm.3886

Yu, W., Shewan, A. M., Brakeman, P., Eastburn, D. J., Datta, A., Bryant, D. M., Fan, Q. W., Weiss, W. A., Zegers, M. M., & Mostov, K. E. (2008). Involvement of RhoA, ROCK I and myosin II in inverted orientation of epithelial polarity. *EMBO Rep*, *9*(9), 923–929. https://doi.org/10.1038/embor.2008.135

Yurchenco, P. D., & Patton, B. L. (2009). Developmental and pathogenic mechanisms of basement membrane assembly. *Current Pharmaceutical Design*, *15*(12), 1277–1294. https://doi.org/10.2174/138161209787846766

Zaha, D. C. (2014). Significance of immunohistochemistry in breast cancer. *World J Clin Oncol*, *5*(3), 382–392. https://doi.org/10.5306/wjco.v5.i3.382

Zhang, L., Valdez, J. M., Zhang, B., Wei, L., Chang, J., & Xin, L. (2011). ROCK Inhibitor Y-27632 Suppresses Dissociation-Induced Apoptosis of Murine Prostate Stem/Progenitor Cells and Increases Their Cloning Efficiency. *PLoS ONE*, *6*(3), e18271. https://doi.org/10.1371/journal.pone.0018271

Zhou, H., & Kramer, R. H. (2005). Integrin engagement differentially modulates epithelial cell motility by RhoA/ROCK and PAK1. *J Biol Chem*, *280*(11), 10624–10635. https**://**doi.org/10.1074/jbc.M411900200

Zilfou, J. T., & Lowe, S. W. (2009). Tumor Suppressive Functions of p53. *Cold Spring Harbor Perspectives in Biology*, *1*(5), a001883–a001883. https://doi.org/10.1101/cshperspect.a001883

8 Acknowledgements

In this paragraph I want to thank several people whose support has been essential for the successful completion of this PhD thesis.

First of all, I want to thank Heiko Lickert who has not only supported me with scientific advice but also backed me up when I needed support to extend my contract. This support has enabled me to finish all experiments and to bring my work to a point where most parts of it could be published. Furthermore, I want to thank Magdalena Götz for letting me finish my PhD thesis at the Institute of Stem Cell Research.

Special thanks also go to Christina Scheel, who has supervised my work, whether up close or from a distance. I am very happy that she has helped me to conduct and finish my research despite all obstacles.

I also want to acknowledge the support of my second thesis committee member, Stefan Stricker. His novel insights have been a great addition to my work.

Moreover, I want to thank all members of the Scheel group as well as my collaborators, particularly the members of the Bausch group, for the great scientific advice and for making our collaboration fun.

Finally, I want to thank my family, friends and my boyfriend for their emotional support and non-scientific advice during the past years.

9 Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die bei der Fakultät für Medizin der TUM zur Promotionsprüfung vorgelegte Arbeit mit dem Titel „Development and characterization of an in vitro assay that elicits invasive potential from human mammary luminal progenitor cells“ am Helmholtz Zentrum München, Institut für Stammzellforschung unter der Anleitung und Betreuung durch Prof. Dr. Heiko Lickert ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 6 und 7 Satz 2 angegebenen Hilfsmittel benutzt habe.

[X] Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

[X] Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

[O] Die vollständige Dissertation wurde in _____ veröffentlicht. Die promotionsführende Einrichtung _____ hat der Vorveröffentlichung zugestimmt.

[X] Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

[O] Ich habe bereits am _____ bei der Fakultät für _____ der Hochschule _____ unter Vorlage einer Dissertation mit dem Thema _____ die Zulassung zur Promotion beantragt mit dem Ergebnis: _____

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst. Mit der Aufnahme meiner personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich [X] einverstanden, [O] nicht einverstanden.

_____, den _____ _____