The role of non-coding RNA and peroxidasin in breast epithelial progenitor cell plasticity and differentiation

Anna Karen Sigurðardóttir

Thesis for the degree of Philosophiae Doctor

Supervisor:
Þórarinn Guðjónsson

Advisor:
Gunnhildur Ásta Traustadóttir

Doctoral committee:
Bylgja Hilmarsdóttir
Eiríkur Briem
Jón Þór Bergþórsson
Sigríður Valgeirsdóttir

June 2022

UNIVERSITY OF ICELAND
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

Hlutverk non-coding RNA og peroxidasins í frumusérhæfingu og greinóttri formgerð í brjóstafrumulínu með stofnfrumueiginleika

Anna Karen Sigurðardóttir

Ritgerð til doktorsgráðu

Umsjónarkennari:
Þórarinn Guðjónsson

Leiðbeinandi:
Gunnhildur Ásta Traustadóttir

Doktorsnefnd:
Bylgja Hilmarsdóttir
Eiríkur Briem
Jón Þór Bergþórsson
Sigríður Valgeirsdóttir

Júní 2022

UNIVERSITY OF ICELAND
SCHOOL OF HEALTH SCIENCES
FACULTY OF MEDICINE

ISBN: 978-9935-9657-5-2

Printing by Háskólaprent.

Reykjavik, Iceland 2022

Fyrir pabba

Ágrip

Brjóstkirtillinn er greinótt líffæri sem skiptist í bleðla og samanstendur af tvískiptri þekju af hollægum þekjufrumum og vöðvaþekjufrumum. Grunnhimnan, sem meðal annars er rík af kollagen IV, skilur þekjuna frá bandvefnum í kring en í honum má finna æðaþel, bandvefsfrumur og ónæmisfrumur. Brjóstkirtill kvenna nýtur sérstöðu þar sem meirihluti þroskaferils hans á sér stað eftir að fósturskeiði er lokið, aðallega á kynþroskaskeiði, en kirtillinn lýkur ekki fullum þroska fyrr en á meðgöngu. Við þetta greinast kirtileiningar í bleðlum enn frekar og mjólkurmyndandi þekjueiningar myndast. Vísir að þessu ferli á sér stað í hverjum tíðahring þegar kirtillinn býr sig undir mögulega meðgöngu og því skiptast á tímabil uppbyggingar og niðurbrots á meðan á frjósemistímabili kvenna stendur. Þessu ferli er viðhaldið af stofnfrumum og þekjuforverum sem sérhæfast í hollægar kirtilþekjufrumur eða vöðvaþekjufrumur ásamt því að viðhalda eigin stofnum. D492 er frumulína með stofnfrumueiginleika sem einangruð var úr eðlilegum brjóstvef. Frumulínan getur myndað frumur sem hafa hvor um sig eiginleika kirtilþekju og vöðvaþekju og myndar þyrpingar með greinótta formgerð í þrívíðri rækt. Fyrri rannsóknir hafa sýnt að báðar svipgerðir eru nauðsynlegar fyrir greinótta formgerð í þrívíðum þyrpingum D492 frumulínunnar. Þar að auki er bandvefsumbreyting þekjufrumanna í kirtlinum nauðsynleg til að frumurnar geti skriðið inn í umlykjandi stoðvef og myndað greinótta formgerð. Ýmsir þættir koma að stýringu greinóttrar formgerðar, bandvefsumbreytingar og frumusérhæfingar í þekju brjóstkirtilsins. Í doktorsverkefni mínu beindi ég sjónum að ncRNA sameindum og grunnhimnuþættinum peroxidasin (PXDN) og notaði D492 frumulínuna ásamt dótturlínunni D492M, sem hefur bandvefseiginleika, til að rannsaka hlutverk þessara þátta í fyrrnefndum ferlum.

Í fyrstu greininni rannsökuðum við hlutverk miR-203a í bandvefsumbreytingu í D492/D492M frumumódelinu og greinóttri formgerð í

D492. Niðurstöður sýndu að tjáning miR-203a eykst í samræmi við myndun greinóttrar formgerðar. Einnig kom í ljós að tjáning miR-203a er lítil sem engin í D492M miðað við D492 en yfirtjáning á miR-203a í D492M leiddi að hluta til viðsnúnings á bandvefsumbreytingu með aukinni tjáningu á luminal og basal markerum. Einnig kom í ljós að í D492M$^{miR\text{-}203a}$ dró mest úr tjáningu PXDN. Við staðfestum síðar að *PXDN* er markgen miR-203a sem binst við bindiset á 3'UTR próteinsins og veldur þar með niðurtjáningu.

Önnur greinin segir frá meginrannsókn doktorsverkefnisins sem beindist að PXDN próteininu, sem myndar krosstengi á milli kollagen IV sameinda, og hlutverki þess í frumusérhæfingu og myndun greinóttrar formgerðar í D492. Yfirtjáning á *PXDN* í D492 sýndi að próteinið ýtti undir ósérhæfða basal svipgerð og dró um leið úr kirtilþekju- og vöðvaþekjusérhæfingu. Þetta leiddi til hindrunar á myndun greinóttrar formgerðar þar sem þörf er á bæði vöðvaþekju og kirtilþekju. Jafnframt kom fram aukin tjáning á umritunarþættinum p63, sem þekkt er að viðheldur basal svipgerð í brjóstaþekju. RNA raðgreining sýndi að auki að gen sem taka þátt í ferlum tengdum sérhæfingu og þroskun þekjuvefja voru með aukna tjáningu en að sama skapi var virkni í bandvefsumbreytingu þekjuvefjar, sem er nauðsynlegur ferill í greinóttri formgerð, verulega skert. Niðurstöður rannsóknarinnar sýndu fram á áður óþekkt hlutverk PXDN í frumusérhæfingu í brjóstkirtlinum en genið hafði fram til þessa ekki verið skoðað í þessu tilliti.

Í þriðju greininni var hlutverk lncRNA sameindarinnar MEG3 í bandvefsumbreytingu og frumusérhæfingu í D492/D492M frumumódelinu kannað. Þar kom í ljós að MEG3 er bundið við bandvefslíka svipgerð en tjáning þess var marktækt hærri í bandvefslíkum frumulínum miðað við þekjufrumulínur. Við yfirtjáningu á *MEG3* í D492 varð svipgerðarbreyting í átt að basal svipgerð ásamt marktækri aukningu á bandvefslíkum eiginleikum og meðfylgjandi aukinni tjáningu á bandvefskenniprótеinum. Þar að auki dró úr greinóttri formgerð í D492^{MEG3}. Við niðurslátt á *MEG3* í D492M varð viðsnúningur á bandvefsumbreytingu og þekjueiginleikar jukust, þó aðeins í átt að kirtilþekjusvipgerð. Þessar niðurstöður voru staðfestar með RNA

raðgreiningu sem sýndi fram á að dregið hafði úr virkni bandvefsumbreytingar samfara niðurtjáninu á bandvefsgenum og aukningu á kirtilþekjugenum.

Til samantektar varpaði ég ljósi á hlutverk ncRNA sameindanna MEG3 og miR-203a í bandvefsumbreytingu, frumusérhæfingu og greinóttri formgerð í D492 og D492M en einnig bar ég kennsl á *PXDN* sem markgen miR-203a. Í framhaldinu sýndi ég fram á að PXDN, sem hingað til hefur ekki verið rannsakað í brjóstaþekju, eykur basal svipgerð í D492 en dregur úr frumusérhæfingu og hindrar myndun greinóttrar formgerðar í D492.

Lykilorð:
Brjóstkirtill, D492, PXDN, ncRNA, EMT, frumusérhæfing

Abstract

The mammary gland is a branched, lobular organ comprised of terminal duct lobular units (TDLUs) that contain bilayered epithelium of inner luminal cells and outer myoepithelial cells. The basement membrane, with collagen IV as the chief component, separates the epithelium from the surrounding stroma which is rich in fibroblasts, microvessels, and immune cells. The human female mammary gland is unique as its main developmental period occurs long after birth during puberty, and full maturity is not reached until pregnancy where branching is increased, and milk-producing units are formed. Furthermore, the gland undergoes cyclical changes during each menstrual cycle in preparation for a possible pregnancy, where the epithelium expands and is subsequently broken down if pregnancy does not occur. This process is maintained by epithelial stem cells and progenitor cells that differentiate into luminal and myoepithelial cells in addition to maintaining their population. D492 is a breast epithelial cell line with stem cell properties, isolated from the suprabasal subpopulation of normal breast epithelium. The suprabasal population is believed to contain mammary stem cells, reflected in the ability of D492 cells to undergo branching morphogenesis when cultured under three-dimensional conditions. These branching structures resemble the TDLUs of the breast and require plasticity of D492 cells which can form populations of cells with luminal and myoepithelial characteristics. In addition to epithelial plasticity, epithelial-to-mesenchymal transition (EMT) is a central process in branching morphogenesis in the breast, as epithelial cells require mesenchymal characteristics to be able to migrate into the surrounding matrix when branches are being formed. A myriad of factors is involved in the control of branching morphogenesis, EMT, and epithelial differentiation within the breast. In my Ph.D. project, I applied the D492 cell model along with its mesenchymal daughter cell line D492M and focused on the role of the ncRNAs MEG3 and miR-203a and the basement membrane factor peroxidasin (PXDN) in these processes.

The first article focused on the role of microRNA (miRNA) miR-203a in EMT in D492 and D492M in addition to branching morphogenesis in D492. Small RNA sequencing revealed that miR-203a expression is increased during branching with increased colony complexity and cell differentiation. Furthermore, we found that miR-203a was significantly less expressed in D492M compared to D492. Overexpression of miR-203a in D492M lead to partial mesenchymal-to-epithelial transition (MET) with upregulation of both luminal and basal markers. We also found that *PXDN* was the most downregulated gene in $D492M^{miR\text{-}203a}$. Subsequently, we confirmed that miR-203a targets the 3‘UTR of PXDN, resulting in post-transcriptional downregulation.

In the second article, I present data which were the main focus of my Ph.D. project, the collagen IV crosslinker PXDN and its role in epithelial differentiation and branching morphogenesis in D492. Overexpression of *PXDN* in D492 eliminated the luminal and myoepithelial phenotype while inducing undifferentiated basal phenotype, thereby negatively affecting plasticity that is characteristic for the D492 cell line. This led to complete inhibition of branching morphogenesis. I also present evidence that induction of basal phenotype in $D492^{PXDN}$ cells is mediated through upregulation of p63, which maintains basal phenotype in breast epithelium. Furthermore, RNA-sequencing reveals significant enrichment in pathways involved in epithelial differentiation and development in $D492^{PXDN}$ while EMT, a significant pathway involved in branching morphogenesis, is significantly downregulated.

In the third article, the role of the long non-coding RNA (lncRNA) MEG3 in EMT and epithelial differentiation in D492 and D492M was explored. Results showed that MEG3 expression was associated with mesenchymal phenotype and was significantly higher in mesenchymal cell lines. Overexpression of *MEG3* in D492 was accompanied by increased basal and mesenchymal features and reduced branching morphogenesis. With knock-down of *MEG3* in D492M partial MET was observed with an increase in epithelial markers but only towards the luminal phenotype. Results from D492M were confirmed

by RNA sequencing which showed that the EMT pathway was downregulated with simultaneous downregulation of mesenchymal genes and upregulation of luminal markers

To summarize, I investigated the role of ncRNAs MEG3 and miR-203a in EMT, epithelial differentiation, and branching morphogenesis in the D492 and D492M cell lines. Furthermore, I identified *PXDN*, a collagen IV crosslinker previously unexplored in the mammary gland, as a target of miR-203a. Subsequently, I showed that PXDN induces basal phenotype and stemness with a simultaneous reduction in differentiation and inhibition of branching morphogenesis in D492. In addition, I showed that this effect was elicited through the p63, presenting a novel link to PXDN.

Keywords:

Mammary gland, D492, PXDN, ncRNA, EMT, cell plasticity

Acknowledgements

The work described in this thesis was carried out at the Stem Cell Research Unit (SCRU), Department of Anatomy, School of Health Sciences, University of Iceland.

I want to start by thanking my supervisors, Þórarinn Guðjónsson and Gunnhildur Ásta Traustadóttir, for their guidance, patience, and encouragement. You supported me when needed, pushed me to do better, and gave me the freedom to develop my path and ideas. Being your student has been an honor and has taught me skills that extend far beyond the lab.

Furthermore, I would like to thank my doctoral committee for their input, guidance, and constructive criticism throughout my Ph.D. project: Bylgja Hilmarsdóttir, Eiríkur Briem, Jón Þór Bergþórsson, and Sigríður Valgeirsdóttir.

I thank Ólöf Guðrún Sigurðardóttir at Keldur for igniting the spark I have for pathology and for giving me time and encouragement to finish writing this thesis.

I also want to thank the past and present members of the SCRU team for their help and exceptional teamwork. I want to give special thanks to my friends; Ari, Jennifer, Hildur, Zuzana, Sophie, Sævar, Snævar, Árni, Arna, and Hildur for all the help, teamwork, cups of coffee, conversations, and laughs. You have provided me with the daily joys of working in a good team. Finally, I want to thank my best friend Bryndís, one of the best people I have ever met.

I also want to express my gratitude to my family for their patience and help. Especially my mother, Bryndís, and my in-laws, Þura and Skúli, for always being ready to step in and help.

To my dear Einar, Elín and Ernir, thank you for your endless love and support, and for always cheering me on.

Finally, I want to thank The Icelandic Center for Research (Rannís), Göngum saman, and the Icelandic Cancer Society research fund for providing funding for this project.

Contents

List of abbreviations

3D	Three-dimensional
3D-rBM	Three-dimensional reconstituted basement membrane
3'UTR	Three prime untranslated region
AGO	Argonaute protein
ALDH	Aldehyde dehydrogenase
BRENCs	Breast endothelial cells
BSA	Bovine serum albumin
cDNA	Complementary DNA
CK	Cytokeratin
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPRa	CRISPR activation
CRISPRi	CRISPR interference
CSC	Cancer stem cells
CTRL	Control
DMEM	Dulbecco's Modified Eagle Medium
DMFS	Distant metastasis free survival
ECM	Extracellular matrix
EMT	Epithelial-to-mesenchymal transition
ER	Estrogen receptor
EXP5	Exportin 5
FBS	Fetal bovine serum
GOBO	Gene Expression-Based Outcome database
GSEA	Gene set enrichment analysis
HEK-293T	Human embryonic kidney 293 cells with SV40-T antigen
HER2	Receptor tyrosine-protein kinase ErbB-2
HMLE	Human mammary epithelial cell line
HPA	Human Protein Atlas
HUVEC	Human umbilical vein endothelial cells
KD	Knock down
KO	Knock out
LEP	Luminal epithelial cells
lncRNA	Long non-coding RNA
MDA-MB-231	Metastatic breast cancer cell line derived at M.D.

	Anderson
MEG3	Maternally expressed gene 3
MEP	Myoepithelial cells
MET	Mesenchymal-to-epithelial transition
miRNA	MicroRNA
miR	MicroRNA
ncRNA	Non-coding RNA
Neu	Receptor tyrosine-protein kinase ErbB-2
NC1	Noncollagenous domain
NX	Normalized expression
OS	Overall survival
PAM50	Prediction analysis of microarray 50
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
P_{CT}	Probability of conserved targeting
PR	Progesterone receptor
pre-miRNA	Precursor miRNA
pri-miRNA	Primary miRNA
PXDN	Peroxidasin
rBM	Reconstituted basement membrane
qRT-PCR	Real-time Quantitative PCR
RISC	RNA-induced silencing complex
RNase III	Ribonuclease III
siRNA	Small interfering RNA
SNP	Single nucleotide polymorphism
TEB	Terminal end bud
TDLUs	Terminal duct lobular units

List of figures

List of tables

List of original papers

The following publications form the basis for this thesis:

I. **MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin.**
Eirikur Briem, Zuzana Budkova, Anna Karen Sigurdardottir, Bylgja Hilmarsdottir, Jennifer Kricker, Winston Timp, Magnus Karl Magnusson, Gunnhildur Asta Traustadottir and Thorarinn Gudjonsson. (Published in Mechanisms of Development in 2019).

II. **PXDN enhances basal phenotype and inhibits branching morphogenesis in breast epithelial progenitor cell line D492.**
Anna Karen Sigurdardottir, Arni Asbjarnarson, Arna Steinunn Jonasdottir, Hildur Run Helgudottir, Thorarinn Gudjonsson and Gunnhildur Asta Traustadottir. (Published in Journal of Mammary Gland Biology and Neoplasia in 2021).

III. **Expression of ncRNAs on the DLK1-DIO3 locus is associated with basal and mesenchymal phenotype in breast epithelial progenitor cells**
Zuzana Budkova, Anna Karen Sigurdardottir, Eirikur Briem, Jon Thor Bergthorsson, Snaevar Sigurdsson, Magnus Karl Magnusson, Gunnhildur Asta Traustadottir, Thorarinn Gudjonsson and Bylgja Hilmarsdottir. (Published in Frontiers in Cell and Developmental Biology in 2020).

IV. **Application of 3D culture assays to study stem cells in breast morphogenesis, extracellular matrix remodeling, and epithelial plasticity**
Anna Karen Sigurdardottir, Bylgja Hilmarsdottir, Thorarinn Gudjonsson and Gunnhildur Asta Traustadotti. (Published in Stem Cell Assays: Methods and Protocols in 2022).

Additional unpublished data will also be presented.

Declaration of contribution

Article I. MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin. In this paper, of which I am the third author, where the role of miR-203a in branching morphogenesis and epithelial-to-mesenchymal transition is investigated, I participated in the confirmation of *PXDN* as a target for miR-203a via luciferase reporter assay.

Article II. Peroxidasin enhances basal phenotype and inhibits branching morphogenesis in breast epithelial progenitor cell line D492. In this paper, of which I am the first author, I designed and performed experiments, analyzed experimental data and RNA sequencing data, performed the statistical analysis, and wrote the manuscript together with my supervisors and co-authors.

Article III. Expression of ncRNAs on the DLK1-DIO3 locus is associated with basal and mesenchymal phenotype in breast epithelial progenitor cells. In this paper, of which I am the second author, I analyzed RNA sequencing data from $D492M^{KD\text{-}MEG3}$ and produced heatmaps and pathway analysis. I also isolated fibroblasts and primary epithelial cells, which I sorted into LEPs and MEPs, from breast tissue and performed qRT-PCR for miRNAs and mRNA.

Article IV. Application of 3D culture assays to study stem cells in breast morphogenesis, extracellular matrix remodeling, and epithelial plasticity. In this protocol paper, of which I am the first author, I wrote the manuscript where I described the methods we use for 3D culture of D492 cell model.

1 Introduction

1.1 Anatomy of the human breast

The human breast is a modified skin gland, divided into collections of terminal duct lobular units (TDLUs) which represent the functional milk producing units of the breast (Figure 1). The TDLUs are acinar structures which interconnect with intralobular ducts that merge into larger interlobular ducts, forming passageways that eventually terminate at the nipple. The TDLUs can be divided into two layers of epithelial cells; an inner layer of luminal epithelial cells (LEPs) which are surrounded by an outer layer of myoepithelial cells (MEPs). The LEPs form the ducts of the gland along with the secretory alveoli where milk production takes place during the lactational period. Finally, in contact with the underlying basement membrane but not protruding into the lumen are the suprabasal cells, believed to be progenitor cells of the breast. The TDLUs are surrounded by collagen rich matrix containing fibroblasts, adipocytes, endothelium and immune cells, which together with the matrix form the breast stroma (Macias and Hinck 2012, Bonnans, Chou et al. 2014).

A central element in the development of the mammary gland is branching morphogenesis, a highly conserved process that can also be observed during embryogenesis in other branched organs such as the lung, kidney and salivary gland (Affolter, Zeller et al. 2009, Huebner and Ewald 2014). However, the breast is unique as extensive branching morphogenesis does not occur until puberty. During this period the TDLUs are formed when hormonal queues facilitate the elongation of the functional epithelium into the surrounding stroma, along with remodeling of the surrounding extracellular matrix (ECM) (Fata, Werb et al. 2004). After puberty the breast undergoes continuous remodeling with periods of epithelial proliferation and differentiation in preparation for pregnancy, followed by epithelial regression and apoptosis if impregnation has not occurred.

Finally, terminal epithelial differentiation and developmental potential is reached during pregnancy and the gland becomes lactation competent. The cycle is completed at weaning when the glandular tissue undergoes involution and the breast returns to its resting state (Hansen and Bissell 2000, Fata, Werb et al. 2004, Macias and Hinck 2012). The extensive morphological and functional development of the breast is believed to be induced by mammary epithelial stem cells that direct both the cyclical changes and the lactation preparation of the gland (Ruan and Kleinberg 1999, Hansen and Bissell 2000, Macias and Hinck 2012).

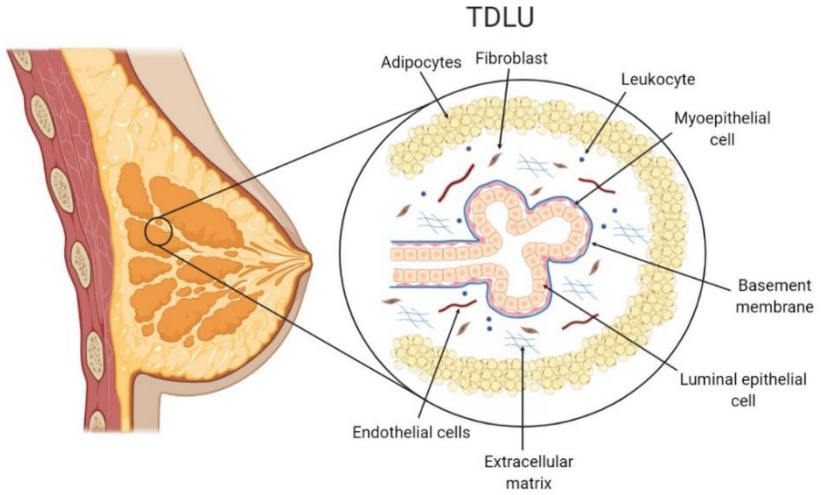


Figure 1. Anatomy of the human breast. The basic functional unit of the breast, the TDLU, can be divided into two compartments. The epithelial compartment contains the functional glandular epithelium with LEPs and MEPs. The stromal compartment supports the functional epithelium and contains adipocytes, leukocytes, endothelial cells and fibroblasts that produce and maintain the ECM. The two compartments are separated by the basement membrane rich in collagen IV.

1.2 Breast epithelial stem cells in the normal gland and breast cancer

Stem cells are defined by self-renewal, differentiation and homeostatic control and the idea of stem cells within the mammary gland, accounting for the gland's dynamic nature, is well established (Dalerba, Cho et al. 2007, Visvader and Stingl 2014). The first evidence of the existence of mammary gland stem cells came as a result of complete epithelial branching structures growing from sections of mouse mammary epithelium that had been transplanted into cleared recipient fat pad (Deome, Faulkin et al. 1959). Later, Kordon and Smith (1998) reported that retrovirally transduced mouse mammary cells produced glandular tissue with the same viral traces as the original cells, indicating extensive clonal outgrowths. Bipotent stem cells were first reported in human mammary glands in 2001 when Stingl et al. reported a population of cells that could be stimulated to produce cells with either luminal phenotype (expressing MUC1, K18, K19 and EpCam) or myoepithelial phenotype (expressing K14). In 2002, Gudjonsson et al. reported $MUC1^-/EpCAM^+$ cells that were able to differentiate into both $K14^+$ and $K19^+$ cells. By 2005, three different populations of human mammary gland progenitor cells had been identified; the bipotent population mentioned before along with lineage restricted luminal progenitor cells and myoepithelial progenitor cells, respectively. Both the bipotent and the lineage restricted luminal progenitor populations expressed EpCam and CD49f, however the luminal restricted population was characterized by positive MUC1 expression not found in the bipotent population (Stingl, Raouf et al. 2005). At the time, a pattern of multi-step hierarchy with increasing differentiation among mammary gland stem cells was beginning to emerge. Finally, in 2007 Villadsen et al. identified ductal epithelial cells that co-expressed K14 and K19 and were able to form TDLU-like structures in three-dimensional reconstituted basement membrane (3D-rBM). Populations with $K14^+/K19^-$, $K14^-/K19^-$ or $K14^-/K19^+$ expression were identified as lineage restricted and were unable to branch *in vitro*. Since then, lineage tracing in mice further

supports the existence of a stem cell hierarchy within the mammary gland (Inman, Robertson et al. 2015).

Breast cancer is the most common type of cancer in women worldwide. It has been estimated that in 2018, 2.1 million women were diagnosed with breast cancer and over 626.000 women died (Bray, Ferlay et al. 2018). Breast cancer is a heterogeneous disease where treatment options and prognosis are largely dependent upon subtype (Harbeck, Penault-Llorca et al. 2019). Based on gene expression signature of 50 genes (PAM50) six different subtypes of breast cancer have been suggested (Figure 2): Basal-like, Luminal A, Luminal B, HER2-enriched, Claudin-low and Normal-like (Perou, Sørlie et al. 2000, Cheang, Martin et al. 2015, Harbeck, Penault-Llorca et al. 2019). However, routine diagnosis and treatment selection are carried out by histological evaluation and immunohistochemistry of four key proteins: Human epidermal growth factor receptor 2 (HER2), estrogen receptor (ER), progesterone receptor (PR) and Ki67, a proliferation marker (Harbeck, Penault-Llorca et al. 2019). It is well known that breast cancer cells can hijack developmental programs and utilize to their benefit.

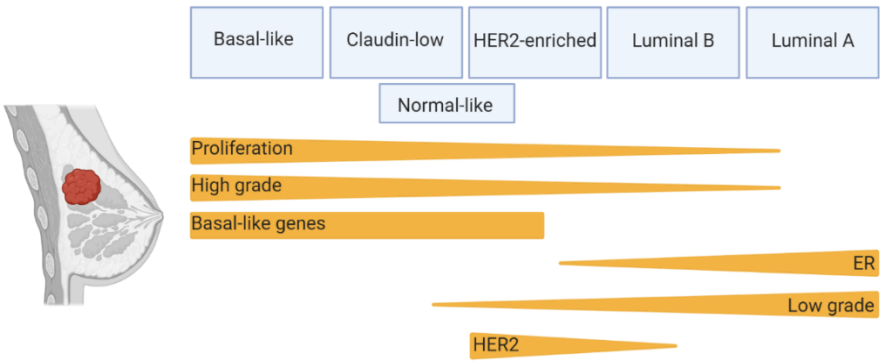


Figure 2. PAM50 classification of breast cancer. PAM50 classifies breast cancer tumors into six different subtypes based on genetic signature. These subtypes vary in major cell phenotype and marker expression status which can significantly affect aggressiveness and response to therapy. *Adapted from Harbeck et al. 2019*.

In recent years there has been increasing evidence that mammary stem cells play a role in breast cancer development and progression and, furthermore, are able to give rise to progeny that contribute to heterogeneity within breast tumors that can significantly impact prognosis and treatment outcome (Dalerba, Cho et al. 2007, Visvader 2009, Palomeras, Ruiz-Martínez et al. 2018). Stem cells harbouring this tumor associated phenotype are commonly referred to as cancer stem cells (CSC), originally identified by Al-Hajj et al. (2003). They reported a limited subpopulation of primary breast cancer cells with the ability to initiate tumor growth in mice. Furthermore, the tumors from the mice were similar in heterogeneity as the original tumors, supporting the presence of stemness in the original cells. Since then, mammosphere assays and cell sorting techniques for CSC markers, such as luminal marker CD24, basal marker CD44 and aldehyde dehydrogenase (ALDH) have been applied to identify cells with CSC properties *in vitro* (Charafe-Jauffret, Monville et al. 2008, Dittmer 2018).

1.3 Epithelial-to-mesenchymal transition

A certain degree of plasticity is necessary for branching morphogenesis to take place within the breast gland. To carry out this task the cells need to lose the cell-to-cell adhesion characteristic for epithelial tissue and acquire a mobile form to invade the surrounding stroma. To fulfil this requirement, epithelial cells undergo epithelial-to-mesenchymal transition (EMT), where cells lose epithelial characteristics while gaining mesenchymal traits, which is a key mechanism of embryonic development and wound healing (Moustakas and Heldin 2007, Yang and Weinberg 2008, Thiery, Acloque et al. 2009). The EMT program in mammary gland development has been extensively described in mice. During the embryonic stage the cells of the mouse mammary bud undergo partial EMT to facilitate the formation of the rudimentary mammary gland, and subsequently re-epithelialize in the reverse process of MET (Nanba, Nakanishi et al. 2001). As in humans, rapid morphogenesis and ductal growth are initiated at the onset of puberty in mice with full developmental potential being reached during pregnancy (Inman, Robertson et al. 2015). Ductal growth, herein elongation and bifurcation,

occurs mainly at the terminal end buds (TEBs), which correspond to the human TDLUs, and is driven by collective cell migration (Ewald, Brenot et al. 2008). Epithelial cells at the tip of the TEBs have mesenchymal characteristics acquired via EMT, which facilitates migration of cells into the mammary fat pad (Ewald, Huebner et al. 2012). Therefore, it is believed that branching morphogenesis and EMT are closely related and even regulated through similar mechanisms.

EMT is marked by phenotypic and functional changes within cells (Figure 3). Mesenchymal cells acquire increased mobility resulting in greater ability for migration and invasion. Furthermore, they display resistance to apoptosis (Hanahan and Weinberg 2011). Downregulation of the epithelial marker E-cadherin is widely considered the hallmark of EMT (Peinado, Olmeda et al. 2007). The suppression of E-cadherin and the induction of EMT is driven by several transcription factors, mainly ZEB1, ZEB2, TWIST1, TWIST 2, SNAI1 and SNAI2. In addition to suppression of epithelial markers these transcription factors induce the expression of mesenchymal markers, most notably N-cadherin, α-SMA, vimentin and fibronectin (Moustakas and Heldin 2007, Georgakopoulos-Soares, Chartoumpekis et al. 2020). Further switch in marker expression can often be noted during EMT where epithelial markers such as epithelial adhesion molecule (EpCAM) and epithelial keratins are downregulated while mesenchymal markers such as alpha-smooth muscle actin (α-SMA) and fibronectin (FN1) are upregulated (Luo, Brooks et al. 2015). Morphological changes also become apparent as cuboidal shape of the epithelial cell changes towards spindle shaped mesenchymal morphology where cellular protrusions assist the cell in eliciting its migratory behavior (Yang and Weinberg 2008).

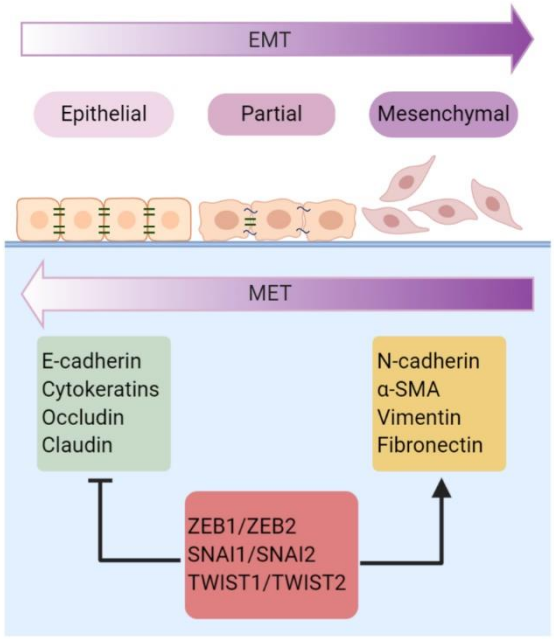


Figure 3. Epithelial-to-mesenchymal transition. In EMT epithelial cells lose the epithelial phenotype and acquire mesenchymal characteristics. EMT is carefully controlled via transcription factors (red box) that directly or indirectly inhibit expression of epithelial markers (green box) and facilitate the upregulation of mesenchymal factors (yellow box). During this process the epithelial cells lose adhesion to neighbouring cells and gain the ability to migrate. Adapted from Dongre and Weinberg 2019.

There is accumulating evidence that EMT can induce stem cell properties in both normal epithelial cells and cancer. Paralogous stem-cell programs have been identified in normal mammary stem cells and tumor initiating cells linked to the activity of SNAI1 and SNAI2 (Ye, Tam et al. 2015). SNAI2 has also been linked to stem cell regulation in the normal mammary gland, where stem cell activity is blocked if SNAI2 is inhibited. Furthermore, transient overexpression of *SNAI2* in LEPs can revert their phenotype to mammary stem cells (Guo, Keckesova et al. 2012). Induction of EMT can also encourage upregulation of stem cell markers and the acquisition of stem cell functional traits, such as increased mammosphere formation (Mani, Guo et al. 2008). The EMT program can also induce stem-like properties to cancer cells (Shibue and Weinberg 2017) which can greatly influence prognosis and

treatment outcome, as cancer stem cells in various tumor types are widely considered to be resistant to therapy (Singh and Settleman 2010, Kotiyal and Bhattacharya 2014, Zhou, Li et al. 2017, Kuşoğlu and Biray Avcı 2019). This is largely due to cancer stem cells contributing to tumor heterogeneity which results in poorly differentiated cell populations. This leads to the formation of different cancer cell subtypes that may differ in their response to therapy, contributing to therapy resistance and increasing the risk of relapse (Li and Bhatia 2011).

1.4 The basement membrane

The functional epithelium of each organ rests upon a thin, acellular matrix that separates the epithelial cells from the underlying connective tissue. This matrix, commonly referred to as the basement membrane, is specifically adapted to the needs of each organ with molecular composition that reflects the functional role of the epithelium resting above it (Halfter, Oertle et al. 2015). The ECM not only maintains tissue integrity and structure, but also plays critical role in cell homeostasis, plasticity as well as signaling (Nelson and Bissell 2006; Xu, Boudreau, and Bissell 2009; Bonnans, Chou, and Werb 2014). Changes in the microenvironment serve the epithelial cells with mechanical and chemical cues which in turn influence cell behavior, thereby linking tissue function and architecture. This effect is mediated through ECM-receptors on the cell surface, which induce conformational changes in the cytoskeleton. The cytoskeleton then mediates information to the nucleus, leading to changes in gene expression (Nelson and Bissell 2006; Fata, Werb, and Bissell 2004).

Anatomically the breast epithelial stem cells rest in direct contact with the basement membrane and the ECM is a key player in determining stem cell fate (Watt and Hogan 2000). A well-known example of this is the terminal differentiation that occurs in keratinocytes when they detach from the basement membrane and move up into the upper layers of the epidermis (Watt 1989), further supported by the fact that ECM protein ligation in suspended keratinocytes can inhibit terminal differentiation (Adams and Watt

1989). Interestingly, skin differentiation and stratification is believed to be mediated through the transcription factor p63, which is expressed in basal cells but becomes repressed as differentiation progresses (Koster and Roop 2004). Crosstalk between epithelial stem cells and the ECM has also been observed. Trappmann et al. (2012) proposed that stem cells exert mechanical force on the collagen fibers they attach to and utilize the feedback to determine degree of differentiation.

The main structural component of the basement membrane is collagen IV, along with other factors such as perlecans, proteoglycans, laminins and nidogens, interwoven between its fibres (Brown, Cummings et al. 2017). Collagen IV is known to preserve the epithelial phenotype, making separation from the surrounding stroma an important function of the basement membrane as collagen I, a main component of the stroma, can induce EMT when in contact with epithelial cells (Zeisberg, Bonner et al. 2001). Furthermore, collagen IV is necessary for the development and maintenance of basoapical polarity of epithelial cells, which is a prerequisite for proper function and homeostasis in normal tissue (Plachot, Chaboub et al. 2009). The same is true for breast epithelial cells. For example, Gudjonsson et al. (2005) showed that normal primary LEPs cultured in collagen I matrix *in vitro* showed reversed polarity that was only corrected in a coculture with normal MEPs. However, when they cultured LEPs in rBM correct polarity was observed. In contrast, carcinoma cells are unable to polarize (van Deurs, Zou et al. 1987), even when provided with a sound rBM *in vitro*, when at the same time normal cells are able to carry out correct polarity, lumen formation, regulation of growth and production of their own basement membrane components when grown in the same type of rBM environment (Gudjonsson, Rønnov-Jessen et al. 2002). Disruption of the collagen IV scaffolding and the basement membrane is a key event in cancer development as it marks the transition from carcinoma *in situ* to invasive disease.

1.5 Peroxidasin

PXDN is a peroxidase enzyme and a member of animal heme-peroxidase family that was originally discovered in *Drosophila melanogaster* (Nelson, Fessler et al. 1994, Cheng, Salerno et al. 2008). The structure of PXDN (Figure 4) is somewhat unique as it contains a catalytic peroxidase domain, leucine-rich repeats, von Willebrand factor- and immunoglobulin domains in addition to structural components that are reminiscent of other ECM proteins (Nelson, Fessler et al. 1994). Bhave et al. (2012) reported that PXDN facilitated the formation of covalent sulfilimine bonds between the C-terminal NC1-domains of collagen IV fibers (Figure 5), thereby confirming the contribution of PXDN to basement membrane synthesis and stabilization. Finally in 2015, Ero-Tolliver et al. found that the catalytic peroxidase domain and the immunoglobulin domains of PXDN were essential for sulfilimine crosslinking. Cysteine residues within the PXDN structure, which are highly conserved among animal species, were shown to facilitate trimerization of PXDN (Soudi, Zamocky et al. 2012). These PXDN-trimers were excreted to the extracellular space where they formed patches which are presumed to be hot-spots for collagen IV crosslinking at the cell surface. Conversely, the monomeric form of PXDN, although being secreted by cells, did not attach to the cell surface and was insufficient in crosslinking collagen IV. This indicates that PXDN trimerization is necessary for collagen IV crosslinking although cells can secrete both protein forms. Furthermore, the presence of PXDN at the cell surface might partake in basement membrane factor binding, which could take place via PXDN motifs that are found in other ECM adhesion proteins (Lázár, Péterfi et al. 2015).

In vivo data regarding the function of PXDN is limited. However, due to its key role in basement membrane homeostasis it is believed to be important in development. In terms of specific cell populations PXDN expression has been found in fibroblasts, endothelial cells and epithelial cells in various organs across diverse animal species (Péterfi, Donkó et al. 2009, Soudi, Zamocky et al. 2012). In pregnant mice, PXDN expression is significantly upregulated in uterine epithelial cells and stromal cells at the location of the

embryo during implantation, indicating an important role of PXDN in basement membrane remodeling during early pregnancy (Jones-Paris, Paria et al. 2017).

Figure 4. Amino acid sequence of PXDN. S: Signal peptide; LRR: Leucine rich repeat domain; Ig: Immunoglobulin domain; Pox: Peroxidase domaine; VWC: Von Willebrand factor C domain. Adapted from Soudi et al. 2015.

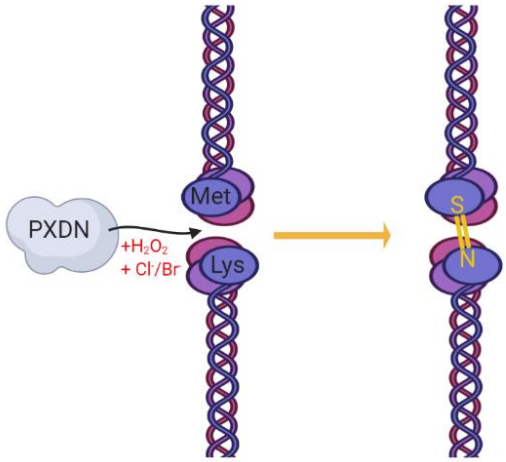


Figure 5. PXDN sulfilimine bond formation. PXDN uses hydrogen peroxide and halide ions to catalize the formation of covalent sulfilimine bonds between NC1 domains of collagen IV. Adapted from Péterfi and Geiszt 2014.

Furthermore, PXDN was recently linked to EMT, an important developmental process like previously discussed, as a target of EMT transcription factor SNAI1 (Sitole and Mavri-Damelin 2018). However, further investigation into the role of PXDN in EMT is lacking. The role of PXDN in disease has only been studied to a limited extent. Mutations in the *PXDN* gene result in severe ocular malformations, mainly in the anterior chamber of the eye. In humans, homozygous mutation presents as congenital corneal opacity and cataract with an increased risk of developing glaucoma, a disease that leads to abnormal intraocular pressure and damage to the optic nerve (Khan, Rudkin et al. 2011). In mice, homozygous *PXDN* mutations cause defected development in the anterior chamber which results in micropthalmia. Furthermore, structural integrity of the lens, likely due to basement membrane disfunction, is compromised which results in lens material leakage and subsequent ocular inflammation (Yan, Sabrautzki et al. 2014). Péterfi et al. (2009) implicated increased transforming growth factor (TGF-β1) mediated PXDN production to the development of fibrotic kidney after induced kidney damage in a mouse model. Altered PXDN expression has also been implicated in neoplastic disease. First, Jayachandran et al. (2016) observed consistently elevated *PXDN* expression in invasive mesenchymal melanoma cells as well as significantly increased expression in metastatic melanoma lesions. Zheng & Liang (2018) then reported significantly worse survival in ovarian cancer patients with high *PXDN* expressing tumors. Furthermore, they observed reduced proliferation as well as suppressed invasion/migration and wound healing potential in the ovarian cancer cell line HEY when *PXDN* was silenced. Elevated *PXDN* expression has also been detected in the microvasculature of primary and metastatic glioma (Liu, Carson-Walter et al. 2010).

The role of PXDN in mammary gland development and breast cancer has not been previously described, beyond the identification of PXDN and other ECM proteins in plasma samples extracted during tumor progression in the inducible HER2/neu mouse model of breast cancer (Pitteri, Kelly-Spratt et al.

2011). Therefore, the article presented in this thesis represents the first functional analysis of PXDN in the mammary gland.

1.6 Non-coding RNAs

The term non-coding RNA (ncRNA) refers to RNA that does not encode a protein, a part of the genome formerly believed to be genetic noise without function (Mattick and Makunin 2006). However, it has come increasingly to light that the ncRNAs serve diverse functions within the cell but most notably they are intricately involved in the regulation of gene expression (Mattick and Makunin 2006, Wei, Huang et al. 2017). Two types of ncRNAs fall within the scope of this thesis as participants in mechanisms involved in mammary gland biology: MiRNAs and lncRNAs.

1.6.1 MicroRNAs

MiRNAs are 22 nucleotide (nt) long, single stranded ncRNAs that play an important role in post-transcriptional gene regulation (Bartel 2004). They were first discovered nearly thirty years ago (Lee, Feinbaum et al. 1993), and have since then been widely studied so that their role as protein repressors has come increasingly to light. MiRNAs have been shown to actively regulate cell fate and phenotype and in recent years they have gained increased attention as participants in carcinogenesis, as context dependent tumor promoters and suppressors, and as biomarkers in various diseases (Lujambio and Lowe 2012, Casey, Sweeney et al. 2016, Ali Syeda, Langden et al. 2020). The principal mechanism of post-transcriptional regulation of miRNAs is complementary base pairing to target mRNA sequence, causing mRNA degradation or transport into p-bodies which leads to translational suppression (Rana 2007).

MiRNA maturation takes place in two steps (Figure 6), starting with primary miRNAs (pri-miRNA) that are cleaved by Drosha, a ribonuclease III enzyme (Rnase III), into hairpin precursor miRNA (pre-miRNA) that are around 60 nucleotides in length (Lee, Ahn et al. 2003). Exportin 5 (EXP5), a nuclear transport receptor, then transports the pre-miRNA from the nucleus

and into the cytoplasm (Yi, Qin et al. 2003), where Dicer, another Rnase III enzyme, receives the pre-miRNA and processes it into a miR-miR* duplex (Zhang, Kolb et al. 2002)

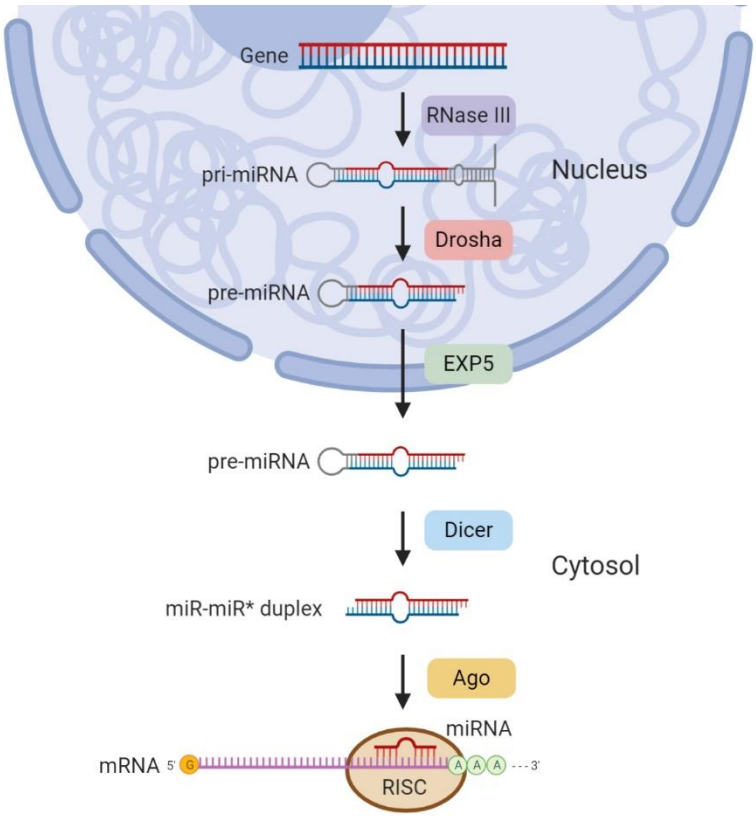


Figure 6. The maturation of miRNAs. The maturation of miRNAs involves a carefully orchestrated process first within the nucleus and then in the cytosol. MiRNA precursors are processed by the Rnase III enzymes Drosha and Dicer and finally with AGO protein, which results in a functioning miRNA that binds to a 3'UTR mRNA target site via RISC. Adapted from Soifer et al. 2008.

Further processing is facilitated by Argonaute (AGO) protein, where the miR* is cleaved off, resulting in a functional miRNA that forms an RNA-induced silencing complex (RISC) along with the associated AGO protein (Ameres and Zamore 2013). A seed-sequence located at the miRNA 5' end determines the binding of the RISC to the target mRNA. Although binding sites can be located anywhere on the mRNA the most potent binding, leading

to the most effective downregulation in expression, occurs at the 3'untranslated region (3'UTR) of the mRNA (Gu, Jin et al. 2009).

Several miRNAs have been linked to EMT regulation, most notably the miR-200 family. Five miRNAs belong to this family, divided into two separate clusters: MiR-200c and miR-141 on chromosome 12 and miR-200a, miR-200b and miR-429 on chromosome 1 (Zaravinos 2015). The only difference between the two groups of miRNAs is one base located in their seed sequence (Hilmarsdottir, Briem et al. 2014). The main role of the miR-200 family in EMT is suppression through targeting of transcription factors ZEB1 and ZEB2 which are suppressors of E-cadherin and inducers of EMT (Burk, Schubert et al. 2008, Gregory, Bert et al. 2008, Park, Gaur et al. 2008). Overexpression of miR-200 family members in cells leads to MET with loss of mesenchymal features and upregulation of epithelial markers along with induction of epithelial phenotype (Korpal, Lee et al. 2008). A reciprocal feedback loop exists between ZEB1/2 and miR-200 family members as ZEB 1 can suppress expression of miR-200c-141 and ZEB1 and ZEB2 induce downregulation of miR-200ba-429 (Bracken, Gregory et al. 2008, Burk, Schubert et al. 2008). The D492 cell model has successfully been applied to study the role of miR-200c-141 in EMT and branching morphogenesis when Bylgja Hilmarsdóttir, a former Ph.D. student at the Stem Cell Research Unit, discovered that miR-200c-141 cluster was significantly downregulated in D492M compared to D492. Upregulation of ZEB1/2 was also confirmed in D492M, in agreement with the known role of the miR-200c-141 as previously discussed. Furthermore, Hilmarsdottir found that overexpression of miR-200c-141 in D492 cells inhibited EMT in coculture with endothelial cells, while overexpression in D492M reversed the cells into MET and induced a luminal epithelial phenotype (Hilmarsdóttir, Briem et al. 2015).

1.6.2 Long non-coding RNA and *MEG3*

Another class of ncRNAs are the lncRNAs, that were discovered with advances in RNA sequencing technology (Mattick and Rinn 2015). The term lncRNA refers to nc RNA that is longer than 200 nt in length (Eades, Zhang

et al. 2014). The lncRNAs are key gene expression regulators and take part in various mechanisms active on the transcriptional, post-transcriptional and translational levels (Nie, Wu et al. 2012, Goff and Rinn 2015, Gloss and Dinger 2016). Unlike other RNA molecules the relationship between the tertiary structure of lncRNAs and their function remains unknown, partly due to their complexity and size (Uroda, Anastasakou et al. 2019). This contributes to the fact that lncRNAs are generally still poorly understood. However, there is increasing evidence for their role in various cell mechanisms such as differentiation, growth, survival, migration and invasion (Mercer, Dinger et al. 2009, Sun, Luo et al. 2013, Di Gesualdo, Capaccioli et al. 2014, Fatica and Bozzoni 2014). Data from Guttman et al. (2009) revealed that lncRNAs are tightly regulated during development and are under control of key transcription factors that take part in embryogenesis. However, recent evidence indicates that lncRNAs are not conserved to the same degree as mRNAs and variation between closely related species has also been noted, although this does not necessarily indicate lack of importance (Gloss and Dinger 2016). There is accumulating evidence that lncRNA play key roles in various diseases. Calin et al. (2007) reported altered expression of many lncRNAs in both leukemia and solid tumors. Later, certain lncRNAs were found to be differentially expressed between primary tumors and metastasis (Gupta, Shah et al. 2010) and dysregulated in response to DNA damage (Huarte, Guttman et al. 2010).

Maternally expressed gene 3 (*MEG3*) is a lncRNA that belongs to the imprinted DLK1-DIO3 locus on human chromosome 14. Incorporated within the locus are three paternally expressed protein coding genes and many maternally expressed non-coding genes, including over 50 miRNAs (Benetatos, Vartholomatos et al. 2014, Dill and Naya 2018, Baulina, Osmak et al. 2019). *MEG3* is controlled by differentally methylated promoter regions within the DLK1-DIO3 locus (McMurray and Schmidt 2012), and is necessary in embryonic neural development where its non-imprinted form directs Polycomb protein chromatin targeting (Mercer, Dinger et al. 2008, Kaneko, Bonasio et al. 2014). However, in adult tissue the imprinted form interacts

with p53 and thereby participates in mechanisms such as apoptosis (Zhou, Zhong et al. 2007). Different splice variants of MEG3 stimulate the p53 pathway to a varying degree which results in selective target gene upregulation (Zhu, Liu et al. 2015). Reports on the involvement of MEG3 in EMT have been conflicting. MEG3 has been linked to EMT induction in glioma cells and in TGF-β mediated EMT in lung cancer cells (Terashima, Tange et al. 2017, Yang, Bian et al. 2020). However, MEG3 repressed EMT in cholangiocarcinoma and cervical cancer (Chen and Qu 2018, Li, Jiang et al. 2019). Likewise, the role of MEG3 in tumorigenesis is also unclear, as various studies have linked it to either tumor suppression or induction (Chunharojrith, Nakayama et al. 2015, He, Luo et al. 2017, Wei and Wang 2017, Ghafouri-Fard and Taheri 2019, Wang, Fu et al. 2019, Zhang and Gao 2019). Therefore, the role of MEG3 is still poorly understood and remains to be further investigated.

1.7 The D492 cell model

The D492 cell line was originally isolated from the suprabasal cell population of the normal breast. The suprabasal cells were isolated from primary cells by sorting according to their characteristic $EpCam^{+}/MUC1^{-}$ marker status, separating them from the $EpCam^{+}/MUC1^{+}$ luminal cells and $EpCam^{-}/MUC1^{-}$ myoepithelial cells. Subsequently, immortalization was performed using the E6/E7 oncogenes from human papilloma virus 16 (Gudjonsson, Villadsen et al. 2002). The assumption of D492 stem cell properties is based on the cell line's ability to form both luminal and myoepithelial cells in monolayer culture as well as branching colonies resembling the TDLUs of the breast when cultured in 3D-rBM (Figure 7.a) (Gudjonsson, Villadsen et al. 2002, Villadsen, Fridriksdottir et al. 2007). Furthermore, a subpopulation of D492 cells is positive for both cytokeratin 14 (CK14) and cytokeratin 19 (CK19), a feature also observed in suprabasal cells of the breast *in situ* and an indicator of stem cell phenotype (Villadsen 2005, Villadsen, Fridriksdottir et al. 2007).

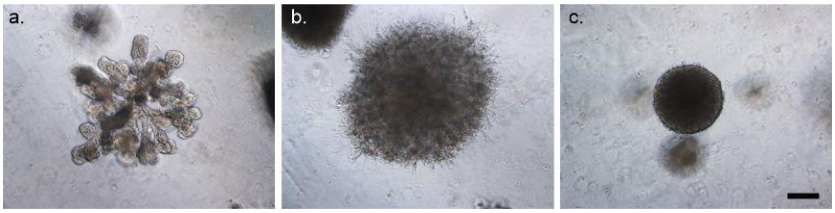


Figure 7. Colony conformation of D492 cells in 3D coculture with endothelial cells. A. D492 cells can form branching colonies reminicent of breast TDLUs when cultured in 3D-rBM. This ability is retained in coculture with endothelial cells. B. In coculture with endothelial cells D492 cells undergo EMT, forming spindle shaped colonies. D492M was isolated from a spindle shaped colony. C. D492 cells also form solid round spheres in both monoculture and coculture with endothelial cells. (Scale bar = 100 µm).

D492 has been used to investigate epithelial-endothelial interactions. Coculture of D492 cells with endothelial cells in 3D-rBM enhances branching morphogenesis and induces EMT, indicated by the creation of spindle shaped colonies (Figure 7.b) alongside the branched and spherical colonies (Figure 7.c) usually observed in D492 monoculture. When isolated, cells from these spindle shaped colonies were found to have increased expression of mesenchymal markers N-cadherin, vimentin and alpha smooth-muscle actin, while epithelial markers such as E-cadherin, keratins and transcription factor p63, were downregulated. Furthermore, mesenchymal traits such as increased anchorage independent growth, apoptosis resistance, increased migration/invasion and increased CD44/CD24 ratio were observed in the cells derived from spindle shaped colonies. Together, these results indicated a switch from an epithelial phenotype to a mesenchymal phenotype and thus this new cell line was named D492M (Sigurdsson, Hilmarsdottir et al. 2011). Together D492 and D492M have successfully been used as an *in vitro* model for EMT in the mammary gland, especially when it comes to the involvement of ncRNAs (Hilmarsdóttir, Briem et al. 2015, Briem, Budkova et al. 2019, Briem, Ingthorsson et al. 2019, Budkova, Sigurdardottir et al. 2020).

All three articles along with the unpublished data presented in this thesis are based upon work with the D492 cell line along with D492M, with supporting data from other breast epithelial cell lines and primary cells. The D492 cell model therefore forms the basis of my Ph.D. project as a

representation of arguably the most dynamic and versatile organ of the body: The mammary gland, an organ that is so important to our existence that it gave an entire class of animals its name, mammals.

2 Aims

Mammary gland development is a process of epithelial and extracellular remodeling from the embryonic stage and throughout the reproductive period, which is under careful regulation of intrinsic and extrinsic signaling pathways. Epithelial plasticity is a vital feature of mammary epithelial cells during branching morphogenesis, where the cells form the network of functional units and interconnecting ducts that comprise the epithelial ductal tree, and refers to the ability of epithelial cells to change phenotype in reaction to various signals. EMT and its reverse process of mesenchymal-to-epithelial transition (MET) are examples of epithelial plasticity that are important in branching morphogenesis. It has come increasingly to light that ncRNAs are important regulatory factors in EMT. MiRNAs can target EMT transcription factors and thereby influence branching morphogenesis. Other types of ncRNAs, the long-non coding RNAs (lnc-RNAs), have also been implicated in EMT regulation, although their role is still not completely understood. The general aim of this thesis was to apply the epithelial progenitor cell line D492 and its mesenchymal daughter cell line D492M to explore the role of ncRNAs miR-203a and MEG3 in EMT and branching morphogenesis. Furthermore, to explore the expression pattern and functional role of the miR-203a target *PXDN* in the mammary gland and epithelial plasticity. For these purposes I used 3D culture methods that capture the ability of the D492 cell line to form branching colonies, thereby mimicking the branching structure of the mammary gland and allowing me to explore the involvement of the aforementioned factors in mammary gland development.

2.1 Specific aims

1. To analyze the involvement of ncRNAs miR-203a and MEG3 in breast morphogenesis, epithelial cell plasticity and EMT.
2. To investigate the expression pattern of PXDN in the human breast and its effect on cell plasticity and branching morphogenesis in D492 cells.

3 Materials and methods

Detailed description of methods can be found in the materials and methods section in each article. However, in this chapter, I specifically describe methods that I used during my studies.

3.1 Cell culture

For a detailed description of methods for culturing of D492 cells please refer to article IV where protocols for monolayer culture, 3D culture and 3D coculture with endothelial cells are covered in detail.

3.2 Lentiviral packaging and transduction

Pseudoviral particles for lentiviral overexpression were produced by transfecting HEK293-T cells with PMDG.2 and psPAX2 plasmids using Turbofect (Thermo Fisher Scientific, R05319). Culture media was without serum and antibiotics. Viral supernatant was collected after 24, 48 and 72 hours and filtered through 0.45 µm pore filter. $1.5x10^6$ D492 cells were plated per T25 flasks, so that they had reached 60-80% confluency 24 hours later, i.e. at the time of transduction. They were then treated with 1 ml of lentiviral particles along with 1 ml of media and 8 µg/ml polybrene. Successfully transduced cells were selected by treating the cells with antibiotics or FACS

Table 1. Summary of plasmids used in lentiviral overexpression of *PXDN*.

Plasmid	Manufacturer, catalog no.
pLenti-C-Myc-DDK-P2A-Puro tagged cloning vector	Origene, PS100092
pLenti-C-Myc-DDK-P2A-Puro lenti ORF clone of Human peroxidasin homolog	Origene, RC224518L3
pMD2.G	Addgene, #12259
psPAX2	Addgene, #12260

based on positive fluorescent signalling. Altered expression of target genes was confirmed via real-time quantitative PCR (qRT-PCR).

3.3 Proliferation assay

Proliferation assay was performed by seeding $5.0x10^3$ cells into 96-well plate pre-coated with collagen I (PureCol Type I Collagen Solution, Advanced Biomatrix, 5005). Cells were incubated and proliferation rate monitored in IncuCyte ZOOM live cell imaging system (Sartorius) until full confluency was reached. Media was changed three times a week.

3.4 Apoptosis assay

Resistance to chemically induced apoptosis was determined by treatment with 10 µg/ml Camptotechin (Sigma-Aldrich, C9911) and 0.1% DMSO as control. Caspase-3/7 reagent (Essen Bioscience, 4440) was used to monitor apoptosis. Cells were incubated and monitored in IncuCyte live cell imaging system as per manufacturer instructions.

3.5 Migration assay

For migration assay in IncuCyte $1.0x10^4$ cells were seeded in quadruplicates onto 96-well ImageLock plates pre-coated with collagen I solution. Cells were grown until confluent after which a wound was formed with Woundmaker. Wound closure was monitored in IncuCyte live cell imaging system until wounds were completely closed and data exported as wound confluency (%).

3.6 Invasion assay

For invasion assay in IncuCyte $1.0x10^4$ cells were seeded in quadruplicates onto 96-well ImageLock plates pre-coated with collagen I solution. Cells were grown until confluency after which a wound was formed with Woundmaker. Matrigel was applied on top of cells immediately after wound formation. Wound closure was monitored in IncuCyte live cell imaging system until wounds were completely closed and data exported as wound density (%).

3.7 Mammosphere assay

Mammosphere formation ability was assessed in 96 well plate format. Plate wells were pre-coated with 20 mg/ml Poly Hema and subsequently dried overnight at 37°C. The next day $5.0x10^3$ cells were seeded in 5% Matrigel per well. Mammosphere formation was monitored for 10 days. Cells were incubated at 37°C and media changed three times per week.

3.8 Flow cytometry

Before performing flow cytometry, $D492^{PXDN}$ and $D492^{empty}$ cells were grown in triplicates in T25 cells until 95% confluent. After trypsinization cell pellets were resuspended in 30 uL of CellWASH (BD Biosciences, 349524). Cell suspensions were incubated with 3 uL of PE Mouse Anti-Human CD24 clone ML5 (BD Biosciences, 560991) and FITC Mouse Anti-Human CD44 clone IM7 (BD Biosciences, 103006), respectively, for 15 minutes at room temperature. Mouse IgG_1 PE (BD Biosciences, 345816) and Mouse IgG_1 FITC (BD Biosciences, 550758) were applied as isotype controls. Cells were then washed with CellWASH and centrifuged at 1500 rpm for 6 minutes. Cell pellets were resuspended in 300 µL of CellWASH and subjected to flow cytometry in Navios Flow Cytometer (Beckman Coulter).

3.9 Aldefluor assay

Before performing Aldefluor assay $D492^{PXDN}$ and $D492^{empty}$ cells were grown in triplicates in T25 cells until 95% confluent. Aldefluor assay was performed by using ALDEFLUOR Kit (Stemcell Technologies, 01700) as per manufacturer instructions. Aldefluor activity was measured in Navios Flow Cytometer (Beckman Coulter).

3.10 Quantitative reverse transcription PCR analysis

Total RNA extraction was performed with Tri-Reagent (Thermo Fisher Scientific, AM9738). Reverse transcription into complementary DNA (cDNA) was performed with LunaScript RT SuperMix Kit (New England BioLabs, E3010L) and qRT-PCR was performed with Luna Universal qPCR Master Mix (New England BioLabs, M3003L). For qRT-PCR analysis of miRNAs, cDNA was produced with

Table 2. Summary of primers used in qRT-PCR analysis of mRNA and miRNA.

Primer	Manufacturer, catalog no.
PXDN	IDT, Hs.PT.58.630748
CK14	IDT, Hs.PT.58.4592110
CK19	IDT, Hs.PT.58.4188708
CK5/6	IDT, Hs.Pt.58.14446018
TP63	IDT, Hs.PT.58.2966111
***MEG3* exon 10-11**	IDT, Hs.PT.58.25190740
CDH3	IDT, Hs.PT.58.39234242
GAPDH	IDT, Hs.Pt.39a.22214836
ZEB1	Thermo Fisher Scientific, Hs00232783_m1
ZEB2	Thermo Fisher Scientific, Hs00207691_m1
SNAI1	Thermo Fisher Scientific, Hs00195591_m1
SNAI2	Thermo Fisher Scientific, Hs00950344_m1
TWIST1	Thermo Fisher Scientific, Hs01675818_s1
GAPDH	Thermo Fisher Scientific, 4326317E
hsa-miR-203a	Exiqon, #205914
hsa-miR-127-3p	Qiagen, YP00204048
hsa-miR-141-3p	Exiqon, #204502
hsa-miR-200c-3p	Exiqon, #204482
hsa-miR-205-5p	Exiqon, #204487
hsa-miR-409-3p	Qiagen, YP00204358
hsa-miR-411-5p	Qiagen, YP00204531
hsa-miR-493-3p	Qiagen, YP00204557
U6 snRNA	Qiagen, YP00203907

universal cDNA synthesis kit II (Exiqon) and qRT-PCR was performed with ExiLENT SYBR Green master mix. A list of primers used for miRNA qRT-PCR can be found in the materials and methods chapters of each paper.

Analysis was performed on Applied Biosystems 7500 Real-Time PCR system and the $2^{-\Delta\Delta Ct}$-method was used to calculate relative expression differences.

3.11 Western blot assay

Cells were washed with ice cold phosphate buffered saline (PBS) before lysis with radioimmunoprecipitation assay (RIPA) buffer containing protease and phosphatase inhibitors (Halt Protease Inhibitor Cocktail, Thermo Fisher Scientific, 78430). Bradford reagent (BioRad, 5000002) was used to measure protein concentration. Protein was separated in equal amounts (5-20 µg) on 10% NuPage Bis-Tris gel (Invitrogen, NP0301PK2) using NuPage MES running buffer (Thermo, Fisher Scientific, NP0002). Protein was transferred with NuPage Transfer buffer (Thermo Fisher Scientific, NP0006-1) to polyvinylidene fluoride (PVDF) membrane (Millipore, IPFL00010). Membrane blocking was conducted with Odyssey TBS blocking buffer (LiCor, 927-500). Incubation with primary antibodies was carried out overnight at 4°C with subsequent incubation with secondary antibodies for 1 hour at room temperature. Protein bands were detected with Odyssey Infrared Imaging System (Li-Cor). Bands were quantified in ImageJ.

3.12 Immunochemistry

Cells were fixed in 4% paraformaldehyde for 1 hour at room temperature. Cells were then rinsed with 1xPBS and incubated with primary antibodies at 4°C overnight. Fluorescent labeling was performed with fluorescent isotype-specific secondary antibodies (Alexa fluor, Thermo Fisher Scientific), DAPI nuclear counterstain (Sigma-Aldrich, #D9542) and Phalloidin (Thermo-Scientific, #A12379). Imaging was performed on Zeiss LSM 5 Pa laser-scanning microscope (Carl Zeiss) and Olympus Fluoview 1200.

3.13 Immunohistochemistry on paraffin-embedded tissue

Tissue slides were boiled twice for 5 minutes each time in citrate buffer. Antigen blocking was performed with 5% bovine serum albuimin (BSA) in 0.1 M Tris buffer. Slides were incubated with primary antibodies for 60 minutes at room temperature. Fluorescent labelling was performed with fluorescent isotype specific secondary antibodies (Alexa fluor, Thermo Fisher Scientific. Imaging was performed on Zeiss LSM 5 Pa laser-scanning microscope (Carl Zeiss) and Olympus fluoview 1200.

Table 3. Summary of antibodies used in Western blot and immunochemistry of cells and paraffin-embedded breast tissue.

Antibody	Manufacturer, Catalog no.
PXDN	A gift from Dr. Miklós Geiszt
K14	Abcam, ab7800
K19	Santacruz, sc-6278
K19	Abcam, ab7754
K5/6	Invitrogen, #180267
p63	Abcam, ab124762
P63	Novocastra, NCL-p63
EpCam	Abcam, ab71916
Actin	Li-Cor, #924-42212
Histone H3	Cell Signaling, #4499

3.14 Transient knock-down with siRNA

Transient knock-down of *TP63* and *PXDN* in D492^{PXDN} and D492 cells, respectively, was performed with target-specific small interfering RNA (siRNA), using Lipofectamine RNAiMAX transfection reagent (Thermo Fisher Scientific, #13778075) as per manufacturer's protocol. Cells were subjected to assays 48 hours after transfection.

Table 4. Summary of siRNA used in transient knock-downs in D492 and D492^{PXDN} cells.

siRNA	Manufacturer, Catalog no.
TP63	Thermo Fisher Scientific, #s16411 lot. AS02HN1T
PXDN	Thermo Fisher Scientific, #4392421 lot. AS02B0C4

3.15 RNA sequencing

Total RNA was isolated from cells in monolayer at 80% confluency and on day 8 of 3D on top of matrix culture. RNA sequencing was performed by DeCode Genetics. RNA transcript expression was quantified in Kallisto version 0.46.1 (Bray, Ferlay et al. 2015) using the Homo Sapiens GRC38 release 96 reference transcriptome (Yates, Achuthan et al. 2019). Differential gene expression was computed with the Sleuth R package v0.30 (Pimentel, Bray et al. 2017). Gene enrichment analyses were performed using the Gene Set Enrichment Analysis (GSEA) database (gsea-msigdb.org) with pre-ranked gene lists and the Hallmark and C5 Ontology gene sets for HALLMARK and GO analyses, respectively (Mootha, Lindgren et al. 2003, Subramanian, Tamayo et al. 2005). Pre-ranking was performed by multiplying p-value with sign of log2-fold change. GSEA results were presented as bar plots generated in Excel or GraphPad Prism. Volcano plots were generated based on p-value and Log2-fold change of selected markers. Heatmaps were generated based on most significantly differentially expressed genes between cell lines. Volcano plots were generated in GraphPad Prism or using the EnhancedVolcano R package v1.8.0 (Blighe

2020). Heatmaps were generated in GraphPad Prism or using the ComplexHeatmap R package v2.7.7 (Gu, Eils et al. 2016).

3.16 Luciferase assay

Synthetic oligonucleotides containing position 106-113 of human *PXDN* 3'UTR with target sequence for hsa-miR203a-3p and control with sequence deletion were cloned into pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega). The correct sequence was verified via DNA sequencing (Eurofins Genomics). As a positive control, the target sequence for miR-203a-3p on p63 3'UTR was cloned into pmirGLO in addition to target site deletion and mismatch sequence for negative control. Luciferase assay was carried out in HEK293T cells in a 96-well plate format. $3.0x10^4$ cells were plated in each well and incubated overnight. Cells were then transfected with hsa-miR203a-3p mimic at a final concentration of 100 nM, with Lipofectamin RNAiMAX in high glucose Dulbecco's Modified Eagle Medium (DMEM) without antibiotics and serum. 24 hours later cells were transfected with 200 ng/well pmirGLO luciferase plasmid, using Lipofectamine 3000. Luciferase activity was assessed 24 hours later using Dual-Glo® Luciferase Assay System. Results were normalized against background firefly luciferase activity.

3.17 Transient transfection with miRNA mimics and inhibitors

Transient transfection of D492M with miRNA mimics and inhibitors was performed by treating cells with 50 pmol of mirVANA mimics of miR-23a-3p (Thermo Fisher Scientific, 4464084, Assay ID MC10644) and miR-29a-3p (Thermo Fisher Scientific, 4464084, Assay ID MC12499), mirVANA inhibitors for miR-23a-3p (Thermo Fisher Scientific, 4464084, Assay ID MH10644) and miR-29a-3p (Thermo Fisher Scientific, 4464084, Assay ID MH12499) along with mirVANA mimic negative control #1 (Thermo Fisher Scientific, 4464058) and mirVANA inhibitor negative control #1 (4464076). RNAiMAX (Thermo Fisher Scientific, 13778075) was used as transfection reagent as per manufacturer instructions.

3.18 Isolation of primary cells from reductionmammoplasties

Primary cells were isolated from breast tissue obtained via reduction mammoplasties (approved by the Icelandic Bioethics Committee VSN-13-057). Luminal epithelial cells and myoepithelial cells were subjected to magnetic labelling with anti-EpCam-conjugated microbeads (Miltenyi Biotech, #130-105-958) and subsequently separated via MiniMACS according to manufacturer's instructions (Miltenyi Biotech). For detailed description of isolation of BRENCs from breast tissue please refer to Article IV. BRENCs were magnetically isolated via the use of anti-CD-31 conjugated microbeads (Miltenyi Biotech, #130-091-935).

3.19 Statistical analysis

Statistical analysis was performed using GraphPad Prism, version 8. Data is presented, unless otherwise stated, as an average of three independent experiments with error bars representing standard deviation. Statistical significance was considered for p values under 0.05. Statistical analysis of each experiment is described in detail in each paper.

3.20 Image graphics

Graphic images were drawn in BioRender, available at www.biorender.com.

4 Results and discussion

In this chapter, I will summarize published and unpublished data acquired during my Ph.D. and discuss in context with published literature in the field. Published articles are included at the end of the thesis.

4.1 Article I: MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin (published in Mechanisms of Development in 2019)

It is well established that miRNAs play an important role in the regulation of branching morphogenesis and EMT through post-transcriptional modulation. In this paper, which was a part of Eiríkur Briem‘s Ph.D. project, we applied the D492 and D492M cell lines, that offer a realistic model of EMT and branching morphogenesis in breast epithelial progenitor cells as was discussed in the introduction to this thesis, to investigate differentially expressed miRNAs by performing small RNA-sequencing in monolayer and 3D culture. A former Ph.D. student at the Stem Cell Research Unit, Bylgja Hilmarsdottir, had previously published data where miRNA expression of D492 and D492M in monolayer was compared (2015). Hilmarsdottir reported significant downregulation of miRNAs related to EMT, in particular the miR-200c/141 cluster, in D492M compared to D492. However, as 3D culture offers microenvironmental input and can better mimic *in vivo* conditions compared to monolayer culture we were curious to see whether similar results would be observed when D492 and D492M were cultured in 3D-rBM.

4.1.1 MiR-203a is significantly downregulated in EMT through other mechanisms than promoter methylation

Small RNA sequencing of D492 and D492M in 3D culture, revealed that miR-203a was the fourth most significantly downregulated miRNA in D492M compared to D492, the other three being miR-200c, miR-141 and miR-205 (Figures 1.A and 1.B in article I). MiR-200c and miR-141, as members of the

miR-200 family, had previously been linked to branching morphogenesis and EMT in the D492 cell line (Hilmarsdóttir, Briem et al. 2015), and are well known inducers of EMT (Gregory, Bert et al. 2008, Park, Gaur et al. 2008, Hilmarsdottir, Briem et al. 2014), like previously discussed. Likewise, both miR-205 and miR-203a had been linked to EMT (Gregory, Bert et al. 2008, Zhang, Zhang et al. 2011, Ding, Park et al. 2013, Chao, Chang et al. 2014).

A few mechanisms exist that regulate miRNA activity within the cell, one of them being epigenetic modulation through methylation of the promoter CpG islands which results in gene silencing (Correia de Sousa, Gjorgjieva et al. 2019). Disruption in promoter methylation of miRNAs with resulting deregulation of miRNA expression is also a known factor in cancer (Daniunaite, Dubikaityte et al. 2017, Varghese, Shukla et al. 2018, Zare, Bastami et al. 2018). With this in mind, we were interested in seeing whether the promoter methylation played a role in the differential expression of the four most downregulated miRNAs in D492M compared to D492: miR-141, miR-200c, miR-205 and miR-203a. Neves et al. (2010) had previously reported that the miR-200c/141 cluster was regulated through promoter methylation during EMT in the normal breast epithelial cell line HMLE and breast cancer cells. Using HumanMethylation450 BeadChip and bisulfite sequencing we found that while miR-200c, miR-141 and miR-205 had increased methylation in D492M, miR-203a was not differentially methylated between the two cell lines (Figures 2.A and 2.B in article I). We therefore concluded that miR-203a downregulation during EMT was due to other regulatory mechanisms than promoter methylation.

4.1.2 MiR-203a in branching morphogenesis and epithelial differentiation

Since both EMT and miRNAs are known participants in branching morphogenesis we were curious to investigate the miRNA profile during branching. To do this we isolated RNA during early branching (day 7), branching (day 14) and late branching (day 21) in D492 cultured in 3D-rBM and performed small RNA-sequencing to analyze differential miRNA expression. This revealed that miR-203a was significantly upregulated

throughout branching morphogenesis, indicating a possible role in epithelial differentiation (Figure 3.A and 3.B in article I). We therefore decided to explore the expression of miR-203a in primary cells as well as other breast epithelial cell lines displaying either epithelial or mesenchymal phenotype. In primary cells, miR-203a had the highest expression in LEP, while expression in MEP was significantly lower. Furthermore, miR-203a expression was significantly higher in $EpCAM^+$ than in $EpCAM^-$ D492 cells (Figure 3.C in article I). Our results were in agreement with previously published data by DeCastro et al. (2013) and Pal et al. (2015) that also reported high miR-203a expression in luminal cells. The luminal restriction of miR-203a might be due to its targeting of the basal factor p63 (Yi, Poy et al. 2008, DeCastro, Dunphy et al. 2013), which induces basal phenotype when its expression is induced in LEPs (Wuidart, Sifrim et al. 2018). MiR-203a might therefore take part in the preservation of the luminal phenotype within the mammary gland.

We also found that while miR-203a was expressed in both LEPs and MEPs, very little expression was detected in fibroblasts and close to none in breast endothelial cells (BRENCs) (Figure 3.C in article I). To investigate whether miR-203a expression was restricted to the epithelial phenotype we performed qPCR in the normal breast epithelial cell line HMLE and its mesenchymal counterpart HMLEmes, as well as in the normal breast epithelial cells MCF10 and the triple negative breast cancer cell line MDA-MB-231. This revealed a significant correlation between miR-203a expression and the epithelial phenotype.

4.1.3 Overexpression of miR-203a leads to partial reversion to epithelial phenotype in D492M

Having established that miR-203a was affiliated with the epithelial phenotype, in D492 and other cell lines, we decided to perform stable overexpression through lentiviral transfection of miR-203a in D492M (Figure 4.A in article I). We found that overexpression of miR-203a affected phenotype both in monoculture and in 3D-rBM in D492M: In monolayer $D492M^{miR\text{-}203a}$ cells were more adherent to each other compared to $D492M^{empty}$. Furthermore, we found that in 3D-rBM $D492^{miR\text{-}203a}$ cells formed more compact 3D colonies,

although they still retained the spindle shaped morphology and mesenchymal protrusions compared to D492M^{empty} (Figure 4.B in article I). With functional experiments, we found that miR-203a overexpression reduced proliferation, migration, invasion and anchorage independent growth in D492M, with a concomitant increase in sensitivity to chemically induced apoptosis, which are traits that are usually affiliated to the epithelial phenotype (Figures 4.C and 4.G-I in article I). Immunostaining revealed that the epithelial marker E-cadherin was slightly increased, while the mesenchymal marker N-cadherin was repressed in D492M$^{miR-203a}$ (Figure 4.D in article I). A follow up by Western blot confirmed this (Figure 4.E in article I). Furthermore, qRT-PCR showed that the EMT marker SNAI2 was significantly downregulated in D492M$^{miR-203a}$ compared to D492M^{empty} (Figure 4.F in article I).

4.1.4 Identification of *PXDN* as a target of miR-203a

As miR-203a affected the phenotype of D492M we were curious to gain a broader view of the effect overexpression of miR-203a had on gene expression. Therefore we decided to perform RNA-sequencing on D492M$^{miR-203a}$ and D492M. Data analysis identified *PXDN* as the most significantly downregulated gene in D492M$^{miR-203a}$ (Figure 5.A in article I), indicating that this gene might be a target of miR-203a. Target analysis of PXDN in the online database TargetScan (Friedman, Farh et al. 2009) revealed that miR-203a likely had a highly conserved 8mer binding site at position 106-113 on PXDN 3‘UTR, as well as two poorly conserved sites 638-645 and 1618-1624 (Figure 8 and Figure 5.B in article I). To investigate further the relationship between *PXDN* and miR-203a we then decided to treat D492M with a miR-203a-3p mimic, which resulted in a significant downregulation of PXDN compared to D492M treated with negative control mimic. Furthermore, we showed that D492M$^{miR-203a}$ cells showed significant PXDN upregulation after treatment with miR-203a inhibitor (Figure 5.C in article I). However, for binding to be confirmed between miRNA and its target 3‘UTR, a luciferase reporter assay had to be performed (Figure 9). Here, we cloned the miR-203a-3p target sequence at the PXDN 3‘UTR (position 106-113) into a luciferase reporter plasmid, using deleted sequence target site as control. We

then transfected HEK-293T cells with each plasmid along with miR-203a-3p mimic, where the cells transfected with the plasmid containing PXDN 3‘UTR sequence showed significantly less luciferase activity than the control (Figure 5.D in article I). This confirmed that miR-203a regulates the expression of PXDN by binding to position 106-113 on the 3‘UTR, which is novel and previously unpublished.

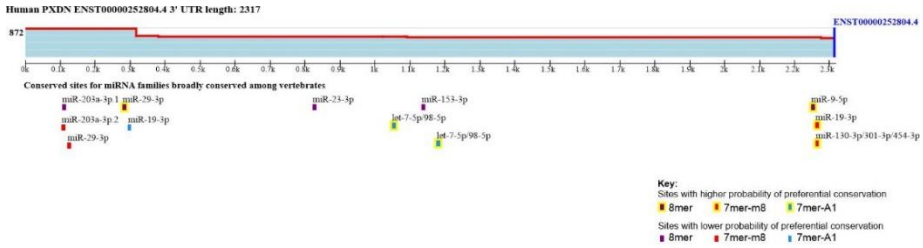


Figure 8. 3‘UTR of PXDN with predicted targeting miRNAs. miR-203a has a high probability of target binding, with 8mer complementary binding site. Image obtained from the TargetScan database v.7.2 (Agarwal, Bell et al. 2015).

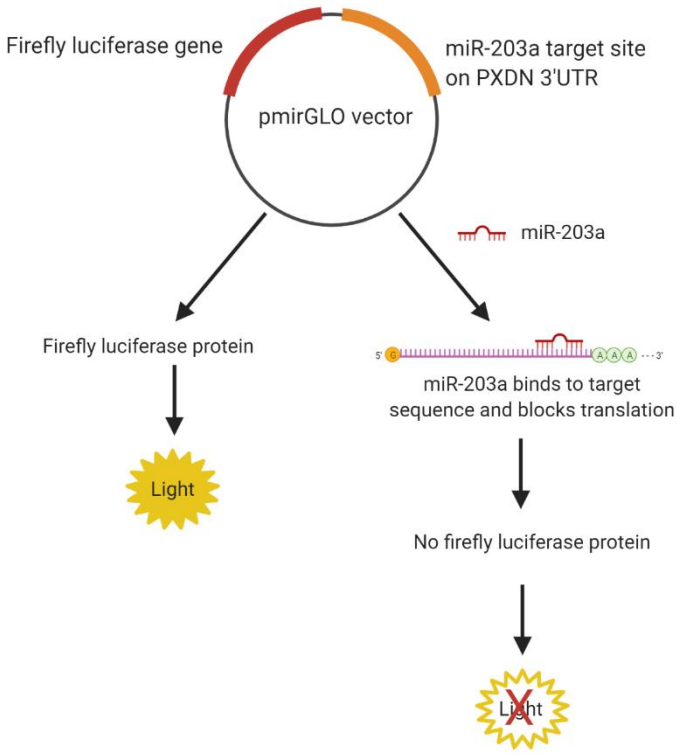


Figure 9. The principle of luciferase reporter assay for the confirmation of miR-203a binding to *PXDN* 3'UTR target site. A pmirGLO plasmid containing the target sequence of the 3'UTR of PXDN was transfected into HEK-293T cells. The cells were then treated with a miR-203a mimic that, upon binding to the PXDN 3'UTR target sequence on the plasmid, inhibited the formation of firefly luciferase protein and thus inhibited light emission (right). Conversely, control samples treated with negative control mimic did not bind to the target sequence, thereby allowing the production of luciferase protein and light emission (left). Adapted from the Promega protocol for pmirGLO vector plasmids, available at www.promega.com.

4.1.5 Regulation of *PXDN* expression via other miRNAs

Information on the mechanisms behind PXDN regulation is limited. Few studies have been published on the subject and miR-203a is the first miRNA that was confirmed to be a suppressor of PXDN via Luciferase Reporter Assay, as is described in article I. However, analysis of TargetScan reveals other miRNAs that likely target the 3‘UTR of PXDN (Table 1). Besides miR-203, two other miRNAs that are significantly differentially expressed between D492 and D492M at day 14 of 3D culture (Figure 1.A. in article I) likely target PXDN: MiR-23-3p and miR-130-3p/5p. Of these miR-130-3p has the highest probability of conserved targeting (P_{CT}) score, which can be interpreted as the highest likelihood of binding. Furthermore, six miRNAs, including miR-203, that possibly target PXDN are significantly differentially expressed during branching morphogenesis in D492 in 3D (Figure 3.A. in article I).

Table 4. MiRNAs with predicted target sequences on conserved *PXDN* 3'UTR sites. (*) MiRNAs that were significantly differentially expressed between D492 and D492M at day 14 of 3D culture. (**) MiRNAs significantly differentially expressed between timepoints during branching morphogenesis in D492 in 3D. (***) MiRNAs with confirmed binding to *PXDN*. P_{CT}: Probability of conserved targeting for a single target site. Data obtained from TargetScan, version 7.2 (www.targetscan.org/vert_72/).

miRNA	3'UTR position	Seed match	P_{CT}
hsa-miR-29-3p	279-286	8mer	0.95
hsa-miR-9-5p	2248-2255	8mer	0.95
hsa-miR-4500	1047-1053	7mer-1A	0.94
hsa-miR-4458	1047-1053	7mer-1A	0.94
hsa-let-7-5p(**)	1047-1053	7mer-1A	0.94
hsa-miR-98-5p	1047-1053	7mer-1A	0.94
hsa-miR-98-5p	1177-1183	7mer-1A	0.94
hsa-let-7-5p(**)	1177-1183	7mer-1A	0.94
hsa-miR-4500	1177-1183	7mer-1A	0.94
hsa-miR-4458	1177-1183	7mer-1A	0.94
hsa-miR-19-3p(**)	2260-2266	7mer-m8	0.83
hsa-miR-454-3p(**)	2261-2267	7mer-m8	0.8
hsa-miR-4295	2261-2267	7mer-m8	0.8
hsa-miR-301-3p	2261-2267	7mer-m8	0.8
hsa-miR-3666	2261-2267	7mer-m8	0.8
hsa-miR-130-3p(*)	2261-2267	7mer-m8	0.8
hsa-miR-19-3p	295-301	7mer-1A	0.58
hsa-miR-23-3p(*)(**)	823-830	8mer	0.53
hsa-miR-130-5p(*)	823-830	8mer	0.53
hsa-miR-29-3p	128-134	7mer-m8	0.45
hsa-miR-203-3p (*)(**)(***)	106-113	8mer	0.23

4.1.6 *PXDN* and mesenchymal traits

Considering that *PXDN* was significantly upregulated in D492M compared to D492 we decided to see what influence silencing of *PXDN* would have in D492M by utilizing *PXDN* siRNA. Interestingly, we saw that this led to a reduction in proliferation and increased sensitivity to apoptosis (Figures 6.A and 6.B in article I). In EMT increased proliferation and resistance to apoptosis is often attributed to the mesenchymal phenotype (Hanahan and Weinberg 2011), and one could therefore imagine that *PXDN* silencing was inducing epithelial traits in the otherwise mesenchymal D492M. Interestingly, Sitole and Mavri-Damelin (2018) reported the opposite effect in two cervical carcinoma cell lines, where EMT inducing treatment with TGFβ caused significant downregulation of *PXDN* with concomitant downregulation of the epithelial marker E-cadherin and upregulation of the mesenchymal markers SNAI1 and vimentin. Furthermore, they showed via luciferase reporter assay that SNAI1 directly represses *PXDN* expression. In this context D492M expresses *SNAI1* while D492 does not (Hilmarsdóttir, Briem et al. 2015). Other publications have also linked *PXDN* to the mesenchymal phenotype. Jayachandran et al. (2016) reported that *PXDN* expression was consistently elevated in mesenchymal-like melanoma cells. When *PXDN* was silenced in these cells a reduction in invasion and migration was reported. Furthermore, *PXDN* expression has also been linked to mesenchymal traits, including increased migration and invasion, in both ovarian and prostate cancer cells (Barnett, Arnold et al. 2011, Zheng and Liang 2018). The linkage between a mesenchymal phenotype and *PXDN* therefore seems real, at least in cancer cells, which is interesting considering the known mechanism of action for the PXDN protein as a crosslinker of collagen IV in epithelial tissue basement membrane.

In conclusion, we found that miR-203a is significantly downregulated in EMT and that downregulation is not due to promoter methylation. Furthermore, miR-203a is also upregulated during branching morphogenesis in D492 in 3D-rBM and is closely affiliated to the epithelial phenotype confirmed via functional assays and increased expression of epithelial markers. We also

confirmed via luciferase reporter assay that miR-203a binds to position 106-113 on *PXDN* 3‘UTR, resulting in significantly reduced luciferase activity in HEK-293T cells transfected with a luciferase reporter plasmid containing the 3‘UTR target site of *PXDN*. Furthermore, preliminary data indicates that miR-29a and miR-23a also target *PXDN,* thereby contributing to its regulation. Finally, *PXDN* silencing in D492M reduced proliferation and increased sensitivity to chemically induced apoptosis, traits typically related to the epithelial phenotype, raising the question of whether *PXDN* might be involved in epithelial cell plasticity.

4.2 Article II: *PXDN* enhances basal phenotype and inhibits branching morphogenesis in breast epithelial progenitor cell line D492 (Published in Journal of Mammary Gland Biology and Neoplasia in 2022)

In article 2, of which I am the first author, I focused on PXDN. I explored *PXDN* expression pattern in the breast and applied the D492 cell line, cultured in monolayer and 3D-rBM, to study its function in epithelial cell plasticity, branching morphogenesis and EMT. Finally, I investigated the possible role of *PXDN* in breast cancer.

4.2.1 Breast tissue expression of *PXDN*

When we identified *PXDN* as a target of miR-203a, its role in breast development and neoplasia had not been studied to any extent. As no published material was available on the subject the relevance and overall expression pattern of *PXDN* in the breast was unknown. In the Human Protein Atlas (HPA) (Uhlén, Fagerberg et al. 2015), RNA expression of *PXDN* had been confirmed in most epithelial tissues except for the upper respiratory tract and oral mucosa, while protein level confirmation was still lacking for all tissue types. Epithelial tissue expression was unsurprising given the known role of PXDN as a crosslinker of collagen IV within the basement membrane, as previously discussed. However, interestingly, female reproductive organs collectively expressed higher levels of *PXDN* compared to other organ systems. On closer inspection this marked difference in expression can mostly be attributed to high levels of *PXDN* in the breast, corresponding to 60.8 NX (tissue expression above normalized expression) while other female reproductive organs (the vagina, ovary, fallopian tube, endometrium, cervix and placenta) each have expression under 20.5 NX (Figure 1.a in article II). To confirm the results from HPA I performed immunohistochemistry for PXDN in paraffin-embedded normal breast tissue. This revealed that protein level expression was present in both the luminal and basal/myoepithelial layer of the epithelium, in fibroblasts and endothelium (Figure 10 and Figure 1.b in article II).

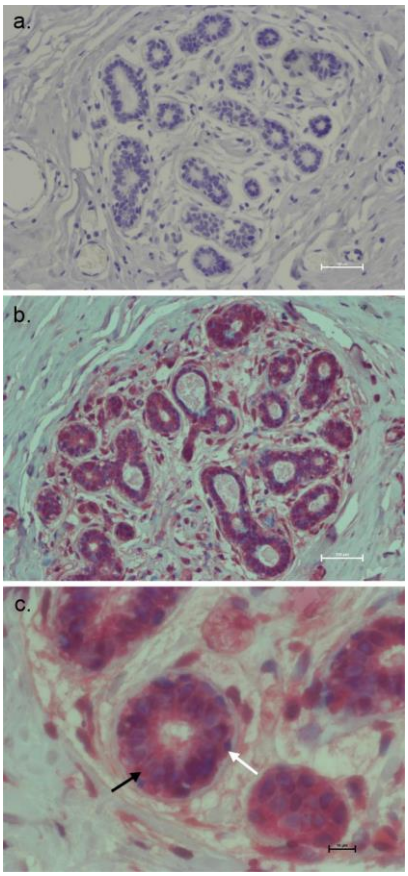


Figure 10. The context of PXDN expression pattern in normal human mammary gland. a. Negative control sample from normal human mammary gland embedded in paraffin, immunostained with hematoxylin. Scale bar = 100 µm. b. Immunohistochemistry for PXDN expression in a TDLU reveals positive (red) staining in epithelium, fibroblasts and endothelial cells. Counterstaining was performed with hematoxylin. Scale bar = 100 µm. c. A closer look at an acinus shows that PXDN is present in both luminal cells (black arrow) and myoepithelial cells (white arrow). Counterstaining performed with hematoxylin. Scale bar = 10 µm.

Next, I decided to investigate the mRNA-level expression of *PXDN* in four main primary cell types isolated from reduction mammoplasties: LEPs, MEPs, fibroblasts and BRENCs (Figure 1.c in article II). Confirmation of EpCam expression in MEPs and LEPs was performed via Western blot (Supplementary figure 1 in article II). The results showed that *PXDN* was expressed in all cell types, although expression was significantly higher in BRENCs compared to epithelial cells. This was not surprising since high concentration of *PXDN* has previously been reported in endothelium (Cheng, Salerno et al. 2008). Furthermore, MEPs expressed lower levels of *PXDN* than LEPs, although the difference was not significant. However, there was considerable variation in expression within MEPs. This could be attributed to individual variation, due to factors such as tissue composition or hormonal status, or that *PXDN* expression can vary within the MEP subpopulation. Interestingly, Western blot showed higher protein level expression of PXDN in MEPs than in LEPs, although this could be attributed to post-translational modifications of the protein (Figure 1.d in article II).

PXDN expression in LEPs is interesting with regards to its role as a crosslinker of collagen IV. Previous research has shown that when cultured in collagen I matrix, without the exogenous collagen IV provided in rBM, LEPs form acini but are unable to polarize correctly in the absence of MEPs. Furthermore, LEPs did not produce their own collagen IV (Gudjonsson, Rønnov-Jessen et al. 2002), which may indicate that MEPs are the primary generators of the basement membrane. This, however, does not exclude the possibility that LEPs can influence the structure of the basement membrane by expressing ECM remodeling factors. Finally, since *PXDN* is expressed in mammary fibroblasts, located within the inter- or intralobular stroma, it would be interesting to see whether increased stromal stiffness in the breast influences *PXDN* expression in the fibroblasts although this falls outside the scope of this thesis. It also raises the question of whether *PXDN* possesses other unknown roles besides crosslinking collagen IV, since collagen IV is exclusively found within the basement membrane and not in the stroma (Khoshnoodi, Pedchenko et al. 2008). Interestingly, overexpression of *PXDN*

in D492 cells did not alter the expression of *COL4A1* and *COL4A5* compared to D492^{empty} (Supplementary figure 2 in article II), which indicates that PXDN is not involved in the regulation of collagen IV.

To confirm the presence of PXDN protein, I decided to perform immunohistochemistry on normal paraffin-embedded breast tissue. This revealed consistent protein signal within all cells of the functional epithelium, in both TDLUs and ducts, as well as in fibroblasts and endothelial cells (Figure 1.b in article II). It was also readily apparent that PXDN expression was not restricted to the basal epithelial compartment, consistent with evenly distributed mRNA-level expression in both LEPs and MEPs. Furthermore, fibrils positive for PXDN were observed within the intralobular and interlobular stroma where collagen IV is not present in normal tissue.

To summarize, both publicly available data along with qRT-PCR results and immunohistochemistry confirm for the first time that PXDN is expressed in normal breast tissue and with this in mind, I decided to study the functional role of PXDN by overexpressing the gene in the D492 cell line.

4.2.2 Overexpression of *PXDN* induces basal epithelial phenotype

To achieve overexpression of *PXDN* in D492 I utilized lentiviral transduction which resulted in 65-fold upregulation in D492^{PXDN} compared to D492^{empty}. I confirmed the overexpression with qRT-PCR, Western blot and immunofluorescent staining (Figure 2.a in article II). When cultured in monolayer the alteration in phenotype in D492^{PXDN} was evident: While D492^{empty} cells had relatively even distribution on the culture surface, D492^{PXDN} cells adhered more to each other and formed patches that eventually merged with increasing confluency. Furthermore, D492^{PXDN} cells were smaller and more homogeneous in appearance compared to D492^{empty} cells (Figure 2.b in article II). Results from qRT-PCR revealed significant upregulation of basal cell markers *KRT14* and *KRT5/6* with simultaneous downregulation in the luminal marker *KRT19* in D492^{PXDN} (Figure 3.a in article II), which I confirmed by Western blot (Figure 3.b in article II).

Immunofluorescent staining of cells in monolayer confirmed the difference in KRT5/6 and KRT14 expression between D492^{PXDN} and D492^{empty}, where overexpression of *PXDN* resulted in increased expression of both markers. However, the difference between the two cell lines was more subtle with regards to KRT19 expression. In fact, most D492^{PXDN} cells display dual expression of KRT14 and KRT19 (Figure 3.c in article II). A previous study identified KRT14-positive basal cells with concurrent KRT19 expression as progenitor cells in the breast (Gudjonsson, Villadsen et al. 2002, Petersen, Gudjonsson et al. 2003).

4.2.3 Branching morphogenesis is inhibited in D492^{PXDN}

Since a distinct feature of the D492 cell line is the ability to form branching structures when cultured in 3D-rBM, I decided to see whether a PXDN induced change in phenotype could also be observed in 3D. For this purpose, I utilized the 3D on-top method and followed the colony formation in IncuCyte, which allowed me to observe cell movement and the morphological timeline in greater detail than the traditional fixed timepoint observation method. This revealed that during the first 24 hours of culture D492^{PXDN}, starting as single cells after plating, were more mobile and clumped together so that colonies had already begun to form the next day. Meanwhile, D492^{empty} remained as single cells. On day 4 of culture, D492^{empty} had formed colonies in the early stages of branching. However, these colonies were considerably smaller than D492^{PXDN} colonies, which also showed no signs of branching and retained a solid round morphology. Four days later, on day 8 of culture, D492^{empty} colonies had reached the same size as D492^{PXDN}, and phenotypic difference was evident (Figure 11). In short, while D492^{empty} colonies retain the branching ability of the original D492 cell line, *PXDN* overexpression is maintained in D492^{PXDN} cells in 3D culture (Figure 4.b in article II) and inhibits branching (Figure 4.a in article II). Furthermore, the observations I made during the first days of the 3D on-top culture indicated that while D492^{PXDN} were able to migrate together to kickstart colony formation, D492^{empty} likely grew colonies from single or few cells which is a more time-consuming process. After quantification of colony morphology, I

found that there was no significant difference in colony size and number between D492^{PXDN} and D492^{empty} on day 8 of culture. However, about 45% of D492^{empty} colonies had branching morphology while all D492^{PXDN} colonies were solid spheroid (Figure 4.c in article II). In summary, early colony formation was enhanced in D492^{PXDN} although colony number remained roughly similar to D492^{empty} at later stages. However, the most important difference lies in the colony morphology as branching morphogenesis is a key feature of the D492 cell line (Briem, Ingthorsson et al. 2019).

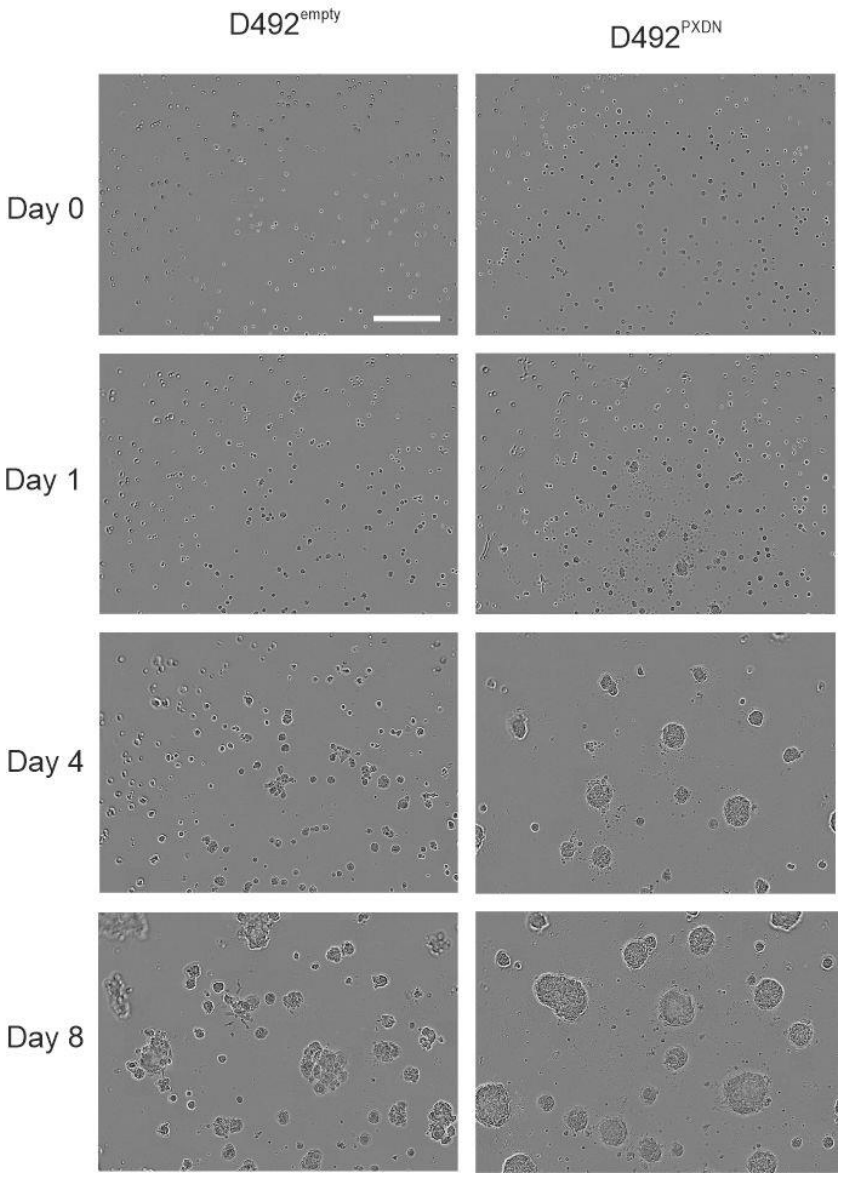


Figure 11. Colony formation in D492^{empty} and D492^{PXDN}. D492^{PXDN} cells show increased affinity towards each other and have started migration and colony formation 24 hours after plating on top of Matrigel. On day 4 D492^{PXDN} colonies are fewer in number although bigger in size while D492^{empty} shows signs of early branching. On day 8 D492^{empty} colonies have reached a similar size to D492^{PXDN} and show branching morphology in combination with spheroid morphology. Branching is inhibited in D492^{PXDN}. Scale bar = 300 µm.

4.2.4 *PXDN* induces migration and invasion in D492

Next, I was interested in seeing whether altered phenotype and morphology of D492^{PXDN} cells were reflected in function as well. Proliferation assay showed little to no significant difference in proliferation between the two cell lines (Figure 5.a in article II). However, D492^{PXDN} cells were significiantly more sensitive to chemically induced apoptosis than D492^{empty} (Supplementary figure 3 in article II). A possible mechanism behind the difference in cell death between D492^{PXDN} and D492^{empty} might be found in the peroxidase activity of PXDN. The peroxidase domain of PXDN utilizes H_2O_2 and yields hypochlorous acid, which is a strong oxidant and an inducer of oxidative stress through the generation of reactive oxygen species (ROS) (Cheng, Salerno et al. 2008, Yuan, Wang et al. 2009, Ma, Zhang et al. 2013). Camptotechin utilizes ROS to provoke cell cycle arrest and apoptosis (Prasad Tharanga Jayasooriya, Dilshara et al. 2018). In migration assay, performed by monitoring cell migration on collagen I coated surface after wound-making, D492^{PXDN} cells closed the wound significantly faster than D492^{empty}. In fact, only 12 hours post-wound-making the wound area of D492^{PXDN} cells had reached around 90% confluence (Figure 5.b in article II). When the assay was stopped at 20 hours post wound-making, D492^{empty} cells had only reached around 70% confluence. Similarly, in invasion assay D492^{PXDN} reached peak relative wound density 24 hours post-wound-making, which D492^{empty} cells had not reached when the assay was stopped at 70 hours (Figure 5.c in article II). Sequence images of the wound closure revealed that wound edges of D492^{PXDN} cells had already merged at 12 hours post wound making (Figure 5.d in article II). As there was no significant difference in proliferation between the two cell lines this pattern of increased migration and invasion in D492^{PXDN} could not be explained by an increased proliferation rate. A closer look at time sequence images from the invasion process revealed that while D492^{empty} cells formed clear filopodia at the wound edge, the edge of D492^{PXDN} cells formed much smaller and finer protrusions that also can be described as filopodia (Figure 5.e in article II). Filopodia are comprised of highly crosslinked and organize branched actin

filaments and have been linked to mesenchymal phenotype (Albuschies and Vogel 2013, Jacquemet, Hamidi et al. 2015, Schaks, Giannone et al. 2019). However, when inspecting the actin filament immunostaining of D492^{empty} and D492^{PXDN}, one notices that D492^{empty} cells at the wound edge resemble mesenchymal cells that are only loosely connected to their neighboring cells. In contrast, D492^{PXDN} cells maintain close cell-cell adhesion. This might indicate that instead of employing EMT to acquire cell motility, D492^{PXDN} cells use collective migration where the epithelial phenotype with intact cell-cell contact is maintained (Ewald, Brenot et al. 2008, Ilina and Friedl 2009).

4.2.5 *PXDN* affects D492 cell plasticity

Due to the profound changes in phenotype and function of D492^{PXDN} cells, I decided to perform RNA-sequencing comparing the cell line to D492^{empty} in monolayer and 3D. Results in monolayer revealed that basal/myoepithelial gene signature was upregulated compared to luminal genes (Figure. 6.a in article II), supporting the results from qRT-PCR. Furthermore, *TP63* was one of the most upregulated genes in D492^{PXDN} overall (Figure 6.b in article II). As monolayer marker expression had indicated that cell plasticity was inhibited in D492^{PXDN} I was interested in seeing whether this phenotype was preserved in 3D culture as lack of branching morphogenesis suggested. Indeed, a similar trend in luminal and basal/myoepithelial marker expression was evident in 3D culture, where *KRT14* and *TP63* were most significantly upregulated (Figure 12). Upregulation of *TP63* in D492^{PXDN} was confirmed with qRT-PCR, Western blot, and immunofluorescent staining (Figure 6.c-e. in article II). To investigate the possibility of a feedback loop between *PXDN* and *TP63* expression I performed *TP63* knock-down in D492^{PXDN}, using two different siRNAs. Although a successful knock-down of *TP63* was reached with both siRNAs (Supplementary figure 4.a in article II), this was not sufficient to significantly affect the expression of *PXDN* (Supplementary figure 4.b in article II) or the expression of *KRT14* and *KRT19* (Supplementary figure 4.c in article II). These findings indicate that PXDN expression induces upregulation of p63 but not vice versa. It also indicates that other basal mediators might be involved in the induction of basal phenotype observed in

D492^{PXDN}. Transient siRNA mediated knock-down of *PXDN* in D492 cells also did not change the expression of *KRT14, KRT19* or *TP63* (Supplementary figure 5 in article II), although this was expected since physiological *PXDN* expression in D492 is very low.

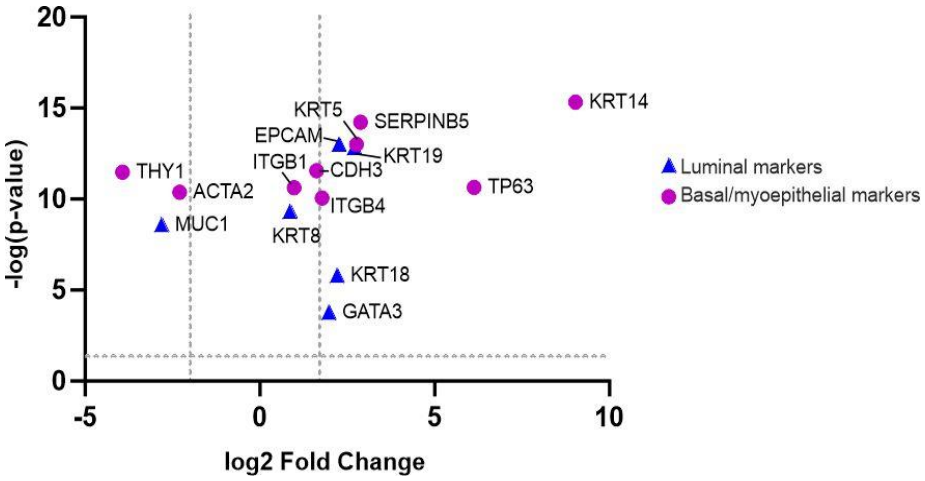


Figure 12. The basal differentiation of D492^{PXDN} cells is maintained in 3D culture. Volcano plot of results from RNA sequencing, showing luminal and basal/myoepithelial markers in D492^{PXDN} vs. D492^{empty} in 3D culture. Compared to monolayer (Fig. 5.a in manuscript II) overall marker trend is similar, indicating that suppression in cell plasticity as an effect of PXDN overexpression is preserved in 3D. Most notably, myoepithelial specific markers THY1 and ACTA2 are both significantly downregulated in D492^{PXDN} cells.

KRT14+/KRT19+ cells have previously been identified as progenitor cells in the breast epithelium (Gudjonsson, Villadsen et al. 2002, Petersen, Gudjonsson et al. 2003). Boecker et al. (2018) proposed a model of mammary gland epithelial cell differentiation based on the expression pattern of key luminal and basal/myoepithelial markers (Figure 13). They suggest that KRT5+/KRT14+/p63+ basal cells represent a progenitor population within the epithelium that can differentiate into either myoepithelial or luminal lineage. A previous study by Villadsen et al. (2007) presented evidence that KRT14+/KRT19+ had stem cell properties as they were able to form branching structures *in vitro*, forming cells with both luminal and myoepithelial characteristics while also maintaining their own population. The KRT5+/KRT14+/KRT19+/p63+ status of D492^{PXDN} cells, confirmed on the

mRNA level by both qRT-PCR and RNA-sequencing and on the protein level by both immunostaining and Western blot, indicates that they have adopted a basal phenotype. Furthermore, when comparing immunostainings of D492^{PXDN} and D492^{empty} it became evident that while D492^{empty} are more heterogeneous in the expression of KRT5, KRT14 and KRT19, which reflects the progenitor cell aspect of the D492 cell line, D942^{PXDN} cells are more homogeneous with regards to the expression pattern of the three markers (Figure 3.c in article II). While this data indicates that D492^{PXDN} cells are losing the ability to form cells with luminal phenotype it raises the question of whether they are still able to form cells with myoepithelial phenotype.

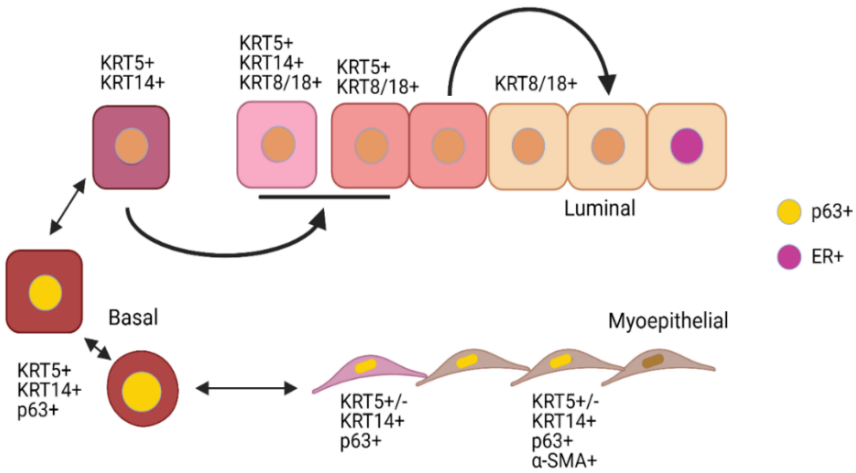


Figure 13. Suggested model of epithelial differentiation in the breast. Model of mammary epithelial cell differentiation based on the expression of luminal markers KRT8 and KRT18, basal/myoepithelial markers KRT5, KRT14 and p63 and myoepithelial specific marker α-SMA, proposed by Boecker et al. (2018). While this study reports that KRT5+KRT14+/- cells are restricted to the luminal epithelium, KRT5+/KRT14+/p63+ cells are defined at basal cells and can differentiate into both luminal and myoepithelial lineages. Figure adapted from (Boecker, van Horn et al. 2018).

To answer this question, I applied the RNA sequencing data to investigate the expression of *THY1* and *ACTA2* (α-SMA), both myoepithelial specific markers (Sapino, Macri et al. 1990, Gudjonsson, Villadsen et al. 2002). In monolayer culture, *THY1* was significantly downregulated while *ACTA2* did

not reach 2-fold downregulation although the difference was significant (Figure 5.a in article II) This pattern was maintained in 3D culture where both markers were significiantly downregulated (Figure 13). One of the defining characteristics of the D492 cell line is the ability to sustain its own population while also yielding THY1+/α-SMA+ expressing cells with myoepithelial characteristics (Gudjonsson, Villadsen et al. 2002). This ability seems to be suppressed in D492^{PXDN} cells in both monolayer and 3D culture, indicating that although *PXDN* overexpression is inducing basal phenotype in D492 cells the cells are losing the ability to form cells with myoepithelial phenotype at the same time. Hilmarsdottir et al. (2015) had previously published data that showed that both the luminal and myoepithelial phenotype, formed due to D492 plasticity, were necessary for branching morphogenesis in 3D-rBM. My data indicates that D492^{PXDN} cells have lost plasticity and only form cells with basal, not luminal or myoepithelial, phenotype which likely contributes to the complete inhibition of branching morphogenesis in 3D culture.

4.2.6 *PXDN* enriches pathways involved in development and differentiation

Since I had seen that basal gene signatures were upregulated in D492^{PXDN} compared to D492^{empty} I was curious to see whether this would be reflected in changes in pathway enrichment. Therefore, I performed GSEA GO-analysis using the C5 ontology gene sets. This revealed that pathways involved in development and differentiation were most significantly enriched in monolayer (Figure 7.a in article II), while metabolic pathways were mainly enriched in 3D (Figure 7.b in article II). I was most interested in the GO_EPITHELIAL_CELL_DIFFERENTIATION pathway as that reflected the experimental results and included *TP63, KRT14* and *KRT5* along with associated genes. Volcano plot of genes involved in the pathway showed that over two-thirds of significantly differentially expressed genes were upregulated, while approximately one-third were downregulated (Figure 7.c in article II). This is also reflected in a heatmap displaying the top 40 significantly differentially expressed genes, where most genes are upregulated in D492^{PXDN} cells compared to D492^{empty} (Figure 7.d in article II).

4.2.7 *PXDN* inhibits EMT in D492 in monolayer and 3D culture

Having observed that branching morphogenesis was inhibited in D492^{PXDN} in addition to the lack of mesenchymal phenotype in cell invasion assay, I asked the question how EMT was influenced by overexpression of *PXDN*. EMT can facilitate migration and invasion in epithelial cells, where the formation of cell protrusions is one of the phenotypical changes observed in cells that have undergone EMT and acquired motility (Sigurdsson, Hilmarsdottir et al. 2011, Oyanagi, Ogawa et al. 2012, Ye and Weinberg 2015, Aiello, Maddipati et al. 2018). I therefore applied the RNA-sequencing data from D492^{empty} and D492^{PXDN} in monolayer and performed GSEA analysis using the HALLMARK dataset. Interestingly, this revealed that EMT was the most downregulated pathway in D492^{PXDN} in both monolayer (Figure 8.a in article t II) and 3D (Figure 8.b in article II), meaning that the genes belonging to this pathway were significantly downregulated in the cell line beyond what would be expected by chance. Next, I looked at the expression pattern of known EMT factors and presented factors that were significantly differentially expressed in a heatmap (Figures 8.a and 8.b, bottom, in article II). 7 out of 12 EMT genes were downregulated in D492^{PXDN} compared to D492^{empty} in monolayer, and 13 out of 16 EMT genes in 3D. Currently, work by Sitole et al. (2018) is the only paper linking *PXDN* to EMT. They demonstrated that *PXDN* was downregulated in cervical carcinoma cell lines after TGFβ1-induced EMT. Furthermore, they showed that TGFβ1 facilitated binding of SNAI1, an EMT-transcription factor (Nieto 2002, Barrallo-Gimeno and Nieto 2005), to the *PXDN* promoter, causing downregulation of *PXDN*. The fact that *SNAI1* is significantly downregulated in D492^{PXDN} might indicate that there is a reciprocal feedback loop between the two genes (Figure 14), pointing to a possible regulatory role of *PXDN* in EMT. *SNAI1* downregulation could also simply be a response to other differentially expressed factors in D492^{PXDN} that collectively suppress EMT. However, further investigation into the precise role of *PXDN* in EMT is beyond the scope of this paper. Although the link between *PXDN* and EMT was not explored further the pathway analysis provides important information when it comes to explaining the complete

inhibition of branching morphogenes. As discussed in the introduction, EMT is an important mechanism in branching morphogenesis.

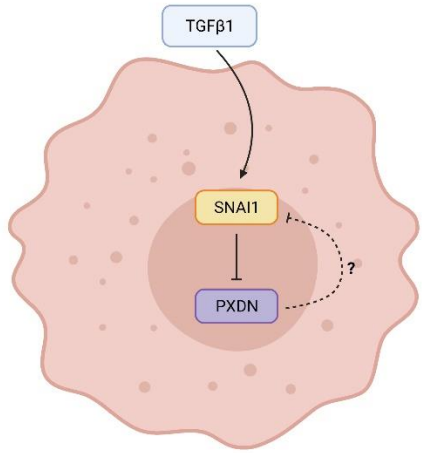


Figure 14. A possible reciprocal feedback loop between SNAI1 and PXDN. SNAI1, induced by TGFβ1, binds to the promoter of *PXDN* (Sitole et al. 2018). *SNAI1* expression is reduced in D492PXDN which might indicate the presence of a reciprocal feedback loop between SNAI1 and PXDN.

4.2.8 *FGFR2* is significantly downregulated in D492^{PXDN}

Although suppression of cell plasticity could account for branching morphogenesis inhibition, I was interested in seeing whether changes in other mechanisms involved in branching were altered in D492^{PXDN}. Most notably, members of the FGF pathway were significantly differentially expressed. The FGF pathway is critical in embryonic and postnatal mammary gland development in the mouse (Mailleux, Spencer-Dene et al. 2002, Veltmaat, Relaix et al. 2006, Parsa, Ramasamy et al. 2008). The pathway includes four receptors (FGFR1-4) and 22 different FGF ligands that can each bind to one or more receptors (Eswarakumar, Lax et al. 2005). Both FGFR1 and FGFR2 are expressed in the mammary gland epithelium during development (Dillon, Spencer-Dene et al. 2004, Lu, Ewald et al. 2008, Pond, Bin et al. 2013), where *FGFR1* deletion has been shown to delay mammary gland growth (Pond, Bin et al. 2013). *FGFR2* is an inducer of luminal

differentiation, which is inhibited in *FGFR2* null basal cells along with branching morphogenesis. Interestingly, RNA-sequencing revealed that *FGFR2* was significantly downregulated in $D492^{PXDN}$ compared to $D492^{empty}$ in both monolayer and 3D (Supplementary figure 7 in article II). Although the relationship between *PXDN* and members of the FGR pathway were not investigated further these results provide grounds for further investigations.

4.2.9 *PXDN* in breast cancer

After searching for *PXDN* in the BreastMark online database (Madden, Clarke et al. 2013) I discovered that women suffering from HER2-positive breast cancers had significantly worse distant metastasis free survival (DMFS) if their tumors had high *PXDN* expression compared to women with low *PXDN* expressing tumors (p=0.00482). Follow up analysis in the online Gene Expression-Based Outcome (GOBO) database (Ringnér, Fredlund et al. 2011) confirmed this (p=0.033) in addition to worse DMFS in basal breast cancers (p=0.02445) (Supplemental figure 6.b in article II). Subsequently, I wanted to know how *PXDN* expression influenced overall survival (OS) in breast cancer patients with HER2 positive tumors since I had seen similar results in both the BreastMark and GOBO databases. GOBO analysis, with stratification according to grade and node status, revealed that high PXDN expression reduced the 10 years survival rate from ~60% to ~25%. Furthermore, PXDN increased the risk of positive lymph node status at diagnosis (Hazard ratio of negative lymph nodes around 50%, p=0.004) (Figure 15.a). In the following analysis, where I stratified results according to ER status and grade, I found that this striking difference in OS could be contributed to ER-negative tumors (0.0309) and not ER-positive tumors where the difference in OS was not significant (0.239) (Figure 15.b-c). Women with ER-negative tumors, that already had high grade tumors or metastatic lesions in regional lymph nodes upon diagnosis, showed an even more drastic reduction in OS if the tumors had high *PXDN* expression.

After analyzing the publicly available data I was interested in whether PXDN was expressed in breast tumors. Therefore, I performed

immunohistochemistry on paraffin-embedded breast tumors of HER2-amplified, basal-like and ER-positive subtypes. Interestingly, I found that PXDN was expressed in all tumors to a similar degree, both in the cancer cells and as fibrils within the cancer stroma (Supplementary figure 6.a in article II). If PXDN is present in all breast tumors it raises questions why women with HER2-amplified tumors have a worse prognosis with high PXDN expression.

The reason why PXDN would decrease survival only in HER2-positive breast cancer is unclear. Breaching of the basement membrane is a major event in carcinogenesis as cancer transitions from localized carcinoma *in situ* to infiltrative disease (Rowe and Weiss 2008). As PXDN essentially stabilizes the basement membrane through collagen IV crosslinking one would think that PXDN would serve as a tumor suppressor by matrix strengthening. However, further analysis with the GOBO database reveals that there is an overwhelming significant correlation ($p<1.0^{-10}$) between PXDN expression and stromal gene signature.

Various factors related to the expression of particular genes can negatively influence survival in breast cancer patients and the result is usually a complex mixture of genetic and environmental inputs. However, most of these interplaying factors can be put into one of two categories, and in some cases both: Factors that contribute to tumor aggressiveness and factors that result in increased resistance to therapy. The reason for this drastic difference in survival only in women with HER2-positive ER-negative tumors remains unclear as no clear-cut connection between PXDN and the HER2 receptor has been made.

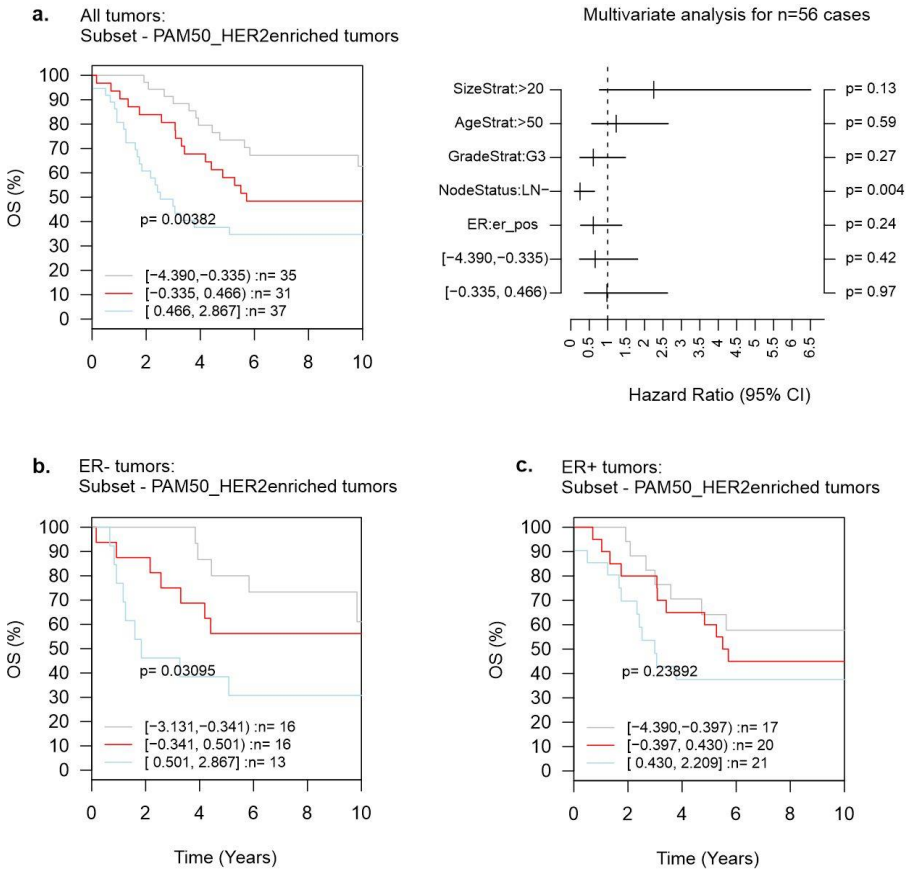


Figure 15. GOBO analysis of *PXDN* expression in breast cancer. a. High expression of PXDN significantly reduces overall survival (OS) in women with HER2-amplified breast cancer (left) and increases the risk of positive lymph node status at diagnosis (right). **b.** The negative effect of PXDN upregulation seen in all HER2-amplified cancers is mainly due to ER-negative tumors (p=0.03095). **c.** The negative effect is not significant in ER-positive tumors (p=0.23892). Data was obtained via the GOBO database (Ringnér, Fredlund et al. 2011) using the PAM50 dataset.

In conclusion, I used publicly available databases, paraffin-embedded breast tissue and primary cells isolated from reduction mammoplasties to analyze the expression pattern of PXDN in the human mammary gland, which suggested that PXDN is present in the whole mammary epithelium as well as in fibroblasts and endothelial cells. Furthermore, I overexpressed *PXDN* in the breast epithelial progenitor cell line D492 and discovered a novel role of *PXDN* in the suppression of branching morphogenesis (Figure 16) by inhibition of cell plasticity and EMT, partially mediated through increased expression of transcription factor p63.

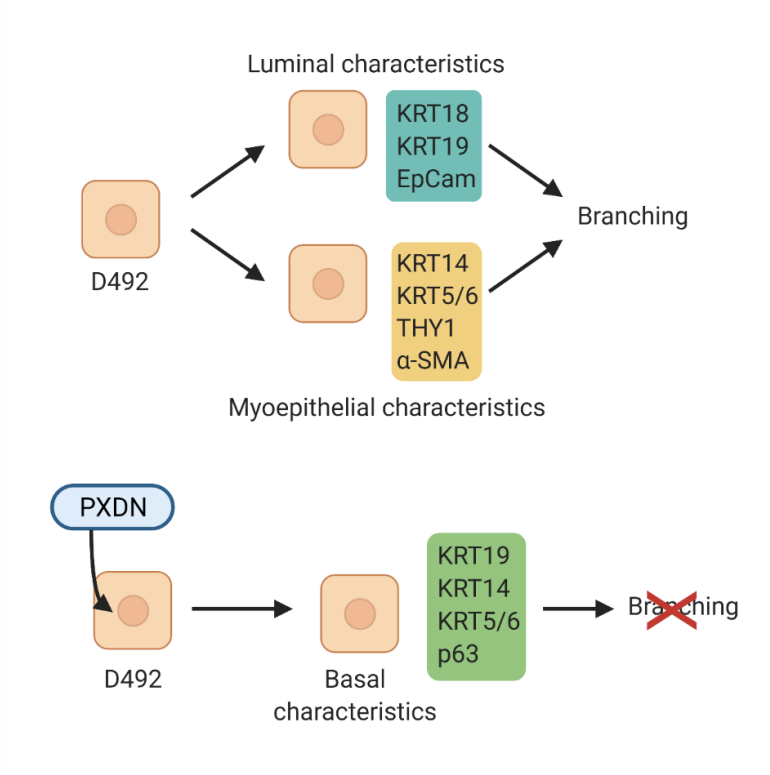


Figure 16. A summary of the effect of *PXDN* overexpression in D492 cells. D492 can form cells with both luminal and myoepithelial characteristics. Both phenotypes are necessary to form branching structures in 3D culture, as was shown by Hilmarsdottir et al. (2015). In D492^{PXDN} cells this plasticity has been inhibited and the cell line yields a population with basal characteristics that is incapable of branching morphogenesis.

4.3 Article III: Expression of ncRNAs on the DLK1-DIO3 locus is associated with basal and mesenchymal phenotype in breast epithelial progenitor cells (published in Frontiers Cell and Developmental Biology in 2020)

In Article III, of which I am the second author, we applied the D492 and D492M cell models to study the lncRNA MEG3 and its effect on cell phenotype and epithelial plasticity. The work on MEG3 is the Ph.D. project of Zuzana Budkova with contributions from other lab members.

4.3.1 Expression of *MEG3* and miRNAs from the DLK1-DIO3 locus is associated with mesenchymal phenotype

MEG3 first came to the attention of the lab after microarray analysis of miRNA expression in D492 and D492M revealed that out of the 25 miRNAs most highly expressed in D492M compared to D492, 15 belonged to the DLK-DIO3 locus (Figure 1.B in Article III), of which *MEG3* is also a member as previously discussed. Furthermore, RNA sequencing revealed that *MEG3* and *MEG8* are among the most differentially expressed lnc-RNAs in D492M (Figure 1.C in Article III). To confirm these results, we selected two miRNAs from each of the maternally expressed clusters in the DLK-DIO3 locus, miR-127 and miR-493 from cluster A and miR-409 and miR-11 from cluster B (Figure 1.D in Article III). We then performed qRT-PCR on D492 and D492M, which revealed that both *MEG3* and each of the miRNAs were significantly upregulated in D492M (Figure 1.E in Article III). Small RNA sequencing on D492 and D492M showed that most miRNAs from the DLK1-DIO3 locus were upregulated in D492M (Supplementary figure 1 in Article III). Furthermore, qRT-PCR performed on material from the epithelial cell line HMLE and its mesenchymal counterpart HMLEmes, also showed significant upregulation of *MEG3* along with the cluster A miRNAs, miR-127 and miR-493, in the mesenchymal line (Figure 1.F in Article III).

Next, we wanted to see whether the expression of *MEG3* and the four miRNAs was also related to the mesenchymal phenotype in primary cells. Indeed, *MEG3* was significantly upregulated in primary mammary fibroblasts

compared to the epithelial compartment, comprised of LEP, MEP and whole organoids. *MEG3* expression was also significantly upregulated in BRENCs which, along with the expression seen in the fibroblasts likely contributed to the increased expression seen in whole tissue (Figure 2.A in Article III). Concurrent upregulation of miR-127, miR-409, miR-411 and miR-493 was also seen in fibroblasts compared to D492, LEP and MEP (Figure 2.B in Article III). GOBO dataset analysis revealed that *MEG3* expression is primarily found within the stromal compartment and correlates with expression of ECM related genes (Supplementary figure 2.A in Article III), and furthermore, MiPanda analysis shows that *MEG3* positively correlates with mesenchymal genes in both normal breast tissue and breast cancer (Supplementary figure 2.B in Article III).

4.3.2 *MEG3* imprinting is not lost in D492M

Given that the DLK1-DIO3 locus is imprinted in adult cells (Zhu, Botticelli et al. 2019), as previously discussed, we were interested in knowing whether upregulation of *MEG3* in D492M occurs due to loss of the expected monoallelic expression. By utilizing pyrosequencing (Harrington, Lin et al. 2013) we observed that on the DNA level in D492M, at the site of a heterozygous single nucleotide polymorphism (SNP) at the *MEG3* locus (C/T, rs4906024), the frequency of the C- and T-alleles was 48% and 52%, respectively, which was close to the expected ratio of 50% for each allele as both D492M and D492 are diploid. However, on the cDNA level the expression of the C allele was 0% and 2% in D492 and D492M, respectively (Figure 3 in Article III). This confirmed the expected monoallelic expression of *MEG3* which led us to the conclusion that increased *MEG3* expression in D492M is not due to loss of imprinting.

By utilizing CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi) we established stable overexpression in D492 ($D492^{MEG3}$) and knockdown in D492M ($D492M^{KD\text{-}MEG3}$), respectively, and control cell lines using sgRNA ($D492^{CTRL}$ and $D492M^{KD\text{-}CTRL}$). Evaluation of the stable cell lines cultured in monolayer revealed no apparent difference in phenotype.

4.3.3 *MEG3* expression in breast cancer

Evidence is emerging that links the mesenchymal and partial EMT phenotype to increased aggressiveness in breast cancers (Hanahan and Weinberg 2011, Chaffer, San Juan et al. 2016, Shibue and Weinberg 2017, Kröger, Afeyan et al. 2019, Jørgensen, Forsare et al. 2020), and others have also reported that EMT and stemness related gene signature can predict the *in vitro* behavior and plasticity of breast cancer cell lines when it comes to therapy resistance (Akbar, Isbilen et al. 2020). Since we saw that *MEG3* expression is correlated with mesenchymal phenotype we wanted to investigate the *MEG3* expression in different breast cancer subtypes and the effect on overall survival (OS) and distant metastasis free survival (DMFS). We found that normal-like tumors had significantly higher expression of *MEG3* ($p=0.0003$) than other breast cancer subtypes, while OS was not affected (Figure 4.A in Article III). The normal-like subgroup, due to its genetic clustering with normal breast tissue, increased stromal compartment and low cellularity has come under increased scrutiny as a misdiagnosis due to its high normal tissue percentage (Peppercorn, Perou et al. 2008, Parker, Mullins et al. 2009, Prat and Perou 2011). As we had found that *MEG3* expression was mainly contained within the breast stroma, in addition to the high stromal component of normal-like tumors we omitted this subgroup from the subsequent survival analysis, which then showed a significant negative effect of high *MEG3* tumor expression ($p=0.01$) on survival in the remaining subgroups. Furthermore, we found that high *MEG3* expression correlated with poorer DMFS in grade 3 tumors ($p=0.00273$) as well as in luminal B tumors ($p=0.04607$) (Figure 4.B in Article III). Luminal B tumors are typically reported as more aggressive than luminal A tumors, with increased expression of proliferative markers, low expression of hormone receptors and higher grade at diagnosis. Overall, luminal B tumors have a similar prognosis as basal-like and HER2-enriched tumors (Ades, Zardavas et al. 2014).

4.3.4 *MEG3* expression influences the expression of miRNAs at the DLK1-DIO3 locus

As previously mentioned, there are two clusters of miRNAs on the maternally expressed DLK1-DIO3 locus. We were interested in how alterations in *MEG3* expression affected the expression of miRNAs in the locus. After confirmation of successful overexpression of *MEG3* in D492 and knockdown in D492M (Figure 5.A in Article III) we performed qRT-PCR for the miRNAs previously mentioned: MiR-127 and miR-493 from cluster A and miR-409 and miR-11 from cluster B. We found that the overexpression of MEG3 in D492 was met with a concurrent increase in the miRNAs. Furthermore, the opposite was seen with MEG3 knockdown in D492M where expression of the four miRNAs was suppressed (Figure 5.B in Article III). When put into context with the results presented in chapter 4.3.1 of this thesis, where the expression of the four miRNAs had a positive correlation with MEG3 expression in both the D492/D492M and HMLE/HLMEmes cell models, there is sufficient evidence to indicate that *MEG3* is involved in the regulation of miRNAs in the DLK1-DIO3 locus, at least in part, or that there is a common regulatory mechanism for both elements. Analysis of expression correlation between *MEG3* and miRNAs, using the Cancer Genome Atlas (TCGA), revealed that 30 out of the 40 miRNAs that positively correlated with *MEG3* belonged to the DLK1-DIO3 locus (Table 1 in Article III). As each of the 54 miRNAs found in the locus have multiple predicted targets it is clear that alterations in *MEG3* expression can significantly affect the cellular phenotype and function.

4.3.5 RNA-sequencing analysis of D492M$^{KD\text{-}MEG3}$

In an attempt to gain insight into the influence *MEG3* manipulation has on gene expression we performed RNA sequencing on D492^{MEG3}, D492M$^{KD\text{-}MEG3}$ and their respective control cell lines. The resulting sequencing data failed to confirm the overexpression of *MEG3* in D492^{MEG3} due to high inter-sample variation. Therefore, we decided to focus on D492M$^{KD\text{-}MEG3}$.

Initially, we generated a volcano plot for the RNA sequencing data from D492M$^{KD\text{-}MEG3}$ and D492M$^{KD\text{-}CTRL}$ to visualize the symmetry and distribution of

the data (Supplementary figure 3.A in Article III). We found that among the 30 most differentially expressed genes were transcripts related to histone modification, metabolism, cell division, differentiation and adhesion as well as the ECM (Supplementary figure 3.B in Article III). Next, we performed Gene Set Enrichment Analysis (GSEA) (Mootha, Lindgren et al. 2003, Subramanian, Tamayo et al. 2005), using the Hallmark dataset as a reference gene list. GSEA enables its users to detect pathway-level patterns in transcriptomic data and can inform whether a given cell mechanism is enriched or not depending on the differential expression and respective statistical significance of the genes involved. This revealed that knockdown of *MEG3* in D492M leads to negative enrichment of EMT (Figure 6.A in Article III). Conversely, we found that expression of luminal epithelial markers was upregulated in $D492M^{KD\text{-}MEG3}$ with concurrent downregulation in mesenchymal markers (Figures 6.B and 6.C in Article III). It has been known for some time that *MEG3* plays a regulatory role in EMT. However, there have been conflicting reports on whether *MEG3* is inducing or suppressing the mesenchymal phenotype, although there is currently more evidence pointing towards the latter (Terashima, Tange et al. 2017, Zhang, Shi et al. 2017, Deng, Fan et al. 2018, Yu, Geng et al. 2018). It was therefore interesting that MEG3 knockdown in D492M was accompanied by reduced EMT and upregulation of luminal markers. Due to these results, we were curious to see whether this would be reflected in cell phenotype and function.

4.3.6 The role of *MEG3* in mesenchymal phenotype and stemness

To confirm the differential expression of mesenchymal markers we had seen in the RNA-sequencing data, we performed qRT-PCR on $D492^{MEG3}$, $D492M^{KD\text{-}MEG3}$ and their respective control cell lines. In $D492^{MEG3}$ mesenchymal markers *SNAI2 and ZEB1* were significantly upregulated compared to $D492^{CTRL}$. Conversely, in $D492M^{KD\text{-}MEG3}$ *SNAI2* and *ZEB1*, along with the additional mesenchymal markers *ZEB2* and *TWIST1* were significantly downregulated compared to $D492M^{KD\text{-}CTRL}$ (Figure 7.A in Article III). These results further confirmed our suspicion that *MEG3* induced mesenchymal properties.

Since *MEG3* knockdown in D492M had also resulted in upregulation of luminal epithelial markers we were curious to see how *MEG3* expression influenced epithelial differentiation in D492 and D492M. This was done by measuring the expression of the luminal marker *KRT19* and basal marker *KRT14*, previously discussed in the context of epithelial cell differentiation, via qRT-PCR and Western blot (Figure 7.B in Article III). This revealed that *KRT19* was significantly downregulated in $D492^{MEG3}$ while *KRT14* was significantly upregulated, compared to $D492^{CTRL}$. This indicates a shift towards a basal phenotype in the D492 cell line when *MEG3* is overexpressed. Conversely, *KRT14* expression was significantly downregulated in $D492M^{KD\text{-}MEG3}$, although no change was observed in the expression of *KRT19*. Furthermore, we found that mRNA expression of the basal markers *TP63* and P-cadherin (*CDH3*) was significantly increased in $D492^{MEG3}$ compared to $D492^{CTRL}$, while significant downregulation of the basal marker *KRT5* was seen in $D492M^{KD\text{-}MEG3}$ compared to $D492M^{KD\text{-}CTRL}$ (Figure 7.D in Article III). On the protein level increased expression of all three markers, CDH3, p63 and KRT5 was confirmed in $D492^{MEG3}$ compared to $D492^{CTRL}$, while a significant decrease was only seen in p63 but not in KRT5 in $D492M^{KD\text{-}MEG3}$ compared to $D492M^{KD\text{-}CTRL}$ (Figure 7.E in Article III). These results indicate that *MEG3* induces differentiation towards the basal phenotype in the D492 model but does not induce EMT (Supplementary figure 4 in Article III). Likewise, *MEG3* knockdown in D492M does not induce MET (Supplementary figure 4 in Article III). This might also indicate that the involvement of *MEG3* in EMT is context dependent and can change according to cell type or even disease state (Figure 17).

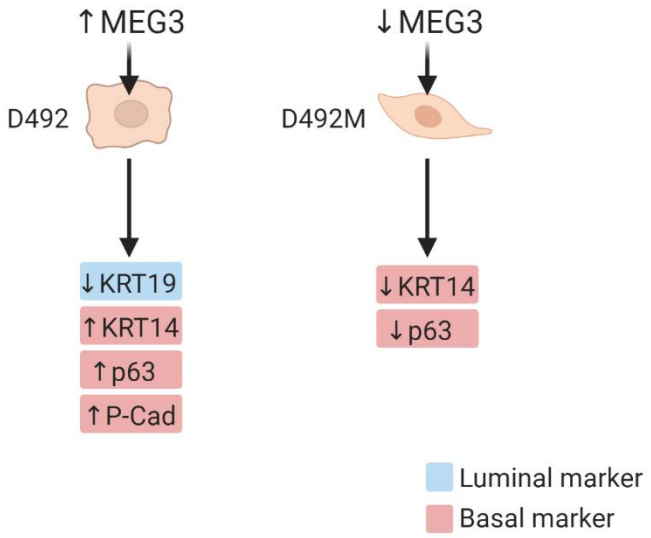


Figure 17. MEG3 effect on phenotype is context dependent. Overexpression of MEG3 in the epithelial D492 induces basal phenotype (left) but does not induce EMT. However, in the mesenchymal D492M, which already expresses higher levels of MEG3, knockdown does not increase the expression of epithelial keratins and basal marker p63 and does not induce MET. On the contrary, basal keratin K14 and p63 are downregulated.

To assess whether the influence of *MEG3* on phenotype would affect the behavior of the cell lines, we performed a series of functional assays. In D492 *MEG3* overexpression decreased sensitivity to chemically induced apoptosis compared to $D492^{CTRL}$ (Figure 8.A left in Article III). Furthermore, we saw an increase in cell migration and colony formation in mammosphere and low attachment assays (Figures 8.B left and 8.C left in Article III). Also, in 3D-rBM co-culture with endothelial cells (HUVECs), $D492^{MEG3}$ showed increased colony size and decreased branching (Figure 8.D in Article III). In contrast, $D492M^{KD\text{-}MEG3}$ had slightly increased susceptibility to chemically induced apoptosis (Figure 8.A right in Article III). Furthermore, $D492M^{KD\text{-}CTRL}$ showed significantly decreased migration as well as reduced colony formation capabilities in mammosphere and low attachment assays (Figures 8.B right and 8.C right in Article III). These results indicate that *MEG3* overexpression induces mesenchymal traits in D492, leading to resistance to apoptosis and

increased motility which has previously been linked to the mesenchymal phenotype, while D492M$^{KD\text{-}MEG3}$ showed the opposite effect (Dontu, Abdallah et al. 2003, Hanahan and Weinberg 2011). Taken together, these results indicate the emergence of a partial EMT-phenotype, induced by *MEG3* expression.

In conclusion, our findings revealed that *MEG3* is associated with mesenchymal phenotype and is upregulated in D492M compared to D492, along with selected miRNAs from two different clusters in the DLK1-DIO3 locus. Furthermore, we found that the observed upregulation in D492M was not due to the loss of imprinting at the locus. Analysis via databases available online also revealed that *MEG3* expression was related to a significantly worse prognosis in breast cancer. Furthermore, overexpression of *MEG3* in D492 enhanced traits related to mesenchymal phenotype and stemness, while *MEG3* knockdown in D492M reduced these factors. Therefore, we postulate that in breast epithelial progenitor cells *MEG3* induces a partial EMT phenotype and increases plasticity. Through these mechanisms, *MEG3* could be involved in breast cancer and contribute to a worse prognosis.

4.4 Unpublished data

4.4.1 A possible link between *PXDN* and stemness

In an attempt to clarify possible links between *PXDN* and breast cancer, we applied flow cytometry assay to evaluate the expression of luminal marker CD24 and basal marker CD44 in $D492^{empty}$ and $D492^{PXDN}$. CD24 and CD44 are commonly used markers for the identification of stem cells, where $CD24^{-}/CD44^{+}$ phenotype has been identified as a feature of cancer stem cells in malignant breast tumors (Al-Hajj, Wicha et al. 2003, Charafe-Jauffret, Monville et al. 2008, Ricardo, Vieira et al. 2011, Dittmer 2018). The majority of $D492^{empty}$ cells, 90% ± SD, are $CD24^{+}/CD44^{+}$ whereas only 30% ± SD of $D492^{PXDN}$ cells are double positive. Moreover, *PXDN* induces a significant shift towards $CD24^{-}/CD44^{+}$ as this fraction goes from 10% ± SD in $D492^{empty}$ up to 50% ± SD in $D492^{PXDN}$. Finally, the $CD24^{-}/CD44^{-}$ population, which is almost not present in $D492^{empty}$, is increased to 15% ± SD in $D492^{PXDN}$ (Figure 18). Ricardo et al. (2011) reported that $CD24^{-/low}/CD44^{+}$ was a feature of basal/mesenchymal phenotype. Furthermore, CD44 expression was significantly correlated with the expression of basal markers. However, they mainly associated $CD24^{-/low}/CD44^{+}$ with basal-like breast cancers. ALDH1 is another commonly used marker for cancer stem cells, which further divides $CD24^{-/low}/CD44^{+}$ cells into $ALDH1^{+}/CD24^{-/low}/CD44^{+}$ and $ALDH1^{-}/CD24^{-/low}/CD44^{+}$ subpopulations that have high tumorigenic and low tumorigenic potential, respectively (Ginestier, Hur et al. 2007, Croker, Goodale et al. 2009). To investigate the ALDH1 activity of $D492^{PXDN}$ compared to $D492^{empty}$ I performed Aldefluor assay. Preliminary results showed that activity of ALDH1 was considerably less in $D492^{PXDN}$, indicating that $D492^{PXDN}$ cells were $ALDH1^{-}$ (Figure 19).

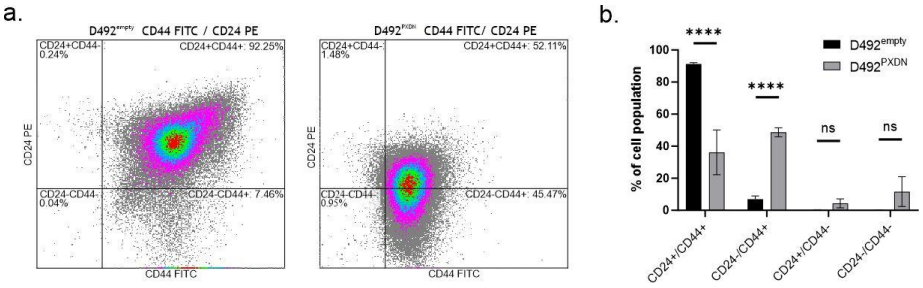


Figure 18. CD24-/CD44+ cell population is enriched in D492^{PXDN}. a. Results from flow cytometry indicated a shift from CD24+/CD44+ in D492^{empty} towards CD24-/CD44+ in D492^{PXDN}. b. A significant decline was observed in the CD24+/CD44+ population of D492^{PXDN} with a concurrent significant rise in the CD24-/CD44+ cell population. Statistical significance was determined by Two-way ANOVA (****p ≤ 0.0001) and data is presented as results of triplicate experiments (mean ± SD).

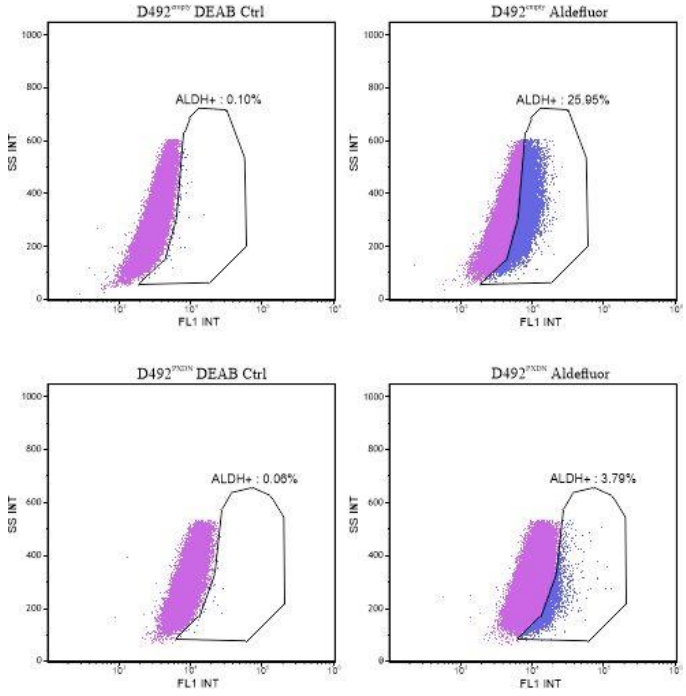


Figure 19. D492^{PXDN} cells have reduced ALDH activity compared to D492^{empty}. Preliminary results from Aldefluor assay indicate that D492^{PXDN} are primarily ALDH^{-} with only 3.79% of the cell population presented as ALDH^{+} compared to 25.95% of D492^{empty} cells.

These results, together with the evidence for basal phenotype in D492^{PXDN} presented in paper #2, might indicate that *PXDN* induces stem-like phenotype. This could partially explain the negative effect high *PXDN* expression has on survival in HER2-positive breast cancers as cancer stem cells are generally considered to be therapy resistant and important contributors to cancer cell heterogeneity (Dalerba, Cho et al. 2007, Visvader 2009, Palomeras, Ruiz-Martínez et al. 2018). However, this does not explain why *PXDN* does not negatively affect survival in basal-like breast cancer, typically considered to be aggressive with a high re-occurrence of lesions (Cleator, Heller et al. 2007, Millikan, Newman et al. 2008, Lund, Trivers et al. 2009). Therefore, the relationship between *PXDN* and features of HER2-positive breast cancers remains a topic in need of further research.

5 Technical considerations

The nature of Ph.D. studies in any field requires the student to acquire skills and theoretical knowledge to solve the problems that come up along the way. Equally important is the ability to critically review one's work to learn from the mistakes made. In this chapter, I will describe the technical obstacles and difficulties I encountered during my Ph.D. project and discuss what I, in hindsight, would have done differently.

5.1 Stable overexpression of *PXDN* in D492 cells

In cell biology, a common method for studying protein function is overexpressing or knocking down the protein coding gene of interest. In all, I spent 18 months out of the first two years of my Ph.D. studies attempting to optimize *PXDN* stable overexpression and knockdown in D492 and D492M, respectively, using the CRISPR-dCas9 method without success. Although this caused frustration and a sense of failure, it taught me valuable lessons in both experimental planning and technical optimizations. When I subsequently decided to try lentiviral overexpression of *PXDN* instead, the careful planning and the attention to detail I had learned during my failures in CRISPR resulted in successful overexpression on the first try.

CRISPR-Cas9 is a method of gene editing that presents scientists with the opportunity to alter the genome in a site-specific manner. The CRISPR-Cas9 method is in short based upon the generation of double-strand breaks of the DNA helix by the Cas9 bacterial nuclease, guided to a specific genomic region by guide-RNA (gRNA). The cell's repair mechanism can imperfectly repair the double-strand break, leading to gene knockout (Qi, Larson et al. 2013, Doudna and Charpentier 2014). At SCRU the CRISPRi (CRISPR interference) and CRISPRa (CRISPR activation) systems are routinely used. These two methods utilize deactivated Cas9 (dCas9) that is unable to generate double strand breaks in the DNA. Instead, the dCas9 can

target specific regions of the genome for editing (Gilbert, Horlbeck et al. 2014). In the CRISPRi method, a Krüppel associated box (KRAB) domain has been fused to dCas9 leading to knockdown of the target gene (Gilbert, Larson et al. 2013). Conversely, the CRISPRa method upregulates gene expression through the application of a transcriptional activator that has been fused to Cas9 (Perez-Pinera, Kocak et al. 2013), which in this case was VP64. Conversely, stable gene overexpression with lentiviral transduction is based on the use of recombinant lentiviral particles which contains an element of the human immunodeficiency virus (HIV) to insert novel genetic material into the genome of the host cell line (Naldini, Trono et al. 2016). Lentiviruses are equipped with the ability to use the host cell to produce and package their genetic material, thereby allowing the insertion of a target gene sequence (Elegheert, Behiels et al. 2018). The downside to the lentiviral method compared to CRISPR is the inability to select the insertion point of the novel genetic sequence, which is therefore inserted at random.

The use of CRISPR-dCas9 is dependent upon optimizing virus production and culture conditions of both the virus producing HEK-293T cells as well as the cell line of interest. In my case, I could rely on the experience of both my co-supervisor and fellow Ph.D. students at the lab that had successfully used the method for their projects. However, despite my best efforts, as previously mentioned, the overexpression of *PXDN* was unsuccessful.

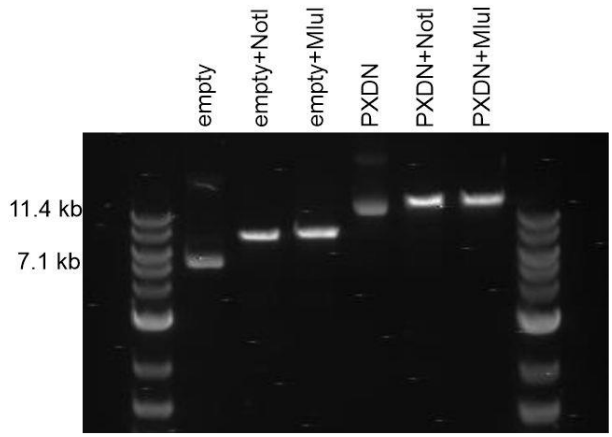


Figure 20. Gel electrophoresis of *PXDN* plasmids. Gel electrophoresis of pLenti-C-Myc-DDK-P2A-Puro empty vector (empty) and pLenti-C-Myc-DDK-P2A-Puro lenti ORF clone of Human peroxidasin homolog (*PXDN*) lentiviral plasmid mini-preps. The plasmids were linearized with NotI and MluI restriction enzymes to determine whether empty and *PXDN* plasmids in the mini-preps were circular and thereby usable. In this case, both empty and *PXDN* plasmids were usable and in supercoiled form, identified due to less friction against the agarose gel and therefore faster run than the linearized forms.

After reviewing my notes from the lab and after having successfully overexpressed *PXDN* using the lentiviral method, I suspect that the root cause of failure in CRISPR lies in the lack of quality control of the plasmids I used as well as mistakes in virus production. With regards to the plasmids, the biggest mistake was not verifying sufficient plasmid quality with gel electrophoresis. Before performing lentiviral transduction, I was able to use the results from gel electrophoresis to carefully choose the best minipreps of plasmids for *PXDN* overexpression as well as the lentiviral packaging plasmids (Figure 20). This helped me realize the difference that can arise in mini-prep quality. During virus production, when transfecting HEK-293T cells with the CRISPR-dCas9-KRAB and CRISPR-dCas9-VP64 plasmids, respectively, I did so by using media supplemented with penicillin-streptomycin and 10% fetal bovine serum (FBS) which is routinely used for HEK-293T cell culture. However, after careful review of available literature, I decided to produce the lentiviral particles without these factors. Although CRISPR was not successful, I was able to utilize my experience of trial and error gained during the process which resulted in the 65-fold overexpression of *PXDN* in D492 cells as discussed in Article 2 and chapter 4.2 of this thesis.

5.2 The search for a reliable PXDN antibody

Only a few commercially available antibodies exist for PXDN and during the initial stages of the project I purchased four of them from the following companies: Abcam, Abnova, Merck-Millipore and Sigma-Aldrich (See Table 5). All antibodies were tested in Western blot according to their product information. However, only the Anti-PXDN/VPO1 antibody from Merck-Millipore had been used in immunostaining of both cells in monolayer and paraffin-embedded breast tissue. Only a few published papers were specified for each antibody as stated on each company's website.

Table 5. A list of commercially available antibodies against PXDN purchased for this project.

Antibody	Manufacturer, Catalog no.
Anti-PXDN/VPO1 – N-terminal	Abcam, ab179663
PXDN polyclonal antibody (A01)	Abnova, H00007837-A01
Anti-PXDN/VPO1 Antibody	Merck-Millipore, AB1675
Monoclonal Anti-PXDN antibody produced in mouse	Sigma- Aldrich, SAB1402973

Initially, when work with PXDN began at our lab, we used PXDN anti-rabbit antibody available by Merck-Millipore to perform immunostaining on paraffin-embedded breast tissue. This staining showed a strong concentration of PXDN in the epithelial cells located within the acini of the breast gland, although immunostaining does not offer the resolution necessary to distinguish LEPs from MEPs. Furthermore, PXDN was also visible in stromal fibroblasts and endothelial cells. Probably the most interesting aspect of these breast samples was positive PXDN staining within the interlobular stroma with concurrent negative staining of collagen IV. This raised the question of whether PXDN might have other functions than strictly crosslinking collagen IV, which is entirely possible, but at the same time, unspecific binding could not be excluded due to the lack of validation of the antibody used.

A different problem arose when immunostaining of PXDN in D492M and D492 was attempted. Here, no staining was visible in either cell line even though PXDN expression in D492M should have led to a positive signal. Later, when attempting to confirm the 70-fold overexpression of PXDN in $D492^{PXDN}$ compared to $D492^{empty}$ observed at the mRNA level via qRT-PCR, immunostaining failed again as there was no visible signal in either cell line. This was even though other markers, such as KRT14 and KRT19, tested at the same time confirmed protein level expression that matched mRNA level expression as seen in qRT-PCR. Attempts were made to treat $D492^{PXDN}$ and $D492^{empty}$ with different concentrations and treatment duration of Brefeldin A, an inhibitor of Golgi secretion, to prevent PXDN secretion into the ECM before immunostaining. However, no signal emerged and the protein-level confirmation of PXDN overexpression in D492 remained without success.

The lack of positive control for PXDN complicated optimization procedures for the antibodies. In the beginning, I contemplated that the reason for insufficient bands and immunostaining signals was due to expression levels in both D492 and D492M being too low for detection. I therefore employed The Human Protein Atlas online database to identify a cell line with high expression of PXDN. RNA sequencing results available in the database revealed that the breast cancer cell line SKBR3 contained levels of PXDN that should be sufficient to detect the protein via Western blot. Although other cell lines expressed higher levels of PXDN I chose SKBR3 due to it being a breast cell line compatible with my experience in cell culture. Furthermore, we already had the cell line available at our lab. Despite that qRT-PCR results indicated that SKBR3 expressed higher levels of PXDN than D492M, no bands were detected on Western blot which indicated that protein amount was not the problem. I then decided that purchase of recombinant PXDN was warranted and started by acquiring recombinant PXDN from Abnova (H00007837-Q01). However, testing on Western blot revealed a band of much smaller size than the 165 kD expected for PXDN. Upon further inspection, I discovered that this particular recombinant protein only contained amino acids 1452-1561, which roughly corresponded to the

position of the von Willebrand factor domain of PXDN. This amino acid sequence contained the target for the Abnova antibody described above but not the targets of the other antibodies.

It remained clear that the main problem regarding protein level detection of PXDN was the questionable reliability of the commercially available antibodies and the lack of positive control. A Hungarian group led by Dr. Miklós Geiszt had published numerous papers on PXDN using a rabbit monoclonal antibody made at their facilities. The published results in their articles of both Western blots and immunostainings were in my opinion more convincing than those published by other authors. For example, the band appearing for PXDN in Western blots was single and of the correct size. Furthermore, immunofluorescent images revealed both cytosolic and membrane expression of the protein (Péterfi, Donkó et al. 2009, Lázár, Péterfi et al. 2015). Dr. Geiszt kindly agreed to give us the antibody to test on $D492^{PXDN}$ and $D492^{empty}$ in Western blot and immunofluorescent staining in both monolayer and 3D. Using whole recombinant PXDN, purchased from OriGene, as positive control I finally got clear bands at 165 kD on Western blot for $D492^{PXDN}$ and $D492^{empty}$ (Figure 2.d in paper #2). Furthermore, immunostaining was also successful in both monolayer and 3D and confirmed the difference in protein level expression of PXDN in $D492^{PXDN}$ and $D492^{empty}$.

By going through the process of optimizing PXDN protein detection I learned that obtaining a reliable positive control should be a priority during antibody testing as judging antibody performance without it is time consuming, costly and likely to fail. It also taught me that reviewing literature and diligent background work are necessary when antibodies are chosen, especially when working with a protein that has not been studied extensively. Furthermore, I learned that reaching out to other scientists that have successfully solved the problem you are facing is well worth the try. This experience proved to be one of two great problems I faced during my Ph.D. project. The other big hurdle was overexpressing PXDN in the D492 cell line.

5.3 Monolayer vs. 3D – Impact on gene expression

The traditional monolayer cell culture is well established and has greatly contributed to biological research in the past. However, this type of culture method has various limitations as cells grown in monolayer on top of coated or uncoated surface lack the microenvironmental input that is present *in vivo* (Souza 2016). This problem is most obvious in primary cells that have been removed from their original scaffolding and placed in monolayer cultures. Under these conditions, primary mammary epithelial cells lose the ability to express organotypic gene signatures which will not be regained unless cells are placed in 3D matrix (Lo, Mori et al. 2012). Like in other fields of organ biology the prevalence of 3D cell culture methods has been on the rise in mammary gland development and cancer research (Sumbal, Budkova et al. 2020). Here, 3D culture methods have proven to be an extremely useful tool as they allow us to mimic the *in vivo* microenvironment by providing cells with the ECM architecture that monolayer cultures lack (Weigelt and Bissell 2008, Weigelt, Ghajar et al. 2014, Simian and Bissell 2017). Traditionally, D492 cells are grown in flasks pre-coated with collagen I and the phenotype under these culture conditions is well documented. I therefore continued culturing $D492^{empty}$ and $D492^{PXDN}$ on collagen I coated surface. However, Matrigel mimics the composition and role of the basement membrane as it contains collagen IV and other basement membrane components (Kleinman and Martin 2005). Therefore, it is possible that collagen IV, along with mechanical stimulus from the 3D environment and perhaps the aid of other ECM factors, increased the expression of *PXDN*. Interestingly, I found that *PXDN* overexpression did not significantly alter the expression of *COL4A1* and *COL4A5* genes (Supplemental figure 1 in article II).

5.4 3D culture: Cells embedded in vs. on top of Matrigel

Having successfully overexpressed *PXDN* in D492, a part of the original characterization of the cell line was to perform 3D embedded assay in Matrigel, which members of the Stem Cell Research Unit have been performing on D492 cells and its daughter cell lines for the past two decades

with good results (Sigurdsson, Hilmarsdottir et al. 2011, Hilmarsdóttir, Briem et al. 2015, Briem, Budkova et al. 2019, Briem, Ingthorsson et al. 2019). Although relatively simple in setup and practice, the method of culturing D492 cells embedded within Matrigel is expensive and time consuming as branching is traditionally observed for up to 21 days before analysis is performed. Furthermore, the risk of failure in the form of ruptured or detached Matrigel is high, as the gel is very delicate and can easily be damaged if not handled with care, for example during change of media which I experienced quite often. Therefore, an alternative and less expensive method would be feasible. In this regard, we have found that culturing D492 cells on top of Matrigel gives results compatible with the 3D embedded assay. This requires up to three times less Matrigel, reducing cost significantly. Furthermore, branching morphogenesis can be observed much sooner, setting the experimental endpoint at 8 days instead of 21 days which reduces workload and cost. When I originally cultured D492^{PXDN} and D492^{empty} using the 3D embedded method I observed that D492^{PXDN} had lost the ability to branch while D492^{empty} retained similar colony morphology seen in D492 (see further in chapter 4.2). Subsequently, I observed that colony morphology remained consistent in both D492^{PXDN} and D492^{empty} when the cells were cultured on top of Matrigel as compared to embedded in Matrigel. Therefore, I decided to utilize the 3D on top method in all subsequent experiments with D492^{PXDN} where 3D culture was needed.

6 Summary and concluding remarks

Branching morphogenesis is a fundamental process in the development of the mammary gland, resulting in the formation of epithelial ducts and functional units with enough surface area to initiate and sustain efficient lactation after pregnancy. Sadly, the functional epithelium within the TDLUs is often the place of origin of breast cancer which, as was discussed in the introduction to this thesis, manipulates normal developmental pathways for its own benefit. Epithelial plasticity, and more specifically EMT and MET, is one of the better-known examples of malignant turn in a previously benign process. During my Ph.D. project I have applied the D492 and D492M cell models, in monolayer and 3D culture, to explore the regulatory role of ncRNAs miR-203a and MEG3, and the function of *PXDN* in branching morphogenesis, epithelial plasticity and EMT that have formed the core of my project.

The first article highlighted the role of miRNAs in breast morphogenesis and EMT. By applying small RNA-sequencing we were able to identify the most differentially expressed miRNAs in D492M compared to D492 in 3D culture and found that miR-203a was one of the most downregulated in EMT. We highlighted the dynamic nature of branching morphogenesis in the D492 cell line by showing temporal changes in miRNA expression at three different time points in 3D culture, where miR-203a was significantly increased over time. By overexpressing miR-203a in the D492M cell line we were able to induce partial MET and by performing RNA-sequencing we identified *PXDN* as a possible target of miR-203a. Finally, we confirmed binding between miR-203a to a target site on *PXDN* 3'UTR by applying Luciferase Assay.

In the second manuscript, which was the focal point of my Ph.D. project I showed via immunohistochemistry of breast tissue and qRT-PCR in primary breast cells that *PXDN* is expressed in breast epithelium, fibroblasts and endothelial cells. With the purpose of investigating the role of *PXDN* in the

epithelium, I applied lentiviral transduction to overexpress *PXDN* in D492 cells. *PXDN* overexpression completely inhibited branching morphogenesis in 3D culture while inducing basal phenotype, partially mediated through p63, and inhibited plasticity in D492 cells that lost the ability to form cells with luminal or myoepithelial characteristics. Furthermore, $D492^{PXDN}$ had increased sensitivity to chemically induced apoptosis as well as increased capability for migration and invasion highlighted by significantly decreased wound closure time. Finally, by performing RNA-sequencing analysis I found that *PXDN* upregulation enriches genes and pathways involved in epithelial development and differentiation while causing reduced expression of EMT genes and inhibiting the EMT pathway which is necessary for branching morphogenesis.

The third article focused on the role of lncRNA MEG3 and ncRNAs on the DLK1-DIO3 locus on phenotype in the D492 and D492M cell lines. We showed that *MEG3* and multiple miRNAs from the DLK1-DIO3 locus are upregulated in EMT in the D492 cell model. We also showed that expression of *MEG3* and four different miRNAs from the locus are associated with the stromal compartment in the breast and found that increased expression of *MEG3* is associated with worse survival in breast cancer. With overexpression and knockdown of *MEG3* in D492 and D492M cells, respectively, miRNAs from the DLK1-DIO3 locus showed concomitant changes in expression as well. RNA-sequencing data revealed that knockdown of *MEG3* in D492M cells decreased the expression of mesenchymal and luminal genes and reduced enrichment of the EMT pathway. Functional assays showed that *MEG3* induced partial EMT when overexpressed in D492 cells, and induced stem cell properties.

Finally, I presented unpublished data which suggests that *PXDN* is relevant in breast cancer as high expression leads to significantly decreased survival in women with HER2-positive breast cancer, despite also being present in ER-positive and triple negative breast tumors as well, as I showed with immunohistochemistry. Furthermore, I presented evidence that *PXDN* increased stemness in D492 cells as indicated by a significant increase in $CD24^{-}/CD44^{+}$ population and decreased activity of ALDH in $D492^{PXDN}$ compared to $D492^{empty}$.

6.1 Future perspective

Our understanding of normal development and developmental pathways has led to great advances in medicine. Herein lies the concept that to battle the enemy one must first know how it operates. This is especially true in cancer where accumulating mutations distort and manipulate normal function, often for the benefit of the tumor. Breast cancer is one of the most prevalent cancer worldwide, affecting women of all ages. Therefore, it is of the utmost importance to identify novel biomarkers and therapeutic targets to aid in the battle against this heterogeneous disease.

The ncRNAs are gaining interest in both development and disease. In particular, the miRNAs, like miR-203a, are involved in a myriad of cellular functions as each of them can target multiple genes. As the data presented in this thesis showed that miRNAs are differentially expressed over time during branching in D492 it would be interesting to see whether this applies to branching morphogenesis *in vivo* as well. Following the branching process in a mouse model and performing RNA sequencing on different time points would provide valuable insight into the temporal regulatory function of miRNAs in mammary development. Furthermore, as EMT aids carcinoma cells in tissue invasion has been linked to increased therapy resistance identifying novel therapeutic targets is integral in the battle against the disease. It is possible that miRNAs that are downregulated during EMT, like miR-203a, could either inhibit EMT from occurring or reverse the process. We showed that miR-203a induced partial MET in D492M, indicating that this could be possible. However, the drawback is that this might be easier said than done *in vivo*, both due to technical limitations and due to unknown effect on other tissues with systemic therapy. It might also be counter-productive to induce MET in cancer cells if metastasis has already occurred, indicating the presence of cancer cells in the bloodstream and lymphatic system, where MET might facilitate the colonization of distant sites, thereby creating new metastatic lesions.

Much is still left to explore with regards to the lncRNAs, as the precise mechanism behind their function is still relatively unknown. We showed that

high *MEG3* tumor expression significantly reduced survival in breast cancer patients. When held in context with the association to the affiliation we presented between *MEG3* and mesenchymal gene signature and the stromal compartment it would be interesting to see whether *MEG3* could be used as an indicator of an increased mesenchymal component or EMT in breast tumors. Furthermore, it would be interesting to see whether it might be possible to inhibit *MEG3* expression in breast tumors, possibly hindering EMT, invasion and metastasis.

Much is still left unknown about the role of *PXDN* in the mammary gland although the results discussed in this thesis indicate that further research is valid. The fact that PXDN is expressed in both layers of the mammary epithelium, within the TDLUs and in ducts indicates that PXDN must play a role in epithelial cell function. On top of this PXDN is highly conserved between species, which together with the data from the HPA presented in this thesis might lead to the conclusion that PXDN must play some fundamental role, whether it would be linked solely to basement membrane stabilization or not. However, Kim et al. (2019) generated *PXDN* knockout mice that showed no significant phenotypic defects other than anopthalmia. Nonetheless, this does not automatically deem PXDN irrelevant in all tissue except the eye as information on the histological appearance of different organs, organ function, disease development and overall lifespan in these knockout mice is still lacking, presenting an opportunity for further research.

As I presented evidence for *PXDN* expression within both layers of the mammary epithelium it is interesting that D492 cells express very low levels of *PXDN*. In the future it would be interesting to explore the existence of *PXDN*-/low subpopulation within the breast epithelium, and whether the phenotype of this subpopulation coincides with plasticity and progenitor cell features. It also raises the question of whether it is indeed not *increased* expression that is more relevant in the mammary gland *in vivo*, but the *loss* of expression. With consideration to the data presented in this thesis, it is an interesting thought whether programmed downregulation of *PXDN* within the epithelium increases cell plasticity and alleviates EMT inhibition, thereby

leading to branching. Therefore, identifying a normal mammary epithelial cell line with high *PXDN* expression and performing knockdown studies might provide more pieces of the puzzle.

As discussed in the results of this thesis the analysis of data available in online databases indicates that high expression of *PXDN* significantly decreases survival in women with HER2-positive/ER-negative breast cancer. The data also shows that in the previously mentioned subgroup of breast cancers women are at a higher risk of having larger tumors and lymph node metastasis and at diagnosis if their primary tumors express high levels of *PXDN*. This may indicate increased aggressiveness is the primary culprit of decreased survival although chemoresistance cannot be excluded. The mechanism behind this is largely unexplored and there is no clear-cut explanation of how a collagen IV-crosslinker can have such an impact on cancer progression. With results from the D492 cell line in mind, it might be possible that high *PXDN* expression induces basal phenotype in breast cancer cells, leading to increased heterogeneity and dedifferentiation which is often associated with increased therapy resistance and worse prognosis. However, why this would only happen in HER2-positive breast cancer remains a mystery.

During my Ph.D. project I have gained valuable experience and insight into the scientific process. I was very fortunate to have the opportunity to apply the cell model D492 which provides a unique *in vitro* opportunity to study branching morphogenesis, EMT and other processes involved in mammary gland development.

References

Adams, J. C. and F. M. Watt (1989). "Fibronectin inhibits the terminal differentiation of human keratinocytes." Nature **340**(6231): 307-309.

Ades, F., D. Zardavas, I. Bozovic-Spasojevic, L. Pugliano, D. Fumagalli, E. d. Azambuja, G. Viale, C. Sotiriou and M. Piccart (2014). "Luminal B Breast Cancer: Molecular Characterization, Clinical Management, and Future Perspectives." Journal of Clinical Oncology **32**(25): 2794-2803.

Affolter, M., R. Zeller and E. Caussinus (2009). "Tissue remodelling through branching morphogenesis." Nat Rev Mol Cell Biol **10**(12): 831-842.

Agarwal, V., G. W. Bell, J. W. Nam and D. P. Bartel (2015). "Predicting effective microRNA target sites in mammalian mRNAs." Elife **4**.

Aiello, N. M., R. Maddipati, R. J. Norgard, D. Balli, J. Li, S. Yuan, T. Yamazoe, T. Black, A. Sahmoud, E. E. Furth, D. Bar-Sagi and B. Z. Stanger (2018). "EMT Subtype Influences Epithelial Plasticity and Mode of Cell Migration." Developmental cell **45**(6): 681-695.e684.

Akbar, M. W., M. Isbilen, N. Belder, S. D. Canli, B. Kucukkaraduman, C. Turk, O. Sahin and A. O. Gure (2020). "A Stemness and EMT Based Gene Expression Signature Identifies Phenotypic Plasticity and is A Predictive but Not Prognostic Biomarker for Breast Cancer." J Cancer **11**(4): 949-961.

Al-Hajj, M., M. S. Wicha, A. Benito-Hernandez, S. J. Morrison and M. F. Clarke (2003). "Prospective identification of tumorigenic breast cancer cells." Proc Natl Acad Sci U S A **100**(7): 3983-3988.

Albuschies, J. and V. Vogel (2013). "The role of filopodia in the recognition of nanotopographies." Sci Rep **3**: 1658.

Ali Syeda, Z., S. S. S. Langden, C. Munkhzul, M. Lee and S. J. Song (2020). "Regulatory Mechanism of MicroRNA Expression in Cancer." Int J Mol Sci **21**(5).

Ameres, S. L. and P. D. Zamore (2013). "Diversifying microRNA sequence and function." Nat Rev Mol Cell Biol **14**(8): 475-488.

Barnett, P., R. S. Arnold, R. Mezencev, L. W. K. Chung, M. Zayzafoon and V. Odero-Marah (2011). "Snail-mediated regulation of reactive oxygen species in ARCaP human prostate cancer cells." Biochemical and Biophysical Research Communications **404**(1): 34-39.

Barrallo-Gimeno, A. and M. A. Nieto (2005). "The Snail genes as inducers of cell movement and survival: implications in development and cancer." Development **132**(14): 3151.

Bartel, D. P. (2004). "MicroRNAs: Genomics, Biogenesis, Mechanism, and Function." Cell **116**(2): 281-297.

Baulina, N., G. Osmak, I. Kiselev, E. Popova, A. Boyko, O. Kulakova and O. Favorova (2019). "MiRNAs from DLK1-DIO3 Imprinted Locus at 14q32 are Associated with Multiple Sclerosis: Gender-Specific Expression and Regulation of Receptor Tyrosine Kinases Signaling." Cells **8**(2): 133.

Benetatos, L., G. Vartholomatos and E. Hatzimichael (2014). "DLK1-DIO3 imprinted cluster in induced pluripotency: landscape in the mist." Cell Mol Life Sci **71**(22): 4421-4430.

Behave, G., C. F. Cummings, R. M. Vanacore, C. Kumagai-Cresse, I. A. Ero-Tolliver, M. Rafi, J. S. Kang, V. Pedchenko, L. I. Fessler, J. H. Fessler and B. G. Hudson (2012). "Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis." Nat Chem Biol **8**(9): 784-790.

Blighe, K., Rana, S., Lewis, M. (2020). "EnhancedVolcano: Publication-ready volcano plots with enhanced coloring and lebeling." R package version 1.8.0, https://github.com/kevinblighe/EnhancedVolcano.

Boecker, W., L. van Horn, G. Stenman, C. Stürken, U. Schumacher, T. Loening, L. Liesenfeld, E. Korsching, D. Gläser, K. Tiemann and I. Buchwalow (2018). "Spatially correlated phenotyping reveals K5-positive luminal progenitor cells and p63-K5/14-positive stem cell-like cells in human breast epithelium." Lab Invest **98**(8): 1065-1075.

Bonnans, C., J. Chou and Z. Werb (2014). "Remodelling the extracellular matrix in development and disease." Nat Rev Mol Cell Biol **15**(12): 786-801.

Bracken, C. P., P. A. Gregory, N. Kolesnikoff, A. G. Bert, J. Wang, M. F. Shannon and G. J. Goodall (2008). "A double-negative feedback loop between ZEB1-SIP1 and the microRNA-200 family regulates epithelial-mesenchymal transition." Cancer Res **68**(19): 7846-7854.

Bray, F., J. Ferlay, M. Laversanne, D. H. Brewster, C. Gombe Mbalawa, B. Kohler, M. Piñeros, E. Steliarova-Foucher, R. Swaminathan, S. Antoni, I. Soerjomataram and D. Forman (2015). "Cancer Incidence in Five Continents: Inclusion criteria, highlights from Volume X and the global status of cancer registration." Int J Cancer **137**(9): 2060-2071.

Bray, F., J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre and A. Jemal (2018). "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries." CA Cancer J Clin **68**(6): 394-424.

Briem, E., Z. Budkova, A. K. Sigurdardottir, B. Hilmarsdottir, J. Kricker, W. Timp, M. K. Magnusson, G. A. Traustadottir and T. Gudjonsson (2019). "MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin." Mech Dev **155**: 34-47.

Briem, E., S. Ingthorsson, G. A. Traustadottir, B. Hilmarsdottir and T. Gudjonsson (2019). "Application of the D492 Cell Lines to Explore Breast Morphogenesis, EMT and Cancer Progression in 3D Culture." J Mammary Gland Biol Neoplasia **24**(2): 139-147.

Brown, K. L., C. F. Cummings, R. M. Vanacore and B. G. Hudson (2017). "Building collagen IV smart scaffolds on the outside of cells." Protein Sci **26**(11): 2151-2161.

Budkova, Z., A. K. Sigurdardottir, E. Briem, J. T. Bergthorsson, S. Sigurdsson, M. K. Magnusson, G. A. Traustadottir, T. Gudjonsson and B. Hilmarsdottir (2020). "Expression of ncRNAs on the DLK1-DIO3 Locus Is Associated With Basal and Mesenchymal Phenotype in Breast Epithelial Progenitor Cells." Front Cell Dev Biol **8**: 461.

Burk, U., J. Schubert, U. Wellner, O. Schmalhofer, E. Vincan, S. Spaderna and T. Brabletz (2008). "A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells." EMBO Rep **9**(6): 582-589.

Calin, G. A., C. G. Liu, M. Ferracin, T. Hyslop, R. Spizzo, C. Sevignani, M. Fabbri, A. Cimmino, E. J. Lee, S. E. Wojcik, M. Shimizu, E. Tili, S. Rossi, C. Taccioli, F. Pichiorri, X. Liu, S. Zupo, V. Herlea, L. Gramantieri, G. Lanza, H. Alder, L. Rassenti, S. Volinia, T. D. Schmittgen, T. J. Kipps, M. Negrini and C. M. Croce (2007). "Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas." Cancer Cell **12**(3): 215-229.

Casey, M. C., K. J. Sweeney, J. A. Brown and M. J. Kerin (2016). "Exploring circulating micro-RNA in the neoadjuvant treatment of breast cancer." Int J Cancer **139**(1): 12-22.

Chaffer, C. L., B. P. San Juan, E. Lim and R. A. Weinberg (2016). "EMT, cell plasticity and metastasis." Cancer Metastasis Rev **35**(4): 645-654.

Chao, C.-H., C.-C. Chang, M.-J. Wu, H.-W. Ko, D. Wang, M.-C. Hung, J.-Y. Yang and C.-J. Chang (2014). "MicroRNA-205 signaling regulates mammary stem cell fate and tumorigenesis." The Journal of clinical investigation **124**(7): 3093-3106.

Charafe-Jauffret, E., F. Monville, C. Ginestier, G. Dontu, D. Birnbaum and M. S. Wicha (2008). "Cancer stem cells in breast: current opinion and future challenges." Pathobiology **75**(2): 75-84.

Cheang, M. C., M. Martin, T. O. Nielsen, A. Prat, D. Voduc, A. Rodriguez-Lescure, A. Ruiz, S. Chia, L. Shepherd, M. Ruiz-Borrego, L. Calvo, E. Alba, E. Carrasco, R. Caballero, D. Tu, K. I. Pritchard, M. N. Levine, V. H. Bramwell, J. Parker, P. S. Bernard, M. J. Ellis, C. M. Perou, A. Di Leo and L. A. Carey (2015). "Defining breast cancer intrinsic subtypes by quantitative receptor expression." Oncologist **20**(5): 474-482.

Chen, X. and J. Qu (2018). "Long non-coding RNA MEG3 suppresses survival, migration, and invasion of cervical cancer." Onco Targets Ther **11**: 4999-5007.

Cheng, G., J. C. Salerno, Z. Cao, P. J. Pagano and J. D. Lambeth (2008). "Identification and characterization of VPO1, a new animal heme-containing peroxidase." Free Radic Biol Med **45**(12): 1682-1694.

Chunharojrith, P., Y. Nakayama, X. Jiang, R. E. Kery, J. Ma, C. S. De La Hoz Ulloa, X. Zhang, Y. Zhou and A. Klibanski (2015). "Tumor suppression by MEG3 lncRNA in a human pituitary tumor derived cell line." Mol Cell Endocrinol **416**: 27-35.

Cleator, S., W. Heller and R. C. Coombes (2007). "Triple-negative breast cancer: therapeutic options." Lancet Oncol **8**(3): 235-244.

Correia de Sousa, M., M. Gjorgjieva, D. Dolicka, C. Sobolewski and M. Foti (2019). "Deciphering miRNAs' Action through miRNA Editing." International journal of molecular sciences **20**(24): 6249.

Croker, A. K., D. Goodale, J. Chu, C. Postenka, B. D. Hedley, D. A. Hess and A. L. Allan (2009). "High aldehyde dehydrogenase and expression of cancer stem cell markers selects for breast cancer cells with enhanced malignant and metastatic ability." J Cell Mol Med **13**(8b): 2236-2252.

Dalerba, P., R. W. Cho and M. F. Clarke (2007). "Cancer stem cells: models and concepts." Annu Rev Med **58**: 267-284.

Daniunaite, K., M. Dubikaityte, P. Gibas, A. Bakavicius, J. Rimantas Lazutka, A. Ulys, F. Jankevicius and S. Jarmalaite (2017). "Clinical significance of

miRNA host gene promoter methylation in prostate cancer." Hum Mol Genet **26**(13): 2451-2461.

DeCastro, A. J., K. A. Dunphy, J. Hutchinson, A. L. Balboni, P. Cherukuri, D. J. Jerry and J. DiRenzo (2013). "MiR203 mediates subversion of stem cell properties during mammary epithelial differentiation via repression of ΔNP63α and promotes mesenchymal-to-epithelial transition." Cell Death Dis **4**(2): e514.

Deng, R., F. Y. Fan, H. Yi, F. Liu, G. C. He, H. P. Sun and Y. Su (2018). "MEG3 affects the progression and chemoresistance of T-cell lymphoblastic lymphoma by suppressing epithelial-mesenchymal transition via the PI3K/mTOR pathway." J Cell Biochem.

Deome, K. B., L. J. Faulkin, Jr., H. A. Bern and P. B. Blair (1959). "Development of mammary tumors from hyperplastic alveolar nodules transplanted into gland-free mammary fat pads of female C3H mice." Cancer Res **19**(5): 515-520.

Di Gesualdo, F., S. Capaccioli and M. Lulli (2014). "A pathophysiological view of the long non-coding RNA world." Oncotarget **5**(22): 10976-10996.

Dill, T. L. and F. J. Naya (2018). "A Hearty Dose of Noncoding RNAs: The Imprinted DLK1-DIO3 Locus in Cardiac Development and Disease." J Cardiovasc Dev Dis **5**(3).

Dillon, C., B. Spencer-Dene and C. Dickson (2004). "A crucial role for fibroblast growth factor signaling in embryonic mammary gland development." J Mammary Gland Biol Neoplasia **9**(2): 207-215.

Ding, X., S. I. Park, L. K. McCauley and C.-Y. Wang (2013). "Signaling between transforming growth factor β (TGF-β) and transcription factor SNAI2 represses expression of microRNA miR-203 to promote epithelial-mesenchymal transition and tumor metastasis." The Journal of biological chemistry **288**(15): 10241-10253.

Dittmer, J. (2018). "Breast cancer stem cells: Features, key drivers and treatment options." Seminars in Cancer Biology **53**: 59-74.

Dongre, A. and R. A. Weinberg (2019). "New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer." Nat Rev Mol Cell Biol **20**(2): 69-84.

Dontu, G., W. M. Abdallah, J. M. Foley, K. W. Jackson, M. F. Clarke, M. J. Kawamura and M. S. Wicha (2003). "In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells." Genes Dev **17**(10): 1253-1270.

Doudna, J. A. and E. Charpentier (2014). "The new frontier of genome engineering with CRISPR-Cas9." Science **346**(6213): 1258096.

Eades, G., Y.-S. Zhang, Q.-L. Li, J.-X. Xia, Y. Yao and Q. Zhou (2014). "Long non-coding RNAs in stem cells and cancer." World journal of clinical oncology **5**(2): 134-141.

Elegheert, J., E. Behiels, B. Bishop, S. Scott, R. E. Woolley, S. C. Griffiths, E. F. X. Byrne, V. T. Chang, D. I. Stuart, E. Y. Jones, C. Siebold and A. R. Aricescu (2018). "Lentiviral transduction of mammalian cells for fast, scalable and high-level production of soluble and membrane proteins." Nat Protoc **13**(12): 2991-3017.

Ero-Tolliver, I. A., B. G. Hudson and G. Bhave (2015). "The Ancient Immunoglobulin Domains of Peroxidasin Are Required to Form Sulfilimine Cross-links in Collagen IV." J Biol Chem **290**(35): 21741-21748.

Eswarakumar, V. P., I. Lax and J. Schlessinger (2005). "Cellular signaling by fibroblast growth factor receptors." Cytokine Growth Factor Rev **16**(2): 139-149.

Ewald, A. J., A. Brenot, M. Duong, B. S. Chan and Z. Werb (2008). "Collective epithelial migration and cell rearrangements drive mammary branching morphogenesis." Dev Cell **14**(4): 570-581.

Ewald, A. J., R. J. Huebner, H. Palsdottir, J. K. Lee, M. J. Perez, D. M. Jorgens, A. N. Tauscher, K. J. Cheung, Z. Werb and M. Auer (2012). "Mammary collective cell migration involves transient loss of epithelial features and individual cell migration within the epithelium." J Cell Sci **125**(Pt 11): 2638-2654.

Fata, J. E., Z. Werb and M. J. Bissell (2004). "Regulation of mammary gland branching morphogenesis by the extracellular matrix and its remodeling enzymes." Breast Cancer Res **6**(1): 1-11.

Fatica, A. and I. Bozzoni (2014). "Long non-coding RNAs: new players in cell differentiation and development." Nat Rev Genet **15**(1): 7-21.

Friedman, R. C., K. K. Farh, C. B. Burge and D. P. Bartel (2009). "Most mammalian mRNAs are conserved targets of microRNAs." Genome Res **19**(1): 92-105.

Georgakopoulos-Soares, I., D. V. Chartoumpekis, V. Kyriazopoulou and A. Zaravinos (2020). "EMT Factors and Metabolic Pathways in Cancer." Frontiers in oncology **10**: 499-499.

Ghafouri-Fard, S. and M. Taheri (2019). "Maternally expressed gene 3 (MEG3): A tumor suppressor long non coding RNA." Biomed Pharmacother **118**: 109129.

Gilbert, Luke A., Max A. Horlbeck, B. Adamson, Jacqueline E. Villalta, Y. Chen, Evan H. Whitehead, C. Guimaraes, B. Panning, Hidde L. Ploegh, Michael C. Bassik, Lei S. Qi, M. Kampmann and Jonathan S. Weissman (2014). "Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation." Cell **159**(3): 647-661.

Gilbert, Luke A., Matthew H. Larson, L. Morsut, Z. Liu, Gloria A. Brar, Sandra E. Torres, N. Stern-Ginossar, O. Brandman, Evan H. Whitehead, Jennifer A. Doudna, Wendell A. Lim, Jonathan S. Weissman and Lei S. Qi (2013). "CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes." Cell **154**(2): 442-451.

Ginestier, C., M. H. Hur, E. Charafe-Jauffret, F. Monville, J. Dutcher, M. Brown, J. Jacquemier, P. Viens, C. G. Kleer, S. Liu, A. Schott, D. Hayes, D. Birnbaum, M. S. Wicha and G. Dontu (2007). "ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome." Cell Stem Cell **1**(5): 555-567.

Gloss, B. S. and M. E. Dinger (2016). "The specificity of long noncoding RNA expression." Biochim Biophys Acta **1859**(1): 16-22.

Goff, L. A. and J. L. Rinn (2015). "Linking RNA biology to lncRNAs." Genome Res **25**(10): 1456-1465.

Gregory, P. A., A. G. Bert, E. L. Paterson, S. C. Barry, A. Tsykin, G. Farshid, M. A. Vadas, Y. Khew-Goodall and G. J. Goodall (2008). "The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1." Nat Cell Biol **10**(5): 593-601.

Gu, S., L. Jin, F. Zhang, P. Sarnow and M. A. Kay (2009). "Biological basis for restriction of microRNA targets to the 3' untranslated region in mammalian mRNAs." Nat Struct Mol Biol **16**(2): 144-150.

Gu, Z., R. Eils and M. Schlesner (2016). "Complex heatmaps reveal patterns and correlations in multidimensional genomic data." Bioinformatics **32**(18): 2847-2849.

Gudjonsson, T., M. C. Adriance, M. D. Sternlicht, O. W. Petersen and M. J. Bissell (2005). "Myoepithelial cells: their origin and function in breast morphogenesis and neoplasia." J Mammary Gland Biol Neoplasia **10**(3): 261-272.

Gudjonsson, T., L. Rønnov-Jessen, R. Villadsen, F. Rank, M. J. Bissell and O. W. Petersen (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**(Pt 1): 39-50.

Gudjonsson, T., R. Villadsen, H. L. Nielsen, L. Rønnov-Jessen, M. J. Bissell and O. W. Petersen (2002). "Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties." Genes Dev **16**(6): 693-706.

Guo, W., Z. Keckesova, J. L. Donaher, T. Shibue, V. Tischler, F. Reinhardt, S. Itzkovitz, A. Noske, U. Zürrer-Härdi, G. Bell, W. L. Tam, S. A. Mani, A. van Oudenaarden and R. A. Weinberg (2012). "Slug and Sox9 cooperatively determine the mammary stem cell state." Cell **148**(5): 1015-1028.

Gupta, R. A., N. Shah, K. C. Wang, J. Kim, H. M. Horlings, D. J. Wong, M. C. Tsai, T. Hung, P. Argani, J. L. Rinn, Y. Wang, P. Brzoska, B. Kong, R. Li, R. B. West, M. J. van de Vijver, S. Sukumar and H. Y. Chang (2010). "Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis." Nature **464**(7291): 1071-1076.

Guttman, M., I. Amit, M. Garber, C. French, M. F. Lin, D. Feldser, M. Huarte, O. Zuk, B. W. Carey, J. P. Cassady, M. N. Cabili, R. Jaenisch, T. S. Mikkelsen, T. Jacks, N. Hacohen, B. E. Bernstein, M. Kellis, A. Regev, J. L. Rinn and E. S. Lander (2009). "Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals." Nature **458**(7235): 223-227.

Halfter, W., P. Oertle, C. A. Monnier, L. Camenzind, M. Reyes-Lua, H. Hu, J. Candiello, A. Labilloy, M. Balasubramani, P. B. Henrich and M. Plodinec (2015). "New concepts in basement membrane biology." Febs j **282**(23): 4466-4479.

Hanahan, D. and Robert A. Weinberg (2011). "Hallmarks of Cancer: The Next Generation." Cell **144**(5): 646-674.

Hansen, R. K. and M. J. Bissell (2000). "Tissue architecture and breast cancer: the role of extracellular matrix and steroid hormones." Endocr Relat Cancer **7**(2): 95-113.

Harbeck, N., F. Penault-Llorca, J. Cortes, M. Gnant, N. Houssami, P. Poortmans, K. Ruddy, J. Tsang and F. Cardoso (2019). "Breast cancer." Nat Rev Dis Primers **5**(1): 66.

Harrington, C. T., E. I. Lin, M. T. Olson and J. R. Eshleman (2013). "Fundamentals of pyrosequencing." Arch Pathol Lab Med **137**(9): 1296-1303.

He, Y., Y. Luo, B. Liang, L. Ye, G. Lu and W. He (2017). "Potential applications of MEG3 in cancer diagnosis and prognosis." Oncotarget **8**(42): 73282-73295.

Hilmarsdottir, B., E. Briem, J. T. Bergthorsson, M. K. Magnusson and T. Gudjonsson (2014). "Functional Role of the microRNA-200 Family in Breast Morphogenesis and Neoplasia." Genes (Basel) **5**(3): 804-820.

Hilmarsdóttir, B., E. Briem, V. Sigurdsson, S. R. Franzdóttir, M. Ringnér, A. J. Arason, J. T. Bergthorsson, M. K. Magnusson and T. Gudjonsson (2015). "MicroRNA-200c-141 and ΔNp63 are required for breast epithelial differentiation and branching morphogenesis." Developmental Biology **403**(2): 150-161.

Huarte, M., M. Guttman, D. Feldser, M. Garber, M. J. Koziol, D. Kenzelmann-Broz, A. M. Khalil, O. Zuk, I. Amit, M. Rabani, L. D. Attardi, A. Regev, E. S. Lander, T. Jacks and J. L. Rinn (2010). "A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response." Cell **142**(3): 409-419.

Huebner, R. J. and A. J. Ewald (2014). "Cellular foundations of mammary tubulogenesis." Semin Cell Dev Biol **31**: 124-131.

Ilina, O. and P. Friedl (2009). "Mechanisms of collective cell migration at a glance." J Cell Sci **122**(Pt 18): 3203-3208.

Inman, J. L., C. Robertson, J. D. Mott and M. J. Bissell (2015). "Mammary gland development: cell fate specification, stem cells and the microenvironment." Development **142**(6): 1028-1042.

Jacquemet, G., H. Hamidi and J. Ivaska (2015). "Filopodia in cell adhesion, 3D migration and cancer cell invasion." Curr Opin Cell Biol **36**: 23-31.

Jayachandran, A., P. Prithviraj, P. H. Lo, M. Walkiewicz, M. Anaka, B. L. Woods, B. Tan, A. Behren, J. Cebon and S. J. McKeown (2016). "Identifying and targeting determinants of melanoma cellular invasion." Oncotarget **7**(27): 41186-41202.

Jones-Paris, C. R., S. Paria, T. Berg, J. Saus, G. Bhave, B. C. Paria and B. G. Hudson (2017). "Embryo implantation triggers dynamic spatiotemporal expression of the basement membrane toolkit during uterine reprogramming." Matrix Biol **57-58**: 347-365.

Jørgensen, C. L. T., C. Forsare, P. O. Bendahl, A. K. Falck, M. Fernö, K. Lövgren, K. Aaltonen and L. Rydén (2020). "Expression of epithelial-mesenchymal transition-related markers and phenotypes during breast cancer progression." Breast Cancer Res Treat **181**(2): 369-381.

Kaneko, S., R. Bonasio, R. Saldaña-Meyer, T. Yoshida, J. Son, K. Nishino, A. Umezawa and D. Reinberg (2014). "Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin." Mol Cell **53**(2): 290-300.

Khan, K., A. Rudkin, D. A. Parry, K. P. Burdon, M. McKibbin, C. V. Logan, Z. I. Abdelhamed, J. S. Muecke, N. Fernandez-Fuentes, K. J. Laurie, M. Shires, R. Fogarty, I. M. Carr, J. A. Poulter, J. E. Morgan, M. D. Mohamed, H. Jafri, Y. Raashid, N. Meng, H. Piseth, C. Toomes, R. J. Casson, G. R. Taylor, M. Hammerton, E. Sheridan, C. A. Johnson, C. F. Inglehearn, J. E. Craig and M. Ali (2011). "Homozygous mutations in PXDN cause congenital cataract, corneal opacity, and developmental glaucoma." Am J Hum Genet **89**(3): 464-473.

Khoshnoodi, J., V. Pedchenko and B. G. Hudson (2008). "Mammalian collagen IV." Microsc Res Tech **71**(5): 357-370.

Kim, H. K., K. A. Ham, S. W. Lee, H. S. Choi, H. S. Kim, H. K. Kim, H. S. Shin, K. Y. Seo, Y. Cho, K. T. Nam, I. B. Kim and Y. A. Joe (2019). "Biallelic Deletion of Pxdn in Mice Leads to Anophthalmia and Severe Eye Malformation." Int J Mol Sci **20**(24).

Kleinman, H. K. and G. R. Martin (2005). "Matrigel: basement membrane matrix with biological activity." Semin Cancer Biol **15**(5): 378-386.

Kordon, E. C. and G. H. Smith (1998). "An entire functional mammary gland may comprise the progeny from a single cell." Development **125**(10): 1921-1930.

Korpal, M., E. S. Lee, G. Hu and Y. Kang (2008). "The miR-200 family inhibits epithelial-mesenchymal transition and cancer cell migration by direct targeting of E-cadherin transcriptional repressors ZEB1 and ZEB2." J Biol Chem **283**(22): 14910-14914.

Koster, M. I. and D. R. Roop (2004). "The role of p63 in development and differentiation of the epidermis." J Dermatol Sci **34**(1): 3-9.

Kotiyal, S. and S. Bhattacharya (2014). "Breast cancer stem cells, EMT and therapeutic targets." Biochem Biophys Res Commun **453**(1): 112-116.

Kröger, C., A. Afeyan, J. Mraz, E. N. Eaton, F. Reinhardt, Y. L. Khodor, P. Thiru, B. Bierie, X. Ye, C. B. Burge and R. A. Weinberg (2019).

"Acquisition of a hybrid E/M state is essential for tumorigenicity of basal breast cancer cells." Proc Natl Acad Sci U S A **116**(15): 7353-7362.

Kuşoğlu, A. and Ç. Biray Avcı (2019). "Cancer stem cells: A brief review of the current status." Gene **681**: 80-85.

Lázár, E., Z. Péterfi, G. Sirokmány, H. A. Kovács, E. Klement, K. F. Medzihradszky and M. Geiszt (2015). "Structure-function analysis of peroxidasin provides insight into the mechanism of collagen IV crosslinking." Free Radic Biol Med **83**: 273-282.

Lee, R. C., R. L. Feinbaum and V. Ambros (1993). "The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14." Cell **75**(5): 843-854.

Lee, Y., C. Ahn, J. Han, H. Choi, J. Kim, J. Yim, J. Lee, P. Provost, O. Rådmark, S. Kim and V. N. Kim (2003). "The nuclear RNase III Drosha initiates microRNA processing." Nature **425**(6956): 415-419.

Li, J., X. Jiang, C. Li, Y. Liu, P. Kang, X. Zhong and Y. Cui (2019). "LncRNA-MEG3 inhibits cell proliferation and invasion by modulating Bmi1/RNF2 in cholangiocarcinoma." J Cell Physiol **234**(12): 22947-22959.

Li, L. and R. Bhatia (2011). "Stem cell quiescence." Clin Cancer Res **17**(15): 4936-4941.

Liu, Y., E. B. Carson-Walter, A. Cooper, B. N. Winans, M. D. Johnson and K. A. Walter (2010). "Vascular gene expression patterns are conserved in primary and metastatic brain tumors." J Neurooncol **99**(1): 13-24.

Lo, A. T., H. Mori, J. Mott and M. J. Bissell (2012). "Constructing three-dimensional models to study mammary gland branching morphogenesis and functional differentiation." J Mammary Gland Biol Neoplasia **17**(2): 103-110.

Lu, P., A. J. Ewald, G. R. Martin and Z. Werb (2008). "Genetic mosaic analysis reveals FGF receptor 2 function in terminal end buds during mammary gland branching morphogenesis." Dev Biol **321**(1): 77-87.

Lujambio, A. and S. W. Lowe (2012). "The microcosmos of cancer." Nature **482**(7385): 347-355.

Lund, M. J., K. F. Trivers, P. L. Porter, R. J. Coates, B. Leyland-Jones, O. W. Brawley, E. W. Flagg, R. M. O'Regan, S. G. Gabram and J. W. Eley (2009). "Race and triple negative threats to breast cancer survival: a population-based study in Atlanta, GA." Breast Cancer Res Treat **113**(2): 357-370.

Luo, M., M. Brooks and M. S. Wicha (2015). "Epithelial-mesenchymal plasticity of breast cancer stem cells: implications for metastasis and therapeutic resistance." Curr Pharm Des **21**(10): 1301-1310.

Ma, Q. L., G. G. Zhang and J. Peng (2013). "Vascular peroxidase 1: a novel enzyme in promoting oxidative stress in cardiovascular system." Trends Cardiovasc Med **23**(5): 179-183.

Macias, H. and L. Hinck (2012). "Mammary gland development." Wiley Interdiscip Rev Dev Biol **1**(4): 533-557.

Madden, S. F., C. Clarke, P. Gaule, S. T. Aherne, N. O'Donovan, M. Clynes, J. Crown and W. M. Gallagher (2013). "BreastMark: an integrated approach to mining publicly available transcriptomic datasets relating to breast cancer outcome." Breast Cancer Res **15**(4): R52.

Mailleux, A. A., B. Spencer-Dene, C. Dillon, D. Ndiaye, C. Savona-Baron, N. Itoh, S. Kato, C. Dickson, J. P. Thiery and S. Bellusci (2002). "Role of FGF10/FGFR2b signaling during mammary gland development in the mouse embryo." Development **129**(1): 53-60.

Mani, S. A., W. Guo, M. J. Liao, E. N. Eaton, A. Ayyanan, A. Y. Zhou, M. Brooks, F. Reinhard, C. C. Zhang, M. Shipitsin, L. L. Campbell, K. Polyak, C. Brisken, J. Yang and R. A. Weinberg (2008). "The epithelial-mesenchymal transition generates cells with properties of stem cells." Cell **133**(4): 704-715.

Mattick, J. S. and I. V. Makunin (2006). "Non-coding RNA." Hum Mol Genet **15 Spec No 1**: R17-29.

Mattick, J. S. and J. L. Rinn (2015). "Discovery and annotation of long noncoding RNAs." Nat Struct Mol Biol **22**(1): 5-7.

McMurray, E. N. and J. V. Schmidt (2012). "Identification of imprinting regulators at the Meg3 differentially methylated region." Genomics **100**(3): 184-194.

Mercer, T. R., M. E. Dinger and J. S. Mattick (2009). "Long non-coding RNAs: insights into functions." Nat Rev Genet **10**(3): 155-159.

Mercer, T. R., M. E. Dinger, S. M. Sunkin, M. F. Mehler and J. S. Mattick (2008). "Specific expression of long noncoding RNAs in the mouse brain." Proc Natl Acad Sci U S A **105**(2): 716-721.

Millikan, R. C., B. Newman, C.-K. Tse, P. G. Moorman, K. Conway, L. G. Dressler, L. V. Smith, M. H. Labbok, J. Geradts, J. T. Bensen, S. Jackson, S. Nyante, C. Livasy, L. Carey, H. S. Earp and C. M. Perou

(2008). "Epidemiology of basal-like breast cancer." Breast cancer research and treatment **109**(1): 123-139.

Mootha, V. K., C. M. Lindgren, K. F. Eriksson, A. Subramanian, S. Sihag, J. Lehar, P. Puigserver, E. Carlsson, M. Ridderstråle, E. Laurila, N. Houstis, M. J. Daly, N. Patterson, J. P. Mesirov, T. R. Golub, P. Tamayo, B. Spiegelman, E. S. Lander, J. N. Hirschhorn, D. Altshuler and L. C. Groop (2003). "PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes." Nat Genet **34**(3): 267-273.

Moustakas, A. and C. H. Heldin (2007). "Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression." Cancer Sci **98**(10): 1512-1520.

Naldini, L., D. Trono and I. M. Verma (2016). "Lentiviral vectors, two decades later." Science **353**(6304): 1101-1102.

Nanba, D., Y. Nakanishi and Y. Hieda (2001). "Changes in adhesive properties of epithelial cells during early morphogenesis of the mammary gland." Dev Growth Differ **43**(5): 535-544.

Nelson, R. E., L. I. Fessler, Y. Takagi, B. Blumberg, D. R. Keene, P. F. Olson, C. G. Parker and J. H. Fessler (1994). "Peroxidasin: a novel enzyme-matrix protein of Drosophila development." Embo j **13**(15): 3438-3447.

Neves, R., C. Scheel, S. Weinhold, E. Honisch, K. M. Iwaniuk, H. I. Trompeter, D. Niederacher, P. Wernet, S. Santourlidis and M. Uhrberg (2010). "Role of DNA methylation in miR-200c/141 cluster silencing in invasive breast cancer cells." BMC Res Notes **3**: 219.

Nie, L., H. J. Wu, J. M. Hsu, S. S. Chang, A. M. Labaff, C. W. Li, Y. Wang, J. L. Hsu and M. C. Hung (2012). "Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer." Am J Transl Res **4**(2): 127-150.

Nieto, M. A. (2002). "The snail superfamily of zinc-finger transcription factors." Nat Rev Mol Cell Biol **3**(3): 155-166.

Oyanagi, J., T. Ogawa, H. Sato, S. Higashi and K. Miyazaki (2012). "Epithelial-mesenchymal transition stimulates human cancer cells to extend microtubule-based invasive protrusions and suppresses cell growth in collagen gel." PLoS One **7**(12): e53209.

Pal, B., Y. Chen, A. Bert, Y. Hu, J. M. Sheridan, T. Beck, W. Shi, K. Satterley, P. Jamieson, G. J. Goodall, G. J. Lindeman, G. K. Smyth and J. E.

Visvader (2015). "Integration of microRNA signatures of distinct mammary epithelial cell types with their gene expression and epigenetic portraits." Breast cancer research : BCR **17**(1): 85-85.

Palomeras, S., S. Ruiz-Martínez and T. Puig (2018). "Targeting Breast Cancer Stem Cells to Overcome Treatment Resistance." Molecules **23**(9).

Park, S. M., A. B. Gaur, E. Lengyel and M. E. Peter (2008). "The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2." Genes Dev **22**(7): 894-907.

Parker, J. S., M. Mullins, M. C. Cheang, S. Leung, D. Voduc, T. Vickery, S. Davies, C. Fauron, X. He, Z. Hu, J. F. Quackenbush, I. J. Stijleman, J. Palazzo, J. S. Marron, A. B. Nobel, E. Mardis, T. O. Nielsen, M. J. Ellis, C. M. Perou and P. S. Bernard (2009). "Supervised risk predictor of breast cancer based on intrinsic subtypes." J Clin Oncol **27**(8): 1160-1167.

Parsa, S., S. K. Ramasamy, S. De Langhe, V. V. Gupte, J. J. Haigh, D. Medina and S. Bellusci (2008). "Terminal end bud maintenance in mammary gland is dependent upon FGFR2b signaling." Developmental Biology **317**(1): 121-131.

Peinado, H., D. Olmeda and A. Cano (2007). "Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?" Nat Rev Cancer **7**(6): 415-428.

Peppercorn, J., C. M. Perou and L. A. Carey (2008). "Molecular subtypes in breast cancer evaluation and management: divide and conquer." Cancer Invest **26**(1): 1-10.

Perez-Pinera, P., D. D. Kocak, C. M. Vockley, A. F. Adler, A. M. Kabadi, L. R. Polstein, P. I. Thakore, K. A. Glass, D. G. Ousterout, K. W. Leong, F. Guilak, G. E. Crawford, T. E. Reddy and C. A. Gersbach (2013). "RNA-guided gene activation by CRISPR-Cas9-based transcription factors." Nat Methods **10**(10): 973-976.

Perou, C. M., T. Sørlie, M. B. Eisen, M. van de Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, D. T. Ross, H. Johnsen, L. A. Akslen, O. Fluge, A. Pergamenschikov, C. Williams, S. X. Zhu, P. E. Lønning, A. L. Børresen-Dale, P. O. Brown and D. Botstein (2000). "Molecular portraits of human breast tumours." Nature **406**(6797): 747-752.

Péterfi, Z., A. Donkó, A. Orient, A. Sum, A. Prókai, B. Molnár, Z. Veréb, E. Rajnavölgyi, K. J. Kovács, V. Müller, A. J. Szabó and M. Geiszt (2009).

"Peroxidasin is secreted and incorporated into the extracellular matrix of myofibroblasts and fibrotic kidney." Am J Pathol **175**(2): 725-735.

Petersen, O. W., T. Gudjonsson, R. Villadsen, M. J. Bissell and L. Rønnov-Jessen (2003). "Epithelial progenitor cell lines as models of normal breast morphogenesis and neoplasia." Cell Prolif **36 Suppl 1**(Suppl 1): 33-44.

Pimentel, H., N. L. Bray, S. Puente, P. Melsted and L. Pachter (2017). "Differential analysis of RNA-seq incorporating quantification uncertainty." Nat Methods **14**(7): 687-690.

Pitteri, S. J., K. S. Kelly-Spratt, K. E. Gurley, J. Kennedy, T. B. Buson, A. Chin, H. Wang, Q. Zhang, C. H. Wong, L. A. Chodosh, P. S. Nelson, S. M. Hanash and C. J. Kemp (2011). "Tumor microenvironment-derived proteins dominate the plasma proteome response during breast cancer induction and progression." Cancer Res **71**(15): 5090-5100.

Plachot, C., L. S. Chaboub, H. A. Adissu, L. Wang, A. Urazaev, J. Sturgis, E. K. Asem and S. A. Lelièvre (2009). "Factors necessary to produce basoapical polarity in human glandular epithelium formed in conventional and high-throughput three-dimensional culture: example of the breast epithelium." BMC Biol **7**: 77.

Pond, A. C., X. Bin, T. Batts, K. Roarty, S. Hilsenbeck and J. M. Rosen (2013). "Fibroblast growth factor receptor signaling is essential for normal mammary gland development and stem cell function." Stem cells (Dayton, Ohio) **31**(1): 178-189.

Prasad Tharanga Jayasooriya, R. G., M. G. Dilshara, I. M. Neelaka Molagoda, C. Park, S. R. Park, S. Lee, Y. H. Choi and G. Y. Kim (2018). "Camptothecin induces G(2)/M phase arrest through the ATM-Chk2-Cdc25C axis as a result of autophagy-induced cytoprotection: Implications of reactive oxygen species." Oncotarget **9**(31): 21744-21757.

Prat, A. and C. M. Perou (2011). "Deconstructing the molecular portraits of breast cancer." Mol Oncol **5**(1): 5-23.

Qi, Lei S., Matthew H. Larson, Luke A. Gilbert, Jennifer A. Doudna, Jonathan S. Weissman, Adam P. Arkin and Wendell A. Lim (2013). "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression." Cell **152**(5): 1173-1183.

Rana, T. M. (2007). "Illuminating the silence: understanding the structure and function of small RNAs." Nat Rev Mol Cell Biol **8**(1): 23-36.

Ricardo, S., A. F. Vieira, R. Gerhard, D. Leitão, R. Pinto, J. F. Cameselle-Teijeiro, F. Milanezi, F. Schmitt and J. Paredes (2011). "Breast cancer

stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype." J Clin Pathol **64**(11): 937-946.

Ringnér, M., E. Fredlund, J. Häkkinen, Å. Borg and J. Staaf (2011). "GOBO: gene expression-based outcome for breast cancer online." PLoS One **6**(3): e17911.

Rowe, R. G. and S. J. Weiss (2008). "Breaching the basement membrane: who, when and how?" Trends Cell Biol **18**(11): 560-574.

Ruan, W. and D. L. Kleinberg (1999). "Insulin-Like Growth Factor I Is Essential for Terminal End Bud Formation and Ductal Morphogenesis during Mammary Development1." Endocrinology **140**(11): 5075-5081.

Sapino, A., L. Macri, P. Gugliotta and G. Bussolati (1990). "Evaluation of proliferating cell types in human and mouse mammary gland by a double immunostaining procedure." Acta Histochem Suppl **40**: 81-84.

Schaks, M., G. Giannone and K. Rottner (2019). "Actin dynamics in cell migration." Essays in biochemistry **63**(5): 483-495.

Shibue, T. and R. A. Weinberg (2017). "EMT, CSCs, and drug resistance: the mechanistic link and clinical implications." Nature Reviews Clinical Oncology **14**(10): 611-629.

Sigurdsson, V., B. Hilmarsdottir, H. Sigmundsdottir, A. J. Fridriksdottir, M. Ringnér, R. Villadsen, A. Borg, B. A. Agnarsson, O. W. Petersen, M. K. Magnusson and T. Gudjonsson (2011). "Endothelial induced EMT in breast epithelial cells with stem cell properties." PLoS One **6**(9): e23833.

Simian, M. and M. J. Bissell (2017). "Organoids: A historical perspective of thinking in three dimensions." J Cell Biol **216**(1): 31-40.

Singh, A. and J. Settleman (2010). "EMT, cancer stem cells and drug resistance: an emerging axis of evil in the war on cancer." Oncogene **29**(34): 4741-4751.

Sitole, B. N. and D. Mavri-Damelin (2018). "Peroxidasin is regulated by the epithelial-mesenchymal transition master transcription factor Snai1." Gene **646**: 195-202.

Soifer, H., J. Rossi and P. Saetrom (2008). "MicroRNAs in Disease and Potential Therapeutic Applications." Molecular therapy : the journal of the American Society of Gene Therapy **15**: 2070-2079.

Soudi, M., M. Paumann-Page, C. Delporte, K. F. Pirker, M. Bellei, E. Edenhofer, G. Stadlmayr, G. Battistuzzi, K. Z. Boudjeltia, P. G. Furtmüller, P. Van Antwerpen and C. Obinger (2015). "Multidomain human peroxidasin 1 is a highly glycosylated and stable homotrimeric

high spin ferric peroxidase." The Journal of biological chemistry **290**(17): 10876-10890.

Soudi, M., M. Zamocky, C. Jakopitsch, P. G. Furtmüller and C. Obinger (2012). "Molecular evolution, structure, and function of peroxidasins." Chem Biodivers **9**(9): 1776-1793.

Souza, A., Ferreir, ICC, Marangoni, K, Bastos, VAF, Goulart, VA (2016). "Advances in Cell Culture: More than a Century after Cultivating Cells." J Biotechnol Biomater **6:221**.

Stingl, J., C. J. Eaves, I. Zandieh and J. T. Emerman (2001). "Characterization of bipotent mammary epithelial progenitor cells in normal adult human breast tissue." Breast Cancer Res Treat **67**(2): 93-109.

Stingl, J., A. Raouf, J. T. Emerman and C. J. Eaves (2005). "Epithelial progenitors in the normal human mammary gland." J Mammary Gland Biol Neoplasia **10**(1): 49-59.

Subramanian, A., P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander and J. P. Mesirov (2005). "Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles." Proceedings of the National Academy of Sciences **102**(43): 15545-15550.

Sumbal, J., Z. Budkova, G. Traustadóttir and Z. Koledova (2020). "Mammary Organoids and 3D Cell Cultures: Old Dogs with New Tricks." J Mammary Gland Biol Neoplasia **25**(4): 273-288.

Sun, L., H. Luo, Q. Liao, D. Bu, G. Zhao, C. Liu, Y. Liu and Y. Zhao (2013). "Systematic study of human long intergenic non-coding RNAs and their impact on cancer." Science China Life Sciences **56**(4): 324-334.

Terashima, M., S. Tange, A. Ishimura and T. Suzuki (2017). "MEG3 Long Noncoding RNA Contributes to the Epigenetic Regulation of Epithelial-Mesenchymal Transition in Lung Cancer Cell Lines." J Biol Chem **292**(1): 82-99.

Thiery, J. P., H. Acloque, R. Y. Huang and M. A. Nieto (2009). "Epithelial-mesenchymal transitions in development and disease." Cell **139**(5): 871-890.

Trappmann, B., J. E. Gautrot, J. T. Connelly, D. G. Strange, Y. Li, M. L. Oyen, M. A. Cohen Stuart, H. Boehm, B. Li, V. Vogel, J. P. Spatz, F. M. Watt and W. T. Huck (2012). "Extracellular-matrix tethering regulates stem-cell fate." Nat Mater **11**(7): 642-649.

Uhlén, M., L. Fagerberg, B. M. Hallström, C. Lindskog, P. Oksvold, A. Mardinoglu, Å. Sivertsson, C. Kampf, E. Sjöstedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A.-K. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P.-H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen and F. Pontén (2015). "Tissue-based map of the human proteome." Science **347**(6220): 1260419.

Uroda, T., E. Anastasakou, A. Rossi, J.-M. Teulon, J.-L. Pellequer, P. Annibale, O. Pessey, A. Inga, I. Chillón and M. Marcia (2019). "Conserved Pseudoknots in lncRNA MEG3 Are Essential for Stimulation of the p53 Pathway." Molecular cell **75**(5): 982-995.e989.

van Deurs, B., Z. Z. Zou, P. Briand, Y. Balslev and O. W. Petersen (1987). "Epithelial membrane polarity: a stable, differentiated feature of an established human breast carcinoma cell line MCF-7." J Histochem Cytochem **35**(4): 461-469.

Varghese, V. K., V. Shukla, S. P. Kabekkodu, D. Pandey and K. Satyamoorthy (2018). "DNA methylation regulated microRNAs in human cervical cancer." Mol Carcinog **57**(3): 370-382.

Veltmaat, J. M., F. d. r. Relaix, L. T. Le, K. Kratochwil, F. d. r. G. Sala, W. van Veelen, R. Rice, B. Spencer-Dene, A. A. Mailleux, D. P. Rice, J. P. Thiery and S. Bellusci (2006). "Gli3-mediated somitic Fgf10 expression gradients are required for the induction and patterning of mammary epithelium along the embryonic axes." Development **133**(12): 2325-2335.

Villadsen, R. (2005). "In search of a stem cell hierarchy in the human breast and its relevance to breast cancer evolution." Apmis **113**(11-12): 903-921.

Villadsen, R., A. J. Fridriksdottir, L. Rønnov-Jessen, T. Gudjonsson, F. Rank, M. A. LaBarge, M. J. Bissell and O. W. Petersen (2007). "Evidence for a stem cell hierarchy in the adult human breast." J Cell Biol **177**(1): 87-101.

Visvader, J. E. (2009). "Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis." Genes Dev **23**(22): 2563-2577.

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

Wang, D., C. W. Fu and D. Q. Fan (2019). "Participation of tumor suppressors long non-coding RNA MEG3, microRNA-377 and PTEN in glioma cell invasion and migration." Pathol Res Pract **215**(10): 152558.

Watt, F. M. (1989). "Terminal differentiation of epidermal keratinocytes." Curr Opin Cell Biol **1**(6): 1107-1115.

Watt, F. M. and B. L. Hogan (2000). "Out of Eden: stem cells and their niches." Science **287**(5457): 1427-1430.

Wei, G. H. and X. Wang (2017). "lncRNA MEG3 inhibit proliferation and metastasis of gastric cancer via p53 signaling pathway." Eur Rev Med Pharmacol Sci **21**(17): 3850-3856.

Wei, J. W., K. Huang, C. Yang and C. S. Kang (2017). "Non-coding RNAs as regulators in epigenetics (Review)." Oncol Rep **37**(1): 3-9.

Weigelt, B. and M. J. Bissell (2008). "Unraveling the microenvironmental influences on the normal mammary gland and breast cancer." Semin Cancer Biol **18**(5): 311-321.

Weigelt, B., C. M. Ghajar and M. J. Bissell (2014). "The need for complex 3D culture models to unravel novel pathways and identify accurate biomarkers in breast cancer." Adv Drug Deliv Rev **69-70**: 42-51.

Wuidart, A., A. Sifrim, M. Fioramonti, S. Matsumura, A. Brisebarre, D. Brown, A. Centonze, A. Dannau, C. Dubois, A. Van Keymeulen, T. Voet and C. Blanpain (2018). "Early lineage segregation of multipotent embryonic mammary gland progenitors." Nat Cell Biol **20**(6): 666-676.

Yan, X., S. Sabrautzki, M. Horsch, H. Fuchs, V. Gailus-Durner, J. Beckers, M. Hrabě de Angelis and J. Graw (2014). "Peroxidasin is essential for eye development in the mouse." Hum Mol Genet **23**(21): 5597-5614.

Yang, J. and R. A. Weinberg (2008). "Epithelial-mesenchymal transition: at the crossroads of development and tumor metastasis." Dev Cell **14**(6): 818-829.

Yang, Z., E. Bian, Y. Xu, X. Ji, F. Tang, C. Ma, H. Wang and B. Zhao (2020). "Meg3 Induces EMT and Invasion of Glioma Cells via Autophagy." Onco Targets Ther **13**: 989-1000.

Yates, A. D., P. Achuthan, W. Akanni, J. Allen, J. Allen, J. Alvarez-Jarreta, M. R. Amode, I. M. Armean, A. G. Azov, R. Bennett, J. Bhai, K. Billis, S. Boddu, J. C. Marugán, C. Cummins, C. Davidson, K. Dodiya, R. Fatima, A. Gall, C. G. Giron, L. Gil, T. Grego, L. Haggerty, E. Haskell, T. Hourlier, O. G. Izuogu, S. H. Janacek, T. Juettemann, M. Kay, I. Lavidas, T. Le, D. Lemos, J. G. Martinez, T. Maurel, M. McDowall, A. McMahon, S. Mohanan, B. Moore, M. Nuhn, D. N. Oheh, A. Parker, A. Parton, M. Patricio, M. P. Sakthivel, A. I. Abdul Salam, B. M. Schmitt, H. Schuilenburg, D. Sheppard, M. Sycheva, M. Szuba, K. Taylor, A.

Thormann, G. Threadgold, A. Vullo, B. Walts, A. Winterbottom, A. Zadissa, M. Chakiachvili, B. Flint, A. Frankish, S. E. Hunt, G. IIsley, M. Kostadima, N. Langridge, J. E. Loveland, F. J. Martin, J. Morales, J. M. Mudge, M. Muffato, E. Perry, M. Ruffier, S. J. Trevanion, F. Cunningham, K. L. Howe, D. R. Zerbino and P. Flicek (2019). "Ensembl 2020." Nucleic Acids Research **48**(D1): D682-D688.

Ye, X., W. L. Tam, T. Shibue, Y. Kaygusuz, F. Reinhardt, E. Ng Eaton and R. A. Weinberg (2015). "Distinct EMT programs control normal mammary stem cells and tumour-initiating cells." Nature **525**(7568): 256-260.

Ye, X. and R. A. Weinberg (2015). "Epithelial-Mesenchymal Plasticity: A Central Regulator of Cancer Progression." Trends Cell Biol **25**(11): 675-686.

Yi, R., M. N. Poy, M. Stoffel and E. Fuchs (2008). "A skin microRNA promotes differentiation by repressing 'stemness'." Nature **452**(7184): 225-229.

Yi, R., Y. Qin, I. G. Macara and B. R. Cullen (2003). "Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs." Genes Dev **17**(24): 3011-3016.

Yu, F., W. Geng, P. Dong, Z. Huang and J. Zheng (2018). "LncRNA-MEG3 inhibits activation of hepatic stellate cells through SMO protein and miR-212." Cell Death Dis **9**(10): 1014.

Yuan, W., Y. Wang, J. W. Heinecke and X. Fu (2009). "Hypochlorous acid converts the gamma-glutamyl group of glutathione disulfide to 5-hydroxybutyrolactam, a potential marker for neutrophil activation." J Biol Chem **284**(39): 26908-26917.

Zaravinos, A. (2015). "The Regulatory Role of MicroRNAs in EMT and Cancer." J Oncol **2015**: 865816.

Zare, M., M. Bastami, S. Solali and M. R. Alivand (2018). "Aberrant miRNA promoter methylation and EMT-involving miRNAs in breast cancer metastasis: Diagnosis and therapeutic implications." J Cell Physiol **233**(5): 3729-3744.

Zeisberg, M., G. Bonner, Y. Maeshima, P. Colorado, G. A. Müller, F. Strutz and R. Kalluri (2001). "Renal Fibrosis: Collagen Composition and Assembly Regulates Epithelial-Mesenchymal Transdifferentiation." The American Journal of Pathology **159**(4): 1313-1321.

Zhang, H., F. A. Kolb, V. Brondani, E. Billy and W. Filipowicz (2002). "Human Dicer preferentially cleaves dsRNAs at their termini without a requirement for ATP." Embo j **21**(21): 5875-5885.

Zhang, J. and Y. Gao (2019). "Long non-coding RNA MEG3 inhibits cervical cancer cell growth by promoting degradation of P-STAT3 protein via ubiquitination." Cancer Cell Int **19**: 175.

Zhang, W., S. Shi, J. Jiang, X. Li, H. Lu and F. Ren (2017). "LncRNA MEG3 inhibits cell epithelial-mesenchymal transition by sponging miR-421 targeting E-cadherin in breast cancer." Biomed Pharmacother **91**: 312-319.

Zhang, Z., B. Zhang, W. Li, L. Fu, L. Fu, Z. Zhu and J. T. Dong (2011). "Epigenetic Silencing of miR-203 Upregulates SNAI2 and Contributes to the Invasiveness of Malignant Breast Cancer Cells." Genes Cancer **2**(8): 782-791.

Zheng, Y.-Z. and L. Liang (2018). "High expression of PXDN is associated with poor prognosis and promotes proliferation, invasion as well as migration in ovarian cancer." Annals of Diagnostic Pathology **34**: 161-165.

Zhou, P., B. Li, F. Liu, M. Zhang, Q. Wang, Y. Liu, Y. Yao and D. Li (2017). "The epithelial to mesenchymal transition (EMT) and cancer stem cells: implication for treatment resistance in pancreatic cancer." Mol Cancer **16**(1): 52.

Zhou, Y., Y. Zhong, Y. Wang, X. Zhang, D. L. Batista, R. Gejman, P. J. Ansell, J. Zhao, C. Weng and A. Klibanski (2007). "Activation of p53 by MEG3 non-coding RNA." J Biol Chem **282**(34): 24731-24742.

Zhu, J., S. Liu, F. Ye, Y. Shen, Y. Tie, J. Zhu, L. Wei, Y. Jin, H. Fu, Y. Wu and X. Zheng (2015). "Long Noncoding RNA MEG3 Interacts with p53 Protein and Regulates Partial p53 Target Genes in Hepatoma Cells." PLoS One **10**(10): e0139790.

Zhu, W., E. M. Botticelli, R. E. Kery, Y. Mao, X. Wang, A. Yang, X. Wang, J. Zhou, X. Zhang, R. J. Soberman, A. Klibanski and Y. Zhou (2019). "Meg3-DMR, not the Meg3 gene, regulates imprinting of the Dlk1-Dio3 locus." Dev Biol **455**(1): 10-18.

Article I

Mechanisms of Development 155 (2019) 34–47

Contents lists available at ScienceDirect

Mechanisms of Development

journal homepage: www.elsevier.com/locate/mod

MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin

Eirikur Briem[a], Zuzana Budkova[a], Anna Karen Sigurdardottir[a], Bylgja Hilmarsdottir[a,d], Jennifer Kricker[a], Winston Timp[e], Magnus Karl Magnusson[b,c], Gunnhildur Asta Traustadottir[a], Thorarinn Gudjonsson[a,b,*]

[a] Stem Cell Research Unit, Biomedical Center, Department of Anatomy, Faculty of Medicine, School of Health Sciences, University of Iceland, Iceland
[b] Department of Laboratory Hematology, Landspitali - University Hospital, Iceland
[c] Department of Pharmacology and Toxicology, Faculty of Medicine, School of Health Sciences, University of Iceland, Iceland
[d] Department of Tumor Biology, The Norwegian Radium Hospital, Oslo, Norway
[e] Department of Biomedical Engineering, Johns Hopkins University, USA

A B S T R A C T

MicroRNAs regulate developmental events such as branching morphogenesis, epithelial to mesenchymal transition (EMT) and its reverse process mesenchymal to epithelial transition (MET). In this study, we performed small RNA sequencing of a breast epithelial progenitor cell line (D492), and its mesenchymal derivative (D492M) cultured in three-dimensional microenvironment. Among the most downregulated miRNAs in D492M was miR-203a, a miRNA that plays an important role in epithelial differentiation. Increased expression of miR-203a was seen in D492, concomitant with increased complexity of branching. When miR-203a was overexpressed in D492M, a partial reversion towards epithelial phenotype was seen. Gene expression analysis of D492M and D492M$^{miR\text{-}203a}$ revealed peroxidasin, a collagen IV cross-linker, as the most significantly downregulated gene in D492M$^{miR\text{-}203a}$. Collectively, we demonstrate that miR-203a expression temporally correlates with branching morphogenesis and is suppressed in D492M. Overexpression of miR-203a in D492M induces a partial MET and reduces the expression of peroxidasin. Furthermore, we demonstrate that miR-203a is a novel repressor of peroxidasin. MiR-203-peroxidasin axis may be an important regulator in branching morphogenesis, EMT/MET and basement membrane remodeling.

Summary statement

MiR-203a is highly upregulated during branching morphogenesis of D492 breast epithelial progenitor cells and is downregulated in its isogenic mesenchymal cell line D492M. Furthermore, miR-203a is a novel suppressor of peroxidasin (PXDN), a collagen IV crosslinking agent. MiR-203a may be an important regulator of branching morphogenesis and basement remodeling in the human breast gland, possibly through its target PXDN.

1. Introduction

Developmental events underlying breast epithelial morphogenesis are closely related to pathways important to cancer progression, *i.e.* epithelial to mesenchymal transition (EMT) and mesenchymal to epithelial transition (MET). Evidence shows that the two distinct epithelial cell lineages that make up branching morphogenesis in the breast, luminal- and myoepithelial cells, originate from common breast epithelial stem cells (Pechoux et al., 1999; Gudjonsson et al., 2002; Villadsen et al., 2007; Petersen and Polyak, 2010). These stem cells are responsible for continuous tissue remodeling throughout the reproductive period, as well as the extensive epithelial expansion and branching morphogenesis seen during pregnancy and lactation. Although, potential stem or progenitor cells have been identified in the human female breast gland there is still limited knowledge about the lineage development in the human breast. In that respect much can be learned from linage tracing studies in the mouse mammary gland. Recent studies have demonstrated that embryonic mammary gland contains homogenous basal progenitor cells that become restricted in the postnatal mammary gland (Lloyd-Lewis et al., 2018). Lilja et al. demonstrated that embryonic multipotent mammary cells become lineage-restricted early in development and that gain of function dictates luminal cell fate specification to both embryonic and basally committed mammary cells (Lilja et al., 2018). In a recent article Pal et al. demonstrated by single cell profiling of four developmental stages in the post-natal gland that the epithelium undergoes large changes in gene expression (Pal et al., 2017). In this paper they show that homogeneous basal-like expression pattern in pre-puberty was distinct to lineage-restricted programs in

* Corresponding author at: Stem Cell Research Unit, Department of Medical Faculty, Biomedical Center, University of Iceland, Vatnsmyrarvegi 16, 101 Reykjavík, Iceland.
E-mail address: tgudjons@hi.is (T. Gudjonsson).

https://doi.org/10.1016/j.mod.2018.11.002
Received 29 March 2018; Received in revised form 1 November 2018; Accepted 23 November 2018
Available online 01 December 2018

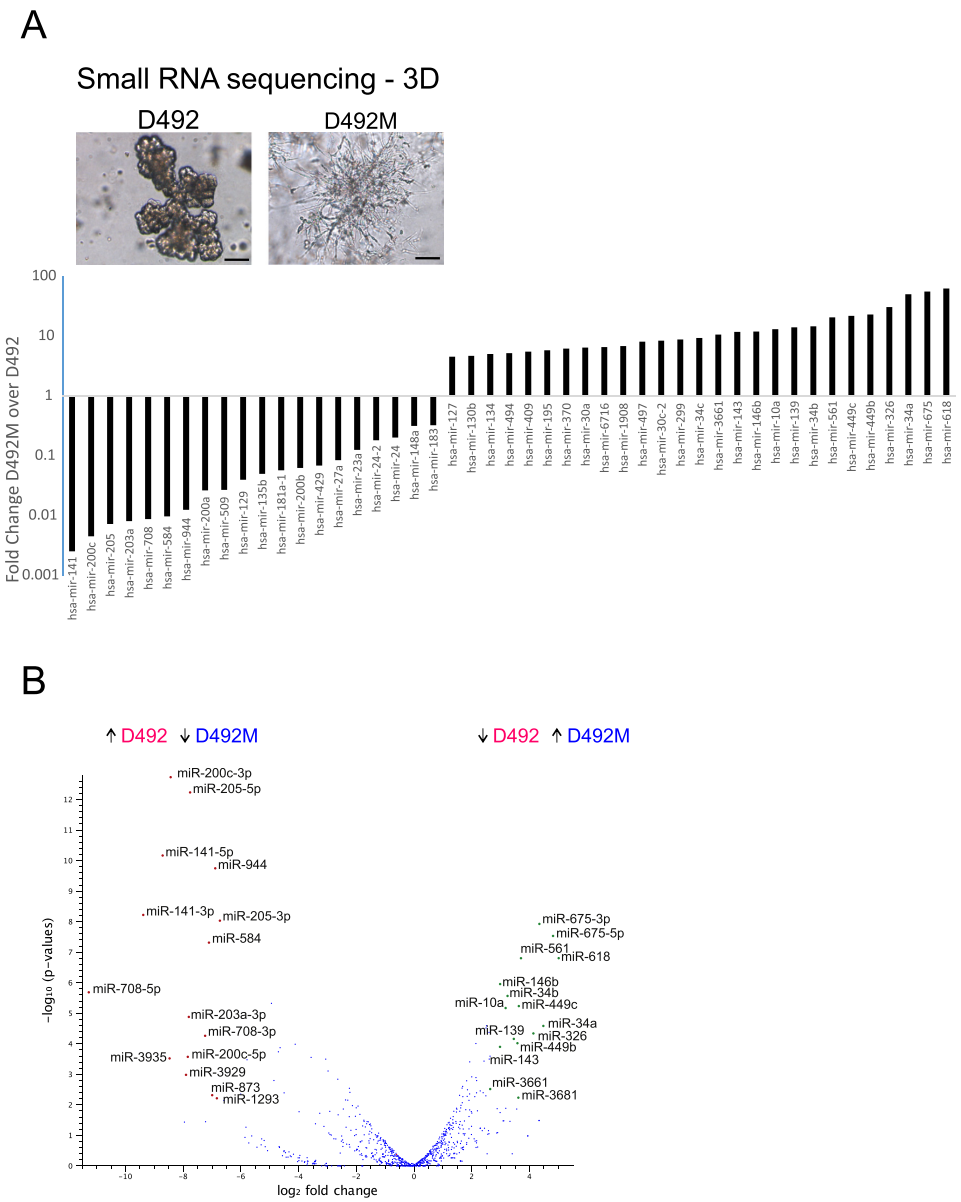


Fig. 1. Profiling miRNAs in the isogenic breast epithelial progenitor cell line D492 and its mesenchymal derivative D492M.
A) Differentially expressed miRNAs between D492M and D492 in 3D culture based on small RNA sequencing. Forty-seven miRNAs are differentially expressed between the mesenchymal (D492M) and epithelial (D492) state at day 14 in 3D culture. Twenty miRNAs are downregulated in D492M and among them are miRNAs that are important for epithelial integrity, such as the miR-200 family and miR-203a. Scale bar = 100 μm.
B) Volcano plot showing differential expressed miRNAs in D492M and D492, depicting statistical significance and high fold change in expression of epithelial associated miRNAs. Volcano plot constructed from small RNA sequencing data using fold change of expression values and p-values, enabling the visualization of the relationship between fold change and statistical significance. More extreme values on the x-axis show increased differential expression and higher values on the y-axis show increased statistical significance. The miR-200c-141 locus along with miR-205, miR-944, miR-584, miR-708 and miR-203a show the most fold change and statistical significance of the miRNAs downregulated in D492M.

puberty. Cellular linage tracing in humans is for obvious reasons not possible. However, transplantation of human breast epithelial cells into cleared mouse mammary fat pad (Lawson et al., 2015; Wronski et al., 2015) and 3D culture (Vidi et al., 2013) have to some extend unraveled the stem cell biology in the human breast gland. D492 is a breast epithelial stem cell line that generates branching *in vivo*-like structures in 3D-rBM (Gudjonsson et al., 2002; Villadsen et al., 2007; Sigurdsson et al., 2011). We have previously shown that when D492 is co-cultured

under 3D conditions, breast endothelial cells markedly stimulate their branching ability, but can also induce an irreversible EMT in the epithelial cells (Sigurdsson et al., 2011). This condition gave rise to the D492M cell line (Sigurdsson et al., 2011). In EMT, different pathways ultimately control transcriptional regulatory factors such as SNAI1, SNAI2, TWIST, ZEB1 and ZEB2 leading to increased expression of mesenchymal and decreased expression of epithelial markers (Moustakas and Heldin, 2007). Downregulation of the epithelial cell-cell adhesion proteins, such as E-cadherin, is one of the hallmarks of EMT (Peinado et al., 2007). In addition, expression of epithelial specific keratins is often greatly reduced, while expression of mesenchymal markers, such as N-cadherin, vimentin, alpha smooth muscle actin and fibronectin are increased in EMT (Moustakas and Heldin, 2007). Enhanced migration, invasion and resistance to apoptosis are also characteristic for cells that have undergone EMT (Hanahan and Weinberg, 2011). The tightly regulated process of cell conversion seen in EMT is also a critical event seen in many cancer types, including breast cancer (Petersen et al., 2001; Mani et al., 2008; Sarrio et al., 2008) and EMT is typically associated with increased aggressiveness and metastatic behavior (Hanahan and Weinberg, 2011).

Interestingly, the mesenchymal transition from D492 to D492M is accompanied by drastic changes in microRNA (miRNA) expression (Hilmarsdottir et al., 2015). MiRNAs have been described as regulators of protein expression through their ability to bind and silence mRNAs, where silencing of certain transcription factors, which are important for gene regulation, may cause marked changes in cell phenotype and fate (Hilmarsdottir et al., 2014). In recent years, miRNAs have been shown to be either tumor promoting or tumor suppressing depending on context and cancer type. For example, the miR-200 family has been linked to both stem cell regulation and cancer progression (Shimono et al., 2009). This family of miRNAs is strongly downregulated in tumors with high metastatic potential (Olson et al., 2009) and miR-200c, a member of this family, is downregulated in human breast cancer stem cells and normal mammary stem cells (Shimono et al., 2009). We have recently shown that the miRNA-200c-141 cluster is predominantly expressed in luminal breast epithelial cells, and that miRNA-200c-141 is highly expressed in D492, but not in D492M. Interestingly, overexpression of miR-200c-141 in D492M restored the luminal epithelial phenotype (Hilmarsdottir et al., 2015).

In this study, we compared the miRNA expression profile of D492 and D492M, when cultured in 3D reconstituted basement membrane matrix (3D-rBM). Small RNA sequencing revealed a striking difference in expression patterns between the two cell lines, with a number of known epithelial miRNAs, including miR-203a, downregulated in D492M. We demonstrated that miR-203a expression increases during branching morphogenesis *in vitro* similar to its expression pattern in mouse mammary gland *in vivo* (Avril-Sassen et al., 2009). Furthermore, overexpression of miR-203a in D492M induced partial phenotypic changes towards MET, reduced cell proliferation, migration, invasion, and increased sensitivity to chemically induced apoptosis. Finally, we identified miR-203a as a novel repressor of peroxidasin (PXDN), an extracellular matrix protein with peroxidase activity and a collagen IV crosslinking agent.

2. Results

2.1. Comparison of miRNA expression in the breast epithelial progenitor cell line D492 and its mesenchymal derivative D492M

D492 and D492M are isogenic cell lines with epithelial and mesenchymal phenotypes, respectively. We have previously shown a profound difference in miRNA expression between these cell lines when cultured in monolayer (Hilmarsdottir et al., 2015). Here, we conducted an expression analysis in 3D culture. Due to its stem cell properties, D492 can generate branching structures in 3D-rBM culture reminiscent of terminal duct lobular units (TDLUs) in the breast. In contrast, D492M form disorganized mesenchymal-like structures with spindle shape protrusions of cells (Fig. 1A, top). To investigate the miRNA expression patterns between D492 and D492M in 3D-rBM we performed small RNA sequencing (Fig. 1A). This revealed 47 differentially expressed miRNAs (> 1.5-fold change, $p < 0.05$ and FDR < 0.1) between D492 from (days 7, 14 and 21) and D492M at day 14 in 3D culture, where 20 of these miRNAs were downregulated in D492M. Among the most profound changes was downregulation of miR-200c, miR-141, miR-205 and miR-203a in D492M (Fig. 1A), all of which have been previously associated with epithelial integrity and repression of EMT (Gregory et al., 2008; Wellner et al., 2009; Feng et al., 2014). These miRNAs were all downregulated with high statistical significance in D492M as shown in the volcano plot (Fig. 1B). The upregulated miRNAs in D492M do not show as clear involvement in EMT as the downregulated ones.

2.2. Downregulation of miR-203a in D492M is not due to methylation of its CpG islands in the promoter area

Methylation of CpG islands in promoter areas is a common event during gene silencing. To further explore what causes downregulation of miRNAs in D492M, we explored the promoter DNA methylation status of differentially expressed miRNAs. We used the HumanMethylation450K BeadChip and looked at CpGs within 5 kb region of the promoter area of differentially expressed miRNAs. Two miRNAs did not have probes within the 5 kb region of the gene but of the 45 which did, 8 miRNAs showed differential DNA methylation between D492 and D492M (Table S2). Our data shows that promoter areas upstream of miR-200c-141 and miR-205 were methylated in D492M, but not in D492 (Fig. 2A). In contrast, the promoter area of miR-203a was not differentially methylated between D492 and D492M (Fig. 2A). Bisulfite sequencing of the promoter area of miR-200c-141, miR-205 and miR-203a confirmed that the promoter of miR-203a was unmethylated in D492M, unlike the promoters of miR-200c-141 and miR-205 (Fig. 2B). This is in contrast to previously published results in human mammary epithelial cells (HMLE) undergoing EMT (Taube et al., 2013) and some metastatic breast cancer cell lines (Zhang et al., 2011), where the promoter of miR-203a becomes methylated. Our data shows that a profound downregulation of miR-203a could be mediated through other mechanisms than DNA methylation, such as histone modifications. Indeed, miR-203a expression has been shown to be suppressed by EZH2 in prostate cancer (Cao et al., 2011). In summary, the repression of miR-203a, in endothelial induced EMT of D492 is not due to promoter methylation.

2.3. MiR-203a shows temporal changes in expression during branching morphogenesis in 3D culture and is associated with luminal breast epithelial cells

The breast gland is a dynamic organ during the reproduction period. In each menstrual cycle the breast epithelium undergoes changes associated with branching morphogenesis and if pregnancy occurs, the branching epithelium expands resulting in maximal differentiation during lactation. The glandular epithelium is then subject to apoptosis after breastfeeding during the involution phase (Javed and Lteif, 2013). Using D492 breast progenitor cells that are capable of generating branching epithelial morphogenesis in 3D-rBM, it is possible to analyze temporal changes during the branching process. To investigate if changes in miRNA expression occur during formation of TDLU–like structures in 3D-rBM, we isolated RNA from 3D-rBM cultures at days 7, 14 and 21, and performed small RNA sequencing.

Fifty-five miRNAs were differentially expressed between individual time points (> 2-fold change, FDR corrected p-value < 0.05), thereof 40 miRNAs were downregulated and 15 miRNAs were upregulated during branching (Fig. 3A). In particular, among the upregulated miRNAs were miR-141 and miR-203, which were also some of the most downregulated miRNAs in D492M (Fig. 1A). Thus, small RNA

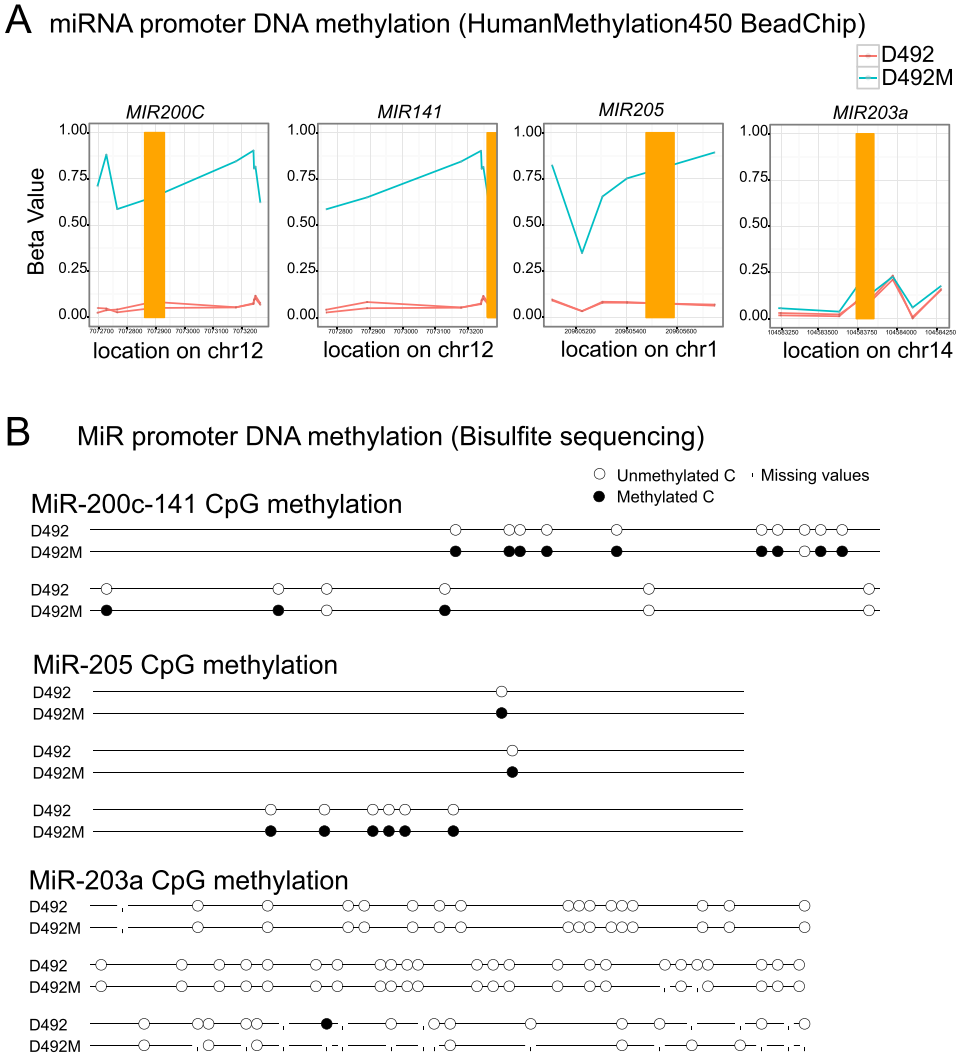
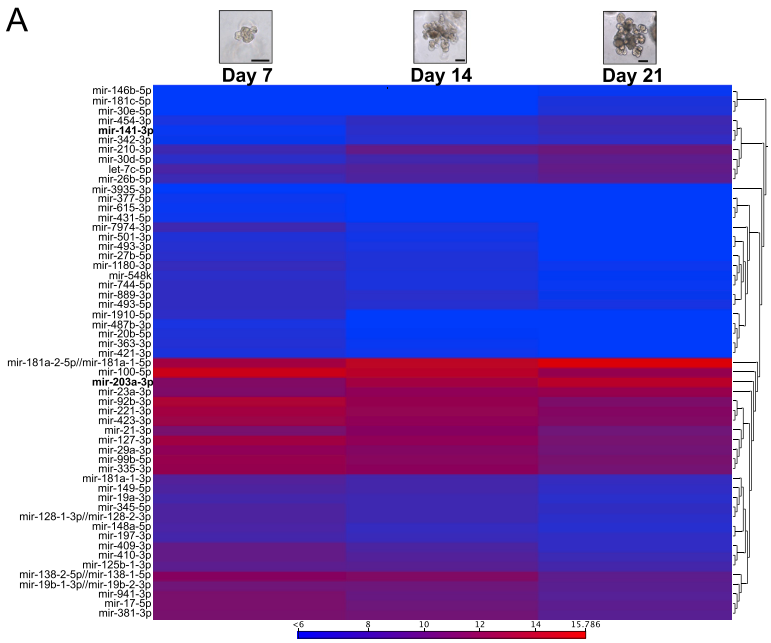


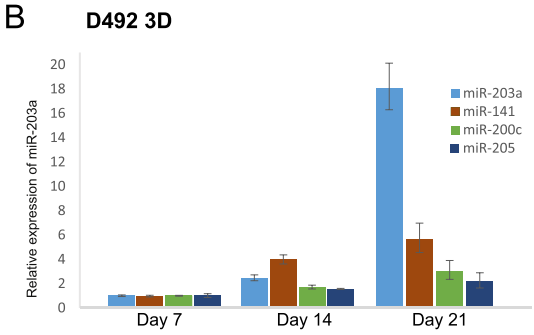
Fig. 2. miRNA promoter methylation of D492 and D492M

A) Differential DNA methylation of D492 (red) and D492M (blue). CpG methylation in the promoter area of differentially expressed miRNAs was investigated using the HumanMethylation450 BeadChip. The yellow bar represents the pre-miRNA and differential methylation is determined by a beta value > 0.2. MiR-200c, miR-141 and miR-205 are differentially methylated with more methylation in D492M, while methylation of miRNA-203a is unchanged. The expression of miR-200c-141 and miR-205 is repressed through DNA methylation, while expression of miRNA-203a is not.

B) Bisulfite sequencing of miRNA promoters in D492 and D492M. The promoters of miR-200c-141, miR-205 and miR-203a are un-methylated in D492. In D492M, the promoter of miR-203a remains un-methylated, while the promoters of miR-200c-141 and miR-205 are methylated.

sequencing, and subsequent verification with RT-qPCR, demonstrates that expression of miR-203a and miR-141 increased between time points and peaked at the late branching stage (day 21) (Fig. 3A–B). In contrast, expression of miR-205 and miR-200c was constant throughout the branching process (Fig. 3B). This indicates that miR-203a and miR-141 may play a role in differentiation of the epithelial cells during branching morphogenesis. The human breast epithelium is composed of two epithelial lineages, the luminal epithelial and the myoepithelial cells. In order to investigate whether miR-203a expression was lineage specific with either luminal- or myoepithelial cells, we sorted primary breast epithelial cells with EpCAM positive magnetic beads, as EpCAM is a luminal associated adhesion molecule that is also highly useful for antibody-based enrichment of luminal epithelial cells. We measured miR-203a expression in EpCAM high and EpCAM low cells, and demonstrated that, miR-203a expression was predominately associated with the EpCAM high luminal epithelial cells (Fig. 3C, left). There was some expression in the EpCAM low/negative myoepithelial cells, but no or little expression in endothelial cells and fibroblast (Fig. 3C, left). Due to the bipotential properties of D492 cells, they are able to generate luminal- and myoepithelial cells and differentiation of D492 branching colonies, into the two epithelial lineages, is clearly demonstrated with immunostaining against luminal- and myoepithelial markers (Gudjonsson et al., 2002). Based on EpCAM sorting of D492 cells into EpCAM high and EpCAM low fractions, we further confirmed that miR-





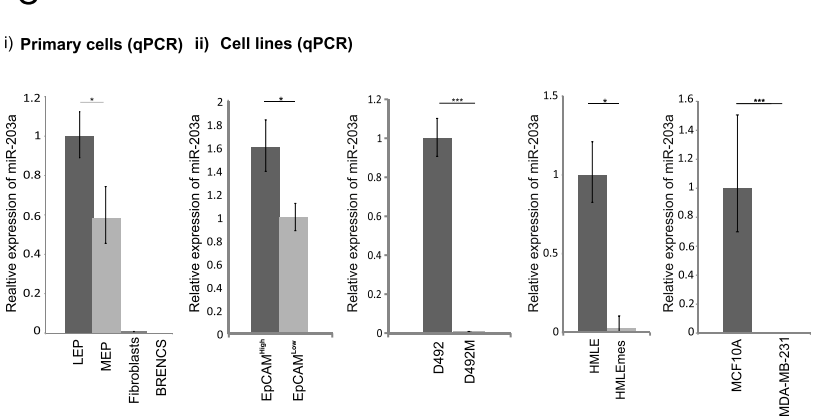


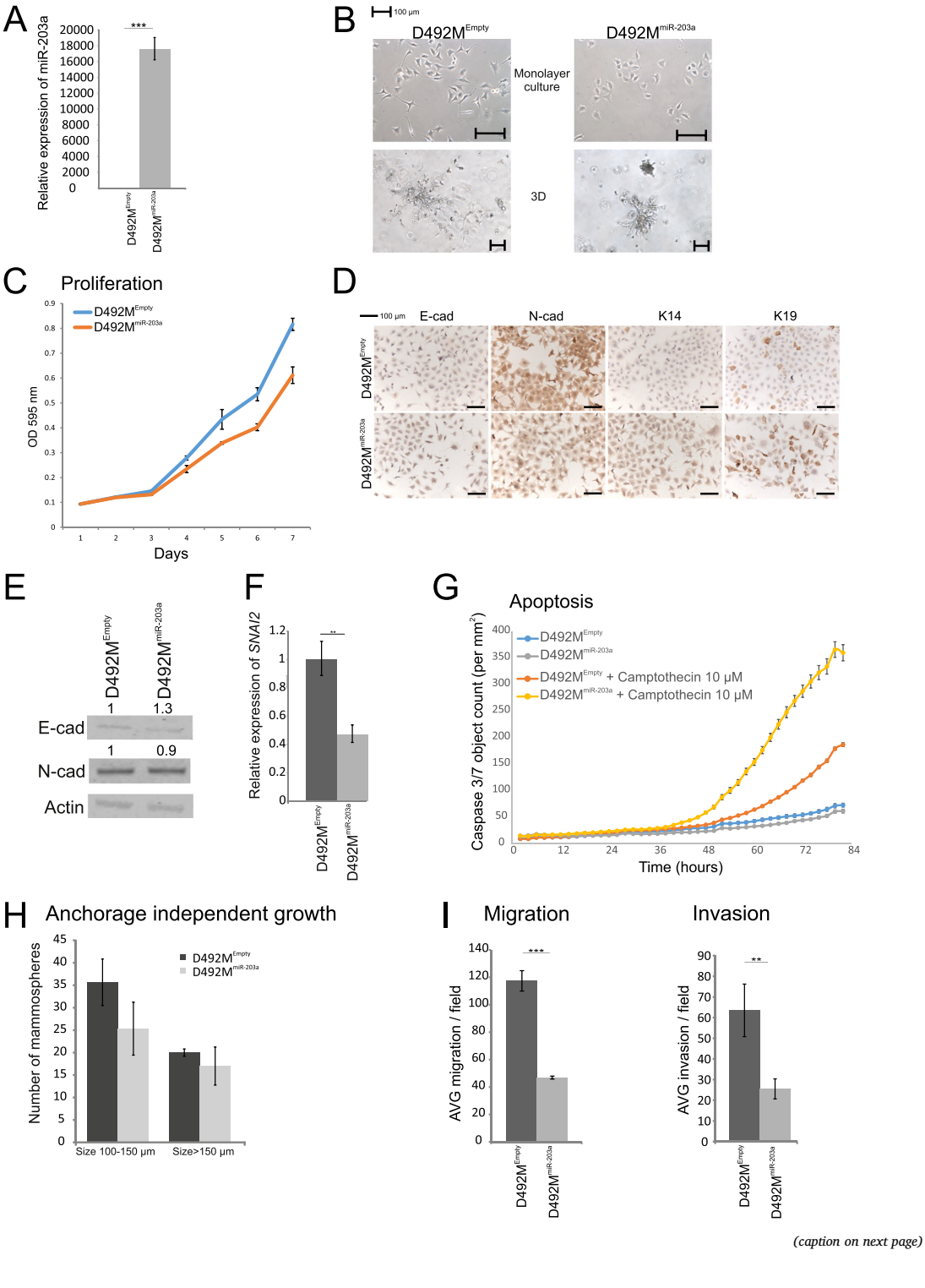
Fig. 3. MiR-203a is differentially expressed during branching morphogenesis in 3D culture.

A) Heatmap showing differentially expressed miRNAs in D492 at different time points during branching morphogenesis. Small RNA-sequencing of D492 at three different time points during branching morphogenesis at day 7, 14 and 21. miR-203a expression increases during branching of D492. Heatmap shows miRNAs with > 2-fold change in expression an FDR corrected p-values < 0.1. Scale bar = 100 µm.

B) Expression of epithelial miRNAs in D492 at different time points during branching. RT-qPCR showing that the expression of miR-203a and miR-141 significantly increases between time points during branching morphogenesis. Interestingly, expression of miR-203a increases markedly in late branching at day 21.

C) i) Expression of miR-203a in breast primary cells and D492 EpCAM sorted cells. In primary cells, miR-203a is expressed in the epithelial cells and is predominantly associated with the luminal cells but is absent in fibroblasts and endothelial cells. In D492 cells, miR-203a expression is predominantly associated with the EpCAM High cells.

ii) MiR-203a expression in breast cell lines. MiR-203a is expressed in breast cell lines with an epithelial phenotype but not with a mesenchymal phenotype. Expression of miR-203a in mesenchymal derivatives of D492 and HMLE, *i.e.* D492M and HMLEmes, is down-regulated with little or no expression in the mesenchymal state. The normal like breast cell line MCF10A has high expression of miR-203a, while the triple negative mesenchymal breast cancer cell line MDA-MB-231 has no expression of miR-203a. Results are displayed as average of three independent experiments (mean ± sd, n = 3).



(caption on next page)

Fig. 4. Overexpression of miR-203a in D492M.
A)Overexpression of miR-203a in D492M. RT-qPCR showing 17,500-fold expression of miR-203a in D492M$^{miR\text{-}203a}$ compared to control.
B) MiR-203a increases adherence of D492M cells. Phase contrast images of monolayer cultures show that miR-203a overexpression in D492M causes cells to adhere more to each other compared to D492M^{Empty} control. Phase contrast images from 3D culture show that D492M$^{miR\text{-}203a}$ form more compact colonies, with some colonies showing protrusion of mesenchymal cells compared to control. Scale bar = 100 μm.
C) MiR-203a reduces proliferation of D492M. Reduced proliferation of D492M shown by staining with crystal violet from day 1–7 in monolayer culture. Results are shown as average of 4 replicates (mean ± SD)
D) DAB staining of monolayer cultures from D492M$^{miR\text{-}203a}$. D492M$^{miR\text{-}203a}$ cells display reduced expression of N-Cadherin but show little change in expression of E-Cadherin. Scale bar = 100 μm.
E) Western blot of E- and N-Cadherin in D492M$^{miR\text{-}203a}$. There is slightly less protein expression of N-Cadherin (10% reduction) but little change in expression of E-Cadherin in D492M$^{miR\text{-}203a}$ indicating reduced mesenchymal characteristics in D492M$^{miR\text{-}203a}$.
F) Expression of SNAI2 is reduced in D492M$^{miR\text{-}203a}$. MiR-203a, which has a known binding site in the 3′-UTR of SNAI2, significantly reduces expression of SNAI2 as determined by qPCR.
G) D492M$^{miR\text{-}203a}$ cells are more sensitive to chemically induced apoptosis. In D492M cells, miR-203a increases sensitivity to apoptosis induced by Camptothecin. Data is analyzed on IncuCyte Zoom and data is displayed as Caspase 3/7 object count/mm^2 (mean ± SEM).
H) Anchorage independent growth is reduced upon miR-203a overexpression. miR-203a reduces stem cell like properties in D492M. D492M$^{miR\text{-}203a}$ has reduced ability to generate mammospheres in low attachment assay compared to D492M^{Empty}.
I) MiR-203a reduces the ability of D492M cells to migrate and invade. D492M$^{miR\text{-}203a}$ has less ability to migrate through trans-well filter than D492M^{Empty} and D492M$^{miR\text{-}203a}$ has reduced capability to invade through Matrigel coated transwell filter. Data is shown as average number of cells per field (mean ± SEM).

203a was more associated with the luminal epithelial population in D492 (Fig. 3C, right). Differential expression of miR-203a between D492 and D492M was confirmed by RT-qPCR (Fig. 3C, right). To determine whether this was in concurrence with other breast cell lines with epithelial and mesenchymal phenotypes, we examined the expression levels of miR-203a in HMLE, its mesenchymal derivative HMLEmes (Fig. 3C, right), the normal cell-derived MCF10A, and the mesenchymal cancer cell line, MDA-MB-231 (Fig. 3C, right). Expression of miR-203a was restricted to all three cell lines with an epithelial phenotype.

2.4. MiR-203a expression reduces the mesenchymal traits of D492M cells

To analyze if the presence of miR-203a affects the phenotype of D492M, we overexpressed miR-203a in D492M (D492M$^{miR\text{-}203a}$), using a lentiviral based transfection system. D492M$^{miR\text{-}203a}$ expressed miR-203a 17,500-fold higher than control (D492M^{Empty}) (Fig. 4A). Interestingly, phase contrast images of monolayer cultures show that miR-203a overexpression in D492M causes increased adherence between cells (Fig. 4B). Furthermore, when D492M is cultivated in 3D-rBM, it forms spindle shaped mesenchymal-like colonies in contrast to D492M$^{miR\text{-}203a}$ that forms more compact colonies although the colonies still show cellular protrusions, typical of mesenchymal cells (Fig. 4B). When the cells were cultured in monolayer, overexpression of miR-203a in D492M reduces the proliferation rate of the cells (Fig. 4C). Furthermore, there was visible reduction in expression of the mesenchymal marker N-cadherin and increased expression of E-cadherin, K14 and K19 in D492M$^{miR\text{-}203a}$, indicating a moderate reduction in mesenchymal characteristics (Fig. 4D–E). The EMT transcription factor SNAI2 is a confirmed target of miR-203a, and miR-203a is in return suppressed by SNAI2 forming a negative feedback loop (Ding et al., 2013). In concordance with SNAI2 being a target of miR-203a, there was reduced expression of SNAI2 when miR-203a was overexpressed in D492M (Fig. 4F). Since the EMT phenotype has been associated with apoptosis resistance and SNAI2 is a known inhibitor of apoptosis (Inoue et al., 2002), we asked if D492M$^{miR\text{-}203a}$ cells were less resistant to chemically induced apoptosis compared to D492M^{Empty}. Indeed, D492M$^{miR\text{-}203a}$ cells were more sensitive than D492M^{Empty} to camptothecin, a chemical inducer of apoptosis (Fig. 4G).

The ability of cells to proliferate as mammospheres under non-adherent conditions has been described as a property of early progenitor/stem cells (Dontu et al., 2003). When D492 and D492M were cultured in low attachment plates both generated mammospheres but D492M generated significantly larger and a higher number of colonies (Sigurdsson et al., 2011). Interestingly, overexpression of miR-203a in D492M negatively impacts anchorage independent growth, resulting in loss of progenitor/stem cell properties indicating that miR-203a overexpression in D492M induces cellular differentiation (Fig. 4H). As mesenchymal cells have increased motility compared to epithelial cells, we analyzed the effect of miR-203a overexpression on the ability of D492M to migrate and invade. D492M$^{miR\text{-}203a}$ cells showed reduced ability to migrate through transwell filters (Fig. 4I, left) and to invade through Matrigel coated transwell filters (Fig. 4I, right) indicating a suppression of mesenchymal characteristics. Collectively, overexpressing miR-203a in D492M induced a reduction of mesenchymal characteristics, caused reduced proliferation, migration and invasion, and increased sensitivity towards chemically induced apoptosis.

2.5. Peroxidasin (PXDN) is a novel target of miR-203a

Due to the phenotypic differences between D492 and D492M in 3D culture and the fact that miR-203a is not expressed in D492M, we decided to compare the transcriptional profiling of D492, D492M and D492M$^{miR\text{-}203a}$, with emphasis on identifying novel targets of miR-203a. Herein, we identified peroxidasin (PXDN), a collagen IV cross-linking agent in the basement membrane, as the most significantly downregulated gene in D492M$^{miR\text{-}203a}$ compared to D492M (Fig. 5A). In addition, we identified three potential binding sites for miR-203a in the 3′-UTR of PXDN, one well conserved and two less conserved (Fig. 5B). To further investigate the potential interaction between miR-203a and PXDN we treated D492M and D492M$^{miR\text{-}203a}$ with miR-203a-mimic and miR-203a-inhibitor, respectively. PXDN expression was reduced (Fig. 5C, left) and increased (Fig. 5C, right) when D492M and D492M$^{miR\text{-}203a}$ were treated with miR-203a mimic and inhibitor, respectively, corroborating that miR-203a regulates PXDN expression. Furthermore, we demonstrated that PXDN expression was directly regulated by miR-203a by carrying out a dual luciferase reporter assay showing that the relative luciferase activity was significantly lower in the PXDN 3′-UTR containing the miR-203a target sequence compared with PXDN 3′-UTR with deleted miR-203a binding sequence, when transfected with miR-203a mimic (***$p < 0.001$) (Fig. 5D). This suggested that miR-203a could directly regulate the expression of PXDN through targeting its 3′-UTR. As a positive control we demonstrated significantly lower relative luciferase activity in p63 3′-UTR containing the miR-203a target sequence compared with p63 3′-UTR with deleted miR-203a binding sequence, when transfected with miR-203a mimic (**$p < 0.01$) (Fig. S1).

In order to investigate whether the observed effect of miR-203a overexpression on D492M is due to repression of PXDN, we knocked down PXDN in D492M using siRNA (Fig. S2) and analyzed the phenotypic changes in monolayer. As with miR-203a overexpression, we see reduced proliferation and increased sensitivity to chemically

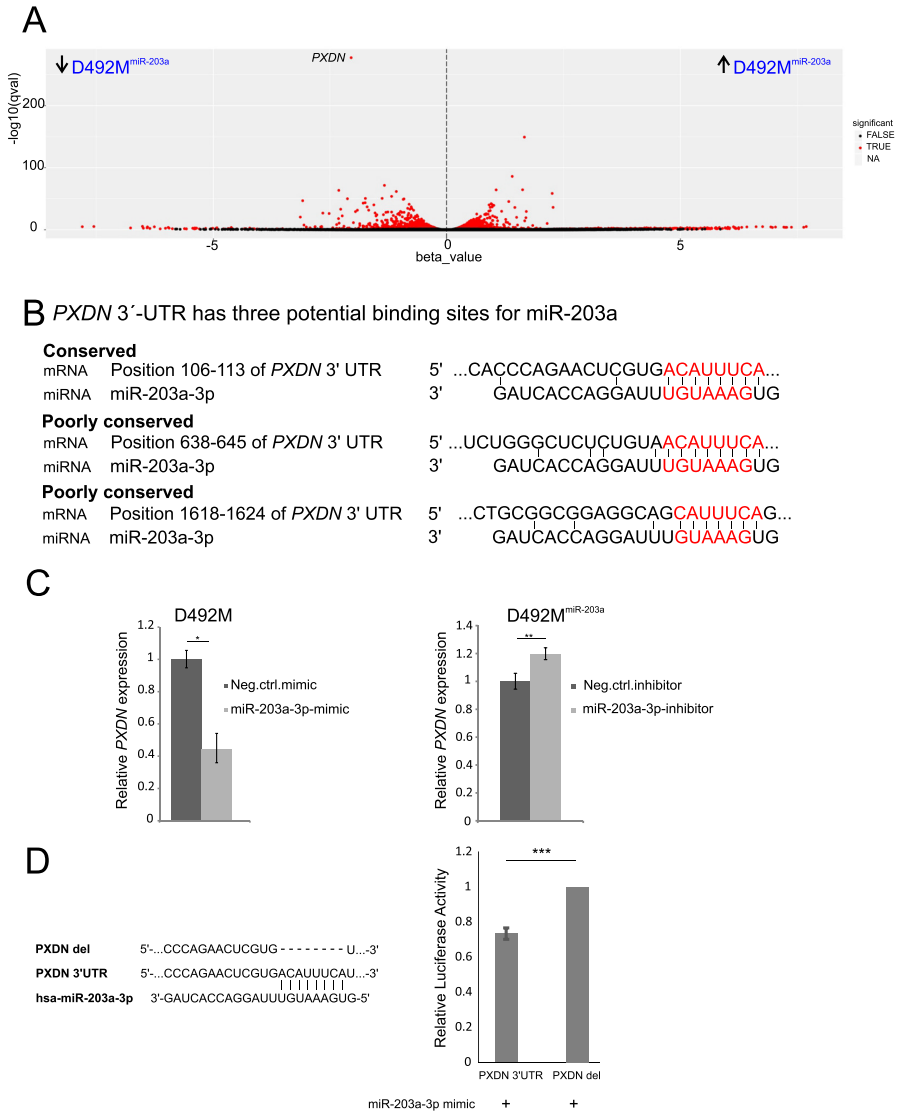


Fig. 5. Gene expression profiling of D492M overexpressing miR-203a.

A) Volcano plot showing differential expressed transcripts in D492M$^{miR\text{-}203a}$ and D492M, depicting statistical significance and high fold change in gene expression. Volcano plot was constructed from RNA sequencing data using beta and q-values from Sleuth analysis (max FDR = 0.05), enabling the visualization of the relationship between differentially expressed transcripts and statistical significance. More extreme values on the x-axis show increased differential expression and higher values on the y-axis show increased statistical significance. PXDN is the most downregulated gene when miR-203a is overexpressed in D492M.

B) PXDN 3′-UTR has three potential binding sites for miR-203a

Bioinformatics analysis reveals three potential binding sites for miR-203a in the 3′-UTR of PXDN a heme-containing peroxidase that is secreted into the extracellular matrix and is involved in extracellular matrix formation. One of the binding sites is conserved and two are poorly conserved.

C) MiR-203a regulates PXDN expression. D492M cells treated with miR-203a mimic have reduced expression of PXDN, whereas D492M$^{miR\text{-}203a}$ cells treated with miR-203a inhibitor have increased expression of PXDN. Cells were treated with miR-203a-3p mimic or inhibitor, respectively, and PXDN expression was measured by qPCR. Results are displayed as average of three independent experiments (mean ± sd, n = 3).

D) MiR-203a directly binds to its target sequence on PXDN 3′UTR. HEK293T cells transfected with miR-203a-3p mimic and pmirGLO plasmid containing the miR-203a binding site and surrounding sequence had reduced luciferase activity compared to cells transfected with the miR-203a-3p mimic and plasmid with the miR-203a binding site deleted, indicating binding of the miR-203a-3p mimic to the target sequence. Results are displayed as average of three independent experiments (mean ± SD, n = 3).

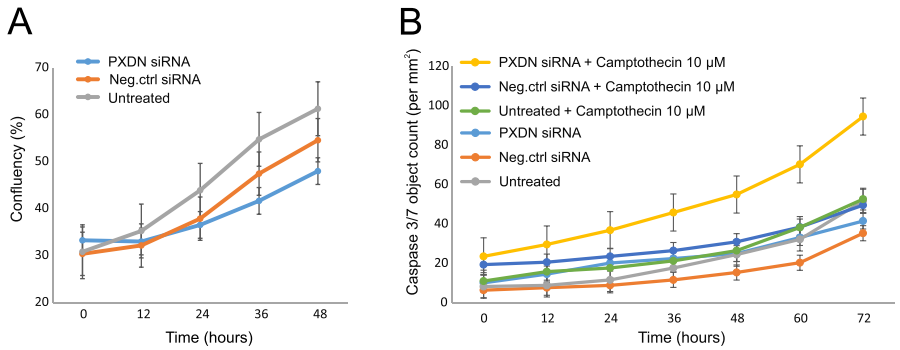


Fig. 6. PXDN repression reduces proliferation and increases sensitivity to chemically induced apoptosis.
A) PXDN silencing in D492M cells reduces proliferation. D492M cells treated with siRNA against PXDN have reduced proliferation rate compared to D492M cells treated with negative control siRNA or untreated D492M cells. Data is analyzed on IncuCyte Zoom and data is displayed as Confluency percentage (mean ± SEM). B) D492M cells with silenced PXDN are more sensitive to chemically induced apoptosis. Camptothecin induced apoptosis is increased in D492M cells with silenced PXDN, compared to cells treated with negative control siRNA and untreated D492M cells. Data is analyzed on IncuCyte Zoom and data is displayed as Caspase 3/7 object count/mm^2 (mean ± SEM).

induced apoptosis upon PXDN silencing (Fig. 6A & B), however no changes were seen in migration nor invasion (data not shown). This suggests that miR-203a mediates some of its effects, at least partly, through repression of PXDN.

Collectively, we have shown that miR-203a is a novel repressor of PXDN, a protein that plays an important role as a collagen IV cross-linking agent in the basement membrane and has been implicated in the EMT process, *i.e.* development, fibrosis and cancer.

3. Discussion

Modeling breast morphogenesis and EMT in 3D culture is an important tool to shed light on cellular and molecular mechanisms behind these developmental events and complements well *in vivo* models. Although animal models, in particular mouse models, have contributed significantly to our understanding of mammary gland morphogenesis and breast cancer, they cannot replace the necessity for using human cells due to the molecular and cellular differences in mammary gland biology between species (Dontu and Ince, 2015). Although, functionally working in similar ways during breast feeding, the mouse and human female mammary gland differ greatly. The human female breast gland is composed of ducts that end in terminal duct lobular units (TDLUs), surrounded by cellular-rich connective tissue. In contrast, the mouse mammary gland is composed of ducts that terminate in end buds, which are surrounded by adipose tissue. Thus, the histological comparison between the mouse and human mammary gland demonstrates large differences that need to be taken into account when studies are designed to explore the interactions between epithelial cells and stroma.

In this study, we have applied two isogenic breast cell lines, D492 and D492M, and 3D culture based on reconstituted basement membrane (rBM) to capture the phenotypic architecture of branching morphogenesis of breast epithelium and EMT, respectively. D492 is a breast epithelial progenitor cell line established from suprabasal epithelium of a normal breast gland (Gudjonsson et al., 2002; Villadsen et al., 2007; Sigurdsson et al., 2011). D492, was established by immortalizing cells from reduction mammoplasty using the E6 and E7 oncogenes from human papilloma virus 16. E6 and E7 target p53 and RB, respectively that without any doubt interfere with number of processes in the cells and this should be kept in mind when designing projects involving this cell line. However, D492 is non-tumorigenic, can generate both luminal and myoepithelial cells and in 3D culture generate structures reminiscent of TDLU-like structures *in vivo* and therefore remains a good tool for studying molecular mechanisms involved in these processes. D492M has a mesenchymal phenotype and is derived from D492 cells that have undergone EMT when co-cultured with breast endothelial cells (Sigurdsson et al., 2011).

In a previous work we demonstrated that overexpression of miR-200c-141 in D492M was sufficient to induce MET in D492M albeit only to luminal epithelial cells. When p63 was overexpressed in D492M containing miR-200c-141, bipotential and branching properties of D492 were restored (Hilmarsdottir et al., 2015). Recently, Wuidart et al. showed that p63 was essential for maintaining the unipotent basal fate of embryonic mouse mammary gland progenitors. Interestingly, sustained p63 expression in luminal epithelial cells reprogrammed cells towards basal cells (Wuidart et al., 2018). D492 has stem cell properties based on its ability to generate luminal and myoepithelial cells in culture and branching structures in 3D culture. It is however, possible that the immortalization using the E6 and E7 oncogenes from human papilloma virus 16 (Gudjonsson et al., 2002) and the 3D cell culture condition may contribute to the cellular plasticity of D492 and thus reflect more embryonic development of the mammary gland than postnatal development.

In this study, we analyzed differential expression of miRNAs in D492 and D492M on day 14 in 3D-rBM culture. Small RNA sequencing analysis revealed 47 differentially expressed miRNAs between D492 and D492M at day 14 in culture, of which 20 miRNAs were downregulated in D492M. Among the most profound changes, was the downregulation of the miR-200 family, miR-205 and miR-203a in D492M. We have previously shown that these miRNAs are also downregulated in D492M when cultured in monolayer (Hilmarsdottir et al., 2015) indicating a comparable expression pattern between 2D and 3D conditions. In the same study we demonstrated that when miR-141 and miR-200c, which are under regulation of the same promoter on chromosome 12 (Hilmarsdottir et al., 2014; Hilmarsdottir et al., 2015), were overexpressed in D492M, a reversal towards epithelial phenotype was induced, albeit only towards the luminal epithelial phenotype (Hilmarsdottir et al., 2015). Unlike miR-200c-141 and miR-205, which are both heavily methylated in D492M, we did not observe any methylation of CpG islands in the promoter area of miR-203a. The unmethylated pattern of the miR-203a promoter in D492M is opposite to what has been described in other breast cell lines with a mesenchymal phenotype (Zhang et al., 2011; Taube et al., 2013). This indicates that miR-203a down regulation in D492M may be through other processes, such as histone modification, or transcription factor repression.

Ectopic overexpression of miR-203a in D492M partially induced epithelial traits. The effect of increased miR-203a expression was

primarily evident in functional assays, where it caused reduced cell proliferation, migration, and invasion. Morphologic and gene expression changes were subtle and indicated only partial MET, as evidenced by reduced expression of N-cadherin, and SNAI2, an EMT transcription factor known to silence expression of miR-203a. Moes et al. demonstrated that overexpression of SNAI1 in MCF-7 breast cancer cells resulted in repression of miR-203 (Moes et al., 2012).

Branching morphogenesis of the D492 cells in 3D-rBM was used to mimic breast development and differentiation. We show that miR-203a expression coincides with increased differentiation status of the breast epithelial cells. In contrast, other epithelial associated miRNAs such as miR-205, miR-200c and miR-141 show constant or minimal fluctuation in expression throughout the branching period (21 days). In a miRNA expression study on mouse mammary gland development, Avril-Sassen et al. (2009) demonstrated that miR-203a showed temporal changes during different phases of mammary gland development. High expression was seen in early development and gestation followed by low expression in lactation and involution stages (Avril-Sassen et al., 2009). Reduced expression of miR-203a in lactation might be due to that fact that lactation requires functionally active p63 positive myoepithelial cells. MiR-203a is a repressor of p63 and therefore, downregulation of miR-203a is probably necessary for full activity of the myoepithelium.

Since D492 cells have stem cell properties, it is possible to separate cells into luminal epithelial and myoepithelial cells (Gudjonsson et al., 2002; Villadsen et al., 2007). Based on EpCAM, a luminal associated epithelial adhesion molecule, we separated D492 cells into two cell populations: EpCAM-high and EpCAM-low. EpCAM-high cells showed significantly higher expression of miR-203a than EpCAM-low cells. This was also confirmed in primary luminal epithelial- and myoepithelial cells. This result is in agreement with DeCastro et al. (DeCastro et al., 2013), where they demonstrated that miR-203a was significantly more expressed in mouse luminal epithelial progenitor cells and differentiated luminal epithelial cells compared to stem/basal cells. They also tested the expression levels of miR-203a in a number of normal and cancerous human breast epithelial cell lines: highest expression was predominantly in cell lines with a luminal epithelial phenotype.

To search for potential target genes for miR-203a we analyzed gene expression of D492M and D492M$^{miR\text{-}203a}$. The most significantly differential expressed gene between D492M and D492M$^{miR\text{-}203a}$ was PXDN. PXDN is a heme-containing peroxidase that produces hypobromous acid (HOBr) to form sulfilimine cross-links to stabilize the collagen IV network in the basement membrane and is believed to be important in normal development (Colon et al., 2017). PXDN has also been linked to EMT and disease conditions such as melanoma invasion and fibrosis (Tindall et al., 2005; Cheng et al., 2008; Peterfi et al., 2009; Liu et al., 2010; Tauber et al., 2010; Barnett et al., 2011; Khan et al., 2011; Bhave et al., 2012; Yan et al., 2014; Ero-Tolliver et al., 2015; Colon and Bhave, 2016; Jayachandran et al., 2016; Sitole and Mavri-Damelin, 2018).

Jayachandran et al. identified PXDN as consistently elevated in invasive mesenchymal-like melanoma cells and it was also found highly expressed in metastatic melanoma tumors. Gene silencing led to reduced melanoma invasion *in vitro* (Jayachandran et al., 2016). Moreover, Peterfi et al. have shown that PXDN is secreted from myofibroblasts and also from fibrotic kidney (Peterfi et al., 2009). Young et al. demonstrated that PXDN expression increases in PIK3CA mutant MCF10A cells and that the PXDN protein is mainly associated with secreted exosomes (Young et al., 2015). When knocking down PXDN in four basal like breast cancer cell lines, all cell lines showed reduced proliferation and when PXDN was knocked down in the luminal cell line MDA-MB-361, which expresses little PXDN, there was also reduced proliferation (Young et al., 2015). The reduced proliferation rate we saw in D492M$^{miR\text{-}203a}$, could be because of PXDN silencing by miR-203a, as we also see reduced proliferation when PXDN is transiently knocked down with silencing RNA. The fact that PXDN is overexpressed in the mesenchymal cell line D492M may link it to the EMT phenotype.

In summary, mammary gland development is mostly a post-natal process, where post-transcriptional regulation *via* miR-203a expression correlates with differentiation stages in the mammary epithelium. Using the D492 breast progenitor cells cultivated in 3D-rBM it is possible to capture the critical aspects of branching morphogenesis. Temporal changes in miR-203a expression during the *in vitro* branching process correlate well with *in vivo* conditions. Although miR-203a repression and overexpression are well documented in a number of studies, we observed only subtle changes towards MET when overexpressed in D492M. Furthermore, we have demonstrated a novel link between miR-203a and PXDN, which is highly expressed in D492M. Collectively, we conclude that miR-203a may be important to retain epithelial phenotype of breast epithelial cells, possibly through its target PXDN.

4. Material and methods

4.1. Cell culture

D492 and D492M cells were maintained in H14 medium as described previously (Sigurdsson et al., 2011). Primary luminal epithelial cells (EpCAM$^+$) and myoepithelial cells (EpCAM$^-$) were isolated from primary culture of breast epithelial cells derived from reduction mammoplasties by magnetic cell sorting (MACS) and maintained in CDM3 and CDM4, respectively, as previously described (Pechoux et al., 1999). Primary human breast endothelial cells (BRENCs) were isolated from breast reduction mammoplasties and cultured in endothelial growth medium (EGM) (Lonza) + 5% FBS (Invitrogen), referred to as EGM5 (Sigurdsson et al., 2006). Growth factor reduced reconstituted basement membrane (rBM, purchased as Matrigel, Corning #354230) was used for 3D cultures. 3D monocultures were carried out in 24-well culture plates (Corning). 1×10^4 D492 cells were suspended in 300 µl of rBM. Co-culture experiments were carried out with either 500, or 1×10^3 cells mixed with 1×10^5–2×10^5 BRENCs. 300 µl of mixed cells/rBM were seeded in each well of a 24-well plate and cultured on H14 (monoculture) or EGM5 (co-culture) for 14–21 days.

Branching, solid and spindle-like structures were isolated from 3D co-cultures with gentle shaking on ice in PBS - EDTA (5 mM) solution as previously described (Lee et al., 2007).

4.2. Small RNA sequencing

Total RNA was isolated from branching and spindle-like colonies from D492 and D492M, respectively using Tri-Reagent (Thermo Fisher Scientific, #AM9738). For D492 branching time points, RNA was isolated from 3D-rBM culture on days 7, 14 and 21 but for D492M RNA was isolated only on day 14. Samples were pooled in triplicates from each time point before small RNA library preparation. Small RNA libraries were prepared using the TruSeq Small RNA Library Kit from Illumina (#RS-200-0012) per manufacturer's protocol. The small RNA libraries were then sequenced using the Illumina MiSeq platform and V2 sequencing chemistry. FASTQ files were generated with MiSeq Reporter (Illumina, San Diego, US-CA). Small RNA sequence analysis was performed using the CLC Genomics Workbench (CLC Bio-Qiagen, Aarhus, Denmark) and miRBase – release 21 was used for annotation. Samples were normalized by totals and counts reported as reads per million. Reads above 29 nt and below 15 nt in length were discarded. Proportion-based statistical analysis was done using the test of Kal (Kal et al., 1999). Hierarchal clustering of features was performed using Log2 transformed expression values, Euclidean distance and single linkage. For the D492 and D492M comparison samples from day 7, 14 and 21 from D492 were compared to D492M sample from day 14.

4.3. DNA isolation and methylation bead chip array

D492 (1×10^4 cells) and D492M (2.5×10^4 cells) were grown in

3D-rBM in triplicate in a 24-well plate for 14 days and colonies extracted from 3D-rBM with gentle shaking on ice in PBS - EDTA (5 mM) solution as previously described (Lee et al., 2007). DNA was extracted using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific, #K182002) and DNA was bisulfite converted using the EZ-96 DNA Methylation-Gold Kit (Zymo Research, #D5007) per manufacturer's protocols. The samples were hybridized to the Infinium HumanMethylation450 BeadChip array (Illumina, # WG-314-1003). Data was analyzed using the minfi Bioconductor package (Aryee et al., 2014).

4.4. Bisulfite sequencing

DNA (0.5–1 μg) was bisulfite converted using the EpiTect Bisulfite Kit (Qiagen, #59104). Target DNA sequences were amplified using nested PCR (see primers in Table S1). Methylation levels were analyzed by sequencing the bisulfite modified promoter regions on a 3130 Genetic Analyzer (Applied Biosystems). Methylation data from bisulfite sequencing was analyzed and visualized using the BiQ Analyzer v2.0 (Bock et al., 2005).

4.5. Quantitative reverse transcription PCR analysis

Total RNA was extracted with Tri-Reagent (Thermo Fisher Scientific, #AM9738) and

reverse transcription performed using random hexamers (Thermo Fisher Scientific, #N8080127) and SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, #18090050). The following primers were used for mRNA qRT-PCR analysis, SNAI2 (Hs00950344_m1) (Thermo Fisher Scientific, #4331182) and GAPDH as endogenous reference gene (Thermo Fisher Scientific, # 4326317E). Maxima Probe/ROX qPCR Master Mix (2×) (Thermo Fisher Scientific, # K0231) was used for TaqMan qRT-PCR analysis.

Quantitative RT-PCR analysis of miRNAs was performed using the universal cDNA synthesis kit II (Exiqon, #203301) and ExiLENT SYBR Green master mix (Exiqon, #203402). The following primer sets from Exiqon were used for miRNA qRT-PCR analysis, hsa-miR-203a (#205914), hsa-miR-141-3p (#204504), hsa-miR-200c-3p (#204482), hsa-miR-205-5p (#204487) and U6 snRNA (#203907) was used as endogenous reference. All qRT-PCRs were performed on the Applied Biosystems 7500 Real-Time PCR system and relative expression differences were calculated with the $2^{\Delta Ct}$ method.

4.6. Cloning of miR-203a into pCDH lentivector

The miR-203a miRNA construct was amplified from D492 genomic DNA using nested PCR with the following outer primers, miR-203a-outer-F 5′-ATCAGTCGCGGGACCTATG-3′ and miR-203a-outer-R 5′-GAATTCCACGGAGTTTCGAG-3′. From miR-203a-outer amplicon, *Eco*RI and *Not*I restriction sites were incorporated with PCR using the following inner primers, miR-203a-EcoRI-F-5′-TAAGCAGAA TTCaggcgagggcgtctaagg-3′ and miR-203a-NotI-R-5′-TGCTTAGCGGC CGCacctcccagcagcacttg-3′. Phusion High-Fidelity DNA Polymerase (NEB, # M0530S) was used for PCR and amplicons were purified using the GeneJET PCR purification kit (Thermo Fisher Scientific, #K0701). Double digestions of miR-203a-inner amplicon and pCDH vector (System Biosciences, #CD516B-2) were performed using EcoRI (NEB, #R0101) and NotI (Thermo Fisher Scientific, #ER0591). The miR-203a-inner amplicon was cloned into pCDH lentivector at insert-to-vector molar ratio 10:1 using T4 DNA ligase (Thermo Fisher Scientific, #15224041). Empty pCDH lentivector and miR-203a-pCDH lentivector were transformed into *E. coli* DH5alpha competent cells and inserts confirmed with colony PCR. Lentivectors were produced in cultures of DH5alpha and isolated using GeneJET Plasmid Miniprep kit (Thermo Fisher Scientific, #K0502). The cloned miR-203a insert sequence was then confirmed with sequencing. The pCDH lentivector has a RFP +Puro fusion containing the T2A element to enable co-expression of balanced levels of RFP and Puro genes. Viral particles were produced in HEK-293 T cells using TurboFect transfection reagent (Thermo Fisher Scientific, #R0531) and virus containing supernatant collected after 48 and 72 h, centrifuged and filtered through 0.45 μm filter. Target cells were transfected with virus titer in the presence of 8 μg/μl polybrene. Stable cell lines and control (empty-lentivector) cells were isolated with puromycin (2 μg/ml) (Thermo Fisher Scientific, # A1113803) followed by flow-sorting (Sony SH800), selecting for RFP expressing cells.

4.7. Immunochemistry

The following primary antibodies were used for DAB staining (Dako, # K3467), E-cadherin (BD, # 610182), *N*-cadherin (BD, # 610921), K14 (Abcam, #ab7800) and K19 (Abcam, #ab7754). Specimens were visualized on a Leica DMI3000 B inverted microscope.

4.8. Western blotting

Equal amounts (5 μg) of proteins in RIPA buffer were separated on NuPAGE™ 10% Bis-Tris Protein Gels (Thermo Fisher Scientific, #NP0302BOX) and transferred to a PVDF membrane (Millipore, #IPFL00010). Antibodies: SNAI2 (Cell Signaling, #9585) and Histone H3 (Cell Signaling, #4499). Secondary antibodies were mouse or rabbit IRDey (Li-Cor) used at 1:20,000 and detected using the Odyssey Infrared Imaging System (Li-Cor). Fluorescent images were converted to gray scale.

4.9. Proliferation assay

Cells were seeded in triplicates in 24-well plates and cultures stopped every 24 h. Cells were fixed in 3.7% formaldehyde in PBS for 10 min, washed once with 1 x PBS, stained with 0.1% crystal violet in 10% ethanol for 15 min, washed four times with water and dried. Density of cells was evaluated by extracting the crystal violet stain in 10% acetic acid and measuring optical density at 595 nm using a spectrometer. In addition, proliferation rate of D492M cells treated with siRNA was analyzed on IncuCyte Zoom (Essen Bioscience) per manufacturer's instructions.

4.10. Apoptosis assay

Resistance to chemically induced apoptosis with 10 μM camptothecin (Sigma-Aldrich, #C9911) was determined using IncuCyte Caspase-3/7 Reagents (Essen Bioscience, #4440) and imaging on IncuCyte Zoom (Essen Bioscience) per manufacturer's instructions.

4.11. Anchorage independence, migration and invasion assays

Anchorage independent growth was determined using 24-well ultra-low attachment plates (Corning, #3473), where triplicates of 500 cells of D492M$^{miR\text{-}203a}$ and D492M^{Empty} were single cell filtered and cultured using EGM5 medium for 9 days.

For migration analysis, triplicates of 10,000 starved D492M$^{miR\text{-}203a}$ and D492M^{Empty} cells were seeded in DMEM/F12, HEPES medium (Thermo Scientific, #31330038) on collagen I (Advanced BioMatrix, #5005-B) coated transwell filters with 8 μm pore size (Corning, #353097) with EGM5 medium in the lower chamber and incubated for 24 h. Filters were rinsed with 1 × PBS and cells on the apical layer wiped off with a cotton swab (Q-tip) and migrated cells fixed in 3.7% formaldehyde in PBS for 10 min, washed with 1 × PBS, stained with 0.1% crystal violet in 10% ethanol for 15 min and washed with water and then dried. Images were acquired and migrated cells counted, 3 images per filter.

Invasion assay was performed using transwell filters with 8 μm pore size (Corning, #353097) that were coated with 100 μl diluted Matrigel

(Corning, #354230) 1:10 in H14 media. D492M$^{miR-203a}$ and D492M^{Empty} (25,000 cells) were seeded in H14 media on top of Matrigel coated filters and H14 + 5% FBS added to the lower chamber and incubated for 44 h. Matrigel was then removed with a cotton swab and washed with PBS. Cells were fixed in 3.7% formaldehyde in PBS for 15 min and washed four times with water. Images were acquired and migrated cells counted, 3 images per filter.

4.12. PolyA mRNA sequencing and analysis

RNA was isolated from 3D and 2D cultures using the Exiqon miRCURY RNA Isolation Kit - Cell and Plant and quality control of RNA samples was performed using BioAnalyzer. Libraries were prepared using polyA mRNA library kit from Illumina and sequenced on a HiSeq sequencer from Illumina. Alignment of reads and annotation was performed using Kallisto (Bray et al., 2016) and differential expression analysis was done using Sleuth (Pimentel et al., 2017).

4.13. Transient transfection with miR-203a mimic

Briefly, D492M cells were separately transfected with 50 pmol of mirVANA miR-203a-3p mimic (Thermo Fisher, #4464066, Assay ID MC10152) and miRNA mimic negative control #1 (Thermo Fisher, #4464058) using RNAiMAX (Thermo Fisher, #13778075) per manufacturer's instructions for mirVana miRNA mimics.

4.14. Transient transfection with miR-203a inhibitor

Briefly, D492M$^{miR-203a}$ cells were separately transfected with 50 pmol of mirVANA miR-203a-3p inhibitor (Thermo Fisher, # 4464084, Assay ID MH10152) and mirVana miRNA Inhibitor, Negative Control #1 (Thermo Fisher, # 4464076) using RNAiMAX (Thermo Fisher, #13778075) per manufacturer's instructions for mirVana miRNA inhibitors.

4.15. Plasmid vector constructs and Luciferase activity assay

Synthetic oligonucleotides containing the hsa-miR203a-3p target sequence of human PXDN 3′-UTR (Position 106–113) or a deletion thereof (Table 1) were cloned into pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega Corporation, Madison, WI, USA). As a positive control we cloned the target sequence of p63 3′-UTR, a known hsa-miR203a-3p target, into pmirGLO, as well as a mismatched version of the p63 target site or a deletion thereof (Fig. S1). Correct sequence and orientation was verified by DNA sequencing (Eurofins Genomics, Ebersberg, Germany).

HEK293T cells were plated in a 96-well plate 3.0×10^4per well and incubated overnight. Cells were first transfected with hsa-miR203a-3p mimics (mirVana™ miRNA Mimics, Thermo Fisher Scientific, #4464070) with final concentration of 100 nM, using Lipofectamine RNAiMAX (Thermo Fisher Scienttific, #13778150) transfection reagent (according to manufacturer's instructions, with minor changes: using serum free and antibiotic-free high glucose DMEM medium instead of Opti-MEM® Medium). 24 h after transfection with miRNA mimics, cells were transfected (according to manufacturer's instructions) with 200 ng/well luciferase plasmids (pmirGLO constructs), using Lipofectamine 3000 (Thermo Fisher Scientific, #L3000015). After 24 h, plasmid-transfection cells were analyzed for luciferase activity using the Dual-Glo® Luciferase Assay System (Promega Corporation, Madison, WI, USA). Results are displayed as normalized firefly luciferase activity (background subtracted firefly luciferase activity/background subtracted renilla luciferase activity) for each construct. For each transfection, luciferase activity was averaged from four replicates, and three independent experiments were performed.

4.16. Transient transfection with siRNA

D492M cells were transfected with 10 nm final concentration of negative control siRNA (SilencerSelect siRNA #4390843, Ambion) and siRNA targeting PXDN (SilencerSelect siRNA #4427037, Ambion), using Lipofectamine RNAiMAX (Thermo Fisher Scienttific, #13778150), according to the manufacturer's protocol. Cells were incubated for 48 h and then the knockdown was confirmed by qRT-PCR using the following primers; Hs.PT.58.630748 (PrimeTime, IDT).

4.17. Statistics

All analyses comprised at least three independent experiments. Two-tailed student *t*-test was used to test significance ($p < 0.05$).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mod.2018.11.002.

Acknowledgements

We thank Erik Knutsen and Isac Lee for their contribution to this work.

Competing interests

No competing interests declared.

Funding

This work was supported by Grants from Landspitali University Hospital Science Fund, University of Iceland Research Fund, and Icelandic Science and Technology Policy - Grant of Excellence: 152144051. 'Göngum saman', a supporting group for breast cancer research in Iceland (www.gongumsaman.is). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. Primary cells were received from reduction mammoplasty after acquiring informed consent from the donor. Approved by the Icelandic National Bioethics Committee VSN-13-057.

Data availability

NCBI GEO: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112306

The data discussed in this publication have been deposited in NCBI's

Table 1
Synthetic oligonucleotides containing the hsa-mir-203a-3p target sequence of human PXDN 3′-UTR/p63 3′-UTR or a deletion thereof.

PXDN 3′-UTR sense	/5Phos/AAACTAGCGGCCGCTAGTCCCAGAACTCGTGACATTTCATT
PXDN 3′-UTR antisense	/5Phos/CTAGAATGAAATGTCACGAGTTCTGGGACTAGCGGCCGCTAGTTT
PXDN del sense	/5Phos/AAACTAGCGGCCGCTAGTCCCAGAACTCGTGTT
PXDN del antisense	/5Phos/CTAGAACACGAGTTCTGGGACTAGCGGCCGCTAGTTT
p63 3′-UTR sense	/5Phos/AAACTAGCGGCCGCTAGTGAATGAGTCCTTGATTTCAAAT
p63 3′-UTR antisense	/5Phos/CTAGATTTGAAATCAAGGACTCATTCACTAGCGGCCGCTAGTTT
p63 del sense	/5Phos/AAACTAGCGGCCGCTAGTGAATGAGTCCTTGAT
p63 del antisense	/5Phos/CTAGATCAAGGACTCATTCACTAGCGGCCGCTAGTTT

Gene Expression Omnibus (Edgar et al., 2002) and are accessible through GEO Series accession number GSE112306.

(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE112306).

References

Aryee, M.J., Jaffe, A.E., Corrada-Bravo, H., Ladd-Acosta, C., Feinberg, A.P., Hansen, K.D., Irizarry, R.A., 2014. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. Bioinformatics 30 (10), 1363–1369.

Avril-Sassen, S., Goldstein, L.D., Stingl, J., Blenkiron, C., Le Quesne, J., Spiteri, I., Karagavriilidou, K., Watson, C.J., Tavare, S., Miska, E.A., Caldas, C., 2009. Characterisation of microRNA expression in post-natal mouse mammary gland development. BMC Genomics 10, 548.

Barnett, P., Arnold, R.S., Mezencev, R., Chung, L.W., Zayzafoon, M., Odero-Marah, V., 2011. Snail-mediated regulation of reactive oxygen species in ARCaP human prostate cancer cells. Biochem. Biophys. Res. Commun. 404 (1), 34–39.

Bhave, G., Cummings, C.F., Vanacore, R.M., Kumagai-Cresse, C., Ero-Tolliver, I.A., Rafi, M., Kang, J.S., Pedchenko, V., Fessler, L.I., Fessler, J.H., Hudson, B.G., 2012. Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. Nat. Chem. Biol. 8 (9), 784–790.

Bock, C., Reither, S., Mikeska, T., Paulsen, M., Walter, J., Lengauer, T., 2005. BiQ analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. Bioinformatics 21 (21), 4067–4068.

Bray, N.L., Pimentel, H., Melsted, P., Pachter, L., 2016. Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. 34 (5), 525–527.

Cao, Q., Mani, R.S., Ateeq, B., Dhanasekaran, S.M., Asangani, I.A., Prensner, J.R., Kim, J.H., Brenner, J.C., Jing, X., Cao, X., Wang, R., Li, Y., Dahiya, A., Wang, L., Pandhi, M., Lonigro, R.J., Wu, Y.M., Tomlins, S.A., Palanisamy, N., Qin, Z., Yu, J., Maher, C.A., Varambally, S., Chinnaiyan, A.M., 2011. Coordinated regulation of polycomb group complexes through microRNAs in cancer. Cancer Cell 20 (2), 187–199.

Cheng, G., Salerno, J.C., Cao, Z., Pagano, P.J., Lambeth, J.D., 2008. Identification and characterization of VPO1, a new animal heme-containing peroxidase. Free Radic. Biol. Med. 45 (12), 1682–1694.

Colon, S., Bhave, G., 2016. Proprotein convertase processing enhances Peroxidasin activity to reinforce collagen IV. J. Biol. Chem. 291 (46), 24009–24016.

Colon, S., Page-McCaw, P., Bhave, G., 2017. Role of Hypohalous acids in basement membrane homeostasis. Antioxid. Redox Signal. 27 (12), 839–854.

DeCastro, A.J., Dunphy, K.A., Hutchinson, J., Balboni, A.L., Cherukuri, P., Jerry, D.J., DiRenzo, J., 2013. MiR203 mediates subversion of stem cell properties during mammary epithelial differentiation via repression of DeltaNP63alpha and promotes mesenchymal-to-epithelial transition. Cell Death Dis. 4, e514.

Ding, X., Park, S.I., McCauley, L.K., Wang, C.Y., 2013. Signaling between transforming growth factor beta (TGF-beta) and transcription factor SNAI2 represses expression of microRNA miR-203 to promote epithelial-mesenchymal transition and tumor metastasis. J. Biol. Chem. 288 (15), 10241–10253.

Dontu, G., Ince, T.A., 2015. Of mice and women: a comparative tissue biology perspective of breast stem cells and differentiation. J. Mammary Gland Biol. Neoplasia 20 (1–2), 51–62.

Dontu, G., Abdallah, W.M., Foley, J.M., Jackson, K.W., Clarke, M.F., Kawamura, M.J., Wicha, M.S., 2003. In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. Genes Dev. 17 (10), 1253–1270.

Edgar, R., Domrachev, M., Lash, A.E., 2002. Gene expression omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30 (1), 207–210.

Ero-Tolliver, I.A., Hudson, B.G., Bhave, G., 2015. The ancient immunoglobulin domains of peroxidasin are required to form sulfilimine cross-links in collagen IV. J. Biol. Chem. 290 (35), 21741–21748.

Feng, X., Wang, Z., Fillmore, R., Xi, Y., 2014. MiR-200, a new star miRNA in human cancer. Cancer Lett. 344 (2), 166–173.

Gregory, P.A., Bert, A.G., Paterson, E.L., Barry, S.C., Tsykin, A., Farshid, G., Vadas, M.A., Khew-Goodall, Y., Goodall, G.J., 2008. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat. Cell Biol. 10 (5), 593–601.

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.

Hanahan, D., Weinberg, R.A., 2011. Hallmarks of cancer: the next generation. Cell 144 (5), 646–674.

Hilmarsdottir, B., Briem, E., Bergthorsson, J.T., Magnusson, M.K., Gudjonsson, T., 2014. Functional role of the microRNA-200 family in breast morphogenesis and neoplasia. Genes 5 (3), 804–820.

Hilmarsdottir, B., Briem, E., Sigurdsson, V., Franzdottir, S.R., Ringner, M., Arason, A.J., Bergthorsson, J.T., Magnusson, M.K., Gudjonsson, T., 2015. MicroRNA-200c-141 and Np63 are required for breast epithelial differentiation and branching morphogenesis. Dev. Biol. 403 (2), 150–161.

Inoue, A., Seidel, M.G., Wu, W., Kamizono, S., Ferrando, A.A., Bronson, R.T., Iwasaki, H., Akashi, K., Morimoto, A., Hitzler, J.K., Pestina, T.I., Jackson, C.W., Tanaka, R., Chong, M.J., McKinnon, P.J., Inukai, T., Grosveld, G.C., Look, A.T., 2002. Slug, a highly conserved zinc finger transcriptional repressor, protects hematopoietic progenitor cells from radiation-induced apoptosis in vivo. Cancer Cell 2 (4), 279–288.

Javed, A., Lteif, A., 2013. Development of the human breast. Semin. Plast. Surg. 27 (1), 5–12.

Jayachandran, A., Prithviraj, P., Lo, P.H., Walkiewicz, M., Anaka, M., Woods, B.L., Tan, B., Behren, A., Cebon, J., McKeown, S.J., 2016. Identifying and targeting determinants of melanoma cellular invasion. Oncotarget 7 (27), 41186–41202.

Kal, A.J., van Zonneveld, A.J., Benes, V., van den Berg, M., Koerkamp, M.G., Albermann, K., Strack, N., Ruijter, J.M., Richter, A., Dujon, B., Ansorge, W., Tabak, H.F., 1999. Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources. Mol. Biol. Cell 10 (6), 1859–1872.

Khan, K., Rudkin, A., Parry, D.A., Burdon, K.P., McKibbin, M., Logan, C.V., Abdelhamed, Z.I., Muecke, J.S., Fernandez-Fuentes, N., Laurie, K.J., Shires, M., Fogarty, R., Carr, I.M., Poulter, J.A., Morgan, J.E., Mohamed, M.D., Jafri, H., Raashid, Y., Meng, N., Piseth, H., Toomes, C., Casson, R.J., Taylor, G.R., Hammerton, M., Sheridan, E., Johnson, C.A., Inglehearn, C.F., Craig, J.E., Ali, M., 2011. Homozygous mutations in PXDN cause congenital cataract, corneal opacity, and developmental glaucoma. Am. J. Hum. Genet. 89 (3), 464–473.

Lawson, D.A., Werb, Z., Zong, Y., Goldstein, A.S., 2015. The cleared mammary fat pad transplantation assay for mammary epithelial organogenesis. Cold Spring Harb Protoc 2015 (12) (pdb.prot078071).

Lee, G.Y., Kenny, P.A., Lee, E.H., Bissell, M.J., 2007. Three-dimensional culture models of normal and malignant breast epithelial cells. Nat. Methods 4 (4), 359–365.

Lilja, A.M., Rodilla, V., Huyghe, M., Hannezo, E., Landragin, C., Renaud, O., Leroy, O., Rulands, S., Simons, B.D., Fre, S., 2018. Clonal analysis of Notch1-expressing cells reveals the existence of unipotent stem cells that retain long-term plasticity in the embryonic mammary gland. Nat. Cell Biol. 20 (6), 677–687.

Liu, Y., Carson-Walter, E.B., Cooper, A., Winans, B.N., Johnson, M.D., Walter, K.A., 2010. Vascular gene expression patterns are conserved in primary and metastatic brain tumors. J. Neuro-Oncol. 99 (1), 13–24.

Lloyd-Lewis, B., Davis, F.M., Harris, O.B., Hitchcock, J.R., Watson, C.J., 2018. Neutral lineage tracing of proliferative embryonic and adult mammary stem/progenitor cells. Development 145 (14).

Mani, S.A., Guo, W., Liao, M.J., Eaton, E.N., Ayyanan, A., Zhou, A.Y., Brooks, M., Reinhard, F., Zhang, C.C., Shipitsin, M., Campbell, L.L., Polyak, K., Brisken, C., Yang, J., Weinberg, R.A., 2008. The epithelial-mesenchymal transition generates cells with properties of stem cells. Cell 133 (4), 704–715.

Moes, M., Le Bechec, A., Crespo, I., Laurini, C., Halavatyi, A., Vetter, G., Del Sol, A., Friederich, E., 2012. A novel network integrating a miRNA-203/SNAI1 feedback loop which regulates epithelial to mesenchymal transition. PLoS One 7 (4), e35440.

Moustakas, A., Heldin, C.H., 2007. Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. Cancer Sci. 98 (10), 1512–1520.

Olson, P., Lu, J., Zhang, H., Shai, A., Chun, M.G., Wang, Y., Libutti, S.K., Nakakura, E.K., Golub, T.R., Hanahan, D., 2009. MicroRNA dynamics in the stages of tumorigenesis correlate with hallmark capabilities of cancer. Genes Dev. 23 (18), 2152–2165.

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. Nat. Commun. 8 (1), 1627.

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.

Peinado, H., Olmeda, D., Cano, A., 2007. Snail, ZEB and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat. Rev. Cancer 7 (6), 415–428.

Peterfi, Z., Donko, A., Orient, A., Sum, A., Prokai, A., Molnar, B., Vereb, Z., Rajnavolgyi, E., Kovacs, K.J., Muller, V., Szabo, A.J., Geiszt, M., 2009. Peroxidasin is secreted and incorporated into the extracellular matrix of myofibroblasts and fibrotic kidney. Am. J. Pathol. 175 (2), 725–735.

Petersen, O.W., Polyak, K., 2010. Stem cells in the human breast. Cold Spring Harb. Perspect. Biol. 2 (5), a003160.

Petersen, O.W., Lind Nielsen, H., Gudjonsson, T., Villadsen, R., Ronnov-Jessen, L., Bissell, M.J., 2001. The plasticity of human breast carcinoma cells is more than epithelial to mesenchymal conversion. Breast Cancer Res. 3 (4), 213–217.

Pimentel, H., Bray, N.L., Puente, S., Melsted, P., Pachter, L., 2017. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods 14 (7), 687–690.

Sarrio, D., Rodriguez-Pinilla, S.M., Hardisson, D., Cano, A., Moreno-Bueno, G., Palacios, J., 2008. Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. Cancer Res. 68 (4), 989–997.

Shimono, Y., Zabala, M., Cho, R.W., Lobo, N., Dalerba, P., Qian, D., Diehn, M., Liu, H., Panula, S.P., Chiao, E., Dirbas, F.M., Somlo, G., Pera, R.A., Lao, K., Clarke, M.F., 2009. Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell 138 (3), 592–603.

Sigurdsson, V., Fridriksdottir, A.J., Kjartansson, J., Jonasson, J.G., Steinarsdottir, M., Petersen, O.W., Ogmundsdottir, H.M., Gudjonsson, T., 2006. Human breast microvascular endothelial cells retain phenotypic traits in long-term finite life span culture. In Vitro Cell. Dev. Biol. Anim. 42 (10), 332–340.

Sigurdsson, V., Hilmarsdottir, B., Sigmundsdottir, H., Fridriksdottir, A.J., Ringner, M., Villadsen, R., Borg, A., Agnarsson, B.A., Petersen, O.W., Magnusson, M.K., Gudjonsson, T., 2011. Endothelial induced EMT in breast epithelial cells with stem cell properties. PLoS One 6 (9), e23833.

Sitole, B.N., Mavri-Damelin, D., 2018. Peroxidasin is regulated by the epithelial-mesenchymal transition master transcription factor Snai1. Gene 646, 195–202.

Taube, J.H., Malouf, G.G., Lu, E., Sphyris, N., Vijay, V., Ramachandran, P.P., Ueno, K.R., Gaur, S., Nicoloso, M.S., Rossi, S., Herschkowitz, J.I., Rosen, J.M., Issa, J.P., Calin, G.A., Chang, J.T., Mani, S.A., 2013. Epigenetic silencing of microRNA-203 is required for EMT and cancer stem cell properties. Sci. Rep. 3, 2687.

Tauber, S., Jais, A., Jeitler, M., Haider, S., Husa, J., Lindroos, J., Knofler, M., Mayerhofer, M., Pehamberger, H., Wagner, O., Bilban, M., 2010. Transcriptome analysis of human

cancer reveals a functional role of heme oxygenase-1 in tumor cell adhesion. Mol. Cancer 9, 200.
Tindall, A.J., Pownall, M.E., Morris, I.D., Isaacs, H.V., 2005. Xenopus tropicalis peroxidasin gene is expressed within the developing neural tube and pronephric kidney. Dev. Dyn. 232 (2), 377–384.
Vidi, P.A., Bissell, M.J., Lelievre, S.A., 2013. Three-dimensional culture of human breast epithelial cells: the how and the why. Methods Mol. Biol. 945, 193–219.
Villadsen, R., Fridriksdottir, A.J., Ronnov-Jessen, L., Gudjonsson, T., Rank, F., LaBarge, M.A., Bissell, M.J., Petersen, O.W., 2007. Evidence for a stem cell hierarchy in the adult human breast. J. Cell Biol. 177 (1), 87–101.
Wellner, U., Schubert, J., Burk, U.C., Schmalhofer, O., Zhu, F., Sonntag, A., Waldvogel, B., Vannier, C., Darling, D., zur Hausen, A., Brunton, V.G., Morton, J., Sansom, O., Schuler, J., Stemmler, M.P., Herzberger, C., Hopt, U., Keck, T., Brabletz, S., Brabletz, T., 2009. The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. Nat. Cell Biol. 11 (12), 1487–1495.
Wronski, A., Arendt, L.M., Kuperwasser, C., 2015. Humanization of the mouse mammary gland. Methods Mol. Biol. 1293, 173–186.
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. Nat. Cell Biol. 20 (6), 666–676.
Yan, X., Sabrautzki, S., Horsch, M., Fuchs, H., Gailus-Durner, V., Beckers, J., Hrabe de Angelis, M., Graw, J., 2014. Peroxidasin is essential for eye development in the mouse. Hum. Mol. Genet. 23 (21), 5597–5614.
Young, C.D., Zimmerman, L.J., Hoshino, D., Formisano, L., Hanker, A.B., Gatza, M.L., Morrison, M.M., Moore, P.D., Whitwell, C.A., Dave, B., Stricker, T., Bhola, N.E., Silva, G.O., Patel, P., Brantley-Sieders, D.M., Levin, M., Horiates, M., Palma, N.A., Wang, K., Stephens, P.J., Perou, C.M., Weaver, A.M., O'Shaughnessy, J.A., Chang, J.C., Park, B.H., Liebler, D.C., Cook, R.S., Arteaga, C.L., 2015. Activating PIK3CA mutations induce an epidermal growth factor receptor (EGFR)/extracellular signal-regulated kinase (ERK) paracrine signaling Axis in basal-like breast Cancer. Mol. Cell. Proteomics 14 (7), 1959–1976.
Zhang, Z., Zhang, B., Li, W., Fu, L., Fu, L., Zhu, Z., Dong, J.T., 2011. Epigenetic silencing of miR-203 upregulates SNAI2 and contributes to the invasiveness of malignant breast Cancer cells. Genes Cancer 2 (8), 782–791.

Article II

Journal of Mammary Gland Biology and Neoplasia (2021) 26:321–338
https://doi.org/10.1007/s10911-021-09507-1

ORIGINAL PAPER

Peroxidasin Enhances Basal Phenotype and Inhibits Branching Morphogenesis in Breast Epithelial Progenitor Cell Line D492

Anna Karen Sigurdardottir[1] · Arna Steinunn Jonasdottir[1] · Arni Asbjarnarson[1] · Hildur Run Helgudottir[1] · Thorarinn Gudjonsson[1,2] · Gunnhildur Asta Traustadottir[1]

Received: 3 May 2021 / Accepted: 13 December 2021 / Published online: 28 December 2021

Abstract
The human breast is composed of terminal duct lobular units (TDLUs) that are surrounded by stroma. In the TDLUs, basement membrane separates the stroma from the epithelial compartment, which is divided into an inner layer of luminal epithelial cells and an outer layer of myoepithelial cells. Stem cells and progenitor cells also reside within the epithelium and drive a continuous cycle of gland remodelling that occurs throughout the reproductive period. D492 is an epithelial cell line originally isolated from the stem cell population of the breast and generates both luminal and myoepithelial cells in culture. When D492 cells are embedded into 3D reconstituted basement membrane matrix (3D-rBM) they form branching colonies mimicking the TDLUs of the breast, thereby providing a well-suited *in vitro* model for studies on branching morphogenesis and breast development. Peroxidasin (*PXDN*) is a heme-containing peroxidase that crosslinks collagen IV with the formation of sulfilimine bonds. Previous studies indicate that *PXDN* plays an integral role in basement membrane stabilisation by crosslinking collagen IV and as such contributes to epithelial integrity. Although *PXDN* has been linked to fibrosis and cancer in some organs there is limited information on its role in development, including in the breast. In this study, we demonstrate expression of PXDN in breast epithelium and stroma and apply the D492 cell line to investigate the role of *PXDN* in cell differentiation and branching morphogenesis in the human breast. Overexpression of *PXDN* induced basal phenotype in D492 cells, loss of plasticity and inhibition of epithelial-to-mesenchymal transition as is displayed by complete inhibition of branching morphogenesis in 3D culture. This is supported by results from RNA-sequencing which show significant enrichment in genes involved in epithelial differentiation along with significant negative enrichment of EMT factors. Taken together, we provide evidence for a novel role of *PXDN* in breast epithelial differentiation and mammary gland development.

Keywords Peroxidasin · p63 · Mammary gland · Branching morphogenesis · D492 · Mammary stem cells

Introduction

The mammary gland, the signature organ of mammals, is responsible for the production of milk to feed the offspring. The gland is a branched structure comprised of excretory ducts and terminal duct lobular units (TDLUs) which contain the functional epithelium along with stroma. The epithelium can be divided into an inner layer of luminal epithelial cells (LEP) and an outer layer of myoepithelial cells (MEP), both believed to be of the same origin [1–3]. Separating the epithelium from the stroma is the basement membrane, a thin layer of extracellular matrix mainly comprised of laminins and collagen IV, that maintains epithelial integrity and induces correct cell polarity [4, 5]. Although the development of the mammary gland begins during embryogenesis it is unique in that it does not complete its developmental process until pregnancy. However, during each oestrus cycle, continuous epithelial remodelling occurs as the gland prepares in prospect of pregnancy [6, 7]. This remodelling is maintained by bipotent epithelial stem cells that, along with lineage restricted progenitor cells, induce branching and epithelial differentiation within the gland [2, 8–10]. Branching morphogenesis is dependent upon epithelial cell plasticity that enables the cells to invade the surrounding

✉ Gunnhildur Asta Traustadottir
gutra@hi.is

[1] Stem Cell Research Unit, Biomedical Center, Department of Anatomy, Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland

[2] Department of Laboratory Haematology, Landspitali – University Hospital, Reykjavik, Iceland

stroma. This can be achieved through mechanisms such as collective migration [11] or epithelial to mesenchymal transition (EMT) [12, 13].

The D492 breast epithelial progenitor cell line was originally isolated from the $EpCAM^{+}/MUC1^{-}$ supra-basal cell population of the human breast epithelium [1]. This population is multipotent and can give rise to cells with both luminal and myoepithelial characteristics when cultured *in vitro*. This ability is preserved in D492 cells which, furthermore, mimic the breast TDLUs by forming branched structures when cultured in 3D reconstituted basement membrane (3D-rBM, Matrigel) [1]. In addition, D492 cells undergo EMT when co-cultured with endothelial cells [14]. Due to its ability to replicate key features of breast development the D492 cells are well suited for *in vitro* modelling and have successfully been used to study branching morphogenesis, epithelial plasticity and EMT [14, 15]. Furthermore, D492 has been used to study the role of non-coding RNAs (ncRNAs) in these processes [16–19].

Recently, we applied D492 and its mesenchymal daughter cell line D492M to study differential expression of miRNAs during EMT. This revealed that miR-203a was one of the most downregulated miRNAs in D492M compared to D492. Subsequent overexpression of miR-203a in D492M led to partial revision to an epithelial phenotype with accompanying functional changes such as decreased proliferation, migration and invasion, and increased susceptibility to chemically induced apoptosis. Interestingly, RNA-sequencing revealed that the most downregulated gene in $D492M^{miR\text{-}203a}$ was the basement membrane factor peroxidasin (*PXDN*) which we subsequently confirmed as a novel target for miR-203a [17]. Peroxidasin (*PXDN*) is an extracellular matrix peroxidase and the only known enzyme that stabilizes the basement membrane through sulfilimine bond formation between collagen IV fibrils [20] which are only found within the basement membrane [21]. *PXDN* loss of function leads to basement membrane destabilisation due to lack of cross-linking, indicating that *PXDN* plays a crucial role in development and tissue homeostasis [20, 22]. *PXDN* is expressed in various tissues such as the heart, colon and liver, and on the cellular level it has been identified in epithelium, vascular endothelium and smooth muscle cells [23]. In melanoma, increased *PXDN* expression has been linked to mesenchymal and invasive phenotype of malignant cells [24]. Furthermore, high *PXDN* expression in ovarian cancer has been linked to increased malignancy and poor patient outcome. Conversely, *in vitro* knock down of *PXDN* in HEY human ovarian carcinoma cell line decreased invasion, migration and proliferation, supporting the link between *PXDN* and malignant phenotype [25]. In the present study, we addressed the largely unexplored functional role of *PXDN* in mammary gland development by overexpressing the *PXDN* gene in the normal epithelial progenitor cell line D492. We demonstrate that overexpression of *PXDN* in D492 cells enforces basal epithelial phenotype and inhibits branching. Furthermore, we show through RNA-sequencing that *PXDN* is involved in epithelial cell differentiation and inhibition of EMT.

Materials and Methods

D492 Cell Culture

D492 were cultured in H14 media containing Dulbecco's Modified Eagle Medium DMEM:F12 (Gibco, #31330), supplemented with penicillin and streptomycin, Insulin (Sigma, #I1882), EGF (Peprotech, #AF-10-15), Transferrin (Sigma, #T1147), NaSel (BD Biosciences, #354201), Estradiol (Sigma, #E2758), Hydrocortisone (Sigma, #H0888), and Prolactin (Peprotech, #100-07). Cells were cultured in flasks precoated with collagen I solution (PureCol Type 1 Collagen Solution, Advanced Biomatrix, #5005-100ML) at 37 °C and 5% CO_2. Media was changed three times a week and cells were passaged when 90% confluent at a ratio of 1:10.

3D Matrix Embedded Culture

3D cultures with cells embedded in matrix were carried out in 48 well plates (Corning, #353078), where $1{,}0x10^4$ cells were resuspended and plated in 150 µl of growth factor reduced reconstituted basement membrane matrix (Matrigel, Corning, #354230) in triplicates. Plate was incubated at 37 °C and in 5% CO_2 for 30 minutes until Matrigel had solidified, and 500 µl of H14 media was gently added on top. Cells were grown for 21 days, and media was changed three times a week.

3D on Top of Matrix Culture

3D with cells on top of matrix was carried out in 96 well plates (Matrigel, Corning, #354230), where 50 µl of Matrigel were plated and then incubated for 30 minutes at 37 °C and in 5% CO_2 until the gel had solidified. Subsequently, $6{,}0x10^3$ cells were seeded on top of the gel in 100 µl of H14 media. Media was changed three times a week by replacing half of old media with equal volume of fresh media. Cells were grown for 8 days, and colony formation and growth were monitored in IncuCyte Live Cell Analysis System (Essen BioScience).

Isolation of Primary Cells from Breast Tissue

Luminal epithelial- and myoepithelial cells were purified after breast organoids from reduction mammoplasty had spread out in primary culture. The isolation was done immunomagnettically

using anti-EpCAM columns. Luminal- and myoepithelial cells separations were carried out by use of the MiniMACS magnetic cell separation system according to the manufacturer's instructions (Miltenyi Biotech) [1]. BRENCs were isolated from adipose tissue derived from reduction mammoplasty which resulted in relatively pure microvessel organoids. The microvasculature component was incubated with anti-CD31 Dynabeads (Invitrogen, #11155D) and isolated on a magnetic concentrator. Microvessels were seeded onto collagen coated flasks, resulting in enriched population of endothelial cells [26]. Fibroblasts appear as single cells upon collagenase treatment of the breast tissue. Purification of fibroblasts was obtained by differential centrifugation of the digest and moving the floating fibroblast to a culture flask [27].

Lentiviral Packaging and Transduction

Lentiviral particles were produced with pLenti-C-Myc-DDK-P2A-Puro cloning vector containing *PXDN* sequence (OriGene, #RC224518L3) along with pMD2.G (a gift from Didier Trono, Addgene plasmid #12259; http://n2t.net/addgene:12259; PRID:Addgene_12259) and psPAX2 (a gift from Didier Trono, Addgene plasmid #12260; http://n2t.net/addgene:12260; PRID:Addgene_12260) packaging plasmids. Plasmids were transfected into HEK-293T cells using TurboFect Transfection Reagent (Thermo Fisher Scientific, #R0533) in antibiotic and serum free high glucose DMEM (Gibco, #31966-021). Supernatant containing lentivirus was collected 48 hours and 72 hours post-transfection. D492 were infected with viral supernatant overnight in the presence of 8 µg/ml Polybrene (Sigma-Aldrich, #TR-1003). *PXDN* transduced cells were subsequently selected with puromycin (Gibco, #A11138-03) at concentration of 2 µg/ml for D492 cells. Control cells were transduced with empty pLenti-C-Myc-DDK-P2A-Puro cloning vector (OriGene, #PS100092).

RNA Isolation and Quantitative RT-PCR Analysis

RNA was isolated with Tri-Reagent (Thermo Fisher Scientific, #AM9738) and 1 µg of each sample RNA was reverse transcribed using LunaScript RT SuperMix Kit (New England BioLabs, #E3010). Quantitative real time PCR was performed using SYBR Green Luna Universal qPCR Master Mix (New England BioLabs, #M3003) as per manufacturer recommendation. Relative expression was determined via calculation of $2^{-\Delta\Delta Ct}$ using ABI 7500 instrument (Applied Biosystems) and GAPDH as endogenous control for gene expression. The following primers were used (purchased from IDT Integrated DNA Technologies) *PXDN* (Hs.PT.58.630748), *CK14* (Hs.PT.58.4592110), *CK19* (Hs.PT.58.4188708), *CK5/6* (Hs.Pt.58.14446018), *TP63* (Hs.PT.58.2966111) and *GAPDH* (Hs.Pt.39a.22214836) as an endogenous control.

Immunochemistry of Cells

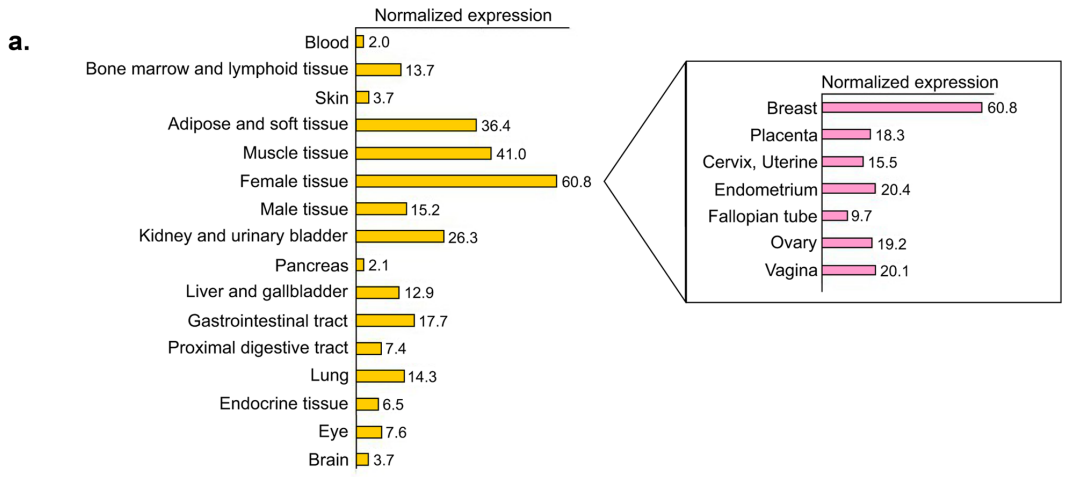
For immunochemistry of cells in monolayer and 3D culture the following antibodies were used: PXDN (a gift from M. Geiszt, Semmelweis University, Budapest) 1:250, KRT19 (Abcam, #ab7754) 1:100, KRT14 (Abcam, #ab7800) 1:100, and p63 (Novocastra, #NCL-p63) 1:100. Fluorescent labelling was performed using fluorescent secondary antibodies (Alexa fluor, Thermo Fisher Scientific). Imaging was performed on Zeiss LSM 5 Pa laser-scanning microscope (Carl Zeiss) and Olympus fluoview 1200.

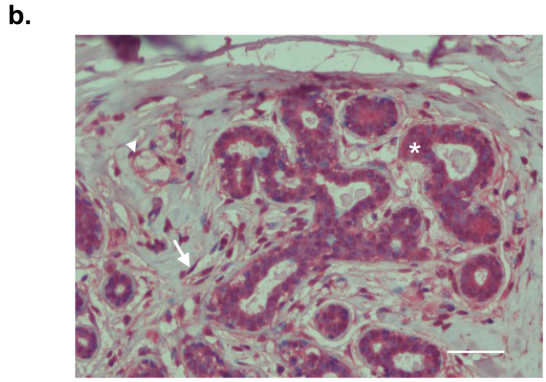
Immunohistochemistry of Paraffin Embedded Mammary and Breast Cancer Tissue

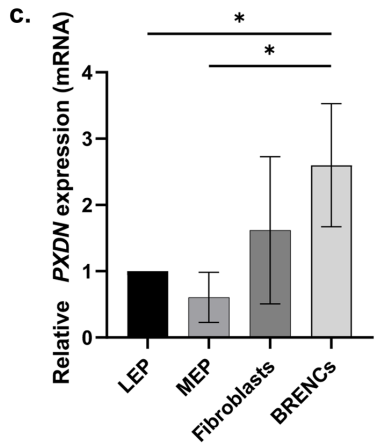
Immunohistochemistry of tissue slides was performed with anti-PXDN (a gift from M. Geiszt, Semmelweis University, Budapest) 1:400. Secondary staining was performed with Fast Red (Sigma, #F4648). Tissue slides were counterstained with hematoxylin. Imaging was performed using Nikon Eclipse Ci microscope.

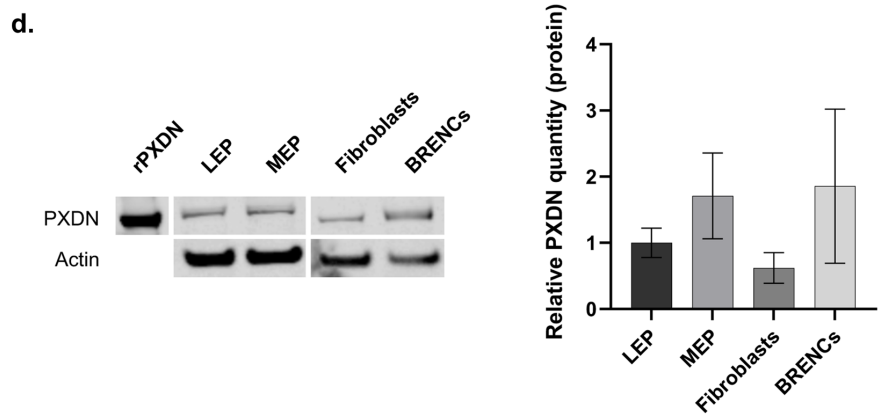
Protein Isolation and Western Blot Assay

Protein was isolated from cells lysed with radio immunoprecipitation assay (RIPA) buffer containing protease and phosphatase inhibitors (Halt Protease Inhibitor Cocktail, Thermo Fisher Scientific, #78430). Bradford reagent (BioRad, #5000002) was used to determine protein concentration. Western blot was performed on 10% NuPage Bis-Tris gels (Invitrogen, #NP0301PK2) in NuPage MES running buffer (Thermo Fisher Scientific, #NP0002). Protein was transferred from gel to Immobilon-FL PVDF membrane (Merck-Millipore, #IPFL00010) with NuPage transfer buffer (Thermo Fisher Scientific, #NP0006-1). Membrane was blocked with Odyssey Blocking Buffer (Li-Cor, #927-500) and incubated overnight at 4 °C with primary antibodies. The following primary antibodies were used for protein detection: PXDN (a gift from M. Geiszt, Semmelweis University, Budapest) 1:1000, KRT19 (Santacruz, #sc-6278) 1:1000, KRT14 (Abcam, #ab7800) 1:1000, KRT5/6 (Invitrogen, #180267) 1:500, EpCam (Abcam, #ab71916) 1:1000, with Actin (Li-Cor, #924-42212) as endogenous control, and in nuclear fraction p63 (Abcam, #ab124762) 1:2000, with Histone H3 (Cell Signaling, #4499). Secondary antibodies used were rabbit or mouse (Li-Cor) and imaging was performed using Odyssey Infrared Imaging System (Li-Cor).


a.
Normalized expression
Blood 2.0
Bone marrow and lymphoid tissue 13.7
Skin 3.7
Adipose and soft tissue 36.4
Muscle tissue 41.0
Female tissue 60.8
Male tissue 15.2
Kidney and urinary bladder 26.3
Pancreas 2.1
Liver and gallbladder 12.9
Gastrointestinal tract 17.7
Proximal digestive tract 7.4
Lung 14.3
Endocrine tissue 6.5
Eye 7.6
Brain 3.7
Normalized expression
Breast 60.8
Placenta 18.3
Cervix, Uterine 15.5
Endometrium 20.4
Fallopian tube 9.7
Ovary 19.2
Vagina 20.1


b.


c.
Relative PXDN expression (mRNA)
0
1
2
3
4
*
*
LEP
MEP
Fibroblasts
BRENCs


d.
rPXDN
LEP
MEP
Fibroblasts
BRENCs
PXDN
Actin
Relative PXDN quantity (protein)
0
1
2
3
4
LEP
MEP
Fibroblasts
BRENCs

◂**Fig. 1** *PXDN* is expressed in both epithelial and stromal compartments of the breast gland. **a** *PXDN* expression is highest in breast tissue. Tissue specific expression data from the Human Protein Atlas showed that *PXDN* levels are higher in breast than other tissues. **b** Immunohistochemistry of PXDN in paraffin embedded normal breast tissue. Positive signal (red) was detected in epithelial cells (example depicted with asterix), fibroblasts (example depicted with arrow) and endothelial cells (example depicted with arrowhead). Tissue slides were counterstained with hematoxylin. Scale bar = 100 µm. **c** *PXDN* was expressed in luminal epithelial cells (LEP), myoepithelial cells (MEP), fibroblasts and breast endothelial cells (BRENCs), isolated from primary tissue from reduction mammoplasties. BRENCs expressed the highest levels of *PXDN* compared to the other three subtypes followed by fibroblasts, as detected by qRT-PCR. Furthermore, *PXDN* was expressed in both LEPs and MEPs which both belong to the epithelial compartment. Statistical significance was determined by One-way ANOVA ($*p \leq 0.05$) and data is presented as an average of three replicated experiments (mean ± SD). **d** Western blot showing protein level expression of PXDN in LEP, MEP, fibroblasts and BRENCs with recombinant PXDN (rPXDN) as positive control. No significant difference was found in PXDN expression between primary cells. Statistical significance was determined by One-way ANOVA and data is presented as an average of three replicated experiments (mean ± SD)

Proliferation Assay

Cells were seeded at $4.0x10^3/cm^2$ in a 6-well plate (Falcon, #353046) format, pre-coated with collagen I solution (PureCol Type 1 Collagen Solution, Advanced Biomatrix, #5005-100ML). Plates were incubated at 37 °C and 5% CO_2. Every 24 hours three wells of each cell line were treated with Trypsin-EDTA, and cells were counted. Proliferation rate was determined from increase in cell number over six consecutive days.

Apoptosis Assay

Apoptosis assay was performed in IncuCyte (Essen Bioscience) as per manufacturer's instructions using 10µM Camptotechin (Sigma-Aldrich, #C9911). Apoptosis was detected by IncuCyte Caspase-3/7 Green Reagent (Essen Bioscience, #4440).

Migration and Invasion Assays

Migration and invasion assays were performed as per manufacturer's instructions in IncuCyte (Essen Bioscience) in a 96 well plate format. $1.0x10^{10}$ cells were seeded per well and allowed to grow until confluent before scratching.

Statistical Analysis

Statistical analysis was performed in GraphPad Prism version 8.3.0. All experiments were performed in three independent experiments.

RNA Sequencing

RNA was isolated from $D492^{empty}$ and $D492^{PXDN}$ cells in monolayer at 80% confluency and on day 8 of 3D on top of matrix culture. RNA sequencing was performed on samples in quintuplets at DeCODE Genetics (Reykjavik, Iceland).

Gene enrichment analyses were performed in GSEA (gsea-msigdb.org) with pre-ranked gene lists and the Hallmark and C5 Ontology gene sets for HALLMARK and GO analyses, respectively. Volcano plots and heatmaps were generated in Graph Pad Prism.

Transient Knock Down of *TP63* and *PXDN*

Transient knock down in $D492^{PXDN}$ cells was performed with siRNA targeting *TP63* (Thermo Fisher Scientific. SiRNA #1: #s16411 lot.AS02HN1T. SiRNA #2: #s502043 lot.AS02043), and in D492 cells with siRNA targeting *PXDN* (Thermo Fisher Scientific, #4392421 lot. AS02B0C4), using Lipofectamine RNAiMAX Transfection Reagent (Thermo Fisher Scientific, #13778075) as per manufacturer's instructions. Cells were incubated for 48 hours after transfection whereafter assays were performed.

Analysis of *PXDN* Expression in Breast Cancer

PXDN expression in different breast cancer subtypes was analysed in the Gene expression-based Outcome for Breast Cancer Online (GOBO) database [28]. Distant metastasis free survival (DMFS) over a period of 10 years was analysed in 3 quantiles in all tumor subtypes.

Results

PXDN is Expressed in Breast Epithelium and Stroma

Data from the Human Protein Atlas database revealed that *PXDN* expression is highest in female tissues compared to other human tissues and among those, *PXDN* expression was highest in the breast (Fig. 1a). In support, immunohistochemistry of paraffin embedded normal breast tissue revealed positive PXDN signal in epithelial cells, fibroblasts and endothelial cells (Fig. 1b). To confirm the expression of *PXDN* in the mammary gland we investigated the expression pattern in four different subtypes of primary cells isolated from reduction mammoplasties: EpCam-sorted luminal epithelial cells (LEPs) and myoepithelial cells (MEPs) (Supplementary Fig. 1), fibroblasts and breast endothelial cells (BRENCs). Quantitative RT-PCR analysis revealed that *PXDN* was

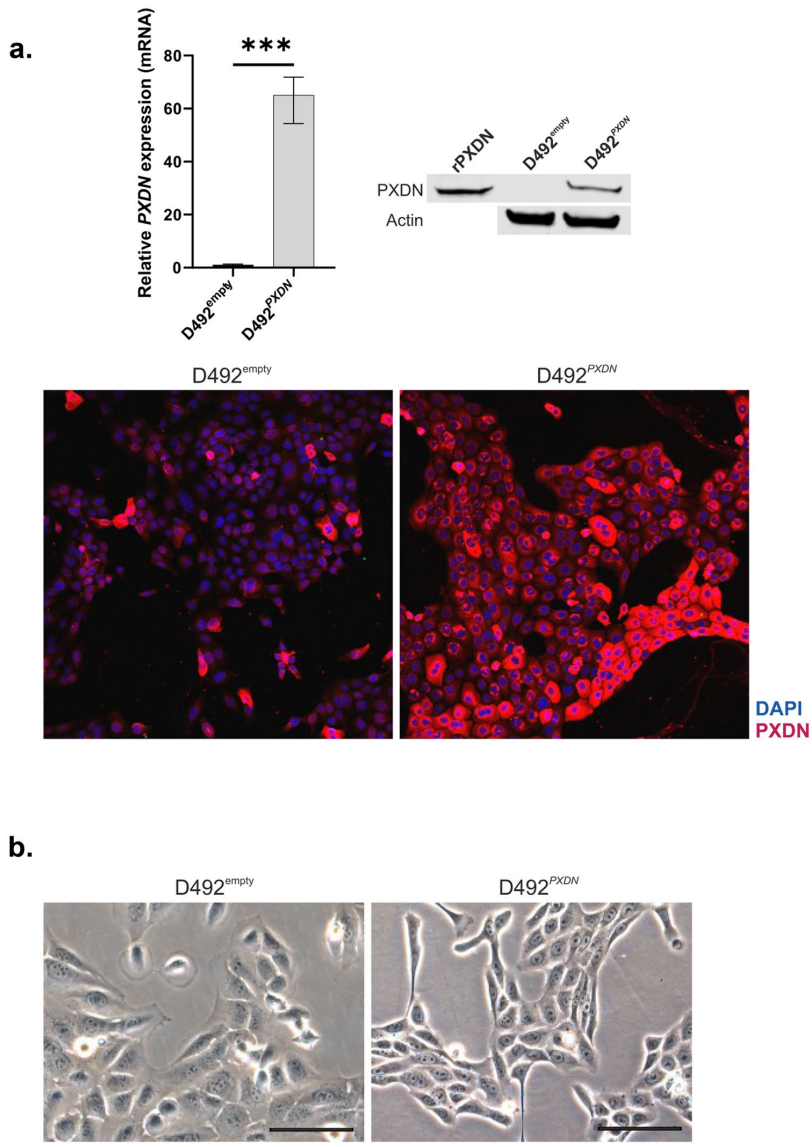
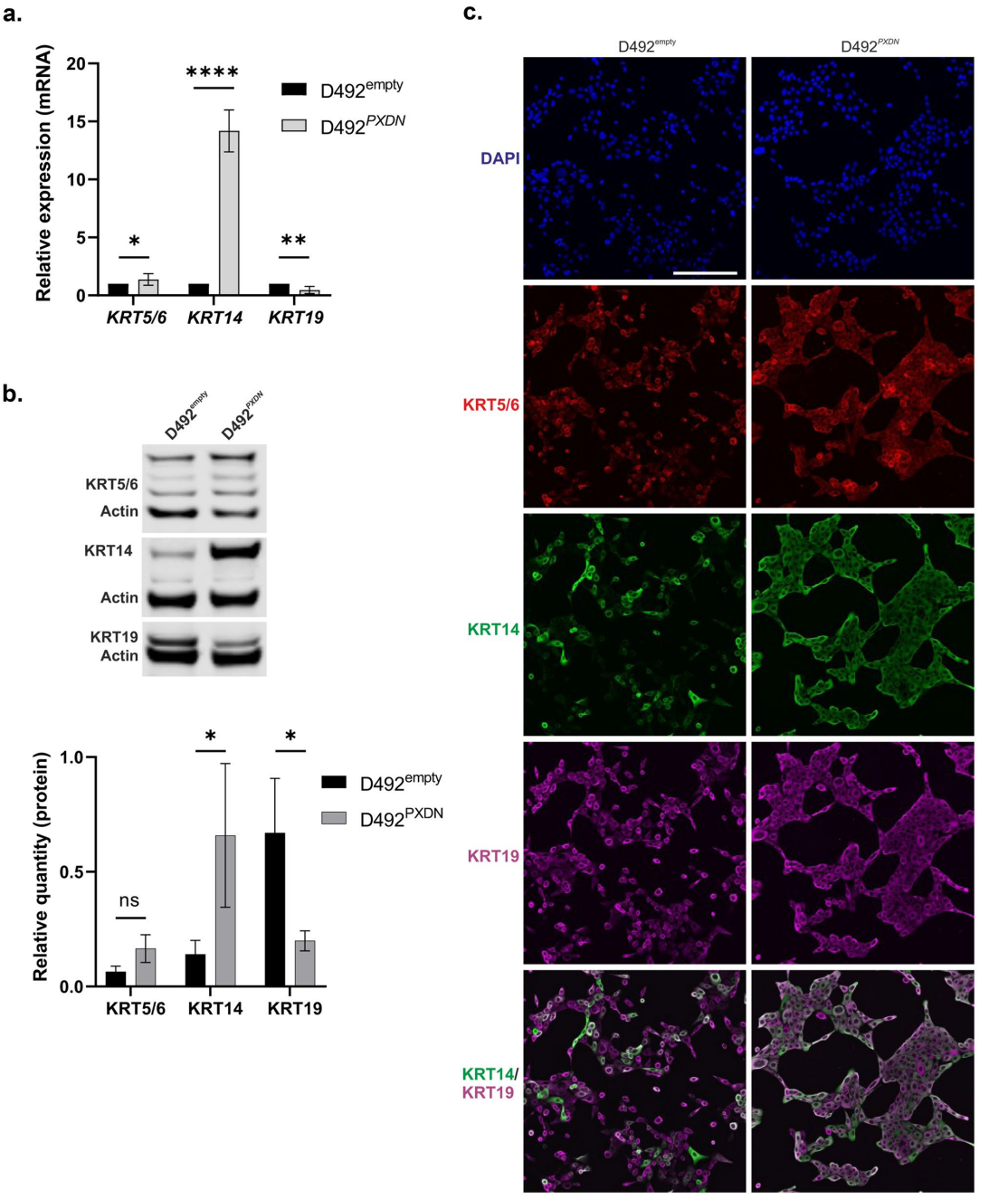


Fig. 2 Confirmation of *PXDN* overexpression in D492 cells and effect on monolayer growth pattern. **a** *PXDN* was successfully overexpressed in D492. qRT-PCR results showed approximately 65-fold upregulation of *PXDN* in D492^{PXDN} compared to D492^{empty}. Results were confirmed by Western blot where recombinant *PXDN* (rPXDN) was used as positive control, and by immunofluorescence staining. Statistical significance in qRT-PCR was determined by unpaired Student's t-test (***$p \leq 0.001$) and data is presented as an average of three replicated experiments (mean ± SD). Scale bar = 125 µm. **b** Overexpression of *PXDN* affected cell phenotype and growth pattern of D492 in monolayer culture. D492^{PXDN} cells were more homogeneous and cuboidal in phenotype compared to D492^{empty}. D492^{PXDN} grew closer together than D492^{empty} in monolayer, forming a dense network as opposed to more even distribution of cells in D492^{empty}. Scale bar = 100 µm

expressed in all cell types, albeit with higher expression in fibroblasts and BRENCs, both belonging to the breast stromal compartment. Considerable expression was also detected in LEPs and MEPs, indicating a role of *PXDN* within the functional epithelium. (Fig. 1c). Western blot confirmed protein level expression in all four primary cell types, although the results did not completely reflect mRNA-level expression (Fig. 1d).


a.
Relative expression (mRNA)
20
15
10
5
0
****
*
**
D492empty
D492PXDN
KRT5/6
KRT14
KRT19
b.
D492empty
D492PXDN
KRT5/6
Actin
KRT14
Actin
KRT19
Actin
Relative quantity (protein)
1.0
0.5
0.0
*
*
ns
D492empty
D492PXDN
KRT5/6
KRT14
KRT19
c.
D492empty
D492PXDN
DAPI
KRT5/6
KRT14
KRT19
KRT14/
KRT19

◀ **Fig. 3** Overexpression of *PXDN* in D492 enforced epithelial basal phenotype. **a** Overexpression of *PXDN* enforced upregulation of basal epithelial cell markers in D492. qRT-PCR results revealed that epithelial basal cell markers *KRT14* and *KRT5/6* were significantly upregulated in D492^{PXDN} compared to D492^{empty}, while luminal epithelial cell marker *KRT19* was significantly downregulated. Statistical significance in qRT-PCR was determined by unpaired Student's t-test (*$p \leq 0.05$, **$p \leq 0.01$, ****$p \leq 0.0001$) and data is presented as an average of three replicated experiments (mean ± SD). Scale bar = 125 µm. **b** Western blot confirmed protein level upregulation of basal markers in D492^{PXDN}. Relative quantification of protein bands confirmed significant upregulation of *KRT5/6*, *KRT14* and *KRT19*. Statistical significance was determined by unpaired Student's t-test (*$p \leq 0.05$) and data is presented as an average of three replicated experiments (mean ± SD). **c** D492^{PXDN} cells coexpressed luminal and basal markers. Immunofluorescence staining revealed that D492^{empty} had both *KRT14-/KRT19+* and *KRT14+/KRT19+* populations. In D492^{PXDN} the *KRT14-/KRT19+* population was lost while *KRT14+/KRT19+* cells were enriched. Scale bar = 125 µm

Overexpression of *PXDN* in D492 Affected Monolayer Growth Pattern, Enforced Basal Epithelial Phenotype and Inhibited Branching Morphogenesis

We have previously reported that while *PXDN* is minimally expressed in D492 it is significantly upregulated in its mesenchymal daughter cell line D492M [17]. As D492 cells are well established as a model for branching morphogenesis in the breast and yield cells with both luminal and myoepithelial characteristics, we decided to apply D492 to investigate the role of *PXDN* in mammary gland development and epithelial differentiation. Therefore, we overexpressed *PXDN* in D492 cells using lentiviral approach. Quantitative RT-PCR revealed approximately 65-fold upregulation of *PXDN* in D492^{PXDN} compared to D492^{empty} and successful

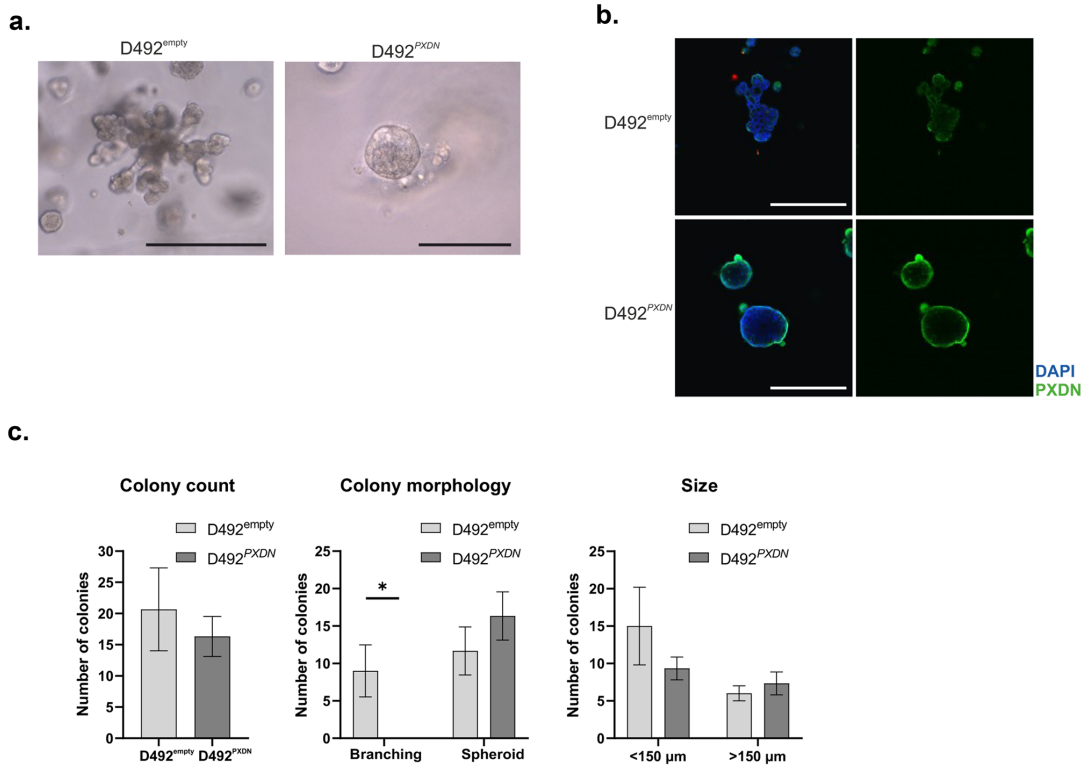
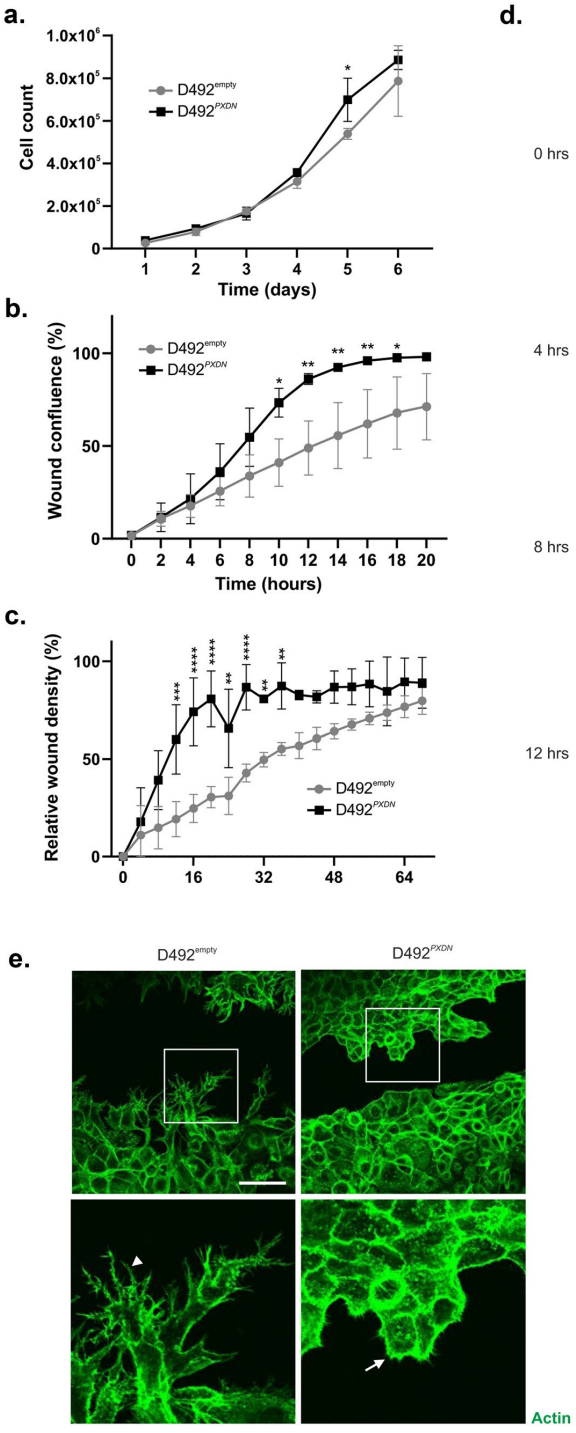


Fig. 4 *PXDN* inhibited branching morphogenesis in D492 cells. **a** When cultured in 3D-rBM for 21 days, D492 cells form branching structures reminiscent of TDLUs in the breast. This ability was retained in D492^{empty}. However, branching morphogenesis was completely inhibited in D492^{PXDN} cells. Scale bar = 125 µm. **b** Overexpression of *PXDN* in D492^{PXDN} was preserved under 3D conditions. Increased secretion of *PXDN* to the ECM was observed in D492^{PXDN} cells cultured in 3D-rBM. *PXDN* was also visible around D492^{empty} colonies but to a lesser extent. Scale bar = 125 µm. **c** Overexpression of *PXDN* significantly altered morphology but not count and size of colonies in 3D-on top culture. No significant difference was seen in colony count or size between to D492^{empty} and D492^{PXDN} cultured on top of Matrigel on day 9. However, branching was inhibited in D492^{PXDN}, which only formed solid spheroid colonies. Scale bar = 150 µm. Statistical significance was determined using unpaired Student's t-test and data is presented as an average of three replicated experiments (mean ± SD)


a.
Cell count
1.0x10^6
8.0x10^5
6.0x10^5
4.0x10^5
2.0x10^5
0
D492^empty
D492^PXDN
1 2 3 4 5 6
Time (days)
b.
Wound confluence (%)
100
50
0
D492^empty
D492^PXDN
0 2 4 6 8 10 12 14 16 18 20
Time (hours)
c.
Relative wound density (%)
100
50
0
D492^empty
D492^PXDN
0 16 32 48 64
d.
D492^empty
D492^PXDN
0 hrs
4 hrs
8 hrs
12 hrs
e.
D492^empty
D492^PXDN
Actin

◂ **Fig. 5** Overexpression of *PXDN* increased proliferation, decreased wound closure time and negatively affected plasticity in D492 cells. **a** *PXDN* did not significantly affect proliferation rate in D492 cells. D492^{PXDN} and D492^{empty} cells proliferated at a similar rate. D492^{PXDN} only proliferated slightly faster on day 5 of the experiment. Statistical significance was determined by two-way ANOVA with multiple comparisons ($*p \leq 0.05$) and data is presented as an average of three replicated experiments (mean ± SD). **b** D492^{PXDN} cells had increased migratory capacity in wound healing. *PXDN* significantly increased migration in D492 cells in wound healing compared to D492^{empty}. Data was analyzed in IncuCyte Zoom and is presented as wound confluence % and an average of three replicated experiments (mean ± SD). Statistical significance was determined by two-way ANOVA with multiple comparisons ($*p \leq 0.05$, $**p \leq 0.01$). **c** D492^{PXDN} cells had increased invasive capacity in wound healing under Matrigel. D492^{PXDN} cells wound closure was significantly accelerated compared to D492^{empty} when Matrigel was applied on top of cells after scratching. Data was analyzed in IncuCyte Zoom and is presented as wound confluence % and an average of three replicated experiments (mean ± SD). Statistical significance was determined by two-way ANOVA with multiple comparisons ($**p \leq 0.01$, $***p \leq 0.001$, $****p \leq 0.0001$). **d** D492^{PXDN} cells closed the wound in 12 hours in invasion scratch assay under Matrigel. D492^{PXDN} cells were able to invade through Matrigel and close scratch wound area in 12 hours. In comparison D492^{empty} cells reached wound closure four and half days later. Scale bar = 300 µm. **e** D492^{PXDN} formed lamellipodia during wound closure under Matrigel. Immunofluorescence staining of actin filaments in wound healing under Matrigel, 24 hours post scratching, revealed that D492^{empty} formed extensive filopodia (example depicted with arrowhead) while D492^{PXDN} cells formed lamellipodia (example depicted with arrow). Furthermore, peripheral actin staining was increased in D492^{PXDN}. Scale bar = 150 µm

upregulation was confirmed with Western blot and immunofluorescence staining of monolayer cell culture (Fig. 2a). In monolayer culture, D492^{PXDN} cells were phenotypically distinguishable from D492^{empty} cells, as they were smaller, cuboidal and more homogeneous in appearance compared to D492^{empty}. Furthermore, they grew in tighter proximity to neighbouring cells than D492^{empty} (Fig. 2b).

Having established phenotypic differences between D492^{PXDN} and D492^{empty}, we asked whether this was reflected by changes in epithelial cell differentiation markers. Indeed, quantitative RT-PCR showed significant upregulation of basal markers *KRT14* and *KRT5/6* in D492^{PXDN}, while luminal marker *KRT19* was significantly downregulated (Fig. 3a). Differential expression of these three markers was confirmed on protein level by Western blotting where significant change in expression of all three markers was observed (Fig. 3b). Immunofluorescence staining of cells cultured in monolayer revealed upregulation of both KRT5/6 and KRT14 in D492^{PXDN} while changes in expression of KRT19 were less clear. However, when comparing the expression pattern of KRT14 and KRT19 it was apparent that all D492^{PXDN} cells co-expressed both markers and that there was a depletion of KRT19+/KRT14- cells (Fig. 3c). This subpopulation of KRT19+/KRT14- cells reflected the luminal population of the breast epithelium [29] while it has been suggested that KRT19+/KRT14+ cells are mammary gland stem cells [2]. The ability to yield populations that are KRT19+/KRT14- or KRT19-/KRT14+ with luminal and myoepithelial characteristics, respectively, in addition to a small KRT19+/KRT14+ subpopulation is characteristic for the D492 cell line and is reflective of its plasticity [15]. The depletion of populations other than KRT19+/KRT14+ in D492^{PXDN} might therefore indicate loss of plasticity. To analyse whether overexpression affected the expression of *COL4A1* and *COL4A5* we decided to evaluate the expression in D492^{empty} and D492^{PXDN} which revealed no observed differences between the cell lines (Supplementary Fig. 2).

We have previously demonstrated that plasticity is necessary for branching morphogenesis to occur in D492 cultured in 3D-rBM [16]. Colonies formed by cells with either luminal or myoepithelial phenotype without plasticity were unable to adjust their phenotype and thus, were also unable to branch [16]. Interestingly, 3D cultures of D492^{empty} and D492^{PXDN} showed that D492^{empty} colonies retained the branching ability of D492, whereas branching was completely inhibited in D492^{PXDN} (Fig. 4a). Immunohistochemical staining of D492^{PXDN} and D492^{empty} 3D colonies revealed that increased expression of *PXDN* was preserved under 3D conditions (Fig. 4b). Colony characterisation confirmed that there was no significant difference between the size and number of D492^{empty} and D492^{PXDN}. Colony morphology reflected previous observation where D492^{empty} formed both spheroid and branching colonies while D492^{PXDN} only formed spheroids (Fig. 4c).

Overexpression of *PXDN* Decreased Wound Closure Times in D492 Cells and Negatively Affected Plasticity

Next, we applied a series of functional assays to clarify whether changes in phenotype, as a result of *PXDN* overexpression, also affected cell behaviour. There was no significant difference in proliferation in D492^{PXDN} compared to D492^{empty}, except at a single timepoint (Fig. 5a). However, *PXDN* promoted migration of D492 cells, evidenced by increased migratory ability of D492^{PXDN} in a wound healing assay, where wound confluence reached near 100% 15 hours after application of the scratch wound. On the contrary, when the experiment was terminated at 20 hours, D492^{empty} confluency was still only near 75% (Fig. 5b) Invasion assay revealed a similar trend where wound edges under Matrigel closed significantly sooner in D492^{PXDN} compared to D492^{empty} (Fig. 5c, d). Immunofluorescence staining of actin showed that D492^{empty} cells formed long, branching filopodia that were absent in D492^{PXDN}. However, D492^{PXDN} formed lamellipodia with stronger peripheral actin signal than D492^{empty} (Fig. 5e). As phenotype can influence response to chemically induced apoptosis, we performed apoptosis assay using Camptotechin. This revealed significant increase in sensitivity to apoptosis in D492^{PXDN} compared to D492^{empty} (Supplementary Fig. 3).

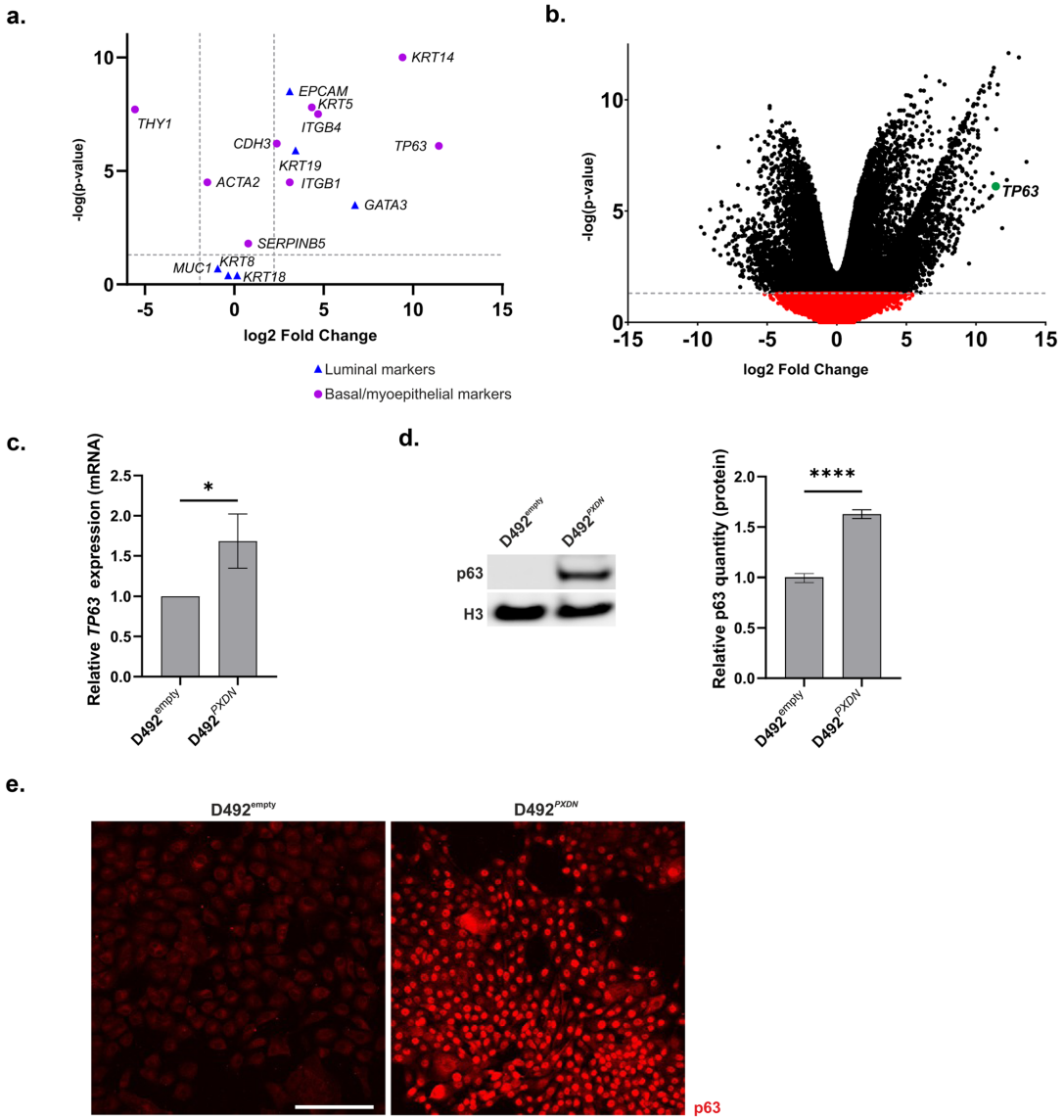
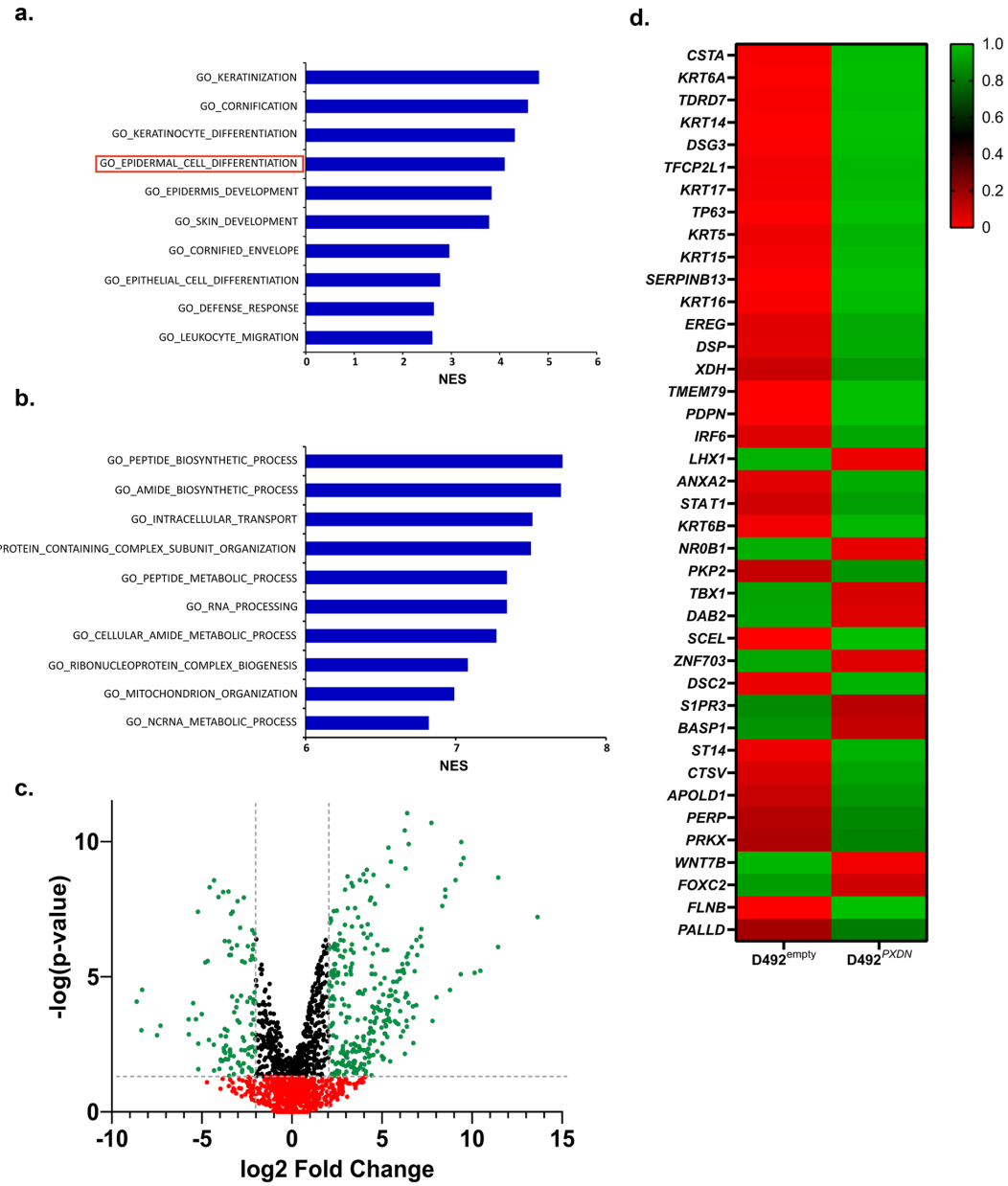


Fig. 6 Basal phenotype was induced in D492^{PXDN} cells. **a** Basal markers were more significantly upregulated in D492^{PXDN} than luminal and myoepithelial markers. Volcano plot depicting differential gene expression of basal, luminal and myoepithelial markers in D492^{PXDN} compared to D492^{empty}. Dotted lines mark statistical significance of p-value=0.05 and log2 fold change above 2 and below -2. **b** *TP63* was one of the most upregulated genes in D492^{PXDN}. Volcano plot showing RNA sequencing results from D492^{empty} and D492^{PXDN} in monolayer. Significantly upregulated genes ($p \leq 0.05$) are represented by black dots and non-signifcant genes are represented by red dots. **c** PXDN overexpression induced upregulation of p63. qRT-PCR revealed significant upregulation of basal cell marker p63 in D492^{PXDN}. Statistical significance in qRT-PCR was determined by unpaired Student's t-test ($*p \leq 0.05$) and data is presented as an average of three replicated experiments (mean ± SD). **d** Protein level upregulation of p63 was confirmed by Western blot. Relative quantification of protein band is displayed to the right. Statistical significance of Western blot was determined by unpaired Student's t-test ($****p \leq 0.0001$) and data is presented as an average of three replicated experiments (mean ± SD). **e** Immunofluorescence staining of D492^{empty} and D492^{PXDN} revealed increased nuclear signal of p63 in D492^{PXDN}. Upregulation of p63 was confirmed by Western blot of nuclear protein. Scale bar = 125 µm



Basal Markers were Significantly Upregulated in D492^{PXDN} cells

To gain insight into how overexpression of *PXDN* influences gene expression we performed RNA sequencing on D492^{empty} and D492^{PXDN} in monolayer and in 3D culture. Results from monolayer showed that *KRT5* and *KRT14* were among the most significantly upregulated genes in D492^{PXDN} compared to D492^{empty}. Further analysis of commonly applied basal, luminal and myoepithelial markers in breast epithelium showed that, compared to luminal and myoepithelial markers, the basal markers were more highly expressed with higher level of significance (Fig. 6a). *GATA3*, an important transcription factor involved in luminal epithelial cell differentiation [30],

◂**Fig. 7** RNA-sequencing revealed enrichment of genes involved in epithelial cell differentiation and development. **a** Developmental and differentiation processes were enriched in D492^{PXDN} in 2D. GSEA results using GO: Gene Ontology dataset showed that processes involved in epithelial development and differentiation were significantly enriched (FDR $q \leq 0.05$), most notably the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set. NES: Normalised Expression Score. **b** Metabolic pathways were most enriched in D492^{PXDN} in 3D. GSEA results using GO: Gene Ontology dataset showed that processes involved in metabolism were most significantly enriched (FDR $q \leq 0.05$). NES: Normalised Expression Score. **c** Approximately two thirds of the genes belonging to the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set, that were significantly differentially expressed, were upregulated in D492^{PXDN}. Volcano plot representing genes in the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set in D492^{PXDN} cells. **d** In the top 40 most significantly differentiated genes belonging to the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set the majority was upregulated in D492^{PXDN}. Heatmap representing the top 40 most significantly differentially expressed genes in the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set. Among them are *KRT6A, KRT5, KRT14* and *TP63*

was significantly upregulated in D492^{PXDN}, although it did not seem to be sufficient in inducing the luminal phenotype (Fig. 6a). Interestingly, the basal marker *TP63* (p63), a known transcription factor for *KRT14* and a key player in epithelial differentiation in the mammary gland [16, 31, 32], was one of the highest upregulated genes in D492^{PXDN} (Fig. 6b). This was confirmed on the mRNA- and protein-level with quantitative RT-PCR (Fig. 6c) and Western blot of nuclear fraction protein (Fig. 6d), respectively. Furthermore, immunofluorescence staining showed increased protein level expression and nuclear localization of p63 in all D492^{PXDN} cells, indicating increased activity (Fig. 6e). To investigate whether the basal shift in phenotype observed in D492^{PXDN} cells was due to *TP63* function we performed transient knock down of *TP63* with siRNA. Successful knock down of *TP63* was confirmed with quantitative RT-PCR and immunofluorescence staining (Supplementary Fig. 4a). Interestingly, the knock down did not significantly affect expression of *PXDN* (Supplementary Fig. 4b) or basal keratins *KRT5/6* and *KRT14* in D492$^{PXDN\text{-KD }TP63}$ compared to D492$^{PXDN\text{-Ctrl}}$. *KRT19* showed a trend towards upregulation in D492$^{PXDN\text{-KD }TP63}$ although the change was not significant. Immunofluorescence staining did not show any protein level change in keratins between D492$^{PXDN\text{-KD }TP63}$ and D492$^{PXDN\text{-Ctrl}}$ (Supplementary Fig 4c). Furthermore, transient knock down of *PXDN* in D492 cells did not significantly alter the expression of *KRT14, KRT19* or *TP63* (Supplementary Fig. 5).

PXDN was Involved in Epithelial Differentiation and Suppressed EMT

Next, as phenotypic shift was already established in D492^{PXDN}, we were interested in seeing what biological processes were affected by overexpression of *PXDN*. We performed pre-ranked Gene Set Enrichment Analysis (GSEA) using the GO: Gene Ontology datasets which showed that various epithelial developmental processes were significantly enriched, especially those related to cornification, keratinization and development of skin and epidermis (Fig. 7a). As it is generally accepted that the mammary gland is a modified sweat gland that arises from pluripotent epidermal stem cells of the skin [33, 34], common developmental pathways between the two organs are not surprising. More interestingly, GO_EPITHELIAL_CELL_DIFFERENTIATION was significantly enriched. In this category, 113 genes were significantly downregulated ($p<0.05$, $log2FC \leq -2.0$) while 300 genes were significantly upregulated ($p<0.05$, $log2FC \geq 2.0$), contributing to the enrichment observed in the results (Fig. 7b). Of the top 40 most highly differentially expressed genes belonging to the GO_EPITHELIAL_CELL_DIFFERENTIATION gene set most were upregulated in D492^{PXDN} (Fig. 7c).

Next, we performed pre-ranked GSEA analysis using the HALLMARK dataset. Monolayer data showed significant negative enrichment of HALLMARK_EPITHELIAL_TO_MESENCHYMAL_TRANSITION. Furthermore, several EMT factors were significantly downregulated in D492^{PXDN} (Fig. 8a). As epithelial to mesenchymal transition (EMT) contributes to epithelial plasticity and is a fundamental process in branching morphogenesis [15, 35], we were interested in whether this would also be the case in D492^{PXDN} 3D colonies, where branching was inhibited. We therefore performed RNA sequencing on D492^{empty} and D492^{PXDN} colonies isolated from 3D-rBM. Results showed that EMT was the second most negatively enriched pathway in D492^{PXDN}. Furthermore, a greater number of EMT factors was downregulated in 3D compared to 2D conditions. (Fig. 8b). These results indicate that *PXDN* is involved in suppression of EMT which might contribute to inhibition of branching morphogenesis.

PXDN in Breast Cancer

Immunohistochemistry of three subtypes of breast cancer (ER-positive, HER2-positive and Triple negative) revealed that PXDN was expressed in cancer cells. Furthermore, four tumors showed PXDN-positive fibrils in the tumor stroma (Supplementary Fig. 6a). Results from the Gene expression-based Outcome for Breast Cancer Online (GOBO) database showed that women with basal and HER2-positive tumors that had high *PXDN* expression had poorer distant metastasis free survival (DMFS) compared to women with low *PXDN*-expressing tumors (Supplementary Fig. 6b).

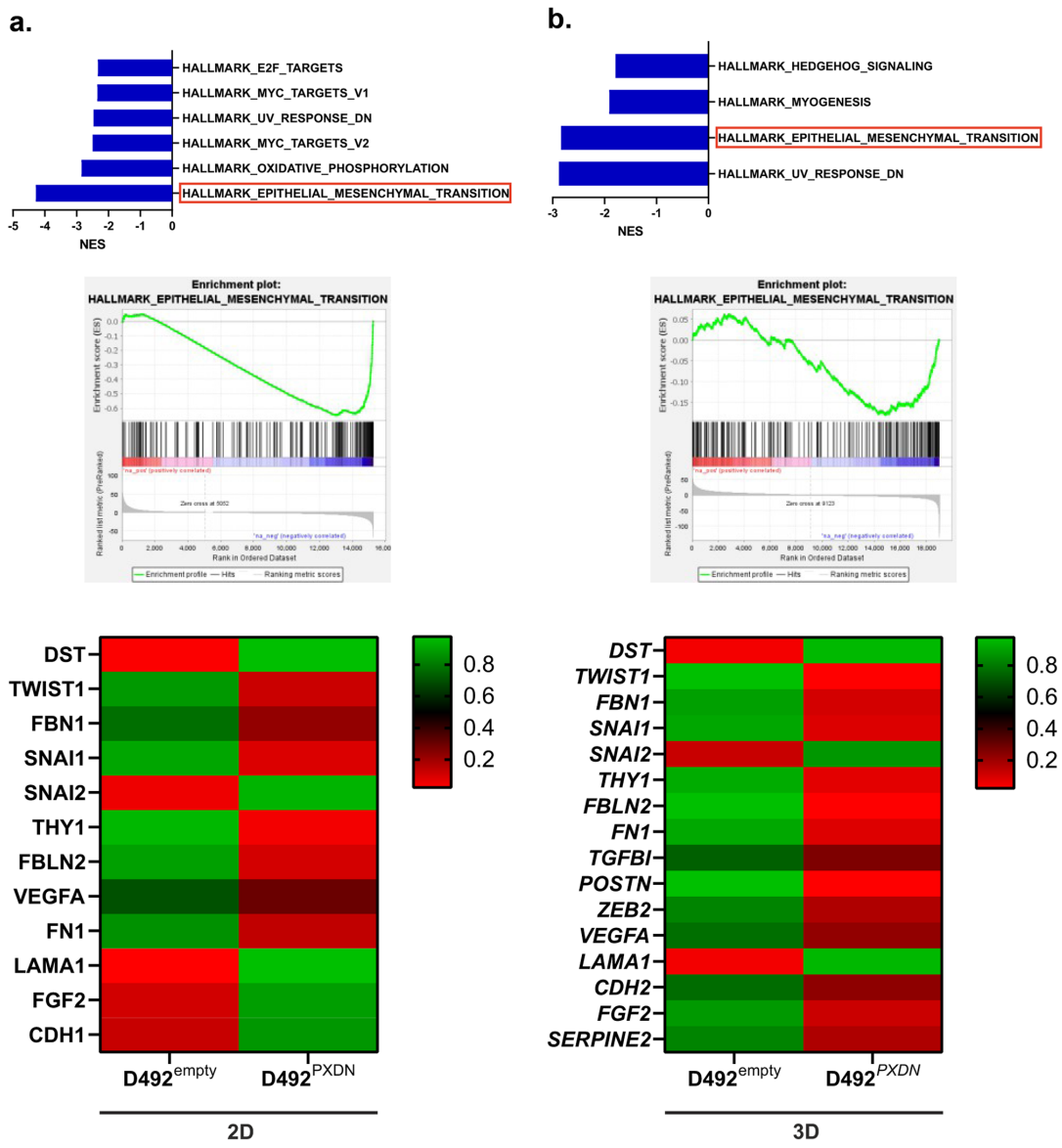


Fig. 8 *PXDN* suppressed epithelial to mesenchymal activity and led to significant downregulation of key EMT factors. **a** EMT was significantly negatively enriched and key EMT factors were significantly downregulated in D492^{PXDN} in 2D culture. GSEA results using the HALLMARK data set revealed significant negative enrichment of HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION gene set (FDR q-value≤0.05). Heatmap depicting key significantly differentially expressed factors involved in EMT in 2D culture. NES: Normalised Expression Score. **b** EMT was significantly negatively enriched and key EMT factors were significantly downregulated in D492^{PXDN} 3D culture. HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION was the second most negatively enriched gene set (FDR q-value≤0.05) in D492^{PXDN} colonies cultured in 3D-rBM. Heatmap depicting key significantly differentially expressed factors involved in EMT in 3D culture. NES: Normalised Expression Score

Discussion

Until now the presence and role of *PXDN* within the human breast has not been explored. In this study we show, for the first time, that *PXDN* is expressed within both the epithelial and stromal compartments of the breast, in LEPs and MEPs as well as fibroblasts and BRENCs, respectively. This was confirmed with immunohistochemistry of paraffin embedded normal breast tissue, where PXDN expression was detected within the epithelium and stroma.

In the past two decades lineage tracing studies have attempted to delineate distinct cell subpopulations within the mammary gland, with the purpose of identifying the hierarchy of stem and progenitor cells responsible for epithelial development and maintenance. Cells belonging to the basal lineage have thus far been recognized as the main instigators of differentiation [8, 10, 36, 37], and the implantation of a single pluripotent stem cell into cleared mouse fat pad can reconstruct a full mammary tree [38]. Villadsen et al. [2] identified the multipotent stem cell populations according to co-expression of KRT14 and KRT19. Earlier on, Yang et al. [39] had reported that p63, a member of the p53 family, was essential for basal cell population maintenance and epithelial development. Furthermore, they showed that p63 deficient mice underwent non-regenerative terminal epithelial differentiation and lacked certain epithelial structures such as the mammary gland. In previous work, members of our group showed that cell plasticity, with the ability to form cells with both luminal and myoepithelial characteristics, was necessary for branching morphogenesis in D492 [16]. Furthermore, they found that along with the mesenchymal repressor miR-200c-141, p63 reverted D492M, a mesenchymal daughter cell line of D492, back to epithelial phenotype. This restored cell plasticity and thus branching morphogenesis in 3D culture. On the contrary overexpression of p63 in primary luminal epithelial cells induced expression of basal markers and basal phenotype [40]. In the present study, we show that overexpression of *PXDN* in the epithelial progenitor cell line D492 induces basal phenotype and leads to loss of plasticity, inhibiting the cell line from flexibly forming cells with luminal and myoepithelial characteristics and thereby inhibiting branching morphogenesis. This is highlighted by co-expression of *KRT14, KRT19* and p63, which reflects the expression pattern of mammary epithelial cell subpopulations that have previously been identified as mammary progenitor cells [41, 42]. Interestingly, transient silencing of *TP63* did not lead to reduced expression of the basal keratins. This indicates that, despite its significant upregulation in D492^{PXDN}, p63 is not the primary driver of the basal phenotype in these cells. It is also unlikely that p63 knock down could reverse plasticity inhibition in D492^{PXDN}, thereby restoring branching, as some level of p63 activity is needed for necessary cell differentiation to occur [16]. This is supported by the fact that p63 knock down in VA10, a bronchial epithelial cell line with basal characteristics that forms differentiated bronchial epithelium in air-liquid interface culture, inhibited epithelial differentiation and induced senescence [43]. The role of other mediators of basal phenotype are to be explored in further studies. Another aspect of branching morphogenesis is the activity of FGFR2, which is critical during mammary gland development in mice along with FGFR1 [44–46]. Lack of *FGFR2* expression in basal cells of the murine epithelium leads to failure in luminal epithelial cell differentiation and inhibition of branching morphogenesis [47, 48]. As *FGFR2* was significantly downregulated in D492^{PXDN}, both in 2D and 3D (Supplementary Fig. 7), this could explain inhibition of both cell plasticity and branching in 3D. However, this remains a subject for future investigations.

While D492^{PXDN} cells attain a basal phenotype, they do not show signs of increased EMT as was demonstrated with RNA sequencing. Physiological levels of *PXDN* in D492 cells are very low (Fig. 2a). However, a previous study showed that *PXDN* expression is significantly upregulated in the mesenchymal D492M cell line [17]. In that study, knock down of PXDN in D492M reduced mesenchymal traits in the cells, resulting in a shift towards epithelial phenotype. This might indicate divergent roles of *PXDN* in D492M, where *PXDN* expression seems to contribute to mesenchymal traits, as opposed to D492 cells where overexpression of *PXDN* results in downregulation of EMT factors. Inhibition of EMT might seem contradictive of the decreased wound closure time observed in the cells during live cell imaging, albeit without the formation of cell protrusions often connected to cell migration [49, 50]. However, epithelial cells can achieve increased motility without losing E-cadherin expression and the epithelial phenotype and without the formation of leading cellular protrusions, by undergoing collective migration [11], which is one of key elements involved in branching morphogenesis in mammary gland development [11, 51, 52]. Collective migration and EMT are two distinct mechanisms that can occur independently although between them there is a spectrum of cell behaviour that can result in a combination of the two [52]. Cheung et al. [53] showed that the basal phenotype, with upregulation of *KRT14* and p63, was essential to collective migration without EMT. Furthermore, induced expression of ΔNp63, an isoform of p63, increased migration in breast cancer cells while inhibiting the formation of cell protrusions [54]. ΔNp63 has also been linked to selective regulation of EMT factors in breast cancer cells, as upregulation of *SNAI2* lead to increased migration while miR-205 mediated inhibition of *ZEB1* and *ZEB2* resulted in EMT suppression [55].

Interestingly, RNA-sequencing of D492^{PXDN} also revealed upregulation of *SNAI2* and downregulation of *ZEB2* which could contribute to the observed phenotype. Cell death is often included in functional analysis of mammary epithelial cell lines as it is often influenced by change in phenotype [56]. A possible mechanism behind the difference in cell death between D492^{PXDN} and D492^{empty} might be found in the peroxidase activity of PXDN. The peroxidase domain of PXDN utilizes H_2O_2 and yields hypochlorous acid, which is a strong oxidant and an inducer of oxidative stress through the generation of reactive oxygen species (ROS) [57–59]. Camptotechin utilizes ROS to provoke cell cycle arrest and apoptosis [60]. Increased expression of PXDN in D492 cells could lead to increased production of ROS which are then readily available to Camptotechin, thereby increasing cell death.

The role of PXDN in breast cancer is previously unexplored. The reason why PXDN only affects DMFS in basal and HER2-positive breast cancers, while also being expressed in estrogen (ER)-positive tumors, is unclear. Triple-negative breast cancers, named by their lack of expression of ER, progesterone and HER2, are typically basal-like and are associated with aggressiveness and poor prognosis [61]. Furthermore, a basal subtype of HER2-positive breast cancers is also related to reduced survival and resistance to therapy [62]. Whether increased expression of *PXDN* could contribute to basal characteristics within these breast cancer subtypes, thereby causing increased metastatic potential, therapy resistance or both, remains a topic for further studies.

Despite being present in both layers of the epithelium *in vivo*, *PXDN* inhibits plasticity in D492 cells and only induces basal phenotype. This cannot be explained effectively with the already known role of *PXDN* as a crosslinker of collagen IV within the extracellular matrix [20], although the basement membrane plays an important regulatory role in epithelial development along with other microenvironmental elements [63, 64]. We therefore propose a novel function of *PXDN* in progenitor cells as an inducer of p63 and basal phenotype and a suppressor of cell plasticity and EMT, providing evidence that *PXDN* might hold a modulatory role in mammary gland development.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10911-021-09507-1.

Acknowledgements We thank Dr. Miklós Geiszt of Semmelweis University in Hungary for kindly gifting the antibody against PXDN used in this paper. We also thank Eygló Gísladóttir at the Institute for Experimental Pathology at Keldur, University of Iceland, for assistance by performing immunohistochemistry of breast tissue and Dr. Snævar Sigurðsson at the University of Iceland Biomedical Center for assistance with RNA-sequencing data processing.

Authors Contribution Conception, design of study and manuscript writing: Anna Karen Sigurdardottir, Thorarinn Gudjonsson and Gunnhildur Asta Traustadottir. Data acquisition: Anna Karen Sigurdardottir, Arni Asbjarnarson, Arna Steinunn Jonasdottir and Hildur Run Helgudottir. Data analysis: Anna Karen Sigurdardottir and Arni Asbjarnarson. Supervision: Thorarinn Gudjonsson and Gunnhildur Asta Traustadottir.

Funding This work was supported by grants from the Icelandic Science and Technology Policy Grant of Excellence #152144051 (T.G.), The Science Fund of the Icelandic Cancer Society (G.A.T.), and 'Göngum saman', a support group for basic research of breast cancer in Iceland (A.K.S., A.S.J.).

Declarations

Ethics Approval Use of primary cells isolated from reduction mammoplasties was approved by the Icelandic National Bioethics Committee VSN-13-057.

Competing Interests Authors declare no conflict of interest.

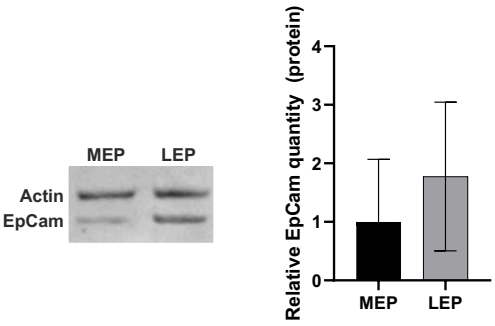
References

1. Gudjonsson T, et al. Isolation, immortalization, and characterization of a human breast epithelial cell line with stem cell properties. Genes Dev. 2002;16(6):693–706.
2. Villadsen R, et al. Evidence for a stem cell hierarchy in the adult human breast. J Cell Biol. 2007;177(1):87–101.
3. Fridriksdottir AJ, et al. Proof of region-specific multipotent progenitors in human breast epithelia. Proc Natl Acad Sci U S A. 2017;114(47):E10102-e10111.
4. Brown KL, et al. Building collagen IV smart scaffolds on the outside of cells. Protein Sci. 2017;26(11):2151–61.
5. Plachot C, et al. Factors necessary to produce basoapical polarity in human glandular epithelium formed in conventional and high-throughput three-dimensional culture: example of the breast epithelium. BMC Biol. 2009;7:77.
6. Macias H, Hinck L. Mammary gland development. Wiley Interdiscip Rev Dev Biol. 2012;1(4):533–57.
7. Javed A, Lteif A. Development of the human breast. Semin Plast Surg. 2013;27(1):5–12.
8. Visvader JE, Stingl J. Mammary stem cells and the differentiation hierarchy: current status and perspectives. Genes Dev. 2014;28(11):1143–58.
9. Inman JL, et al. Mammary gland development: cell fate specification, stem cells and the microenvironment. Development. 2015;142(6):1028–42.

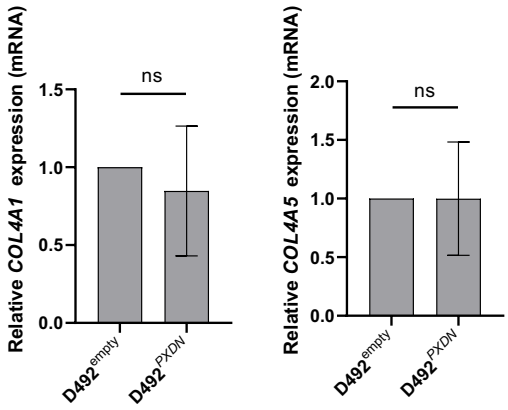
10. Fu NY, et al. Stem Cells and the Differentiation Hierarchy in Mammary Gland Development. Physiol Rev. 2020;100(2):489–523.
11. Ewald AJ, et al. Collective epithelial migration and cell rearrangements drive mammary branching morphogenesis. Dev Cell. 2008;14(4):570–81.
12. Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. Science. 2013;342(6159):1234850.
13. Thiery JP, et al. Epithelial-mesenchymal transitions in development and disease. Cell. 2009;139(5):871–90.
14. Sigurdsson V, et al. Endothelial induced EMT in breast epithelial cells with stem cell properties. PLoS One. 2011;6(9):e23833.
15. Briem E, et al. Application of the D492 Cell Lines to Explore Breast Morphogenesis, EMT and Cancer Progression in 3D Culture. J Mammary Gland Biol Neoplasia. 2019;24(2):139–47.
16. Hilmarsdóttir B, et al. MicroRNA-200c-141 and ΔNp63 are required for breast epithelial differentiation and branching morphogenesis. Developmental Biology. 2015;403(2):150–61.
17. Briem E, et al. MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin. Mech Dev. 2019;155:34–47.
18. Hilmarsdottir B, et al. Functional Role of the microRNA-200 Family in Breast Morphogenesis and Neoplasia. Genes (Basel). 2014;5(3):804–20.
19. Budkova Z, et al. Expression of ncRNAs on the DLK1-DIO3 Locus Is Associated With Basal and Mesenchymal Phenotype in Breast Epithelial Progenitor Cells. Front Cell Dev Biol. 2020;8:461.
20. Bhave G, et al. Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis. Nat Chem Biol. 2012;8(9):784–90.
21. Khoshnoodi J, Pedchenko V, Hudson BG. Mammalian collagen IV. Microsc Res Tech. 2008;71(5):357–70.
22. Colon S, Page-McCaw P, Bhave G. Role of Hypohalous Acids in Basement Membrane Homeostasis. Antioxid Redox Signal. 2017;27(12):839–54.
23. Péterfi Z, et al. Peroxidasin is secreted and incorporated into the extracellular matrix of myofibroblasts and fibrotic kidney. Am J Pathol. 2009;175(2):725–35.
24. Jayachandran A, et al. Identifying and targeting determinants of melanoma cellular invasion. Oncotarget. 2016;7(27):41186–202.
25. Zheng YZ, Liang L. High expression of PXDN is associated with poor prognosis and promotes proliferation, invasion as well as migration in ovarian cancer. Ann Diagn Pathol. 2018;34:161–5.
26. Sigurdsson V, et al. Human breast microvascular endothelial cells retain phenotypic traits in long-term finite life span culture. In Vitro Cell Dev Biol Anim. 2006;42(10):332–40.
27. Rønnov-Jessen L, Petersen OW, Bissell MJ. Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. Physiol Rev. 1996;76(1):69–125.
28. Ringnér M, et al. GOBO: gene expression-based outcome for breast cancer online. PLoS One, 2011. 6(3): p. e17911.
29. Péchoux C, et al. Human mammary luminal epithelial cells contain progenitors to myoepithelial cells. Dev Biol. 1999;206(1):88–99.
30. Chou J, Provot S, Werb Z. GATA3 in development and cancer differentiation: cells GATA have it! Journal of cellular physiology. 2010;222(1):42–9.
31. Romano RA, et al. An active role of the DeltaN isoform of p63 in regulating basal keratin genes K5 and K14 and directing epidermal cell fate. PLoS One. 2009;4(5):e5623.
32. Romano RA, Birkaya B, Sinha S. A functional enhancer of keratin14 is a direct transcriptional target of deltaNp63. J Invest Dermatol. 2007;127(5):1175–86.
33. Hennighausen L, Robinson GW. Signaling pathways in mammary gland development. Dev Cell. 2001;1(4):467–75.
34. Howard B, Ashworth A. Signalling pathways implicated in early mammary gland morphogenesis and breast cancer. PLoS Genet. 2006;2(8): e112.
35. Friedl P, Mayor R. Tuning Collective Cell Migration by Cell-Cell Junction Regulation. Cold Spring Harb Perspect Biol. 2017;9(4).
36. Van Keymeulen A, et al. Distinct stem cells contribute to mammary gland development and maintenance. Nature. 2011;479(7372):189–93.
37. Visvader JE, Lindeman GJ. Mammary stem cells and mammopoiesis. Cancer Res. 2006;66(20):9798–801.
38. Shackleton M, et al. Generation of a functional mammary gland from a single stem cell. Nature. 2006;439(7072):84–8.
39. Yang A, et al. p63 is essential for regenerative proliferation in limb, craniofacial and epithelial development. Nature. 1999;398(6729):714–8.
40. Yalcin-Ozuysal O, et al. Antagonistic roles of Notch and p63 in controlling mammary epithelial cell fates. Cell Death Differ. 2010;17(10):1600–12.
41. Boecker W, et al. Spatially correlated phenotyping reveals K5-positive luminal progenitor cells and p63–K5/14-positive stem cell-like cells in human breast epithelium. Lab Invest. 2018;98(8):1065–75.
42. Boecker W, et al. Multicolor immunofluorescence reveals that p63- and/or K5-positive progenitor cells contribute to normal breast epithelium and usual ductal hyperplasia but not to low-grade intraepithelial neoplasia of the breast. Virchows Arch. 2017;470(5):493–504.
43. Arason AJ, et al. deltaNp63 has a role in maintaining epithelial integrity in airway epithelium. PloS One. 2014;9(2):e88683–e88683.
44. Dillon C, Spencer-Dene B, Dickson C. A crucial role for fibroblast growth factor signaling in embryonic mammary gland development. J Mammary Gland Biol Neoplasia. 2004;9(2):207–15.
45. Parsa S, et al. Terminal end bud maintenance in mammary gland is dependent upon FGFR2b signaling. Developmental Biology. 2008;317(1):121–31.
46. Lu P, et al. Genetic mosaic analysis reveals FGF receptor 2 function in terminal end buds during mammary gland branching morphogenesis. Developmental Biology. 2008;321(1):77–87.
47. Pond AC, et al. Fibroblast growth factor receptor signaling is essential for normal mammary gland development and stem cell function. Stem Cells. 2013;31(1):178–89.
48. Zhang X, et al. FGF ligands of the postnatal mammary stroma regulate distinct aspects of epithelial morphogenesis. Development. 2014;141(17):3352–62.
49. Mitchison TJ, Cramer LP. Actin-based cell motility and cell locomotion. Cell. 1996;84(3):371–9.
50. Prasad M, Montell DJ. Cellular and molecular mechanisms of border cell migration analyzed using time-lapse live-cell imaging. Dev Cell. 2007;12(6):997–1005.
51. Ilina O, Friedl P. Mechanisms of collective cell migration at a glance. J Cell Sci. 2009;122(Pt 18):3203–8.
52. Campbell K, Casanova J. A common framework for EMT and collective cell migration. Development. 2016;143(23):4291.
53. Cheung KJ, et al. Collective Invasion in Breast Cancer Requires a Conserved Basal Epithelial Program. Cell. 2013;155(7):1639-1651.
54. Westcott JM, et al. ΔNp63-Regulated Epithelial-to-Mesenchymal Transition State Heterogeneity Confers a Leader-Follower Relationship That Drives Collective Invasion. Cancer Res. 2020;80(18):3933–44.
55. Dang TT, et al. ΔNp63α Promotes Breast Cancer Cell Motility through the Selective Activation of Components of the Epithelial-to-Mesenchymal Transition Program. Cancer Res. 2015;75(18):3925–35.
56. Hanahan D, Robert A. Weinberg, Hallmarks of Cancer: The Next Generation. Cell. 2011;144(5):646–74.

57. Ma QL, Zhang GG, Peng J. Vascular peroxidase 1: a novel enzyme in promoting oxidative stress in cardiovascular system. Trends Cardiovasc Med. 2013;23(5):179–83.
58. Cheng G, et al. Identification and characterization of VPO1, a new animal heme-containing peroxidase. Free Radic Biol Med. 2008;45(12):1682–94.
59. Yuan W, et al. Hypochlorous acid converts the gamma-glutamyl group of glutathione disulfide to 5-hydroxybutyrolactam, a potential marker for neutrophil activation. J Biol Chem. 2009;284(39):26908–17.
60. Prasad Tharanga Jayasooriya RG, et al. Camptothecin induces G(2)/M phase arrest through the ATM-Chk2-Cdc25C axis as a result of autophagy-induced cytoprotection: Implications of reactive oxygen species. Oncotarget. 2018;9(31):21744-21757.
61. Rakha EA, Reis-Filho JS, Ellis IO. Basal-like breast cancer: a critical review. J Clin Oncol. 2008;26(15):2568–81.
62. Sagara Y, et al. Effectiveness of neo-adjuvant systemic therapy with trastuzumab for basal HER2 type breast cancer: results from retrospective cohort study of Japan Breast Cancer Research Group (JBCRG)-C03. Breast Cancer Res Treat. 2018;171(3):675–83.
63. Vidi PA, Bissell MJ, Lelièvre SA. Three-dimensional culture of human breast epithelial cells: the how and the why. Methods Mol Biol. 2013;945:193–219.
64. McCave EJ, et al. The normal microenvironment directs mammary gland development. J Mammary Gland Biol Neoplasia. 2010;15(3):291–9.

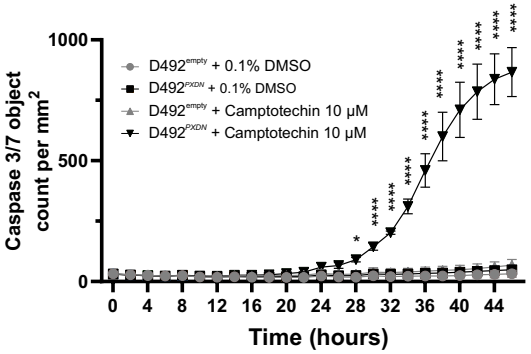
Supplementary fig. 1



Supplementary fig. 2

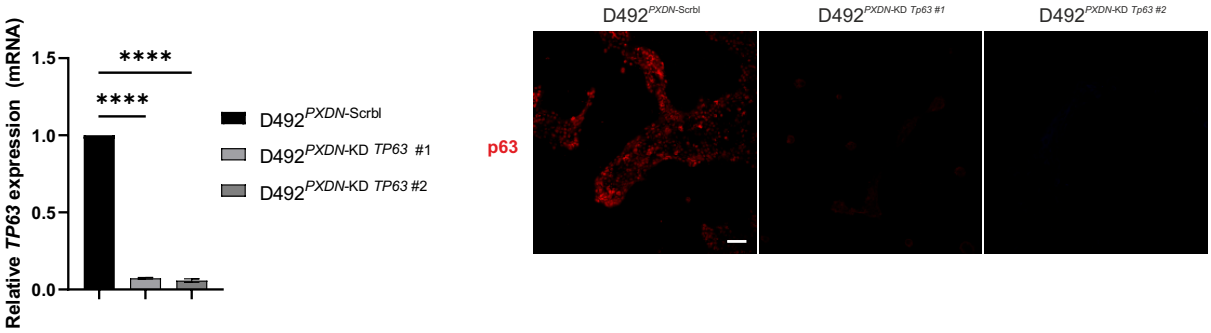


Supplementary fig. 3

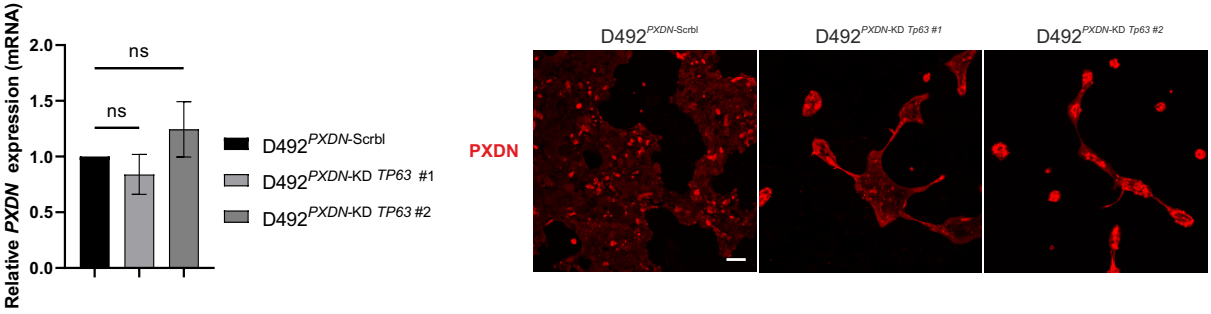


Supplementary fig. 4

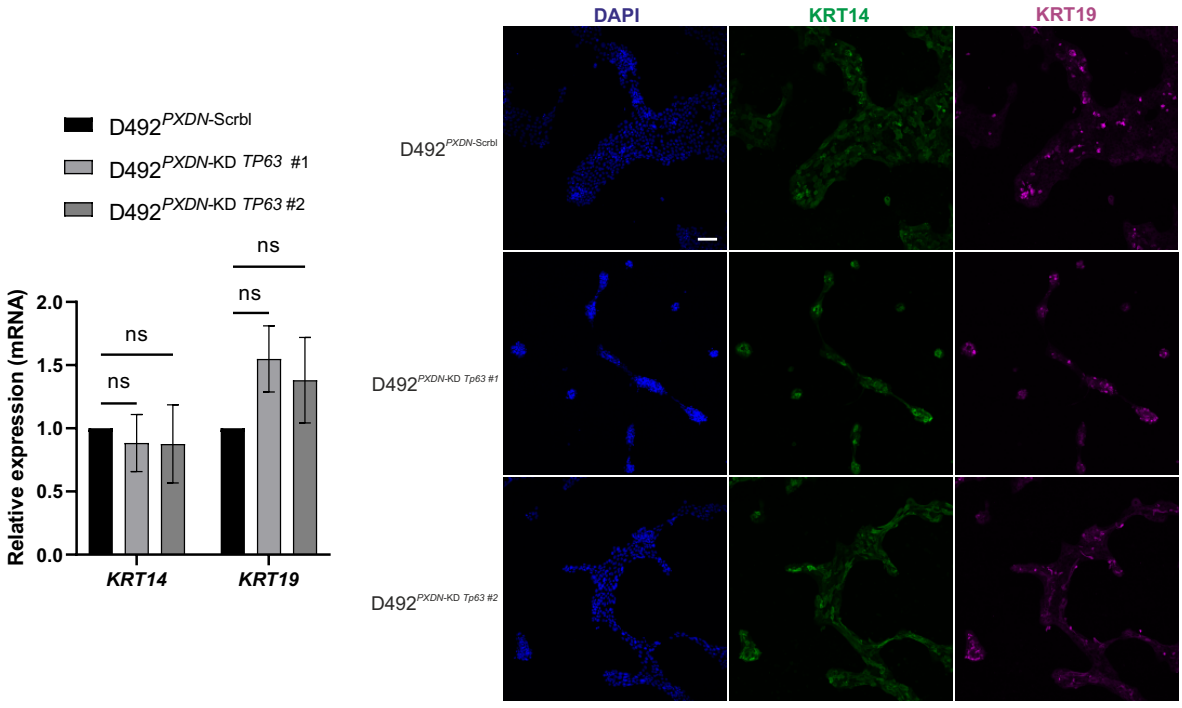
a.



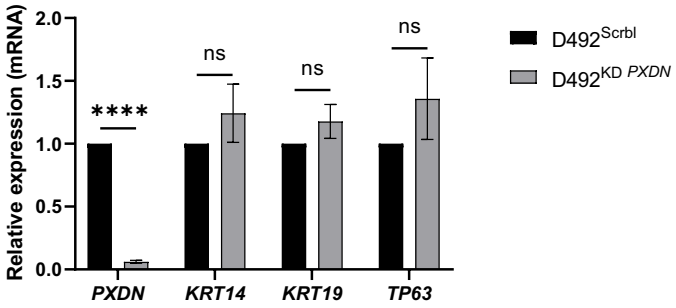
b.



c.

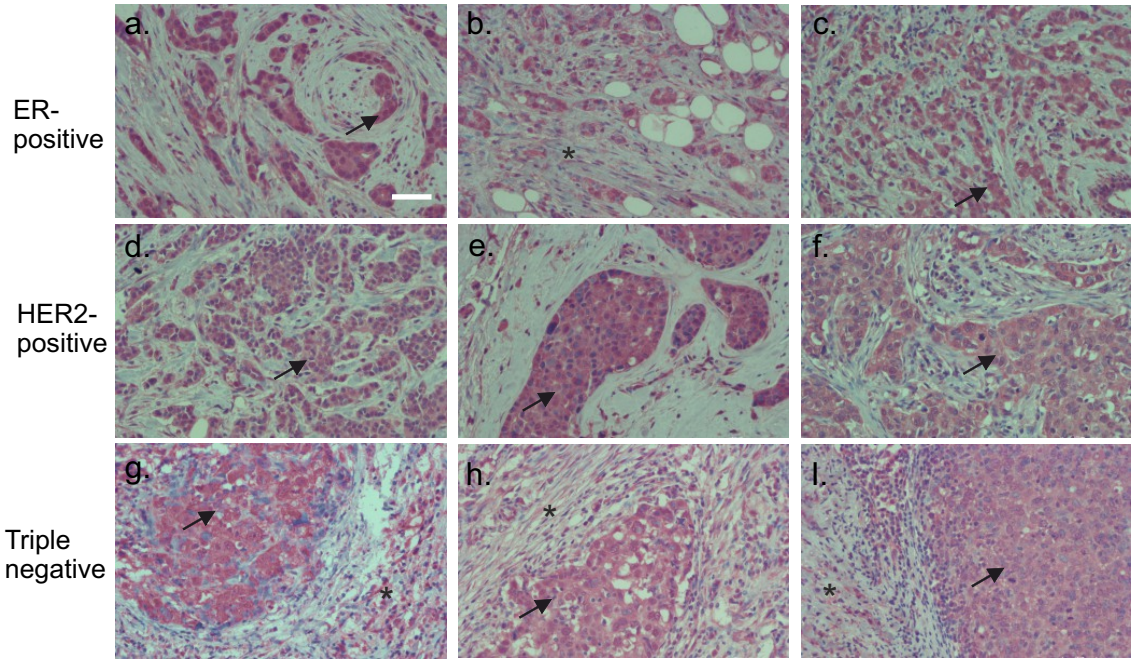


Supplementary fig. 5

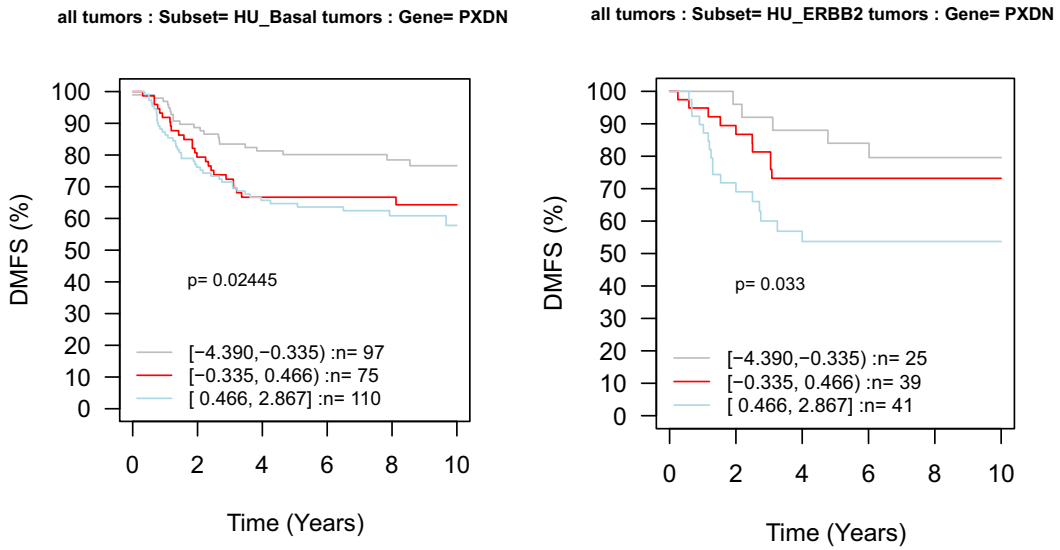


Supplementary fig. 6

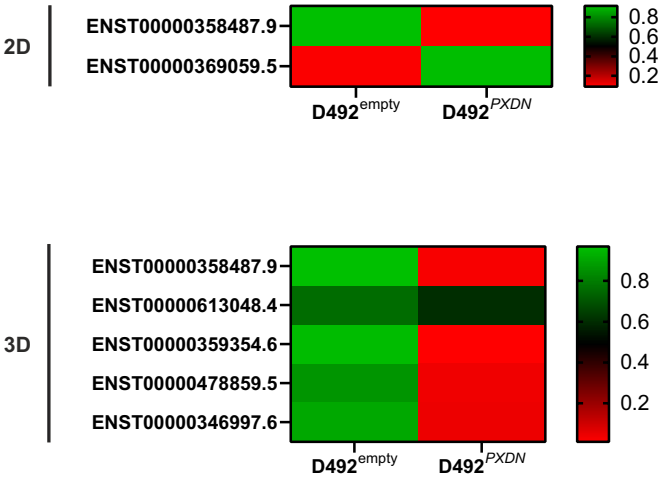
a.



b.



Supplementary fig. 7



Supplementary fig. 1 Western blot confirming magnetic sorting of EpCam-positive (LEP) and EpCam-negative (MEP) primary mammary epithelial cells.

Supplementary fig. 2 Overexpression of *PXDN* did not signficantly alter expression of collagen IV genes. Expression of *COL4A1* and *COL4A5* did not change significantly in D492^{PXDN} compared to D492^{empty}. Statistical significance in qRT-PCR was determined by unpaired Student's t-test and data is presented as an average of three replicated experiments (mean ± SD).

Supplementary fig. 3 D492^{PXDN} cells were more sensitive to chemically induced apoptosis. *PXDN* increased sensitivity to chemically induced apoptosis in D492 cells when treated with Camptotechin. Data was analyzed in IncuCyte Zoom and is presented as Caspase 3/7 object count per mm^2. Statistical significance was determined using two-way ANOVA followed by Tukey's post hoc analysis (*$p \leq 0.05$, ****$p \leq 0.0001$) and data is presented as an average of three independent experiments (mean ± SD, n=3).

Supplementary fig. 4 Transient knock down of *TP63* in D492^{PXDN} did not significantly affect expression of *PXDN* or keratins

a. Confirmation of transient knock down of *TP63* in D492^{PXDN} with two different siRNAs with qRT-PCR and immunofluoresence staining. Statistical significance was determined using One-way ANOVA (****$p \leq 0.0001$) and data is presented as an average of three replicated experiments (mean ± SD). Scale bar = 100 µm.
b. *TP63* knock down did not significantly alter expression of *PXDN* as measured via qRT-PCR and immunofluoresence staining did not show a difference in PXDN signal intensity. Statistical significance was determined using One-way ANOVA and data is presented as an average of three replicated experiments (mean ± SD). Scale bar = 100 µm.
c. *TP63* knock down did not significantly affect expression of *KRT14 or KRT19* as measured via qRT-PCR and immunofluoresence staining did not show a difference in signal intensity. Statistical significance was determined using One-way ANOVA and data is presented as an average of three replicated experiments (mean ± SD). Scale bar = 100 µm.

Supplementary fig. 5 Transient siRNA knock down of *PXDN* in D492 did not significantly affect expression of keratins or *TP63*. qRT-PCR confirmed siRNA mediated transient knock down of *PXDN* in D492 cells and also revealed no significant change in expression of *KRT14, KRT19* and *TP63.* Statistical significance in qRT-PCR was determined by unpaired Student's t-test (****$p \leq 0.0001$) and data is presented as an average of three replicated experiments (mean ± SD).

Supplementary fig. 6 *PXDN* expression in breast cancer

a. PXDN was expressed in three different breast cancer subtypes: ER-positive, HER2-positive and Triple negative. Immunohistochemistry on paraffin embedded tumor tissue revealed positive PXDN signal (red) in all cancer cells (examples depicted with arrows) in all tumors. Tumors b., g., h. and i. also had PXDN-positive fibrils in the intratumor stroma (asterix). Tissue slides were counterstained with hematoxylin. Scale bar = 100 µm.

b. Women with basal tumors and HER2-positive breast cancers that express high levels of *PXDN* had significantly lower distant metastatsis free survival (DMFS) than women with low *PXDN* expressing tumors. Data from the GOBO database.

Supplementary fig. 7 Heatmaps showing differential expression of *FGFR2* transcripts in D492^{empty} and D492^{PXDN} in 2D and 3D. In 2D, one of two significantly differentially expressed transcripts was downregulated in D492^{PXDN}. However, all transcripts were downregulated in D492^{PXDN} in 3D.

Article III

frontiers in Cell and Developmental Biology

ORIGINAL RESEARCH
published: 16 June 2020
doi: 10.3389/fcell.2020.00461

Expression of ncRNAs on the DLK1-DIO3 Locus Is Associated With Basal and Mesenchymal Phenotype in Breast Epithelial Progenitor Cells

*Zuzana Budkova[1], Anna Karen Sigurdardottir[1], Eirikur Briem[1], Jon Thor Bergthorsson[2], Snævar Sigurdsson[1], Magnus Karl Magnusson[3], Gunnhildur Asta Traustadottir[1], Thorarinn Gudjonsson[1,2] and Bylgja Hilmarsdottir[1,4]**

[1] Stem Cell Research Unit, Biomedical Center, Department of Anatomy, Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland, [2] Department of Laboratory Hematology, Landspitali – University Hospital, Reykjavik, Iceland, [3] Department of Pharmacology and Toxicology, Faculty of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland, [4] Department of Pathology, Landspitali – University Hospital, Reykjavik, Iceland

OPEN ACCESS

Edited by:
Zuzana Koledova,
Masaryk University, Czechia

Reviewed by:
Lone Rønnov-Jessen,
University of Copenhagen, Denmark
Giulia Ricci,
University of Campania Luigi Vanvitelli, Italy
Zhongxin Lu,
Huazhong University of Science and Technology, China

***Correspondence:**
Bylgja Hilmarsdottir
bylgjahi@landspitali.is

Specialty section:
This article was submitted to Stem Cell Research, a section of the journal Frontiers in Cell and Developmental Biology

Received: 12 December 2019
Accepted: 18 May 2020
Published: 16 June 2020

Citation:
Budkova Z, Sigurdardottir AK, Briem E, Bergthorsson JT, Sigurdsson S, Magnusson MK, Traustadottir GA, Gudjonsson T and Hilmarsdottir B (2020) Expression of ncRNAs on the DLK1-DIO3 Locus Is Associated With Basal and Mesenchymal Phenotype in Breast Epithelial Progenitor Cells. Front. Cell Dev. Biol. 8:461. doi: 10.3389/fcell.2020.00461

Epithelial-to-mesenchymal transition (EMT) and its reversed process mesenchymal-to-epithelial transition (MET) play a critical role in epithelial plasticity during development and cancer progression. Among important regulators of these cellular processes are non-coding RNAs (ncRNAs). The imprinted DLK1-DIO3 locus, containing numerous maternally expressed ncRNAs including the lncRNA maternally expressed gene 3 (*MEG3*) and a cluster of over 50 miRNAs, has been shown to be a modulator of stemness in embryonic stem cells and in cancer progression, potentially through the tumor suppressor role of *MEG3*. In this study we analyzed the expression pattern and functional role of ncRNAs from the DLK1-DIO3 locus in epithelial plasticity of the breast. We studied their expression in various cell types of breast tissue and revisit the role of the locus in EMT/MET using a breast epithelial progenitor cell line (D492) and its isogenic mesenchymal derivative (D492M). Marked upregulation of ncRNAs from the DLK1-DIO3 locus was seen after EMT induction in two cell line models of EMT. In addition, the expression of *MEG3* and the maternally expressed ncRNAs was higher in stromal cells compared to epithelial cell types in primary breast tissue. We also show that expression of *MEG3* is concomitant with the expression of the ncRNAs from the DLK1-DIO3 locus and its expression is therefore likely indicative of activation of all ncRNAs at the locus. *MEG3* expression is correlated with stromal markers in normal tissue and breast cancer tissue and negatively correlated with the survival of breast cancer patients in two different cohorts. Overexpression of *MEG3* using CRISPR activation in a breast epithelial cell line induced partial EMT and enriched for a basal-like phenotype. Conversely, knock down of *MEG3* using CRISPR inhibition in a mesenchymal cell line reduced the mesenchymal and basal-like phenotype of the cell line. In summary our study shows that maternally expressed ncRNAs are markers of EMT and suggests that *MEG3* is a novel regulator of EMT/MET in breast tissue. Nevertheless, further studies are needed to fully dissect the molecular pathways influenced by non-coding RNAs at the DLK1-DIO3 locus in breast tissue.

Keywords: DLK1-DIO3 locus, *MEG3*, ncRNAs, epithelial plasticity, breast progenitor cells

INTRODUCTION

Breast cancer is the most common cancer in women and the second most common cancer overall (Ghoncheh et al., 2016). Despite major advances in diagnosis and treatment of cancer in recent years, metastasis and development of resistance to cancer therapies continues to be a challenge, causing over 90% of all cancer-related deaths (Ben-Jacob et al., 2012). A major contributing factor to metastasis and drug resistance is the heterogeneity and plasticity of the cells within tumors (Dagogo-Jack and Shaw, 2018). Epithelial-to-mesenchymal transition (EMT), is a developmental process that can be hijacked by cancer cells (Zeisberg and Kalluri, 2004; Moustakas and Heldin, 2007; Radisky et al., 2007). Generally, cells undergoing EMT, acquire increased migration and invasive properties and show increased resistance to apoptosis (Robson et al., 2006; Cao et al., 2016). Through these processes, EMT is considered a major mediator of phenotypic plasticity in cancer cells, metastatic formation and drug resistance (Mani et al., 2008; Scheel and Weinberg, 2012; Ansieau, 2013; Nieto et al., 2016; Lu and Kang, 2019). Recently, hybrid E/M (or partial EMT) cells have been shown to have even more metastatic and stem cell potential compared to the full epithelial or mesenchymal phenotype (Pastushenko et al., 2018). A reversed program, mesenchymal-to–epithelial transition (MET) is considered to facilitate colonization in secondary sites and reverse the plastic mesenchymal phenotype back to an epithelial state (Lu and Kang, 2019). This, however, is debated and further studies will increase our knowledge of the role of EMT/MET in cancer progression and metastasis.

EMT can be initiated through intrinsic factors such as expression of EMT related transcription factors (SNAI1, SNAI2, TWIST1, ZEB1, or ZEB2), cadherin switch from E-cadherin (CDH1) to N-cadherin (CDH2) or through epigenetic mechanisms. It can also be brought on by extrinsic factors derived from the microenvironment, such as secreted soluble factors: transforming growth factor-β (TGF-β), epidermal growth factor (EGF), fibroblast growth factors (FGFs), hepatocyte growth factor (HGF) or Wnt signaling factors (Moustakas and Heldin, 2007; Peinado et al., 2007; De Craene and Berx, 2013; Wang and Zhou, 2013; Williams et al., 2019).

Non-coding RNAs (ncRNAs) are among intrinsic regulators of EMT (Zaravinos, 2015). It is increasingly apparent that the ncRNAs are crucial in normal development and disease, but its mechanistic mode of action is largely unknown (Liz and Esteller, 2016). The two major classes of non-coding RNAs are long non-coding RNA (lncRNAs) and microRNA (miRNAs). Accumulating evidence suggests that lncRNAs function in a broad range of cellular processes such as cell growth, survival, migration, invasion and differentiation (Mercer et al., 2009; Sun et al., 2013; Di Gesualdo et al., 2014; Fatica and Bozzoni, 2014). LncRNAs are defined by the size of their transcripts and are longer than 200 nucleotides (nt), with no protein-coding function (Eades et al., 2014). Unlike microRNAs, lncRNAs are poorly conserved, but function in a regulatory network at the transcriptional, post-transcriptional, and translational level. miRNAs are 22 nt long RNA molecules that regulate expression post-transcriptionally primarily by binding to three prime untranslated region (3′UTR) of target genes (Bartel, 2009).

The imprinted DLK1-DIO3 locus located on chromosome 14 contains three paternally expressed protein-coding genes (*DLK1*, *RTL1*, *DIO3*) and numerous maternally expressed non-coding genes, including the lncRNA maternally expressed gene 3 (*MEG3*), and a cluster of over 50 miRNAs (Zhang et al., 2010; Dill and Naya, 2018; Baulina et al., 2019; Li et al., 2019).

The DLK1-DIO3 locus has been described as an important contributor to pluripotency and stemness in embryonic stem cells (ESCs) (Kaneko et al., 2014). It discriminates between mouse induced pluripotent stem cells (iPCS) and mouse ESCs, where genes from the locus were strongly repressed in iPSC clones compared to ES clones (Liu et al., 2010; Stadtfeld et al., 2010a). Furthermore, activation of maternally expressed genes from the locus is a strong indicator of the developmental potential of iPSC (Kang et al., 2009). miRNAs from the DLK1-DIO3 locus have been shown to promote pluripotency by inhibition of differentiation and stimulation of self-renewal in mouse ES cells (Moradi et al., 2017) and were found to be increased in tumor-originating cancer cells from lung adenocarcinoma (Valdmanis et al., 2015).

MEG3 is a potential tumor suppressor gene in several cancer types, mainly through the observation that *MEG3* expression is lower in various tumor tissues compared with non-tumor tissues of the same origin (Sheng et al., 2014; Sun et al., 2014, 2016; Yin et al., 2015; Chak et al., 2017; Molina-Pinelo et al., 2018). The tumor suppressor role of *MEG3* is ascribed to stabilization of p53 with inhibition of proliferation and promotion of apoptosis (Zhang et al., 2003, 2010; Zhou et al., 2007; Wang et al., 2012; Sun et al., 2016).

MEG3 was reported to positively regulate EMT in lung (Terashima et al., 2017) and ovarian (Mitra et al., 2017) cancer. Furthermore, *MEG3* has been shown to contribute to the development of osteosarcoma through increased migration, invasion and decreased apoptosis (Wang and Kong, 2018). Higher levels of *MEG3* were detected in plasma from colorectal cancer patients compared with non-cancerous controls (Liu et al., 2019).

D492 is a primary breast epithelial cell line, immortalized with the E6 and E7 oncogenes from the human papilloma virus 16 (Gudjonsson et al., 2002). Therefore, the p53 protein, which mediates the previously described tumor suppressor role of *MEG3*, is repressed in this cell line. D492 can generate both luminal and basal/myoepithelial cells in monolayer and 3D culture, expressing luminal or myoepithelial keratins such as keratin 19 and keratin 14, respectively. Furthermore, when D492 cells are co-cultured with endothelial cells, they, can generate spindle-shaped colonies with EMT phenotype. D492M (mesenchymal) was established from one such spindle-shaped colony (Sigurdsson et al., 2011). D492M is a phenotypically stable EMT cell line. It has lost epithelial markers such as keratins, E-cadherin and TP63, and gained expression of mesenchymal markers such as N-cadherin (Sigurdsson et al., 2011; Hilmarsdottir et al., 2015). D492M has acquired classical properties of cancer stem cells, such as increased CD44/CD24 ratio, anchorage independent growth, resistance to apoptosis

and increased migration/invasion (Sigurdsson et al., 2011). D492 serves as a model for branching morphogenesis and together D492 and D492M represent a unique EMT model of isogenic cell lines with an epithelial and mesenchymal phenotype, respectively (Briem et al., 2019b). The ability of D492 to undergo mesenchymal transition upon endothelial stimulation makes it a valuable cell model to study EMT induced by extrinsic factors, although it is important to note that neither D492 nor D492M are tumorigenic in mice.

In this study, we describe a new role for the DLK1-DIO3 locus in EMT and phenotypic plasticity of breast cells. Following EMT in breast epithelial cell lines, expression of the ncRNAs at the DLK1-DIO3 locus was increased. In addition, *MEG3* was highly expressed in stromal cells in breast tissue and its expression correlated with decreased survival in breast cancer. Moreover, increased expression of the ncRNAs at the DLK1-DIO3 locus in a breast epithelial progenitor cell line promoted cellular plasticity and induced partial EMT. Collectively, our study provides a further understanding of the role of the DLK1-DIO3 locus in cellular phenotype of breast cells and might provide important insight into novel therapeutic targets aimed at overcoming heterogeneity and therapy resistance in breast cancer.

MATERIALS AND METHODS

Cell Lines

Both D492 and D492M were cultured in H14 medium, as described previously (Gudjonsson et al., 2002; Sigurdsson et al., 2011) in flasks coated with collagen I (Advanced BioMFatrix, 5005-B). HEK-293T cell were cultured in Dulbecco's Modified Eagle Medium (DMEM), high glucose, GlutaMAX (TM), pyruvate (Gibco, 31966), supplemented with 10% Fetal bovine serum (FBS), penicillin and streptomycin (Gibco, 15140-122). Primary Human umbilical vein endothelial cells (HUVECs) were obtained from Landspitali, University Hospital in Reykjavik, Iceland, (with informed consent, approved by Landspitali Ethical Committee No. 35/2013), cultured in Endothelial Growth Medium 2 (EGM2) media (Lonza, CC-3162) supplemented with growth factors and 5% FBS, further referred to as EGM5 medium as previously described (Sigurdsson et al., 2011). HMLE (Elenbaas et al., 2001) is epithelial progenitor cell line, from which was derived mesenchymal cell line HMLEmes after stable induction of EMT-TF (Mani et al., 2008). HMLE and HMLEmes were cultured in chemically defined HMLE media, containing DMEM/F12 with penicillin and streptomycin and growth factors Insulin (Sigma, I1882) 10 μg/ml, EGF (Peprotech, AF-100-15) 10 ng/ml, Hydrocortisone (Sigma, H0888) 500 ng/ml.

Primary human luminal-epithelial cells (LEP), myoepithelial cells (MEP), breast endothelial cells (BRENCs) and fibroblast were isolated from breast reduction mammoplasties (with informed consent, approved by the Icelandic National Bioethics Committee VSN-13-057) as previously described (Sigurdsson et al., 2011) and maintained in chemically defined medium 3 (CDM3) and chemically defined medium 4 (CDM4) as previously described (Pechoux et al., 1999; Ingthorsson et al., 2010). All cells were maintained in an incubator with 5% CO_2 at 37°C.

3D Cultures/Mammosphere Assays

3D cultures were carried out in a 48-well plate format (Corning, 353078) in growth factor reduced reconstituted basement membrane rBM (further referred to as Matrigel, Corning, 354230). 5–10 × 10^3 cells were seeded in 150 μl of Matrigel per well. Plate was incubated in 5% CO_2 at 37°C for 15 min to solidify the Matrigel and then 300 μl of H14 media was added on top. The cells were grown for 3 weeks and pictures were taken on day 1, 7, 14, and 21. Cell culture media was changed three times per week. The colonies were quantified at day 14. The total number of cells was converted into percentage.

For co-culture experiments, 0.5 × 10^3 of the epithelial cells were co-cultured with 1 × 10^5 of endothelial cells (HUVECs) and were resuspended in 150 μl of Matrigel. Plate was incubated in 5% CO_2 at 37°C for 15 min to solidify the Matrigel and then 300 μl EGM5 media was added on top. HUVECs cultured in Matrigel are viable, however, quiescent, having supporting role in the epithelial cells' proliferation. The effect of *MEG3* was quantified by counting all colonies bigger than 100 μm.

Total RNAseq and Analysis of the Data

The gene microarray expression analysis from D492 and D492M was published previously from our group by Sigurdsson and colleagues (Sigurdsson et al., 2011) and the total RNA-sequencing comparing D492 and D492M was published by Halldorsson and colleagues (Halldorsson et al., 2017).

The RNA was extracted using Tri-Reagent (Thermo Fisher Scientific, AM9738) from 5 replicates for each cell line. Whole Transcriptome Sequencing of $D492M^{KD-CTRL}$ and $D492M^{KD-MEG3}$ was performed in deCODE genetics (Reykjavik, Iceland). RNA sequencing reads were mapped to the reference genome (Ensembl primary assembly, version GRCh38) using STAR version 2.6.1 (Dobin et al., 2013). The program htseq-count (Anders et al., 2015) was used to quantify how many reads match each gene in an annotation file (Ensemble version GRCh38.96). The data from htseq-count was imported into R (R Development Core Team, 2015) and differential expression (DE) analysis on $D492M^{KD-CTRL}$ vs $D492^{KD-MEG3}$ was performed using DESeq2 (Love et al., 2014). Prior to DE analysis, genes with expression less than two reads were discarded. *P*-values were corrected for multiple testing using the false discovery rate (FDR) method. To compare gene expression from $D492M^{KD-CTRL}$ vs $D492M^{KD-MEG3}$ a volcano plot was generated. *P* value cut-off of 0.05 was applied. Volcano plot over all data ($p < 0.05$) was made in R using the EnhancedVolcano package from BioConductor. The top ten most upregulated and downregulated genes according log2 fold change were labeled. Gene Set Enrichment Analysis (GSEA) was applied to identify enrichment of gene signatures. Comparative analysis was investigated using the "Hallmark" database. The list of significantly expressed pathways is presented as a bar plot.

Quantitative RT-PCR Analysis

Total RNA was extracted with Tri-Reagent (Thermo Fisher Scientific, AM9738). 1 μg of RNA of each sample was reverse transcribed into complementary DNA (cDNA), using Random

Hexamers (Thermo Fisher Scientific, N8080127) and SuperScript IV Reverse Transcriptase (Thermo Fisher Scientific, 18090-200) kit and subjected to quantitative real time PCR (qRT-PCR) using Sybr Green dye Luna® Universal qPCR Master Mix (NEB, M3003L) or TaqMan probes Luna® Universal Probe qPCR Master Mix (NEB, M3004L) according to manufacturer's protocol. *GAPDH* was used as control for gene expression. For assaying the relative expression of each gene, the $2^{-\Delta\Delta Ct}$ was determined using an ABI 7500 instrument (Applied Biosystems).

List of Primers

TaqMAN: *ZEB1* (Thermo Fisher Scientific, Hs00232783_m1), *ZEB2* (Thermo Fisher Scientific, Hs00207691_m1), *SNAI1* (Thermo Fisher Scientific, Hs00195591_m1), *SNAI2* (Thermo Fisher Scientific, Hs00950344_m1), *TWIST1* (Thermo Fisher Scientific, Hs01675818_s1), *GAPDH* (Thermo Fisher Scientific, 4326317E).

SYBR Green: *KRT14* (IDT, Hs.PT.58.4592110), *KRT19* (IDT,Hs.PT.58.4188708), *MEG3* ex 10-11 (IDT, Hs.PT.58.25190740), *GAPDH* (IDT, Hs.PT.39a.22214836), *KRT5* (IDT, Hs.PT.58.14446018), *TP63* (IDT, Hs.PT.58.2966111), *CDH3* (IDT, Hs.PT.58.39234242).

Small RNAseq

The Microarray of small RNA data was published previously by our group by Hilmarsdottir and colleagues (Hilmarsdottir et al., 2015) and the small RNAseq data was published previously by Briem and colleagues (Briem et al., 2019a).

miRNA qRT PCR

Total RNA was extracted with Tri-Reagent (Thermo Fisher Scientific, AM9738). The RNA was reverse transcribed using miRCURY LNA RT Kit (Qiagen, 339340) for cDNA synthesis reactions, according to manufacturer's protocol. Quantitative RT-PCR analysis of miRNAs was performed using miRCURY LNA SYBR Green PCR Kit (Qiagen, 339346), according to manufacturer's protocol. Gene expression levels were quantified using primers for: hsa-miR-127-3p (Qiagen, YP00204048), hsa-miR-409-3p (Qiagen, YP00204358), hsa-miR-411-5p (Qiagen, YP00204531), hsa-miR-493-3p (Qiagen, YP00204557). Normalization was done with U6 snRNA (Qiagen, YP00203907). The $2^{-\Delta\Delta Ct}$ was used determined using ABI 7500 instrument (Applied Biosystems) to calculate the relative expression of each gene.

Allele Specific Expression Analysis

Total RNA was extracted with Tri-Reagent (Thermo Fisher Scientific, AM9738) and reverse transcription done using 1 μg of DNase I-treated total RNA using random hexamers (Thermo Fisher Scientific, N8080127) and SuperScript II Reverse Transcriptase (Thermo Fisher Scientific, 18064022) according to the manufacturer's instructions. PCR primers were designed using Primer3 and Pyrosequencing primers were designed using PyroMark Assay Design 2.0 (Qiagen). The reverse PCR primer had a 5′-biotin modification and was HPLC-purified. Primers were synthesized by IDT 5′-TGGCCTTTTCTTCTCCTGAA, 5′-/5Biosg/TGACACATGGAAAGCACCAT and sequencing primer 5′-TCCGGGGTTACTGCCCT-3′. Polymerase chain reactions were performed in 50 μl using 10 ng of diluted cDNA or 10 ng of DNA, 1 U DreamTaq DNA polymerase (Fermentas, EP0701), 1X PCR buffer, 200 μM of dNTPs and 0,5 μM of each PCR primer. The following PCR protocol was used: 94°C for 2 min, followed by 50 cycles of 94°C for 1 min, 60°C for 1 min, 72°C for 1 min and 72°C for 9 min. To check the quality of the amplification, PCR products were analyzed by gel electrophoresis. Pyrosequencing were sequenced using the PyroMark Q24 system (Qiagen), following the manufacturer's instructions. For the ASE SNP, DNA and RNA (cDNA) were pyrosequenced simultaneously. The proportions of individual alleles for the SNP were obtained using the PyroMark Q24 software version 1.0.10 (Qiagen). Genomic DNA from D492M was examined to confirm the heterozygosity.

Clinical Cohort

RNA from breast cancer patients (diagnosed in the years 1987–2003) and relevant patient data was obtained from the Department of Pathology Landspitali – The National University Hospital of Iceland. Informed consent was obtained from patients involved in this study according to the national guidelines. The study was approved by The Icelandic Data Protection Commission (2001/523 and 2002/463) as well as the National Bioethics Committee of Iceland (VSN-11-105-V2). 119 samples were used in the study assigned to the following subgroup: 33 luminal A, 24 luminal B, 22 Basal, 12 ErbB2, 10 Normal and 18 not classified. cDNA was synthesized from 2 μg of total RNA using Random Hexamers primers (Thermo Fisher Scientific, N8080127) and RevertAid First Strand cDNA Synthesis Kit (Thermo Fisher). *MEG3* mRNA expression level was measured with the previously described qRT-PCR primers and TBP (Applied Biosystems, 4326322E) was used as a reference gene.

Western Blot Assay

Cells were washed with cold Phosphate Buffered Saline (PBS) and lysed in radio immunoprecipitation assay (RIPA) buffer with phosphatase and protease inhibitors (Halt Protease Inhibitor Cocktail, Thermo Fisher Scientific, 78430) for 10 min on ice and scraped with cell scraper. Protein concentration was measured using Bradford reagent (BioRad, 5000002). Equal amounts of protein (5–15 μg) were separated on 10% NuPage Bis-Tris gels (Invitrogen, NP0301PK2) with NuPage MES running buffer (Thermo Fisher Scientific, NP0002) and transferred with NuPage Transfer buffer (Thermo Fisher Scientific, NP0006-1) to polyvinylidene fluoride (PVDF) membranes Millipore Imobilion-FL transfer membrane, pore size 0,45 μM (Millipore, IPFL00010). The membranes were blocked with Odyssey Blocking buffer (TBS) (LiCor, 927-500) and incubated with primary antibodies overnight at 4°C. List of antibodies: keratin 14 (KRT14; Abcam, Ab15461), keratin 19 (KRT19; Abcam, Ab7754), P-cadherin (CDH3; Cell signaling, CS2130), tumor protein p63 (TP63, Abcam, Ab124762), keratin 5/6 (KRT5/6; Invitrogen, 180267), Actin (Licor, 926-42212). Actin was used as loading control. Secondary antibodies were mouse or rabbit IRDey (Li-Cor

926-32213, 926-32212, respectively) used at 1:10.000 for 1 h at room temperature (RT) and detected and quantified using the Odyssey Infrared Imaging System (Li-Cor Fluorescent signal was detected by Odyssey image system (Li-Cor) and converted to gray scale.

Cell Migration Assay

Cell migration was examined by using trans-well Boyden chambers with 8 μm pore size (Corning, 353097). Briefly, 3×10^3 cells were resuspended in 250 μl H14 medium and seeded on the trans-well inserts in 24-well plate (Corning, 353047). H14 media with 10% FBS was added to the lower chamber, below filter. Cells were incubated for 48 h in 5% CO_2 at 37°C. Non-migratory cells from the upper part of the filter were removed with cotton swab and washed 3 times with 1× PBS. The filters were then fixed with methanol and stained with DAPI (Sigma, D9542-1MG). Cells were photographed in three random fields EVOS FL Auto 2 imaging system (ThermoFisher). Pictures were analyzed with ImageJ Software.

Low Attachment Assay

Anchorage independent growth was examined using 24-well ultra-low attachment plates (Corning, 3473). Briefly, D492 and D492M cells were single cell filtered and 500 cells/well were seeded into EGM5 media and cultured for 9 days. The growth of colonies was quantified under the microscope, counting all the colonies bigger than 40 μm.

Apoptosis Assay

Resistance to chemically induced apoptosis was examined by inducing the cells with 10 μM camptothecin [CPT, Sigma-Aldrich, C9911)] in 96-well plate format (Corning, 353072). and quantified using IncuCyte Caspase-3/7 Reagents (Essen Bioscience, 4440) on IncuCyte Zoom (Essen Bioscience) according to the manufacturer's instructions.

Lentivirus Packaging and Transfection

The packaging of lentiviral expression constructs into pseudoviral particles, was performed with the psPAX2 (Addgene, 12260) and PMDG.2 (Addgene, 12259) plasmids using Turbofect (Thermo Fisher Scientific, R05319) in HEK-293T cells. The supernatant was harvested after 48 and 72 h and filtered through 0,45 μm pore filter. For infection, cells were plated on T25 flasks, so they were 70–80% confluent following day and were infected with 1 ml of viral particles and 1 ml of fresh media in the presence of 8 μg/ml polybrene. Lentivirus-transduced cells were selected with antibiotics or sorted by FACS (Sony SH800), based on fluorescent dye to obtain stable pool of clones. The altered expression of *MEG3* was determined by qRT-PCR.

The list of lentiviral expression constructs (plasmids) used in the study and their selection marker (with final concentration in case of antibiotics): pLenti_sgRNA(MS2)_zeo (Zeocin Invitrogen 4 μl/ml), pLenti_dCas9-VP64_Blast (Blasticidin, 2 μg/ml), pLenti_dCas9-KRAB_mCherry (mCherry fluorescence), SAM MS2-P65-HSF1 Plasmids (Hygromycin 1 μl/ml).

CRISPRi/CRISPRa

To perform CRISPRi and CRISPRa, two vectors were used. First, vector with dCAS9 with effector domain KRAB (pLenti_dCas9-KRAB_mCherry, Genscript) and VP64 (pLenti_dCas9-VP64_Blast, Genscript) effector domain for CRISPRi and CRISPRa, respectively, was incorporated, using lentiviral transfection. Subsequently, vector with designed gRNA targeting specific site of our gene of interest *MEG3* was incorporated, in second round of lentiviral transfection. In case of gain of function studies with CRISPRa, one additional helper plasmid SAM (SAM MS2-P65-HSF1 Plasmids, Genscript) was used to further increase activation.

The sequence of gRNA for overexpression of *MEG3*: Guide 1: GCTCTCCGCCGTCTGCGCTA, the sequence of gRNA for downregulation of *MEG3*: Guide 2: GCGGGTGAGGGATCCTCTCGT, the sequence of gRNA for negative control: GCTTAGTTACGCGTGGACGA were cloned into pLenti_sgRNA(MS2)_zeo (Genscript).

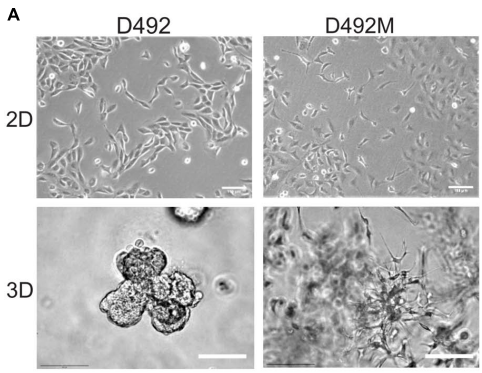
Statistical Analysis

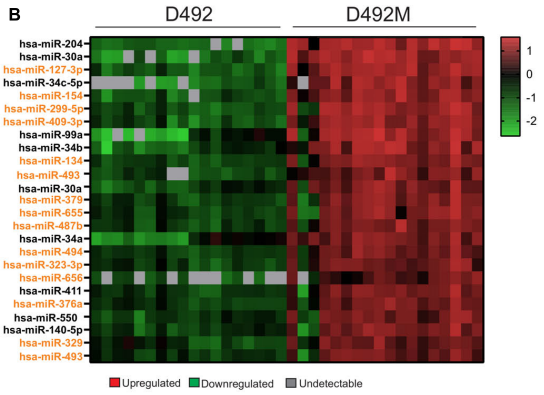
Statistical differences of qRT-PCRs (**Figures 1E,F**, **Figures 5A,B**, and **Figures 7A–C**) and functional assay (**Figures 8B–E**) between samples were assessed with unpaired Student *t*-test. Statistical differences in **Figure 8A** was calculated using multiple unpaired Student *t*-test per row. Statistical differences of quantifications of western blots (**Figures 7B,C**) among samples were assessed using one-way ordinary ANOVA, followed by Tukey's multiple comparison test. Statistical differences in **Figure 4A** (left) was calculated using Kruskal Wallis Test (one-way ANOVA on ranks). Statistical analysis of qRT-PCRs in **Figure 2** were assessed with One-way ANOVA with Dunnett's multiple comparisons test. All statistical analyses were performed in GraphPad Prism. *P*-values below 0,05 were considered significant ($^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$).

RESULTS

MEG3 Is Highly Expressed in Cell Lines With a Mesenchymal Phenotype and in the Stromal Compartment of Breast Tissue

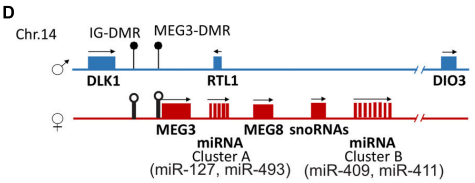
D492 and D492M are isogenic cell lines with stem cell and mesenchymal properties, respectively. D492 cells acquire cuboidal shape in 2D culture, and form branching structures in 3D culture, akin to terminal duct lobular units (TDLU) in the breast. In contrast, D492M is elongated and spindle-shaped in 2D culture and in 3D culture it forms irregular mesenchymal-like colonies (**Figure 1A**). We have previously shown that MIR203a and the MIR200 family are downregulated in D492M and their expression is essential for the epithelial phenotype (Hilmarsdottir et al., 2015; Briem et al., 2019a). Of miRNAs upregulated in D492M, the miRNAs at the DLK1-DIO3 locus are prominent. A microarray analysis of miRNA expression demonstrated that 15 of the 25 most highly expressed miRNAs in D492M compared to D492 belong to the DLK1-DIO3





C

	lncRNA ID	D492M vs D492 - Fold change	D492M vs D492 FDR p-value correction	D492 - Normalized expression values	D492M - Normalized expression values
1	*HIF1A-AS2*	304.7	0.000	0.08	23.49
2	*MIR381HG*	217.3	0.077	0.04	8.38
3	*MIR155HG*	207.7	0.092	0.04	8.01
4	*RP11-175K6.1*	68.1	0.347	0.08	5.25
5	*LINC01551*	45.8	0.000	0.69	31.76
6	*RP11-102K13.5*	38.3	0.043	0.27	10.34
7	*MEG3*	17.9	0.000	10.64	190.67
8	*AC016831.7*	16.3	0.253	0.42	6.89
9	*MEG8*	12.0	0.000	5.05	60.60
10	*CASC19*	11.1	0.095	0.93	10.29



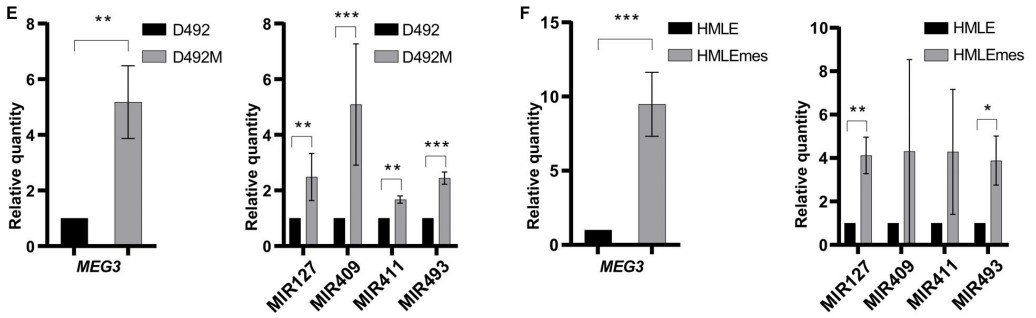


FIGURE 1 | The DLK1-DIO3 locus is upregulated in breast epithelial stem cells undergoing EMT. **(A)** D492 and D492M generate branching and mesenchymal structures in 3D culture, respectively. In 2D culture, D492 is cuboidal in shape and D492M is more spindle shaped. Scale bar = 100 μm. **(B)** Majority of the top upregulated miRNAs in D492M are from the DLK1-DIO3 locus. Microarray heat map showing top 25 upregulated miRNAs in D492M compared to D492. 15 of them are from the DLK1-DIO3 locus (highlighted in orange). **(C)** LncRNAs from the DLK1-DIO3 locus are among the most upregulated in D492M. RNAseq data showing top ten differentially expressed lncRNA, with lncRNAs from DLK1-DIO3 locus (*MEG3* and *MEG8*) highlighted in orange. **(D)** Schematic figure of the DLK1-DIO3 locus. The DLK1-DIO3 locus is located on chromosome 14 and is imprinted. It contains three paternally expressed protein coding genes (*DLK1*, *RTL1*, and *DIO3*) and many maternally expressed non-coding genes, among them lncRNAs (*MEG3* and *MEG8*) and over 50 miRNAs, among them MIR127 and MIR493 located in cluster A and MIR409 and MIR411 located in cluster B and numerous C/D-box-containing small nucleolar RNAs (snoRNAs). DMR – differentially methylated region, filled circles represent methylated DMRs, and unfilled represent unmethylated DMRs. **(E)** Upregulation of selected ncRNA from the DLK1-DIO3 locus verified with qRT-PCR. Graphs showing higher expression of *MEG3* in D492M compared to D492 (left) and higher expression of four representative miRNAs (MIR127, MIR493, MIR409, and MIR411) at the DLK1-DIO3 locus in D492M compared to D492 (right). Results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $n = 3$. **(F)** ncRNAs from the DLK1-DIO3 locus are upregulated in HMLEmes. qRT-PCR showing higher expression of *MEG3* and representative miRNAs (MIR127, MIR493, MIR409, and MIR411) from DLK1-DIO3 locus in HMLEmes compared to HMLE. Results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $n = 3$.

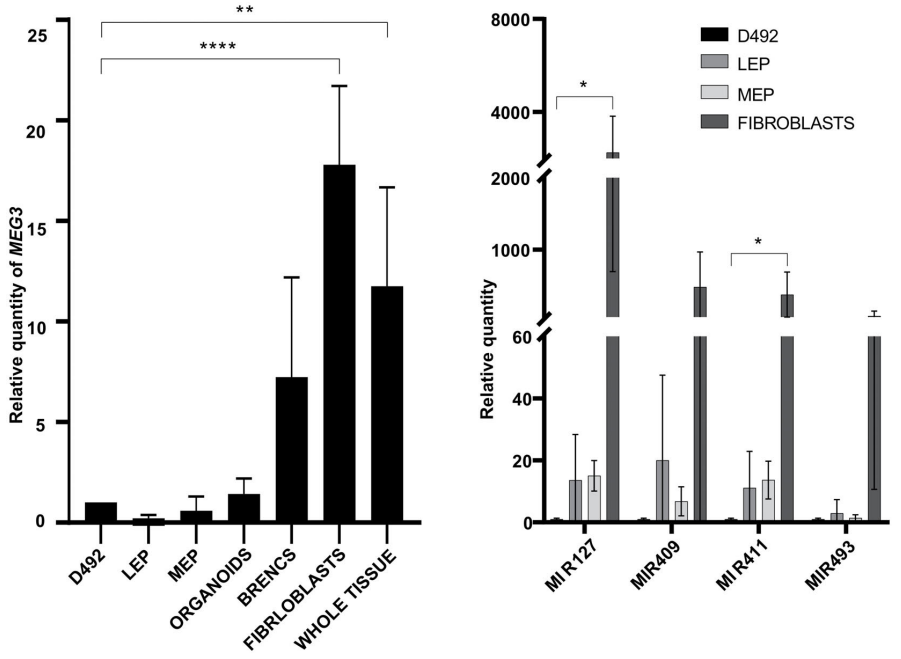


FIGURE 2 | The ncRNA from the DLK1-DIO3 locus are highly expressed in stromal cells and whole tissue compared to epithelial cells. qRT-PCR showing *MEG3* expression is higher in breast stromal cells (fibroblasts) and whole tissue than in D492 (left). (LEP – luminal epithelial cells, MEP – myoepithelial cells, BRENCs – breast endothelial cells). Results shown as mean ± SD. One-way ANOVA with Dunnett's multiple comparisons test was used to test significance: $^{**}p \leq 0.01$; $^{****}p \leq 0.0001$; $n = 3$. Expression of representative miRNAs MIR127 and MIR411 at the DLK1-DIO3 locus is higher in breast fibroblast than in D492 (right). LEP – luminal epithelial cells, MEP - myoepithelial cells. Results shown as mean ± SD. One-way ANOVA with Dunnett's multiple comparisons test was used to test significance: $^{*}p \leq 0.5$; $n = 3$.

locus (**Figure 1B**). Furthermore, small RNA sequencing revealed that 33 of the miRNAs belonging to the DLK1-DIO3 miRNA cluster have more than 1,5-fold increased expression in D492M compared to D492 (**Supplementary Figure 1**). Moreover, total RNA sequencing of D492 and D492M, revealed that *MEG3* and *MEG8* are amongst the most upregulated lncRNAs in D492M (**Figure 1C**). The non-coding part of the DLK1-DIO3 locus consists of maternally expressed lncRNAs *MEG3* and *MEG8* and miRNAs grouped into two clusters (**Figure 1D**). To confirm the sequencing results, we selected four representative miRNAs from the DLK1-DIO3 locus, two from each cluster (MIR127 and MIR493 from cluster A, MIR409 and MIR411 from cluster B). These miRNAs as well as the lncRNA *MEG3* had higher expression, as revealed by qRT-PCR, in D492M compared to D492 (**Figure 1E**). In another isogenic EMT cell model, HMLE (epithelial) and HMLEmes (mesenchymal variant) both *MEG3* and the representative miRNAs were more highly expressed in HMLEmes compared to HMLE (**Figure 1F**). Thus, our data suggests that increased *MEG3* expression is not a stochastic event but consistently associates with EMT induction in breast epithelial cell lines.

Next, we analyzed the expression of *MEG3* and miRNAs from the DLK1-DIO3 locus in primary cells from three healthy donors. We found that the expression of *MEG3* is higher in purified stromal cells (fibroblasts) than in epithelial cells (D492, luminal epithelial cells, myoepithelial cells and organoids; **Figure 2**, left). Interestingly, expression of *MEG3* in whole breast tissue lysates is closer to fibroblast expression levels than epithelial cells (**Figure 2**, left). This finding is most likely explained by the richness of stroma in normal breast tissue, whereas organoids contain only the epithelial cells. A similar pattern is seen with the four representative miRNAs, where MIR127 and MIR411 have higher expression in fibroblasts compared to their expression in D492 (**Figure 2**, right).

We next acquired a list of genes correlated the expression of MEG3 using the GOBO (Gene expression-based Outcome for Breast Cancer Online) dataset and submitted the list to DAVID (the database for annotation, visualization and integrated discovery, version 6.7) (Huang et al., 2009a,b) to identify pathways associated with *MEG3*. Herein, the expression of *MEG3* correlates with expression of extracellular matrix genes, which are in line with the observations of a high expression of *MEG3* in cells found in the stromal compartment (**Supplementary Figure 2A**). Using analysis of publicly available NGS data using MiPanda (Niknafs et al., 2018) we found positive correlation of *MEG3* with

common EMT markers in normal breast and breast cancer (**Supplementary Figure 2B**). Many of these have a correlation coefficient > 0.3 (Spearman correlation) which is considered a fair positive correlation (Chan, 2003). Interestingly, even more genes are positively correlated to *MEG3* expression in breast cancer as compared to normal breast tissue (**Supplementary Figure 2B**).

Collectively, the lncRNA *MEG3* and miRNAs from DLK1-DIO3 locus are highly expressed in the mesenchymal compartment compared to epithelial breast tissue and their expression positively correlate with numerous mesenchymal genes and EMT markers.

MEG3 Is Imprinted in Both D492 and D492M

The DLK1-DIO3 locus is imprinted and regulated by DNA methylation (Cui et al., 2018). Using pyrosequencing (Harrington et al., 2013) covering a heterozygous SNP (C/T) in *MEG3* (rs4906024) we confirmed monoallelic expression of *MEG3* in both D492 and D492M, with expression in both cell lines being from the T allele (**Figure 3**). As both cell lines are diploid at the *MEG3* locus on a DNA level a C/T ratio of 50% is expected which is consistent with the 48% C-allele prominence observed. On the mRNA deviation from expected monoallelic expression was not detected as results showed zero C allele expression in D492 and 2% in D492M. Hence, increased expression of *MEG3* in D492M is not caused by loss of imprinting. The expression remains monoallelic confirming that the increased expression originates from the non-imprinted allele.

Increased Expression of *MEG3* Is Negatively Correlated With Survival of Breast Cancer Patients

EMT has been suggested to promote metastatic behavior of epithelia-originating cancer (Felipe Lima et al., 2016) and, in addition, our data shows association of *MEG3* expression with the mesenchymal phenotype. We therefore investigated *MEG3* expression levels in different subtypes of breast cancer. We have evaluated the expression of *MEG3* in clinically well-defined breast tumors. Herein, normal like (NL) breast tumors had significantly higher expression of *MEG3* with a *p*-value of 0.0003 (**Figure 4A**, left). Survival analysis of all tumor samples showed reduced, but not significant overall survival in patients with high *MEG3* expression. However, as the normal-like tumors have in recent years been subjected to scrutiny as a possible misclassification due to low tumor cellularity and thus, high proportion of normal tissue. In light of our results showing high expression of *MEG3* in breast stromal tissue, and uncertainty that measured *MEG3* expression in the normal-like subgroup is representative of the primary tumor, we omitted NL breast tumors from the survival analysis (Elloumi et al., 2011; Prat and Perou, 2011; Yersal and Barutca, 2014). The results show significant worse overall survival of patients with high *MEG3* expression (**Figure 4A**, right). Corroborating our findings, using the GOBO database (Ringnér et al., 2011)[1], we found that high *MEG3* expression reduces distant metastasis free

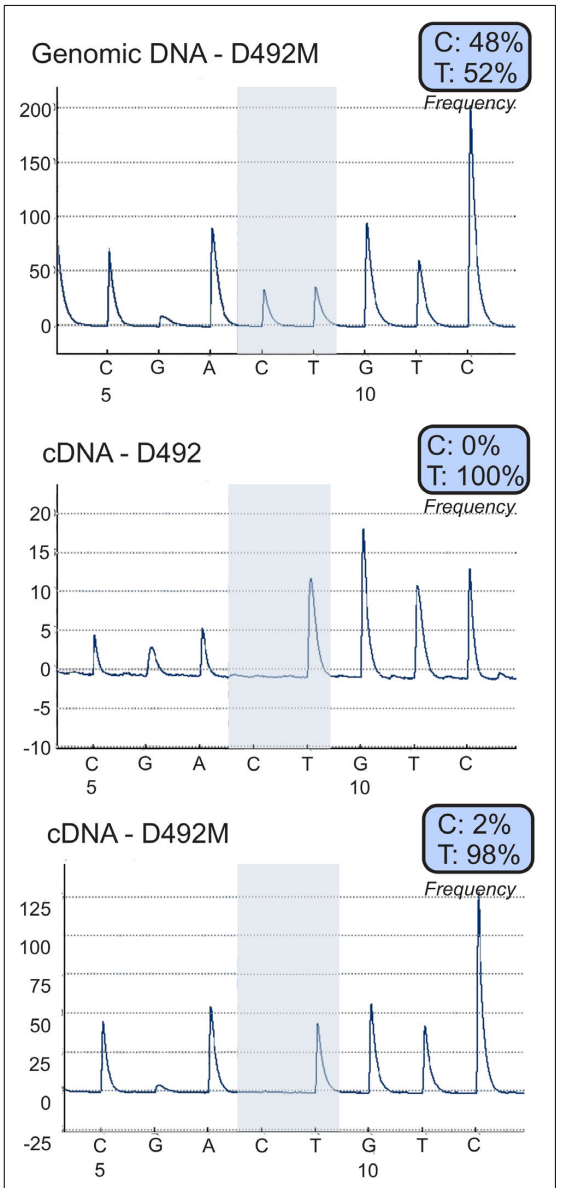


FIGURE 3 | Increased expression of *MEG3* in D492M is not caused by loss of imprinting. Allele specific expression analysis using pyro-sequencing shows that the increased expression of *MEG3* in D492M is not caused by loss of imprinting, but increased expression of the already expressed allele, the expression remains mono-allelic. Analyzed sequence: GAGCAC/TGTCCCA.

[1] http://co.bmc.lu.se/gobo/

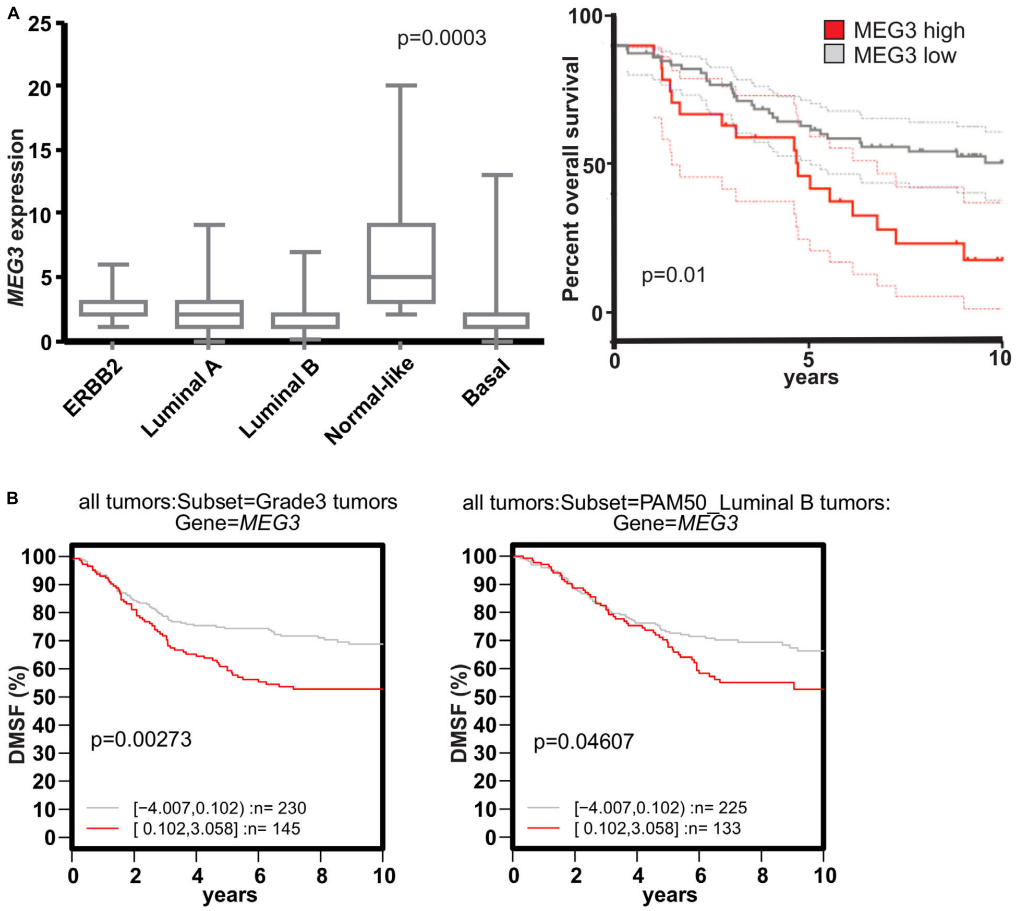


FIGURE 4 | *MEG3* expression negatively correlates with breast cancer prognosis. **(A)** High *MEG3* expression decreases overall survival in breast tumors. qRT-PCR shows that *MEG3* expression (Numbers of tumors per group: ErBB2 = 12, Luminal A = 33, Luminal B = 24, Normal = 10, Basal = 22) is significantly higher in normal-like (NL) breast cancer (left; $p = 0.0003$). High MEG3 expression is correlated with low overall patient survival (NL tumors omitted). Kruskal Wallis Test (or one-way ANOVA on ranks) was used to test the significance (right; $p = 0.01$). **(B)** High *MEG3* expression decrease distant metastasis free survival in grade three and Luminal B type tumors. Kaplan-Meier plot showing data from the online GOBO database: high expression of *MEG3* decrease DMSF (distant metastasis free survival), of poorly differentiated (grade 3) tumors (left; $p = 0.00373$) and luminal B type tumor (right; $p = 0.04607$).

survival (DMSF) of patients with poorly differentiated (grade 3) tumors (**Figure 4A**, left) and patients with luminal B tumors (**Figure 4B**, right).

Increased Expression at the DLK1-DIO3 Locus Contributes to the Basal and Mesenchymal Phenotype

To explore the functional role of *MEG3* in D492 and D492M, we established sublines with altered expression of *MEG3*. Using the CRISPRa approach (Cheng et al., 2013), we generated a D492 cell line with stable overexpression of *MEG3* ($D492^{MEG3}$). A control cell line was generated using a scrambled sgRNA ($D492^{CTRL}$). Furthermore, we used the CRISPRi approach (Gilbert et al., 2013; Qi et al., 2013), to generate knockdown of *MEG3* in D492M ($D492M^{KD-MEG3}$) and a control cell line was generated using scrambled sgRNA ($D492M^{KD-CTRL}$). The increase of *MEG3* expression was about seven-fold in $D492^{MEG3}$ compared to $D492^{CTRL}$ as determined by qRT-PCR (**Figure 5A**, left). Downregulation of *MEG3* in $D492M^{KD-MEG3}$ was more prominent, with about 20-fold reduced expression compared to $D492M^{KD-CTRL}$ (**Figure 5A**, right). Having established stable overexpression and downregulation of *MEG3* in D492 and D492M, we re-evaluated the epithelial/mesenchymal phenotypes of D492 and D492M, respectively. Based on phase contrast images, no obvious

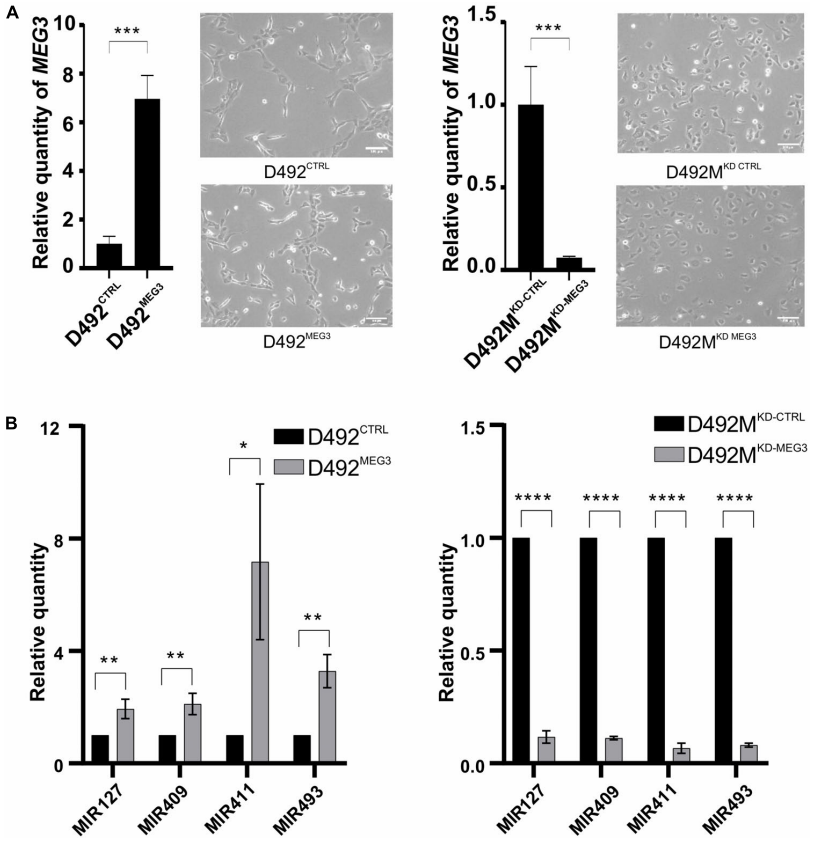


FIGURE 5 | Concomitant expression of non-coding RNAs from DLK-DIO3 locus with *MEG3*. **(A)** Overexpression and knock-down of *MEG3* in D492 and D492M, respectively. qRT-PCR confirming upregulation of *MEG3* in D492 ($D492^{MEG3}$) compared to D492 with scrambled control ($D492^{CTRL}$; left). Phase contrast pictures of $D492^{CTRL}$ and $D492^{MEG3}$ (below). qRT-PCR confirming knock-down of *MEG3* in D492M ($D492M^{KD-MEG3}$) compared to D492M with scrambled control ($D492M^{KD-CTRL}$; right). Results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{***}p \leq 0.001$; $^{****}p \leq 0.0001$; $n = 3$. Phase contrast pictures of $D492M^{KD-CTRL}$ and $D492M^{KD-MEG3}$ (below). Scale bar = 100 μm. **(B)** miRNAs form the DLK1-DIO3 locus are upregulated with overexpression of *MEG3* and downregulated with knock-down of *MEG3*. qRT-PCR shows increased expression of four representative miRNAs from the DLK1-DIO3 locus in $D492^{MEG3}$ compared to $D492^{CTRL}$ (left) and their decreased expression in $D492M^{KD-MEG3}$ compared to $D492M^{KD-CTRL}$ (right). Results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{*}p \leq 0.05$ $^{**}p \leq 0.01$; $^{****}p \leq 0.0001$; $n = 3$.

difference in phenotype could be seen between $D492^{MEG3}$ and $D492^{CTRL}$ or $D492M^{KD-MEG3}$ and $D492M^{KD-CTRL}$ (**Figure 5A**, below). Interestingly, expression of the representative miRNAs located on the DLK1-DIO3 locus is increased in $D492^{MEG3}$ compared to $D492^{CTRL}$, to similar levels as seen in D492M (**Figure 5B**, left). Conversely, the expression of representative miRNAs is downregulated in $D492M^{KD-MEG3}$ compared to $D492M^{KD-CTRL}$ (**Figure 5B**, right). Thus, it appears, that the expression of miRNAs from the DLK1-DIO3 locus is concomitant with *MEG3* expression. To test, if that holds true, we used the cBioPortal and explored correlation of *MEG3* with miRNAs using data on invasive breast cancer from the Cancer Genome Atlas (TCGA) (Cancer Genome Atlas Network, 2012) we found that of 40 miRNAs that had positive correlation over 0,3 (Person score) with *MEG3*, 30 were located at the DLK1-DIO3 locus (with other miRNAs from the locus not being in the dataset; **Table 1**). This suggests that *MEG3* may be used as a marker for the expression of ncRNAs from the DLK1-DIO3 locus.

Next, we conducted RNA sequencing of our cell lines with stably altered expression levels of *MEG3* focusing on the analysis of $D492M^{KD-CTRL}$ vs $D492M^{KD-MEG3}$. There were 1235 significantly differentially expressed genes, with symmetric distribution over genes downregulated and upregulated in $D492M^{KD-MEG3}$, shown in the volcano plot ($p < 0.05$; **Supplementary Figure 3A**), with the list of top 30 up and down-regulated genes in $D492M^{KD-MEG3}$ (**Supplementary Figure 3B**). To identify unifying biological them from RNA-sequencing data, we performed Gene Set enrichment analysis (GSEA).

TABLE 1 | MiRNAs from the DLK1-DIO3 locus positively correlate with *MEG3* expression.

Correlated gene	Location	Pearson score	*P*-value	Spearman score	*P*-value
MIR-154/154*	Chr14	0.49	7.46E−20	0.44	1.27E−15
MIR-134/134	Chr14	0.48	2.09E−18	0.44	1.22E−15
MIR-199B/3P	Chr9	0.47	4.66E−18	0.47	8.03E−18
MIR-199A-1/3P; MIR-199A-2/3P[#]	Chr19; Chr1	0.47	4.66E−18	0.47	8.15E−18
MIR-127/3P	Chr14	0.46	4.18E−18	0.45	3.89E−16
MIR-136/136	Chr14	0.46	1.08E−16	0.41	2.15E−13
MIR-431/431*	Chr14	0.46	7.09E−17	0.39	5.23E−12
MIR-539/539	Chr14	0.44	1.66E−15	0.43	1.21E−14
MIR-199A-1/5P; MIR-199A-2/5P[#]	Chr19; Chr1	0.43	1.14E−14	0.39	2.18E−12
MIR-382/382	Chr14	0.42	1.53E−14	0.41	2.71E−13
MIR-199B/5P	Chr9	0.42	2.47E−14	0.40	1.06E−12
MIR-214/214*	Chr1	0.42	6.45E−14	0.39	4.35E−12
MIR-409/3P	Chr14	0.42	3.41E−14	0.37	3.34E−11
MIR-369/3P	Chr14	0.41	1.06E−13	0.37	5.93E−11
MIR-127/5P	Chr14	0.41	1.46E−13	0.36	9.70E−16
MIR-495/495	Chr14	0.4	1.07E−12	0.39	3.84E−12
MIR-758/758	Chr14	0.4	3.97E−13	0.36	1.09E−10
MIR-381/381	Chr14	0.39	3.10E−12	0.39	2.80E−12
MIR-485/3P	Chr14	0.39	2.12E−12	0.39	5.07E−12
MIR-125B-1/125B; MIR-125B-2/125B[#]	Chr11; Chr21	0.39	3.92E−12	0.37	6.52E−11
MIR-337/3P	Chr14	0.39	1.80E−12	0.37	7.17E−11
MIR-493/493*	Chr14	0.38	1.15E−11	0.37	2.83E−11
MIR-369/5P	Chr14	0.38	1.68E−11	0.33	2.98E−11
MIR-379/379	Chr14	0.37	5.82E−11	0.33	4.49E−09
MIR-370/370	Chr14	0.37	2.51E−11	0.32	1.29E−08
MIR-214/214	Chr1	0.36	1.88E−10	0.35	5.19E−10
MIR-708/708	Chr11	0.35	4.91E−10	0.35	3.60E−10
MIR-432/432	Chr14	0.35	5.39E−10	0.32	1.80E−08
MIR-409/5P	Chr14	0.35	3.48E−10	0.31	6.03E−08
MIR-323/3P	Chr14	0.35	6.01E−10	0.30	1.65E−07
MIR-376C/376C	Chr14	0.34	1.40E−09	0.35	7.42E−10
MIR-889/889	Chr14	0.34	1.57E−09	0.30	1.50E−07
MIR-493/493	Chr14	0.34	1.07E−09	0.29	5.38E−07
MIR-487B/487B	Chr14	0.33	7.98E−09	0.30	1.50E−09
MIR-655/655	Chr14	0.33	6.70E−09	0.30	1.90E−07
MIR-410/410	Chr14	0.33	4.36E−09	0.29	5.07E−07
MIR-184/184	Chr15	0.31	3.50E−08	0.34	2.68E−09
MIR-411/411	Chr14	0.31	2.83E−08	0.29	3.79E−07
MIR-654/3P	Chr14	0.31	3.64E−08	0.26	4.37E−06
MIR-22/22	Chr17	0.3	1.66E−07	0.29	2.72E−07

Out of 40 miRNAs that positively correlate with MEG3 in breast cancer (with correlation over 0.3), 30 are from the DLK1-DIO3 locus, highlighted in orange (TCGA, Nature 2012 data set). [#]Due to sequence similarities, these two miRNAs are indistinguishable in the sequencing data used.

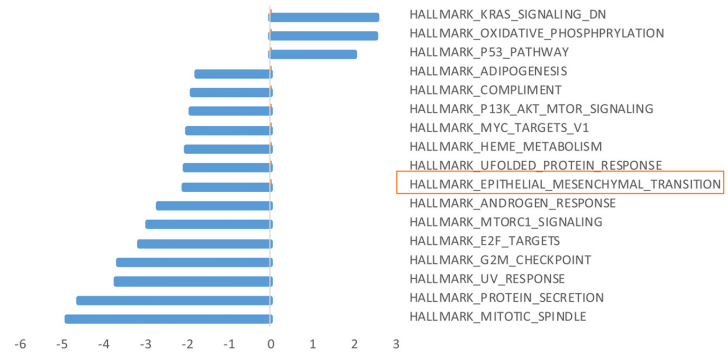
These gene sets consist of the defined gene lists, based on biological knowledge about biochemical pathways and co-expression data. Using the Hallmark dataset, one of the significantly, downregulated set of genes in $D492M^{KD-MEG3}$ was the epithelial-mesenchymal transition gene set, with normalized enrichment score (NES) of −2.03 and False discovery rate (FDR) q = 0.023 (**Figure 6A**). These genes define epithelial-mesenchymal transition, as in wound healing, fibrosis and metastasis. The genes belonging to this gene set are overrepresented toward the top of the ranked list, based on fold change of $D492M^{KD-CTRL}$ vs $D492M^{KD-MEG3}$ (**Figure 6B**, right). A manually curated list of mesenchymal genes from the Hallmark EMT dataset that are downregulated in $D492M^{KD-MEG3}$ is shown in **Figure 6B**, left. Further analysis of the RNA sequencing data of $D492M^{KD-MEG3}$ vs $D492M^{KD-CTRL}$, using common literature-based markers of breast tissue has showed that luminal epithelial markers *GATA3* and *MUC1* are upregulated, while myoepithelial *KRT14*, mesenchymal *VIM*, *ZEB2*, *SNAI2*, *LAMA1*, *CDH2*, and stem cell *MME*, *CTNNB1* are downregulated with knock down of *MEG3* (**Figure 6C**).

Expression of mesenchymal and basal markers was additionally confirmed on RNA level by qRT-PCR and on protein level western blot. Most of the core EMT-related transcription factors (EMT-TF) were affected by *MEG3*. $D492^{MEG3}$ has increased expression of *SNAI2* compared to $D492^{CTRL}$ (**Figure 7A**, left). On the other hand, $D492M^{KD-MEG3}$ has decreased expression of *SNAI2*, *ZEB1*, *ZEB2* and *TWIST1* compared to $D492M^{KD-CTRL}$ (**Figure 7A**, right). Luminal cytokeratin 19 (*KRT19*) and basal/myoepithelial cytokeratin 14 (*KRT14*) are also affected by manipulation of *MEG3* expression levels. Thus, $D492^{MEG3}$ shows increased *KRT14* and decreased *KRT19* expression compared to $D492^{CTRL}$ on both mRNA (**Figure 7B**, left) and protein level (**Figure 7C**, left). $D492M^{KD-MEG3}$ shows decreased *KRT14* expression compared to $D492M^{KD-CTRL}$ (**Figure 7B**, right). Furthermore, $D492^{MEG3}$ shows increased expression of other myoepithelial markers such as *CDH3* (P-cad), *TP63* or *KRT5* compared to $D492^{CTRL}$ as determined both at mRNA (**Figure 7D**, left) and protein level (**Figure 7E**). Also, $D492M^{KD-MEG3}$ shows decreased expression of myoepithelial markers *KRT5* on mRNA level (**Figure 7D**, right) and of TP63 on protein level (**Figure 7E**, middle) compared to $D492M^{KD-CTRL}$. This suggests that *MEG3* expression induces a shift toward a basal/myoepithelial phenotype. However, our cell lines with stably altered expression of *MEG3* do not show a significant switch in E-cadherin (CDH1) to N-cadherin (CDH2) expression (**Supplementary Figure 4**), which may explain why there are no clear changes in morphology.

MEG3 Induces Mesenchymal Properties and Stemness

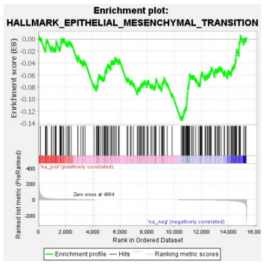
As *MEG3* has previously been ascribed to have a role in pluripotency and stemness (Stadtfeld et al., 2010b; Kaneko et al., 2014), we asked how *MEG3* manipulation affects mesenchymal and stem cell properties of D492 and D492M. The expression of both aldehyde dehydrogenase (*ALDH1A3*) and integrin alpha 6 (*ITGA6*; **Supplementary Figure 6**), markers of stemness, is increased in $D492^{MEG3}$ compared to $D492^{CTRL}$. Next, we employed several functional assays to assess the effect of *MEG3* levels in D492 and D492M on mesenchymal and stem cell properties. $D492^{MEG3}$ is more resistant to chemically induced apoptosis than $D492^{CTRL}$ (**Figure 8A**). Migration can be assessed *in vitro* using the wound healing assay or by trans-well migration

A



B

Gene	Name	Log2 FC	p-value
POSTN	Periostin	-2.45	$3.69x10^{-10}$
DST	Dystonin	-1.57	$2.21x10^{-9}$
IGFBP3	Insulin like grow th factor binding protein 3	-1.28	$4.99x10^{-19}$
ITGAV	Integrin alpha-V	-1.25	$1.62x10^{-42}$
ITGB1	Integrin subunit beta 1	-1.23	$1.48x10^{-56}$
VCAN	Versican	-1.19	$1.44x10^{-5}$
FERMT2	Fermitin family homolog 2	-1.07	$1.16x10^{-5}$
LAMA1	Laminin subunit alpha 1	-1.01	$2.82x10^{-26}$
COL12A1	Collagen type XII alpha 1 chain	-0.98	$9.17x10^{-19}$
CDH2	N-cadherin	-0.97	$1.21x10^{-24}$
SERPINE2	Serpine 2	-0.97	$2.42x10^{-18}$
FN1	Fibronectin 1	-0.95	$3.13x10^{-19}$
LAMC1	Laminin subunit gamma 1	-0.92	$2.61x10^{-25}$
LAMA3	Laminin subunit alpha 3	-0.92	$6.25x10^{-28}$
SERPINE1	Serpine 1	-0.9	$3.10x10^{-15}$
FBN2	Fibronectin 2	-0.89	$6.38x10^{-18}$
LAMC2	Laminin subunit gamma 2	-0.88	$1.66x10^{-22}$
VEGFC	Vascular endothelial grow th factor C	-0.74	$1.86x10^{-17}$
FGF2	Fibroblast grow th factor 2	-0.65	$3.47x10^{-11}$
SNAI2	Snail family transcriptional repressor 2	-0.63	$3.32x10^{-8}$
VIM	Vimentin	-0.49	$3.04x10^{-7}$
COL4A1	Collagen type IV alpha 1 chain	-0.48	$1.19x10^{-4}$



C

Gene	Name	Log2 FC	Marker	p-value
GATA3	GATA3	0.42	Luminal epithelial	$8.19x10^{-4}$
MUC1	Mucin 1	0.83	Luminal epithelial	$1.97x10^{-6}$
CDH2	N-cadherin	-0.97	Mesenchymal	$1.21x10^{-24}$
LAMA1	Laminin subunit alpha 1	-1.01	Mesenchymal	$2.82x10^{-26}$
SNAI2	Snail family transcriptional repressor 2	-0.63	Mesenchymal	$3.32x10^{-8}$
ZEB2	Zinc finger E-box binding homeobox 2	-0.43	Mesenchymal	$3.76x10^{-6}$
VIM	Vimentin	-0.49	Mesenchymal	$3.04x10^{-7}$
CTNNB1	Catenin beta-1	-0.3	Stem cell	$4.1x10^{-4}$
MME	Membrane metallo-endopeptidase	-2.97	Stem cell	$1.19x10^{-72}$

FIGURE 6 | Knock down of *MEG3* in D492M decrease mesenchymal markers. **(A)** Epithelial-mesenchymal transition gene set is enriched pathway in $D492M^{KD-MEG3}$ (highlighted in orange). Bar plot of Gene Set Enrichment Analysis (GSEA) with Hallmark dataset showing all significantly (False discovery rate-FDR $q \leq 0.05$) enriched pathways in $D492M^{KD-MEG3}$. Gene set Epithelial-mesenchymal transition has normalized enrichment score (NES) of –2.06 and FDR $q = 0.014$. **(B)** Knock-down of *MEG3* correlates with downregulation of mesenchymal genes relevant for breast cells. Enrichment plot showing the Enrichment Score (ES) of the genes in the Hallmark gene set Epithelial-mesenchymal transition in $D492M^{KD-MEG3}$. The genes are overrepresented toward the top the ranked list of $D492M^{KD-MEG3}$ (right). Table with relevant genes for breast cells, significantly ($p \leq 0.05$) deregulated, from the Hallmark gene list epithelial-mesenchymal transition showing Log2 fold change (FC; left). **(C)** Luminal epithelial markers (*GATA3* and *MUC1*) are upregulated, while myoepithelial (*KRT14*), mesenchymal (*CDH2*, *LAMA1*, *SNAI2*, *VIM*, and *ZEB2*) and stem cell (*CTNNB1* and *MME*) are downregulated with knock down of *MEG3*. Genes from literature-based list of markers significantly ($p \leq 0.05$) differentially expressed in $D492M^{KD-MEG3}$ vs $D492M^{KD-CTRL}$.

where the cells migrate toward a chemo-attractant. In the wound healing assay, $D492^{MEG3}$ has slightly increased migration rate compared to $D492^{CTRL}$, while $D492M^{KD-MEG3}$ has decreased migration rate compared to $D492M^{KD-CTRL}$ (**Supplementary Figure 5A**). In the trans-well migration assay, $D492^{MEG3}$ has about two-fold increased migration rate compared to $D492^{CTRL}$ and $D492M^{KD-MEG3}$ has reduced migration rate compared to $D492M^{KD-CTRL}$ (**Figure 8B**). *MEG3* manipulation, however, did not affect invasion in a transwell invasion assay (**Supplementary Figure 5B**). We performed

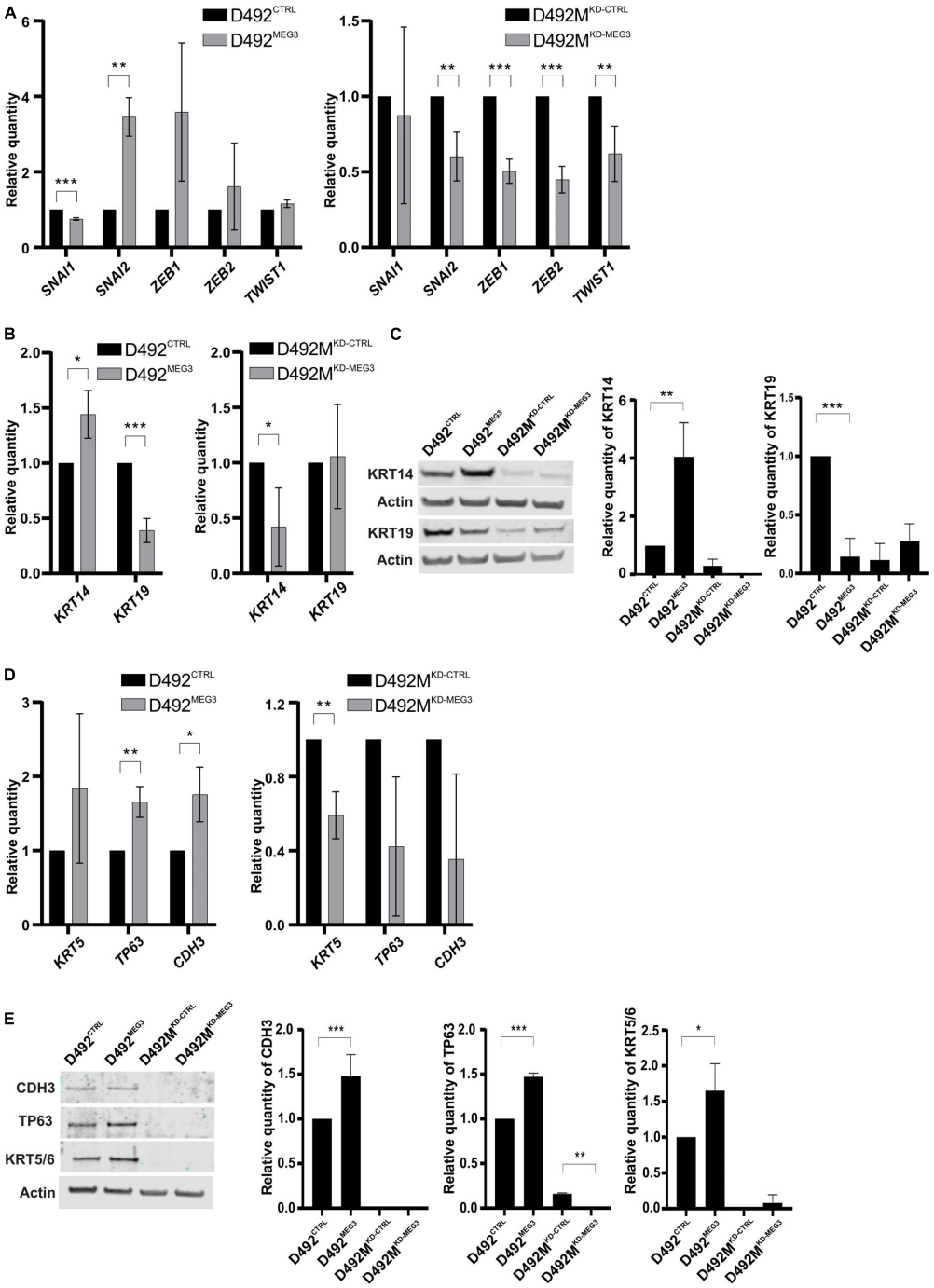


FIGURE 7 | Continued

FIGURE 7 | *MEG3* induce partial EMT. **(A)** *MEG3* increases expression of EMT transcription factors in D492 and the knock down of *MEG3* decrease expression of TF-EMT in D492M. qRT-PCR showing $D492^{MEG3}$ increased expression of transcription factors (TF) *SNAI2* compared to $D492^{CTRL}$ (left) and decreased expression of EMT related TF *SNAI2*, *ZEB1*, *ZEB2* and *TWIST1* in $D492M^{KD-MEG3}$ compared to $D492M^{KD-CTRL}$ (right). Results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $n = 3$. **(B)** *MEG3* increases expression of myoepithelial marker *KRT14* and decrease expression of luminal epithelial marker *KRT19* on mRNA level. qRT-PCR showing $D492^{MEG3}$ has increased expression of *KRT14* and decrease expression of *KRT19* compared to $D492^{CTRL}$ (left). $D492M^{KD-MEG3}$ has decreased expression of *KRT14* compared to $D492M^{KD-CTRL}$ (right). qRT-PCR results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{*}p \leq 0.05$; $^{***}p \leq 0.001$; $n = 3$. **(C)** qRT-PCR results confirmed on protein level. Representative pictures of western blot (WB) with its quantification (below). $D492^{MEG3}$ has increased protein level of KRT14 and decreased protein level of KRT19 compared to $D492^{CTRL}$. WB results shown as mean ± SD. One-way ordinary ANOVA, followed by Tukey's multiple comparison test was used to test significance: $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $n = 3$. **(D)** *MEG3* increase expression of myoepithelial markers *TP63* and *CDH3* and knock-down of *MEG3* decrease expression of myoepithelial marker *KRT5*, on mRNA level. qRT-PCR showing $D492^{MEG3}$ has increased expression of *TP63* and *CDH3* compared to $D492^{CTRL}$ (left). $D492M^{KD-MEG3}$ has decreased expression of *KRT5* compared to $D492M^{KD-CTRL}$ (right). qRT-PCR results shown as mean ± SD. Unpaired *t*-test was used to test significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $n = 3$. **(E)** qRT-PCR results confirmed on protein level. Representative pictures of western blot (WB) with its quantification (below). $D492^{MEG3}$ has increased protein level of CDH3 (P-cad), TP63 (p63) and KRT5 compared to $D492^{CTRL}$. $D492M^{KD-MEG3}$ has decreased protein level of TP63 compared to $D492M^{KD-CTRL}$. WB results shown as mean ± SD. One-way ordinary ANOVA, followed by Tukey's multiple comparison test was used to test significance: $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$; $n = 3$.

mammosphere assays in rBM (reconstituted basement membrane, Matrigel) (**Figure 8C**) and in low attachment plates (**Figure 8D**), with comparable results. $D492^{MEG3}$ increases the formation of colonies compared to $D492^{CTRL}$ while $D492M^{KD-MEG3}$ decreases the formation of colonies compared to $D492M^{KD-CTRL}$. In addition, we co-cultured $D492^{MEG3}$ with endothelial cells (HUVECs) and observed increased size of colonies and less branching compared to $D492^{CTRL}$ (**Figure 8E**). Finally, manipulation of *MEG3* levels slightly affected proliferation rate of $D492M^{KD-MEG3}$ compared to $D492M^{KD-CTRL}$ (**Supplementary Figure 5C**).

DISCUSSION

In this study, we show that ncRNAs from the DLK1-DIO3 locus are highly expressed in stromal/mesenchymal cells in the breast and positively correlate with the expression of EMT genes in breast tissue. *MEG3* expression was monoallelic in both D492 and D492M and gain and loss of function studies have shown concomitant expression of *MEG3* with miRNAs from the DLK1-DIO3 locus, indicating that *MEG3* could be used as a marker for the expression of the non-coding RNAs from the locus. *MEG3* expression was shown to be negatively correlated with survival of breast cancer patients, particularly with the luminal B subtype. Furthermore, we demonstrate that enhanced *MEG3* expression accompanied by increased expression of the ncRNAs at the DLK1-DIO3 locus, contributes to partial EMT more correctly referred to as epithelial plasticity, seen by increased expression of EMT related TFs, increase of basal/mesenchymal markers and enhanced properties such as migration, resistance to apoptosis and clonogenic capacity.

We used an isogenic breast cell line model to study the expression pattern and functional role of ncRNAs, both miRNAs and lncRNAs, in EMT. Of interest was the largest miRNA locus in the human genome and the lncRNA *MEG3*, both within the DLK1-DIO3 imprinted region on chromosome 14. The non-coding part of the DLK1-DIO3 locus has higher expression in cells with mesenchymal phenotype (D492M) compared to cells with epithelial phenotype (D492). These results were validated in primary breast tissue and in another cellular model of EMT. Furthermore, we have shown that *MEG3* expression correlates with expression of extracellular matrix proteins, which are secreted by cells with a mesenchymal phenotype, and with mesenchymal genes in breast tissue. Data from pyrosequencing demonstrate that the expression of *MEG3* is monoallelic in both D492 and D492M indicating that the increased expression of *MEG3* in D492M is not due to loss of imprinting. We have shown that *MEG3* negatively correlates with survival of luminal B breast cancer patients and patients with grade 3 breast cancer. This is in line with a recent study where high expression of *MEG3* was identified to be a negative prognostic marker for breast cancer (Yao et al., 2019).

Many studies suggest *MEG3* as a tumor suppressor, largely due to the observation that *MEG3* expression is lower in tumor tissue compared to normal tissue (Sheng et al., 2014; Sun et al., 2014, 2016; Yin et al., 2015; Chak et al., 2017; Molina-Pinelo et al., 2018). Our data demonstrates that *MEG3* expression levels are comparable in whole normal breast tissue and in stroma (fibroblasts), however, the expression of *MEG3* in epithelial cells is much lower. There was considerable variation of the *MEG3* expression in breast tissue samples that could be partially due to different proportions of subset of fibroblasts associated with ducts vs TDLUs. There are studies confirming existence of, for instance, two distinct functionally specialized lineages of lobular vs ductal fibroblast (Morsing et al., 2016) or myoepithelial cells (Fridriksdottir et al., 2017), which could be identified by specific marker expression. Importantly, relative proportions of stromal and epithelial compartment are different in normal and cancerous human breast tissue. Breast cancers arise in vast majority from epithelial cells, with TDLUs being the predominant site of breast tumor occurrence (Tabar et al., 2014). Therefore, it would be expected that expression of *MEG3* is higher in normal breast tissue, as it comprises relatively more stromal cells compared to breast cancer tissue. In line with this, expression of *MEG3* from whole breast tissue is distorted as proportions of stroma vs epithelia in normal/cancer tissue are different, resulting in misleading interpretations. Using RNA only from unsorted normal tissue will mainly represents expression of stromal cells. Therefore, it is crucial to use a proper control when comparing expression of genes in normal vs tumor tissue. Single-cell RNA-sequencing or sorted stromal and epithelial cells would give more informative results as it would enable distinctions between epithelial and stromal tissue

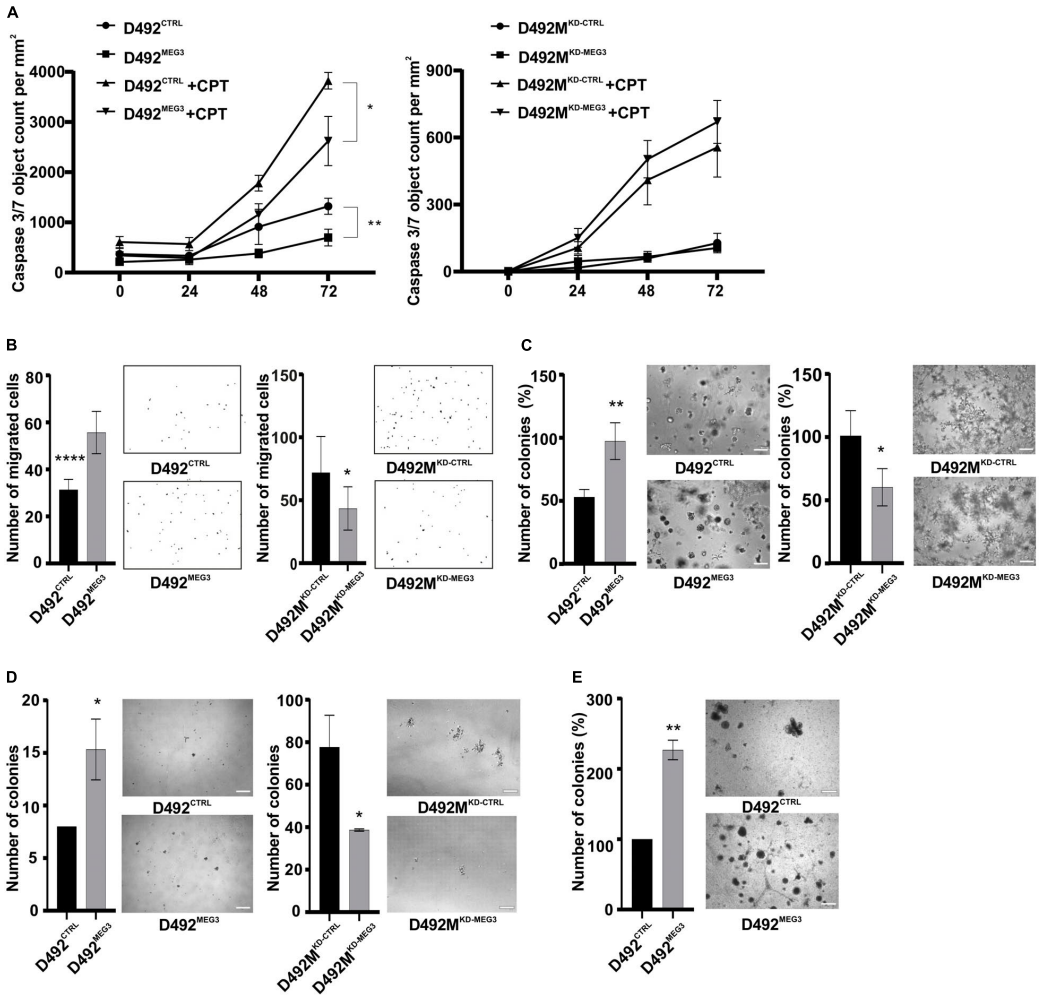


FIGURE 8 | *MEG3* increases stem cell properties. **(A)** *MEG3* increase resistance to chemically induced apoptosis. Apoptosis assay: D492^{MEG3} is more resistant to chemically induced apoptosis compared to D492^{CTRL} (left). Data is analyzed on Incucyte Zoom and displayed as Caspase 3/7 object count/mm^2. Results are shown as mean ± SD. Multiple unpaired Student *t*-test per row was used to test significance at 72-h time-point: *$p \leq 0.05$; **$p \leq 0.01$; $n = 6$. **(B)** *MEG3* increases migration and knock down of *MEG3* decreases migration through trans-well filters. Migration assay: D492^{MEG3} has increased migration rate compared to D492^{CTRL} (left), with representative pictures on side. D492M$^{KD-MEG3}$ migrates less compared to D492M$^{KD-CTRL}$ (right), with representative pictures on side. Quantification of number of migratory cells, analyzed using ImageJ software, results shown as mean ± SD. Unpaired *t*-test was used to test significance: *$p \leq 0.05$; ****$p \leq 0.0001$; $n = 6$. **(C)** *MEG3* increases clonogenic capacity in 3D culture in rBM (reconstitute basement membrane). Mammosphere assay: D492^{MEG3} has higher clonogenic capacity compared to D492^{CTRL} (left). D492M$^{KD-MEG3}$ has lower clonogenic capacity compared to D492M$^{KD-CTRL}$ (right). Data shown as % mean ± SD, *$p \leq 0.05$; **$p \leq 0.01$; $n = 3$. Scale bar = 500 μm. **(D)** *MEG3* increases clonogenic capacity in low attachment assay. D492^{MEG3} increases the formation of colonies compared to D492^{CTRL} (left), with representative pictures on side. D492M$^{KD-MEG3}$ decreases the formation of colonies compared to D492M$^{KD-CTRL}$ (right), with representative pictures on side. Results shown as % of mean ± SD. Unpaired *t*-test was used to test significance: *$p \leq 0.05$; **$p \leq 0.01$; $n = 3$. Scale bars = 200 μm. **(D)** Co-culture of D492^{MEG3} with endothelial cells (HUVECs) increases number and size of colonies and forms less branching compared to D492^{CTRL}. Results shown as % of mean ± SD. Unpaired *t*-test was used to test significance: **$p \leq 0.01$; $n = 3$. Representative pictures scale bars = 200 μm.

compartments. In this paper we show that *MEG3* expression negatively correlates with survival in breast cancer, particularly in grade three tumors and the luminal B subtype. However, our study does not determine if the high *MEG3* expression represents increased stromal infiltration in the tumors or elevated expression in cancer cells.

Another reason for classifying *MEG3* as tumor suppressor is its action on stabilization of p53 (Ghafouri-Fard and Taheri, 2019). However, inactivation of p53 is a frequent event in cancer, estimated to have about 50 % occurrence (Gasco et al., 2002; Marine et al., 2006; Haupt and Haupt, 2017). The percentage is even higher, when the inactivation in p53's regulatory pathways is considered (Joerger and Fersht, 2016). Therefore, the use of cell lines which lack active p53, such as D492 and D492M, offers a different approach, more relevant for studying breast cancer signaling pathways, to study the role of DLK1-DIO3. The role of p53 in the cell is that of a tumor suppressor, impacting acts in proliferation, cell cycle and genomic stability (Mercer, 1992). In D492 cell lines, as could be expected, we did not observe effect on cell proliferation. Recently, Uroda and colleagues' stated, that cell cycle arrest by *MEG3* is exclusively p53-dependent, (Uroda et al., 2019), in line with our suggestions that *MEG3* can have a different role in cells lacking p53. Collectively, these observations could explain the conflicting results about role of *MEG3* in tumors.

Many imprinted genes are located in clusters regulated by a differentially methylated regions (DMRs) (Bartolomei and Ferguson-Smith, 2011). In our study targeting the *MEG3* promoter, we have observed concomitant expression of *MEG3* with other miRNAs from the DLK1-DIO3 locus. Our data may support previous studies showing that the *MEG3* promoter controls expression of all maternally expressed genes from the DLK1-DIO3 locus (Tierling et al., 2006; Ioannides et al., 2014; Sanli et al., 2018). Zhu et al. (2019) have shown that the MEG3-DMR overlaps with the *MEG3* gene promoter and any deletion in this region inactivates both MEG3-DMR and the *MEG3* gene. Their data shows, that it is the MEG3-DMR, not the *MEG3* gene, which regulates imprinting (and expression). Therefore, by targeting the *MEG3* promoter at the MEG3-DMR all the non-coding RNAs at the DLK1-DIO3 locus are inactivated. *MEG3* expression can be considered as a marker for the expression of other ncRNAs at the locus.

Cellular plasticity, an important contributor to heterogeneity and drug resistance in breast cancer can be conveyed through EMT/MET (Liu et al., 2014). Partial EMT (p-EMT) may reflect cellular plasticity better than full-EMT and consequently, cells possessing this state adapt more easily to a new environment, which is necessary for cancer cell invasion and metastasis (Thiery, 2002; Tam and Weinberg, 2013; Lambert et al., 2017). Notably, a recent report highlights the importance of the intermediate stages of EMT for the intravasation of tumor cells and for metastasis formation in experimental breast or skin tumors (Pastushenko et al., 2018). Similarly, another study showed that cancer cells might only reach an intermediate EMT stage allowing for increased motility, while keeping its cellular plasticity (Brabletz et al., 2018). It has also been observed that full mesenchymal phenotype (EMT), has a low capacity to form metastasis compared to p-EMT (Schmidt et al., 2015). The essential criteria for aggressive behavior does not need to be a particular phenotype, but rather enhanced cellular plasticity, as is also observed for hybrid E/M cells (Grosse-Wilde et al., 2015). Thus, EMT may be viewed as a trans-differentiation process where epithelial and mesenchymal cells interconvert by passing through an intermediate "stem-like" state (Grosse-Wilde et al., 2018).

EMT is a complex process and meta-analysis indicates that there are possibly different types of EMT (Liang et al., 2016). We have shown, that by manipulating *MEG3* expression, and thus changing the expression of the non-coding genes at the DLK1-DIO3 locus, the majority of these EMT related TFs are affected, indicating an important role of the ncRNAs the DLK1-DIO3 locus in the EMT process. One of the most typical hallmarks of EMT is downregulation of *CDH1* (E-cadherin) and epithelial-specific keratins (Peinado et al., 2007). Altered expression of *MEG3* does not lead to change of E-cadherin expression and therefore *MEG3* may have induced only a partial EMT phenotype. However, it has been shown, that cells with p-EMT phenotype display concomitant expression of epithelial and mesenchymal markers (Armstrong et al., 2011) and loss of E-cadherin is not a prerequisite for EMT (Hollestelle et al., 2013). Cells undergoing collective migration have hybrid EMT phenotype characterized by E-cadherin expression, which helps to maintain cell–cell contacts (Friedl et al., 2012; Aceto et al., 2015). Furthermore, we have shown that altered expression of *MEG3* revealed distinct luminal and myoepithelial marker expression. Increased expression of *KRT14* and decreased expression *KRT19* indicate increased myoepithelial differentiation, which has been connected to a partial EMT phenotype (Petersen et al., 2001). Study on collective migration revealed *KRT14* as a key regulator of metastasis (Cheung et al., 2016) and the same applied for collective invasion, which was facilitated by subpopulation of cells expressing *KRT14* (Cheung et al., 2013). The observed increase of myoepithelial/basal differentiation in cells with higher expression of *MEG3* was supported with altered expression of other markers such as *KRT5*, *TP63*, and *CDH3*.

A key characteristic defining breast stem cells is the ability to form of mammospheres (Dontu et al., 2003; Grosse-Wilde et al., 2015). Morel and colleagues confirmed that human mammary epithelial cells undergoing EMT exhibited better mammosphere-forming capabilities (Morel et al., 2008) and Shimono et al. have shown that mammosphere-forming activity is abrogated in both normal and malignant mammary stem cells when the EMT program is shut down (Shimono et al., 2009). In this study phenotypic differences upon altered MEG3 expression were more prominent in 3D than in 2D cell culture, where *MEG3* increased mammosphere formation ability and slightly decreases branching potential in 3D culture. Furthermore, we have shown increased expression of *ALDH1A3* and *ITGA6*, in cells with overexpression of *MEG3*, supporting role of *MEG3* in stemness.

We propose that increased expression of *MEG3*, and thus increased expression of the ncRNAs at the DLK1-DIO3 locus, in D492 leads to partial EMT phenotype/enhanced plasticity, seen by molecular changes with increased mesenchymal and myoepithelial/basal genes and increased migration and resistance to apoptosis. In contrast, the repression of *MEG3*, and the maternally imprinted ncRNAs, in D492M leads to decreased mesenchymal and basal gene expression and decreased migration and resistance to apoptosis. Nguyen-Ngoc et al.

also demonstrated, that motility can occur in cells that retain an epithelial molecular signature (Nguyen-Ngoc et al., 2012). This supports our observation, that manipulation of *MEG3* expression did not affect the morphological phenotype, but rather affected the functional phenotype. These characteristic properties of cells undergoing EMT were originally proposed to occur in breast cancer by Mani and colleagues (Mani et al., 2008), showing that stem-like and p-EMT properties share many characteristics, such as increased migration, resistance and survival (Creighton et al., 2009; Armstrong et al., 2011; Hanahan and Weinberg, 2011).

Increased understanding of branching morphogenesis in the breast and the regulation of EMT and MET may hold the key for future development of methods and drugs that neutralize the invading properties of cancer cells. Currently, there is need for biomarkers to accurately monitor the EMT/MET process that may improve treatment. Prognostic value of *MEG3* in human malignancies remains controversial and requires further investigation. Our results and conflicting data from the literature suggest that *MEG3* has a complex role in breast tissue.

DATA AVAILABILITY STATEMENT

The RNAseq data for this article has been submitted to GEO, with the GEO accession number GSE142268, see here: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE142268

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Icelandic National Bioethics Committee VSN-13-057 and VSN-11-105-V2. The Icelandic Data Protection Commission (2001/523 and 2002/463) Landspitali Ethical Committee No. 35/2013. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

MM, TG, JB, ZB, EB, GT, and BH: conceptualization and design of the study. ZB, EB, JB, AS, and BH: data acquisition. ZB, AS, EB, GT, SS, and BH: data analysis. ZB, GT, TG, and BH: drafting the manuscript. All authors participated in data interpretation, revision of the manuscript and approved the final version to be published.

FUNDING

This work was supported by Grants from Landspitali University Hospital Science Fund, University of Iceland Research Fund, Icelandic Science and Technology Policy Council Research Fund no. 1103010061, Icelandic Science and Technology Policy – Grant of Excellence: 52144051, "Vísindasjóður Krabbameinsfélagsins" (Icelandic Cancer Society Science Fund) 2017 and 'Göngum saman', a supporting group for breast cancer research in Iceland (www.gongumsaman.is). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Rosa B. Barkardottir, Adalgeir Arason, Bjarni A. Agnarsson, and Oskar Thor Johannsson for their contribution to this work by providing RNA samples from breast cancer patients and pathological and clinical information. We would also like to thank Gudrun Johannesdottir for excellent technical support.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcell.2020.00461/full#supplementary-material

REFERENCES

Aceto, N., Toner, M., Maheswaran, S., and Haber, D. A. (2015). En route to metastasis: circulating tumor cell clusters and epithelial-to-mesenchymal transition. *Trends Cancer* 1, 44–52. doi: 10.1016/j.trecan.2015.07.006

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq–a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31, 166–169. doi: 10.1093/bioinformatics/btu638

Ansieau, S. (2013). EMT in breast cancer stem cell generation. *Cancer Lett.* 338, 63–68. doi: 10.1016/j.canlet.2012.05.014

Armstrong, A. J., Marengo, M. S., Oltean, S., Kemeny, G., Bitting, R. L., Turnbull, J. D., et al. (2011). Circulating tumor cells from patients with advanced prostate and breast cancer display both epithelial and mesenchymal markers. *Mol. Cancer Res.* 9, 997–1007. doi: 10.1158/1541-7786.mcr-10-0490

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. *Cell* 136, 215–233. doi: 10.1016/j.cell.2009.01.002

Bartolomei, M. S., and Ferguson-Smith, A. C. (2011). Mammalian genomic imprinting. *Cold Spring Harb. Perspect. Biol.* 3:a002592.

Baulina, N., Osmak, G., Kiselev, I., Popova, E., Boyko, A., Kulakova, O., et al. (2019). MiRNAs from DLK1-DIO3 imprinted locus at 14q32 are associated with multiple sclerosis: gender-specific expression and regulation of receptor tyrosine kinases signaling. *Cells* 8:133. doi: 10.3390/cells8020133

Ben-Jacob, E., Coffey, D. S., and Levine, H. (2012). Bacterial survival strategies suggest rethinking cancer cooperativity. *Trends Microbiol.* 20, 403–410. doi: 10.1016/j.tim.2012.06.001

Brabletz, T., Kalluri, R., Nieto, M. A., and Weinberg, R. A (2018). EMT in cancer. *Nat. Rev. Cancer.* 18, 128–134. doi: 10.1038/nrc.2017.118

Briem, E., Budkova, Z., Sigurdardottir, A. K., Hilmarsdottir, B., Kricker, J., Timp, W., et al. (2019a). MiR-203a is differentially expressed during branching morphogenesis and EMT in breast progenitor cells and is a repressor of peroxidasin. *Mech. Dev.* 155, 34–47. doi: 10.1016/j.mod.2018.11.002

Briem, E., Ingthorsson, S., Traustadottir, G. A., Hilmarsdottir, B., and Gudjonsson, T. (2019b). Application of the D492 cell lines to explore breast morphogenesis, EMT and cancer progression in 3D culture. *J. Mammary Gland Biol. Neoplasia* 24, 139–147. doi: 10.1007/s10911-018-09424-w

Cancer Genome Atlas Network (2012). Comprehensive molecular portraits of human breast tumours. *Nature* 490, 61–70. doi: 10.1038/nature11412

Cao, Z., Livas, T., and Kyprianou, N. (2016). Anoikis and EMT: lethal "Liaisons" during cancer progression. *Crit. Rev. Oncog.* 21, 155–168. doi: 10.1615/critrevoncog.2016016955

Chak, W. P., Lung, R. W., Tong, J. H., Chan, S. Y., Lun, S. W., Tsao, S. W., et al. (2017). Downregulation of long non-coding RNA MEG3 in nasopharyngeal carcinoma. *Mol. Carcinog.* 56, 1041–1054. doi: 10.1002/mc.22569

Chan, Y. H. (2003). Biostatistics 104: correlational analysis. *Singapore Med. J.* 44, 614–619.

Cheng, A. W., Wang, H., Yang, H., Shi, L., Katz, Y., Theunissen, T. W., et al. (2013). Multiplexed activation of endogenous genes by CRISPR-on, an RNA-guided transcriptional activator system. *Cell Res.* 23, 1163–1171. doi: 10.1038/cr.2013.122

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

Cheung, K. J., Padmanaban, V., Silvestri, V., Schipper, K., Cohen, J. D., Fairchild, A. N., et al. (2016). Polyclonal breast cancer metastases arise from collective dissemination of keratin 14-expressing tumor cell clusters. *Proc. Natl. Acad. Sci. U.S.A.* 113, E854–E863.

Creighton, C. J., Li, X., Landis, M., Dixon, J. M., Neumeister, V. M., Sjolund, A., et al. (2009). Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13820–13825. doi: 10.1073/pnas.0905718106

Cui, X., Yi, Q., Jing, X., Huang, Y., Tian, J., Long, C., et al. (2018). Mining prognostic significance of MEG3 in human breast cancer using bioinformatics analysis. *Cell. Physiol. Biochem.* 50, 41–51. doi: 10.1159/000493956

Dagogo-Jack, I., and Shaw, A. T. (2018). Tumour heterogeneity and resistance to cancer therapies. *Nat. Rev. Clin. Oncol.* 15, 81–94. doi: 10.1038/nrclinonc.2017.166

De Craene, B., and Berx, G. (2013). Regulatory networks defining EMT during cancer initiation and progression. *Nat. Rev. Cancer* 13, 97–110. doi: 10.1038/nrc3447

Di Gesualdo, F., Capaccioli, S., and Lulli, M. (2014). A pathophysiological view of the long non-coding RNA world. *Oncotarget* 5, 10976–10996. doi: 10.18632/oncotarget.2770

Dill, T. L., and Naya, F. J. (2018). A hearty dose of noncoding RNAs: the imprinted *DLK1-DIO3* locus in cardiac development and disease. *J. Cardiovasc. Dev. Dis.* 5:37. doi: 10.3390/jcdd5030037

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Dontu, G., Abdallah, W. M., Foley, J. M., Jackson, K. W., Clarke, M. F., Kawamura, M. J., et al. (2003). In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells. *Genes Dev.* 17, 1253–1270. doi: 10.1101/gad.1061803

Eades, G., Zhang, Y. S., Li, Q. L., Xia, J. X., Yao, Y., and Zhou, Q. (2014). Long non-coding RNAs in stem cells and cancer. *World J. Clin. Oncol.* 5, 134–141.

Elenbaas, B., Spirio, L., Koerner, F., Fleming, M. D., Zimonjic, D. B., Donaher, J. L., et al. (2001). Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. *Genes Dev.* 15, 50–65. doi: 10.1101/gad.828901

Elloumi, F., Hu, Z., Li, Y., Parker, J. S., Gulley, M. L., Amos, K. D., et al. (2011). Systematic bias in genomic classification due to contaminating non-neoplastic tissue in breast tumor samples. *BMC Med. Genomics* 4:54. doi: 10.1186/1755-8794-4-54

Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi: 10.1038/nrg3606

Felipe Lima, J., Nofech-Mozes, S., Bayani, J., and Bartlett, J. M. (2016). EMT in breast carcinoma-a review. *J. Clin. Med.* 5:65. doi: 10.3390/jcm5070065

Fridriksdottir, A. J., Villadsen, R., Morsing, M., Klitgaard, M. C., Kim, J., Petersen, O. W., et al. (2017). Proof of region-specific multipotent progenitors in human breast epithelia. *Proc. Natl. Acad. Sci. U.S.A.* 114, E10102–E10111. doi: 10.1073/pnas.1714063114

Friedl, P., Locker, J., Sahai, E., and Segall, J. E. (2012). Classifying collective cancer cell invasion. *Nat. Cell Biol.* 14, 777–783. doi: 10.1038/ncb2548

Gasco, M., Shami, S., and Crook, T. (2002). The p53 pathway in breast cancer. *Breast Cancer Res.* 4, 70–76.

Ghafouri-Fard, S., and Taheri, M. (2019). Maternally expressed gene 3 (MEG3): a tumor suppressor long non coding RNA. *Biomed. Pharmacother.* 118:109129. doi: 10.1016/j.biopha.2019.109129

Ghoncheh, M., Pournamdar, Z., and Salehiniya, H. (2016). Incidence and mortality and epidemiology of breast cancer in the world. *Asian Pac. J. Cancer Prev.* 17, 43–46. doi: 10.7314/apjcp.2016.17.s3.43

Gilbert, L. A., Larson, M. H., Morsut, L., Liu, Z., Brar, G. A., Torres, S. E., et al. (2013). CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell* 154, 442–451. doi: 10.1016/j.cell.2013.06.044

Grosse-Wilde, A., Fouquier D'herouel, A., Mcintosh, E., Ertaylan, G., Skupin, A., Kuestner, R. E., et al. (2015). Stemness of the hybrid Epithelial/Mesenchymal state in breast cancer and its association with poor survival. *PLoS One* 10:e0126522. doi: 10.1371/journal.pone.0126522

Grosse-Wilde, A., Kuestner, R. E., Skelton, S. M., Macintosh, E., D'herouel, A. F., Ertaylan, G., et al. (2018). Loss of inter-cellular cooperation by complete epithelial-mesenchymal transition supports favorable outcomes in basal breast cancer patients. *Oncotarget* 9, 20018–20033. doi: 10.18632/oncotarget.25034

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

Halldorsson, S., Rohatgi, N., Magnusdottir, M., Choudhary, K. S., Gudjonsson, T., Knutsen, E., et al. (2017). Metabolic re-wiring of isogenic breast epithelial cell lines following epithelial to mesenchymal transition. *Cancer Lett.* 396, 117–129. doi: 10.1016/j.canlet.2017.03.019

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

Harrington, C. T., Lin, E. I., Olson, M. T., and Eshleman, J. R. (2013). Fundamentals of pyrosequencing. *Arch. Pathol. Lab. Med.* 137, 1296–1303. doi: 10.5858/arpa.2012-0463-ra

Haupt, S., and Haupt, Y. (2017). P53 at the start of the 21st century: lessons from elephants. *F1000Res.* 6:2041. doi: 10.12688/f1000research.12682.1

Hilmarsdottir, B., Briem, E., Sigurdsson, V., Franzdottir, S. R., Ringner, M., Arason, A. J., et al. (2015). MicroRNA-200c-141 and Np63 are required for breast epithelial differentiation and branching morphogenesis. *Dev. Biol.* 403, 150–161. doi: 10.1016/j.ydbio.2015.05.007

Hollestelle, A., Peeters, J. K., Smid, M., Timmermans, M., Verhoog, L. C., Westenend, P. J., et al. (2013). Loss of E-cadherin is not a necessity for epithelial to mesenchymal transition in human breast cancer. *Breast Cancer Res. Treat.* 138, 47–57.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009a). Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* 37, 1–13. doi: 10.1093/nar/gkn923

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009b). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Ingthorsson, S., Sigurdsson, V., Fridriksdottir, A. Jr., Jonasson, J. G., Kjartansson, J., Magnusson, M. K., et al. (2010). Endothelial cells stimulate growth of normal and cancerous breast epithelial cells in 3D culture. *BMC Res. Notes* 3:184. doi: 10.1186/1756-0500-3-184

Ioannides, Y., Lokulo-Sodipe, K., Mackay, D. J., Davies, J. H., and Temple, I. K. (2014). Temple syndrome: improving the recognition of an underdiagnosed chromosome 14 imprinting disorder: an analysis of 51 published cases. *J. Med. Genet.* 51, 495–501. doi: 10.1136/jmedgenet-2014-102396

Joerger, A. C., and Fersht, A. R. (2016). The p53 pathway: origins, inactivation in cancer, and emerging therapeutic approaches. *Annu. Rev. Biochem.* 85, 375–404. doi: 10.1146/annurev-biochem-060815-014710

Kaneko, S., Bonasio, R., Saldana-Meyer, R., Yoshida, T., Son, J., Nishino, K., et al. (2014). Interactions between JARID2 and noncoding RNAs regulate PRC2 recruitment to chromatin. *Mol. Cell* 53, 290–300. doi: 10.1016/j.molcel.2013.11.012

Kang, L., Wang, J., Zhang, Y., Kou, Z., and Gao, S. (2009). iPS cells can support full-term development of tetraploid blastocyst-complemented embryos. *Cell Stem Cell* 5, 135–138. doi: 10.1016/j.stem.2009.07.001

Lambert, A. W., Pattabiraman, D. R., and Weinberg, R. A. (2017). Emerging biological principles of metastasis. *Cell* 168, 670–691. doi: 10.1016/j.cell.2016.11.037

Li, J., Shen, H., Xie, H., Ying, Y., Jin, K., Yan, H., et al. (2019). Dysregulation of ncRNAs located at the DLK1DIO3 imprinted domain: involvement in urological cancers. *Cancer Manag. Res.* 11, 777–787. doi: 10.2147/cmar.s190764

Liang, L., Sun, H., Zhang, W., Zhang, M., Yang, X., Kuang, R., et al. (2016). Meta-analysis of EMT datasets reveals different types of EMT. *PLoS One* 11:e0156839. doi: 10.1371/journal.pone.0156839

Liu, H., Ye, D., Chen, A., Tan, D., Zhang, W., Jiang, W., et al. (2019). A pilot study of new promising non-coding RNA diagnostic biomarkers for early-stage colorectal cancers. *Clin. Chem. Lab. Med.* 57, 1073–1083. doi: 10.1515/cclm-2019-0052

Liu, L., Luo, G. Z., Yang, W., Zhao, X., Zheng, Q., Lv, Z., et al. (2010). Activation of the imprinted Dlk1-Dio3 region correlates with pluripotency levels of mouse stem cells. *J. Biol. Chem.* 285, 19483–19490.

Liu, S., Cong, Y., Wang, D., Sun, Y., Deng, L., Liu, Y., et al. (2014). Breast cancer stem cells transition between epithelial and mesenchymal states reflective of their normal counterparts. *Stem Cell Rep.* 2, 78–91. doi: 10.1016/j.stemcr.2013.11.009

Liz, J., and Esteller, M. (2016). lncRNAs and microRNAs with a role in cancer development. *Biochim. Biophys. Acta* 1859, 169–176. doi: 10.1016/j.bbagrm.2015.06.015

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550.

Lu, W., and Kang, Y. (2019). Epithelial-Mesenchymal plasticity in cancer progression and metastasis. *Dev. Cell* 49, 361–374. doi: 10.1016/j.devcel.2019.04.010

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

Marine, J. C., Francoz, S., Maetens, M., Wahl, G., Toledo, F., and Lozano, G. (2006). Keeping p53 in check: essential and synergistic functions of Mdm2 and Mdm4. *Cell Death Differ.* 13, 927–934. doi: 10.1038/sj.cdd.4401912

Mercer, T. R., Dinger, M. E., and Mattick, J. S. (2009). Long non-coding RNAs: insights into functions. *Nat. Rev. Genet.* 10, 155–159. doi: 10.1038/nrg2521

Mercer, W. E. (1992). Cell cycle regulation and the p53 tumor suppressor protein. *Crit. Rev. Eukaryot. Gene Expr.* 2, 251–263.

Mitra, R., Chen, X., Greenawalt, E. J., Maulik, U., Jiang, W., Zhao, Z., et al. (2017). Decoding critical long non-coding RNA in ovarian cancer epithelial-to-mesenchymal transition. *Nat. Commun.* 8:1604.

Molina-Pinelo, S., Salinas, A., Moreno-Mata, N., Ferrer, I., Suarez, R., Andres-Leon, E., et al. (2018). Impact of DLK1-DIO3 imprinted cluster hypomethylation in smoker patients with lung cancer. *Oncotarget* 9, 4395–4410. doi: 10.18632/oncotarget.10611

Moradi, S., Sharifi-Zarchi, A., Ahmadi, A., Mollamohammadi, S., Stubenvoll, A., Gunther, S., et al. (2017). Small RNA sequencing reveals Dlk1-Dio3 locus-embedded MicroRNAs as major drivers of ground-state pluripotency. *Stem Cell Rep.* 9, 2081–2096. doi: 10.1016/j.stemcr.2017.10.009

Morel, A. P., Lievre, M., Thomas, C., Hinkal, G., Ansieau, S., and Puisieux, A. (2008). Generation of breast cancer stem cells through epithelial-mesenchymal transition. *PLoS One* 3:e2888. doi: 10.1371/journal.pone.0002888

Morsing, M., Klitgaard, M. C., Jafari, A., Villadsen, R., Kassem, M., Petersen, O. W., et al. (2016). Evidence of two distinct functionally specialized fibroblast lineages in breast stroma. *Breast Cancer Res.* 18:108.

Moustakas, A., and Heldin, C. H. (2007). Signaling networks guiding epithelial-mesenchymal transitions during embryogenesis and cancer progression. *Cancer Sci.* 98, 1512–1520. doi: 10.1111/j.1349-7006.2007.00550.x

Nguyen-Ngoc, K. V., Cheung, K. J., Brenot, A., Shamir, E. R., Gray, R. S., Hines, W. C., et al. (2012). ECM microenvironment regulates collective migration and local dissemination in normal and malignant mammary epithelium. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2595–E2604.

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

Niknafs, Y. S., Pandian, B., Gajjar, T., Gaudette, Z., Wheelock, K., Maz, M. P., et al. (2018). MiPanda: a resource for analyzing and visualizing next-generation sequencing transcriptomics data. *Neoplasia* 20, 1144–1149. doi: 10.1016/j.neo.2018.09.001

Pastushenko, I., Brisebarre, A., Sifrim, A., Fioramonti, M., Revenco, T., Boumahdi, S., et al. (2018). Identification of the tumour transition states occurring during EMT. *Nature* 556, 463–468. doi: 10.1038/s41586-018-0040-3

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

Peinado, H., Olmeda, D., and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* 7, 415–428. doi: 10.1038/nrc2131

Petersen, O. W., Lind Nielsen, H., Gudjonsson, T., Villadsen, R., Ronnov-Jessen, L., and Bissell, M. J. (2001). The plasticity of human breast carcinoma cells is more than epithelial to mesenchymal conversion. *Breast Cancer Res.* 3, 213–217. doi: 10.1186/bcr298

Prat, A., and Perou, C. M. (2011). Deconstructing the molecular portraits of breast cancer. *Mol. Oncol.* 5, 5–23. doi: 10.1016/j.molonc.2010.11.003

Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P., et al. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* 152, 1173–1183. doi: 10.1016/j.cell.2013.02.022

R Development Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Radisky, D. C., Kenny, P. A., and Bissell, M. J. (2007). Fibrosis and cancer: do myofibroblasts come also from epithelial cells via EMT? *J. Cell. Biochem.* 101, 830–839. doi: 10.1002/jcb.21186

Ringnér, M., Fredlund, E., Häkkinen, J., Borg, Å., and Staaf, J. (2011). GOBO: gene expression-based outcome for breast cancer online. *PLoS One* 6:e17911. doi: 10.1371/journal.pone.0017911

Robson, E. J., Khaled, W. T., Abell, K., and Watson, C. J. (2006). Epithelial-to-mesenchymal transition confers resistance to apoptosis in three murine mammary epithelial cell lines. *Differentiation* 74, 254–264. doi: 10.1111/j.1432-0436.2006.00075.x

Sanli, I., Lalevee, S., Cammisa, M., Perrin, A., Rage, F., Lleres, D., et al. (2018). Meg3 non-coding RNA expression controls imprinting by preventing transcriptional upregulation in cis. *Cell Rep.* 23, 337–348. doi: 10.1016/j.celrep.2018.03.044

Scheel, C., and Weinberg, R. A. (2012). Cancer stem cells and epithelial-mesenchymal transition: concepts and molecular links. *Semin. Cancer Biol.* 22, 396–403. doi: 10.1016/j.semcancer.2012.04.001

Schmidt, J. M., Panzilius, E., Bartsch, H. S., Irmler, M., Beckers, J., Kari, V., et al. (2015). Stem-cell-like properties and epithelial plasticity arise as stable traits after transient Twist1 activation. *Cell Rep.* 10, 131–139. doi: 10.1016/j.celrep.2014.12.032

Sheng, X., Li, J., Yang, L., Chen, Z., Zhao, Q., Tan, L., et al. (2014). Promoter hypermethylation influences the suppressive role of maternally expressed 3, a long non-coding RNA, in the development of epithelial ovarian cancer. *Oncol. Rep.* 32, 277–285. doi: 10.3892/or.2014.3208

Shimono, Y., Zabala, M., Cho, R. W., Lobo, N., Dalerba, P., Qian, D., et al. (2009). Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. *Cell* 138, 592–603. doi: 10.1016/j.cell.2009.07.011

Sigurdsson, V., Hilmarsdottir, B., Sigmundsdottir, H., Fridriksdottir, A. J., Ringner, M., Villadsen, R., et al. (2011). Endothelial induced EMT in breast epithelial cells with stem cell properties. *PLoS One* 6:e23833. doi: 10.1371/journal.pone.0023833

Stadtfeld, M., Apostolou, E., Akutsu, H., Fukuda, A., Follett, P., Natesan, S., et al. (2010a). Aberrant silencing of imprinted genes on chromosome 12qF1 in mouse induced pluripotent stem cells. *Nature* 465, 175–181. doi: 10.1038/nature09017

Stadtfeld, M., Maherali, N., Borkent, M., and Hochedlinger, K. (2010b). A reprogrammable mouse strain from gene-targeted embryonic stem cells. *Nat. Methods* 7, 53–55. doi: 10.1038/nmeth.1409

Sun, L., Li, Y., and Yang, B. (2016). Downregulated long non-coding RNA MEG3 in breast cancer regulates proliferation, migration and invasion by depending on p53's transcriptional activity. *Biochem. Biophys. Res. Commun.* 478, 323–329. doi: 10.1016/j.bbrc.2016.05.031

Sun, L., Luo, H., Liao, Q., Bu, D., Zhao, G., Liu, C., et al. (2013). Systematic study of human long intergenic non-coding RNAs and their impact on cancer. *Sci. China Life Sci.* 56, 324–334. doi: 10.1007/s11427-013-4460-x

Sun, M., Xia, R., Jin, F., Xu, T., Liu, Z., De, W., et al. (2014). Downregulated long noncoding RNA MEG3 is associated with poor prognosis and promotes cell proliferation in gastric cancer. *Tumour Biol.* 35, 1065–1073. doi: 10.1007/s13277-013-1142-z

Tabar, L., Dean, P. B., Yen, A. M., Tarjan, M., Chiu, S. Y., Chen, S. L., et al. (2014). A proposal to unify the classification of breast and prostate cancers based on the

anatomic site of cancer origin and on long-term patient outcome. *Breast Cancer* 8, 15–38. doi: 10.4137/bcbcr.S13833

Tam, W. L., and Weinberg, R. A. (2013). The epigenetics of epithelial-mesenchymal plasticity in cancer. *Nat. Med.* 19, 1438–1449. doi: 10.1038/nm.3336

Terashima, M., Tange, S., Ishimura, A., and Suzuki, T. (2017). MEG3 long noncoding RNA contributes to the epigenetic regulation of epithelial-mesenchymal transition in lung cancer cell lines. *J. Biol. Chem.* 292, 82–99. doi: 10.1074/jbc.m116.750950

Thiery, J. P. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat. Rev. Cancer* 2, 442–454. doi: 10.1038/nrc822

Tierling, S., Dalbert, S., Schoppenhorst, S., Tsai, C. E., Oliger, S., Ferguson-Smith, A. C., et al. (2006). High-resolution map and imprinting analysis of the Gtl2-Dnchc1 domain on mouse chromosome 12. *Genomics* 87, 225–235. doi: 10.1016/j.ygeno.2005.09.018

Uroda, T., Anastasakou, E., Rossi, A., Teulon, J. M., Pellequer, J. L., Annibale, P., et al. (2019). Conserved pseudoknots in lncRNA MEG3 are essential for stimulation of the p53 pathway. *Mol. Cell* 75, 982–995.e9. doi: 10.1016/j.molcel.2019.07.025

Valdmanis, P. N., Roy-Chaudhuri, B., Kim, H. K., Sayles, L. C., Zheng, Y., Chuang, C. H., et al. (2015). Upregulation of the microRNA cluster at the Dlk1-Dio3 locus in lung adenocarcinoma. *Oncogene* 34, 94–103. doi: 10.1038/onc.2013.523

Wang, P., Ren, Z., and Sun, P. (2012). Overexpression of the long non-coding RNA MEG3 impairs in vitro glioma cell proliferation. *J. Cell. Biochem.* 113, 1868–1874. doi: 10.1002/jcb.24055

Wang, Y., and Kong, D. (2018). Knockdown of lncRNA MEG3 inhibits viability, migration, and invasion and promotes apoptosis by sponging miR-127 in osteosarcoma cell. *J. Cell. Biochem.* 119, 669–679. doi: 10.1002/jcb.26230

Wang, Y., and Zhou, B. P. (2013). Epithelial-mesenchymal transition—a hallmark of breast cancer metastasis. *Cancer Hallm.* 1, 38–49. doi: 10.1166/ch.2013.1004

Williams, E. D., Gao, D., Redfern, A., and Thompson, E. W. (2019). Controversies around epithelial-mesenchymal plasticity in cancer metastasis. *Nat. Rev. Cancer* 19, 716–732. doi: 10.1038/s41568-019-0213-x

Yao, Y., Zhang, T., Qi, L., Zhou, C., Wei, J., Feng, F., et al. (2019). Integrated analysis of co-expression and ceRNA network identifies five lncRNAs as prognostic markers for breast cancer. *J. Cell. Mol. Med.* 23, 8410–8419. doi: 10.1111/jcmm.14721

Yersal, O., and Barutca, S. (2014). Biological subtypes of breast cancer: prognostic and therapeutic implications. *World J. Clin. Oncol.* 5, 412–424.

Yin, D. D., Liu, Z. J., Zhang, E., Kong, R., Zhang, Z. H., and Guo, R. H. (2015). Decreased expression of long noncoding RNA MEG3 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer. *Tumour Biol.* 36, 4851–4859. doi: 10.1007/s13277-015-3139-2

Zaravinos, A. (2015). The regulatory role of microRNAs in EMT and cancer. *J. Oncol.* 2015:865816.

Zeisberg, M., and Kalluri, R. (2004). The role of epithelial-to-mesenchymal transition in renal fibrosis. *J. Mol. Med.* 82, 175–181. doi: 10.1007/s00109-003-0517-9

Zhang, X., Rice, K., Wang, Y., Chen, W., Zhong, Y., Nakayama, Y., et al. (2010). Maternally expressed gene 3 (MEG3) noncoding ribonucleic acid: isoform structure, expression, and functions. *Endocrinology* 151, 939–947. doi: 10.1210/en.2009-0657

Zhang, X., Zhou, Y., Mehta, K. R., Danila, D. C., Scolavino, S., Johnson, S. R., et al. (2003). A pituitary-derived MEG3 isoform functions as a growth suppressor in tumor cells. *J. Clin. Endocrinol. Metab.* 88, 5119–5126. doi: 10.1210/jc.2003-030222

Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., et al. (2007). Activation of p53 by MEG3 non-coding RNA. *J. Biol. Chem.* 282, 24731–24742. doi: 10.1074/jbc.m702029200

Zhu, W., Botticelli, E. M., Kery, R. E., Mao, Y., Wang, X., Yang, A., et al. (2019). Meg3-DMR, not the Meg3 gene, regulates imprinting of the Dlk1-Dio3 locus. *Dev. Biol.* 455, 10–18. doi: 10.1016/j.ydbio.2019.07.005

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Article IV

Chapter 26

Application of 3D Culture Assays to Study Breast Morphogenesis, Epithelial Plasticity, and Cellular Interactions in an Epithelial Progenitor Cell Line

Anna Karen Sigurdardottir, Bylgja Hilmarsdottir, Thorarinn Gudjonsson, and Gunnhildur Asta Traustadottir

Abstract

Capturing breast morphogenesis and cancer progression in 3D culture using cell lines with stem cell properties can greatly increase understanding of the underlying mechanisms involved in these processes, highlighting the importance of the culture method. D492 is a breast epithelial progenitor cell line that provides a model for branching morphogenesis when cultured in 3D reconstituted basement membrane matrix (rBM). Along with its derivate cell lines D492M and D492HER2, D492 also serves as a robust model for epithelial to mesenchymal transition (EMT) and tumorigenicity, respectively. Here, we describe the routine maintenance and application of the D492 cell lines in 3D culture for the study of branching morphogenesis, EMT and epithelial-endothelial interaction.

Key words D492 cell lines, 3D culture, Branching morphogenesis, Breast endothelial cells (BRENCs), Matrigel®

1 Introduction

During the reproductive period, breast epithelial stem cells maintain cellular remodeling in the female breast gland by inducing branching morphogenesis in each menstruation cycle. However, the human female breast gland does not reach its full developmental potential until pregnancy [1]. Breast epithelial stem cells, or their immediate progenitor cells, are believed to be the cells of origin for breast cancer [2], but signals from the microenvironment and stroma are increasingly being recognized as an important factor in both normal breast morphogenesis and malignant growth [3–5]. Three-dimensional (3D) cell culture systems have long been acknowledged as an integral part of mammary gland and breast cancer research (reviewed in [6]). Nevertheless, 3D culture is gaining increased recognition as an essential tool in relation to drug

Nagarajan Kannan and Philip Beer (eds.), *Stem Cell Assays: Methods and Protocols*, Methods in Molecular Biology, vol. 2429, https://doi.org/10.1007/978-1-0716-1979-7_26,

discovery, drug testing, and tissue engineering by more closely mimicking the native environment of the cells within the body compared to conventional 2D culture systems.

D492 is an epithelial progenitor cell line established from tissue obtained from reduction mammoplasty [7]. A suprabasal population of $MUC1^{-}/EpCAM^{+}$ cells was isolated from the primary breast tissue, expanded in culture and subsequently immortalized using lentiviral transduction of E6 and E7 oncogenes from human papilloma virus type 16, targeting and silencing the P53 and retinoblastoma (Rb) oncogenes, respectively [7]. The D492 cell line can yield both luminal- and myoepithelial cells in culture and generates branching structures resembling the terminal duct lobular units of the breast (TDLUs) when cultured in 3D reconstituted basement membrane matrix (rBM). Furthermore, D492 cells show cellular plasticity and respond to changes in the microenvironment, dramatically increasing their branching potential when embedded in an rBM co-culture with endothelial cells [8]. Under co-culture conditions the endothelial cells stay as nonproliferative yet metabolically active single cells, while a subpopulation of D492 cells undergoes EMT as evidenced by formation of spindle shaped colonies in between the epithelial branching colonies (Fig. 1). It was from one of these spindle-shaped colonies that the D492M mesenchymal cell line was isolated [8]. Together D492 and D492M provide a robust model for EMT (reviewed in [9]). Moreover, D492 (epithelial) and D492M (mesenchymal-like) are isogenic nontumorigenic cell lines, whereas the D492HER2 cell line, generated by overexpression of the HER2 oncogene in D492, forms tumors in vivo and shows partial EMT [10]. In addition to being used as a model for branching morphogenesis and EMT the D492/D492M/D492HER2 cell lines have been employed to explore epithelial–endothelial interactions [11, 12]. Upon request, D492 and derivative cell lines are freely available for academic use.

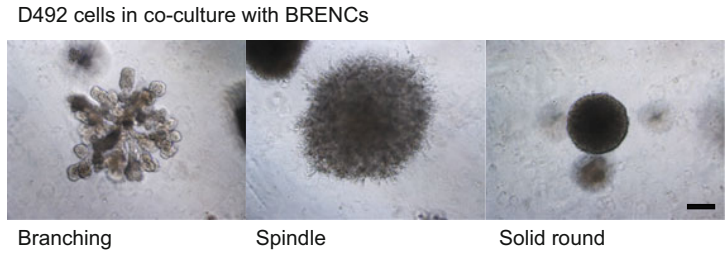


Fig. 1 D492 forms branching, spindle and solid round colonies in co-culture with endothelial cells. In co-culture with breast endothelial cells in Matrigel, D492 cells form three morphologically different colony types: branching, spindle and solid round. Whereas, the branching and solid round colonies have epithelial characteristics, the spindle shaped colonies consist of mesenchymal-like cells that have undergone epithelial-to-mesenchymal transition (EMT) (Scale bar = 100 μm)

2 Materials

Use only sterilized utensils and material for breast tissue isolation and for cell culture. Utensils can be sterilized by autoclaving, whereas solutions may be sterilized by autoclaving or filtration (0.22 μm filter).

2.1 General Maintenance of D492 Cell Lines

1. Pure-Col Type I collagen solution, 3 mg/mL.
2. Cell culture flasks or dishes.
3. Phosphate buffered saline (PBS) without calcium and magnesium.
4. Trypsin–EDTA solution: 0.25% w/v trypsin and 1 mM EDTA in 1 × PBS.
5. Soybean Trypsin Inhibitor (STI), 10 mg/mL.
6. H14 medium [13] consisting of Dulbecco's Modified Eagle's Medium–Nutrient Mixture F-12 (DMEM:F12) with HEPES and L-glutamine, supplemented with; 50 IU/mL penicillin, 50 μg/mL streptomycin, and the following growth factors: 250 ng/mL insulin, 10 ng/mL EGF, 10 μg/mL transferrin, 2.6 ng/mL NaSel, 10^{-10} M estradiol, 500 ng/mL hydrocortisone, and 0.15 IU prolactin.

2.2 Isolation of Breast Endothelial Cells (BRENCs) from Reduction Mammoplasties

1. Breast tissue from reduction mammoplasty (*see* **Note 1**).
2. Polystyrene petri dishes, nontreated, 100 mm, sterile.
3. Scalpels.
4. Forceps.
5. Collagenase solution: DMEM:F12 supplemented with 900 U/mL collagenase-IA.
6. Rotary incubator shaker.
7. Anti-CD31 magnetic microbeads.
8. Magnetic cell sorter.
9. Endothelial Cell Growth Medium BulletKit™ (EGM™) (*see* **Note 2**).
10. Heat-inactivated fetal bovine serum (FBS).

2.3 Preparation of D492 3D Assays

1. Ice in a small container.
2. 24-well cell culture dishes, flat bottom.
3. Hemocytometer.
4. Corning® Matrigel® Basement Membrane Matrix (Corning, 354230) (*see* **Note 3**).
5. H14 medium for monoculture or Endothelial Cell Growth Medium BulletKit™ (EGM™) (*see* **Note 2**), for co-culture, supplemented with 5% heat-inactivated fetal bovine serum (FBS).

2.4 Isolation of 3D Structures

1. 5 mM EDTA, ice cold.
2. Small spatula.
3. Shaker.
4. Ice.

2.5 Immunostaining of 3D Structures

1. 1 × PBS.
2. 4% paraformaldehyde.
3. 10% FBS.
4. IF buffer consisting of distilled H_2O supplemented with; 0.15 M NaCl, 5 mM EDTA, 20 mM HEPES, and 0.1% Triton X-100.
5. Primary and secondary antibodies.
6. DAPI or phalloidin.
7. Distilled water.
8. Spatula.
9. Microscope glass slide.
10. Coverslips.
11. FixoGum.

3 Methods

3.1 General Maintenance of D492 Cell Lines

3.1.1 Collagen Coating of Culture Vessels

1. Add 1 mL Pure-Col Type I collagen solution, 3 mg/mL, to 44 mL 1×PBS and mix.
2. Add 3 mL collagen I solution to a 25 cm^2 culture vessel (*see* **Note 4**).
3. Incubate for 30 min at 37 °C, 1 h at room temperature or overnight at 4 °C (*see* **Note 5**).
4. Remove collagen I solution and rinse the culture vessel with 1× PBS. Remove the PBS and plate cells.

3.1.2 Retrieving D492 Cell Lines from Liquid Nitrogen

1. Thaw one vial of D492 cells in a 37 °C water bath or by the warmth of your hands.
2. When small ice crystals remain, the remaining ice is dissolved by pipetting up and down with a 1 mL pipette. Subsequently transfer cells to a 15 mL tube containing 9 mL 1× PBS (*see* **Note 6**).
3. Pellet cells at 200 × *g* for 5 min. Remove supernatant, resuspend in 4 mL of prewarmed H14 medium, and plate in a 25 cm^2 flask.

3.1.3 Maintaining D492 Cell Lines

1. D492 cells are routinely maintained in collagen I coated 25 cm^2 flasks at 37 °C, 5% CO_2 (*see* **Note 7**).
2. Cells are cultured in H14 medium prewarmed to 37 °C. Medium is changed 3 times per week.
3. D492 cells are routinely frozen in 10% DMSO, 40% FBS, and 50% DMEM:F12 when 70–80% confluent (*see* **Note 8**).

3.1.4 Passaging D492 Cell Lines

1. Passage cells when they are 90% confluent.
2. Aspirate medium and rinse the flask once with 1× PBS.
3. Add 1 mL of Trypsin-EDTA solution to each 25cm^2 flask.
4. Incubate at 37 °C until cells are detached (*see* **Note 9**).
5. Stop trypsinization by adding 20 μL of STI to the flask.
6. Add 4 mL 1× PBS and gently pipet up and down.
7. Transfer to a 15 mL tube and centrifuge at 200 × *g* for 3 min.
8. Aspirate supernatant and resuspend cell pellet in 1 mL prewarmed H14 media.
9. Cells are subcultured at 1:10 ratio for D492M and 1:20 ratio for D492 and D492HER2 (*see* **Note 10**).

3.2 Isolation of Breast Endothelial Cells (BRENCs) from Reduction Mammoplasties (Fig. 2)

Day 1

1. Place fresh pieces of breast tissue in a sterile petri dish inside the cell hood (*see* **Notes 11** and **12**).
2. The breast tissue is composed of adipose tissue as well as epithelial and stromal components (*see* **Note 13**). To facilitate digestion of the tissue, carefully chop the tissue into small pieces (~1 mm) using two opposing scalpel blades (*see* **Note 14**). Place about 30 mL of tissue in a 75 cm^2 flask (*see* **Note 15**).
3. Add 15 mL of prepared collagenase solution to each 75 cm^2 flask.
4. Digest on a rotary shaker (60 rpm) at 37 °C overnight.

Day 2

1. Transfer digested adipose tissue into 50 mL tubes (*see* **Note 16**).
2. Centrifuge at 200 × *g* for 40 s.
3. After centrifugation the liquid will have separated into layers (*see* **Note 17**). Discard the oil from the top and transfer the brown liquid phase, containing endothelial cells, into a new 50 mL tube.
4. Centrifuge at 200 × *g* for 5 min.

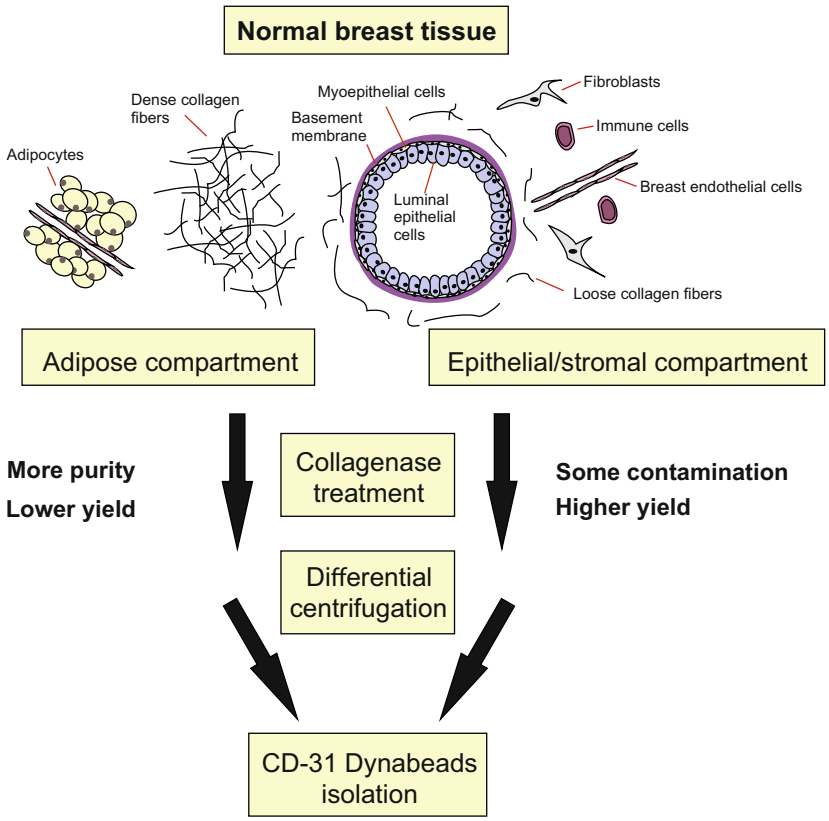


Fig. 2 Isolation of breast endothelial cells (BRENCs) from reduction mammoplasties. Human breast stroma is comprised of intralobular loose connective tissue, dense collagenous interlobular stroma and adipose tissue. Isolation of BRENCs from the adipose compartment gives more purity but lower cell yield compared to isolation from the epithelial/stromal compartment, where cellular contamination is more frequent. After treatment with collagenase and subsequent differential centrifugation, BRENCs can be isolated via magnetic sorting using CD-31 dynabeads

5. Remove supernatant. Wash once with 20 mL 1× PBS, spin down as previously and remove the supernatant.
6. Incubate pellet containing microvessel organoids with anti-CD31 magnetic microbeads as per manufacturer instructions and separate CD31 positive endothelial cells from the supernatant using a magnetic system.
7. Seed cells into 25 cm^2 flasks precoated with collagen I in EGM™ media supplemented with 30% FBS.
8. Maintain endothelial cells in EGM™ media supplemented with 30% FBS until the first passage and then reduce serum concentration to 5% (*see* **Note 18**).

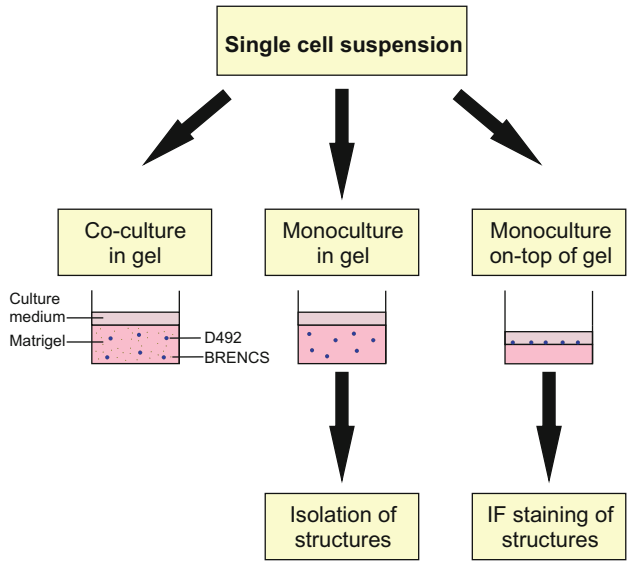


Fig. 3 Three-dimensional culture assays to study breast morphogenesis, epithelial plasticity and cellular interactions. D492 cells can be cultured in Matrigel® in monoculture or in co-culture with BRENCs. Alternatively, D492 cells can be cultured in monoculture on-top of gel, where branching colonies take less time to form and less gel is used

3.3 Preparation of D492 3D Assays (Fig. 3)

3.3.1 D492 Monoculture in Matrigel®

1. Remove an aliquot of Matrigel® from −20 °C and thaw on ice for 2–3 h or overnight (*see* **Note 19**).
2. Place appropriate number of cryovials (v-shaped bottom) on ice.
3. Detach cells (*see* Subheading 3.1.4) and resuspend in a small volume. Count cells using a hemocytometer. Add 10,000 cells (*see* **Note 20**) to a cryovial. For volume >10 μL centrifuge at 200 × *g* for 5 min and aspirate supernatant.
4. Gently resuspend cell pellet in 300 μL Matrigel®. Pipet up and down 6 times to mix (*see* **Note 21**).
5. Quickly transfer Matrigel®–cell mix to a well of a 24-well plate.
6. Let Matrigel®–cell mix set in an incubator at 37 °C for 30 min.
7. Gently add 500 μL H14 medium on top of gel (*see* **Note 22**).
8. Change media three times a week by removing 250 μL of old media and gently adding 250 μL of fresh media (*see* **Note 23**).
9. Record progress of cells by phase contrast imaging. Full branching of D492 cells is reached within 2 weeks (*see* **Note 24**).

3.3.2 D492 Co-culture with Endothelial Cells in Matrigel®

Same experimental setup as for D492 3D monoculture assay (*see* Subheading 3.3.1), with the following exceptions.

1. Add 500 D492 cells and 100,000 endothelial cells to a cryovial before spinning and removal of supernatant (*see* **Note 25**).
2. Use EGM™ medium supplemented with 5% FBS instead of H14.
3. As with the 3D monocultures, full branching will be completed in 14 days. The endothelial cells in the culture are viable throughout the 2 weeks but will not proliferate in the Matrigel®.

3.3.3 Isolation of 3D Colonies in Matrigel®

1. Remove medium.
2. Add 1 mL of ice cold 5 mM EDTA to a 15 mL tube (per sample).
3. Add 100 μL of 5 mM EDTA to each well.
4. Use a small spatula to separate the gel from the edges of the well.
5. Use a spatula to move the gel into the 15 mL tube.
6. Add 100 μL of 5 mM EDTA to the well and aspirate the remaining gel and add to the 15 mL tube.
7. Fill up to 5 mL with 5 mM EDTA in the 15 mL tube. Place the tube on ice and on a shaker.
8. After 20 min, check whether the gel is dissolved. If gel is not dissolved, continue shaking on ice for another 20 min (*see* **Note 26**).
9. Use structures either for isolation of cell populations, protein/RNA/DNA extraction or for immunofluorescent staining.

3.3.4 D492 Monoculture on-Top of Matrigel® (See **Note 27**)

1. Remove an aliquot of Matrigel® from −20 °C and thaw on ice for 2–3 h or overnight.
2. Place 50 μL of Matrigel® per well in a 96 well plate. Incubate at 37 °C for 30 min.
3. Detach cells (*see* Subheading 3.1.4) and resuspend in a small volume. Count cells using a hemocytometer.
4. Plate 2500 cells per well in 200 μL H14 media (*see* **Note 28**).
5. Change media three times per week by gently removing 100 μL of media and adding 100 μL of fresh media instead (*see* **Note 29**).
6. Evaluate colony morphology, size, and number, 8 days after plating (*see* **Note 30**).
7. Proceed with RNA/protein isolation and/or immunostaining (*see* **Note 31**).

3.3.5 Immunostaining of 3D Structures on-Top of Matrigel®

1. Gently remove medium from the wells.
2. Rinse wells 3 times with 1× PBS by holding the plate on its side and letting the liquid seep down the side of the well.
3. Fix cells with 4% formaldehyde in 1× PBS for 30 min at room temperature.
4. Repeat **step 2** (*see* **Note 32**).
5. Block for 10 min in IF buffer with 10% FBS.
6. Repeat **step 2**.
7. Incubate with primary antibody in IF buffer for 1 h at room temperature or overnight at 4 °C.
8. Rinse wells with IF buffer, three times, as described in **step 2**. One quick wash and subsequently twice for 5 min.
9. Incubate with fluorescent secondary antibody and DAPI or phalloidin in IF buffer for 30 min at room temperature.
10. Repeat **step 8**.
11. Rinse with distilled water.
12. Aspirate liquid and take a small spatula and scoop up the Matrigel® with colonies from each well.
13. Place Matrigel® with colonies on a glass slide (*see* **Note 33**).
14. Mount and cover with glass coverslip. Seal sides of coverslip with FixoGum and let dry at room temperature in the dark (*see* **Note 34**).
15. Proceed with imaging.

4 Notes

1. Breast tissue can be kept at 4 °C in DMEM:F12 for a few hours before the isolation procedure is carried out.
2. All growth factors are provided with the kit; however, the gentamicin–amphotericin B mixture, also provided with the kit, is not used.
3. In order to avoid repeated freeze–thaw cycles, Matrigel® should be aliquoted and kept at −20 °C.
4. For other types of culture vessels add enough collagen I solution to cover the surface.
5. Flasks can be kept at 4 °C with the collagen I solution for up to 2 weeks.
6. D492 cell lines are sensitive to the DMSO in the freezing media. Therefore, it is important to remove the DMSO before the cells are plated.
7. D492 cells are very sensitive to CO_2-levels and humidity in the incubator. Growth will halt at suboptimal growth conditions.

8. Cells from one 25 cm^2 flask can be divided into three cryovials.
9. The incubation time varies between cell types, D492 will detach after 3–5 min while D492M and D492HER2, having a more mesenchymal phenotype, need about 1 min of trypsinization. If D492 cells are overconfluent or starved, detaching the cells can be difficult. Then it is possible to facilitate detachment by either rinsing once with the trypsin–EDTA solution before 5 min trypsinization, or by trypsinizing twice and collecting all supernatant in the same tube.
10. Maintaining the cells at this ratio will require the cells to be split once a week.
11. All work with primary breast tissue should be carried out at sterile conditions.
12. The cutting of the tissue may be facilitated by adding a few mL of 1× PBS to the petri dish.
13. BRENCs are found in abundance and are most readily isolated from the adipose tissue of the breast. Isolation from the epithelial and stromal components results in higher cell yield but often has cellular contamination, mainly from fibroblasts.
14. Macroscopic examination of the breast tissue reveals two kinds of tissue: White strands contain the epithelium and stroma while yellow sections contain mostly adipose tissue.
15. Our main source of primary tissue is from breast reduction surgery. As there is plenty of tissue from each surgery we aim to process at least 60 mL of tissue in two 75 cm^2 flask.
16. All pieces of tissue should be digested at this point. If not, we recommend adding 6–12 h to the digestion time.
17. The liquid has now been separated to three layers: Top layer is adipose tissue, the pellet contains organoids and big veins and the middle phase contains microvessels and single cells. From here it is possible to isolate endothelial cells and fibroblasts.
18. Breast endothelial cells can be passaged up to 12 times. During the first passages, fibroblasts may be present in the endothelial culture. Fibroblasts take longer than endothelial cells to detach when subcultured and can thereby be gradually eliminated from the culture by stopping the trypsinization process when 50–60% of the cells have been detached.
19. Matrigel® should be kept on ice at all times until it is in the plate with the cells. Keep ice in a small container inside the hood while working.
20. The exact composition of Matrigel® can vary from lot to lot. It is essential to assess colony formation in every new batch of Matrigel®, by testing different cell numbers. If too many cells are seeded, colony growth may be prevented by steric

hindrance. To reduce the need to test the Matrigel® frequently it is possible to test 3–4 batches of Matrigel® at the same time and subsequently order multiple vials from the best lot that was tested. When Matrigel® is tested we set up 3D monocultures for D492 ranging from 5000 cells/300 μL–30,000 cells/300 μL of Matrigel®. A suitable batch will support the growth of branching 3D colonies.

21. Avoid generating bubbles when pipetting Matrigel®.
22. After cells are seeded in Matrigel®, the first 24 h are critical for their growth. If growth of 3D organoids is problematic it is possible to add EGM™ media with 5% FBS on top of the 3D cultures overnight and replace with H14 media the subsequent day.
23. The 3D assay can be scaled down to 96 well plate where 100 μL of Matrigel® are used and 100 μL of media. Also, it is possible to scale down to an 8 well chamber slide where 50 μL of gel and 200 μL of media are used. The number of cells is scaled down in line with the amount of Matrigel® used.
24. Rarely cells reach the bottom of the cell culture plate and grow in monolayer on the plastic, making it difficult to properly visualize the colonies within the gel.
25. Instead of breast endothelial cells (BRENCs), human umbilical vein endothelial cells (HUVECs) can be used.
26. The gel is dissolved when all colonies move freely in the EDTA solution.
27. In a 3D assay on-top of Matrigel®, less Matrigel® is used compared to 3D assay in gel and it takes less time for branching colonies to form. The 3D monoculture protocol on-top of Matrigel® is adapted from [14].
28. The exact number of cells may need to be adjusted.
29. Colonies growing on-top of Matrigel® can easily detach. Therefore, liquid must be carefully aspirated from the wells, both during change of media and subsequent treatments of the culture, such as during immunostaining.
30. Colonies form much faster when grown on-top of Matrigel® as compared to when they are cultured in the gel.
31. When isolating RNA or protein from 3D cultures on-top of Matrigel® please follow the same procedure as is described for isolation of 3D colonies embedded in Matrigel® (*see* Subheading 3.3.3).
32. Fixed colonies can be kept in IF buffer at 4 °C for a few days. Be careful to not let them dry out.

33. It is not necessary to transfer the Matrigel® with colonies to a glass slide if a culture vessel with high optical quality, such as μ-Slide 8 well from Ibidi, is used instead of a 96-well plate.
34. If the Matrigel® is too thick, consider adding spacers between the glass slide and the coverslip.

Acknowledgments

This work was supported by grants from Landspitali University Hospital Science Fund, University of Iceland Research Fund, and Icelandic Science and Technology Policy—Grant of Excellence: 152144051, "Göngum saman," a supporting group for breast cancer research in Iceland (www.gongumsaman.is) and the Science Fund of the Icelandic Cancer Society.

References

1. Macias H, Hinck L (2012) Mammary gland development. Wiley Interdiscip Rev Dev Biol 1(4):533–557. https://doi.org/10.1002/wdev.35
2. Visvader JE (2009) Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev 23(22): 2563–2577. https://doi.org/10.1101/gad.1849509
3. Acerbi I, Cassereau L, Dean I, Shi Q, Au A, Park C, Chen YY, Liphardt J, Hwang ES, Weaver VM (2015) Human breast cancer invasion and aggression correlates with ECM stiffening and immune cell infiltration. Integr Biol (Camb) 7(10):1120–1134. https://doi.org/10.1039/c5ib00040h
4. Tomko LA, Hill RC, Barrett A, Szulczewski JM, Conklin MW, Eliceiri KW, Keely PJ, Hansen KC, Ponik SM (2018) Targeted matrisome analysis identifies thrombospondin-2 and tenascin-C in aligned collagen stroma from invasive breast carcinoma. Sci Rep 8(1): 12941. https://doi.org/10.1038/s41598-018-31126-w
5. Quail DF, Joyce JA (2013) Microenvironmental regulation of tumor progression and metastasis. Nat Med 19(11):1423–1437. https://doi.org/10.1038/nm.3394
6. Sumbal J, Budkova Z, Traustadottir GA, Koledova Z (2020) Mammary organoids and 3D cell cultures: old dogs with new tricks. J Mammary Gland Biol Neoplasia. https://doi.org/10.1007/s10911-020-09468-x
7. Gudjonsson T, Villadsen R, Nielsen HL, Ronnov-Jessen L, Bissell MJ, Petersen OW (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
8. Sigurdsson V, Hilmarsdottir B, Sigmundsdottir H, Fridriksdottir AJ, Ringner M, Villadsen R, Borg A, Agnarsson BA, Petersen OW, Magnusson MK, Gudjonsson T (2011) Endothelial induced EMT in breast epithelial cells with stem cell properties. PLoS One 6(9):e23833. https://doi.org/10.1371/journal.pone.0023833
9. Briem E, Ingthorsson S, Traustadottir GA, Hilmarsdottir B, Gudjonsson T (2019) Application of the D492 cell lines to explore breast morphogenesis, EMT and cancer progression in 3D culture. J Mammary Gland Biol Neoplasia 24(2):139–147. https://doi.org/10.1007/s10911-018-09424-w
10. Ingthorsson S, Andersen K, Hilmarsdottir B, Maelandsmo GM, Magnusson MK, Gudjonsson T (2016) HER2 induced EMT and tumorigenicity in breast epithelial progenitor cells is inhibited by coexpression of EGFR. Oncogene 35(32):4244–4255. https://doi.org/10.1038/onc.2015.489
11. Morera E, Steinhauser SS, Budkova Z, Ingthorsson S, Kricker J, Krueger A, Traustadottir GA, Gudjonsson T (2019) YKL-40/CHI3L1 facilitates migration and invasion in HER2 overexpressing breast epithelial progenitor cells and generates a niche for capillary-like network formation. In Vitro Cell Dev Biol Anim 55(10):838–853. https://doi.org/10.1007/s11626-019-00403-x

12. Steinhaeuser SS, Morera E, Budkova Z, Schepsky A, Wang Q, Rolfsson O, Riedel A, Krueger A, Hilmarsdottir B, Maelandsmo GM, Valdimarsdottir B, Sigurdardottir AK, Agnarsson BA, Jonasson JG, Ingthorsson S, Traustadottir GA, Oskarsson T, Gudjonsson T (2020) ECM1 secreted by HER2-overexpressing breast cancer cells promotes formation of a vascular niche accelerating cancer cell migration and invasion. Lab Invest 100(7): 928–944. https://doi.org/10.1038/s41374-020-0415-6
13. Blaschke RJ, Howlett AR, Desprez PY, Petersen OW, Bissell MJ (1994) Cell differentiation by extracellular matrix components. Methods Enzymol 245:535–556. https://doi.org/10.1016/0076-6879(94)45027-7
14. Lee GY, Kenny PA, Lee EH, Bissell MJ (2007) Three-dimensional culture models of normal and malignant breast epithelial cells. Nat Methods 4(4):359–365. https://doi.org/10.1038/nmeth1015