Improved orthotopic and metastatic breast cancer models incorporating key elements of the tumour microenvironment enabling patient-relevant drug testing

NIOVI NICOLAOU, M.SC.

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

July 2015

ABSTRACT

Breast cancer (BC) is the most commonly diagnosed cancer in the UK and the second leading cause of cancer-related deaths in women worldwide. Limited treatment is available to patients with metastatic BCs, which are resistant to therapy. There is a strong link between metastatic BC and Epithelial Mesenchymal Transition (EMT). There is a need to target EMT, in order to prevent metastasis for which the therapy is not effective. Therefore, there is a need for improved pre-clinical models that will allow studying EMT in real-time. This study aimed to refine current *in vivo* models of BC, by incorporating inducible bioluminescent reporter(s) that will allow real-time read-out of EMT. The cell-lines used were classified according to their expression of EMT markers and cancer stem cell (CSC) profile. Inducible bioluminescent EMT reporters based on the S100A4 or E-cadherin or N-cadherin promoter driving firefly luciferase were constructed and used to transduce the cell lines. The activity of the S100A4 reporter was validated in hypoxic conditions in MCF-7 and BT549 cells, where it was shown that induction of S100A4-generated bioluminescence was associated with upregulation in *S100A4*. *In vivo* it was shown that human mesenchymal stem cells promoted the tumour growth of MCF-7 cells and/or induced EMT-related traits, including downregulation/loss of E-cadherin, upregulation of Vimentin and/or upregulation of S100A4. The activity of the S100A4 reporter was tested in a similar *in vivo* model, where it was shown that S100A4-generated bioluminescence correlated with protein expression of S100A4. Taken all the data together, we generated an inducible bioluminescent EMT reporter that can detect the baseline levels and changes in the S100A4 expression. The model can be potentially used for giving a real-time read out of EMT/early stages of EMT and subsequently for testing novel drugs targeted at the onset of EMT.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisors, Dr Anna Grabowska and Dr David Onion, for their aspiring guidance and support throughout my PhD project. I am thankful for providing me with their valuable insights, constructive advice and sharing truthful and illuminating views on different aspects related to the project. I am also very grateful for their enthusiasm, kindness and friendly support throughout the years.

I would also like to thank my former supervisor, Dr Nektaria Papadopoulou, for her help and support during the initial stages of my PhD. I want to express my appreciation for helping me settling in and teaching me how to perform various techniques involved.

Furthermore, I want to thank my internal assessor, Dr Cinzia Allegrucci, for correcting my reports, giving me valuable advice, and guiding me in the right direction.

I want to say special thanks to Miss Pamela Collier, Dr Phil Clarke and Miss Alison Ritchie for their constant technical support and respected advice throughout the project. I am truly grateful for their help, and all the knowledge and skills they have passed on to me.

Also, I would like to thank all the other members of our group for their help with various aspects of my PhD. Specifically I want to express my thanks to Miss Marian Meakin and Miss Alison Mackie for their help with the *in vivo* studies, and our lab aid, Pete, for his daily assistance. Also I want to thank Miss Beverly Bramley and Miss Donna Canon for their administrative support. Finally, I want to say a big thank you to all the PhD students and other members of staff, Luana, Eroje, Laura, Francesca, Mansi, John, Matt, and especially Maria, for their kind collaboration, technical help, and for making this journey easier and more enjoyable.

I want to say a massive thank you and express my love to all my family. I am extremely grateful to my lovely mum and dad for supporting me throughout the years, being there for me in every step of this journey, and

always believing in me, even when I did not believe in myself. I want to thank my amazing sister for always being there for me to put a smile on my face, and helping me get through any difficult situation that I had to face throughout the years. Also, I want to thank my lovely brother for his love and support.

Finally, I want to say a special thank you to my wonderful husband-to-be for loving and supporting me in every possible way, and being there for me every single day throughout the years.

I want to dedicate my thesis to my fiancé, my family, and loved ones who might not be with us, but they are always in my heart.

TABLE OF CONTENTS

TABLE OF FIGURES

LIST OF TABLES

GLOSSARY

ALDH=Aldehyde dehydrogenase
BC=Breast cancer
BCSCs=Breast CSC
BrCAFs=Breast cancer CAFs
BSA=Albumin from bovine serum
CAFs=Cancer associated fibroblasts
cDNA=Complementary DNA
CSCs=Cancer stem cells
dNTPs=Deoxynucleotides
EGFR=Epidermal growth factor receptor
EMT=Epithelial-mesenchymal transition
EpCAM=Epithelial adhesion molecule
ER=Oestrogen receptor
FC=Flow cytometry
FSP1(S100A4)=Fibroblast specific protein
HBSS=Hank's balanced salt solution
HIF=Hypoxia-inducible factor
hMSCs=Human MSCs
IF=Immunofluorescence
IgG=Immunoglobulin G
IHC=Immunohistochemistry
IL=Interleukin
IRES=Internal ribosome entry site
mAb=Monoclonal antibody
MET=Mesenchymal-epithelial transition
MFP=Mammary fat pad
mRNA=Messenger RNA
MSCs=Mesenchymal stem cells
O/N=Overnight
PBS=Phosphate buffered saline
PCR=Polymerase chain reaction
PDX=Patient-derived xenograft
PR=Progesterone receptor
qPCR=Quantitative PCR
qRT-PCR=Quantitative reverse transcriptase PCR
RT=Room temperature
SC=Subcutaneous
TAM=Tumour associated macrophages
TF=Transcription factor
TGF=Tumour growth factor
TME=Tumour microenvironment
TNF=Tumour necrosis factor
VEGF=Vascular endothelial growth factor

CHAPTER 1: INTRODUCTION

1.1 Breast Cancer

1.1.1 Incidence

Despite the many advances in breast cancer (BC) therapy including improvements in the local management of breast cancer [1], implementation of screening, and introduction of adjuvant treatments [2], breast cancer is still the most commonly diagnosed cancer in the UK and the leading cause of cancer-related death in females worldwide [3]. In the UK, BC is the second most common cause of cancer death in women, accounted for 11,633 deaths in 2010, 1% of which were in men (latest statistics were obtained from www.cancerresearchuk.org). A new study revealed that most deaths occur in women who were not regularly screened [4]. Failure to detect and treat primary cancers at early stages could result in metastatic breast cancer, in which the cancer has spread to regional and distant sites including lymph nodes, lung, liver, brain and bones [5], and there is currently no effective treatment available.

1.1.2 BC subtypes and clinical implications

The clinical scenario is further complicated by the heterogeneity of breast cancers and their diversity in responsiveness to treatment. In an early study, Perou, C.M., et al. (2000) [6] analysed gene expression in specimens from human breast tumours, and demonstrated two different cancer types which were either oestrogen receptor positive (ER+) or negative (ER-). Analysis of the “intrinsic gene expression” signatures of BCs, allowed further classification of breast cancer into five molecular subgroups defined as luminal A, luminal B, HER2+, basal-like and normal-like [7, 8]. The different subgroups could be explained by their cellular origin, originating from either luminal or basal/myoepithelial progenitors within the normal breast tissue (reviewed in [9]), although the cellular origin of BC is still unclear. Luminal subtypes are ER+ and PR+ and they express luminal markers. While Luminal A tumours do not express HER2, Luminal B tumours are either HER2+ or HER2- [10]. The HER2+ subgroup is characterised by HER2 expression and may lack the

expression of ER and PR. The basal-like tumours are further subdivided into Basal-A and Basal-B tumours and are characterised by the lack of ER, PR and HER2. They are also known as “triple-negative” tumours. Basal-B tumours are also characterised by a ‘claudin-low’ molecular portrait. While Basal-A tumours have mixed luminal/basal features, Basal-B/ claudin-low tumours express markers that define the basal lineage. Finally, the normal-like group shows a genetic profile that resembles the normal breast tissue. Different subgroups have different clinical outcomes, and depending on the molecular profile discussed above, a different therapeutic approach is taken; for example hormonal therapy can be used to treat those cancers that express the ER receptors. Limited treatment is available to patients with triple-negative breast cancers [11].

1.1.3 Triple-negative cancers

Triple-negative cancers, show a more aggressive phenotype compared to other subtypes [7, 12-14], and several studies have associated them with poor prognosis, high incidence of metastasis, and high rate of disease recurrence after therapy [11]. Although effective targeted therapies have been developed to target those patients with hormone receptor-positive tumours, the targeted therapy for triple-negative cancers is not effective. As triple-negative cancers share features with BRCA-1-related cancers (cancers carrying a mutation of the BRCA1 tumour suppressor gene) including high nuclear grade, cytokeratin 5/6 expression, Ki-67 staining, ER negativity and epidermal growth factor receptor (EGFR) expression, a shared therapeutic approach is suggested [15]. However, triple-negative cancers show reduced sensitivity to chemotherapy. This can be explained by their aggressive phenotype: the morphological appearance of these tumours is consistently characterised by high proliferation, high rate of genetic aberration, high grade, central necrosis, frequent apoptotic cells and an advancing border of invasion [7, 12-14]. Furthermore, resistance to therapy can be explained by the fact that Basal-B and BRCA1-related cancers are enriched for cells with stem-cell like characteristics [16, 17],

and have enhanced invasive properties and a mesenchymal gene signature associated with epithelial-mesenchymal transition (EMT) [18].

1.2 Cancer Stem Cells (CSCs)

1.2.1 The Cancer Stem (CSC) Model

The cancer stem cell (CSC) hypothesis has become a major focus in cancer research. It might explain why cancers are not curable, and promises new therapeutic strategies that target only the malignant cells [19].

Cells within the tumour population demonstrate different proliferative and differentiative capacities known as tumour heterogeneity, which can be explained by two different models: the CSC [20, 21] and the clonal evolution model [22, 23]. The clonal evolution model proposes that most malignancies arise from a single normal cell through accumulation of mutations. This cell is characterised by genetic and epigenetic instabilities that are selected and passed to the offspring that also acquire mutations, forming a mass of genetically varied cancer cells. The CSC model comprises two distinct but related components [24]. The first proposes that cancers originate in cell populations that preserve or acquire the properties of normal stem cells to self-renew. These cells can be mutated normal stem cells, transit-amplifying cells, or differentiated cells (Figure 1). The second component holds that cancers are organised in a hierarchical fashion, in which stem cells both drive the malignant process and generate non-renewing cells that form the tumour bulk. The two models are not mutually exclusive in those cancers that are organised hierarchically into populations of tumorigenic and non-tumorigenic cells; in these cells clonal evolution occurs in the CSCs.

Evidence for the clonal evolution model came from different studies including (a) cytogenetic studies that showed that all cells within tumours had the same karyotype, suggesting that cells arose from the same origin, (b) studies of the isoenzymes of glucose-6-phosphate

dehydrogenase in neoplasms in heterozygous women indicated that the same member of the X chromosome pair is functional in all cells of a given tumour, indicating that cells descended from a single precursor and (c) studies that showed that the immunoglobulin produced by plasma cell tumours has in almost every case the homogeneity characteristic of a single clone [22]. Evidence for the CSC model arose from studies in leukaemia; for example, Dick and colleagues showed that adult human AML stem cells can be purified from patients and even though they comprised a very small proportion of AML cells, they were the only cells that could form tumours in immunocompromised mice [21]. This proved that not all the cells have the same tumourigenic capacity. Nonetheless, data over the past few years support the existence of CSCs in solid malignancies. The first evidence of CSCs in solid tumours came from breast cancer, and subsequently CSCs were isolated from brain tumours, colorectal and pancreatic carcinomas [25].

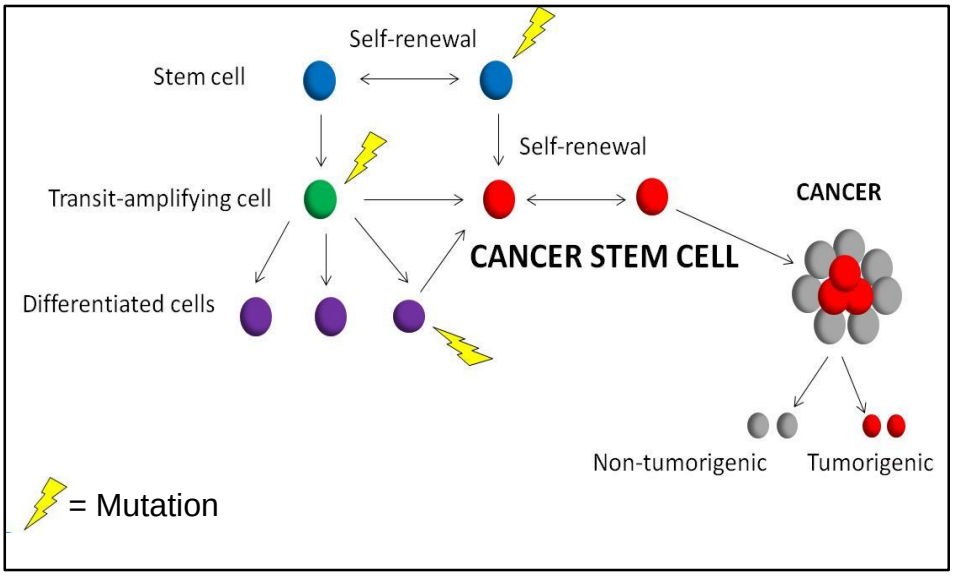


Figure 1. The CSC model. Mutated stem cells, transit-amplifying cells or differentiated cells acquire the ability to self-renew and form tumours, composed of both tumorigenic and non-tumourigenic cells.

CSCs share properties with normal stem cells: they can self-renew through symmetrical or asymmetrical divisions, they can differentiate, and colonise other parts of the body [26]. It is suggested that CSCs are resistant to treatment for many reasons including their low proliferative

rate, quiescence, high efficiency of DNA repair, high levels of anti-apoptotic molecules, and high levels of ATP-binding cassette drug pumps [26]. This could explain why local recurrence is so common after therapy, and unfortunately no effective treatment currently exists for metastatic breast cancer.

1.2.3 Breast CSCs (BCSCs)

Breast cancer stem cells have been identified based on the expression of cell surface markers. In an early study, Al-Hajj M. et al (2003) [27] isolated a cell population from breast cancer xenografts, able to form tumours in immunodeficient mice. This population, which only comprised 1-35% of the total cell population, expressed the markers $CD44^{+}/CD24^{-/low}/lin^{-}$. In contrast to tens of thousands of cells with alternative profiles, as few as 100 cells with this profile formed tumours in NOD/SCID mice and they were able to reinitiate tumours in these mice and retained this ability; each time cells within this population were passaged they were able to form new tumours containing additional $CD44^{+}CD24^{-/low}$ cells. Subsequent studies showed that the $CD44^{+}/CD24^{-/Low}$ phenotype was characterised by radiation-resistance [28] and enhanced invasive properties [29]; it was enriched in basal-like cell lines [29-31] and BRCA1 related cancers [32]. Aldehyde dehydrogenase 1 (ALDH1) has also been used to enrich for BCSCs; cells positive for ALDH1 have increased tumorigenicity, and patients with increased ALDH1 expression are resistant to treatment and have poor outcomes [33]. The main role of ALDHs is to detoxify aldehydes, that would otherwise be cytotoxic, through their oxidation to carboxylic acids (reviewed in [34]). Furthermore, they have other functions such as ester hydrolysis and production of retinoic acid. The exact role of ALDH1 in CSCs still remains unknown, but it is thought to have a role in early differentiation of stem cells, through its role in oxidizing retinol to retinoic acid. CSC markers are also used in combination; for example, the $CD44^{+}/CD24^{-/low}/lin^{-}$ ALDEFLUOR-positive population was highly tumorigenic and could initiate tumours in immunodeficient mice. BCSCs are also enriched for the EpCAM (epithelial cell adhesion molecule, ESA) [30]. The

$CD44^+/CD24^-/ESA^+$ phenotype in basal cancers represents cells capable of initiating tumours in immunodeficient mice.

Emerging research suggests that BCSCs can be generated through the process of epithelial-mesenchymal transition (EMT) [35-37].

1.3 EPITHELIAL-MESENCHYMAL TRANSITION

1.3.1 What is EMT?

EMT is a complex biological process through which polarized epithelial cells are subjected to various changes, enabling them to acquire a mesenchymal cell phenotype [38]. Hallmarks of EMT include disruption of cell-cell adherence, loss of apico-basal polarity, extracellular matrix (ECM) remodelling, enhanced invasiveness and increased migratory capacity. At the molecular level, loss of expression of the cell adhesion molecule, E-cadherin, and upregulation of mesenchymal markers such as N-cadherin, Vimentin and Fibronectin are observed. During EMT, cells acquire a spindle-like morphology, resistance to apoptosis and stem-like traits. This process is, however, reversible, as cells can re-differentiate into epithelial cells through mesenchymal-epithelial transition (MET), or into other cell types, after migration to distant sites [39].

1.3.2 Types of EMT

EMT can occur in three different biological settings that carry very distinct functional consequences and is therefore classified into three types based on the biological context under which it occurs [40].

1.3.2.1 Type 1

Most adult tissues and organs arise from several rounds of EMT and MET referred to as primary, secondary and tertiary EMT [41]. These type 1-EMTs are responsible for the transition of epithelial cells to motile mesenchymal cells during implantation, embryo formation, gastrulation and neural crest migration. Subsequently, these cells generate

secondary epithelia in mesodermal and endodermal organs via MET. At the biochemical level, EMT associated with gastrulation is regulated by the Wnt signalling, with the TFs Sox, Snail, FoxD3 being involved in EMT during embryonic development. In addition, signalling pathways mediated by Wnts, FGFs, BMPs, c-Myb and Msx-1 induce EMT during neural crest formation.

132.2.2 Type 2

Type 2 EMT is associated with organ fibrosis, wound healing and tissue regeneration [40]. Organ fibrosis occurs in a variety of fibrotic tissues, mediated by accumulation of myofibroblasts that secrete a number of inflammatory signals (collagen, laminins, elastin and tenascins) which are in turn deposited as fibres and eventually lead to organ failure. It has been shown that type 2 EMT generates myofibroblasts [42]. In fact, EMT-induced organ fibrosis has been demonstrated in liver, lung, kidney and intestine. Similar processes to EMT occur as a physiological response to injury and tissue regeneration. For example, keratinocytes at the border of the wound recapitulate part of the EMT process during wound healing [42]. Furthermore, the ovarian surface epithelium undergoes an EMT-like process during postovulatory wound healing in each menstrual cycle.

1.3.2.3 Type 3

Type 3 EMT is associated with cancer progression and metastasis [40]. Many *in vitro* and *in vivo* studies have showed that cancer cells acquire a mesenchymal phenotype and express mesenchymal markers [42]. It is now clear that EMT is responsible for the acquisition of cell invasiveness and metastatic events. The invasive step involves epithelial cells that lose their polarity and detach from the basement membrane and undergo EMT (Figure 2). Through EMT, epithelial cancer cells at the invasive border of primary tumours become motile and invade and metastasise though the circulation, and finally generate metastatic lesions through MET to distant tissues or organs [40]. Even though it is now widely accepted that EMT is involved in metastatic events, its contribution in other events may also be highly relevant to cancer progression. EMT

confers resistance to cell death and senescence, immunosuppression, resistance to therapy and finally it generates cells with stem cell properties [40, 41].

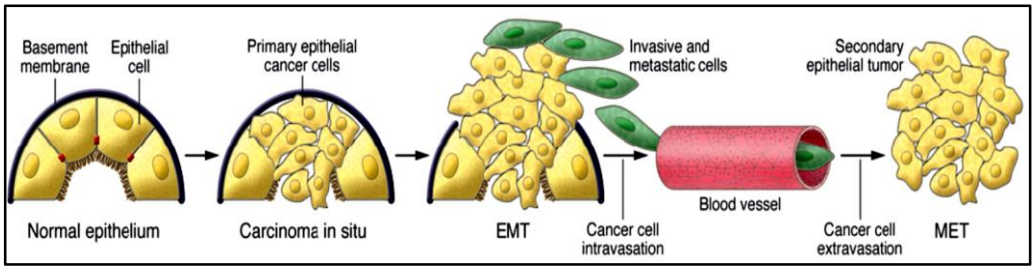


Figure 2. The contribution of EMT in cancer progression and metastasis. Progression from normal epithelium to invasive carcinoma goes through several stages. The invasive step involves epithelial cells losing their cell polarity and detaching from the basement membrane, where cell-ECM and signalling networks are being altered. The next step involves EMT (and other processes such as angiogenesis) that facilitate the tumour growth. Through EMT cells enter the circulation, and reach distant sites, where a secondary tumour might form, which may involve MET.

Picture reprinted (permission granted) from:
The basics of epithelial-mesenchymal transition
Raghu Kalluri, Robert A. Weinberg
Published June 1, 2009
Citation Information: *J Clin Invest.* 2009;119(6):1420-1428. doi:10.1172/JCI39104.

1.3.3 Molecular mechanisms of EMT/EMT biomarkers

Complex networks are engaged in order to facilitate EMT, including activation of transcription factors, changes in the expression of microRNAs, re-organisation and expression of cytoskeletal proteins, production of ECM-degrading enzymes and expression of cell-surface proteins [40-46]. The first step of EMT involves the disassembly of epithelial cell-cell contacts (including tight, adherent and gap functions) and the loss of cell polarity. During this process, epithelial genes are repressed and mesenchymal genes are activated. A hallmark of EMT is the downregulation of E-cadherin, which is involved in the stabilisation of adherens junctions. The downregulation of E-cadherin is concomitant with the increased expression of the mesenchymal cadherin, N-cadherin. This results in ‘cadherin switch’, which promotes invasion and migration.

The changes in the expression of the epithelial and mesenchymal genes are regulated through the activation of master transcription factors (TFs). The main TFs involved are SNAIL, TWIST and ZEB factors, which repress various epithelial genes and activate several mesenchymal genes (Figure 2). The involvement of these TFs is cell- and tissue-specific, but often they cooperate with each other. The activity of the TFs involved in EMT, and, therefore, the EMT progression is regulated by different signalling pathways that are interlinked to facilitate a full EMT response (Figure 3). For example, the RAS-RAF-MEK-ERK MAPK signalling cascade contributes to EMT, by upregulating TFs involved in cell motility and invasion. Another example is the Wnt signalling pathway; Wnt activates β-catenin which in turns activates several EMT-inducers including Slug, Twist and Goosecoid. Wnt and tyrosine kinase receptors (RTKs) can also regulate Snail through GSK3b. Similarly hypoxia, TGF-β, and Notch signalling pathway regulate a number of transcription factors that induce EMT through transcriptional regulation of E-cadherin, such as Twist, ZEB1, ZEB2, Snail, TCF3, FOXC1, and FOXC2. Coordination of EMT programs also results in upregulation of the expression of extracellular matrix (ECM) proteins, proteases, and other ECM remodelling enzymes. Furthermore, there is a crosstalk among the transcription factors that regulate EMT and their miRNAs level. For example, increased expression of RBFOX2 during EMT, promotes differential splicing of different transcripts (e.g. FGFR2), whose expression plays an important role in EMT promotion. Another example is the increased expression of miR-200 during EMT, which results in increasing the levels of the ZEB TFs.

The different factors involved in EMT, are used in the literature as biomarkers, to define cells undergoing EMT. The most common biomarkers used to define/detect EMT are listed in Figure 5. As shown, some of them are common between the three different types of EMT, while others are restricted to specific types.

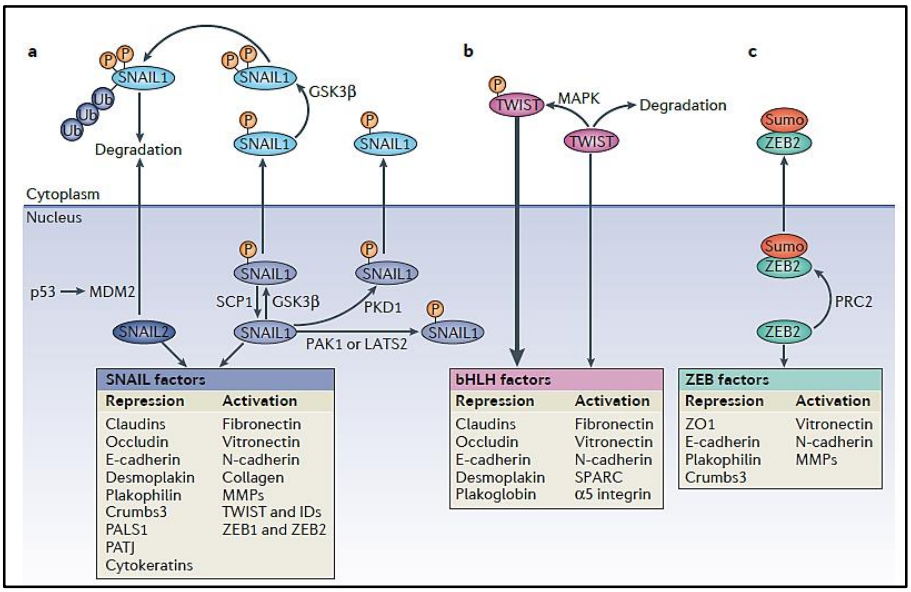


Figure 3. Main EMT transcription factors and their downstream targets. SNAIL, TWIST and ZEB are the main TFs involved in EMT initiation and progression, through repression of epithelial markers such as E-cadherin and activation of mesenchymal markers such as Vimentin.

Picture reprinted (Licence number 3558171111103) from:
Molecular mechanisms of epithelial-mesenchymal transition
Samy Lamouille, Jian Xu, Rik Derynck
Published Feb 21, 2014
Citation information: Nature Reviews Molecular Cell Biology **15**, 178–196 (2014) doi:10.1038/nrm3758

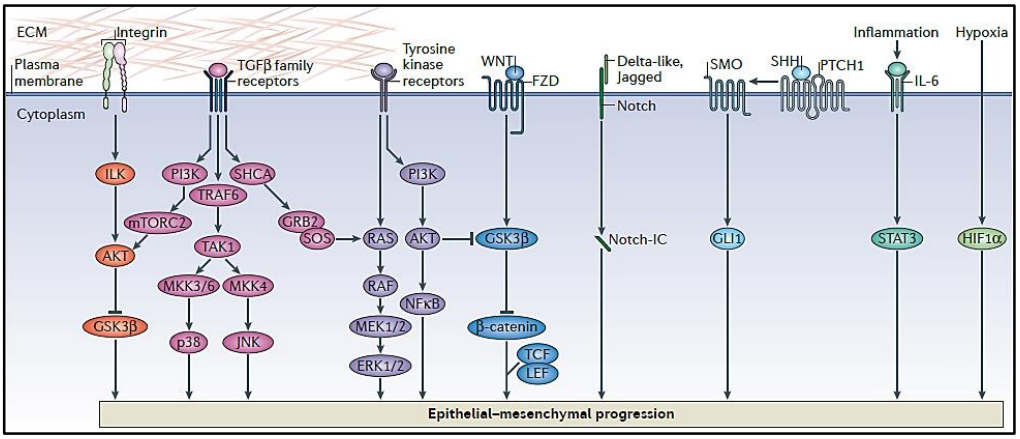


Figure 4. Signalling pathways involved in EMT progression. The TGFβ, Ras-Raf-MEK-ERK, Wnt and Notch signalling pathways, as well as inflammation and hypoxia contribute to EMT progression, through regulation of different TFs.

Picture reprinted (Licence number 3558171111103) from:
Molecular mechanisms of epithelial-mesenchymal transition
Samy Lamouille, Jian Xu, Rik Derynck
Published Feb 21, 2014
Citation information: Nature Reviews Molecular Cell Biology **15**, 178–196 (2014) doi:10.1038/nrm3758

Acquired markers		Attenuated markers	
Name	**EMT type**	**Name**	**EMT type**
Cell-surface proteins			
N-cadherin	1, 2, 3	E-cadherin	1, 2, 3
OB-cadherin	3	ZO-1	1, 2, 3
α5β1 integrin	1, 3		
αVβ6 integrin	1, 3		
Syndecan-1	1, 3		
Cytoskeletal markers			
FSP1	1, 2, 3	Cytokeratin	1, 2, 3
α-SMA	2, 3		
Vimentin	1, 2		
β-Catenin	1, 2, 3		
ECM proteins			
α1(I) collagen	1, 3	α1(IV) collagen	1, 2, 3
α1(III) collagen	1, 3	Laminin 1	1, 2, 3
Fibronectin	1, 2		
Laminin 5	1, 2		
Transcription factors			
Snail1 (Snail)	1, 2, 3		
Snail2 (Slug)	1, 2, 3		
ZEB1	1, 2, 3		
CBF-A/KAP-1 complex	2, 3		
Twist	1, 2, 3		
LEF-1	1, 2, 3		
Ets-1	1, 2, 3		
FOXC2	1, 2		
Goosecoid	1, 2		
MicroRNAs			
miR10b	2	Mir-200 family	2, 3
miR-21	2, 3		

Figure 5. Biomarkers for EMT. The table lists the different factors involved in the different types of EMT that are commonly used in the literature to define/detect cells undergoing EMT.

Picture adapted from:
Biomarkers for epithelial-mesenchymal transitions
Michael Zeisberg and Eric G. Neilson
Published June 1, 2009
Citation information: J Clin Invest. **119(6)**, 1429-1437–196 (2009)
doi: 10.1172/JCI36183

1.3.4 EMT in BC

In vitro and *in vivo* studies have provided much evidence of a role for EMT in breast cancer progression and metastasis, resistance to therapy and local recurrence. These include the expression of EMT-related markers, association of EMT markers with breast cancer histological grade, microRNA regulation of EMT in breast cancer, and the generation of breast cancer stem cells through EMT [45].

EMT and its associated factors correlate with histological grade of breast cancer. Specifically, microarray gene expression profiles of breast cancer patients have demonstrated that up-regulation of genes involved in EMT is associated with poorly differentiated tumours [47].

Changes in the histopathology of breast cancers associated with EMT correlate with high tumour grade, increased mitosis, and oestrogen/progesterone receptor negativity [48]. Moreover, high levels of the EMT-inducer Twist in patient samples associate with high-grade invasive carcinoma and chromosomal instability that promotes EMT and results in aggressive phenotype [49]. Furthermore, elevated levels of the EMT-inducer Snail correlate with higher risk of relapse and local recurrence in mouse models [50]. Other EMT markers (*e.g.* down-regulation of E-cadherin, up-regulation of N-cadherin *etc.*) are associated with basal-like cancers [51].

Association of expression of EMT markers and aggressive phenotypes have been demonstrated in breast cancer. Basal-like breast cancer tumours show high expression of EMT-related genes and EMT-inducers, including high expression of N-cadherin, Vimentin, ACTA2, laminin, fascin, and FOXC2 and low expression of E-cadherin [51, 52]. Such tumours have increased ability to metastasise to other distant sites and are associated with poor prognosis.

Furthermore regulation of EMT by miRNAs has been observed in BC [45]. For example, Inhibition of the miR-200 and miR-250 family induces EMT in BC through upregulation of ZEB1 and ZEB2 [53]. In addition, Twist1 regulates miR-10b expression which is found to be expressed in

metastatic breast cancer cells and correlate with tumour progression [54]. Also, ectopic overexpression of mi-RNA74a in BC cell lines promoted EMT and metastasis through activation of the Wnt signalling pathway [55]. Furthermore, mi-RNA9 in BC has been shown to be associated with EMT, the stem cell phenotype and cancer progression [56].

1.3.5 The link between BCSCs and EMT

Emerging research suggests that there is a direct link between EMT and CSCs [35-37, 51, 52, 57-59]. EMT plays a critical role in cancer metastasis and local recurrence, which is tightly linked with the biology of CSCs. Furthermore, a link between EMT and stemness 'attributes' have been observed in different studies. EMT and BCSCs are major areas of interest in breast cancer research. Most studies follow a common procedure: EMT is induced in human breast cancer cell lines (e.g. by over-expression of Twist, Snail or TGF-β), cells are isolated based on co-expression of EMT and BCSC markers, and the tumorigenic and metastatic capacities of these cells are investigated using common assays (e.g. mammosphere assay) and animal models.

Mani, S.A., et al. (2008) [52] provided the first evidence linking EMT to BCSCs by showing that induction of EMT in immortalized human mammary epithelial cells (HMLEs) by ectopic expression of Twist or Snail results in the acquisition of a mesenchymal phenotype and in the expression of stem cell markers. These cells were able to form mammospheres, soft agar colonies, and tumours more efficiently, compared to cells that have not undergone EMT. This was also supported by other groups [37, 57] and in mouse models [35]. Furthermore, EMT and stem cell markers were found to be associated with basal-like cancers which are more prone to metastasise [51]. Other evidence came from the similar miRNA expression profile in those cells undergoing EMT and BCSCs [36, 58]. Other interesting data showed that the $CD44^{+}/CD24^{-/low}$ breast cancer cells showed a 'claudin-low' molecular subtype which is characterised by the expression of EMT- related

markers [59]. In addition, a recent study has shown that CD146, an EMT transcriptional inducer, generates cells with CSC properties and is associated with poor outcome, high tumour stage, and triple-negative breast cancers [60].

When studying EMT and BCSCs, the role of the tumour microenvironment should be taken into consideration, as both are regulated by its components.

1.4 The influence of the tumour microenvironment (TME) in cancer progression, EMT and BCSCs

1.4.1 The TME

In the TME there exists a complex interplay between the developing tumor and the host, which is thought to be mediated by three systems: angiogenesis/ vasculogenesis, cell-cell and cell-matrix interactions, and host immune response to tumor [61]. It is composed of multiple cell types (immune cells, inflammatory cells, stromal cells etc.) that all play a vital role in tumour growth, angiogenesis, invasion and metastasis and drug resistance (reviewed in [62]) (Figure 5). Similarly, the breast tumour microenvironment is comprised of a variety of cells that secrete a number of cytokines and growth factors that activate signaling pathways, promoting tumour progression and metastasis [63]. Examples include the cancer-associated fibroblasts (CAFs) and other stromal cells (e.g. macrophages) that can promote tumour onset and progression in different ways (reviewed in [64]). One of these ways is through promotion of EMT.

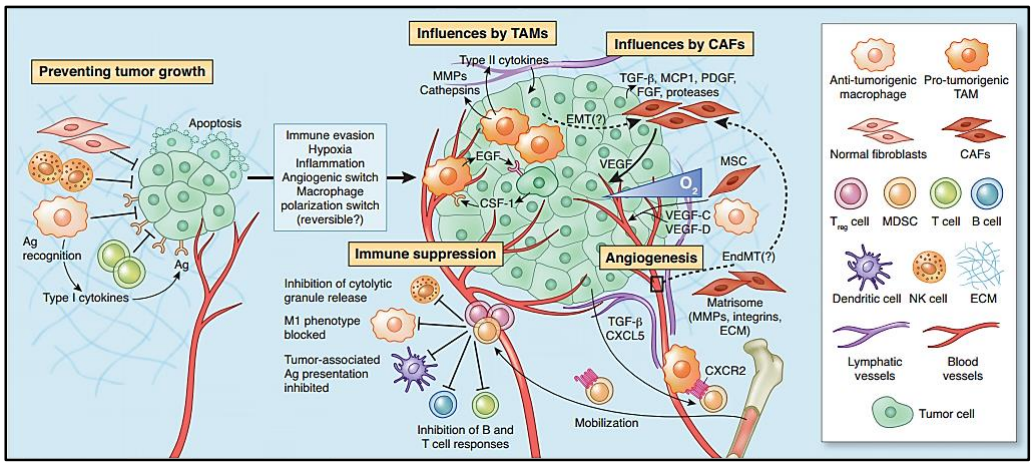


Figure 6. The tumour microenvironment. The TME is composed of multiple cells types that support tumour progression. (Picture reprinted from DF, Q. and J. JA, 2013).

Picture reprinted (Licence number 3558170750281) from:
Microenvironmental regulation of tumor progression and metastasis
Daniela F Quail, Johanna A Joyce
Published Nov 7, 2013
Published June 1, 2009
Citation Information: Nature Medicine 19, 1423–1437(2013)doi:10.1038/nm.3394

1.4.2 Influence on EMT and BCSCs

During EMT, there is a communication between the extracellular matrix, the immune cells and fibroblasts, and interactions with host components influence cancer progression [65]. Inflammatory cells, hypoxia and stem cells can induce the EMT process [66].

The inflammatory cells within the tumour microenvironment are fundamental for invasion and metastasis [67]. Tumour Associated Macrophages (TMAs) and their released cytokines (e.g. TNF-α, IL1) are key mediators of all steps of these processes. For example, EMT was induced in MCF-7 cells treated with TNFα, as demonstrated by

decreased expression of E-cadherin, increased expression of Vimentin and Snail, and increased migratory capacities [68].

Another factor in the induction of EMT is hypoxia (the limited availability of oxygen and nutrients in the TME as a result of tumour growth) [66]. Hypoxia-induced EMT has been demonstrated in BC [69, 70]. Moreover, cancer cells can acquire stem-like properties under hypoxic conditions. It was demonstrated that hypoxia-selected subpopulation from human breast cancer lines could initiate tumours in immunodeficient mice and showed both stem-like and EMT phenotypes [71].

Another possible inducer of EMT within the tumour microenvironment is the mesenchymal stem cells (MSCs). MSCs are recruited from the bone marrow to the tumour stroma; a process that is mediated by different chemokines, growth factors, and receptors found in the tumour microenvironment such as CCL2, EGF, CCL5, CXCL8 and CD44 [10]. Once in tumour sites, they can promote tumour growth by, for example, acting as precursors to CAFs, which in turn secrete a variety of growth factors, inflammatory cytokines and proteinases, all of which contribute to tumour progression. In a xenograft model, it has been shown that bone marrow-derived MSCs can cause cancer cells to metastasis via secretion of CCL5, which acts on the cancer cells to enhance their motility, invasion and metastasis [72]. Furthermore, there is growing evidence that MSCs stimulate EMT in the breast tumour microenvironment and therefore may facilitate tumour growth and metastasis; co-culturing MSCs with BC cells resulted in phenotypical changes of the tumour cells associated with EMT and constant downregulation of E-cadherin [73].

Like EMT, BCSCs are also regulated by the players of the tumour microenvironment, through complex cytokine and growth factor networks [63]. MSCs, CAFs and TAMs are recruited to the tumour microenvironment and secrete IL-6, IL-8 and CXCL7, which activate the STAT3/NF-κB signalling pathway leading to self-renewal of BCSCs. MSC/BCSC interactions have been confirmed to exist in breast cancer patients by immunohistochemical analysis of biopsies [74] and serum levels of IL-1 and IL-8 have been associated with poor clinical outcome

of breast cancer patients [75, 76]. CAFs can also promote proliferation of tumour cells through secretion of SDF-1, the level of which have been associated with poor survival in breast cancer patients [77, 78].

The aforementioned are a few examples demonstrating the important role of the TME in cancer progression and metastasis, including the effect on EMT and BCSCs. Therefore, it is important that pre-clinical models of BC incorporate key elements of the TME in order to recapitulate human tumours as closely as possible, for patient-relevant drugs to be developed.

1.5 Studying/modelling BC and EMT

1.5.1 Cell lines

Human BC cell lines *in vitro* and as xenografts in mice have been extensively used to study/model BC [79]. Human BC cell lines have been used to study the genes and signalling pathways that regulate proliferation, apoptosis and migration during breast cancer progression [80]. They are easily propagated, relatively tractable to genetic manipulations, give reproducible and quantifiable results, can be grown as xenografts, and are more relevant to human disease compared to rodent cells. However, no single cell line is truly representative of human tumours [81]. Furthermore, many of the cell lines have been derived from late-stage tumours and pleural effusion, and, thus, they may be representative of the most malignant phenotype that can be cultured [79]. A major limitation of human breast cancer cell lines is that they are cultured on plastic, and the conditions used to propagate these cells create an environment that differs from the breast microenvironment. When grown in a 2D culture, they do not allow the study of tumour-stromal cell interaction. Methods of *in vitro* cell culture that mimic components found in human cancers have been developed, including 3D

cultures and co-cultures with stromal cells (e.g. human MSCs) [73, 82-84]. Furthermore, growing breast cancer cell lines as xenografts enables investigation of cancer in the *in vivo* environment, which includes the interactions between the developing tumour and stromal cells [85].

1.5.2 Mouse models

Xenografts have given valuable insights about the role of the microenvironment, genetic alterations that contribute to tumour progression, and the events involved in metastasis [79]. For example, a xenograft model has been used to identify genes involved in BC progression and metastasis, which can have therapeutic implications and prognostic significance [86]. Briefly, a poorly metastatic and a highly metastatic BC cell line were injected in the MFP of immunodeficient mice and the tumours and metastases formed were stained for various proteins (e.g. bFGF). This, together with cDNA microarray analysis aided the identification of genes and proteins involved in BC metastasis. Xenograft models have several advantages including (a) they use human rather than mouse cell lines, which makes them more suitable for patient-relevant drug testing (b) the tumours are formed quickly, are easily observable and can be followed, and (c) they are easy to use [87, 88]. A principal restriction is that cell lines transplanted may not be representative of the breast cancers observed in the clinic. Furthermore, xenografts are established in immunocompromised mice, and the absence of an immune system markedly affects tumour progression and metastasis [89]. Also, when the tumour cells are injected subcutaneously into the mouse, this microenvironment may change the growth and metastatic potential of the cells. While orthotopic transplantation of cells into the mammary fat pad should provide a more favourable environment, again there are differences between the mouse and human mammary stroma [90]. Finally, even though orthotopic models of BC metastasis to liver, brain, bone and lymph nodes have been reported [91-95], most models are poorly metastatic or they tend to metastasise only to the lungs [96, 97]. In contrary to xenografts from BC cell lines, patient-derived xenografts (PDXs) retain the morphology, heterogeneity and genomic

profile from the original patient and they have been proved to be more effective for predicting patient-relevant drug response (reviewed in [98]). But again, the PDXs have limitations, including the lack of an intact immune system in the host mice and the difference between the human and mouse microenvironment. For this reason, efforts have been made to humanize the BC xenograft models, by for example adding human stromal fibroblasts [99] or through implantation of human breast epithelial organoids [100].

Others, have used syngeneic models (mouse tumour growing in mouse strain) to study BC, as they take into account anti-tumour immune response, cancer-stromal interactions, and maintain the mammary microenvironment, and have been shown to produce more efficient cancer progression and metastasis than xenografts [101, 102]. On the other hand, the tumour cells in syngeneic models are rodent, not human, so the question is whether the data from studies using such models can translate to humans.

1.5.3 Imaging

Tumour monitoring/ measurement of xenograft tumours is usually being performed by a calliper; a quick, non-invasive and cheap way [103, 104]. Other than this “gold standard” technique, a lot of other methods have been developed over the years to measure tumour including ultra sound [105], CT scan [106] and MRI [107]. Furthermore, techniques have been developed that allow live imaging/tracking of single cells, which are based on optical (fluorescent) reporters, including intravital microscopy, multiphoton microscopy and implantation of tissue window devices in the mice [108].

In order to more specifically study different pathways and genes, involved in cancer progression and metastasis, a variety of reporter systems have been developed, which rely on the expression of a reporter-encoding gene driven by the promoter of interest (reviewed in [109]). The reporter activity, which is most commonly detected by fluorescent or bioluminescent imaging, can be then correlated with promoter activation.

One main advantage of fluorescent reporters is that the detection results from the intrinsic chromophore structure, which allows continuous monitoring of gene expression on a cell-by-cell basis. A main advantage of bioluminescent imaging is the high sensitivity. On the other hand, both methods have drawbacks. In the case of fluorescent reporters, most fluorescent proteins have slow turnover rates, which cause the signal to accumulate over time, and therefore they are not ideal for studying fast activation/inactivation of promoters. A main disadvantage of the bioluminescent reporters, is that the signal detected is an indication of the total quantity of the reporter expressed by the whole cell population, which means that it is not possible to know whether only a subset of cells are expressing the reporter. Since reporters of this kind are based on promoter activity/gene expression, it is important that gene regulation is considered. Gene regulation is a biological process essential for all living organisms, which is achieved through interaction of regulatory elements with DNA sequences within the genes that they regulate [110]. It is controlled by various mechanisms including transcription, post-transcriptional modification and processing, translation, transcript stability and localization. Therefore, when using an optical reporter to detect/image changes in gene expression, key players of gene regulation such as TF binding sites within the gene promoter and other regulatory elements (e.g. enhancer elements) should be considered.

As discussed here, EMT is associated with BC progression, enrichment for therapy-resistant CSCs, and metastasis, for which there is currently no effective therapy. Therefore, there is a need for drugs to be developed to target those cells that are undergoing EMT, in order to prevent them from metastasising. Drug screening is most often performed in mouse models. Despite the fact that many BC metastasis models have been reported, as discussed previously, no model that detects EMT and/or early stages of metastasis has been reported. Therefore, in order to target this process, there is a need for improved models that will allow studying EMT in real-time in *in vivo* models.

1.6 Hypothesis & Aims

It was hypothesised that the process of EMT and/or early metastasis can be visualised/ tracked *in vivo* in real-time by the use of the appropriate optical reporters.

The aim of the research within this thesis was to produce effective EMT reporters using the expression vector pLVX (Clontech), and validate them in stably transduced basal and luminal BC cell-lines, both *in vitro* and *in vivo.* The reporter(s) will allow a real-time read-out of EMT in an orthotopic mouse model, which could be eventually used to test novel drugs developed to target the process of EMT and prevent invasion and metastasis, where the current treatment is not effective.

1.7 Objectives

The first objective was to develop inducible bioluminescent EMT reporters, consisting of an inducible firefly luciferase driven by the promoters of EMT genes (S100A4 or E-cadherin or N-cadherin).

The second objective was to assess the activity of the reporters in *in vitro* models i.e. their ability to detect changes in the genes was assessed.

The final objective was to develop *in vivo* models of EMT and/or metastasis and to validate the activity of the reporters in these models.

CHAPTER 2: Materials and Methods

2.1 Molecular cloning

Kits for molecular cloning were purchased from Sigma Aldrich (Poole, UK) or Qiagen (Manchester, UK). The restriction enzymes, gel loading dye and DNA ladder were purchased from New England Biolabs (Hitchin, UK). The TOPO cloning vector and bacteria were purchased from Invitrogen (Paisley, UK). Other manufacturers are as stated.

2.1.1 Plasmids

Different plasmid vectors have been used for amplifying the gene of interest, cloning the gene/promoter of interest etc. All the plasmids used, including relevant figures, are described here.

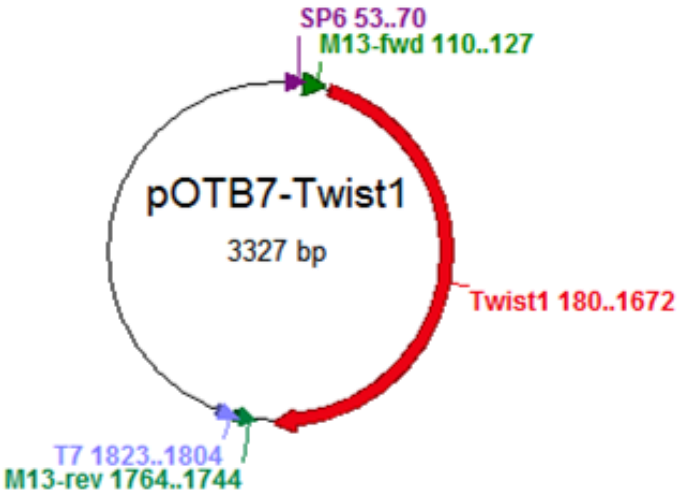


Figure 7. pOTB7-Twist1 plasmid. Location of relevant features: (a) M13 Forward priming site: 110-127, (b) Twist1 (Human TWIST1, exon 1 and 2): 1171-1672, (c) M13 Reverse priming site. Purchased from Geneservice (Nottingham, UK).

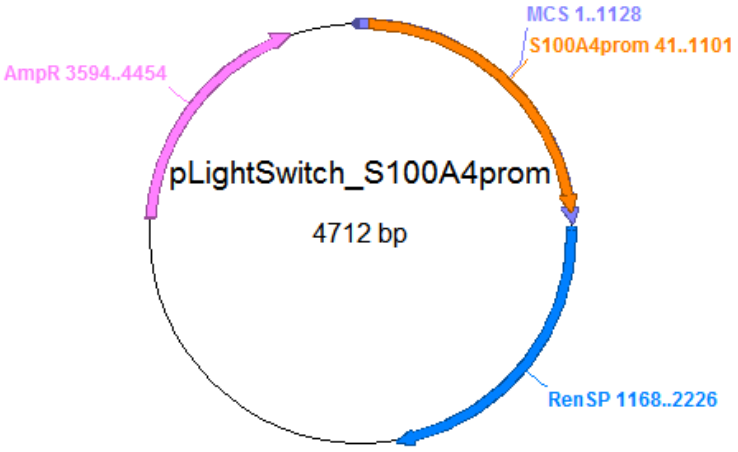


Figure 8. pLightSwitch_S100A4promoter-Renilla. Location of relevant features: (a) Multiple cloning site (MCS): 1-1128, (b) S100A4 promoter (Human): 41-1101, (c) Renilla luciferase: 1168-2226 and (d) Ampicillin resistance gene: 3594-4454. Purchased from SwitchGear Genomics (Belgium, Germany).

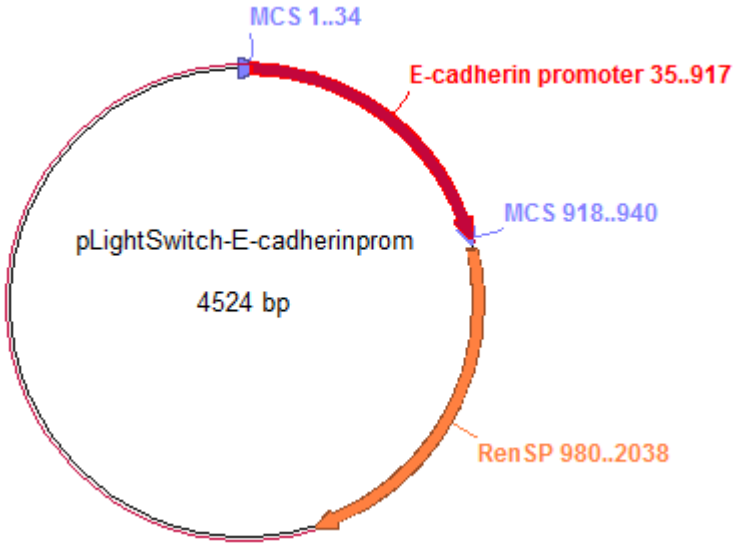


Figure 9. pLightSwitch_E-cadherin promoter-Renilla. Location of relevant features: (a) Multiple cloning site (MCS): 1-940, (b) E-cadherin promoter (Human): 35-917 and (c) Renilla luciferase: 980-2038. Purchased from SwitchGear Genomics.

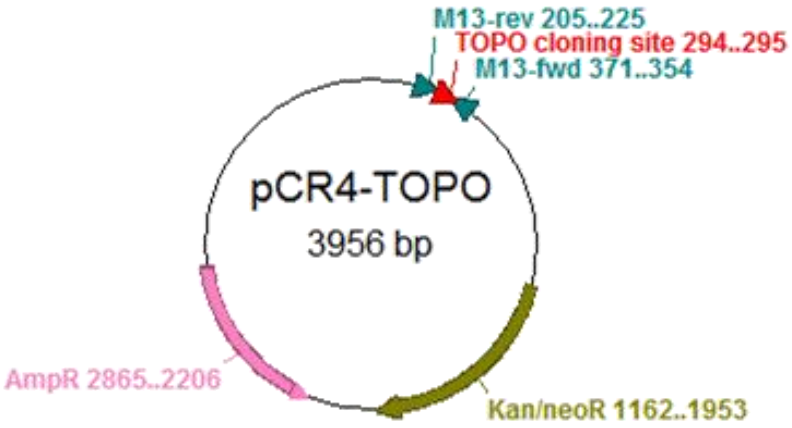


Figure 10. pCR®4-TOPO® plasmid. Location of relevant features: (a) M13 Reverse priming site: 205-225, (b) TOPO® cloning site: 294-295, (c) M13 Forward priming site: 354-371, and (d) AmpR ampicillin resistance gene: 2206-2856. Purchased from Invitrogen.

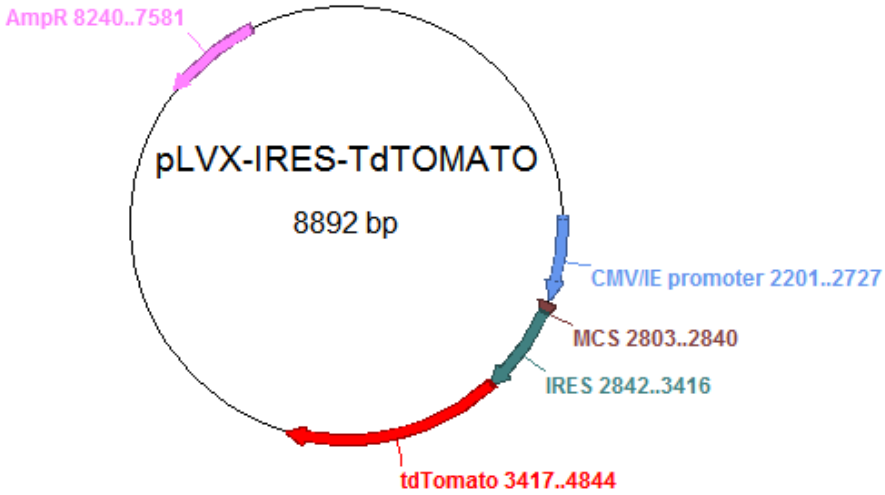


Figure 11. pLVX-IRES-tdTomato. Location of relevant features: (a) CMV IE (human cytomegalovirus immediate early promoter): 2185–2787, (b) MCS: 2803–2840, (c) IRES (encephalomyocarditis virus internal ribosome entry site): 2842–3416, (d) tdTomato: 3417–4844 and (e) AmpR (ampicillin resistance gene): 7578–8574. Purchased from Clontech (Saint-Germain-en-Laye, France).

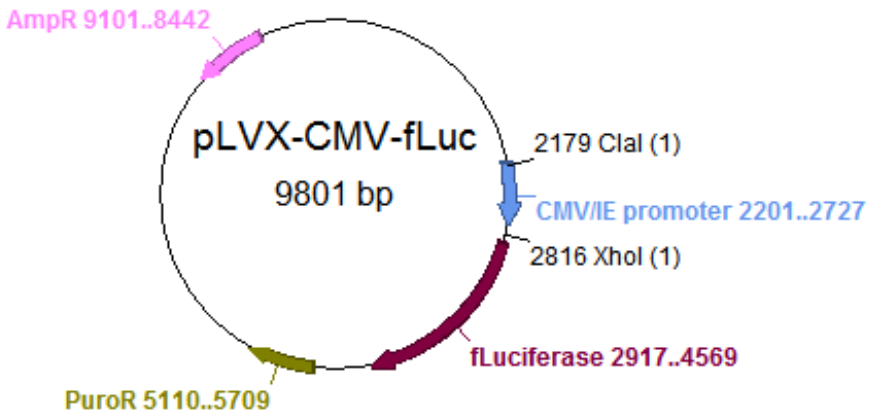


Figure 12. pLVX-CMV-fLuc. It was made by Teresa Coughlan based on 'pLVX-tdTomato'. The CMV_{IE} promoter of this plasmid was replaced with the EMT promoters for the purposes of this study, in order to construct the inducible bioluminescent EMT reporters. Location of relevant features: (a) ClaI restriction site: 2179, (b) CMV: 2201-2727, (c) XhoI restriction site: 2816, (d) Firefly Luciferase: 2917-4569, (e) PuroR: 5110-5709 and (f) AmpR: 8442-9101.

2.1.2 Genomic DNA extraction

DNA was isolated from mammalian cells using the GenElute™ Mammalian Genomic DNA Miniprep Kit, according to manufacturer's instructions. Briefly, mammalian cells were lysed with a salt-containing buffer to ensure denaturation of macromolecules. DNA was bound to the spin column membrane and the remaining lysate was removed by centrifugation. A filtration column was used to remove cell debris. After washing to remove contaminants, the DNA was eluted with 100µl distilled water.

2.1.3 Polymerase Chain Reaction (PCR)

2.1.3.1 Primers

Primers were designed using the OligoPerfect™ Designer tool (Invitrogen, Paisley, UK) and their specificity was checked using the NCBI/ Primer-BLAST tool. Table 1 summarises the sequence of the primers.

2.1.3.2 HotStarTaq® DNA Polymerase (Qiagen)

The reaction mix was composed of 1x PCR buffer, 1x Q solution, dNTP mix (200µM of each), 250nM of forward and reverse primer, 1.25U/rnx of polymerase, 15ng of plasmid DNA and distilled water (up to 50µl). The cycling conditions were set up as described in Table 2.

Table 1. Primers used for the amplification of PCR products

PRIMER	SEQUENCE (5'-3') AND FEATURES
Twist1 Forward	EcoRI Twist 1 (exon1) CCGG^AATTCGAGGTATAAGAGCCTCCAAGTCTG
Twist1 Reverse	BamHI Twist 1 (exon 2) CGCG^GATCCCCCCTCAGAGGAAGGATGA
S100A4 promoter Forward	ClaI S100A4 promoter CCAT^CGATAACTCCCCTATTT
S100A4 promoter Reverse	XhoI S100A4 promoter GATC^TCGAGCAACACAACTCACCAAT
E-cadherin promoter Forward	ClaI E-cadherin promoter GCGTACAT^CGATAAATTAGGCTGCTAGCTCAG
E-cadherin promoter Reverse	XhoI E-cadherin promoter GCAAATCTCGAGAATGCGTCCCTCGCA
N-cadherin promoter Forward	ClaI N-cadherin promoter AAATTTAT^CGATGGCTCTAGGGGCTGGATT
N-cadherin promoter Reverse	XhoI N-cadherin promoter GGTTGGC^TCGAGTGTTGTTCGGGCGTGTAA

Table 2. Cycling conditions for PCR amplification using the HotStarTaq DNA polymerase.

		Additional comments
Initial activation step	15 min 95°C	HotStarTaq DNA polymerase was activated by this heating step
3-step Cycling		
Denaturation	30 sec 94°C	
Annealing	30 sec Gradient: 50°C to 68°C (Twist1) 47°C to 60°C (S100A4 promoter) 50°C to 70°C (E-cadherin promoter)	The gradient covered a variety of temperatures below and above Tm of primers
Extension	1 min 72°C	For PCR products longer than 1 kb, an extension time of approximately 1 min per kb DNA was used
Number of cycles	30	
Final extension	10 min 72°C	

2.1.3.4 Terra™ PCR Direct Polymerase Mix (Clontech)

The reaction mix was composed of 1x Terra PCR Direct Buffer (with Mg2+ and dNTP, 0.3µM of the forward and reverse primer, 100ng of template

DNA, 1.25 of the Terra PCR Direct Polymerase Mix and distilled water (up to 50µl). The cycling conditions were set up as follows:

3-Step PCR (for amplification of standard targets < 2 kb):

98°C 2 min

98°C 10 sec
50°C-68°C 15 sec
68°C 1 min/kb
} 30 cycles

2.1.4 Agarose gel electrophoresis

Agarose gels were made 0.8% agarose (Eurogentec, Southampton, UK) dissolved in 150ml of 1x Tris/Acetate/EDTA (TAE) electrophoresis buffer. 0.8% is suitable for size ranges of 500bp to 10,002bp, which cover the size of all the PCR products amplified in this study. Gel loading dye (6x) was added to the samples and the 1kb DNA ladder was used as a reference to estimate the size of DNA molecules (Figure 12). The nucleic acid stain SafeView (NBS Biologicals, Huntingdon, UK), was used for the detection of DNA. After loading samples, DNA was separated using a constant voltage of 80V for approximately 1 hour or until the loading dye had reached the end of the gel.

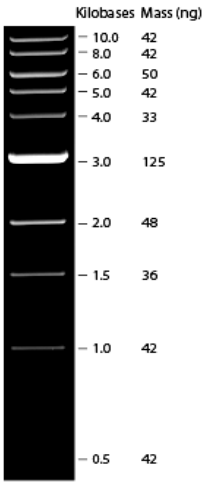


Figure 13. 1kb DNA ladder. The DNA includes fragments range from 0.5-10.0 kilobases (kb).

2.1.5 PCR CLEAN-UP

PCR products were purified using QIAquick PCR purification kit, according to manufacturer's recommendations.

2.1.6 TOPO Cloning (TOPO TA CLONING® FOR SEQUENCING)

The reaction mixture for the TOPO® cloning reaction was composed of 1µl of salt solution, 10ng/µl of the TOPO vector, variable amount of PCR product (or no PCR product for the empty control) and distilled water (up to 6µl). The reaction incubated for 5 minutes at RT and stored at -20˚C. The amount of the PCR product was calculated as follows: ng of PCR product need (**1**:1 Insert: Vector) = **1** x Length of insert in bp x (ng of vector/length of vector in bp)

2.1.7 Transformations

Plasmids were transformed into SUBCLONING EFFICIENCY™ DH5α™ competent cells. The cells were thawed on ice, gently mixed with the pipette tip, and 50µl of cells for each transformation was aliquoted into a 1.5ml microcentrifuge tube. 1µl-5µl (1-10ng) of DNA was added to the cells and mixed gently. For the pUC19 control DNA, 2.5µl (250pg) was added to the cells and mixed. The cells were incubated on ice for 30 minutes. After that, the cells were heat shocked for 20 seconds in a 42°C water bath without shaking, and immediately transferred on ice for 2 minutes to allow cooling down. 950µl of pre-warmed Luria Broth (LB) medium was added to each tube, and cells were incubated at 37°C for 1 hour shaking. Different volumes from each transformation were spread on pre-warmed agar plates containing ampicillin (100µg/ml). Plates were incubated overnight (O/N) at 37°C.

2.1.8 Plasmid minipreps

Plasmids were isolated from bacteria using the GenElute™ plasmid miniprep kit, according to manufacturer's recommendations. Briefly, bacterial cells were harvested via centrifugation, subjected to a modified alkaline-SDS lysis procedure and the DNA adsorbed onto silica in the presence of high salts. Contaminants were then removed by a simple

wash step. Bound DNA was eluted in 30µl water. DNA yield and quality were determined using the NANODROP 2000 Spectophotometer (THERMO SCIENTIFIC, Loughborough, UK). Quality of DNA was determined by the ratio of spectrophotometric absorbance of the sample at 260nm to that of 280nm (A260/A280 ratio). Values above 1.6 were considered acceptable.

2.1.9 Plasmid maxipreps

Plasmids were isolated from bacteria using the QIAGEN Plasmid Maxi Kit, according to manufacturer's recommendations. DNA yield was determined using the NANODROP 2000 Spectophotometer and DNA quality was determined by the A260/A280 ratio. Values above 1.6 were considered acceptable.

2.1.10 Cleavage of DNA with restriction endonucleases

The PCR products and the vectors were digested with restriction endonucleases with compatible ends. Digestions were also performed for selection of clones with the insert incorporated. The enzymes used in this study are the following: EcoRI (G^AATTC), BamHI (G^GATCC), ClaI (AT^CGAT) and XhoI (C^TCGAG). The digestion reactions were composed of 150-200ng of DNA, 1µl of each enzyme, 2µl of the appropriate buffer (EcoRI NEB or NEBuffer 4 + BSA) and distilled water (up to 20µl). The reactions were incubated at 37°C for 2 hours or ON.

2.1.11 Gel extraction (QIAquick gel extraction kit)

After separation on agarose gel, the DNA fragment was visualised using a 312nm UV transilluminator (MODEL TVC-312A) and the relevant band cut out. It was extracted using QIAquick gel extraction kit according to manufacturer's instructions. DNA was eluted in water, and yield was determined using the NANODROP 2000 Spectophotometer.

2.1.12 Ligations

The gel-extracted digested PCR products and vectors were ligated using the DNA ligation kit (Agilent Technologies, Santa Clara, USA). The

following mix was prepared for different insert: vector ratios and incubated at 4°C ON:

	Vector + Insert	Vector only control
Digested vector	50ng	50ng
Digested insert	-	*
10 x Ligase Buffer	1µl	1µl
10 mM rATP (pH 7.5)	1µl	1µl
T4 DNA ligase (4 U/µl)	0.5µl	0.5µl
H_2O	Up to 10µl	Up to 10µl

* Calculated as follows:

ng of PCR product need (**1**:1 Insert: Vector) = **1** x Length of insert in bp x (ng of vector/length of vector in bp)

ng of PCR product need (**3**:1 Insert: Vector) = **3** x Length of insert in bp x (ng of vector/length of vector in bp)

ng of PCR product need (**6**:1 Insert: Vector) = **6** x Length of insert in bp x (ng of vector/length of vector in bp)

2.1.13 DNA sequencing

The DNA sequence of the plasmid constructs was analysed using a 3130xl ABI PRISM Genetic Analyzer by the DNA sequencing facility (University Of Nottingham). The sequencing primers for each construct are listed in Table 3. The primers were either provided by the sequencing facility (M13), or designed for the purposes of this project as explained above (See section 3.1.3.1). The sequence was aligned to

the original sequence of the plasmid construct using the ApE (A plasmid Editor) software Version 1.1

Table 3. Sequencing primers for plasmid constructs. F=Forward, R=Reverse, M=Middle primer

CONSTRUCT	PRIMER	SEQUENCE (5'-3')
pCR4-TOPO-Twist1	M13 F	TGGCCGTCGTTTTACA
	M	GGCGCCCCGCTCTTCTCC
	M13 R	CAGGAAACAGCTATGACCATG
pCR4-TOPO-S100A4prom	M13 F	TGGCCGTCGTTTTACA
	M	CCCCCTAGCTTTTGTGTCAC
	M13 R	CAGGAAACAGCTATGACCATG
pCR4-TOPO-Ncadprom	M13 F	TGGCCGTCGTTTTACA
	M	AGAGCCAGCGAGCCTTTT
	M	CTAGACGGTCGTGGGCCG
	M13 R	CAGGAAACAGCTATGACCATG
pCR4-TOPO-Ecadprom	M13 F	TGGCCGTCGTTTTACA
	M13 R	CAGGAAACAGCTATGACCATG
pLVX-Twist1-IRES-tdTom	F	TATATAAGCAGAGCTCGTTTAG TGAAC
	M	GGCGCCCCGCTCTTCTCC
	R	TATGGTGGAAAATAACATATAG ACAAACG
pLVX-S100A4prom-fLuc	F	GGACCCGACAGGCCCGAA
	M	CCCCCTAGCTTTTGTGTCAC
	R	GCTTCATGGCTTTGTGCAG
pLVX-Twist1-IRES-tdTom	F	TATATAAGCAGAGCTCGTTTAG TGAAC
	M	GGCGCCCCGCTCTTCTCC
	R	TATGGTGGAAAATAACATATAG ACAAACG
pLVX-Ncadprom-fLuc	F	TCAAAATTTTCGGGTTTATTACA GG
	M	AGAGCCAGCGAGCCTTTT
	M	CTAGACGGTCGTGGGCCG
	R	GAACGCTCATCTCGAAGTACTC
pLVX-Ecadprom-fLuc	F	GGGGGGTACAGTGCAGGG
	R	TGGGCCCTTCTTAATGTTTTT

2.2 CELL-BASED STUDIES

The chemicals used were purchased from Sigma Aldrich (Poole, UK), unless otherwise stated. All the NuPAGE® products for western blot were purchased from Invitrogen (Paisley, UK).

2.2.1 Cell culture conditions

The cells used in the study were incubated at 37°C in a humidified 5% CO_2 incubator. They were split in ratios ranging from 1:2 to 1:5. Briefly, they were detached from the flask using sterile trypsin/EDTA solution, they were centrifuged at 415g for 5 minutes, and the pellet was re-suspended in the appropriate medium. Table 4 summarises the source and the growth medium used for each cell line. The lentivirally transduced cell lines were selected weekly with 4µg/ml puromycin in the appropriate growth media. The amount of puromycin for each cell line was previously determined in the lab.

2.2.2 Cell plating

For plating, cells were washed with HBSS, detached with Trypsin/EDTA, collected with medium and pelleted for 5 minutes at 415g, re-suspended in culture medium and counted using Trypan Blue (Life Technologies). For counting 50µl of re-suspended cells were mixed with 50µl of Trypan Blue and counted using a haemocytometer.

2.2.3 Alamar Blue readings

Cell growth was monitored with Alamar Blue® (Life Technologies). Alamar Blue was diluted in TGA media, at 1:10 dilution, and 100µl/well was added. The cells were incubated at 37˚C for 60 minutes exactly and the readings were performed on the FlexStationII^{384} (Molecular Devices) at λex=560nm, λem=588nm, and λcutoff=570. Cell viability was plotted as relative fluorescent units (RFU).

Table 4. Source and growth conditions of cell lines.

CELL LINE	SOURCE	GROWTH MEDIUM
MCF-7	NCI60 or ECACC	Phenol red-free RPMI with 10% Fetal bovine serum (FBS) and 200nM l-glutamine
MDAMB231	NCI60	RPMI with 10% FBS and 200nM l-glutamine
BT549	NCI60	RPMI with 10% FBS, 200nM l-glutamine and 5µg/mL insulin
BrCAFs Age of patient:67 Tumour grade:2 Type: Invasive lobular carcinoma ER/PR and Her2 status: ER8/8, PR8/8, Her2 negative	Kindly provided by Professor Val Speirs (University of Leeds)	DMEM, high glucose, GlutaMAX™ Supplement, GIBCO (Paisley, UK)
Human Mesenchymal Stem Cells-bone marrow (HMSC-bm)	ScienCell (Carlsbad, USA)	Mesenchymal Stem Cell Medium (MSCM), supplemented with 5% FBS and 1% MSC growth supplement

2.2.4 Stable transductions

Lentiviruses were produced in-house using the Lenti-X™ Lentiviral Expression System (Clontech) according to manufacturer's instructions. The quality of the lentivirus stock was assessed using Clontech's Lenti-X GoStix. The cell-lines were transduced with the lentiviral vector at a MOI (multiplicity of infection) of 1 or 3 IFUs (infectious unit)/ cell using the same system, according to manufacturer's instructions.

2.2.5 Bioluminescent readings

Bioluminescent readings of the cells transduced with luciferase-expressing constructs were taken using the Ivis Spectrum from Caliper Life Sciences. D-Luciferin (Caliper Life Sciences) was added to the cells to a final concentration of 150µg/ml and firefly luciferase expression was measured as average radiance (photons/sec/cm2).

2.2.6 Flow cytometry

Flow cytometry is based on the principles of light scattering, light excitation and emission of fluorochromes, in order to obtain specific multi-parameter data from cells. The cells were separated according to their size, granularity and fluorochrome emission.

For each antibody, 50,000 cells were harvested and re-suspended in 50µl PBS 2% BSA. The antibody was added at the appropriate concentrations and mixed. Cells were stained with CD44 (Alexa Fluor® mouse Anti-human CD44 Clone: IM7) only, CD24 (BD Pharmingen™ PE Mouse Anti-Human CD24) only, and with both antibodies (dual staining). Cells were also stained with the isotype controls (Alexa Fluor® 488 Rat IgG2b, κ Isotype and PE Mouse IgG2α, κ Isotype) to the same final concentration as the matched test antibody. Also, unstained cells were used for the analysis. After addition of the antibodies, cells were incubated on ice for 30 minutes in the dark. They were washed with PBS 2% BSA and were fixed with 500µl PBS 3.7% Formaldehyde. At least 10,000 cells per test were analysed using a FC500 flow cytometer (Beckman Coulter) and data was analysed using Weasel Software Version 3.0.2 (The Walter and Eliza Hall Institute of Medical Research, Melbourne Australia).

2.2.7 Western Blot

2.2.7.1 Preparation of protein samples

Cells were washed in cold PBS (pH 7.6). 1ml of PBS was added to the flask and the cells were detached using a cell scraper. The cells were transferred to a 1.5ml micro centrifuge tube and pelleted at 417g for 5 minutes. The cell pellet was washed again in PBS and pelleted. 50-100µl lysis buffer was added to the cells and the cells were mechanically resuspended by pipetting up and down. Lysis buffer was composed of RIPA buffer, 1/100 phosphatase inhibitor cocktail 2 and 3 and 1/50 protease inhibitor. The cells were lysed on ice for 20 minutes. The lysates were centrifuged at 3600g for 5 minutes at 4°C to pellet the cell debris. The protein lysate was transferred to a clean 1.5ml micro centrifuge tube. The cells were stored at -80°C or used immediately. The protein concentration was determined by a Bradford assay (section 2.2.5.6) and 20µg of protein was mixed with 4µl of 4x NuPAGE® LDS Sample Buffer, 2µl of 10x NuPAGE® Sample Reducing Agent and RIPA buffer (up to 16µl). The samples were heated for 5 minutes at 96°C, cooled down on ice and centrifuged to collect any condensation.

2.2.7.2 Preparation of apparatus for electrophoresis

The combs were removed from the gels (NuPAGE® Novex® 4-12% Bis-Tris Protein Gels, 1.0 mm, 12 well) and the wells were rinsed with 1 x SDS buffer (NuPAGE® MOPS SDS Running Buffer (20X) from Life Technologies diluted in distilled water). The gels were placed in the electrophoresis tank and 1 X SDS buffer was added to the wells and tank. The protein samples were loaded to the wells. Two different size markers were also added. The markers were the SeeBlue® Plus2 Pre-stained Protein Standard and MagicMark™ XP Western Protein Standard from Life Technologies. 500µl of NuPAGE® Antioxidant was added in to top of the tank and the gel was run at 80V initially and then 200V, until the samples were at the bottom of the gel.

2.2.7.3 Transfer of samples

The gels were removed from the tank and washed and transferred to the iBlot Anode (iBlot® Transfer Stack, nitrocellulose), placed in the iBlot gel transfer device. A soaked blotting paper was added on top and then the iBlot Cathode was placed on top of that. Finally, a sponge was added on the lid of the device and the transfer was performed for 7 minutes.

2.2.7.4 Binding of antibodies

The membranes were transferred to 15ml tubes and they were blocked for 1 hour in 5ml of 3.5% blocking buffer (Membrane blocking agent, GE Healthcare) diluted in TBS with 0.1% Tween (TBS-Tween) for 1 hour at room temperature on the roller. The blocking buffer was replaced with the primary antibody diluted in blocking buffer at the appropriate concentration and incubated O/N in the cold room on the roller, followed by 1 hour incubation at RT on the roller. The blots were washed with the wash solution, 1 x TBS-Tween, (3 x 5 minutes). The wash solution was replaced by the secondary antibody which was incubated for 1 hour at RT on the roller. The blots were washed with TBS-Tween (3 x 5 minutes). The antibodies used are listed in Table 5. B-actin was used as a reference protein.

2.2.7.5 Imaging

The blots were visualised in the ChemiGenius^2 Bio imaging system, after addition of Amersham ECL Prime Western Blotting Detection Reagent (GE Healthcare). Different exposure times were used for each protein as appropriate.

Table 5. Antibodies used for western blot.

Primary antibody	Secondary antibody	MW (kDa)	Dilution
TWIST1 monoclonal antibody (M03), clone 2F8, Mouse, Anti-human, Abnova	Polyclonal Goat Anti-Mouse Immunoglobulins/HRP, Dako (Cambridgeshire, UK), 1:2000	22	1:500
E-Cadherin (24E10) Rabbit mAb, Anti-human, Cell Signalling Technology	Anti-rabbit IgG, HRP-linked Antibody, Cell Signalling Technology (1:1000)	135	1:1000
Vimentin (D21H3) XP® Rabbit mAb, Reacts with human, mouse, rat and monkey, Cell Signalling Technology	Anti-rabbit IgG, HRP-linked Antibody, Cell Signalling Technology (1:1000)	57	1:100
Mouse monoclonal to β-actin (AC-15), abcam (Cambridge, UK)	Polyclonal Goat Anti-Mouse Immunoglobulins/HRP, Dako, 1:2000	40	1:2000

2.2.7.6 Bradford assay

BSA (Albumin from bovine serum) was diluted in distilled water to prepare the following protein standards: 0.25mg/ml, 0.50mg/ml. 0.75mg/ml, 1mg/ml, 1.15mg/ml, 1.25mg/ml and 1.50mg/ml. 5µl of each standard, of the protein and of a negative control (lysis buffer; section 2.2.5.1) was added to a 96 well clear, polystyrene plate (Corning incorporated, CoStar, London, UK) in triplicates. 250µl of Bradford reagent was added to each

well, and the plate was gently mixed and incubated for 10 minutes at RT. The absorbance at 595nm was measured using an illuminometer (MRX Revelation, Thermo Labsystems, Cheshire, UK). The absorbance of the protein standards was used to plot a standard curve. The equation of linear regression was calculated in Graph Pad Prism, and the concentration of the protein samples was determined based on this, after subtracting the absorbance value of the lysis buffer (negative control).

2.2.8 Immunofluorescence

2.2.8.1 Cell plating

Cells were harvested and counted as before (section 2.2.5.6). Coverslips and tweezers were dipped in absolute ethanol for 5-10 minutes before use. After drying, they were placed on the bottom of the plate. 100µl of cells (~400,000 cells/ ml) were carefully added, and the plate was transferred to the incubator. Each well was topped up with normal medium when the cells were attached to the coverslip.

2.2.8.2 Staining

The next day, the medium was carefully removed and the cells were washed twice with PBS. The cells were fixed in 1ml of 3.7% formaldehyde for 30 minutes at RT. The cells were washed twice with PBS. The slides were covered with 1ml of permeabilsation solution (2% BSA, 0.2% Triton X-100 and 0.005% Tween-20 in PBS) for 15 minutes at RT. The permeabilisation solution was removed and the wells and the cells were stained with the primary antibody diluted in 100µl of permeabilisation buffer for 2 hours at RT. A piece of parafilm was added on the coverslips to avoid evaporation. The primary antibody was washed three times with PBS and the cells were incubated with the secondary antibody (conjugated with a fluorochrome) diluted in 100µl of permeabilisation solution for 1 hour in the cold room. After 3 washes with PBS, the coverslips were mounted on poly-L-lysine coated glass slides (Polysine® slides, Thermo Scientific, Loughborough, UK) with Prolong gold

(Prolong® Gold antifade reagent with DAPI, Invitrogen) and left to harden for 24 hours in the cold room. For each primary antibody, an isotype control and a 'secondary only' control were included. The antibodies used are listed in Table 6. The pictures of staining were taken with the Nikon Eclipse Ti-U microscope.

Table 6. Antibodies used for immunofluorescence.

Primary antibody	Secondary antibody	Isotype control	Dilution
E-cadherin, Mouse, Anti-human, DAKO (Clone NCH-38)	Goat anti-mouse Alexa Fluor 488, Invitrogen, used at 1:50 dilution	Mouse IgG1, DAKO, matched to the concentration of the primary antibody	1:25
Vimentin, mouse, anti-human, DAKO, clone V9	Goat anti-mouse Alexa Fluor 488, Invitrogen, used at 1:50 dilution	Mouse IgG1, DAKO, matched to the concentration of the primary antibody	1:50

2.2.9 Gene expression analysis

RNA was isolated from cells, cDNA was synthesised and gene expression levels were assessed using quantitative PCR (qPCR).

2.2.9.1 RNA isolation

RNA was extracted using the TRI-reagent isolation protocol. Harvested cells were re-suspended in 1mL TRI-reagent and allowed to stand at R/T for 5 minutes to ensure complete dissociation of nucleoprotein

complexes. 200µl of chloroform were added per mL of TRI-reagent and samples were vigorously shaken for 15 seconds. After standing at RT for 15 minutes, cells were centrifuged at 3600g for 15 minutes at 4°C. Centrifugation separated the mixture into 3 phases: a red organic phase (containing protein), an interphase (containing DNA), and a colourless upper aqueous phase (containing RNA). The aqueous phase was transferred to a fresh tube containing 500µl isopropanol and mixed gently. The sample was allowed to stand at R/T for 10 minutes and then centrifuged at 3600g for 15 minutes at 4°C. The supernatant was removed and the RNA pellet was washed with 1mL 70% ethanol. After centrifugation at 3600g for 5 minutes at 4°C, the supernatant was removed and the samples were re-centrifuged for an additional minute. Excess ethanol was removed without disturbing the pellet. The RNA pellet was air-dried and re-suspended in 50µl nuclease-free water.

2.2.9.2 cDNA synthesis

cDNA was synthesised from the mRNA isolated from the cells using the SuperScript™ II Reverse Transcriptase (Invitrogen). For each cDNA made, a cDNA negative that contained all the reagents, except the reverse transcriptase and primers, was also prepared. This allowed the degree of genomic contamination in each RNA preparation to be checked. In order to check contamination of any of the reagents used, a cDNA control was included in which a complete cDNA synthesis was carried out using water instead of RNA.

Two 0.5mL tubes were labelled cDNA+ve and cDNA-ve. 1µl of random hexamers (stock diluted 1:3) and 1µl water were added to the cDNA+ve and cDNA-ve respectively. 10µl of RNA was added to the tubes and mixed. The samples were incubated at 70°C for 10 minutes.

The following master mixes were made:

	cDNA+ve (x1)	cDNA-ve (x1)
Water	3µl	3.4µl
5xBuffer	4µl	4µl
0.1M DTT	2µl	2µl
10mM dNTPS	0.6µl	0.6µl
SuperScript	0.4µl	-

10µl of the appropriate reaction mix was added to each RNA sample. The samples were mixed and incubated for 10 minutes at 25°C and 1 hour at 42°C. The enzyme was heat inactivated for 5 minutes at 95°C. The samples were stored at -20°C or used immediately.

2.2.9.3 Quantitative PCR (qPCR)

2.2.9.3.1 The reaction

The qPCR reactions were carried out in 96 well plates (Applied Biosystems) containing 5µl cDNA and 15µl reaction mixture. The reaction mixture contained 10µl of 2X Master Mix (SYBR green I and polymerase, Applied Biosystems), 4 µl nuclease-free water and 1µl primers (forward and reverse primers mixed, usually 5µM each). The reactions were performed in a StepOnePlus Real Time PCR machine (Applied Biosystems) using the following reaction cycle protocol: one holding stage (95˚C for 10 minutes), one cycling stage (40 cycles at 95˚C for 15 sec and 60˚C for 60 sec) and one melt curve stage (95˚C for 15 minutes, 60˚C for 1 minute and 95˚C for 15 minutes). For every sample analysed a negative cDNA control (H_2O instead of cDNA) was run to detect any contamination during the preparation.

2.2.9.3.2 Primers

The sequences of the primers are listed in Table 7.

Table 7. Sequence of primers used in qRT-PCR.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
HPRT	ATTATGCTGAGGATTTGGAAAGGG	GCCTCCCATCTCCTTCATCAC
VEGF	CCAAGTGGTCCCAGGCTGCA	TGGATGGCAGTAGCTGCGCT
TWIST1	CAAGCTGAGCAAGATTCAGACCC	AGACCGAGAAGGCGTAGCTGA
S100A4	GGATGGATGGTGTCCT	CATCTGTCCTTTTCCCCAAGAAG
CDH1	GAACAGCACGTACACAGCCCT	GCAGAAGTGTCCCTGTTCCAG
CDH2	GACGGTTCGCCATCCAGAC	TCGATTGGTTTGACCACGG
VIM	AAAACACCCTGCAATCTTTCAGA	CACTTTGCGTTCAAGGTCAAGAC
SNAI2	CTGCGGCAAGGCGTT	GCAGTGAGGGCAAGAAAAAGG
ZEB1	GAAAGGAAGGGCAAGAAATCCT	TGCATCTGACTCGCATTCATC

2.2.9.3.3 Data analysis

The data obtained from the qPCR were analysed based on the cycle threshold (Ct) value, the minimum number of cycles necessary in order to detect SYBR green fluorescence incorporated into the DNA. The higher the Ct value, the lower the gene expression level. All samples were analysed in triplicates and each Ct value was normalised to the

triplicate Ct values obtained for the housekeeping gene Hypoxanthine-guanine phosphoribosyl transferase (HPRT) expression in the same cDNA sample. Gene expression was expressed as $2^{-\Delta Ct}$ [111] in which $\Delta C_t = C_t$ target gene-C_t reference gene (HPRT).

2.2.10 Cell culture assays

2.2.10.1 Hypoxia

$1x10^4$ cells/ml were seeded in black 24-well plates in a total volume of 1ml, in quadruplicates. Cells were cultured in normal oxygen (normoxia) conditions (see section 2.2.1) or in low oxygen conditions (1%) in a humidified hypoxia chamber ($Invivo_2$ Hypoxia Workstation 400, RUSKINN) for 4 days. At the end of the assay, bioluminescent readings were performed (see sections 2.2.3) and the activity of the reporter was calculated as follows: (Average signal of EMT reporter in hypoxia / Average signal of control vector in hypoxia) / (Average signal of EMT reporter in normoxia / Average signal of control vector in normoxia). Following bioluminescent readings, the medium was aspirated and the cells were collected with 1ml of TRI-reagent (all four wells in 1ml) for subsequent gene expression analysis.

2.2.10.2 Culture with EMT inducing media supplement

$5x10^3$ cells were seeded in 10cm culture dish in 6ml of culture medium with or without 1X (60µl) StemXVivo EMT Inducing Media Supplement (R&D systems, Abingdon, UK). The medium was replenished after 3 days. After 5 days the cells were collected with TRI-reagent and/or were lysed as described before (section 2.2.5).

2.2.10.3 3D cultures

Cultrex® Basement Membrane Extract (BME) was defrosted on ice overnight. Cells were pelleted for 5 minutes at 835g in phenol free, serum free RPMI supplemented with 1% l –glutamine (Tumour growth assay medium; TGA), and resuspended in TGA medium; an equal amount of BME was added and mixed. These steps were performed on ice, to avoid

solidification of BME. In a 96 well plate black with clear bottom, 100µl per well ($2x10^4$ cells/well, 4-6 wells per day) were added with the plate on 37˚C plate warmer, to allow immediate setting of the BME. The plate was placed in the incubator and after about 6 hours (when the BME was completely settled) 50µl of TGA was added on top. Cell growth was monitor with Alamar Blue (section 2.2.3). Cells were collected with TRI-reagent using a blue tip, in which the tip was cut.

2.2.10.4 Co-cultures with MSCs

Approximately 3300 MCF-7 cells per well were mixed with approximately 1750 MSCs (2:1 ratio) or 1750 MCF-7 cells and plated on coverslips in 6-well plates in a total volume of 3ml MCF-7 medium (Table 4). For each condition (MCF-7 only or MCF-7 + MSCs), 5 wells were included for the IF analysis: 1) No staining, 2) Isotype control, 3) Secondary antibody only, 4) E-cadherin and 5) Vimentin. The cells were cultured for 7 days. At the end of the assay, the co-cultures were analysed for the expression of E-cadherin and Vimentin by IF, as described before (section 2.2.6).

2.2.10.5 Cultures with IL6 supplementation

$1x10^4$ cells were seeded in 10cm culture dishes with or without 50ng/ml recombinant human IL-6 (Life Technologies) in a total volume of 6ml MCF-7 medium (Table 4). The cells were incubated for 7 days. The medium was replenished every 2-3 days. At the end of the experiment, cells were collected with TRI-reagent for gene expression analysis.

2.2.10.6 Trans-wells with BrCAFs

Cells were collected and counted as before (2.2.8.1) in the appropriate medium. $2x10^3$ MCF-7 cells were plated in the well of the transwell plates (Corning incorporated, CoStar) in 600µl MCF-7 medium (Table 4). Cells were added in the insert in their growth medium (MCF-7 medium or CAFs medium; Table 4) as follows: $2x10^3$ MCF-7 (1:1 ratio; control), $6x10^3$ MCF-7 (1:3 ratio; control), $2x10^3$ BrCAFs (1:1 ratio), and $6x10^3$ BrCAFs (1:3 ratio).

2.3 *IN VIVO* STUDIES

The studies were conducted under the UK Home Office Licence number PPL 40/3559. NCRI guidelines for the welfare and use of animals in cancer research, LASA good practice guidelines, and FELASA working group on pain and distress guidelines were followed. To see an example of a full study protocol, including adverse effects, see appendix 1. The chemicals used were purchased from Sigma Aldrich, unless otherwise stated.

2.3.1 Cell maintenance

On day of initiation of the study, cells were harvested from semi-confluent monolayers. Cells with viability of >90% were re-suspended, for *in vivo* administration, in sterile standard formulation matrigel (BD Matrigel™ Matrix Basement Membrane) or in culture medium at $2x10^6$ cells/100µl/mouse. In the studies in which MSCs were co-implanted, $1x10^6$ MSCs were mixed with $2x10^6$ MCF-7 cells or BT549 and re-suspended in 100µl of sterile standard formulation matrigel. Cells were aliquoted into sterile tubes and transported on ice to the *in vivo* facility for initiation within 30mins of arrival. The same cell number of MSCs was used for re-injecting with MSCs in some studies.

2.3.2 Animals (species, strain, sex, identification and diet)

MF1 female nude mice (HsDOla:MF1-*Foxn1*^{nu}) at 4 to 6 weeks of age were obtained from Harlan Laboratories (Leicester, UK). Female Rag2$^{-/-}$ γc$^{-/-}$ mice (C;129S4-*Rag2*$^{tm1.1Flv}$ *Il2rg*$^{tm1.1Flv}$/J) were bred in house under a breeding licence (PPL40/3576). Each animal was earmarked and assigned a cage number which appeared on the data sheets. Ear markings were as follows; pl (plain or no marking), 1r (1 notch in right ear), 1l (1 notch in left ear), 1r1l, 2r etc. For the mice injected with MCF-7 cells, estradiol (Sniff, Germany) was added to their diet.

2.3.3 Tumour initiation (site of implantation)

Mice were injected with 100µl of the cells in the mammary fat pad (MFP); under the skin of the left upper nipple of the mouse. Otherwise, they were injected with the cells subcutaneously, under the skin on the left flank of the mice. For those studies where a “MSCs boost” was included, 100µl of MCF-7/MSCs mixture were injected at initiation and then an additional 100µl of MSCs (same number as at initiation) were injected intratumorally upon tumour formation.

2.3.4 Monitoring/tumour measurement

Tumour establishment was monitored by calliper measurements weekly. The length and width of tumours were measured, and the tumour volumes were calculated as follows: (length x ($width^2$))/2. Tumour growth was monitored weekly, or once every two weeks, by optical imaging, carried under anaesthesia in an IVIS Spectrum. For the bioluminescence imaging, D-luciferin (Caliper Life Sciences) was injected in the mice and the imaging was carried out 15 minutes after injection. The amount/volume of D-luciferin added dependent on mouse weight (for example 200µl of 15mg/ml D-luciferin were injected in a mouse weighted 20g, 250µl were injected in a mouse weighted 25g and so on). The imaging was carried out at different exposure times (1 second, 10 seconds, 30 seconds and 60 seconds). To analyse the bioluminescence imaging, the most representative image (i.e. not overexposed) was selected, regions of interest were drawn around the tumours, and the signal was calculated as total flux (photons/second) using the Living Image (4.3.1 build) software. For the fluorescent (mCherry) imaging, pictures were taken at three different excitation/emission filters (570ex/520em, 605ex/660em and 430ex/660em) to allow accounting for any crossover between the different filters. Regions of interests were drawn around the tumours and at a different site on the mice, for background control, and the total flux was calculated (605ex/660em). The total flux was calculated, and the background signal (autofluorescence) was subtracted from each value. To check for any spill-over between the different emission filters, the background signal

from the mice at other exposures/emissions was also subtracted. To avoid crosstalk between the high levels of bioluminescence and weaker levels of fluorescent signal, the fluorescent reads were performed prior to luminescent reads.

2.3.5 Termination

The studies continued until the tumours reached the maximum size ($1.2cm^2$), as specified by the Home Office licence, or other clinical signs necessitate removal of a particular mouse for the study. After the final imaging, the mice were sacrificed and the terminal imaging was carried out with the skin and muscle layers open to expose the primary site and adjacent lymphatics, as well as distant organs. The lymph nodes and spleens were resected and imaged individually. After imaging, the tumours and other organs of interest were weighed and preserved in fixative solution and/or in RNAlater (Ambion® Life Technologies).

2.3.6 *Ex vivo* analysis

2.3.6.1 Tissue and cell fixation, processing and sectioning

2.3.6.1.1 Fixation

The tissues were immersed in at least 20 volumes of fixative solution (Neutral Buffered Formalin, Thermo Scientific) for at least 24 hours before processing. For storage, the samples were transferred to 70% methanol. Cells cultured *in vitro*, were fixed in 500µl of 2% paraformaldehyde (diluted in PBS) for 10 minutes. They were mixed with 1% agarose and placed into 1.5ml micro centrifuge tubes containing an agarose plug. After the cells had settled, the plug was removed and placed in 70% methanol ON.

2.3.6.1.2 Processing

The tissues (or cells) were placed in embedding cassettes with disposable lid, and processed in the tissue processor (TP1020, Leica Microsystems UK). The chemical-station layout of the tissue processor was set as follows:

1. 70% Methanol diluted in distilled water
2. 90% Methanol diluted in distilled water
3. 100% Methanol
4. 100% Methanol
5. 100% Methanol
6. 100% Methanol
7. 100% Methanol
8. Xylene
9. Xylene
10. Xylene
11. Paraffin wax (VWR BDH Prolabo, West Sussex, UK)
12. Paraffin wax

The tissues were placed in stations 1-10 for 60 minutes at RT under vacuum and agitation and in stations 11 and 12 for 20 minutes at 60◦C under vacuum and agitation. After the processing cycle, the cassettes were removed and placed in the embedding station was reservoir (EG1160 Leica Microsystems UK), which was pre-warmed to 65◦C. Molten paraffin wax was dispensed in pre-heated moulds, and the tissues were placed into the wax and oriented as required. The moulds were allowed to set/cool and the cassette was removed for trimming the excess wax with a pen knife.

2.3.6.1.3 Sectioning

The tissues were chilled at around 4◦C for optimum cutting. The sectioning was performed using a microtome (LEICA RM2135). The blocks were fixed into the chuck and were trimmed in at 10-micron intervals until the required depth was attained. Tissues were cut at 4-microns and the tissue sections were allowed to form a continuous ribbon, which was transferred to a water bath (distilled water at 37◦C) using forceps. The tissue was allowed to stretch out and then transferred to poly-L-lysine coated glass slides (Polysine® slides, Thermo Scientific). The tissue was allowed to air dry ON before use.

2.3.6.2 Immunohistochemistry

All sections were de-waxed in 2 xylene baths for 5 minutes each, then rinsed in 2 100% methanol baths, for 1 minute each, and then rinsed in running tab water for 1 minute. After the staining process, the sections were incubated with DAB (1 drop of iquid DAB+Chromogen per ml of DAB+Substrate buffer) for approximately 3 minutes. After the sections were rinsed in running tap water, they were counterstained with Mayer's Haemalum for 3 minutes and then rinsed in running tap water. The sections were dehydrated in 3 100% methanol baths (1 minute each) and cleared in 2 xylene baths (5 minutes each). The slides were mounted to glass coverslips with DPX. Pictures were taken with the LEICA DMLB microscope (Leica Microsystems).

2.3.6.2.1 H and E staining

Following de-waxing, sections were incubated with pre-filtered Mayer's Haemalum (RAYMOND A LAMB Ltd., Eastbourne, UK) for 3 minutes. Sections were rinsed in running tab water for 1 minute and then incubated with pre-filtered Eosin (RAYMOND A LAMB Ltd.) for 3 minutes. The sections were stained with DAB, counterstained with Haemalum, dehydrated, cleared and mounted, as described above.

2.3.6.2.2 Cytokeratin, E-cadherin and Vimentin staining

Following de-waxing, the sections were immersed in cold citric acid from SIGMA® (10mM in distilled water, pH6) and heated in a microware at 98°C for 30 minutes (antigen retrieval step for E-cadherin) or 20 minutes (antigen retrieval step for Vimentin), or incubated with Proteinase K (Dako) for 5 minutes at RT (antigen retrieval step for cytokeratin).After that, the sections were immediately quenched in running tap water for 1 minute. Sections were incubated for 5 minutes in peroxidase block and then washed in PBS (2 times, 5 minutes each). The primary antibodies E-cadherin (mouse anti-human E-cadherin, clone NCH38, Dako) and Vimentin (anti-human Vimentin, clone V9, Dako) and the isotype control (mouse IgG1, Dako) were mixed mouse with biotinylation reagent and PBS, and incubated for 15 minutes to a final concentration of. 5 minutes

before use, the blocking reagent was added to the mix. The volumes were calculated using the ARKulator (Dako). The sections were incubated with the primary antibody or the control for 30 minutes at room temperature. The sections were rinsed in PBS (2 times, 3 minutes each) and then incubated with the secondary avidin:biotin peroxidase for 15 minutes. The sections were again rinsed in PBS and then incubated with DAB and rinsed. The sections stained with E-cadherin, were intensified in copper sulphate (0.5% $CuS0_4$ and 0.9% NaCl in distilled water) for 10 minutes. The sections were stained with counterstained with Haemalum, dehydrated, cleared, mounted, and imaged, as described above. The reagents used were provided in the form of a kit (Dako ARK™ (Animal Research Kit), Peroxidase for Mouse Primary Antibodies).

2.2.6.2.4 S100A4 staining

After de-waxing, were immersed in sodium citrate buffer (10mM in distilled water, pH6) and heated in a microware at 98°C for 30 minutes and left to cool down to RT for 30 minutes (antigen retrieval step). The sections were rinsed in TBS with Tween (137mM Sodium Chloride, 5mM Trizma, 0.1% Tween-20 in distilled water). The next steps were performed sections in a humidified chamber. The endogenous peroxidase was inactivated by covering the tissues with 3% hydrogen peroxide (diluted in 1XTBS) for 5 minutes. The slides were washed in TBS-Tween. The sections were incubated with the blocking solution (10% swine serum (Vector labs, Peterborough, UK) diluted in PBS), for 1 hour. The primary antibody (rabbit polyclonal anti-humanS100A4 antibody, EPR2761(2), C-term, GeneTex) was diluted in 5% swine serum (in PBS) at a final concentration of approximately 3µg/ml. The sections were incubated with the primary antibody or the control (rabbit serum, Vector labs) in the cold room ON. To avoid evaporation, the sections were covered with a piece of parafilm. The sections were washed with TBS-Tween and incubated with the secondary antibody (polyclonal swine anti-rabbit immunoglobulins/HRP, Dako), which was diluted in 5% swine serum at a final concentration of 10µg/ml, for 1 hour at RT. The sections were washed in TBS-Tween for 5 minutes and incubated with DAB. The

sections were then stained counterstained with Haemalum, dehydrated, cleared, mounted, and imaged, as described above.

2.3.6.3 Tumour homogenisation for gene expression analysis

Tumours stored in RNAlater were mechanically homogenised using a mortar and pestle and collected with TRI-reagent. RNA extraction, cDNA synthesis and qRT-PCR were performed normally as previously described (section 2.2.7).

2.4 STATISTICAL ANALYSIS

Statistical analysis of the data was performed using Graph Pad Prism Version 6 (San Diego, California). The linearity of bioluminescent signal against cell number was analysed by linear regression (R^2 coefficient). The correlation between baseline bioluminescent signal and baseline gene expression levels in the transduced cell lines, and other correlation analyses, was analysed by the Pearson's correlation coefficient. The significant changes in gene expression or bioluminescent signal between different conditions (e.g. normoxia versus hypoxia) was analysed by a ratio paired t-test. Changes in gene expression were also assessed by an unpaired t-test when necessary. Finally, differences between two groups (e.g. average tumour volumes) were also assessed by an unpaired t-test.

CHAPTER 3: RESULTS 1 – Development and validation of the activity of EMT reporters *in vitro*

This chapter describes the development and validation of the activity of EMT reporters *in vitro*.

The cell lines used in this study were initially selected based on previous studies performed in-house and the literature that defined MCF-7 as epithelial (luminal) and MDA-MB-231 and BT549 as mesenchymal (Basal B/triple-negative) [112]. MCF-7 cells were included in the study to be used as a model for inducing EMT, in order to test the EMT reporters upon induction of EMT, both *in vitro* and *in vivo*. The mesenchymal cell lines, MDA-MB-231 and BT549, were included in the study to be used in *in vivo* metastatic models, in order to test the activity of EMT reporters before and after metastasis. Metastasis of MDA-MB-231 in mouse models has been previously reported [92, 95, 113]. To the author's knowledge, no previous *in* vivo models of spontaneous metastasis using the BT549 cells have been reported; therefore a metastatic *in vivo* model of BT549 would be a novel model. As BT549 belong to the same subtype as the MDA-MB-231 cells, it is expected that they will be highly tumourigenic and metastatic. Finally, as the selected cell lines represent different subtypes of BC (luminal and basal/triple negative) they will be used to assess the activity and the patient-relevance of the EMT reporters in the different BC subtypes.

In the first part of the chapter, to confirm the phenotype of the cell lines, epithelial or mesenchymal, and to determine which cell line(s) were the best to use to test the EMT reporters, the baseline expression of common EMT biomarkers (reviewed in [114]) was analysed in the cell lines. The markers included the epithelial marker E-cadherin, which is downregulated in EMT, and the mesenchymal markers N-cadherin, Vimentin, ZEB1, Twist, Slug and S100A4 (FSP1), which are acquired in EMT. The baseline expression of the EMT markers was detected by qPCR (gene expression) and western blot (protein expression).

It has been reported that EMT generates cells with stem-like properties [52], therefore the "CSC population" of the cell lines was investigated, to further characterise the cell lines and to determine which cell lines were best to use to test enrichment of cells with stem-like properties upon

induction of EMT. The markers used were CD44 and CD24, to detect the $CD44^{+}/CD24^{low/-}$ population, which has been previously defined as CSCs in BC [29-31]. To identify this population, the cells were stained with CD44 and CD24, and the reported $CD44^{+}/CD24^{low/-}$ population was quantified by flow cytometry, which has been used for identifying the $CD44^{+}/CD24^{low/-}$ population in different BC cell lines [29-31].

The second part of the chapter describes the development and validation of inducible EMT reporters driven by the promoters of the EMT genes/markers (S100A4, E-cadherin or N-cadherin) upstream of a bioluminescent reporter, firefly luciferase. Bioluminescent reporters have been extensively used for studying gene expression at the transcriptional level [115] and were therefore used for the purposes of this study to measure the changes in the reporter transcription levels. As the reporters would be eventually tested *in vivo*, choosing bioluminescent rather than fluorescent reporters would have advantages, including high sensitivity and low background levels, which would allow the detection of a smaller number of cancer cells [116, 117]. This would allow detecting even a small proportion of cells in which changes in the activity of the reporters will occur.

Briefly, the reporters were constructed by conventional molecular cloning (information on molecular cloning can be found at http://www.moleculercloning.com/). Stable cell lines expressing the reporters were generated, using a commercially available lentiviral expression system (Lenti-X™ Lentiviral Expression System). The linearity between the cell number and the bioluminescent signal was analysed, and the ability of the reporters to detect the baseline levels of the reporter genes was investigated. Finally, the ability of the reporters to detect the changes in the expression of the reporter genes was explored.

The planned reporters were based on a lenti-viral expression vector. Lentiviral vectors have been widely used in research as a tool of transferring genetic material in mammalian cell lines, and will be used in this study due to their ability to integrate in both dividing and non-dividing cells and sustain long-term expression [118].

A bioluminescent reporter measuring the transcription of Vimentin has been previously reported, which was used to measure EMT in BC cell lines [83]. In this study, the reporter was used to screen for EMT inhibitors in MDAMB231 3D spheroids using a small compound library by measuring inhibition of Vimentin transcription/decrease in the bioluminescence signal. In this thesis, bioluminescent EMT reporters were developed in order to detect EMT in real-time, that could be subsequently be used for screening for inhibitors of EMT in early and late stages of EMT.

The S100A4 reporter was used in this study, in order to detect EMT at early stages and mark the cells with prometastatic properties. S100A4 (also known as fibroblast specific protein 1;FSP1) is a small ubiquitous calcium-binding protein that is involved in increasing cell mobility to allow invasion and migration [119]. It targets different proteins that are involved in cytoskeletal rearrangement and cell motility and it has many extracellular functions, including remodelling of the extracellular matrix, cell motility, differentiation and survival [119, 120]. S100A4 has been shown to have an early role in EMT [121] and has been used as an early marker of EMT [122, 123]. In cancer, EMT has been linked to metastasising cells, and it was therefore hypothesised that some of the prometastatic roles of S100A4 are due to its role in EMT (reviewed in [124]). There is evidence in the literature linking S100A4 and EMT in cancer, such as its inverse correlation with epithelial markers and its ability to activate transcription factors involved in the process of EMT. In breast cancer, S100A4 correlates with experimental metastasis, and metastasis and poor prognosis in BC patients [124].

E-cadherin and N-cadherin are both common markers of EMT, as they mark fundamental changes in the cytoskeleton and cell motility characteristic for EMT, and will be therefore used in this study to investigate these changes. Cadherins are transmembrane cell-cell adhesion molecules that have different functions in morphogenesis, including the control of cell contact and cell signalling (reviewed in [125]). E-cadherin has an important role in epithelial cells [126], while N-cadherin

is a neural cadherin [127]. Downregulation of E-cadherin is a hallmark of EMT and the downregulation of E-cadherin is concomitant with the upregulation of N-cadherin, resulting in a ‘cadherin switch’ [128]. This switch results in increasing motility and invasion of the cells, by weakening the affinity of cells epithelial cell interaction and acquisition of the ability to interact with mesenchymal cells.

The activity of the reporters was validated *in vitro* in stable cell lines expressing the reporters. The ability of the reporters to detect changes in the gene expression of the reporter genes was investigated in a model in which activation or repression of the reporter genes would be later induced. This induction was via hypoxia (culture of the cells in low oxygen culture conditions). During hypoxia, the Hypoxia Inducible Factor (HIF)-1 is activated and controls different processes including tumour proliferation and apoptosis and activation or repression of different pathways (reviewed in [129]). Briefly, in normoxia HIF-1α protein binds to the E3 ubiquitin ligase VHL complex leading to its degradation, while in hypoxia HIF-1α protein is stabilised and translocate to the nucleus to bind to HIF-1β, and together they bind to target genes (reviewed in [130]). A major role of HIF-1 is to promote angiogenesis by directly binding to the VEGF (vascular endothelial growth factor) promoter leading to its expression, which in turn promotes the migration of endothelial cells towards a hypoxic area. As VEGF is a downstream target of HIF-1, VEGF will serve as a hypoxia marker in this study, to confirm that hypoxia was induced in the different assays. As the activity of HIF-1α is dependent on protein stabilisation as discussed above, analysing the gene expression levels of HIF-1α wouldn’t be relevant in this study; upregulation of the gene will not necessarily mean activation of the protein and downstream targets. S100A4 has also been reported as a direct target of HIF-1, which increased its expression [131, 132]. Therefore, the activity of the S100A4 reporter was assessed in hypoxia, where it was expected that the gene expression of S100A4 would be increased, resulting in increase in the bioluminescent signal. HIF-1 also regulates the expression of E-cadherin and N-cadherin indirectly via the activation of different TFs (e.g. TWIST)

[133-135]. Thus, hypoxia was used as a model to induce changes in the gene expression of E-cadherin and N-cadherin and test the activity of the reporters. The changes in the gene levels of the EMT genes were assessed by qPCR. The changes in the bioluminescent signal were assessed by luciferase assays, where the signal from the reporter line was normalised to the signal from the control line (cells transduced with a vector that allows constitutive expression of firefly luciferase). This allowed normalisation of the differences in the growth rate of the cells in the different conditions, accounting for any factors that could have affected luciferase expression (oxygen, ATP, cell density, viability and pH). Emission of light, upon administration of the D-luciferin, requires oxygen and ATP [136] and it has been shown that changes in the oxygen levels and pH can significantly reduce the bioluminescent signal in *in vivo* models [137].

3.1 Choice of cell lines

3.1.1 Gene and protein levels of EMT markers

3.1.1.1 Gene expression of EMT markers

The baseline gene expression of EMT markers was analysed in the cell lines. RNA was extracted from the cells, cDNA was synthesised and the gene expression was determined by qPCR. The average $2^{-\Delta CT}$ values are summarised in Table 8. The $2^{-\Delta CT}$ values for each gene were plotted for all the cell lines, to allow comparison across all cell types for each gene (Figure 13). MCF-7 expressed higher levels of the epithelial marker *CDH1* and lower expression of the mesenchymal markers compared to MDA-MB-231 and BT549. MDA-MB-231 expressed higher levels of *SNAI2* and *S100A4* compared to BT549. BT549 expressed higher levels of *CDH2*, *VIM*, *ZEB1* and *TWIST1* compared to MDA-MB-231.

3.1.1.2 Protein expression of EMT markers

The protein expression of the most common EMT markers reported in the literature, E-cadherin and Vimentin, was analysed in the cell lines. Cell lysates were collected, the protein concentration was determined by a Bradford assay, and protein expression was detected by a Western Blot. MCF-7 cells expressed high levels of E-cadherin and undetectable levels of Vimentin (Figure14). E-cadherin could not be detected in either MDA-MB-231 or BT549 (Figure 14), while both the cell lines expressed Vimentin, with MDA-MB-231 showing higher expression compared to BT549.

Table 8. Baseline gene expression of EMT markers in BC cell lines. The values represent the average $2^{-\Delta CT}$ values, from three or more biological replicates, as analysed by qPCR.

		CELL LINE		
		MCF-7	**MDAMB231**	**BT549**
GENE	***CDH1***	20.4520	0.0046	0.0355
	CDH2	0.0005	0.0064	3.1332
	VIM	0.0260	34.7090	56.4600
	ZEB1	0.0037	0.6154	0.7251
	TWIST1	0.0001	0.0006	2.0179
	SNAI2	0.0004	0.3752	0.3224
	S100A4	0.1602	2.5949	1.3967

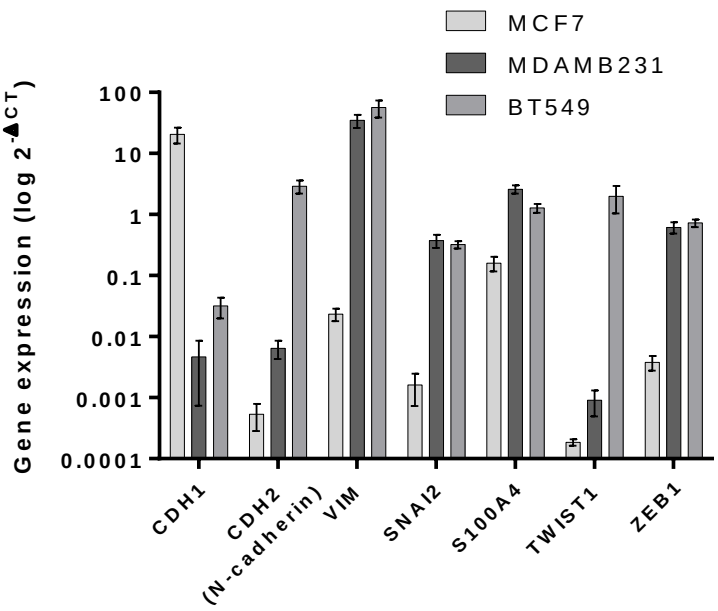


Figure 15. Gene expression of EMT markers in BC cell lines. The graph represents the average $2^{-\Delta CT}$ values from three or more biological replicates, as analysed by qPCR.

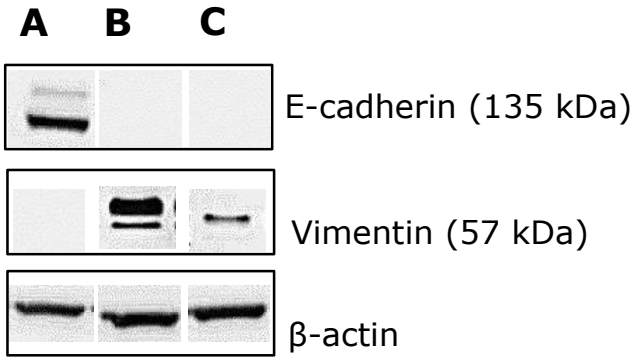


Figure 14. Western blot of E-cadherin and Vimentin in BC cell lines. Analysis of E-cadherin (135kDa) and Vimentin (57kDa) are shown in the top two lanes in MCF-7 (**A**), MDA-MB-231 (**B**) and BT549 (**C**). B-actin (42kDa) was used as protein loading control.

3.1.2 CSC POPULATIONS

The cell lines were stained with CD44 and CD24, and the reported CSC population ($CD44^{+}/CD24^{low/-}$) was quantified using flow cytometry. The cells were identified according to their size (Forward Scatter; FS) and their granularity (Side Scatter; SS). The appropriate compensations were applied based on single colour staining and level of background fluorescence measured on matched isotype controls for both test antibodies for each individual cell line (Figure 15A, 15C &15E). The compensations were applied individually for each cell line, to account for differences in the staining efficiency and autofluorescence.

MCF-7 cells displayed uniform low level (dim) CD44 staining combined with uniform high CD24 staining (Figure 15B). Conversely, MDA-MB-231 displayed uniformly high CD44 staining combined with negative CD24 staining (Figure 15D). BT549 contained two populations, both of which expressed CD44, but 61.4% also expressed CD24 and 39.6% were $CD44^{+}$/CD24 negative (Figure 15F). The CSC population was low in the MCF-7 (0.97%), higher in the BT549 (38.6%) and very high in the MDA-MB-231 (98.5%). The percentage of $CD44^{+}/CD24^{low/-}$ staining is summarised in Table 9.

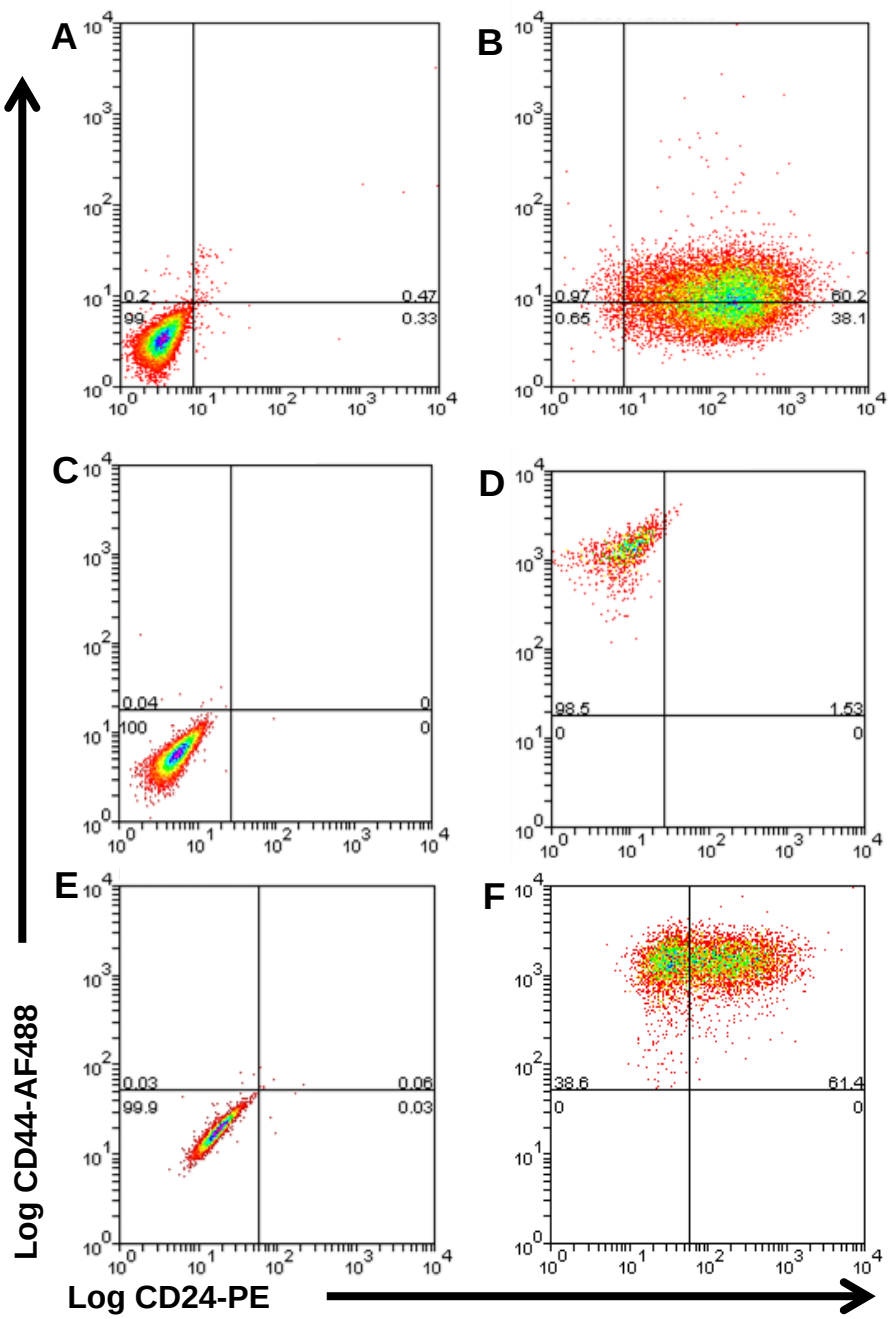


Figure 16. Flow cytometry analysis of BC cell lines stained with CD44 and CD24 and their isotype controls. A. MCF-7 isotype controls. **B.** MCF-7 CD44/CD24 dual staining. **C.** MDA-MB-231 isotype controls. **D.** MDA-MB-231 dual staining. **E.** BT549 isotype controls. **F.** BT549 dual staining. The percentage of cells in each Quadrant is indicated. N=2 (Figure shows representative example)

Table 9. Summary of flow cytometry data of BC cell lines stained with CD44 and CD24. The numbers represent an average of two replicates.

CELL LINE	$CD44^{+}/CD24^{low/-}$
MCF-7	0.97%
MDA-MB-231	98.5%
BT549	38.6%

Summary and discussion

MCF-7 expressed high levels of the epithelial marker E-cadherin (both at gene and protein levels) and did not express the mesenchymal markers. Conversely, MDA-MB-231 and BT549 did not express E-cadherin but expressed higher levels of the mesenchymal markers compared to the MCF-7. The data confirmed the 'Epithelial, E-cadherin-expressing' phenotype of MCF-7 and the 'Mesenchymal' phenotype of MDA-MB-231 and BT549 [112]. Also, high expression of E-cadherin and no or low expression of N-cadherin, Vimentin, Twist, Slug and ZEB1 in MCF-7 has previously been reported [112]. As S100A4 is involved in the early stages of EMT [138], it is expected that it is not expressed in epithelial cell lines like MCF-7.

MDA-MB-231 had the highest CSC population, as defined by $CD44^{+}/CD24^{low/-}$, followed by the BT549 and finally the MCF-7 which had a very small CSC population. The data showed that the CSC population was 0.97% in MCF-7, 98.5% in MDA-MB-231 and 38.6% in BT549. Previous studies have shown similar data, where it was demonstrated that the $CD44^{+}/CD24^{-}$ population was 0% in MCF-7 [29-31], between 80 and >90% in MDA-MB-231 [29-31] and ~36% in BT549 [31]. It is important to note that other CSC markers have also been identified in BC, including ALDH1, CD133 and ESA [16, 30, 33, 139]. Also, it has been proposed that no single CSC markers exists and that it is more relevant to focus on the fundamental properties of the stem cells (e.g. tumour initiation) rather than in the expression of a single CSC marker [140].

The planned EMT reporters will measure the transcriptional activity of S100A4 or N-cadherin or E-cadherin promoters. E-cadherin is expected to be downregulated upon EMT, while N-cadherin and S100A4 are expected to be upregulated. As MCF-7 cells were confirmed as epithelial and their CSC population is very low, they will be used as a model to induce EMT and enrich for CSCs, and test the reporters upon induction of EMT. The MDA-MB-231 and BT549 were defined as mesenchymal with high CSC populations. Therefore, they will be used to test the S100A4 and N-cadherin reporters in *in vivo* metastatic models, in order to test the activity of the reporters in the cells before they metastasise. As these cell lines do not express E-cadherin, they will not be used to test that reporter; downregulation of E-cadherin will not be possible and no upregulation of E-cadherin is expected in EMT or metastasis.

3.2 Development and validation of reporters

3.2.1 Molecular cloning of EMT reporters

3.2.1.1 Optimisation of cloning strategy

In the attempt to generate the various plasmid constructs used in this study, the cloning method was designed and optimised as appropriate. Briefly, the gene of interest was amplified from plasmid or genomic DNA, gel extracted and purified, cloned into a TOPO vector for DNA sequencing, digested from the TOPO vector with the appropriate restriction enzymes, ligated with the digested cloning vector, and the ligation products were transformed into bacteria from which the DNA was isolated and sequenced. The primers used for the amplification were carefully designed to introduce unique and compatible restriction enzyme sequences, as well as additional bases upstream and downstream of the restriction sites for sufficient digestion of the PCR products, as suggested by New England Biolabs. During this process different kits were used (e.g. different polymerases to amplify the different genes) and many of the steps involved were optimised/amended including, the PCR

conditions, the incubation times of digestions, the intensity and time of exposure to UV light during gel extraction, and the amount of ligation products transformed into the bacteria, More details about each construct will be described in the next section.

3.2.1.2 PCR amplification

The S100A4 promoter (1078bp) was amplified from the commercially available plasmid, pLightSwitch_S100A4prom (Figure 7), using the HotStarTaq® DNA polymerase, at different annealing temperatures ranging from 47.1˚C and 59.1˚C (Figure 16A). The promoter was identified by the company using sophisticated algorithms for gene model construction and transcription site prediction. The N-cadherin promoter (2434bp)/regulatory sequence was amplified from genomic DNA extracted from MCF-7 (ENSG00000170558: 1312-3721bp), using the 3-Step Terra™ Polymerase Mix at three different annealing temperatures; 62.5˚C, 65.2˚C and 66.8˚C (Figure 16B). The sequence of the amplification primers has been previously reported [141]. Finally, the E-cadherin reporter (882bp) was amplified from the commercially available plasmid pLightSwitch_Ecadprom (Figure 8), using the HotStarTaq® DNA polymerase at different annealing temperatures ranging from 50.2˚C and 68.7˚C (Figure 16C). As with the S100A4 promoter, the E-cadherin promoter was identified by the company. The primers used to amplify the promoters, were designed to introduce *ClaI* and *XhoI* restriction sites at the 5’ and 3’ of the promoters accordingly, for subsequent replacement of the ‘*ClaI/XhoI*’ CMV promoter in the pLVX-CMV-fLuc plasmid (Figure 11).

3.2.1.3 TOPO cloning

Bright and clear PCR products (indicating high yield and purity) were purified and cloned into a TOPO vector (indicated by arrows; Figure 16). Following bacterial transformations of the TOPO vectors, a few colonies were selected and the plasmids were isolated using plasmid minipreps. The plasmids were digested with *ClaI* and *XhoI*, to select those clones with the insert incorporated, for subsequent DNA sequence analysis

(Figure 17). Digest of the clones was expected to reveal different size fragments depending on whether the PCR product was successfully cloned into the TOPO vector or not.

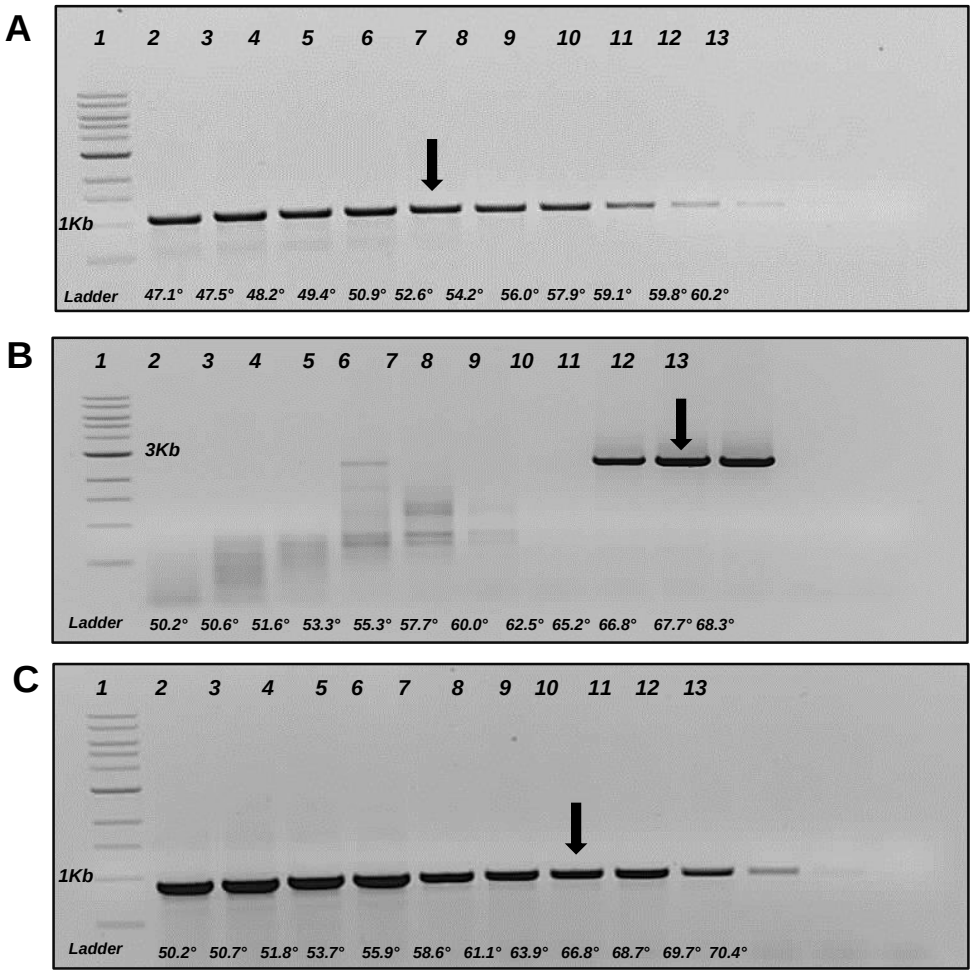


Figure 17. PCR amplification of the EMT promoters. A. Amplification of S100A4 promoter (1073 bases). Lanes: 1 = 1 kb DNA ladder, 2-13 = PCR products at different annealing temperatures (Gradient: 47.0°C – 60.0°C). Size of the product = 1078 bases. **B.** Amplification of N-cadherin promoter (2410 bases). Lanes: 1 = 1 kb DNA ladder, 2-13 = PCR products at different annealing temperatures (Gradient: 50.0°C – 68.0°C). **C.** Amplification of E-cadherin promoter (883 bases). Lanes: 1 = 1 kb DNA ladder, 2-13 = PCR products at different annealing temperatures (Gradient: 50.0°C – 70.0°C). The arrows indicated the PCR products that were used to generate the plasmid constructs. Agarose gels 0.8%.

The restriction digests showed that only three out of the six clones had incorporated the S100A4 promoter, while all the clones had incorporated the N-cadherin and E-cadherin promoters (Figure 17). The clones were sequenced and the sequence was aligned to the original sequences of the designed plasmid constructs. The alignment of the S100A4 clones showed that, unlike the rest of the clones, one of the clones contained only a single mismatch (A>T; position in the promoter 1045bp), and it was therefore selected for constructing the reporter. Similarly, for the N-cadherin promoter, the clone that contained a single mismatch (T>C; position in the promoter 912bp) was selected for constructing the reporter. No evidence were found in the literature suggesting that these mutations could change the activity of the promoters i.e. none of the known TF binding sites (will be discussed later) were affected by these mutations. The clone selected for constructing the E-cadherin reporter had one single mismatch from the original sequence (C>T; position in promoter 854bp), which was previously reported to be a variant (Ensemble Genome Browser).

3.2.1.4 Construction of reporters

The pLVX-CMV-fLuc plasmid (Figure 19D) was digested with *ClaI* and *XhoI*, so that the CMV promoter was removed (data not shown), and the plasmid DNA was gel-extracted. The digested plasmid was ligated with the EMT promoters, which were digested and gel extracted from the TOPO vectors with the same enzymes, to replace the CMV promoter. Following transformations, colonies were selected, the plasmids were isolated and following diagnostic digests (Figure 18), clones were selected for DNA sequencing, which confirmed that the EMT reporters were successfully generated (Figure 19). Throughout the thesis the constructs will be referred to as S100A4, E-cadherin and N-cadherin reporters. For the DNA sequence of the reporter constructs, see appendices 2-4.

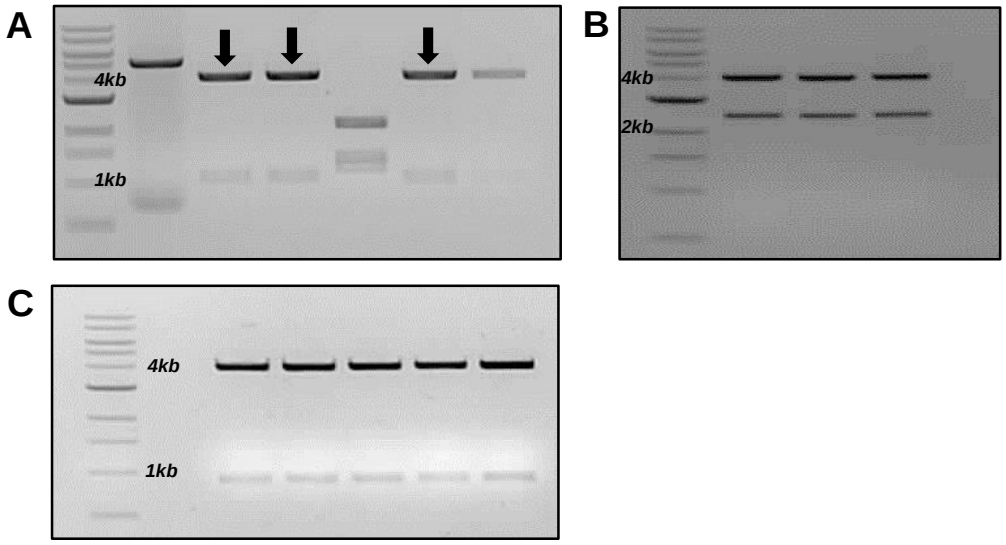


Figure 18. Diagnostic digests of TOPO clones with ClaI/XhoI. A. Digests of TOPO-S100A4 clones. Expected size fragments: (a) Empty vector=3956bp and (b) With insert=3967+1067. **B.** Digests of TOPO-N-cadherin clones. Expected size fragments: (a) Empty vector=3956bp, (b) With insert=3974+2416. **C.** Digest of TOPO-E-cadherin clones. Expected size fragments: (a) Empty vector=uncut and (b) With insert=3974+889. DNA ladder = 1kb DNA ladder. Agarose gels 0.8%.

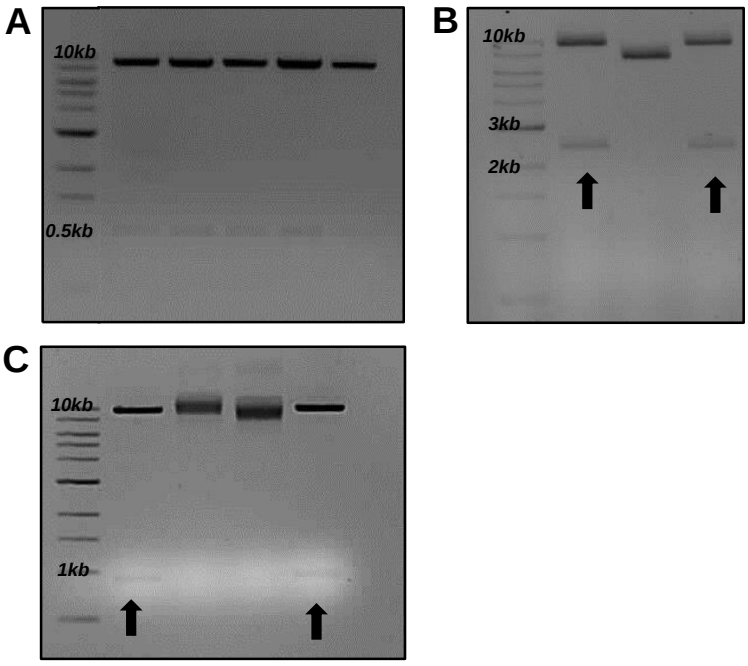


Figure 19. Diagnostic digests of clones of plasmid constructs of EMT reporters with ClaI and XhoI. **A.** Digests of S100A4 reporter construct. Expected size fragments: (a) Empty vector=9970bp, (b) With insert 9164+1067bp and (c) Insert in opposite direction=10231. **B.** Digests of N-cadherin reporter construct. Expected size fragments: (a) Empty vector=9165bp, (b) With insert 9164+2419bp and (c) Insert in opposite direction=11580. **C.** Digests of E-cadherin reporter construct. Expected size fragments: (a) Empty vector=uncut, (b) With insert 9164+889bp and (c) Insert in opposite direction=uncut. DNA ladder=1kb ladder. Agarose gels 0.8%.

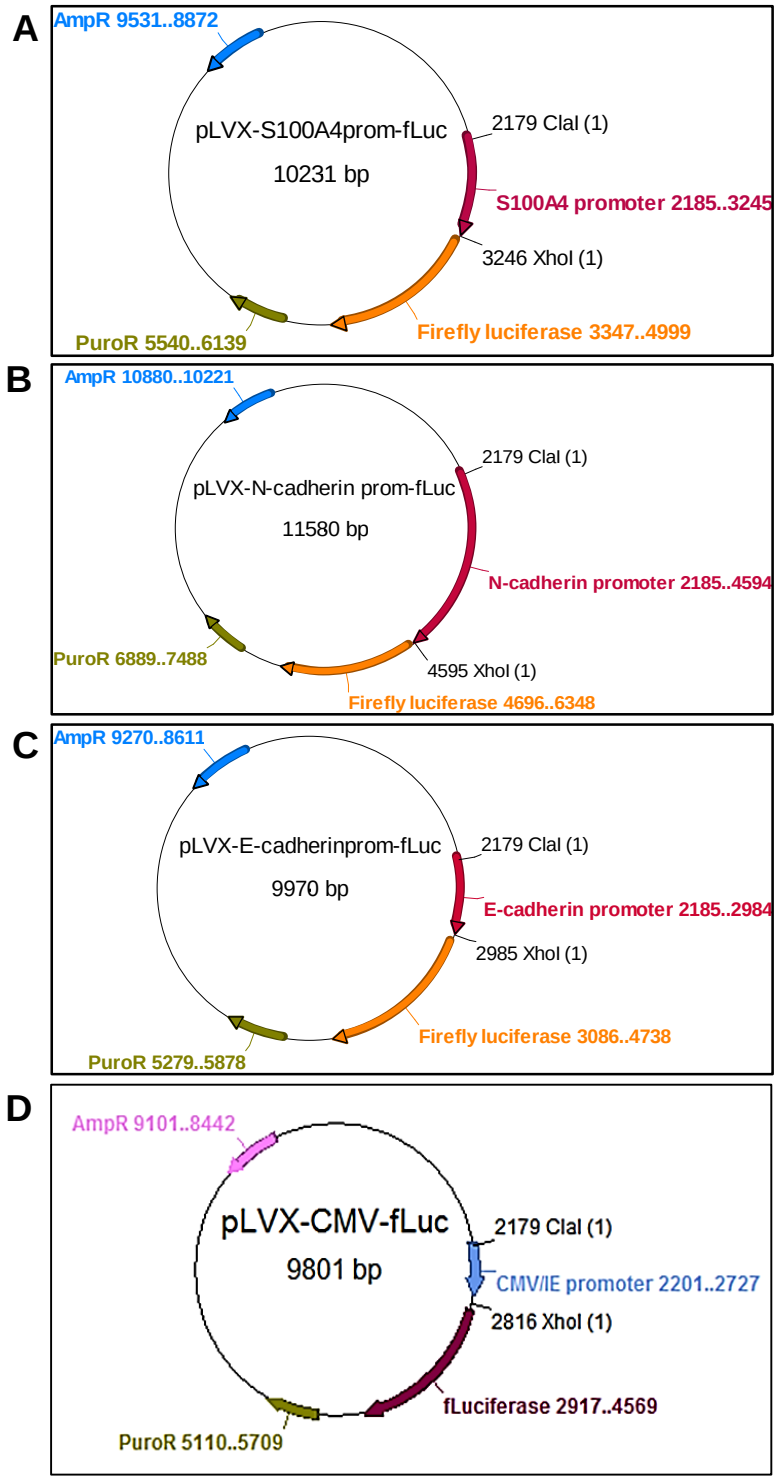


Figure 20. **Plasmid constructs of EMT reporters**. **A.** S100A4 reporter (pLVX-S100A4prom-fLuc). **B.** N-cadherin reporter (pLVX-N-cadprom-fLuc). **C.** E-cadherin reporter (pLVX-E-cadprom-fLuc). The constructs are based on 'pLVX-CMV-fLuc' (**D**). Expression of firefly luciferase is under the control of the S100A4 (A), or N-cadherin (B) or E-cadherin (C) human promoters. They were made by removing the 'ClaI/XhoI' CMV fragment from the original vector, and replacing it with the 'ClaI/XhoI' promoters digested from the TOPO vectors.

3.2.2 Generation of stable cell lines expressing the EMT reporters

Cell lines stably expressing the EMT reporters or the control vector, pLVX-CMV-fLuc (Figure 19D), were generated by lentiviral transductions in-house. Stable transductants were selected weekly with 4µg/ml puromycin; the amount of puromycin was previously determined in-house. To test the baseline activity of the reporters and assess whether the signal increased proportionately to the cell number, the correlation between the bioluminescent signal and cell number was tested. Cells were plated in black 24 well plates at different cell densities (serial dilutions) and following addition of D-luciferin, the bioluminescent signal was detected with the IVIS Spectrum, and the cell number against bioluminescent signal was plotted. The same cell numbers were seeded and the same batch of D-luciferin was used, for subsequent comparisons. Linear regression analysis confirmed that the bioluminescent signal emitted from the cells was proportional to the cell number in all the cell lines tested (Figure 20). As expected, the signal emitted from the cell lines expressing the control vector was higher than the signal detected in the cell lines expressing the EMT reporters; the control vector allows constitutive expression of luciferase, while the signal from the EMT reporters is inducible and dependent on the expression of the reporter genes. Throughout the thesis, the stable cell lines expressing the control vector and the EMT reporters will be referred to as control and reporter lines respectively. The cell lines will be denoted by /fLuc, /S100A4, /E-cadherin or /N-cadherin.

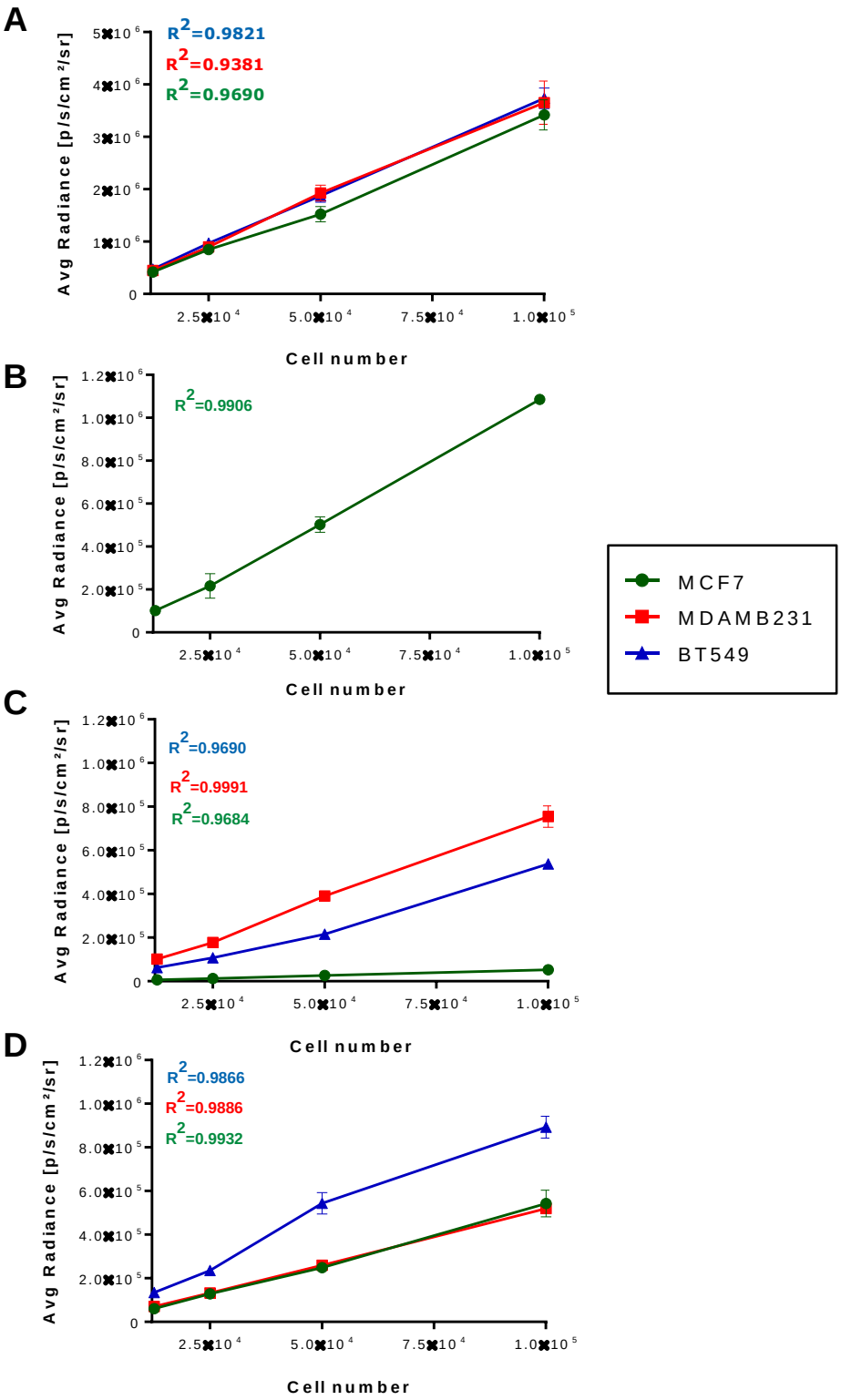


Figure 21. Linear regression of bioluminescent signal and cell number of control and reporter lines. A. Control cell lines. **B.** MCF-7/E-cadherin. **C.** S100A4 reporter lines. **D.** N-cadherin reporter lines. Red=MDA-MB-231, blue=BT549 and green=MCF-7. R^2=coefficient determination. N=2.

3.2.3 Initial correlation of signal levels with gene levels

The next step was to analyse whether the baseline bioluminescent signals corresponded to the baseline levels of gene expression of the reporter genes. First, the baseline levels of the reporter genes in the reporter lines were compared with the baseline levels of gene expression in the wild type cell lines (Table 8), to ensure that the transductions did not affect their expression. As shown in Table 10, the gene levels were very similar between the wild type and the transduced cell lines. Once it was confirmed that the expression of the reporter genes was not affected, the baseline bioluminescent signal previously detected at a particular cell density ($1x10^4$ cells/ml) across the reporter lines (Figure 20), was plotted according to the reporter gene (Figure 21A), and the patterns observed were compared with the baseline levels of gene expression in the reporter lines (Table 10), plotted in the same way (Figure 21B).

Table 10. Comparison of baseline gene expression of reporter genes between wild type and reporter lines. The values represent average 2-ΔCT values from three or more biological replicates.

Gene	**Cell line**	
	MCF-7 wild type	**MCF-7/S100A4**
S100A4	0.1602	0.3720
	MCF-7 wild type	**MCF-7/E-cadherin**
***CDH1* (E-cadherin)**	20.4520	20.9816
	MCF-7 wild type	**MCF-7/N-cadherin**
***CDH2* (N-cadherin)**	0.0005	0.0001
	MDA-MB-231 wild type	**MDA-MB-231/S100A4**
S100A4	2.5949	2.7171
	MDA-MB-231 wild type	**MDA-MB-231/N-cadherin**
***CDH2* (N-cadherin)**	0.0064	0.0055
	BT549 wild type	**BT549/S100A4**
S100A4	1.3967	1.0379
	BT549 wild type	**BT549/N-cadherin**
***CDH2* (N-cadherin)**	3.1332	2.2276

It was observed that the highest signal amongst the N-cadherin reporter lines was observed in BT549 cells, and this corresponded to the highest baseline gene expression of N-cadherin amongst the cell lines. Low signals were detected in the MCF-7 and MDA-MB-231, which corresponded to low gene expression levels. Similarly, the higher signal amongst the S100A4 reporter lines was observed in the MDA-MB-231, followed by the BT549 and finally the MCF-7. The same pattern was observed for the baseline gene levels of S100A4 in these cell lines. Looking at each cell line individually, it was shown that the MCF-7/E-cadherin cells emitted the highest bioluminescent signal compared to the MCF-7/S100A4 and MCF-7/N-cadherin. This corresponded with high baseline gene levels of E-cadherin and low gene levels of N-cadherin and S100A4. In the MDA-MB-231, the signal was high in the MDA-MB-231/S100A4 and lower in the MDA-MB-231/N-cadherin cell lines, which corresponded to high baseline gene levels of S100A4 and low levels of N-cadherin. Finally, in the BT549, high signal was observed in the BT549/N-cadherin and low signal in the BT549/S100A4, which again corresponded with high baseline gene levels on N-cadherin and lower levels of S100A4.

To further validate the reporter lines, the bioluminescent levels detected were correlated with the baseline gene levels of the matched reporter genes across all the reporter lines generated. A significant positive correlation was observed, as analysed by the Pearson correlation coefficient (Figure 22).

In summary, the analysis revealed that the EMT reporters can detect the baseline gene levels of the reporter genes.

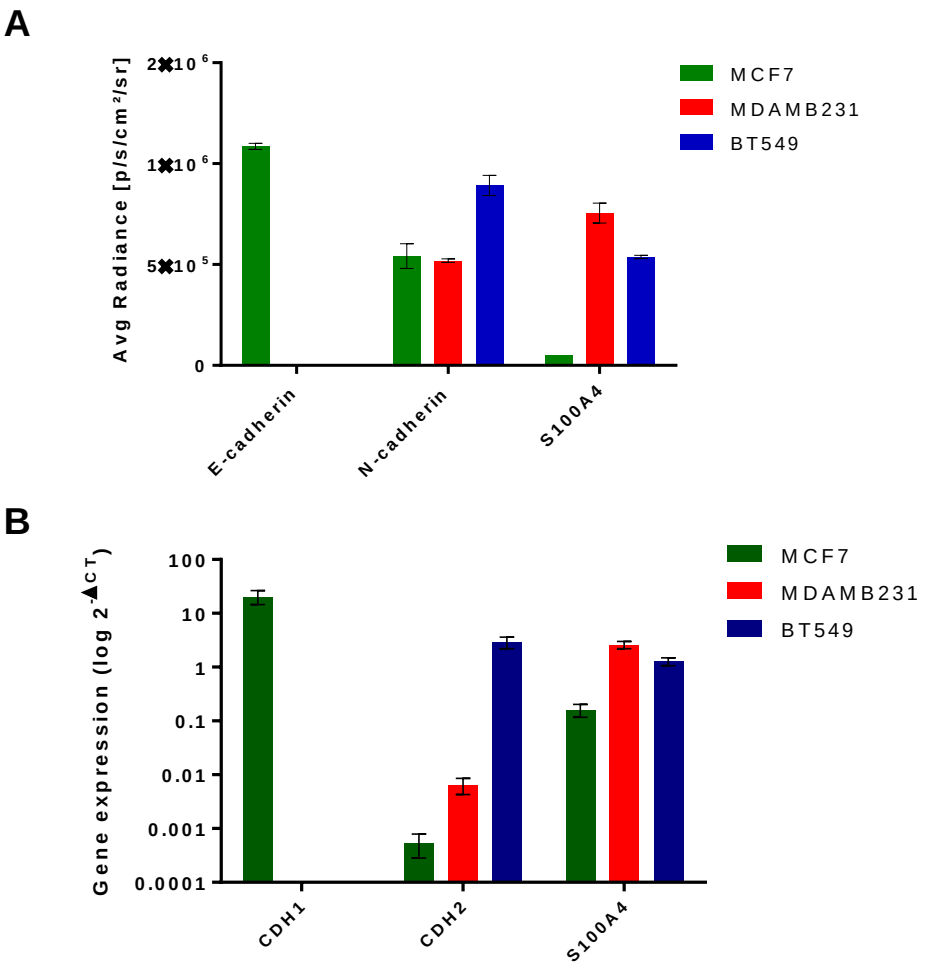


Figure 22. Comparison of baseline bioluminescent levels and gene levels in reporter lines. (A) Baseline bioluminescent levels of reporter lines at the same cell density ($1x10^4$). Each bar represents the mean Average Radiance ± SEM previously detected (Figure 20B-D) **(B)** Baseline gene expression levels of reporter genes in reporter lines. Each bar represents the mean $2^{-\Delta CT}$ values± SEM previously detected (Table 10).

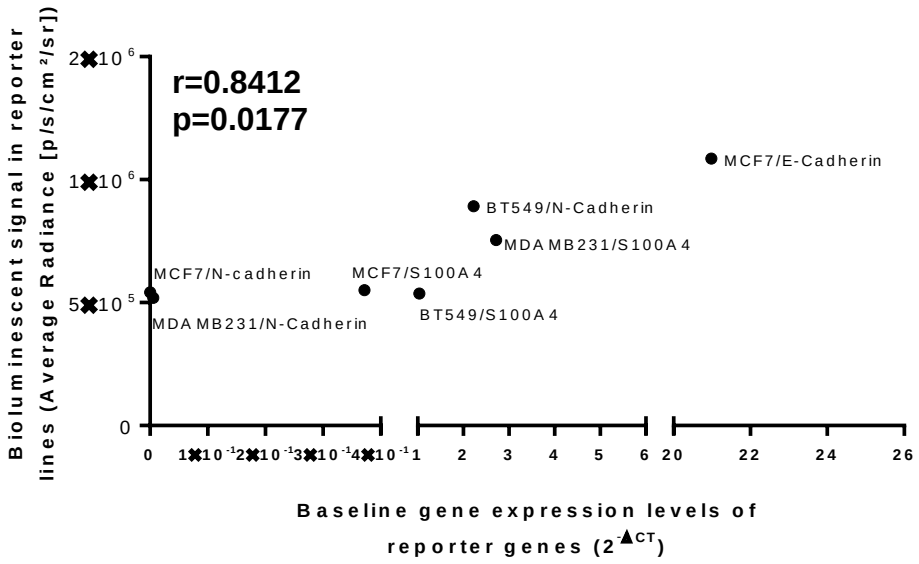


Figure 23. Correlation of baseline bioluminescent signal and baseline gene expression levels or reporter genes in reporter cell lines. Each point represents the average signal and average gene levels of the matched reporter gene in individual cell lines previously identified (Figure 21). r=0.8411, P=0.0177.

3.2.4. Validation of the activity of the reporters *in vitro*

The next step was to analyse whether the EMT reporters would successfully identify changes in the reporter genes. The reporter and control lines were seeded in black 24-well plates and cultured in normal (normoxia) or hypoxic conditions (hypoxia; 1% O_2) for four days. The experiments were repeated at least four times. An extra replicate was included when the results were not consistent and/or comparable. At the end of the experiment, D-luciferin was added to the wells and the bioluminescent signal was detected using the IVIS Spectrum. The cells were then collected with TRI-Reagent and, following RNA extraction and cDNA synthesis, the gene expression of VEGF and of the reporter genes was analysed by qPCR. The changes in the bioluminescent signal in reporter lines between normal conditions and hypoxia were assessed relatively to the signal in the control lines as follow: “Average signal in reporter line in hypoxia / Average signal in control line in hypoxia” or “Average signal in reporter line in normoxia / Average signal in control

line in normoxia". The changes of the genes and bioluminescent signal between normoxia and hypoxia were plotted for each individual biological replicate, and the significance of the changes amongst the replicates was assessed by a ratio paired t-test. The different biological replicates were plotted and analysed individually in order to account for the variability between the experiments arising from various factors including differences in the sensitivity of the D-luciferin (different batch or vial) and different cell numbers (pipetting errors and different growth rates).

3.2.4.1 S100A4 reporter

In the MCF-7/S100A4 cells, VEGF was significantly upregulated in all the replicates (Figure 23A), confirming that hypoxia was induced. The S100A4 gene expression and the bioluminescent signal was also significantly upregulated in all the replicates (Figures 23B and 231C). Similar data were observed in the BT549/S100A4 cells (Figure 24). In the MDA-MB-231/S100A4 cells, VEGF was only upregulated in three replicates, showing that hypoxia was not induced in all the biological replicates (Figure 25A). As a result, the changes in the S100A4 gene and the bioluminescent signal were not consistent (Figure 25B and 25C).

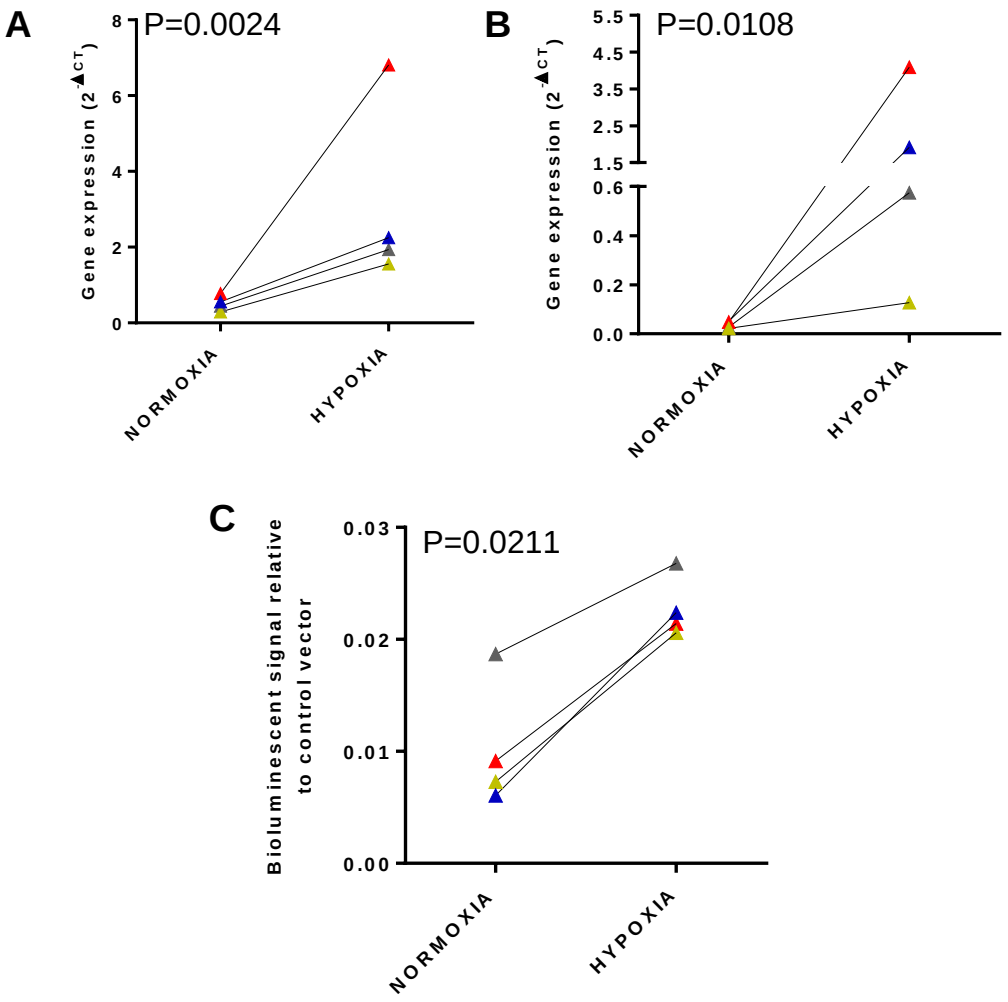


Figure 24. Validation of the S100A4 reporter in MCF-7 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of S100A4 in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=4

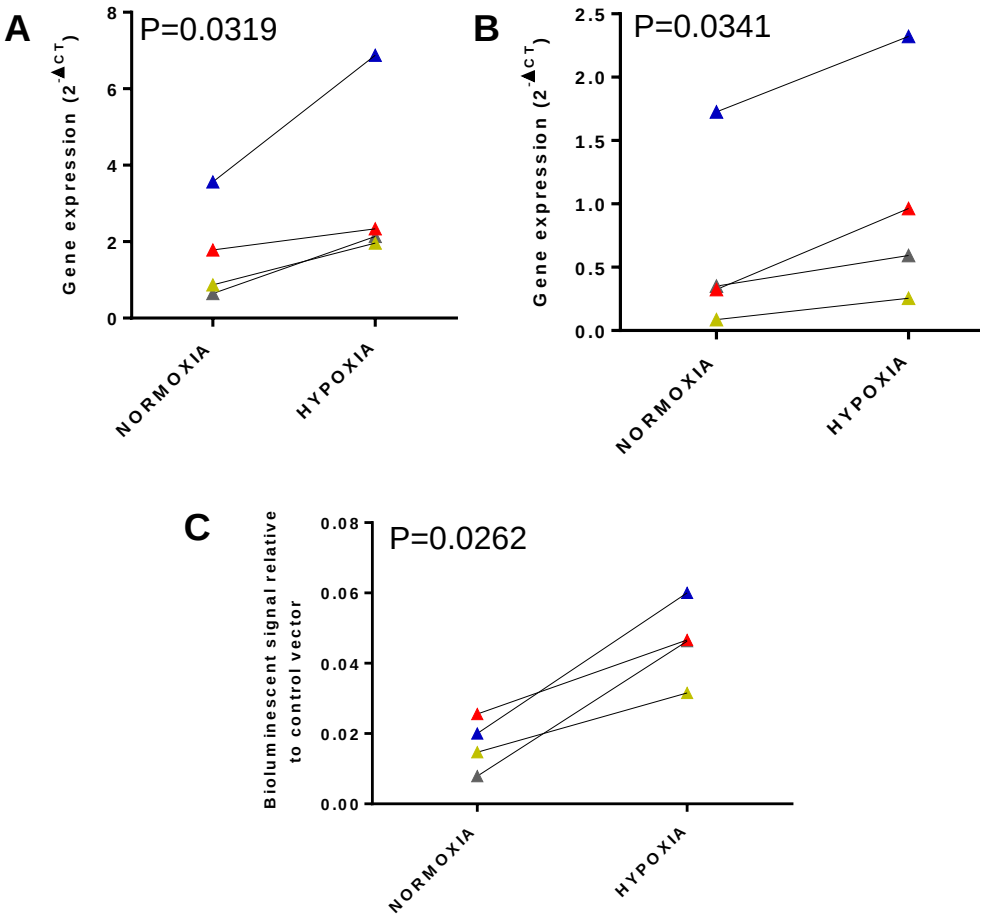


Figure 25. Validation of the S100A4 reporter in BT549 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of S100A4 in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=4

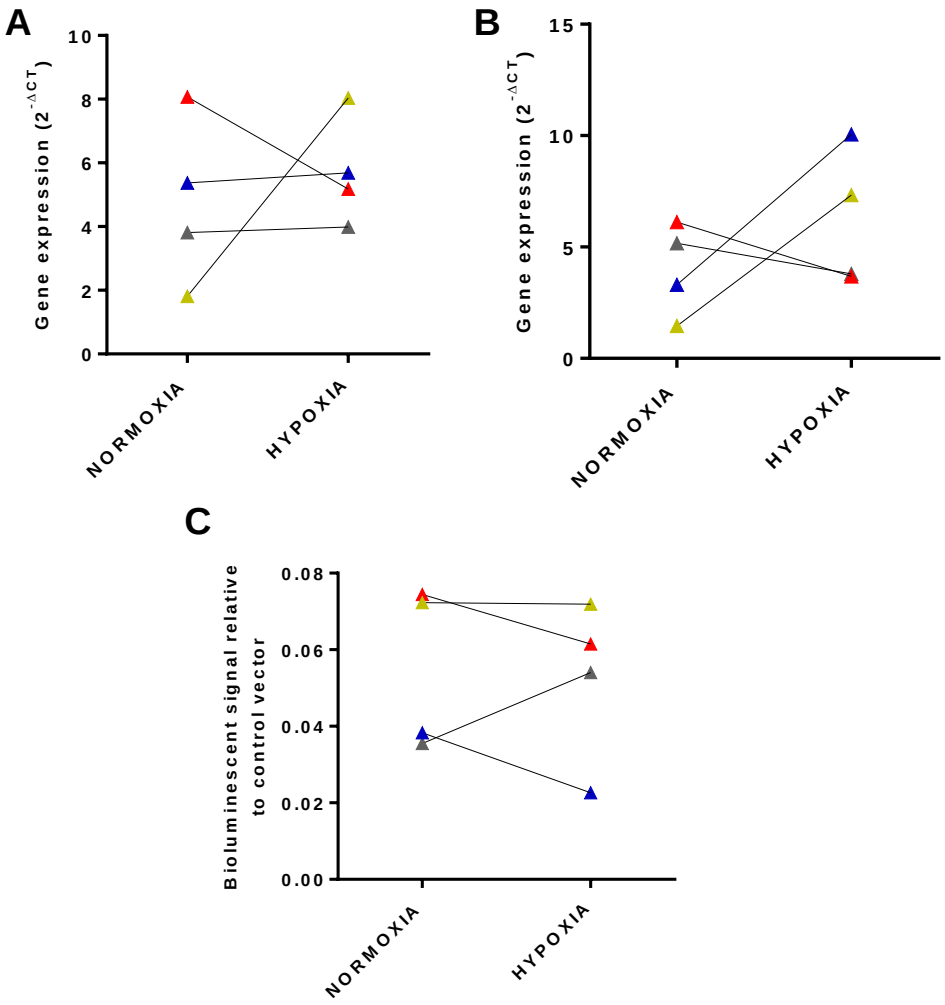


Figure 26. Validation of the S100A4 reporter in MDA-MB-231 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of S100A4 in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/ Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=4

3.2.4.2 E-cadherin reporter

VEGF was significantly upregulated (Figure 26A), confirming that hypoxia was induced in all the replicates. The changes in the E-cadherin expression and the bioluminescent signal were not consistent (Figure 26B and 26C). In three of the replicates the direction of the change was the same for the gene and the signal. Specifically, when the gene was downregulated (blue point), the signal was also downregulated and when the gene was upregulated (yellow and red points), the signal was also upregulated. In the other two replicates (grey and green points), the gene was downregulated but the signal was upregulated.

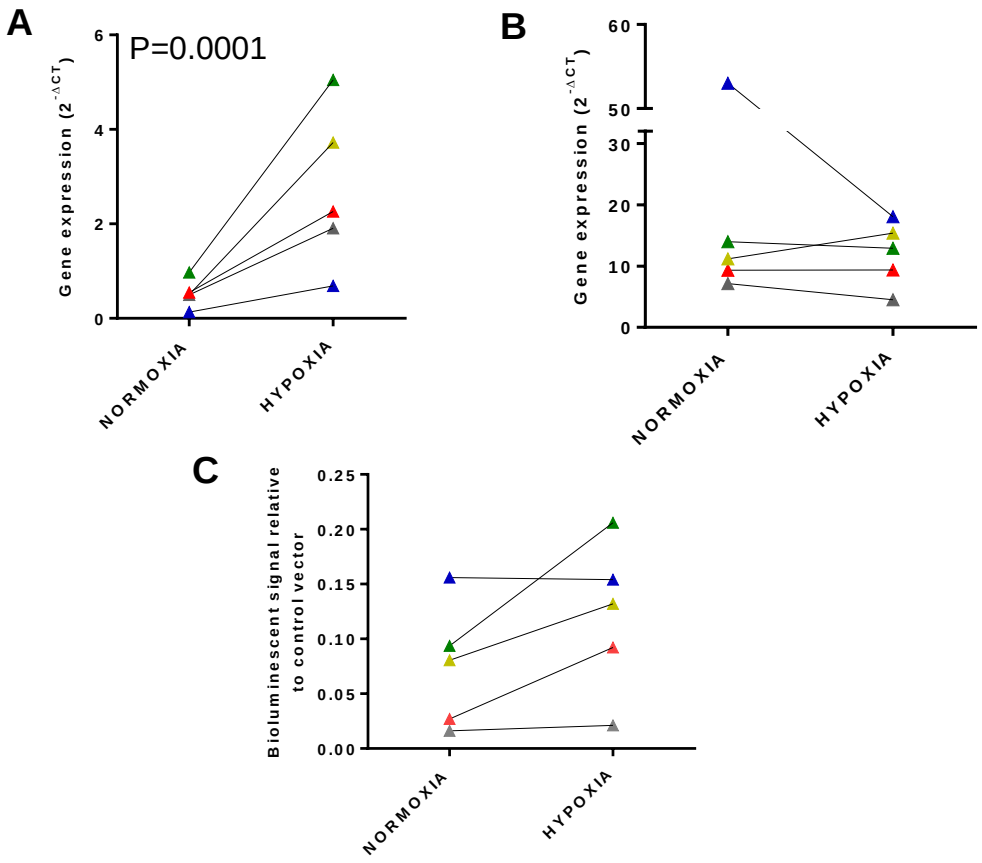


Figure 27. Validation of the E-cadherin reporter in MCF-7 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of E-cadherin in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/ Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=5

3.2.4.3 N-cadherin reporter

In the MCF-7/N-cadherin cells, VEGF was significantly upregulated (Figure 27A). The changes in the gene and the signal were not consistent (Figure 127B-C). As shown, the direction of the change was the same for the gene and the signal in four of the replicates. Specifically, when the gene was downregulated (blue point), the signal was also downregulated and when the gene was upregulated (yellow, red and green points), the signal was also upregulated. On the contrary, in one of the replicates (grey point), the gene was upregulated, but the signal was downregulated.

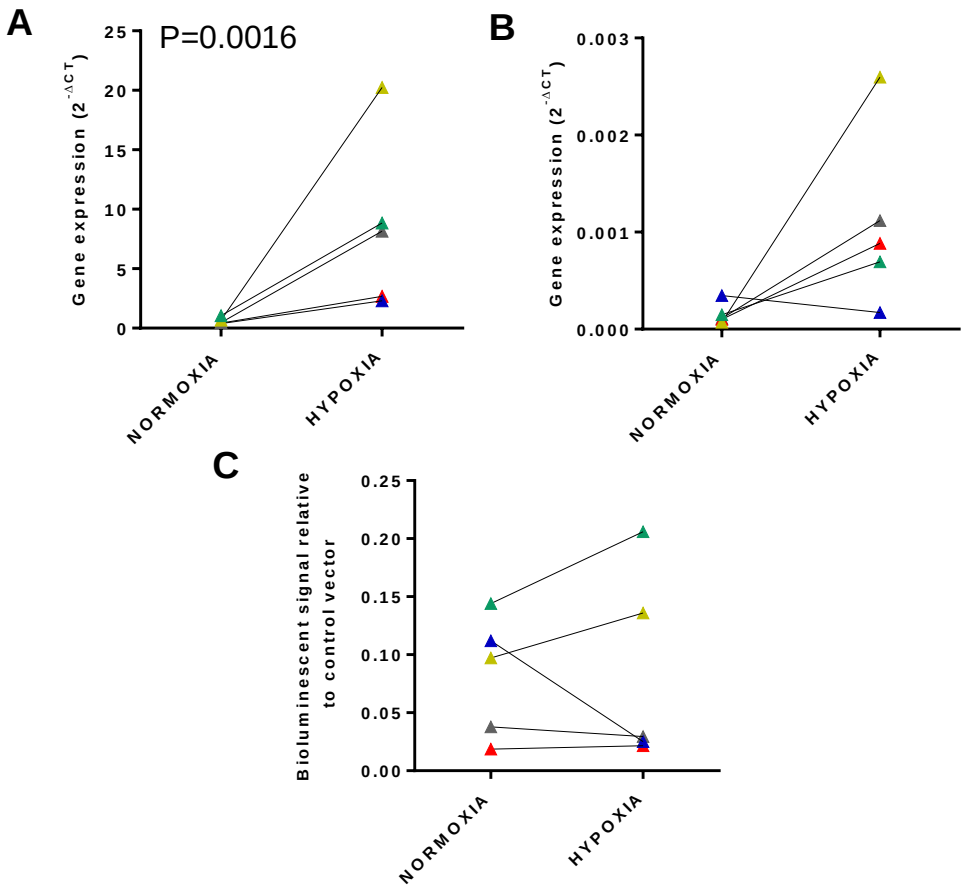


Figure 28. Validation of the N-cadherin reporter in MCF-7 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of N-cadherin in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/ Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=5

In the BT549/N-cadherin cells VEGF was significantly upregulated (Figure 28A). Neither the N-cadherin gene nor the bioluminescent signal was significantly upregulated (Figure 28B-C), and the direction of change in the gene was comparable with the change in the signal only in two biological replicates (yellow and red point).

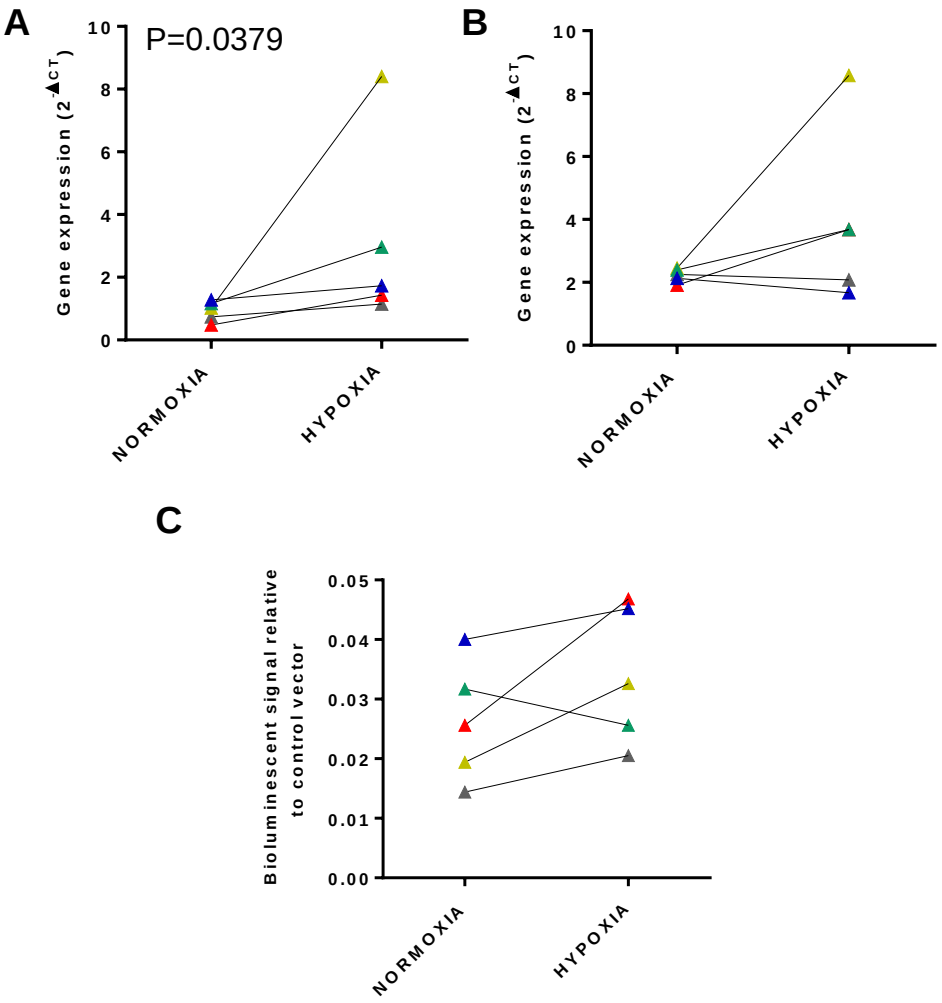


Figure 29. Validation of the N-cadherin reporter in BT549 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of N-cadherin in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/ Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=5

Finally, in the MDA-MB-231/N-cadherin cells, VEGF was not significantly upregulated (Figure 29A) showing that hypoxia was not induced in all the biological replicates. As would be anticipated, the gene and signal changes were neither consistent nor significant (Figure 29B-C).

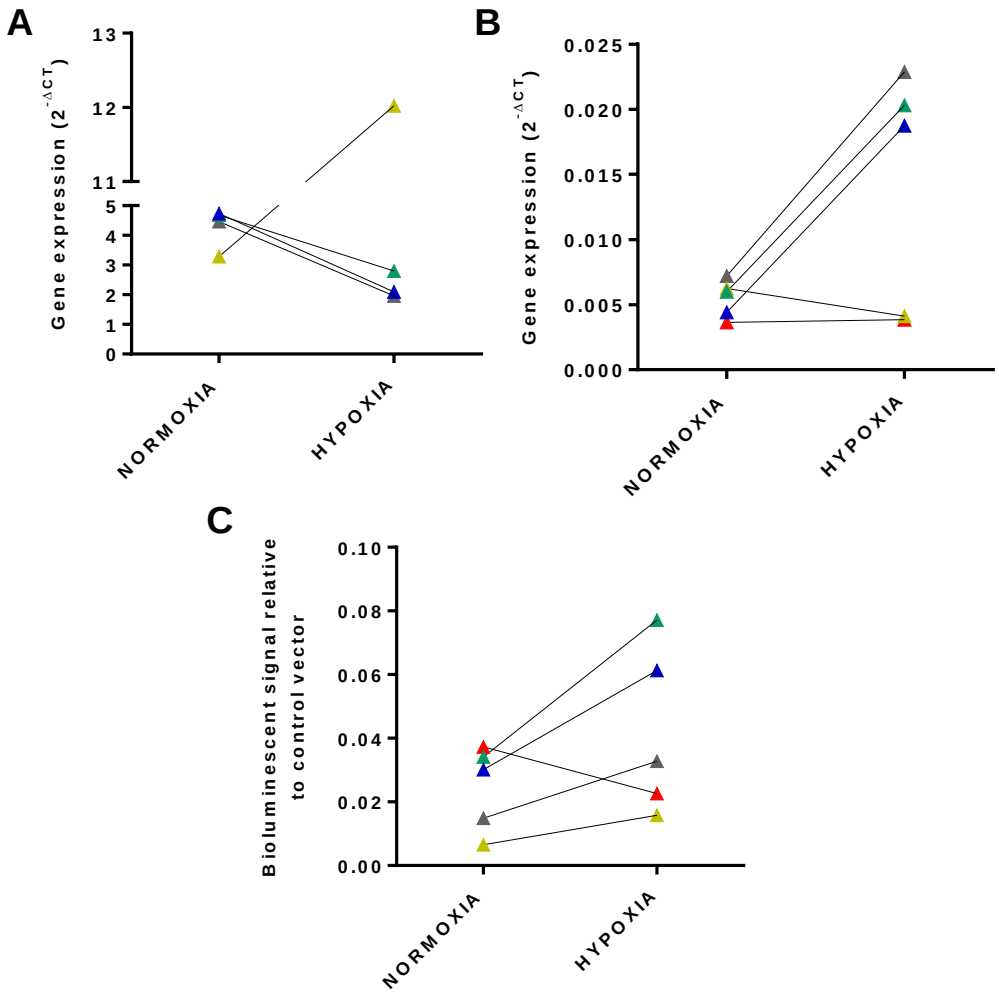


Figure 30. Validation of the N-cadherin reporter in MDA-MB-231 in hypoxia. A. Gene expression of VEGF in normoxia and hypoxia. **B.** Gene expression of N-cadherin in normoxia and hypoxia. **C.** Reporter signal in normoxia and hypoxia calculated as follows: Average reporter signal/ Average control signal. Each point represents a different biological replicate (colour matched in all the graphs). The significance of the different changes between normoxia and hypoxia was assessed by a ratio paired t-test. N=5

In summary, the analysis showed that the S100A4 reporter in MCF-7 and BT549 could effectively detect the direction of the changes in the reporter genes, and that was statistically significant. To investigate the ability of the reporter to detect the degree of changes, the fold change in the gene levels was correlated with the fold change in the bioluminescent signal between normoxia and hypoxia. No significant correlation was found in either the MCF-7/S100A4 or the BT549/S100A4 (Figure 30), as assessed with the Pearson correlation coefficient. This implied that the fold change in the bioluminescent signal did not correspond to the fold change in the gene levels.

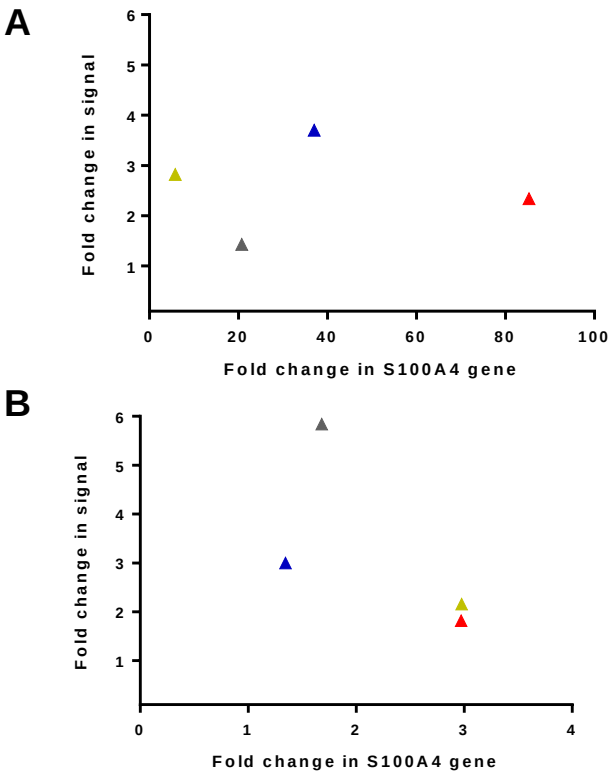


Figure 31. Further validation of the S100A4 reporter in hypoxia. Correlation of the fold change in the bioluminescent signal and the gene levels of the reporter genes between normoxia and hypoxia in **(A)** MCF-7/S100A4 and **(B)** BT549/S100A4. The fold change in the signal was calculated as follows: (Average signal in reporter line in hypoxia / Average signal in control line in hypoxia) / (Average signal in reporter line in normoxia / Average signal in control vector in normoxia). The change in the gene levels was expressed as a relative gene expression ($2^{-\Delta\Delta CT}$).Correlation was assessed by the Pearson correlation coefficient. N=4.

Summary and discussion

Inducible bioluminescent EMT reporters measuring the transcription of S100A4, E-cadherin or N-cadherin alongside firefly luciferase were successfully generated by conventional molecular cloning.

Stable cell lines expressing the reporters or the control vector were generated by lenti-viral transductions, and it was confirmed that the bioluminescent signal was proportional to the cell number.

The baseline levels of the bioluminescent signal in the reporter lines significantly correlated with the baseline gene expression of the corresponding reporter genes. This confirmed that the EMT reporters can detect the baseline levels of the reporter genes.

To validate the activity of the reporters, the ability of the reporters to detect changes in the gene expression of the reporter genes was assessed in hypoxic assays:

Induction of hypoxia was assessed by analysing the gene expression of the hypoxia marker, VEGF, a downstream target of HIF-1. It was shown that VEGF was significantly upregulated in all the MCF-7 and BT549 reporter lines, but not in the MDA-MB-231 cell lines. The expression of VEGF is dependent on all the different conditions involved in the activation and stabilisation of HIF-1 (reviewed in [142]). These include, amongst others, growth factors and cytokines such as EGF, FGF2 and insulin. Furthermore, it was shown that the Ras-Raf-MEK-ERK pathway is essential for the *trans*-activation of the VEGF promoter. Substitution mutations in the EGF receptor and *KRAS* have been previously reported in the MDA-MB-231 cells (COSMIC), suggesting a possible explanation as to why VEGF was not activated in these cells. Furthermore, it has been shown that the levels of HIF-1 expression are dependent on the cell density in the MDA-MB-231 cells [143], suggesting that these cells should have been seeded at a higher cell density in order to allow activation of HIF-1, and subsequent activation of VEGF. In order for HIF-1 to bind to target genes, its two subunits, HIF-1α and HIF-1β, need to form dimers [130]. It has been reported that HIF-1α was unable

to form heterodimers with HIF-1β in MDA-MB-231 [144], which is another possible explanation as to why VEGF was not upregulated in this cell line. Furthermore, the baseline gene levels of VEGF are higher compared to the other cell lines, implying that there is already a degree of HIF-1 activity within the cell line, and suggesting that possibly all the HIF-1α proteins present are already activated and no more activation would be expected. Finally, the synthesis, stability and degradation of HIF-1 are not only regulated due to changes in oxygen levels, but also through other mechanisms that are independent of oxygen levels, including growth factors (e.g. IGF) and acetylation status [130, 145].

The ability of the reporters to detect changes in the reporter genes was assessed by comparing the direction of change of the gene with the direction of change of the bioluminescent signal. The activity of the S100A4 reporter was validated in the MCF-7/S100A4 and BT549/S100A4 lines, where it was shown that both the S100A4 gene and the bioluminescent signal were significantly upregulated in hypoxic conditions. Upregulation of the S100A4 gene was expected in hypoxia, since S100A4 has been shown to be a direct target of HIF-1 [131, 132]. Five hypoxia response elements (HRE) sequences have been detected in the S100A4 gene and in fact, an HRE sequence (5' CAC GC 3') is present in the promoter of the S100A4 reporter, suggesting the direct regulation of the S100A4 reporter by HIF-1 (annotated in Figure 31). It is important to mention though, that the fact that the sequence is present does not necessarily mean that is functional, and that many other TFs control the gene expression of S100A4, as will be discussed later on.

The changes of the E-cadherin and N-cadherin genes were not consistent in hypoxia i.e. in some biological replicates the gene was upregulated and in others it was downregulated. Unlike S100A4, the expression of N-cadherin in hypoxia is not regulated directly by HIF-1, but through other TFs (Twist1, Snail, TCF3 and others) which are directly regulated by HIF-1 and induce EMT, in which N-cadherin upregulation is observed [134, 135]. Thus, changes in the N-cadherin expression, at least the ones associated with hypoxia-induced EMT, will be dependent

on the activation of factors other than HIF-1 or VEGF. Two HRE sequences (5' GGA ATC AGA ACC GTG CAG GTC CCA T 3' and 5' GTG GCC GGC AGG TGA ACC CTC A 3'), previously detected in the E-cadherin promoter [146], are present in the promoter of the E-cadherin reporter (annotated in Figure 32). It would be therefore expected that E-cadherin is directly regulated by HIF-1. In BC it has been shown that HIF-1 binding to the HRE of the E-cadherin promoter was enhanced by the transcriptional co-activator with PDZ-binding motif (TAZ) or the WWdomain-containing oxidoreductase (Wwox) [146], suggesting that other factors are needed for HIF-1 to interact with and downregulate E-cadherin, which might not have been activated in all the different experiments performed in this study. But again, it is not certain that the HRE sequence in the E-cadherin promoter is functional. Furthermore, many other TFs are involved in regulating the E-cadherin expression, including TFs that upregulate the gene expression (reviewed in [147]) (annotated in Figure 32). For example, SP1, a protein that has been shown to be activated in MCF7 in hypoxia [148], is a positive regulator of E-cadherin [147]. In fact, the sequence for the SP1 binding site (5' GGG GGG CGG TGC 3') [149], is present in the promoter of the E-cadherin reporter. Assuming that the sequence is functional, it could explain why E-cadherin expression was upregulated in some biological replicates

The inducible bioluminescent EMT reporters were designed so that the expression of the luciferase gene would be driven by the promoters of the S100A4, E-cadherin or N-cadherin genes. But, the direction of the changes in the E-cadherin and N-cadherin genes was not always comparable with the direction of the change of the bioluminescent signal. Furthermore, no significant correlation between the fold change in the gene and the fold change in the bioluminescent signal was observed in the cell lines expressing the S100A4 reporter. The factors that could have affected the expression of luciferase were controlled in this study, by normalising the signal in the reporter lines to the control lines, as discussed before. Therefore, the data could have been compromised due

to artifacts in the quantification of the gene expression of the reporter genes. Possible explanations for this are the following:

a) The gene expression of the reporter genes was analysed by qPCR, which quantified the expression of the reporter genes found in the cell lines tested. It was anticipated that when the promoters of the reporter genes in the genome were activated or repressed, then the promoter in the EMT reporters would have the same fate. But, this might not have been the case, since the promoters in the reporters may be missing regulatory elements found in the promoters of the genes in the genome, which could have affected their gene expression. For instance, S100A4 expression has been shown to be regulated by C/EBPα [150], and c-Myb [150], whose binding sites are not found in the promoter of the S100A4 reporter; they are found approximately 240 and 220 bases upstream the promoter respectively (annotated in Figure 31). Similarly, the gene expression of E-cadherin has been shown to be regulated by LUN [151] and CREB1 [149], and the gene expression of N-cadherin has been shown to be regulated by, Sp1/Sp3 and MZF-1 [152], whose binding sites are not found in the promoter of the E-cadherin and N-cadherin reporter respectively (annotated in Figure 32 & 33). It is worth noting though, that variation of the binding sites of these TFs might be present in the promoters of the EMT reporters, and that other TFs, whose role has not yet been studied, might be playing an important role in the regulation of the reporter genes. It is not expected that all of the TFs mentioned will be relevant to BC and/or hypoxia, but it is clear that the control of gene expression is not dependent on a single factor.

b) Post-transcriptional RNA modification and processing, transcript stability and localisation, are also very important for gene regulation (reviewed in [153]). Briefly, various sequence elements found within an mRNA can determine its stability and turnover by associating with other proteins or small non-coding RNAs. Examples include the internal ribosome entry sites (IRES) element, the c-Jun amino terminal kinase (JNK) responsive element (JRE) and the A+U rich elements (AREs). The turnover of mRNAs is controlled by various stimuli, one of which is

hypoxia (reviewed in [154]), in which various RNA-binding proteins (RBPs) facilitate changes in the mRNA stability and turnover. Therefore, mRNA stability and turnover in hypoxia can explain the variations observed in the gene changes (direction and fold change) in the various biological replicates. Also, it can explain why the changes in the reporter genes did not compare/ correlate with the signal, since different sequence elements that control the mRNA stability and turnover are present in the mRNA of the genes in the genome from the ones present in the transcribed luciferase mRNA in the EMT reporters.

c) qPCR detected the gene expression levels of the reporter genes in all the cells seeded. Even though the cells were selected prior to setting up the experiments, the selection process might not have been completely efficient resulting in the survival and expansion of untransduced cells. If that was the case, the bioluminescent signals would have been detected only from the transduced cells whereas qPCR would have detected the gene expression in all the cells. This would have resulted in bigger fold changes in the gene expression than in the bioluminescent signal; a phenomenon which was frequently observed in this study. One way to check this would be the use of fluorescent EMT reporters in order to check the percentage of transduced cells by fluorescent microscopy. The reporters could express either a fluorescent reporter or both a fluorescent and bioluminescent reporter. The latter could be achieved by transducing the cells with two different plasmids (i.e. the bioluminescent EMT reporter and a plasmid that allows constitutive expression of a fluorescent marker) or by a single plasmid construct that allows the simultaneous expression of the two proteins (fluorescent marker and luciferase). Another way to test for this would be to identify the copy number of the virus in the cells by absolute qPCR and normalise the gene expression data to the amount of virus present.

In further *in vitro* and *in vivo* experiments the S100A4 reporter in the MCF7 and BT549 (i.e. MCF-7/S100A4 and BT549/S100A4 cell lines) will be validated upon induction of EMT *in vitro*, and/or in *in vivo* models, in which changes in the S100A4 expression (gene and/or protein) will be

correlated with generated bioluminescent levels. Since the activity of the E-cadherin and N-cadherin reporters was not validated in this setting, other methods to induce changes in the E-cadherin and N-cadherin genes will be investigated in the next chapter. Once a model in which consistent changes in the gene expression will be identified, the activity of the reporter will be tested and the reporters will be validated upon induction of EMT. Testing the E-cadherin and N-cadherin reporters were made depending on data in the following chapters i.e. if significant and consistent changes in their expression (gene and/or protein) will were observed in *in vitro* and/or *in vivo* models of EMT.

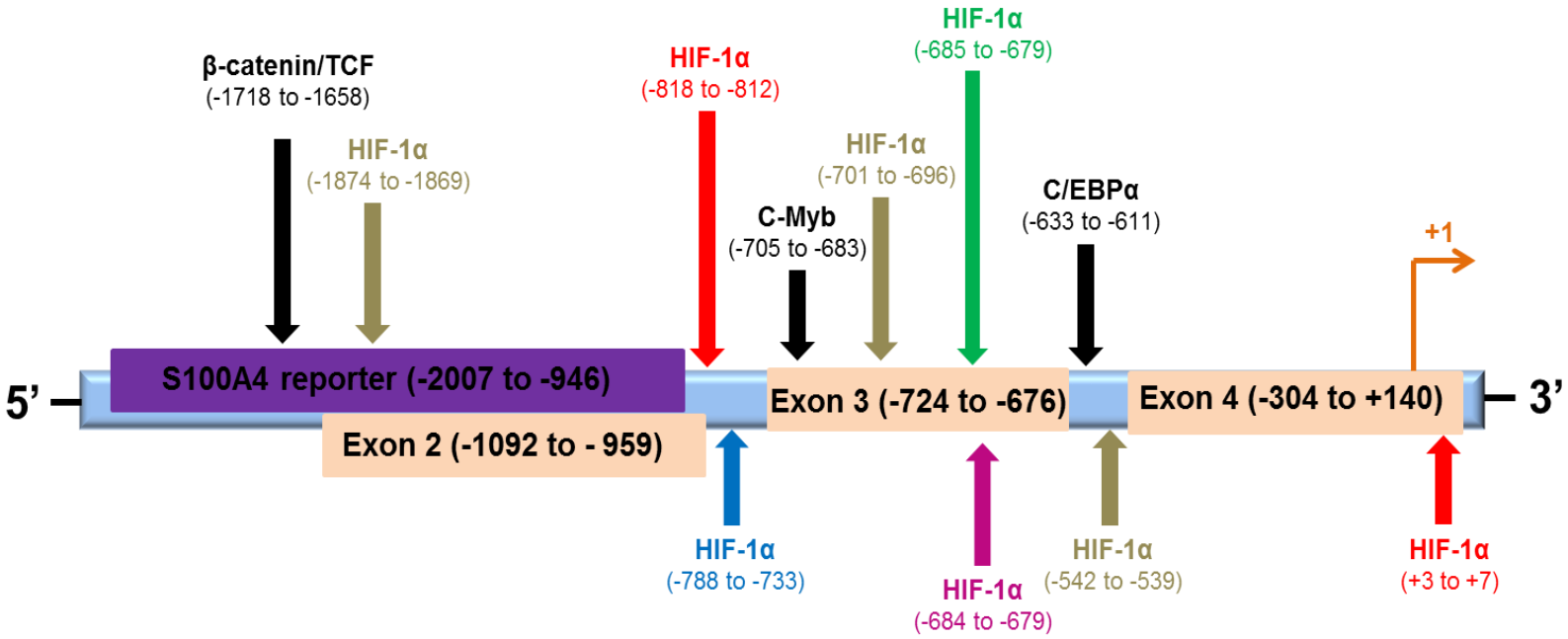


Figure 32. Diagrammatic representation of the human *S100A4* gene displaying the location of the S100A4 promoter and other relevant features. The blue box shows the gene of interest located at chromosome 1, location 153711315-153717838. The purple box show the sequence of the S100A4 promoter, cloned into the S100A4 reporter. The orange boxes demonstrate the different exons of the gene. Anything between the exons is introns. The orange arrow represents the translational start site/start codon (ATG), as identified in ensembl.org, where the +1 is "T" in the "ATG". The other arrows represent the location of TF binding sites, as identified in the literature. The locations of the TFs were calculated manually, having as reference the start codon. The HIF-1α sequences are colour coded i.e. each colour represents a different HRE sequence.

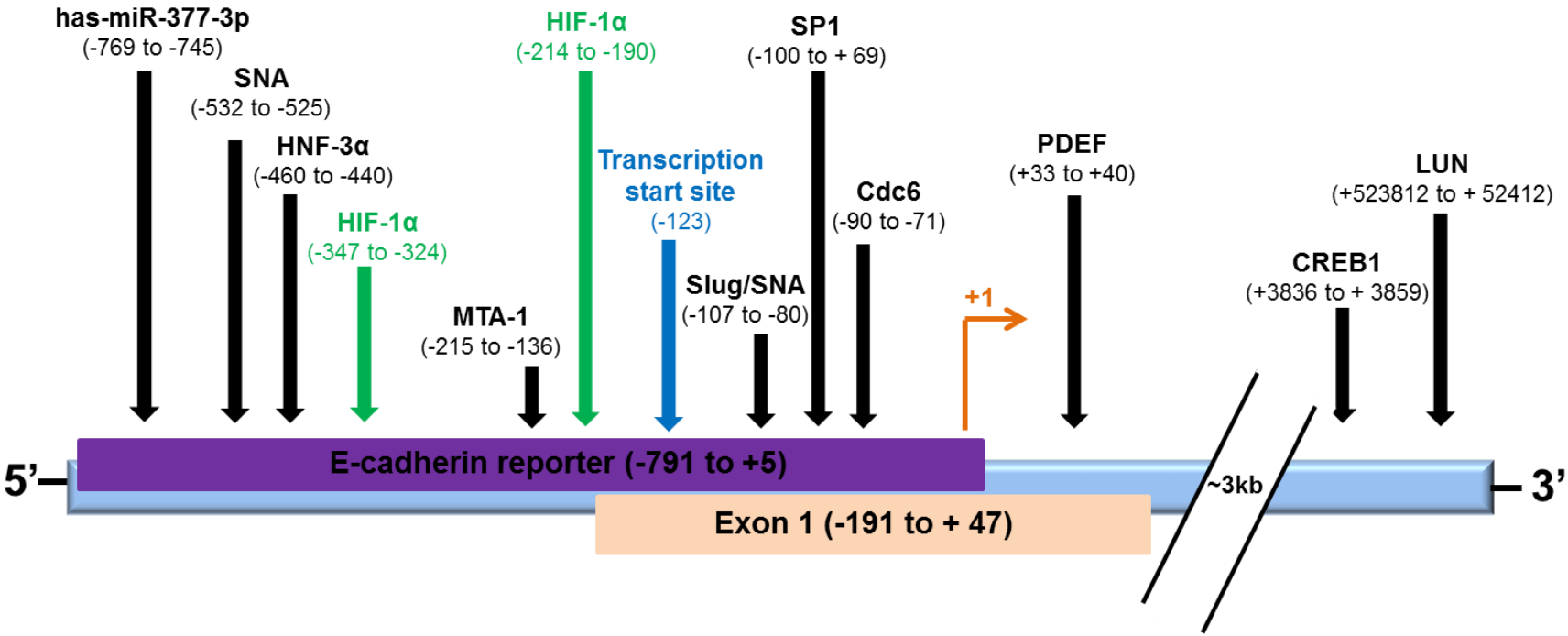


Figure 33. Diagrammatic representation of the human *CDH1* gene displaying the location of the promoter and other relevant features. The blue box shows the gene of interest located at chromosome 16, location 16:68737225-68835548:1. The purple box show the sequence of the *CDH1* promoter, cloned into the E-cadherin reporter. The orange box demonstrates exon 1 of the gene. The orange arrow represents the translational start site/start codon (ATG), as identified in ensembl.org, where the +1 is "T" in the "ATG". The blue arrow shows the transcription start site as identified in YN, L., et al., 2005. The other arrows represent the location of TF binding sites, as identified in the literature. The locations of the TFs were calculated manually, having as reference the start codon. The HIF-1α sequences are emphasised.

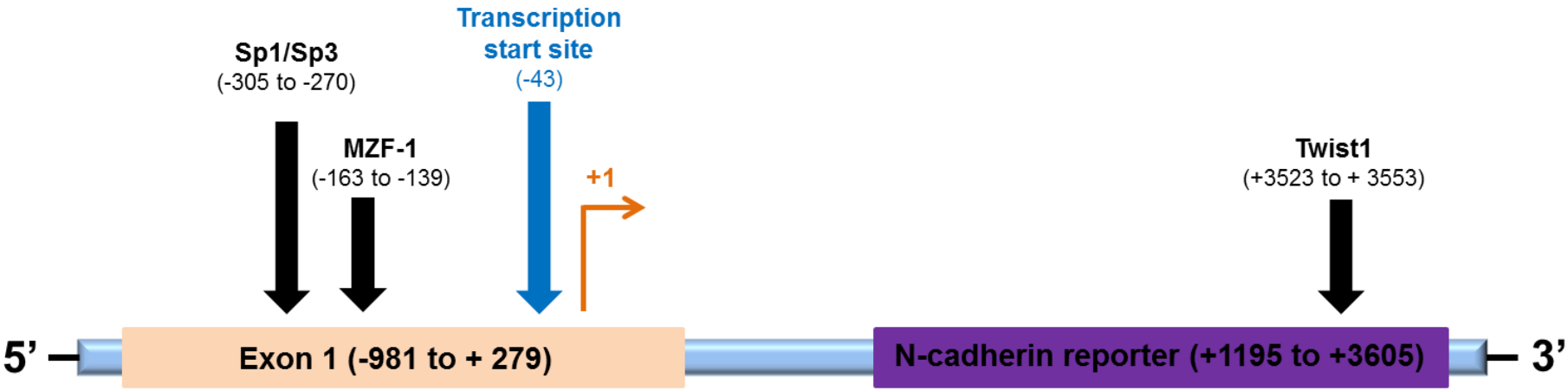


Figure 34. Diagrammatic representation of the human *CDH2* gene displaying the location of the promoter and other relevant features. The blue box shows the gene of interest located at chromosome 18, location 27,950,966-28,177,446. The purple box show the sequence of the *CDH2* promoter/regulatory sequence, cloned into the N-cadherin reporter. The orange box demonstrates exon 1 of the gene. The orange arrow represents the translational start site/start codon (ATG), as identified in ensembl.org, where the +1 is “T” in the “ATG”. The blue arrow shows the transcription start site as identified in S, L.M., F. O, and M. PJ, 2005. The other arrows represent the location of TF binding sites, as identified in the literature. The locations of the TFs were calculated manually, having as reference the start codon. The HIF-1α sequences are emphasised.

CHAPTER 4: RESULTS 2 – Development of an *in vitro* EMT model

This chapter describes the different methods tested to induce EMT (or EMT-like traits) in epithelial cell lines, in order to establish an *in vitro* EMT model, which could be used to test the activity of EMT reporters upon induction of EMT. Furthermore, the work in this chapter aimed to find a model, in which significant and consistent changes in the expression of E-cadherin and N-cadherin would be achieved, in order that it could be used for validating the E-cadherin and N-cadherin reporters.

The different methods used were an adaptation of the methods reported to induce EMT and/or enhance for metastatic and stem-like properties in the current literature.

Induction of EMT was assessed by analysing the gene expression or protein levels of EMT markers, and the morphological changes of the cells. Briefly, during EMT, cells undergo changes in their shape (acquire a spindle-like, “fibroblastic” shape), lose their expression of epithelial markers (e.g. E-cadherin), and acquire a mesenchymal phenotype (motility and expression of mesenchymal markers e.g. Vimentin) (reviewed in [46]). The morphology of the cells was assessed by microscopy, and the expression of EMT markers (attenuation/loss or upregulation) was analysed by qPCR (gene levels) and immunofluorescence (IF) staining (protein levels).

The first method used to induce EMT was the overexpression of Twist1. Twist1 is a main TF driving EMT, whose expression downregulates epithelial markers like E-cadherin and activates mesenchymal markers like N-cadherin (reviewed in [46]). Twist-expressing vectors have previously been reported to induce EMT in BC cells [52, 141, 155]. A plasmid construct allowing the simultaneous expression of Twist1 and a red fluorescent protein (tdTomato) was constructed by molecular cloning. The vector was designed to express the two proteins from a bicistronic mRNA transcript utilising an internal ribosome entry site (IRES) [156] that allowed tdTomato (downstream of TWIST1) to be used as an indicator of transduction efficiency and a marker for selection. MCF-7 cells expressing the Twist1 expression vector, or a control plasmid constitutively expressing tdTomato, were generated by lenti-viral

transductions and stable transductants were purified by Flow Cytometric Cell Sorting.

Induction of EMT was also attempted by culturing the cells in hypoxia. Cells are typically cultured in atmospheric oxygen levels, but when cultured in hypoxia, they have been shown to undergo EMT through a variety of mechanisms, including upregulation of Twist1, Snail and the Notch signalling pathway (reviewed in [157]). Furthermore hypoxia can help in stabilising the mesenchymal state of the cells by activating positive-feedback loops and by maintaining the expression of the various EMT-inducing TFs. 65-70% sub confluent MCF-7 cells cultured in 3% O_2 underwent morphological changes associated with EMT and lost the expression of E-cadherin after 72 hours [69]. In another instance, BC cells cultured in 1% O_2, lost their cell-cell contact, became spindle-like and acquired expression of the mesenchymal marker Vimentin after 50 hours [70].

Another method used was the culture of the cells in a commercially available EMT-Inducing Media Supplement, which contains a cocktail of EMT-inducing recombinant proteins and neutralizing antibodies for the straightforward induction of EMT. Induction and analysis of EMT in different cell lines using this product has been described [158]. The supplement is designed to enhance the Wnt and TGF-β signalling and to block E-cadherin-based cell adhesion. Loss of E-cadherin is a hallmark of EMT and both TGF and Wnt signalling pathway have been shown to induce EMT (reviewed in [46]). In addition, the collaboration of the TGFβ, canonical and non-canonical Wnt signalling to induce EMT in BC and maintain the acquired mesenchymal phenotype of the cells in an autocrine signalling, has been previously described [159]. The supplement specifically contains recombinant human Wnt-5α, recombinant human TGF-β1, anti-human E-cadherin, anti-human Secreted Frizzled Related Protein-1 (sFRP-1) and anti-human Dickkopf-1 (Dkk-1). Both sFRP-1 and Dkk-1 downregulate the Wnt signalling pathway; the former inhibits Wnt receptor binding [160] and the latter is a Wnt antagonist [161].

In order to induce EMT, cells were also cultured in a 3D assay. A 3D multicellular tumour growth assay (TGA) has been developed in order to mimic the *in vivo* scenario of tumours [82]. As opposed to a 2D monolayer culture, in which the cells are cultured in plastic, the 3D culture mimics different conditions present in human cancers, including cell-matrix interactions, stiffness, acidic pH, dimensionality, and secreted factors that affect cell function, behaviour and drug response. The cells were cultured in a laminin-rich Basement Membrane Extract (BME) gel, which was derived from murine chondrosarcoma that was grown in mice and therefore includes tumour matrix and mouse immune modulators. 3D cultures and 3D-related systems have been used for induction of EMT in BC cells, where it was shown that cells have undergone morphological changes and acquired mesenchymal markers [83, 162].

In a further attempt to induce EMT, the MCF-7 cells were co-cultured with human bone marrow-derived mesenchymal stem cells (MSCs). MSCs are non-haematopoietic cells that can self-renew and differentiate into different cell types (reviewed in [163]). Through their interaction with the cells, including BC cells, and the secretion of paracrine factors and cytokines, they have an effect both on the morphology and proliferation of the cells. MSCs are recruited by tumours and they have a significant role in tumour progression and metastasis. Co-cultures of bone marrow and human placenta-derived MSCs with BC epithelial cells induced traits associated with EMT [73, 84]. In this study, MSCs expressing a red fluorescent protein (mCherry) were used. EMT was assessed by IF staining of E-cadherin and Vimentin.

It has been shown that the cytokine Interleukin-6 (IL-6) induced EMT in BC [164], therefore culture of cells in a medium supplemented with IL-6 was used in this study to induce EMT. Elevated levels of IL-6 have been shown in BC patients and it correlated with disease stage, metastasis and poor prognosis. The authors showed induction of EMT in MCF-7 cells exposed to 50ng IL-6 in a 3D system and in MCF-7 cells constitutively expressing ectopic IL-6. In a different paper, IL-6-induced EMT promoted the generation of BC stem-like cells [165].

Another method used is the indirect co-culture of epithelial cells with BC associated fibroblasts (BrCAFs), in which the epithelial cells were seeded in a well as normally and the BrCAFs were seeded in transwell insert. The insert contains a microporous membrane that will allow secreted factors to migrate and interact with the epithelial cells. CAFs, a distinct cell type, are found in the tumour stroma and facilitate tumour progression through secretion of inflammatory cytokines, growth factors and proteinases [166]. It has been shown that MCF-7 and MCF10A cells have undergone EMT, after being exposed to CAF-conditioned medium, as assessed by increased expression of Vimentin, and enhanced invasive abilities [167]. Furthermore, MCF-7 cells cultured in CAF-conditioned medium have undergone EMT, which was dependent on the secretion of IL6 [168].

Finally, some of the methods were used in combination, for example 3D cultures were performed in hypoxia.

4.1 Overexpression of Twist1

4.1.1 Molecular cloning of the Twist1 expression vector

TWIST1 was amplified from plasmid DNA in order to construct a lentiviral vector that allows constitutive dual expression of Twist1 and tdTomato. The sequence covered exon 1 and part of exon 2 of the gene, which gives rise to the full transcript which in turns translates to a 202 amino acid protein. *TWIST1* (1492bp) was amplified from the pOTB7-Twist1 plasmid (Figure 6) using the HotStarTaq® DNA polymerase (Figure 34). The annealing temperatures used ranged from 50.2°C to 68.3°C; the product was successfully amplified at 50.2°C and 50.8°C. The primers used for the amplification, were designed to introduce restriction sites for *EcoRI* and *BamHI* at the 5' and 3' end of *TWIST1* accordingly. The amplified product (indicated by arrow; Figure 34) was purified and cloned into the pCR4-TOPO vector (Figure 9), in order to confirm its sequence before

proceeding to generating the expression vector. Following transformation of the TOPO vectors, plasmids were isolated. In order to select those clones with the gene incorporated, diagnostic digests were performed. The plasmids were digested with *EcoRI* and *BamHI* (data not shown) and it was shown that all the clones had incorporated *TWIST1.* This was confirmed by DNA sequencing. A clone with no mutations introduced was selected for generating the expression vector. *TWIST1* was then digested from the TOPO vector with EcoRI/BamHI. pLVX-IRES-tdTomato (Figure 10) was also digested using the same enzymes, so that *TWIST1* would be cloned in to the Multiple Cloning Site (MCS) of this plasmid. The digested *TWIST1* and vector were gel extracted and ligated. The ligation products were transformed into DH5α and plasmids were isolated. Following diagnostic digests of the clones with *EcoRI* and *BamHI* (Figure 35), the clone with the insert incorporated (Indicated by arrow; Figure 35) was selected for DNA sequencing. The analysis the DNA sequence showed that the Twist1 reporter was successfully constructed (Figure 36). For the full DNA sequence, see appendix 5.

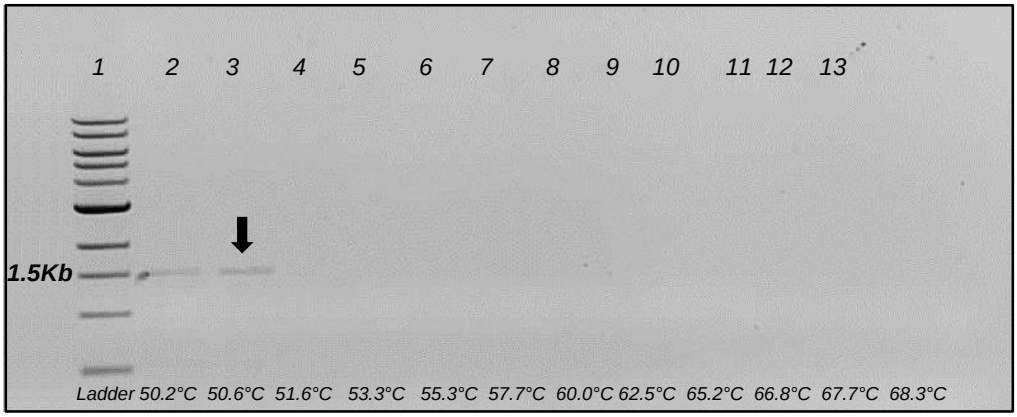


Figure 35. Amplification of *TWIST1* with HotStarTaq® DNA Polymerase. Lane 1=1 kb DNA ladder. Lanes 2-13=Samples at different annealing temperatures (Gradient: 50.2°C – 68.3°C). Size of the product=1492 bases. Agarose gel=0.8%.

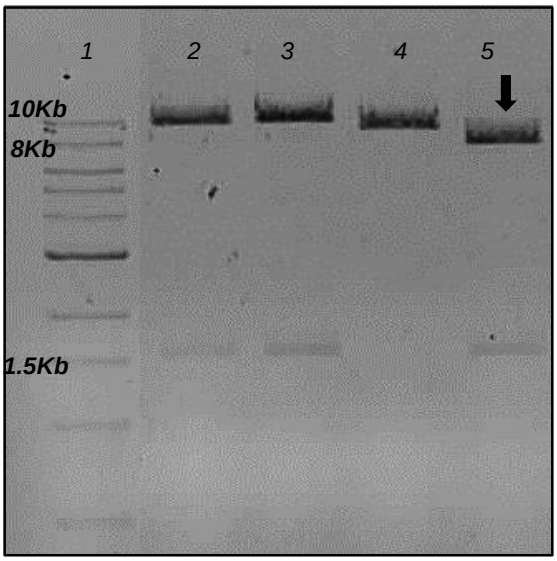


Figure 36. Diagnostic digests of plasmids with *EcoRI* and *BamHI*. Lane 1=kb DNA ladder. Lanes 2-5=digested 'pLVX-Twist1-IRES-tdTomato' (Twist1 expression vector) clones. Expected size of fragments of the clones with the insert incorporated=8860+1499. Gel=0.8%.

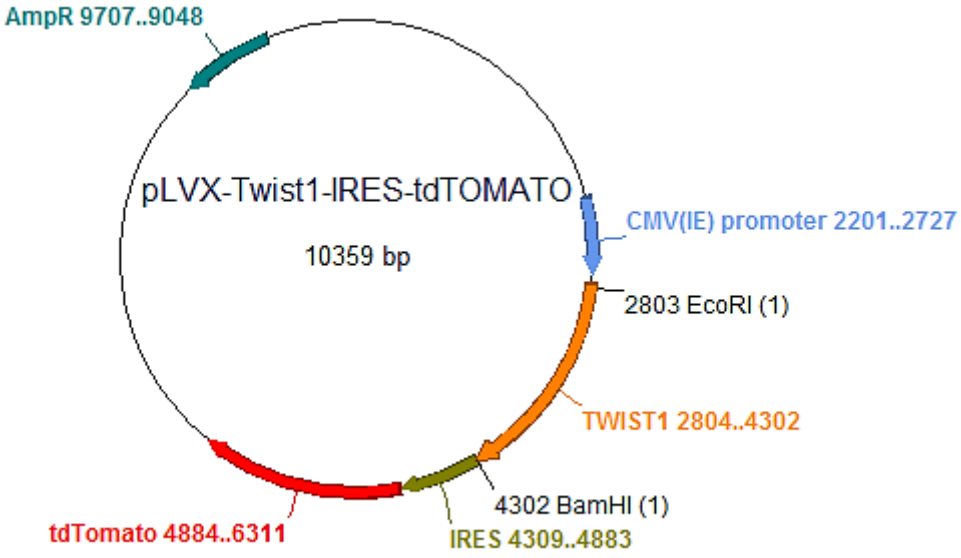


Figure 37. pLVX-Twist1-IRES-tdTomato plasmid construct (Twist1 expression vector). This construct is based on 'pLVX-IRES-tdTomato' (Figure 11). Expression of Twist1 is constitutively driven by the CMV promoter. It was made by removing the inserting *EcoRI*/*BamHI* '*TWIST1*' into the MCS of *EcoRI*/*BamHI* 'pLVX-IRES-tdTomato'.

4.1.2 Stable cell lines expressing the vector

The Twist1 expression vector was packaged into lentiviral particles and MCF-7 cells expressing the Twist1 reporter or a control vector (pLVX-CMV-tdTomato) were generated in-house by lenti-viral transductions, using the Lenti-X™ Lentiviral Expression system from Clontech. Stable transductants were selected by Flow Cytometric Cell Sorting (Beckman Coulter MoFlo XDP Cell Sorter). The sorting efficiencies ranged from 90-95%. Pictures of the cells after sorting are shown in Figure 37. To confirm that the Twist1 was successfully overexpressed in the stable cell lines and that the expression was maintained, the gene expression and protein levels of Twist1 were analysed at different cell passage numbers. RNA was isolated from the transduced cell lines, cDNA was synthesised and the gene expression of *TWIST1* was analysed by qPCR. It was shown that Twist1 was significantly up-regulated in the cells expressing the Twist1 vector compared to the control cell lines at the gene level, as assessed by a ratio paired t-test (Figure 38). The average gene expression in the cells expressing the Twist1 vector was about 560 times higher than the expression of *TWIST1* in the cells expressing the control vector.

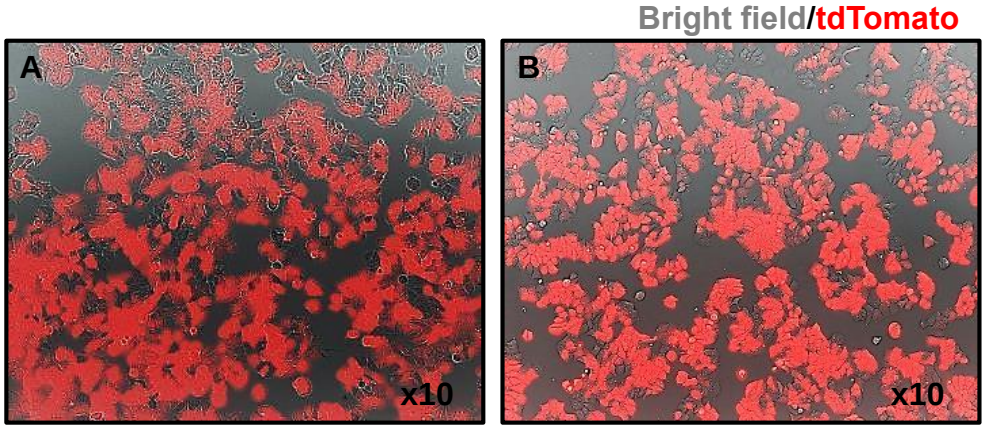


Figure 38. Pictures of transduced MCF-7 cells. Pictures showing the cells expressing the control vector, pLVX-CMV-tdTomato **(A)**, or the Twist1 expression vector, pLVX-Twist1-IRES-tdTom **(B).** Pictures are a merged image of pictures taken in brightfield and with DsRed filter. The red fluorescence (tdTomato) is emitted from the vectors and represents the cells that have successfully uptaken the vectors. Pictures were taken at x10 magnification.

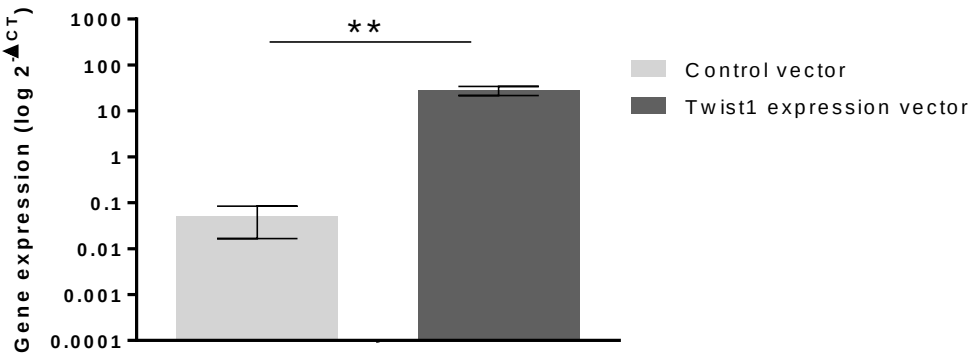


Figure 39. Gene expression of *TWIST1* in MCF-7 stably expressing the Twist1 expression vector or a control vector. Each bar represents the mean gene expression ($2^{-\Delta CT}$) ± SEM on a logarithmic scale. Statistical significance was assessed with a ratio paired t-test. *P≤0.05 and **P≤0.01. N=4.

To test whether the protein was overexpressed, cell lysates were collected and the protein expression was analysed by western blot. Cell lysates from the BT549 and MSCs cells were included in the analysis to serve as a positive control for Twist1 expression, as both cell lines have been shown to express Twist1, at least at the gene levels, in previous in-house experiments. B-actin was used as a reference/ loading protein. In the western blot gel, multiple bands were detected. To overcome this, different optimisations have been attempted including use of fresh lysates from earlier passaged cells, use of fresh ingredients (e.g. fresh reducing agent), use of different antibodies, different concentrations of the antibody and longer incubation times, but in all of the experiments the results were similar. One representative example is shown in Figure 39. The expected size of the protein was 21kDa. Single bands sized ~40kDa were detected in the MSCs and BT549 (Figure 39 lanes A&B). In the MCF-7 control cells (cells transduced with the control vector) three main bands were detected sized, 14kDa, ~21kDa and ~40kDa (Figure 39 lane C). The same sized bands were detected in the MCF-7 cells (Figure 39 lane D), where the ~40kDa band appeared to be much more

concentrated in these cells, compared to the control cells. In addition, extra bands of >40kDa were detected in these cells.

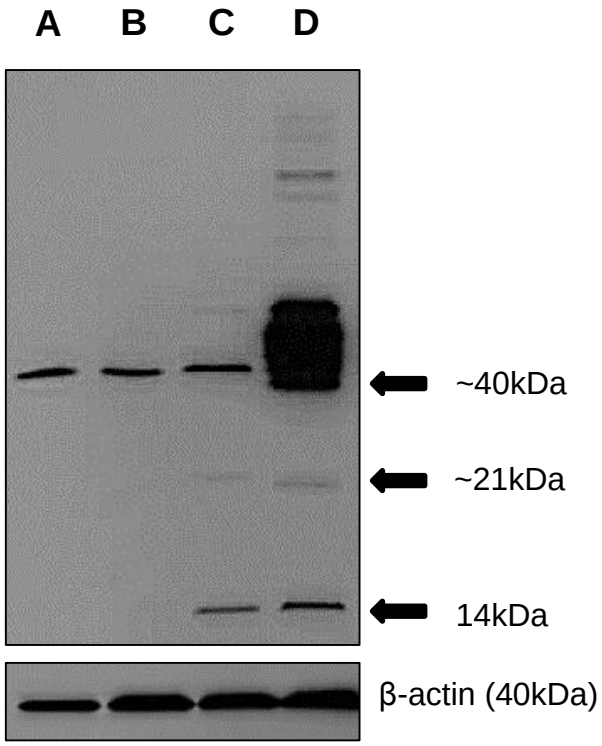


Figure 40. Western blot analysis of Twist1 in MCF-7 cells expressing the Twist1 expression vector. The different cells, MSCs (A), BT549 (B), MCF-7 control (C) and MCF-7/ Twist1(D) were stained with anti-human Twist1 (21kDa). The arrows indicate the size of the bands. B-actin (40kDa) was used as a reference protein.

4.1.3 Analysis of EMT-like traits in cell lines expressing the vector

Even though, the protein level of Twist1 was not determined in the cells, EMT-like traits were examined. No changes in the morphology of the cells were observed between the cells expressing the control vector and the Twist1 vector. Furthermore, while the *TWIST1* expression was upregulated (average fold change 3500), no changes in the gene expression of the EMT markers were observed (Figure 40).

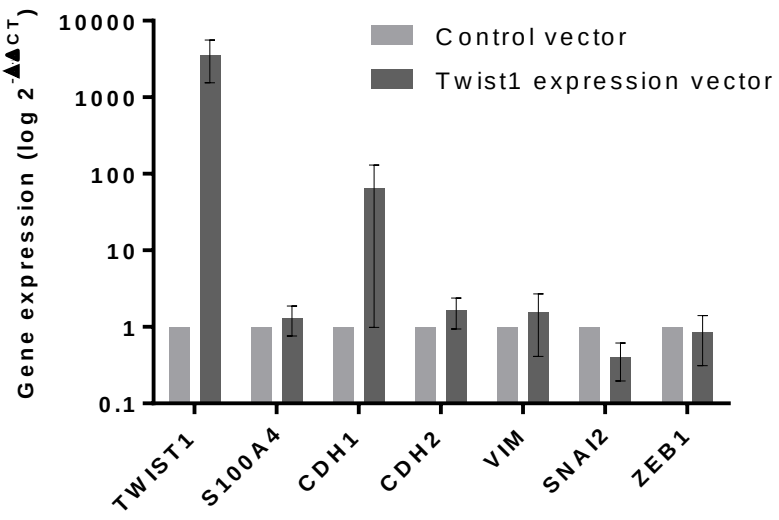


Figure 41. Relative gene expression of EMT markers in MCF-7 cells expressing the control vector or the Twist1 expression vector. Each bar represents the relative mean gene expression ($2^{-\Delta\Delta CT}$) ± SEM of matched biological replicates, on a logarithmic scale.

4.2 Hypoxia

Cells were seeded in a 24-well plate and cultured in a hypoxia chamber (1% O_2) or in normoxia (21% O_2) for four days. At the end of the assay the morphology of the cells was assessed by microscopy and the cells were collected for gene expression analysis of the EMT markers. The experiment was repeated three times. No obvious morphological changes associated with EMT were observed in the NCI60 MCF-7 cells (a representative example is shown in Figure 41). In these cells *VEGF* was upregulated in all biological replicates (average fold increase 3.2), implying a degree of hypoxia (Figure 42). The upregulation of *VEGF* was statistically significant as assessed by a ratio paired t-test (Figure 43A). *S100A4* was also upregulated in all biological replicates (average fold increase 9.76; Figure 42), but the difference was not statistically significant, as assessed by a ratio paired-test (Figure 43B). *VIM* was also upregulated in all biological replicates (average fold increase 1.2; Figure

42). While *VIM* was significantly upregulated (Figure 43C), the changes in the actual $2^{-\Delta CT}$ values were relatively small to be considered substantial (i.e. from 0.008 to 0.009, from 0.009 to 0.011 and from 0.01 to 0.013). No consistent changes were observed in the other EMT genes tested (Figure 42).

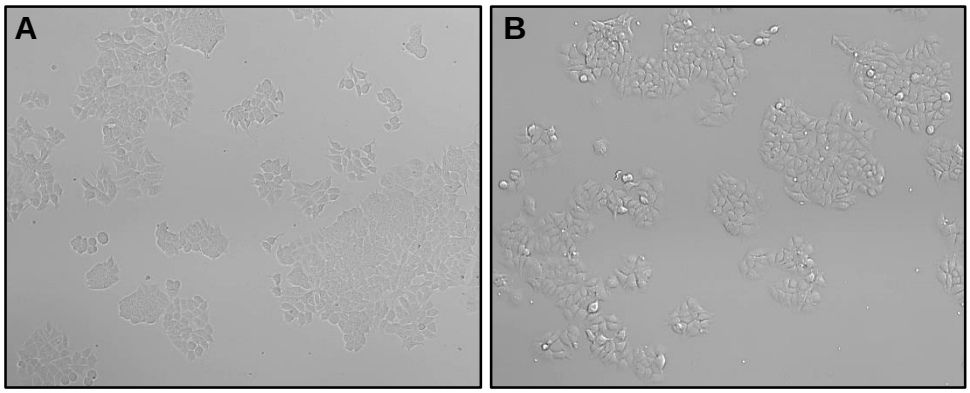


Figure 42. Pictures of MCF-7 cells cultured in normoxia or hypoxia. **A.** MCF-7 (NCI60) normoxia. **B.** MCF-7 (NCI60) hypoxia. Pictures were taken at x10 magnification.

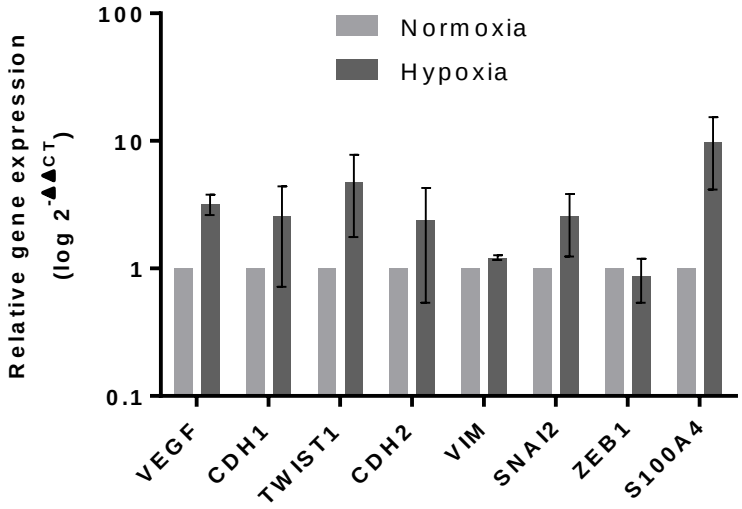


Figure 43. Relative gene expression of EMT markers in hypoxic assay of MCF-7 cells. Each bar represents the mean relative gene expression ($2^{-\Delta\Delta CT}$) ± SEM of matched biological replicates on a logarithmic scale. N=3. Pictures were taken at x10 magnification.

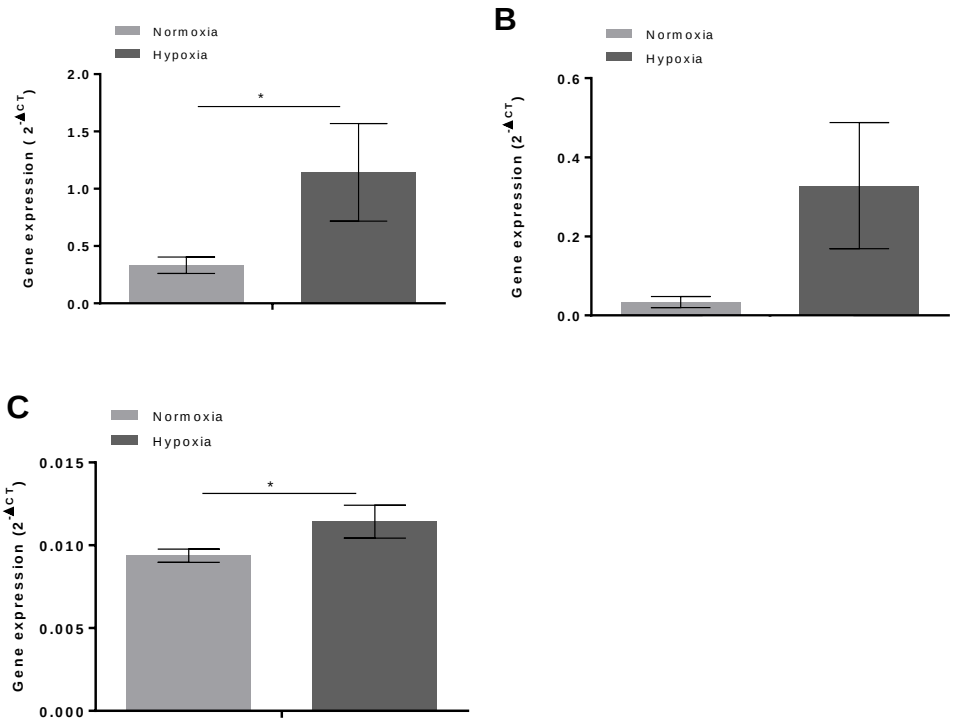


Figure 44. Gene expression of *VEGF*, *S100A4* and *VIM* in hypoxic conditions. Each bar represents the mean gene expression ($2^{-\Delta CT}$) ± SEM of matched biological replicates for *VEGF* (**A**), *S100A4* (**B**) and *VIM* (**C**). Significance was assessed by a ratio paired t-test. *P≤0.05. N=3.

4.3 EMT-inducing media supplement

Cells were cultured in their normal medium or in a medium containing the commercially available EMT-inducing media supplement. The recommended cell density (0.9 - 1.0 x 10^4 cells/cm^2) resulted in cells being over-confluent at the end of the assay; therefore, the cell density was optimised (5x10^3 cells per dish). In this experiment, the MCF-7 cells from NCI60 and ECACC were included. The latter were included in the study at later stages, in order to investigate the EMT potential of different MCF-7 cells, and use as alternatives.

The MCF-7 cells stably expressing the Twist1 vector (MCF-7/Twist1) were also included. It was shown before, that even though the *TWIST1* gene was overexpressed in these cell lines, EMT was not induced. Since,

it was not possible to determine for certain whether Twist1 protein was overexpressed and/or activated, it was assumed that the protein was inactive and that was the reason why EMT was not induced. It has been shown, that TGFβ phosphorylates Twist1 on Ser^{68}, which is required for EMT induction [169] and that Wnt-1 upregulates Twist1 [170]. Since the EMT inducing media supplement was designed to induce the Wnt and TGF-β signalling, these cells were included in this experiment in an attempt to phosphorylate/activate Twist1.

At the end of the experiment, the morphology of the cells was assessed by microscopy and the cells were collected for gene expression analysis of the EMT markers. The experiment was repeated 3 times for the NCI60 MCF-7 cells, 2 times for the ECACC MCF-7 cells and 1 time for the MCF-7/Twist1 cells. For the MCF-7/Twist1 cells, half of the cells were collected for gene expression analysis and half of the cells were lysed for western blot. No morphological changes associated with EMT were observed in either of the cell lines tested (a representative example is shown in Figure 44). In the NCI60 MCF-7 cells, a small downregulation of *CDH1* was observed (average fold change 0.82), but no consistent changes were observed in the mesenchymal genes (Figure 45A). Upregulation of *S100A4*, *CDH2* and *SNAI* was observed in the ECACC MCF-7 cells (average fold change 4.5, 12.4 and 5.2 respectively; Figure 45B). As no morphological changes and no changes in the epithelial gene *CDH1* were observed in these cells, the experiment was not pursued further. In the MCF-7/Twist1 cells, *TWIST1* was not upregulated (Figure 45C). The proteins from the two different samples appeared exactly the same on a western blot, suggesting that the media supplement did not have an effect on the expression of Twist1 (Figure 46). A small downregulation of *CDH1* was observed in these cells (fold change 0.64), but the mesenchymal markers were either downregulated or retained similar levels of expression (Figure 45C). Since no changes were observed in the expression of Twist1 (gene and protein levels), or the morphology of the cells, or the mesenchymal markers tested, the experiment using these cells was not pursued further.

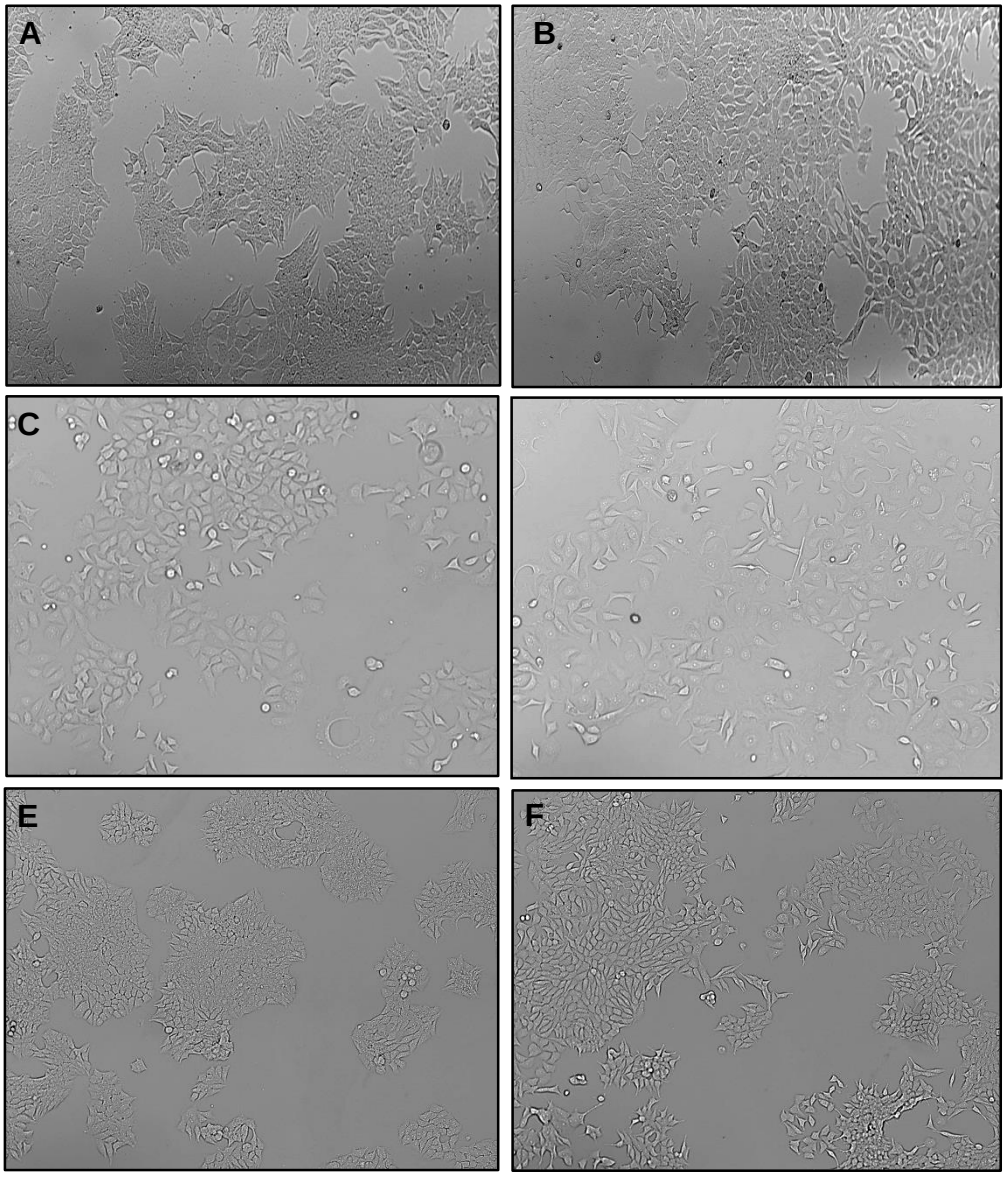


Figure 45. Pictures of cells cultured in EMT inducing media supplement. NCI60 MCF-7 cells without supplement **(A)** and with supplement **(B).** ECACC MCF-7 cells without supplement **(C)** and with supplement **(D).** MCF-7/Twist1 cells without supplement **(E)** and with supplement **(F)**. Pictures were taken at x10 magnification.

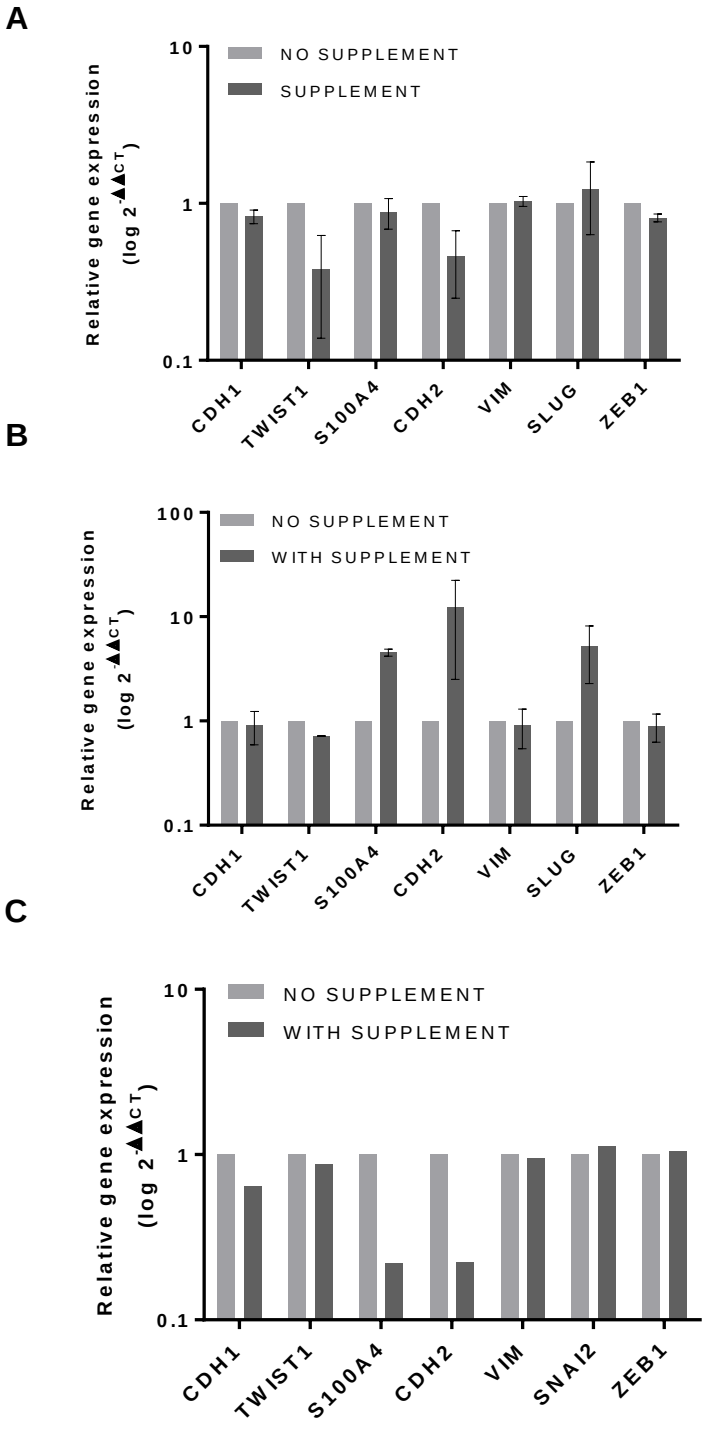


Figure 46. Relative gene expression of EMT markers in cells cultured in a medium containing the EMT-inducing media supplement. The graphs show the analysis of EMT markers in NCI60 MCF-7 (n=3) **(A)**, ECACC MCF-7 (n=2) **(B)** and MCF-7/Twist1 (n=1) **(C)** cells. Each bar represents mean $\mathbf{2}^{-\Delta\Delta CT}$ values ± SEM of matched biological replicates on a logarithmic scale.

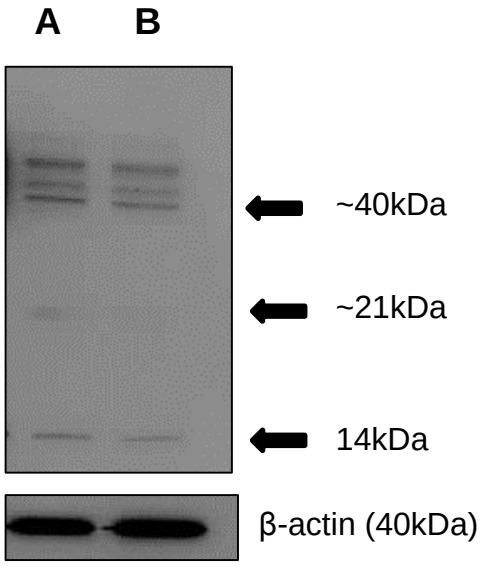


Figure 47. Western blot analysis of Twist1 in MCF-7/Twist1 cells, cultured with EMT-inducing media supplement. The cells cultured without the supplement (A) or with the supplement (B), were stained with anti-human Twist1 (21kDa). The arrows indicate the size of the bands. B-actin (40kDa) was used as a reference protein.

4.4 3D cultures

The NCI60 MCF-7 cells were seeded a 96-well plate in quadruplicate and cultured in a 3D TGA for up to 7 days both in normoxia and hypoxia. The experiment was performed once. The cell density, the concentration of the BME, and the length of the assay had been previously determined in-house. The growth of the cells was monitored every 1-2 days with the Alamar Blue (AB) cell viability assay. As shown in Figure 47, the cell number increased over the first six days of the assay and remained stable/started declining by the end of the assay. Furthermore, the growth rate of the cells was similar in normoxia and hypoxia.

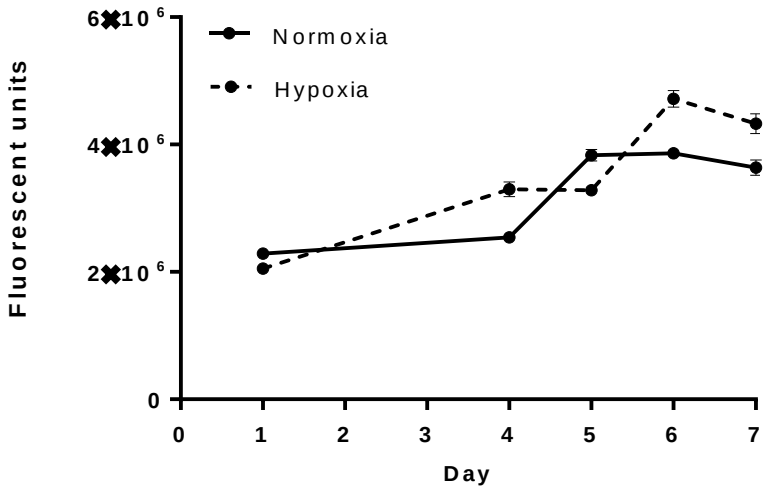


Figure 48. Growth curves of cells cultured in a 3D TGA assay in normoxia and hypoxia. Each point shows the average fluorescent units, as detected by AB ± SEM. N=1.

The gene expression of the EMT markers and the hypoxia marker (*VEGF*) at the end of the assay (Day 7) was compared with the gene expression at the first day of the assay (Day 1). *VEGF* was upregulated both in normoxia and hypoxia (fold change 2.9 and 29.7 respectively), suggesting a degree of hypoxia in both conditions (Figure 48). *VIM* was upregulated in normoxia (fold change 11.1) and *S100A4* was upregulated in hypoxia (fold change 30.3) (Figure 48). Even though, upregulation of the mesenchymal markers was observed, upregulation of the epithelial marker *CDH1* was also observed both in normoxia and hypoxia (fold change 3 and 1.9 respectively) therefore, the experiment was not pursued further.

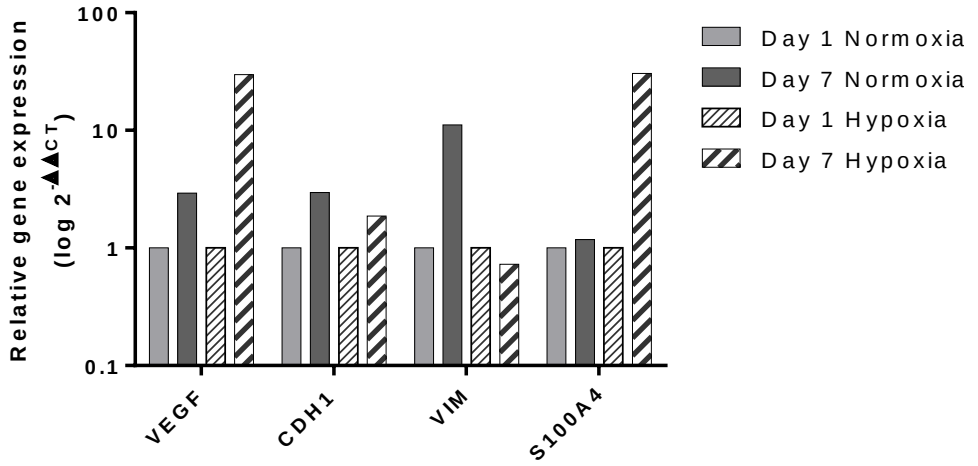


Figure 49. Relative gene expression of EMT markers in cells cultured in a 3D TGA assay in normoxia and hypoxia. Each bar shows the $2^{-\Delta\Delta CT}$ value. N=1.

4.5 Co-cultures with MSCs

NCI60 MCF-7 cells were cultured with (co-culture) or without (mono-culture) mCherry MSCs on coverslips in 6-well plates for 10 days, at a ratio of 2MCF-7:1MSCs. The cell density and length of the assay was optimised, so that the cells reached ~80% confluence at the end of the assay (suitable for immunofluorescence (IF) staining). At the end of the experiment the morphology of the cells was assessed by microscopy and the protein expression of E-cadherin and Vimentin was assessed by IF. In this experiment, the changes in the EMT markers was assessed at the protein level (by IF) rather than by gene expression, in order to avoid separating the cells and risking ending up with a cell population that is not completely pure.

The appropriate controls were included to control for nonspecific, background and autofluorescence, and the pictures were taken at the same exposure times for each dye.

MCF-7 cells cultured in mono-cultures expressed high levels of membranous E-cadherin (Figure 49A&B). MCF-7 cells co-cultured with hMSCs also expressed E-cadherin (Figure 49C&E), showing that the expression of E-cadherin was not downregulated as it would be expected

upon EMT induction. While high expression of Vimentin was observed in the mCherry MSCs, no expression/upregulation of Vimentin was observed in the MCF-7 cells co-cultured with MSCs (Figure 50). Furthermore, from both pictures, it can be seen that the morphology of the MCF-7 cells remained epithelial-like. In summary, EMT-related traits were not induced in this setting. The experiment was performed twice; the results were the same at both times. For all the controls and individual images, see appendices 6-9.

Dapi/FITC (E-cadherin)/ HcRed (mCherry MSCs)

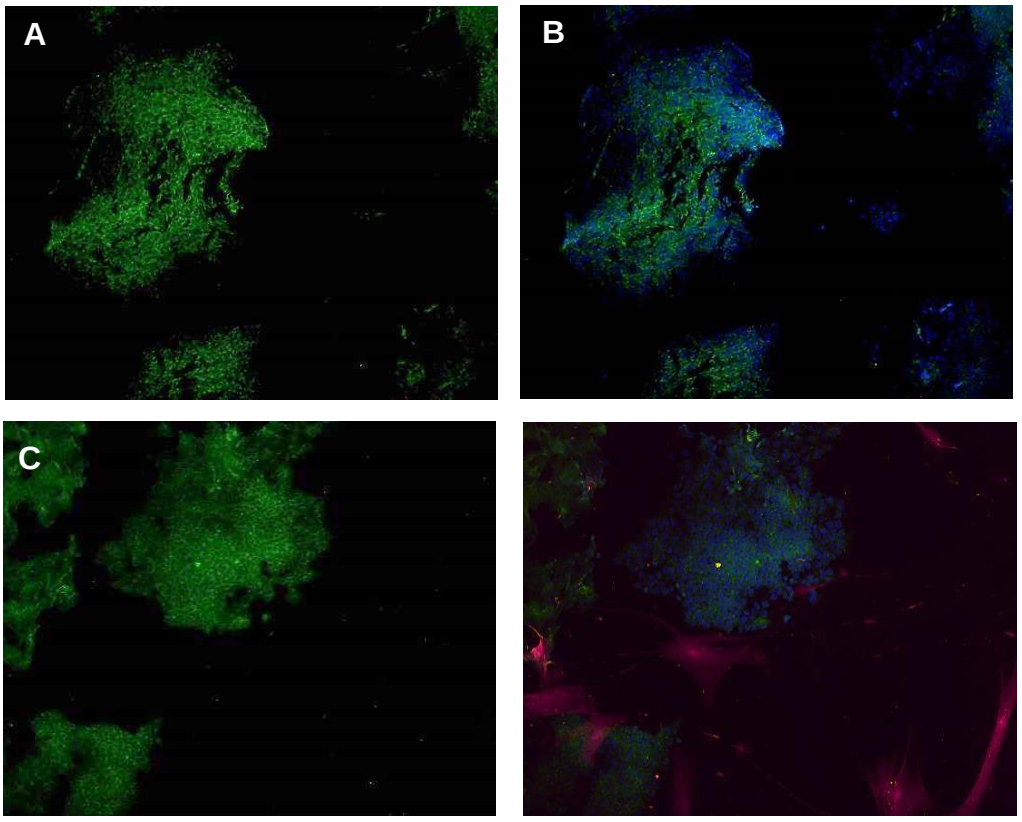


Figure 50. IF staining of co-cultures for E-cadherin. (A) E-cadherin staining in mono-cultures. **(B)** Merge picture of staining in mono-cultures. (**C**) E-cadherin staining in co-cultures. **(D)** Merge image of staining in co-cultures Blue (Dapi) = nuclei, green (FITC) = E-cadherin, and pink (HcRed) = mCherry MSCs. Pictures were taken at x10 magnification.

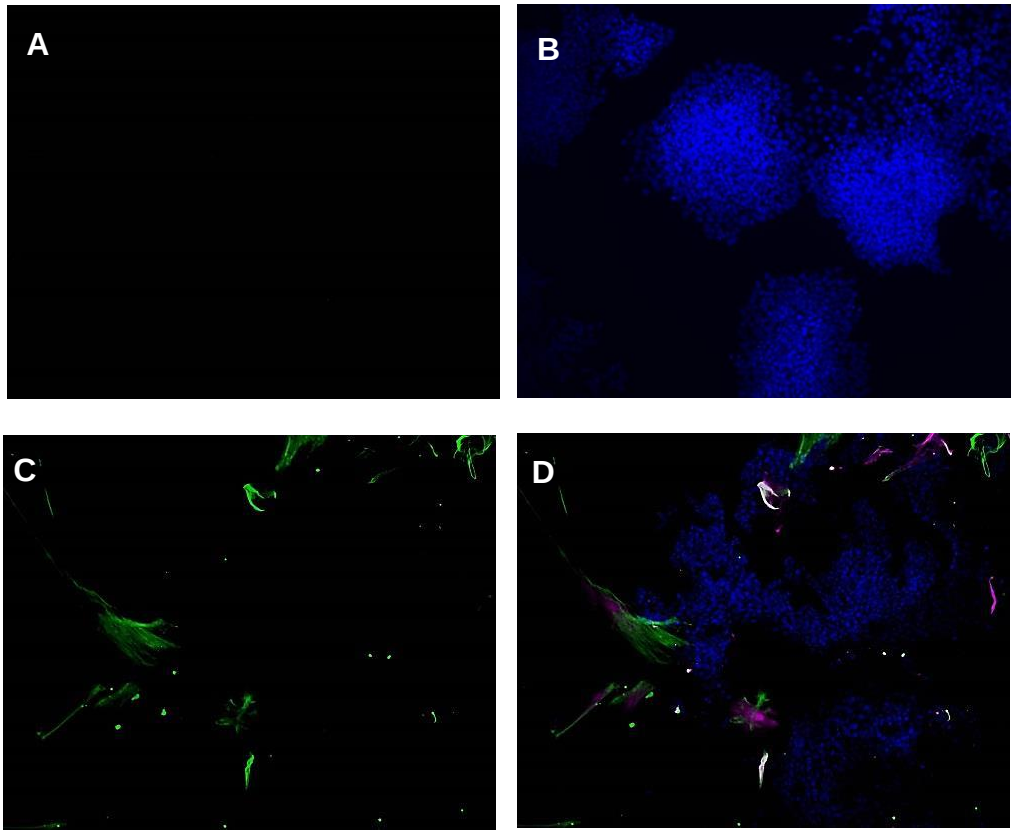


Figure 51. IF staining of co-cultures for Vimentin. (A) Vimentin staining in mono-cultures. **(B)** Merge picture of staining in mono-cultures. **(C)** Vimentin staining in co-cultures. **(D)** Merge picture of staining in co-cultures. Blue (Dapi) = nuclei, green (FITC) = E-cadherin, and pink (HcRed) = mCherry MSCs. Pictures were taken at x10 magnification.

4.6 Supplementation with IL-6

Cells were seeded at low densities and cultured in normal medium or medium supplemented with 50ng/ml human recombinant IL-6 for up to 7 days. The experiment was performed twice (the gene expression of EMT markers was analysed once, while the morphology of the cells twice). The medium, with or without IL-6, was replenished every 2 or 3 days. No morphological changes were observed in either the NCI60 MCF-7 or the ECACC MCF-7 cells (representative examples are shown in Figure 51).

TWIST1 and *CDH2* were upregulated in the NCI60 MCF-7 cells (fold change 6.9 and 2.6 respectively; Figure 52A). However, the difference in the actual $2^{-\Delta CT}$ values was small to be considered substantial; therefore, the experiment was not pursued further (i.e. from 0.0042 in normoxia to 0.0291 in hypoxia for *TWIST1* and 0.00007 to 0.0002 for *CDH2*). The rest of the EMT markers did not change. In the ECACC MCF-7 cells, *S100A4* was slightly upregulated (fold change 2.5), but the other markers remained unchanged (Figure 52B).

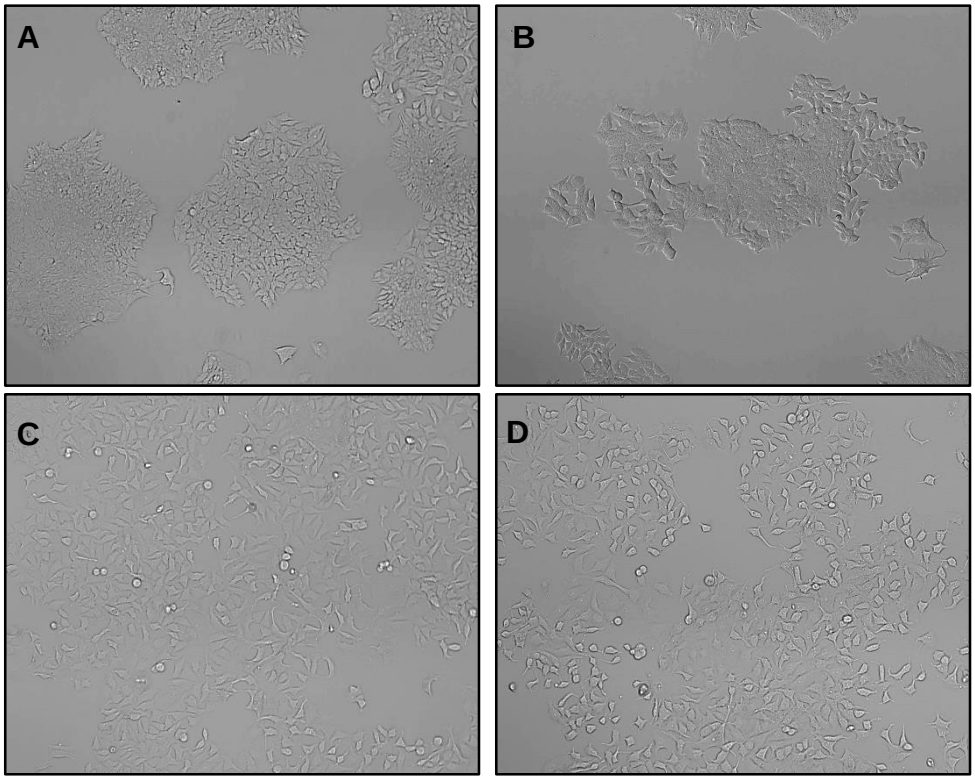


Figure 52. Pictures of cells cultured in medium supplemented with IL6. A. NCI60 MCF-7 cells no IL6. **B.** NCI60 MCF-7 cells with IL6**. C.** ECACC MCF-7 cells no IL6. **D.** ECACC MCF-7 cells with IL6. Pictures were taken at x10 magnification. Pictures were taken at x10 magnification.

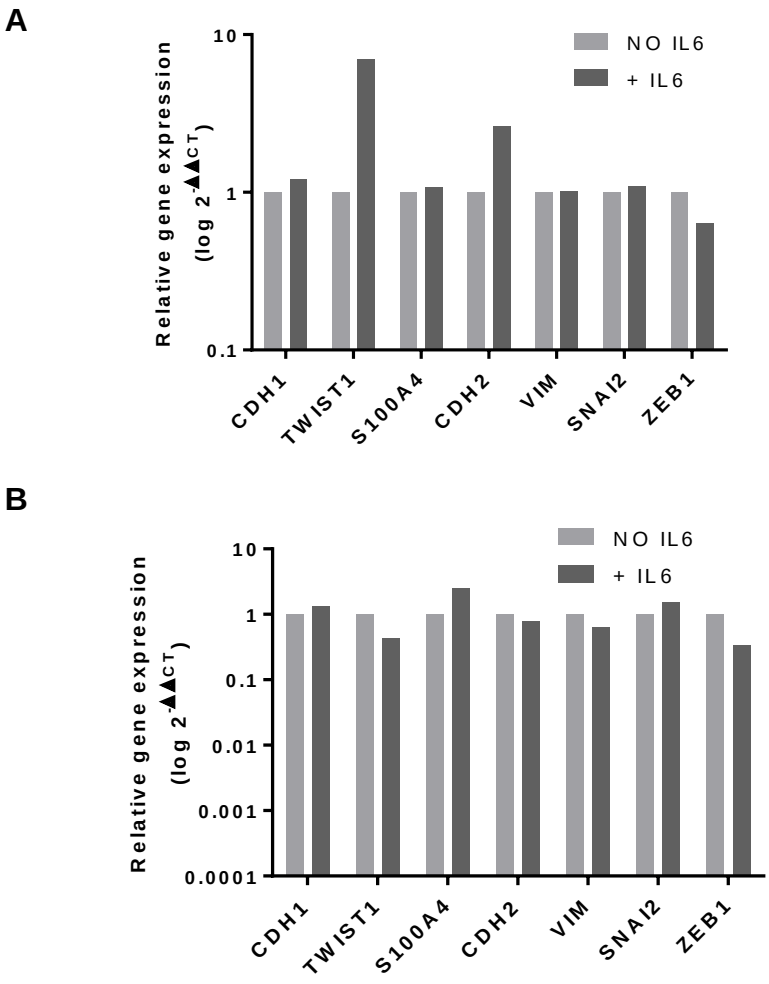


Figure 53. Relative gene expression of EMT markers in cells cultured with IL-6. Gene expression analysis in NCI60 MCF-7 cells **(A)** and in ECACC MCF-7 cells **(B).** Each bar represents the relative $\mathbf{2^{-\Delta\Delta CT}}$ values. N=1.

4.7 Indirect co-cultures with BrCAFs

NCI60 MCF-7 cells were seeded in a trans-well 24-well plate and either MCF-7 or BrCAFs were seeded in the insert in the appropriate culture medium. While the number of the MCF-7 cells in the well was kept constant, the number of cells in the insert was altered so that different cell ratios were used. The ratios used included the following: (a) 1MCF-

7:1MCF-7, (b) 1MCF-7:1BrCAFs, (c) 1MCF-7:3MCF-7 and (d) 1MCF-7:3BrCAFs. The experiment was performed twice (the gene expression of EMT markers was analysed once, while the morphology of the cells twice). The cell density and morphology of the cells was monitored daily by microscopy and at day 10, the cells were collected for analysis of the EMT genes. No obvious morphological changes associated with EMT were observed (representative images are shown in Figure 53) and no substantial differences in the expression of the EMT markers were detected (Figure 54).

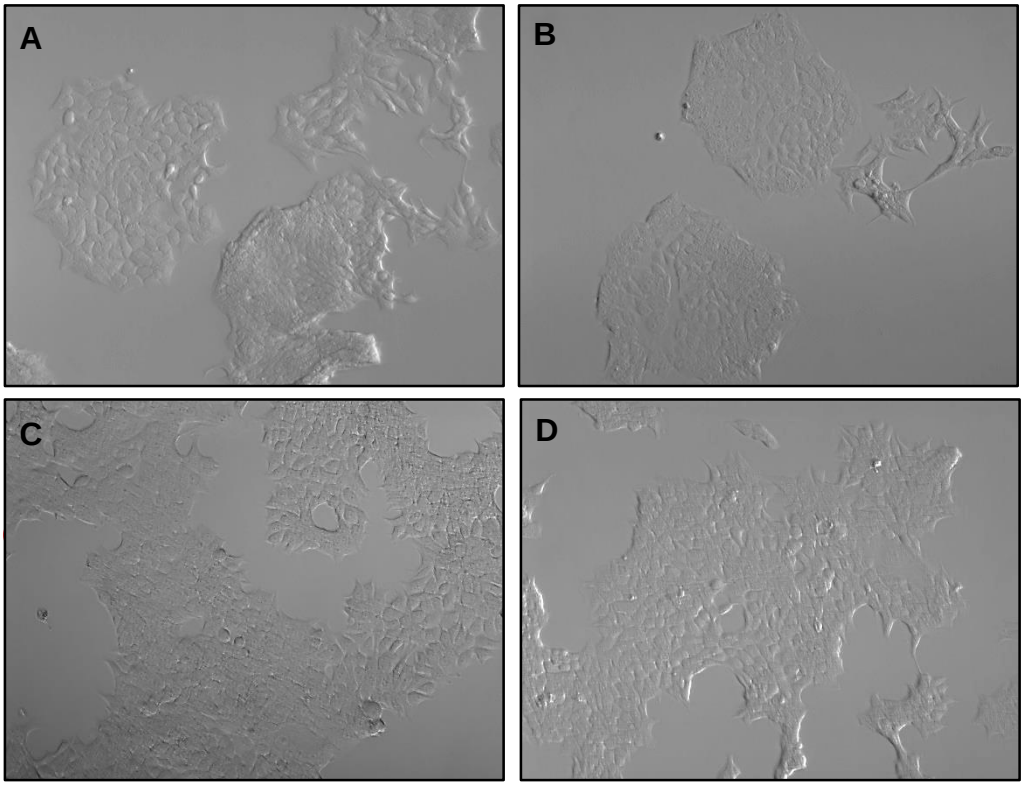


Figure 54. Pictures of MCF-7 cells cultured in trans-well plates with BrCAFs. The pictures show the MCF-7 cells in the well cultured with MCF-7 cells in the insert at 1MCF-7:1MCF-7 **(A)** or 1MCF-7:3MCF-7 **(B)** ratios or with BrCAFs in the insert at 1MCF-7:1BrCAFs **(C)** or 1MCF-7:3BrCAFs **(D)** ratios. Pictures were taken at x10 magnification.

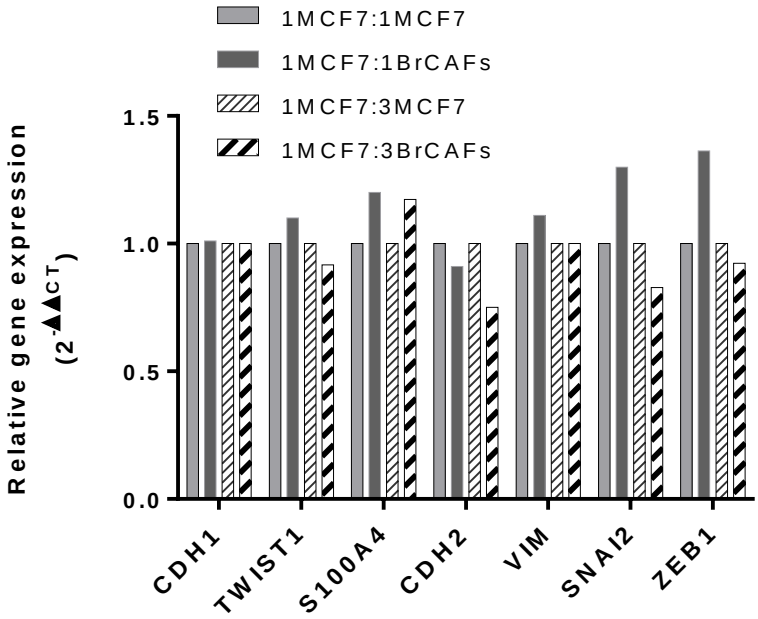


Figure 55. Relative gene expression of EMT markers in MCF-7 cells cultured in transwell plates with BrCAFs, or MCF-7 at different ratios. Each bar represents the $2^{-\Delta\Delta CT}$ value (1MCF7:1BrCAFs relative to 1MCF7:1MCF7 and 1MCF7:3BrCAFs relative to 1MCF7:3MCF7). N=1.

Summary and discussion

In this chapter, different methods were attempted to induce EMT-related traits, including change in the morphology of the cells and change in the expression of EMT markers. While in some of the methods attempted, some of the mesenchymal genes were upregulated, none of the techniques resulted in a complete EMT phenotype. Therefore, it was impossible to test the activity of the EMT reporters upon induction of EMT *in vitro*. Furthermore, no consistent changes in E-cadherin or N-cadherin were observed; hence, a different system to validate the activity of the E-cadherin and N-cadherin *in vitro* was not established.

The different methods used were an adaptation of the methods used in the literature, or were based on commercial products, suggesting that technical/methodological aspects could have affected the induction of

EMT. These include differences in the cell density, length of the assay, dimensionality, and the origin, batch and passage number of the cell lines used. Furthermore, some specific issues arose. For example, when attempting to overexpress Twist1 in MCF-7 cells, the gene was successfully upregulated but it was not shown whether the protein was upregulated and/or activated.

The process of EMT is controlled by different regulatory networks, which are interconnected (reviewed in [171]). In addition to different TFs activating different signalling pathways, non-coding RNAs, differential splicing, and translational and post-translational control which affect the proteins' stability and localisation, also regulate the EMT program. It is clearly demonstrated that EMT is a complicated process and hence not straightforward to achieve experimentally. Also, it can be referred that maybe only some of these regulatory mechanisms were activated by the methods attempted, but were not sufficient to induce a complete EMT alone.

Finally, human tumours and cell lines are heterogeneous, composed of cell with different plasticity and tumourigenic potential (reviewed in [172]). This means that not all cells will behave the same way; for example some cells are responsive to drugs and others are resistant. As a consequence, in breast cancer, and in most cancers, EMT rarely occurs homogeneously [173]. Only a portion of cells will undergo EMT, which implies that maybe EMT was induced in a few cells, but was not detectable.

EMT was not induced *in vitro*, due to all the aforementioned issues. This further stresses the need of improved models and detection of EMT, both *in vitro* and *in vivo*, which is the aim of this study.

CHAPTER 5: RESULTS 3 – *In vivo* model(s) development

This chapter will describe the studies performed, in order to establish *in vivo* model(s) of EMT and/or metastasis, for testing the activity of the EMT reporters in real-time. In the previous chapters, the activity of the S100A4 reporter was validated *in vitro* in MCF-7 and BT549 cells, therefore, these studies were performed using these two cell lines. The work aimed to test the tumourigenecity of these cell lines, induction of EMT, changes in the expression of the reporter markers and metastatic potential.

As earlier discussed, EMT is thought to be involved in the early stages of metastatic BC, where cells in the invasive border of the tumour undergo EMT which allows them to metastasise. It is not clear though when EMT is happening, which cells are involved and what happens to those cells after they metastasis, and to our knowledge this has not been shown in real-time. This project aimed to study this by the use of EMT reporter that will allow this real-time read out of EMT. MCF-7 cells are not thought to be metastatic, but they were chosen for this study in an attempt to experimentally induce EMT in an *in vivo* setting and capture the early stages of EMT. As previously discussed, the MSCs play an important role in tumour progression and metastasis and MSC-induced EMT-related traits have been reported in BC cell lines *in vitro*. So, MSCs were used to induce EMT in the low metastatic MCF-7 cells. Furthermore, the use in human-derived MSCs makes the xenograft model more clinically relevant.

Mouse models have been extensively used for studying human cancer, and they comprise one of the best models due to different features including small size, easy breeding, and most importantly, extensive physiological and molecular similarities to humans (reviewed in [174]). Different mouse models exist, including genetically engineered mice (transgenic or endogenous) and xenografts. In this study, xenograft models were established, in which the cell lines were propagated in immunodeficient/immunocompromised mice in ectopic or orthotopic sites. One main advantage of using xenograft models is the fact that they use cell lines derived from human cancers, and are hence a good model

for testing the efficacy of drugs against human cancers. In immunodeficient/immunocompromised mice, various breast cancer cells lines have been xenotransplanted and shown to be able to form tumours and metastasise into other organs, including lungs, liver, brain, bone and lymph nodes (reviewed in [175]). In this project two different mouse strains were used; the HsdOla:MF1-*Foxn1*^{nu} (aka MF1 nude mice) and the C;129S4-*Rag2*$^{tm1.1Flv}$ *Il2rg*$^{tm1.1Flv}$/J (aka Rag2$^{-/-}$ γc$^{-/-}$ mice). The MF1 nude mice carry the *nude* (*nu/nu*) mutation and are therefore hairless and completely lack a thymus [176]. Therefore, they are T-cell deficient and their lymphocyte population is composed mostly of B-cells. They have a very poor response to thymic-dependent antigens, including failure to reject xenografts. Rag-2 mice carry a germline mutation in which most of the *RAG-2* (recombinant-activating gene) coding region is depleted, whose protein function is to initiate V(D)J recombination, which is important for the maturation of pre-B and pre-T cells [177]. The particular strain also has mutations in the X-linked *IL2RG* gene, which encodes the common γ chain (γc) of the leukocyte receptors for interleukin-2 and multiple other cytokines that regulate the growth and maturation of lymphocytes, and its mutation results in severe combined immunodeficiency (SCID) [178]. Hence, these mice lack T, B and NK cells; their immunodeficiency makes them particularly useful as transplant hosts.

To support/maintain the growth of the cell lines, the cells were mixed with matrigel prior to implantation, and, in the case of the studies with the MCF-7 cells, the mice were receiving oestrogen (estradiol). Matrigel is a soluble form of basement membrane matrix extracted from Engelbreth-Holm-Swarm (EHS) mouse tumour, which is rich in proteins of the extracellular matrix (laminin, collagen IV and others) and growth factors (FGF, TGFβ and others) [179]. Tumour growth enhancement of a variety of BC cell lines, including MCF-7, with the addition of matrigel has been reported [180]. In the same study, it was shown that sustained growth of MCF-7 xenografts was maintained in the presence of oestrogen, confirming the well-documented hormone dependency of these cells.

While mouse models are the main tools for screening for drug efficacy, most drugs fail when tested in clinical trials in humans [181-183]. We anticipate that a major reason for this, is the fact that the mouse models lack the relevant tumour microenvironment, as the tumour microenvironment plays a critical role in tumour progression and metastasis (reviewed in [62]). To ensure a more relevant microenvironment, the human cell lines were implanted orthotopically in the MFP of the mice. Furthermore, we aimed to generate more humanised pre-clinical models, by co-implanting the human breast cancer cell lines with cells of the human stroma i.e. human mesenchymal stem cells (MSCs). MSCs are important for tumour progression, as they contribute to tumour vascularisation, EMT and metastasis (reviewed in [62]). Promotion of breast cancer formation and metastasis by human MSCs in mouse models, have been described [72, 184].

Tumour establishment was monitored manually by caliper measurements, a commonly used easy and non-invasive technique. Tumour growth was monitored by bioluminescent imaging; cell lines were transduced to express a plasmid construct that allowed constitutive expression of firefly luciferase, which was detected upon administration of its substrate, D-luciferin. As previously discussed, bioluminescent imaging is highly sensitive, and is suitable for detecting small numbers of cancer cells [116, 117]. Furthermore, the use of optical imaging has many advantages including non-invasive, quick and real-life monitoring of disease and drug response that minimizes the need of termination of the study at different time points for end-stage analysis, and hence reduces the number of animals being used (reviewed in [185]). This study not only contributed to the “reduction” of animals being used, but also to “refinement” (continual review of improvements in experimental design, techniques and husbandry to minimise adverse effects and improve welfare was undertaken), which comprise 2 of the 3Rs (https://www.nc3rs.org.uk/).

To test whether EMT-related traits were induced, the xenografts were fixed for immunohistochemical analysis of EMT markers. Bigger tumours

were half fixed and half used for gene expression analysis of the EMT markers.

5.1 Determining the effect of site of implantation on growth and phenotype of murine Xenograft models

5.1.1. Bioluminescent MCF-7 tumour EMT / metastasis pilot in an orthotopic mammary fatpad model

The study aimed to compare the ability to generate MCF-7 tumours at the implantation site, induction of EMT-related traits and metastatic growth via different initiation routes: MFP and sub-cutaneous (SC). Furthermore, this first study, aimed to assess the different methodologies used and to optimise the different analyses performed (e.g. correlation between bioluminescent signal and tumour weights), as a reference for subsequent studies.

5.1.1.1 Tumour growth

Bioluminescent MCF-7 cells were implanted in female MF1 nude mice in the MFP or SC and allowed to form primary xenografts. Xenograft formation was assessed by measuring the tumour volume, calculated using a calliper. Tumour growth was monitored by bioluminescent imaging every 1-2 weeks using the IVIS Spectrum. The MCF-7 cells were able to generate primary growth both at the MFP and SC sites after 1 week (Figure 55). At midpoint of the study, two mice per group were terminated; the mice with the bigger tumour volume in each group (mouse 4 from group 1 and mouse 6 from group 2) and the mice with the second smallest tumour (mouse 2 from group 1 and mouse 8 from group 2), in order to assess EMT at different time points i.e. midpoint of the study (day 67) and end of study (day 134).

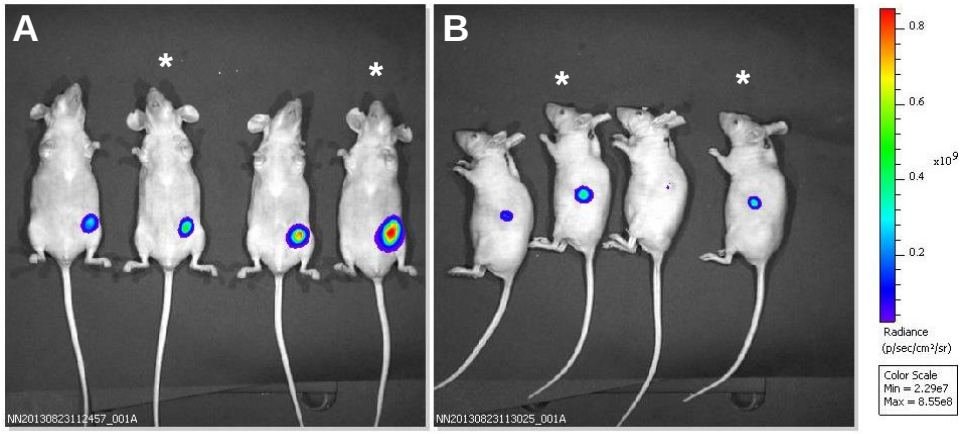


Figure 56. Representative images showing xenograft formation of bioluminescent MCF-7 cells. A. Group 1 = MFP; mice 1-4 (left to right) **B.** Group 2 = SC; mice 5-8 (left to right). Bioluminescent signal was measured as total flux (photons/second). The colour scale ranged from $2.297x10^7$ (blue) to $8.55x10^8$ (red). Pictures were taken at day 67 of the study (midpoint) using the IVIS Spectrum after administration of D-luciferin. The mice terminated at midpoint are demonstrated by a start.

The tumour volume in the two groups, as measured by calliper, followed the same pattern; primary xenografts were initially formed and then the tumour growth reached a plateau (Figure 56A). The difference in tumour volumes between the two groups was analysed using two-way ANOVA for each day. No significant differences in tumour growth were observed between the two groups. Looking at each mouse individually, the tumour volumes were similar in the two groups (ranged from 89.89 mm^3 to 182.95 mm^3 at termination point), with the exception of mouse 4 in which the tumour volume reached the size of 506.25 mm^3 (Figure 56B). Tumour growth was also monitored by bioluminescent readings, after administration of D-luciferin. The readings were misleading, due to variations of sensitivity of D-luciferin between the different time points. This was overcome in subsequent studies, by using the same batch of luciferin throughout the study.

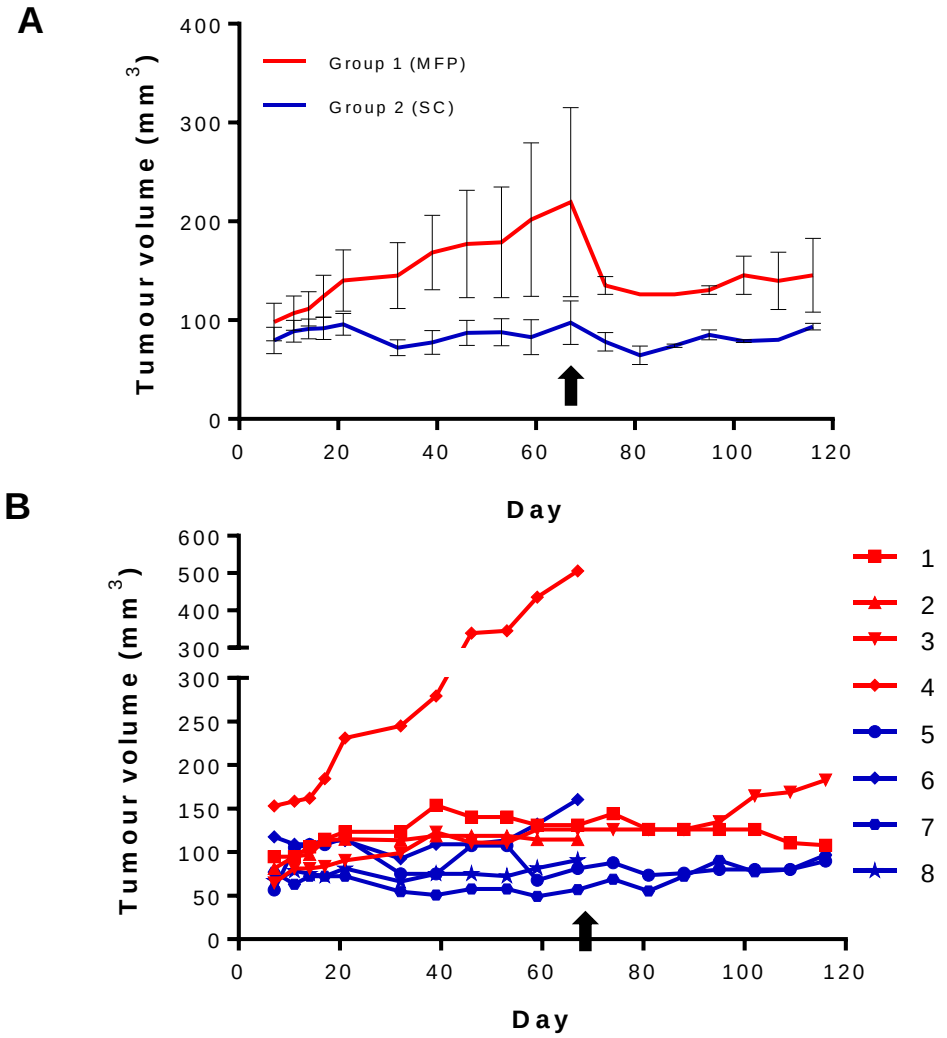


Figure 57. Tumour growth of bioluminescent MCF7-cells in MF1 nude mice injected in the MFP (group 1) or SC (group 2). A. Mean tumour volumes (mm3) ± SEM in group 1 (red line) and Group 2 (blue line) as measured by a calliper. B. Tumour volumes (mm3) of individual mice (1-8) from each group (red lines=group 1and blue lines=group 2). The arrow indicates midpoint of the study where mice 2, 4, 6 and 8 were terminated.

Prior to termination, D-luciferin was administered, and the mice were imaged as normal. Following imaging, the mice were sacrificed, the xenografts were removed, weighed, and bioluminescent signal was measured (Figure 57A&B). Significant correlation was observed between the bioluminescent signal and tumour weights (Figure 57C), as determined by Pearson correlation coefficient (r=0.9576, *p*=0.0002).

Significant correlation was also observed between the tumour volumes (measured prior to removal) and the bioluminescent signal of the removed tumours (r=0.9850, p<0.0001) (Figure 57B). The tumours were later fixed or cut in half (half was fixed for IHC staining and half was preserved in RNAlater for gene expression analyses).

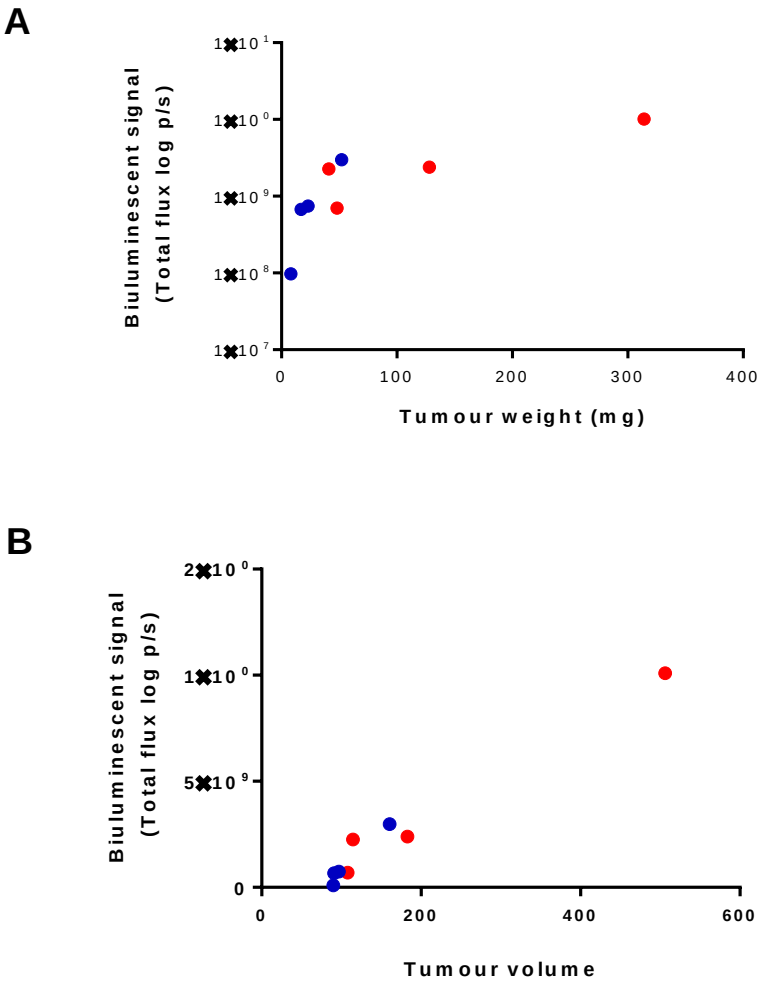


Figure 58. Correlation of bioluminescence signal from tumours removed at termination of the study with tumour weight and tumour volumes. A. Correlation of tumour weights (mg) and bioluminescent signal (total flux p/s). Each point represents the weight and signal of an individual xenograft. (r=0.9170, p=0.0002). **B.** Correlation of tumour volumes (mm^3) and bioluminescent signal. (r=0.9850, p<0.0001). N=8 (4 per group). Red=Group 1. Blue=Group 2.

5.1.1.2 Metastasis

The mice were imaged after intermediate dissection in order to detect any signal in metastatic sites. No signal was detected in distant organs. The lymph nodes and spleens were removed and imaged individually. No signal, indicating metastasis, was detected in the spleens (Figure 58).

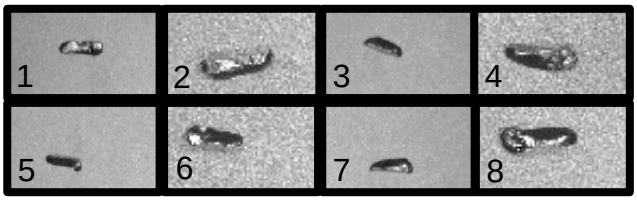


Figure 59. Bioluminescent images of spleens removed. Numbers indicate the mouse number. 1-4=MFP group and 5-6=SC group. Images were taken using the IVIS Spectrum, upon termination of the study.

Bioluminescent signal was only detected in the left LNs removed from the MFP group (Figure 59A, 59C, 59E & 59G). No signal was detected in the left LNs removed from the SC group or from the right LNs removed from either group (Figure 59). For those lymph nodes in which signal was detected, H&E staining was performed to reveal the structure of the lymph nodes and identify any tumour cells invading the lymph nodes. Normal structures were identified including capsule, adipose tissue and blood vessels (Figure 60), but no tumour cells were identified in either of the LNs. A representative example is shown in Figure 60; for all the other lymph nodes see appendix 10. The lymph nodes were also stained with anti-human cytokeratin, which only stains the tumour cells [186], in case the few tumour cells that had invaded the LNs were not detectable by H&E staining, but again no tumour cells were detected (data not shown). The left LNs are located very close to the MFPs (implantation site) and it is possible that the bioluminescent signal detected in the LNs was from MCF-7/fLuc cells that were not successfully removed, or perhaps tumour cells in transit state.

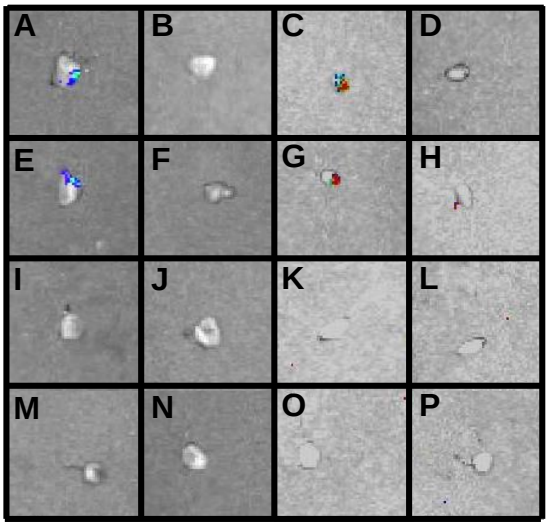


Figure 60. Bioluminescent images of LHS and RHS LNs. Mouse 1, left **(A)** and right **(B)**. Mouse 2, left **(C)** and right **(D)**. Mouse 3, left **(E)** and right **(F)** Mouse 4, left **(G)** and right **(H)**. Mouse 5, left **(I)** and right **(J)** Mouse 6, left **(K)** and right **(L)**. Mouse 7, left **(M)** and right **(N)**. Mouse 8, left **(O)** and right **(P)**.

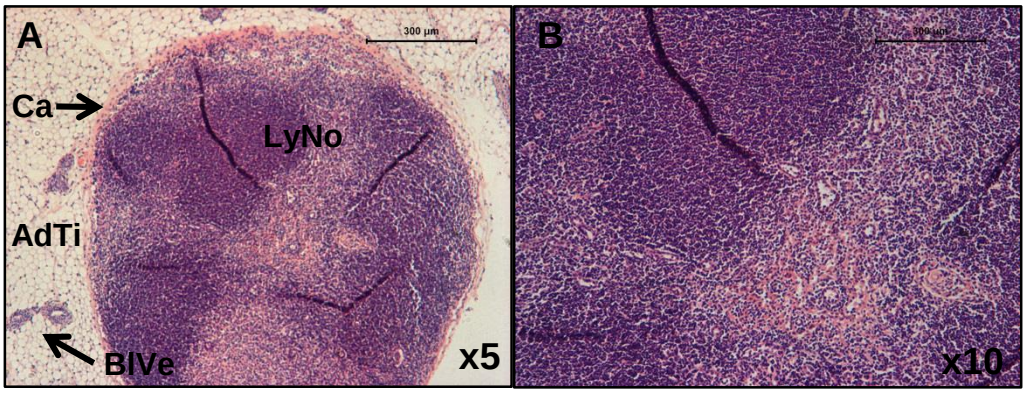


Figure 61. H&E staining of lymph node 2 (LHS). A. x5 magnification. **B.** x10 magnification. Ca = capsule, LyNo = Lymph node, AdTi = Adipose tissue, BlVe = Blood vessel.

5.1.1.2 *Ex vivo* analysis of EMT traits

In order to check whether EMT was induced, the expression of the EMT markers E-cadherin and Vimentin was analysed in primary xenografts, both at the protein level by IHC and at the gene level by qPCR (in the available samples).

5.1.1.2.1 E-cadherin protein and gene expression

The xenografts were stained with anti-human E-cadherin, which is expected to be reduced or lost upon induction of EMT, or an isotype control. As a reference for the staining (intensity and pattern), bioluminescent MCF-7 cells, cultured in normal cell culture conditions, were fixed and stained with the same antibody. The staining of the cells showed that most of the cells were positive for E-cadherin expression, and that the staining was located in the cell junctions and in the cytoplasm (Figure 61A). The same pattern was observed in *in vitro* cultured cells, when stained with the same antibody and imaged by IHC or IF. The staining confirmed that all the cells that were in contact expressed E-cadherin, and that the staining was predominately membranous (Figure 62B). All xenografts, with the exception of xenograft 8 which was mostly negative, stained positive for E-cadherin (Figure 62; for the isotype controls see appendix 11). The intensity and location of the staining was variable across the different xenografts (e.g. staining in xenograft 4 was located near the edges and was strong, while the staining in xenograft 1 was located mostly in the core of the tumour and the intensity was weak). Unlike the MCF-7 cells cultured *in vitro*, not all the cells in the xenografts were positive. To assess the difference in staining between the two groups, the amount of E-cadherin staining both the percentage area stained and the H-Score. Pictures were taken (at x20 magnification), 6 for each slide, using a Leica colour microscope. Images of low, medium and high levels of E-cadherin staining were selected in order to use these as a baseline for setting the ImageScope software (algorithm “Positive Pixel Count v9”). The boundaries had been set (high positive=190, positive=130 and low positive=100) and the program was used to analyse all the pictures. Figure 63 shows the different boundaries, including negative tissues and necrotic areas i.e. red=high, orange=medium, yellow=low, blue=negative tissue and white=no tissue. Areas of necrosis were manually excluded from the analysis. The percentage of positive areas varied from 11% to 45%. No significant difference was observed between the two groups or the different time points of the study, as

determined by an unpaired t-test (Figure 64A-C). The H-scores varied from 14 to 69. No significant difference was observed in the H-scores between the two groups or the different time points of the study (Figure 64D-F).

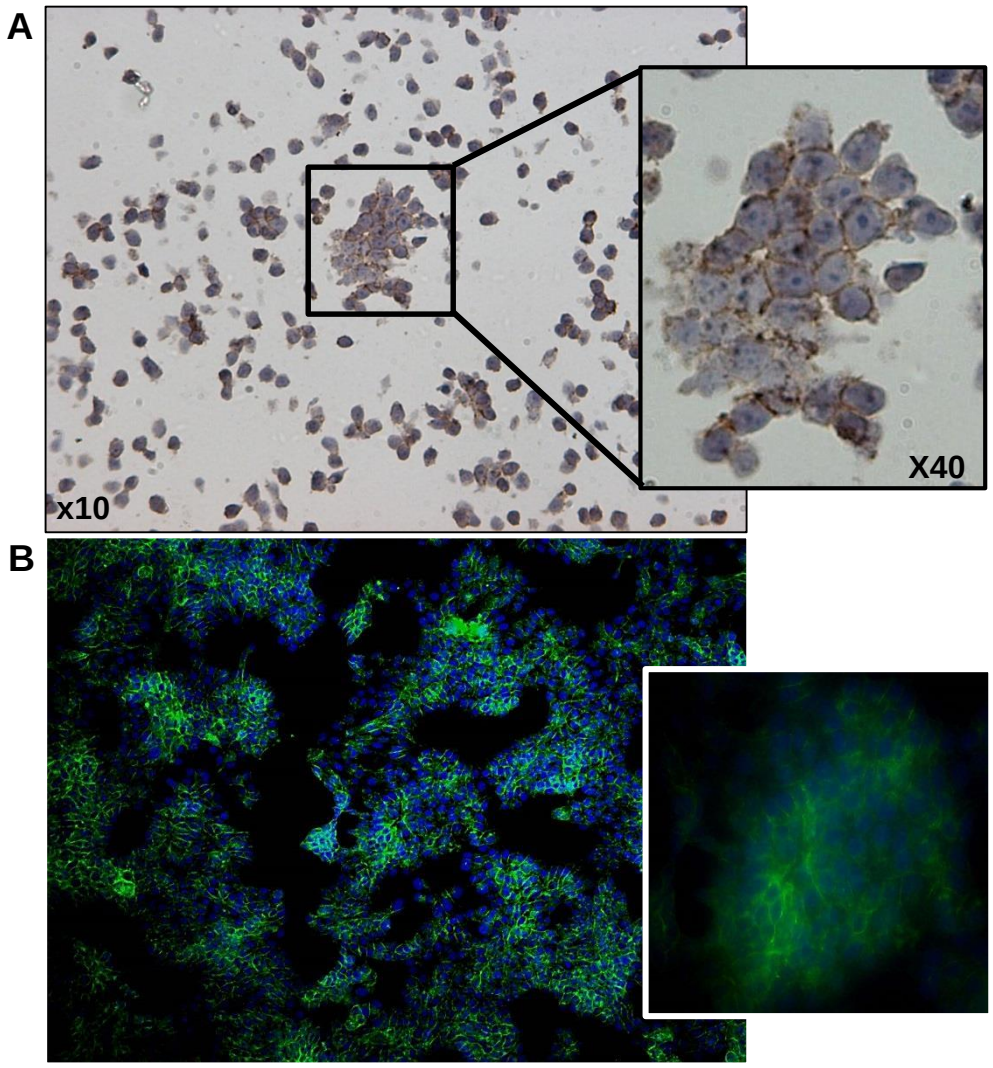


Figure 62. E-cadherin staining of bioluminescent MCF-7 cells. A. IHC staining of MCF-7 cells fixed, suspended in agarose and embedded in paraffin with anti-human E-cadherin. **B.** Immunofluorescence staining of MCF-7 cells cultured in cover slip, stained with anti-human E-cadherin (green) and Dapi (blue). Pictures were taken using the Nikon Eclipse Ti microscope. Pictures were taken at 10x (left) and 40x (right) magnifications.

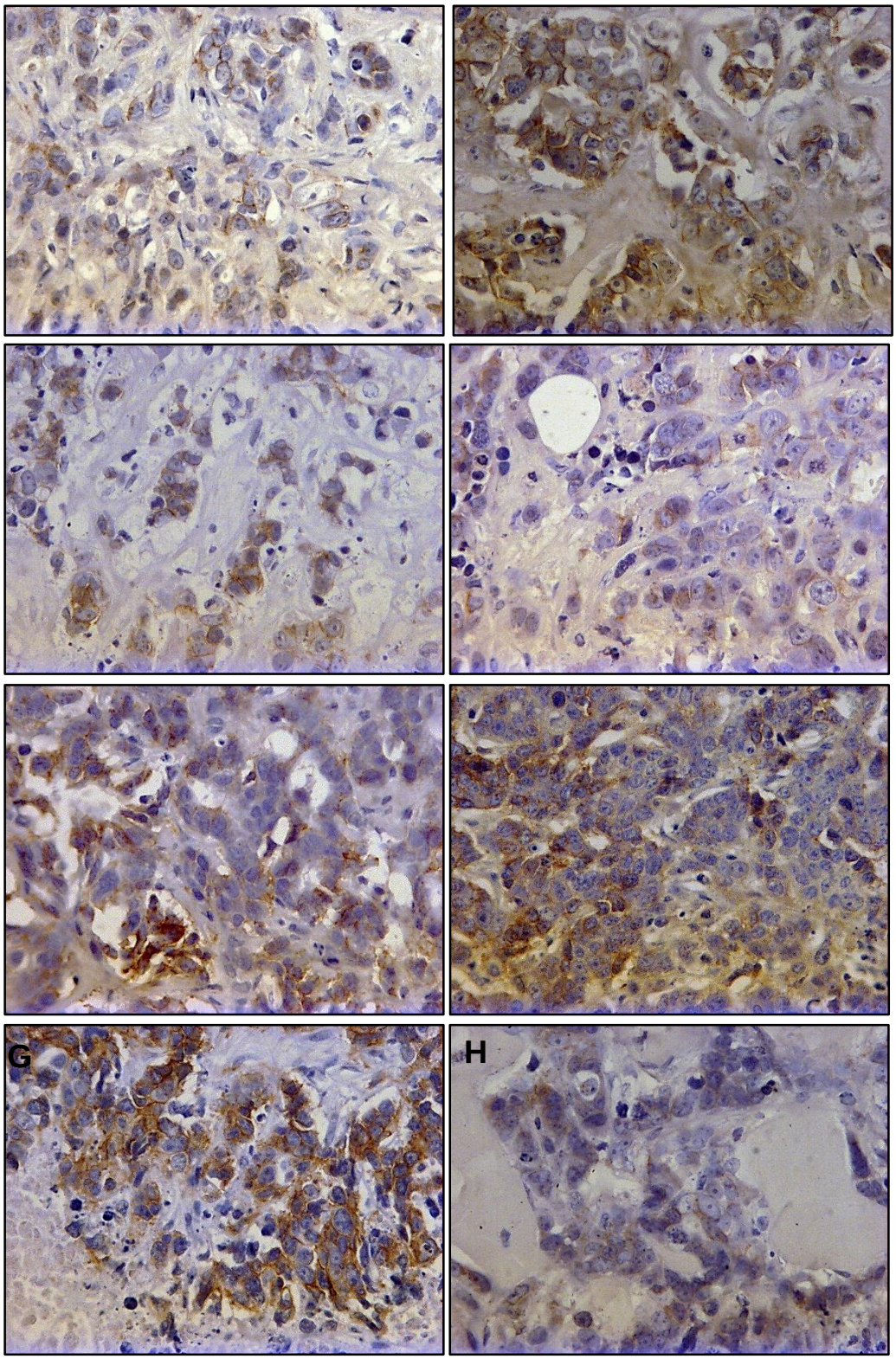


Figure 63. IHC staining of E-cadherin in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/fLuc cells in the MFP). **E-H** Mice 5-8 from group 2 (MCF-7/fLuc cells SC). Pictures taken at x40 magnification.

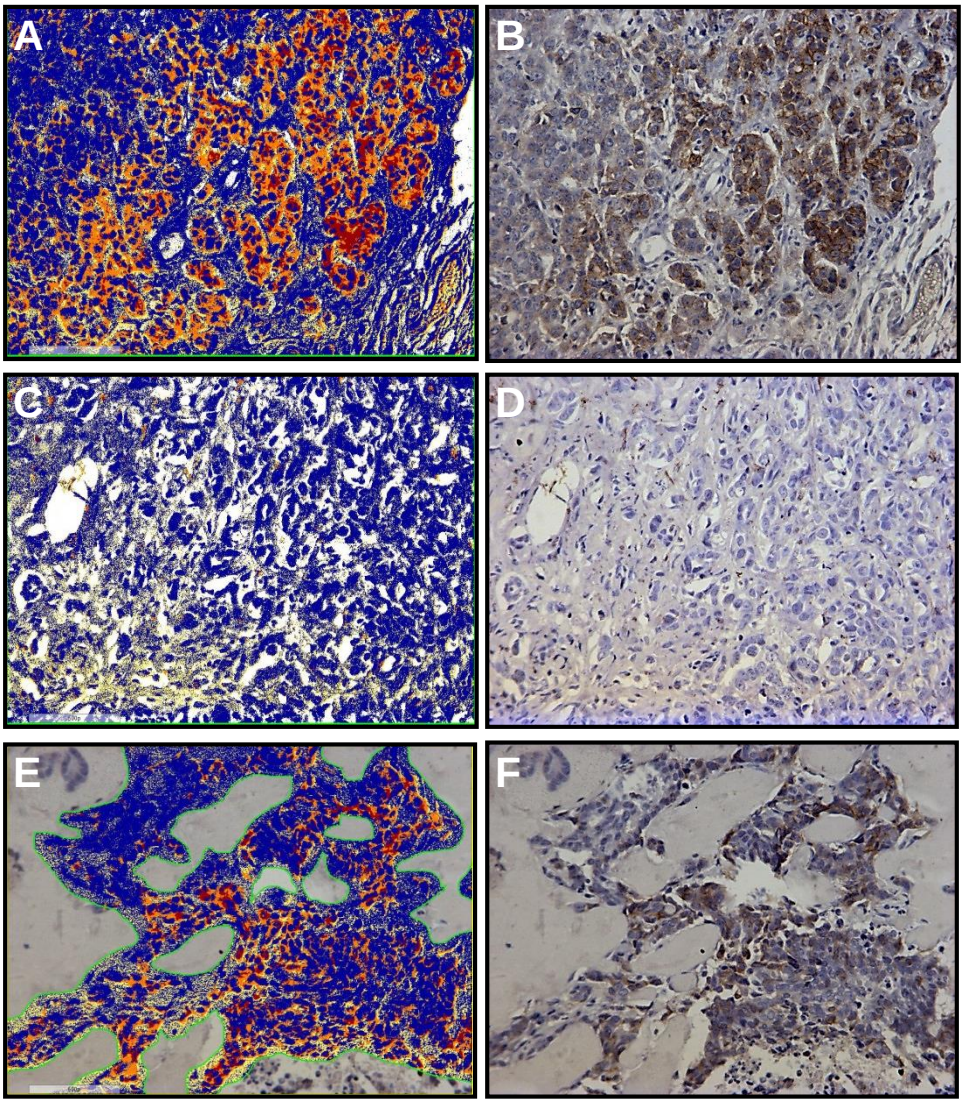


Figure 64. Images showing analysis of E-cadherin staining using the ImageScope Software. Example of staining used to determine the boundaries between positive and negative (A), where red cells represent high staining intensity, orange cells represent medium staining intensity, blue cells represent negative density and yellow cells represent low intensity staining, matched to the original image (B). Boundaries were set as follows: high positive=190, positive=130 and low positive=100. Example of staining used to determine the boundaries between negative cells, shown in blue, and no tissue, shown in white (C), matched to the original image (D). Example showing manual exclusion of areas of necrosis (marked with green line) matched to the original image (F).

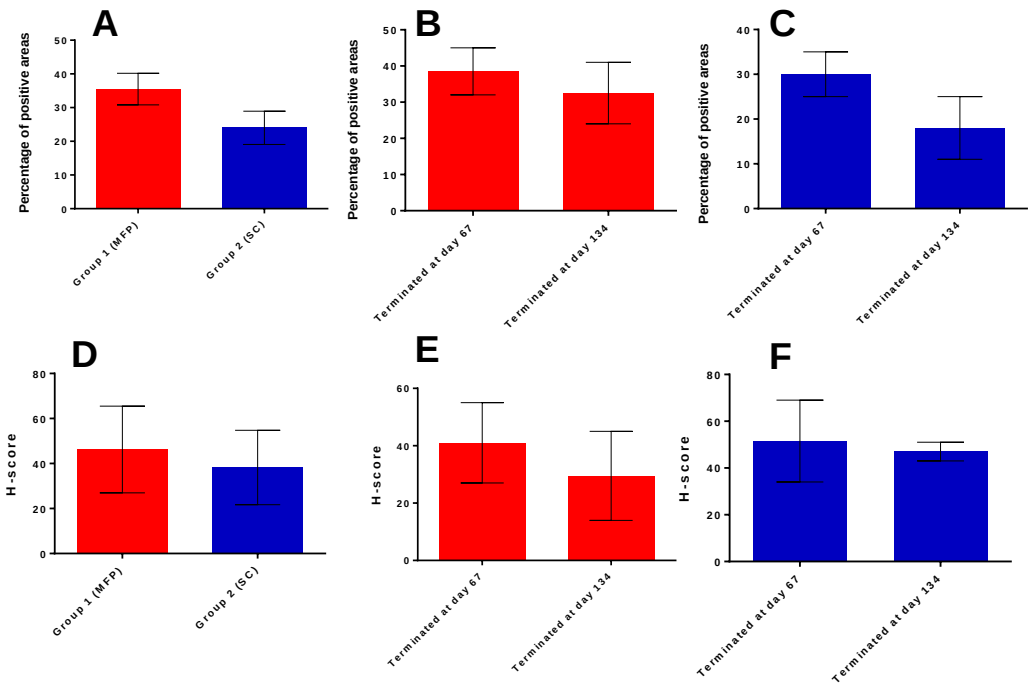


Figure 65. Percentage of E-cadherin positive areas and H-scores in xenografts. A. Average percentage in all mice in group 1 and group 2 (N=4 per group). **B.** Average percentage in mice terminated at day 67 (N=2) and at day 134 (N=2) from group 1. **C.** Average percentage in mice terminated at day 67 (N=2) and at day 134 (N=2) from group 2. **D.** Average H-scores in all mice (N=4 per group). **E.** Average H-scores in mice terminated at day 67 (N=2) and at day 134 (N=2) from group 1. **F.** Average H-scores in mice terminated at day 67 (N=2) and at day 134 (N=2) from group 1. Bars represent mean values. Error bars show ±SEM. Red bars = group 1 and blue bars = group 2.

The xenografts stored in RNAlater (available xenografts only xenografts 2, 3 and 4 from group 1) were mechanically homogenised using a mortar and pestle and collected with TRI-Reagent, the RNA was extracted, cDNA was synthesised and the gene expression levels of E-cadherin were analysed by qRT-PCR. Gene expression was compared with the baseline E-cadherin expression of the bioluminescent MCF-7 cells cultured *in vitro*. The gene expression was similar in the xenografts 2, 3 and 4 ($2^{-\Delta CT}$ values = 3.3, 3.0 and 3.19 respectively) and between 2.84 and 3.12 folds lower than the control MCF7-fLuc cells (Figure 65). This

showed that E-cadherin was downregulated when the cells were grown *in vivo*.

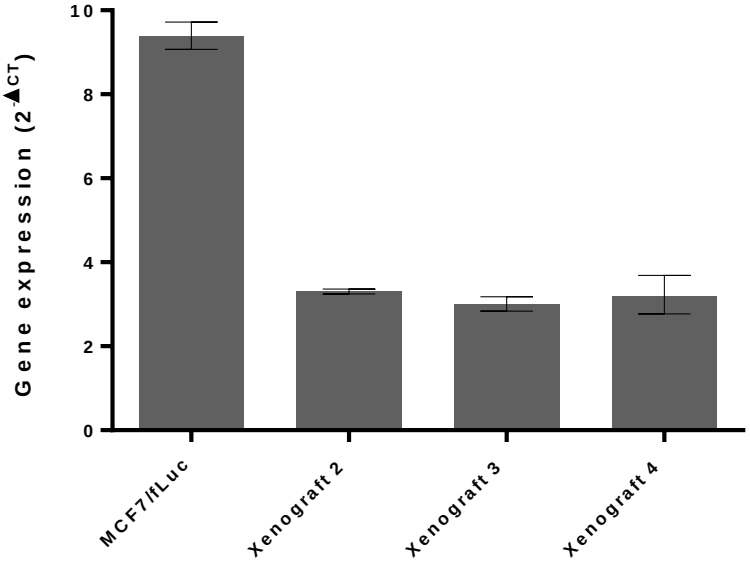


Figure 66. E-cadherin gene expression in mouse xenografts and control cells (bioluminescent MCF-7 cells). Bars represent the average $2^{-\Delta CT}$ values and lower and upper limits as follows: MCF-7/fLuc = 9.40, Xenograft 2 = 3.34, Xenograft 3 = 3.00 and Xenograft 4 = 3.19. N=1.

5.1.1.2.2. Vimentin protein and gene expression

The xenografts were stained with anti-human Vimentin which is upregulated upon induction of EMT. At this point, there was no more tissue left from xenograft 8, therefore it was not included in the analysis. Tissue from human colorectal cancer was used as a positive control for the staining (Figure 66A). All xenografts were negative for Vimentin. For the isotype controls see appendix 12. The gene expression of Vimentin was analysed and was compared with the baseline Vimentin expression of the bioluminescent MCF-7 cells. The $2^{-\Delta CT}$ value in the MCF-7 cells was 0.008 and the values ranged from 0.01 to 0.08 in the xenografts (Figure 67). The extremely low $2^{-\Delta CT}$ values indicate that the expression of Vimentin remained very low when the cells were implanted *in vivo*.

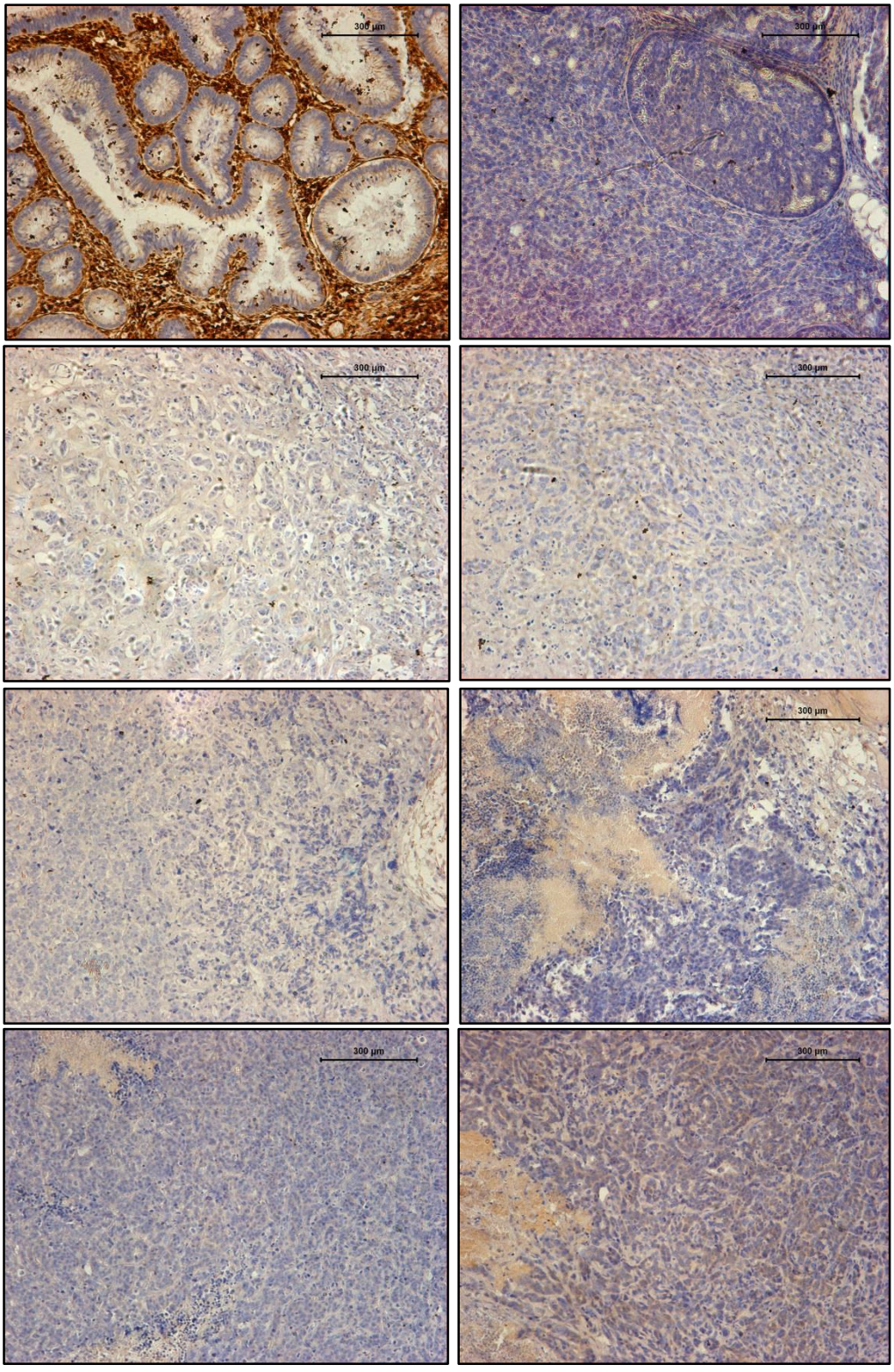


Figure 67. IHC analysis of human anti-Vimentin in mouse xenografts. A. Positive control (human colorectal tumour). **B-E** xenografts 1-4 from group 1. **F-H.** Xenografts 5-7 from group 2.

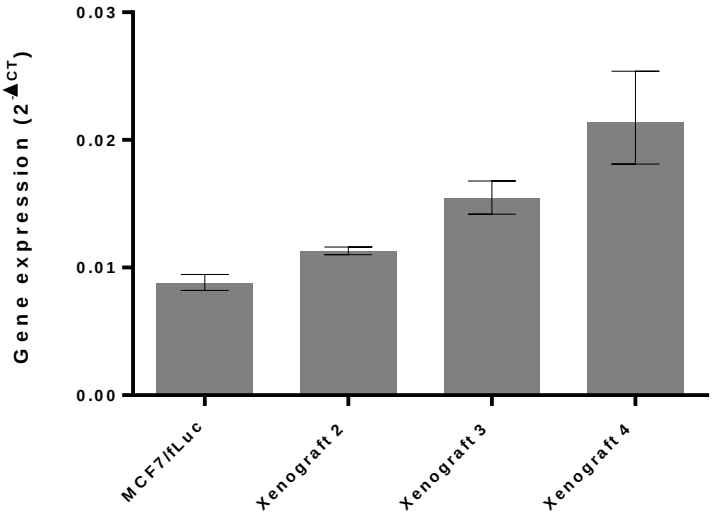


Figure 68. Vimentin gene expression in Xenografts and control (bioluminescent MCF-7 cells). Bars represent the average $2^{-\Delta CT}$ values and lower and upper limits as follows: MCF-7/fLuc = 0.008, Xenograft 2 = 0.011, Xenograft 3 = 0.015 and Xenograft 4 = 0.021. N=1.

5.1.2. Bioluminescent BT549 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model

The study aimed to compare the ability to generate BT549 xenograft at implantation site and metastatic growth via different initiation routes; MFP and sub-cutaneous (SC), as well as to test the effect of matrigel on tumour growth. As explained previously, the tumour microenvironment plays an important role in tumour growth; therefore, the study aimed to determine whether there will be a difference in growth between orthotopic and SC sites. Furthermore, the use of matrigel has been shown to support the growth of BC cell lines, so the study aimed to explore whether matrigel would have an effect on tumour growth.

Bioluminescent BT549 cells were implanted in female MF1 nude mice in the MFP (in culture medium) or SC (in culture medium or matrigel) and allowed to form primary xenografts. Xenograft formation was monitored by calliper measurements, and tumour growth was visualised by bioluminescent imaging every 1-2 weeks using the IVIS Spectrum. As shown in Figure 68, signal was detected at the beginning of the study, which then decreased and reached a plateau, in all the groups. As shown in the image of the bioluminescent reading at the end of the study, little signal was detected in the mice (Figure 69). In summary, the BT549 cells did not form tumours in either of the implantation sites tested, and the matrigel had no effect on growth. The tumour volumes were mainly unmeasurable throughout the study, showing that tumours were not formed. Therefore, no tissue was available at the end of the study for further examination.

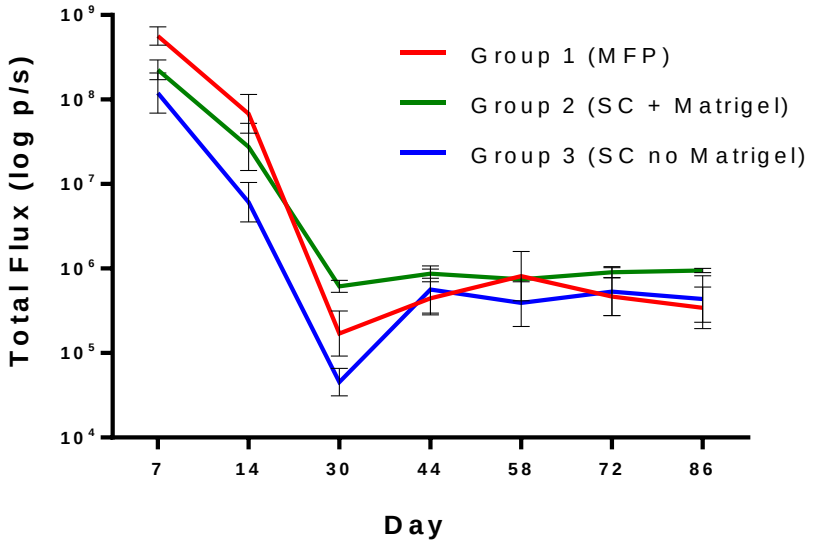


Figure 69. Bioluminescent readings of BT549 cells implanted in MF1 nude mice. Each line represents a different group. Each point represents the average bioluminescent signal (total flux) on a logarithmic scale ± SEM. Readings were taken after administration of D-luciferin throughout the stuy. N=3 per group.

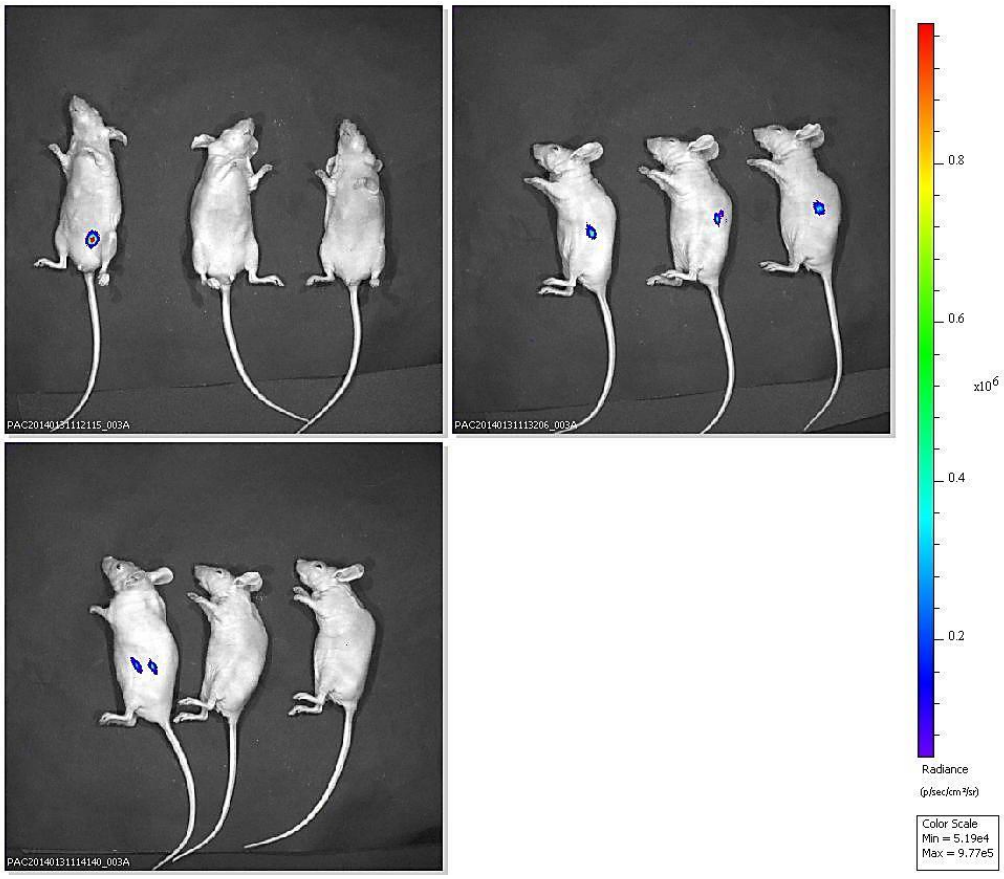


Figure 70. Representative images showing bioluminescent readings of BT549 cells in nude mice. A. Group 1 = MFP; mice 1-3 (left to right) **B.** Group 2 = SC with matrigel; mice 4-6 (left to right). C. Group 3 = SC no matrigel; mice 7-9 (left to right). Bioluminescent signal was measured as total flux (photons/second). The colour scale ranged from $5.19x10^4$ (blue) to $9.77x10^5$ (red). Pictures were taken at the end of the study after administration of D-luciferin.

5.2 Assessing the effect of co-implanting epithelial with mesenchymal cells on growth, metastasis and EMT

Tumour formation of bioluminescent MCF-7 was previously observed, but no substantial tumour growth was observed. Furthermore, no EMT or metastasis was observed, and no tumour formation at all was observed with the BT549 cells. As MSCs have been shown in the literature to promote tumour growth, BT549 were co-implanted with hMSCs and the effect of the MSCs on tumour growth was assessed. Although no significant difference was previously observed between the MFP and SC initiation, the MFP site was chosen for the subsequent studies, as it should provide a more relevant TME.

5.2.1 BT549 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model – Rag2-/- mice via co-implantation with MSCs

The study aimed to test the ability to generate BT549 tumour at implantation site and metastatic growth via co-implantation with MSCs, as well as to test the effect of MSCs on the expression of S100A4. As the nude mice were no longer available at this point, the Rag2 mice were used.

5.2.1.1 Tumour growth and metastasis

BT549 cells were implanted in the MFP site of Rag2-/- mice either alone or with MSCs. The tumour growth was monitored as before. As shown in Figure 70, bigger tumours were formed when co-implanted with MSCs. In mouse 4 (“BT549 + MSCs” group), tumours were formed early in the study and continue to grow until termination of the study. In mouse 3 (“BT549 + MSCs” group), the tumours were formed and started growing later in the study. On the other hand, small-sized tumours were formed in the “BT549 only” group (mice 1 and 2), and the tumour volume decreased until the end of the study; in mouse 2 the tumour volume was undetectable by the end of the study. The study was terminated as the

mice did not recover from anaesthesia. This study showed that BT549s require the support of MSCs to form tumours.

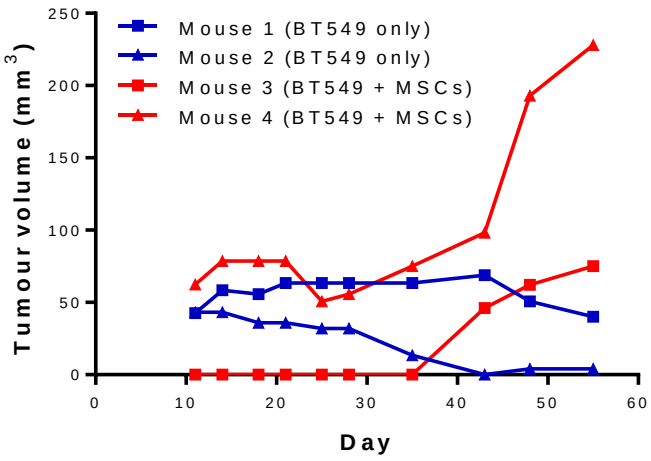


Figure 71. Tumour volume formation of bioluminescent BT549 in Rag2-/- nice. Each line represents a different mouse. Blue lines represent mice from group 1 (BT549 only). Red lines represent mice from group 2 (BT549 + MSCs). Each point represents a calliper tumour volume (mm^3) measurement.

5.2.1.2 S100A4 expression

The activity of the S100A4 reporter had been validated in the BT549 cells (chapter 3). Briefly, under hypoxic conditions the S100A4 gene expression was upregulated, and this was associated with upregulation of the bioluminescent signal. Therefore, the expression of S100A4 was analysed in the BT549 mouse xenografts at the end of the study, to assess whether the BT549 cells expressed S100A4 at the protein levels. It was shown that most of the cells (in the xenografts from the “BT549 + MSCs” group) expressed S100A4 (Figure 71). For the isotype controls see appendix 13. As no tumour formation was observed in “BT549 only” group (or the xenografts were very small), it was not possible to stain for S100A4.

In order to assess whether the S100A4 expression was due to the effect of MSCs, and to reproduce the effect of MSCs on growth, the study was repeated. In this study the imaging was less frequent in order to minimise the stress and avoid earlier termination of the study. The study also aimed to analyse the expression of S100A4 at different stages of tumour growth (early and late stages of tumour growth), in order to assess whether a change on the S100A4 expression was due to tumour progression and/or the effect of MSCs and determine whether the S100A4 reporter would be tested in BT549 mouse xenografts. The tumours did not grow in this study, even when the cells were co-implanted with MSCs; therefore it was not possible to stain for S100A4.

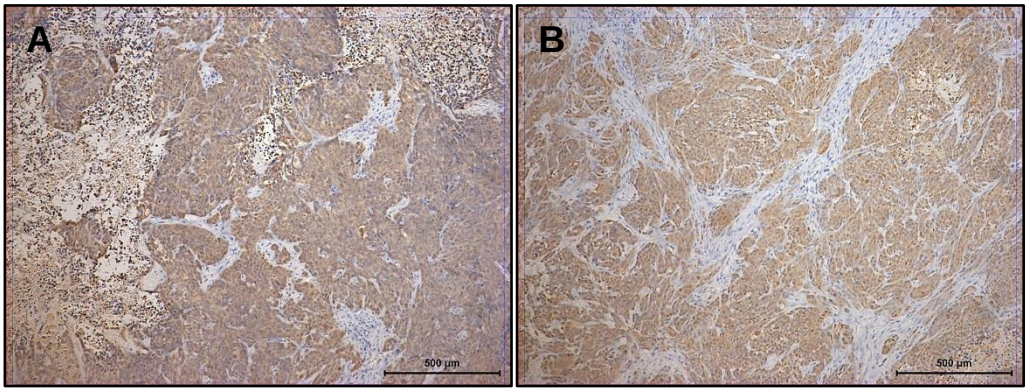


Figure 72. IHC analysis of S100A4 in the BT549 mouse xenografts co-implanted with MSCs. A. Xenograft from mouse 1. **B.** Xenograft from mouse 2.

5.2.2 Bioluminescent MCF-7 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model – nude mice

This study aimed to assess the effect of human MSCs in MCF-7 tumour formation, EMT and metastasis. Bioluminescent MCF-7 cells were implanted in the MFP of female MF1 nude mice with or without MSCs in matrigel, at a ratio of 2MCF-7:1MSCs. This ratio has been previously used in the lab in 3D assays, where it was shown that MSCs promoted the growth of MCF-7 cells.

5.2.2.1 Tumour growth and metastasis

The MCF-7 cells, implanted alone in the MFP, established tumours at the beginning of the study and then the tumour growth reached a plateau (Figure 72A-B), but as before, the tumours did not grow well. Bigger tumours were formed and the growth was significantly promoted when the MSCs were co-implanted with the MSCs, as assessed by a two-way ANOVA test (Figure 72). At the end of the study the mice were imaged after intermediate dissection in order to detect any signal in metastatic sites. The lymph nodes and spleens were removed and imaged individually. No signal, indicating metastasis, was detected in the spleens, lymph nodes or other sites (data not shown).

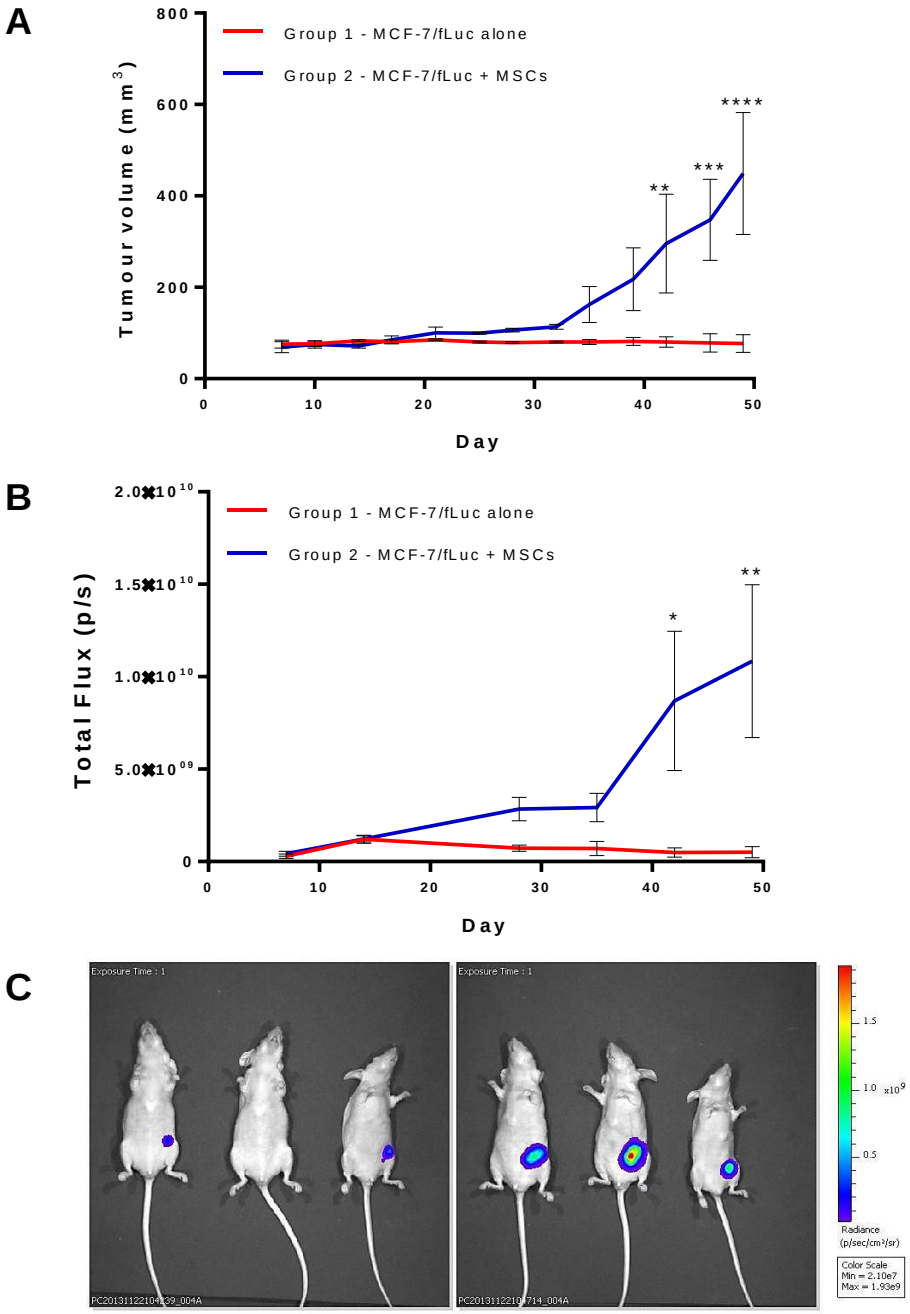


Figure 73. Tumour formation and tumour growth of bioluminescent MCF-7 cells in MFP mouse models. A. Tumour volumes (mm^3) in group 1 (MCF-7 alone) and group 2 (MCF-7 + MSCs). **B.** Bioluminescent imaging (total flux; photons/second) in group 1 and group 2 as measured after administration of D-luciferin. Each bar represents the average volume (A) or average total flux (B) ± SEM. N=3 per group. Statistical analysis was performed with two-way Anova. *P≤0.05, **P≤0.01 and ***P≤0.001. **C.** Representative image of bioluminescent imaging showing mice 1-6 (from left to right); mice 1-3 = group 1 and mice 4-6= group 2**.** The colour scale ranged from $2.10x10^7$ (blue) to $1.93x10^9$ (red).

5.2.2.2 *Ex vivo* analysis of EMT traits

In order to check whether EMT was induced, or changes in the expression of the reporter markers were detected, the xenografts were stained with E-cadherin, Vimentin and S100A4. At the same time, the tumour histology was examined. The gene expression of E-cadherin and Vimentin was also analysed in the available xenografts. The gene expression of S100A4 was not analysed, because of the lack of primer sequences that allow species-specific detection of S100A4, as there is a high homology between the mouse and the human gene (93% homology according to http://atlasgeneticsoncology.org/).

5.2.2.2.1 E-cadherin protein and gene expression

As shown in Figure 73, the xenografts from group 1 (MCF-7 cells only, Figure 73A-C) were composed of few cancer cells and contained large areas of necrosis, while the xenografts from group 2 (MCF-7 with MSCs, Figure 73D-F) were composed mainly of cancer cells (Figures 73, 75, 76 and 78).

The xenografts were stained with anti-human E-cadherin or an isotype control to detect any unspecific/background staining. No background staining was observed (see appendix 14 for pictures of the isotype controls). The cancer cells in xenografts from group 2 had lower expression of E-cadherin compared with the cells in xenografts from group 1 (Figure 73D-F). The gene expression of the available xenografts from group 2 was analysed and compared with the gene expression of E-cadherin of the bioluminescent MCF-7 cells. The gene expression in the xenografts 4 and 5 was 6 and 5 times lower than the expression of the MCF-7 cells accordingly (Figure 74); an effect that was greater than when the MCF-7 cells were implanted in the MFP alone (Figure 65). In summary, co-implanting MSCs with MCF-7 resulted in lower E-cadherin expression, both at the gene and protein levels.

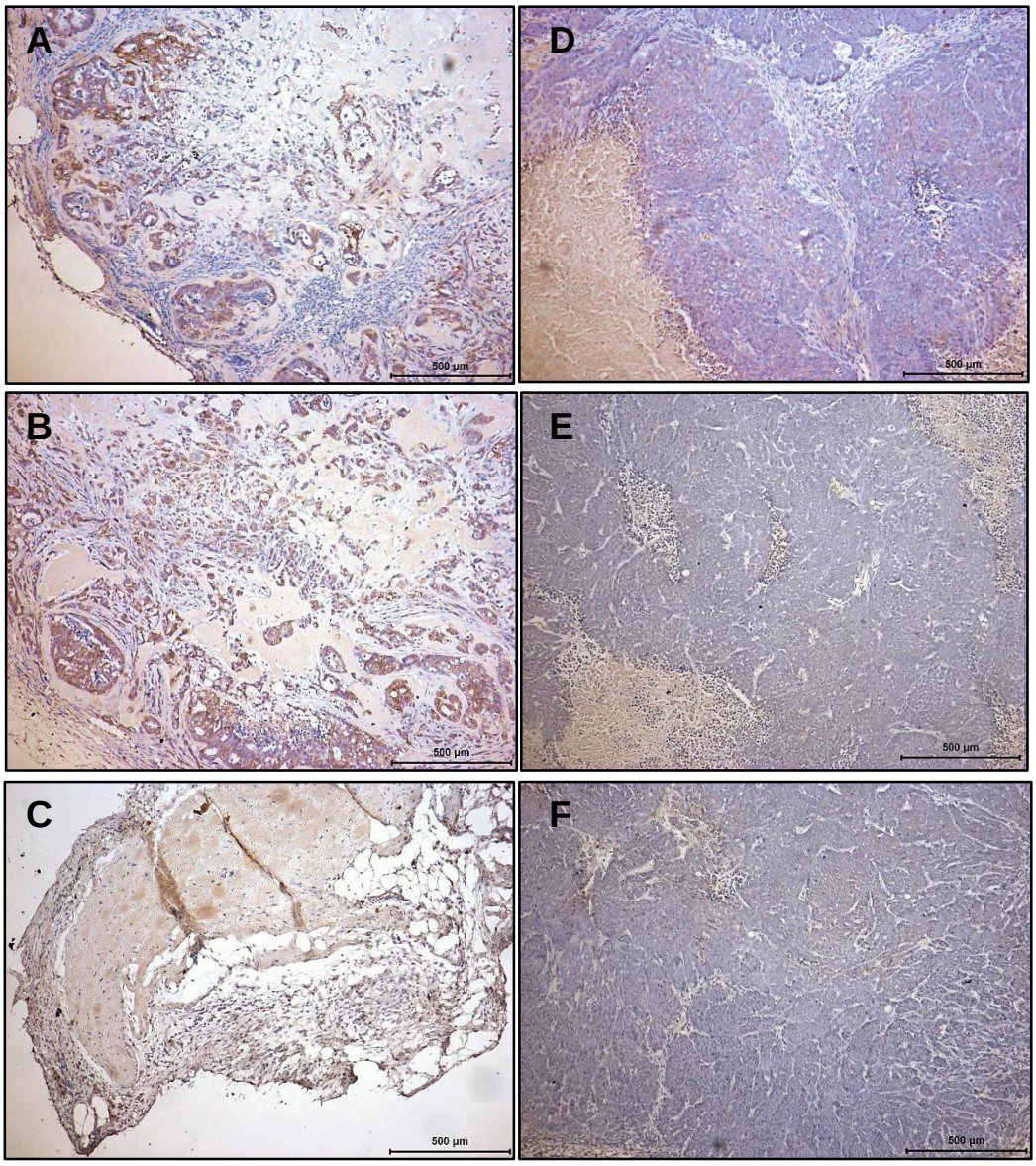


Figure 74. IHC analysis of E-cadherin in MCF-7 mouse xenografts. A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-6 (Group 2; MCF-7 + MSCs).

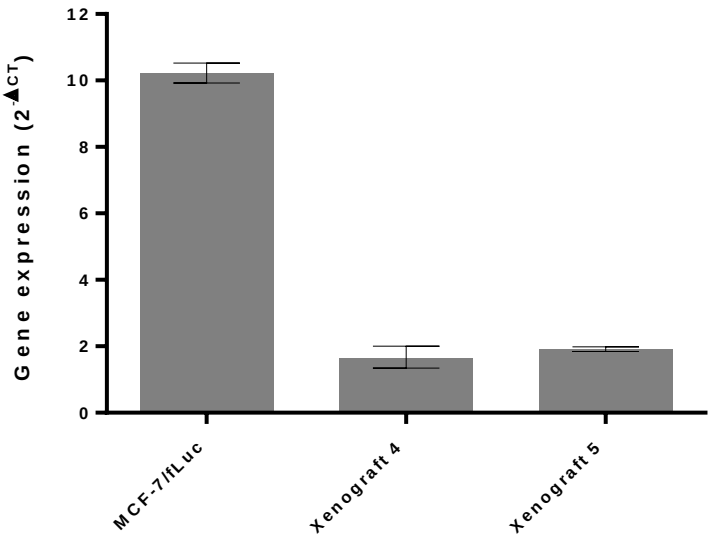


Figure 75. E-cadherin gene expression in mouse xenografts and control cells (bioluminescent MCF-7 cells). Bars represent the average $2^{-\Delta CT}$ values and lower and upper limits as follows: MCF-7/fLuc = 10.22 xenograft 4 = 1.64 and xenograft 5 = 1.91. N=1.

5.2.2.2.2 Vimentin protein and gene expression

The xenografts were stained with anti-human Vimentin, or an isotype control. No background staining was observed (see appendix 15 for pictures of the isotype controls). While none of the xenografts from group 1 expressed Vimentin (Figure 75A-C), patches of strong Vimentin expression was observed in all the xenografts from group 2 (Figure 75D-E). To ensure that Vimentin was expressed by the tumour cells (MCF-7) and not by any residual MSCs, staining on sequential sections of Vimentin and Cytokeratin, which reacts with human breast epithelial tissue [186], was performed. The antibody used reacts with both luminal and basal cytokeratins including CK5, CK6, CK8, CK17 and possibly CK19. The matched pictures are shown in Figure 76. It is shown that Vimentin staining is localised in areas that were also stained with cytokeratin, confirming that Vimentin was expressed in the tumour cells.

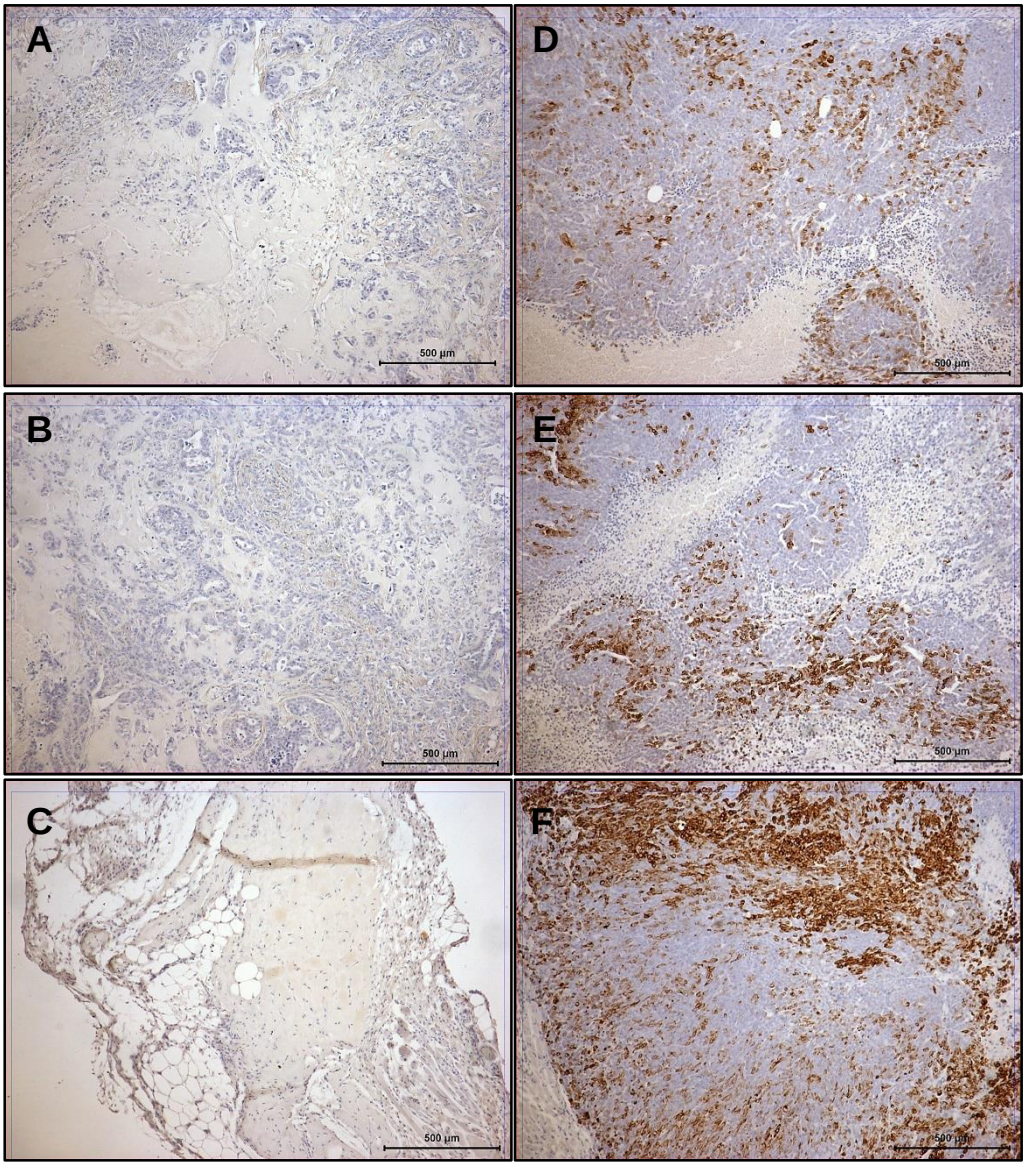


Figure 76. IHC analysis of Vimentin in mouse xenografts. A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-5 (Group 2;MCF-7 + MSCs).

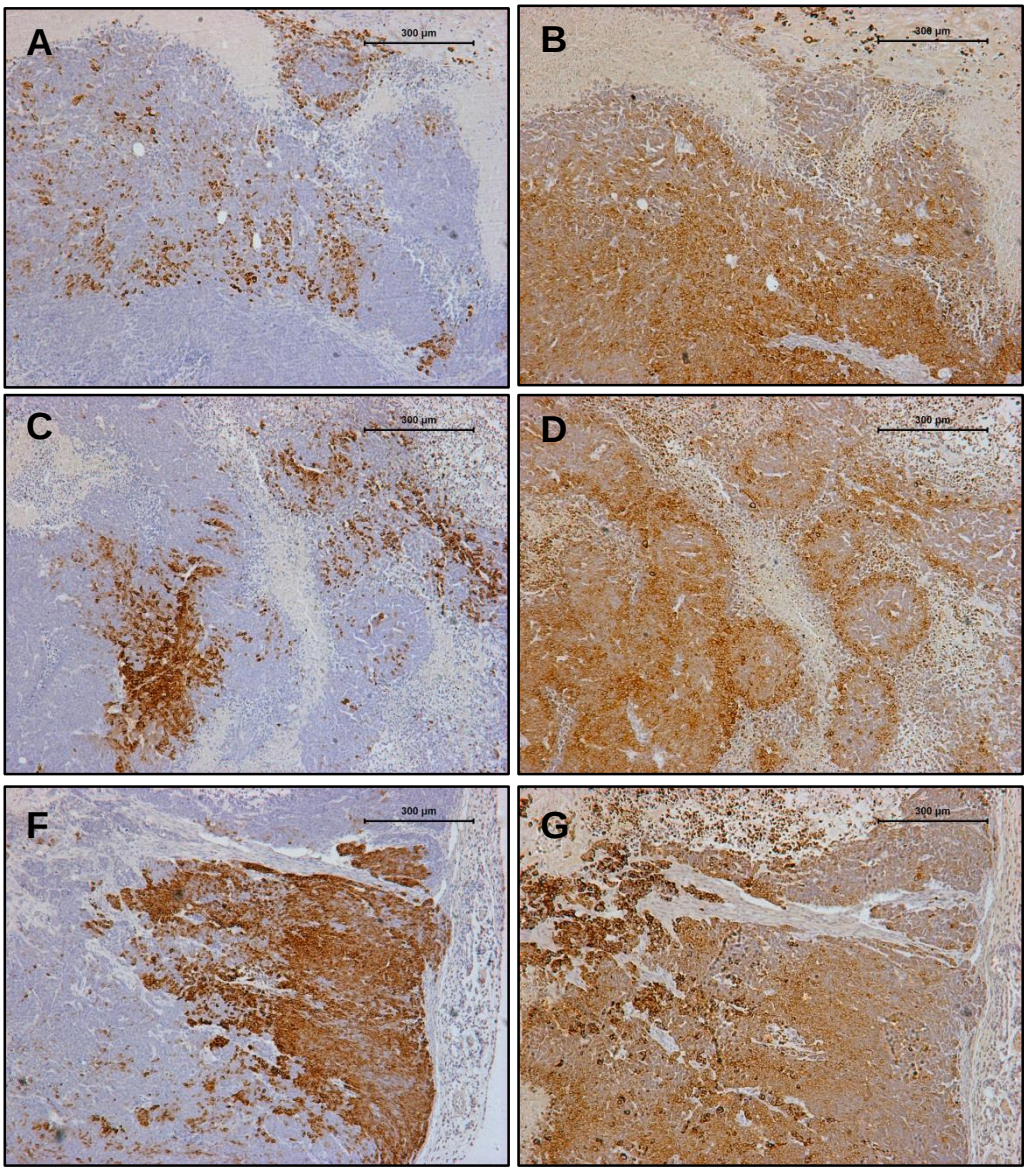


Figure 77. Matched pictures of IHC staining of Vimentin and Cytokeratin in mouse xenografts. A. Vimentin staining in xenograft **4. B.** Cytokeratin staining in xenograft 4. **C.** Vimentin staining in xenograft 5. **D.** Cytokeratin staining in xenograft 5. **F.** Vimentin staining in xenograft 6. **G.** Cytokeratin staining in xenograft 6.

The gene expression of Vimentin was higher in xenografts 4 and 5 (group 2) compared to the control cells (*in vitro*-cultured MCF-7/fLuc cells) (Figure 77). The expression was 76 times higher in xenograft 4, and 165 times fold higher in xenograft.

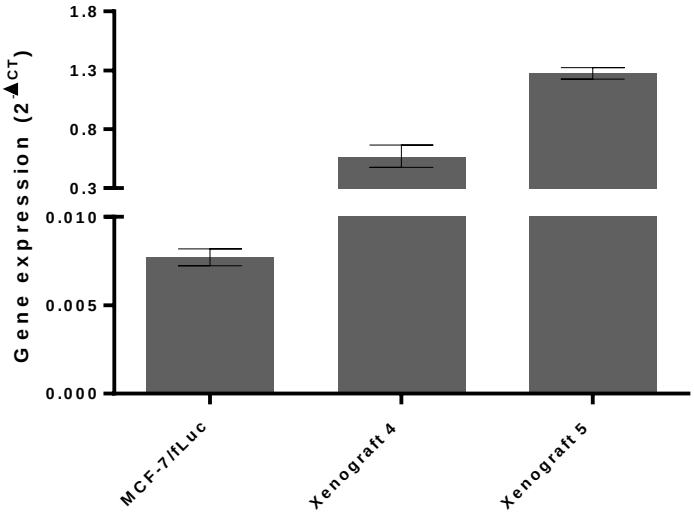


Figure 78. Vimentin gene expression in Xenografts and control (bioluminescent MCF-7 cells). Bars represent the average $2^{-\Delta CT}$ values and lower and upper limits as follows: MCF-7/fLuc = 0.0076, xenograft 4 = 0.5627 and xenograft 5 = 1.2740. N=1.

5.2.2.2.3 S100A4 protein expression

The xenografts were also stained with anti-human S100A4 or a control antibody. No background staining was observed (see appendix 16). No staining was observed in the xenografts from group 1 (Figure 22A-C). Most of the cells from the xenografts in group 2 expressed S100A4 (Figure 78D-F). The staining was nuclear, cytoplasmic and membranous.

In summary, MSCs promoted tumour growth of MCF-7 cells and induced EMT-like traits, including lower expression of E-cadherin, and higher expression of Vimentin and S100A4. On the other hand, MSCs did not result in metastasis. This model would be ideal to test the activity of the EMT reporters *in vivo*. Unfortunately, for reasons that were not in the

author's control, the MF-1 nude mice were no longer available for use. Therefore the study was repeated in a different mouse strain, in an attempt to generate similar data.

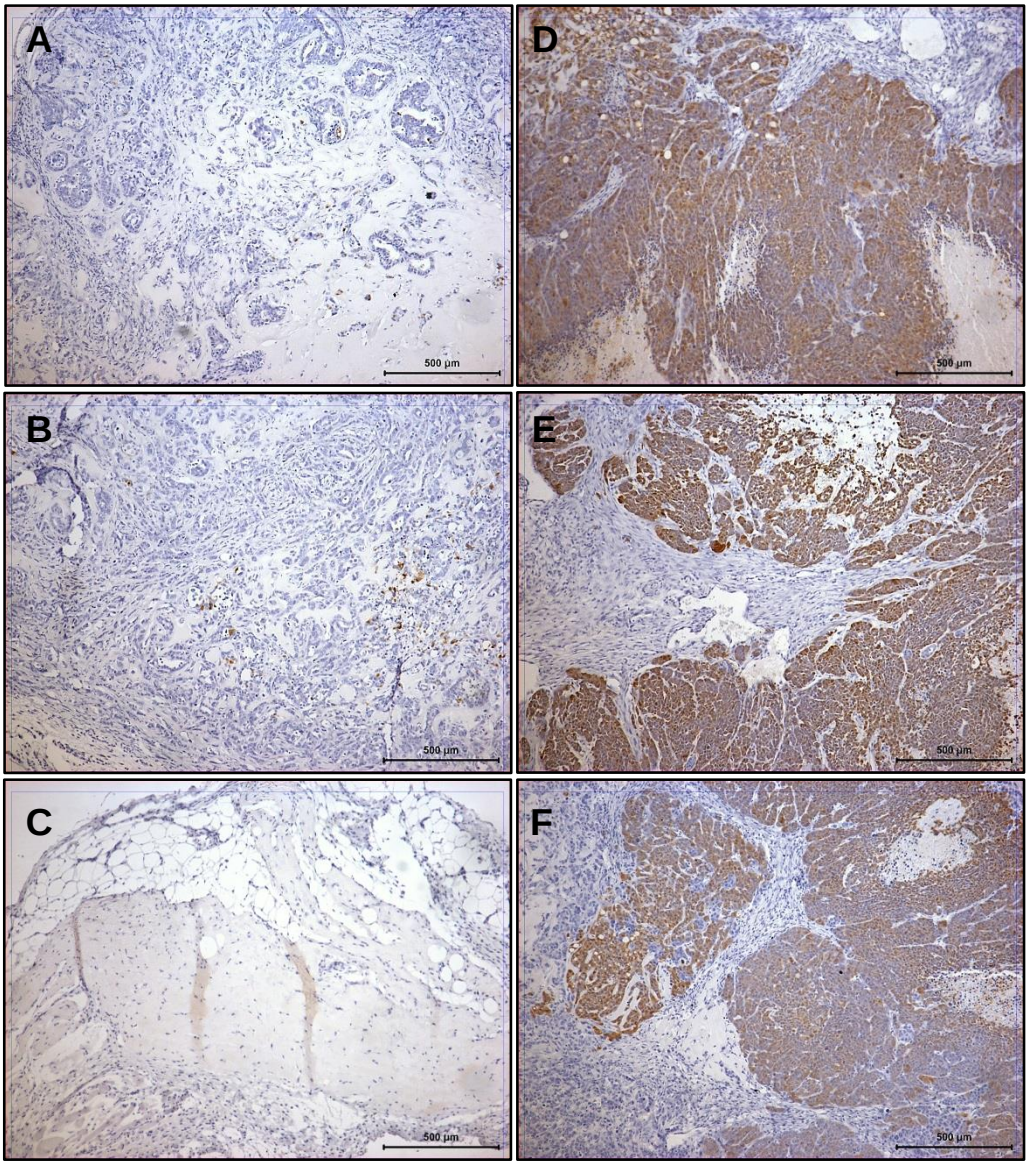


Figure 79. IHC analysis of S100A4 in mouse xenografts. A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-6 (Group 2; MCF-7 + MSCs).

5.2.3 Bioluminescent MCF-7 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model – Rag2$^{-/-}$ γc$^{-/-}$ mice

This study was performed in order to test whether the data produced in the nude mice would also be observed in the Rag2$^{-/-}$ γc$^{-/-}$ mice, before proceeding to test the EMT reporters. As previously, the mice received oestrogen.

5.2.3.1 Tumour growth and metastasis

The bioluminescent MCF-7 cells were implanted in the MFP site of the mice either alone or with MSCs at the same cell number and MCF7:MSC ratio used before. The tumour formation and growth was monitored by calliper measurements and bioluminescent imaging every 1-2 weeks as before. As shown in Figure 79, the MSCs had an effect on tumour formation/growth very early in the study. An increase in the tumour volumes and bioluminescent signal was observed early in the study (around day 20) (Figure 79). The bioluminescent signal in group 2 then reached a plateau until the end of the study, while a drop in the signal in group 1 was observed by the end of the study. A drop in the tumour volumes was observed in both groups. In other words, the tumours did not grow well and metastasis was not observed.

5.2.3.2 *Ex vivo* analysis of EMT traits

The expression of E-cadherin, Vimentin and S100A4 was analysed in the xenografts by IHC.

5.2.3.2.1 E-cadherin protein expression

E-cadherin expression was detected in all the xenografts from both groups (Figure 80). The effect observed before (lower E-cadherin expression in the “MCF-7 + MSCs” group) was not reproduced in this study. For the isotype controls see appendix 17.

5.2.3.2.2 Vimentin protein expression

The xenografts were stained for Vimentin. A positive control was included in the analysis to ensure that the staining was successful (Figure 81A).

No Vimentin expression was detected in either of the xenografts (Figure 81C-E). The effect observed before (higher expression of Vimentin in the “MCF-7 + MSCs” group) was not observed with this model. For the isotype controls see appendix 18.

5.2.3.2.3 S100A4 protein expression

The xenografts were stained for S100A4. A positive control was included in the analysis to ensure that the staining was successful (Figure 82A). No S100A4 expression was detected in either of the xenografts (Figure 82C-E). The effect observed before (upregulation of S100A4 in the “MCF-7 + MSCs” group) was not reproduced in this study. For the isotype controls see appendix 19.

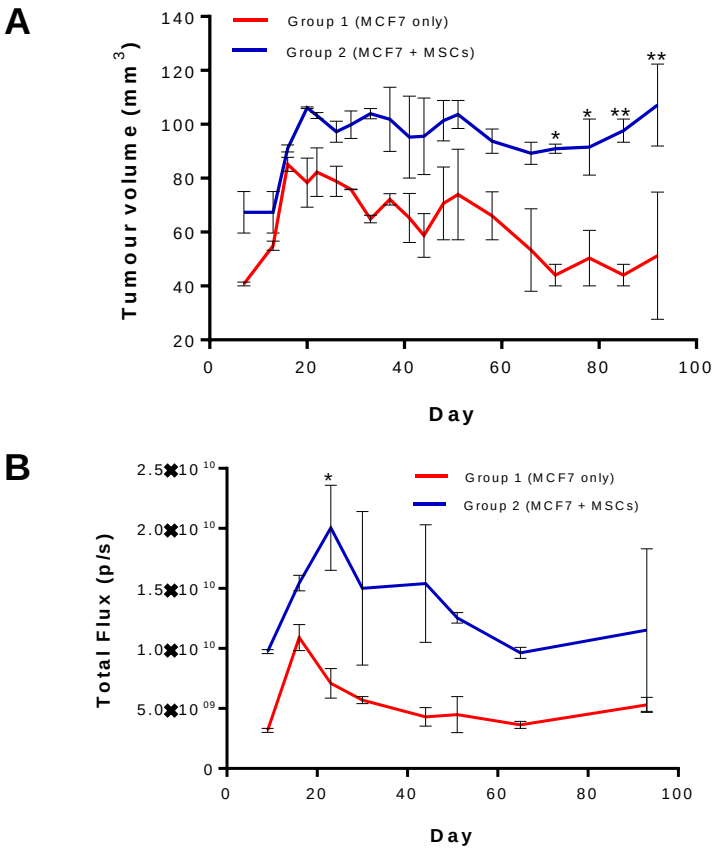


Figure 80. Tumour formation and tumour growth of bioluminescent MCF-7 cells in MFP mouse models. A. Tumour volumes (mm^3) in group 1 (MCF-7 only) and group 2 (MCF-7 + MSCs). **B.** Bioluminescent imaging (total flux; photons/second) in group 1 and group 2. Each bar represents the average volume (A) or average total flux (B) ± SEM. N=2 per group. Statistical analysis was performed with two-way Anova. *P≤0.05 and **P≤0.01.

In summary, MSCs did not induce any EMT-like traits in the xenografts formed in this strain of mice. Furthermore, the growth promotion effect observed in this study was not as great as observed before (Figure 72). This confirms that the EMT-like traits observed in the study with the MF1 nude mice, were associated with MSC-induced tumour growth. As in the study with the MF1 nude mice an effect of the MSCs on tumour growth was observed at the early stages of the study (Figure 72), it was hypothesised that the MSCs may have depleted early in the study and therefore did not promote tumour growth. Thus, an additional study was designed in which MSCs would be co-implanted with the MCF-7 and injected again a few days after the initiation of the study ("MSCs boost").

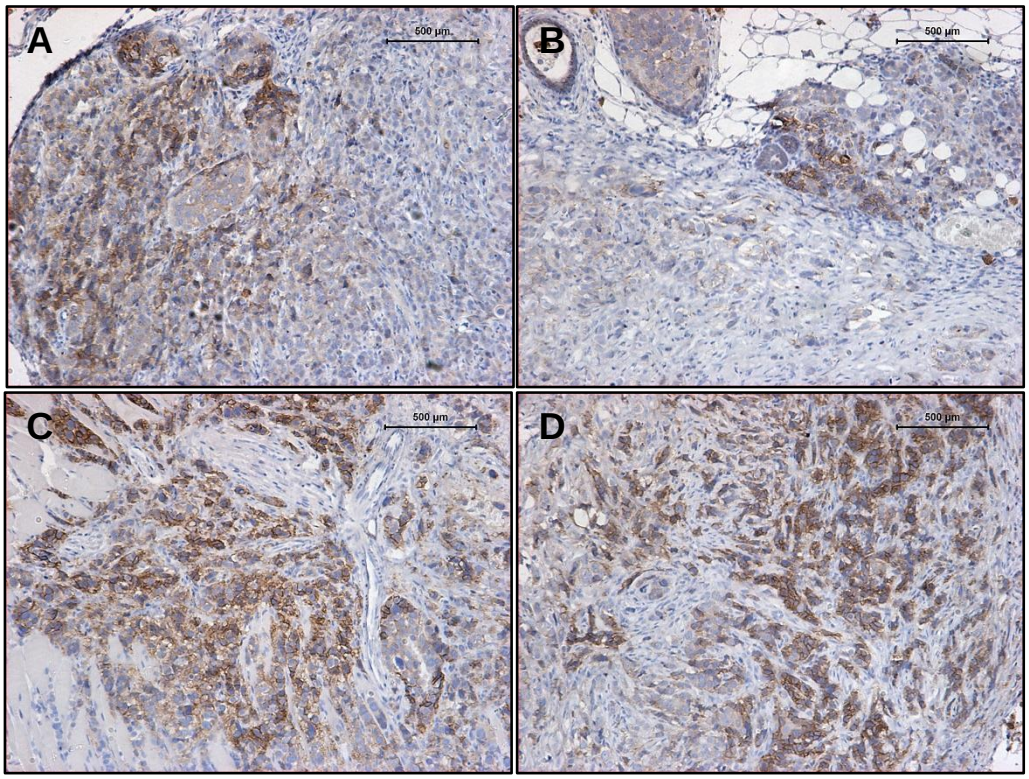


Figure 81. IHC analysis of E-cadherin in mouse xenografts. A-B. Xenografts from mice 1-2 (Group 1; MCF-7 alone). **C-D.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs).

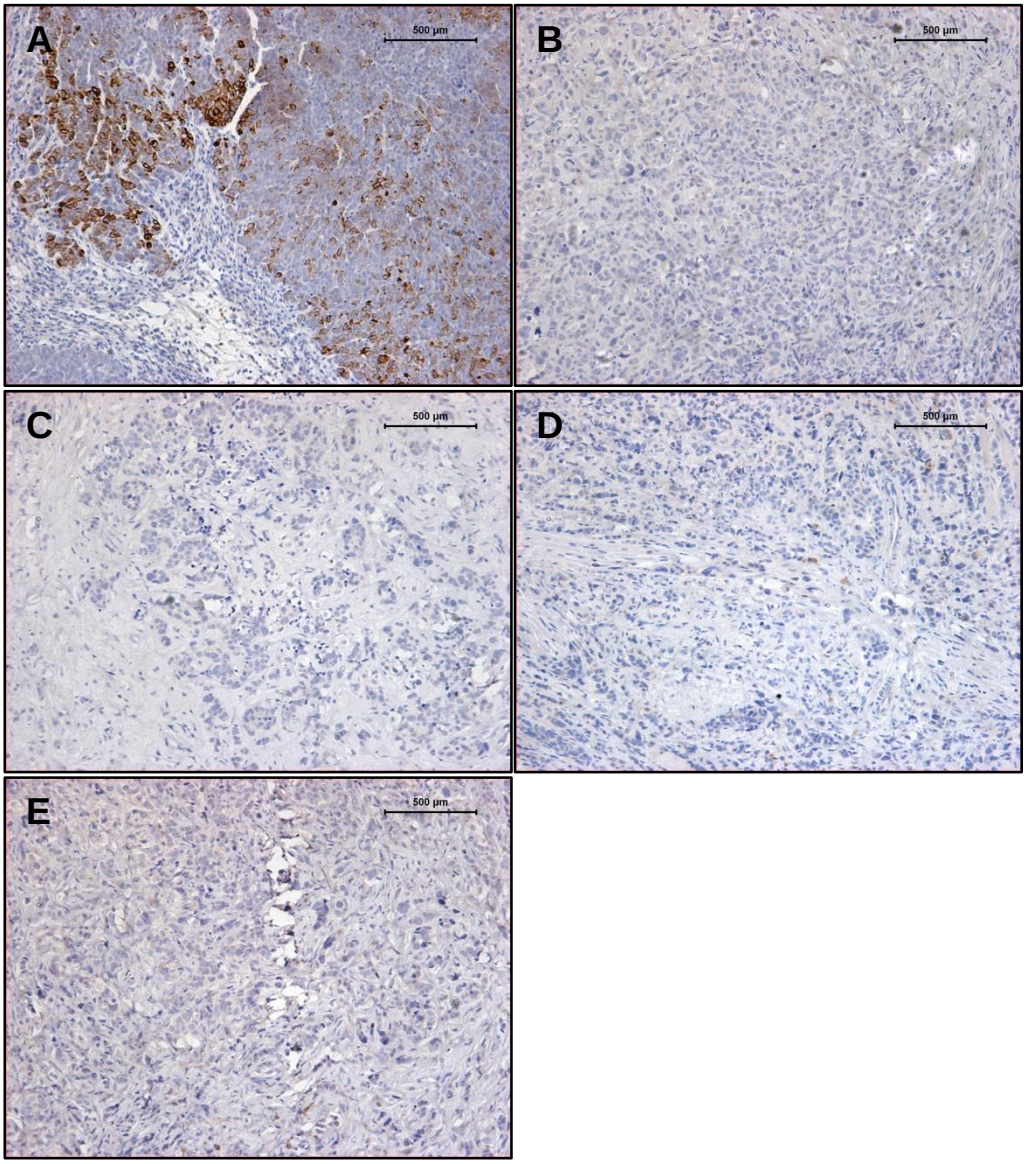


Figure 82. IHC analysis of Vimentin in mouse xenografts. A. Positive control (xenograft 6, Figure 81A). **B-C.** Xenografts from mice 1-2 (Group 1; MCF-7 alone). **D-E.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs).

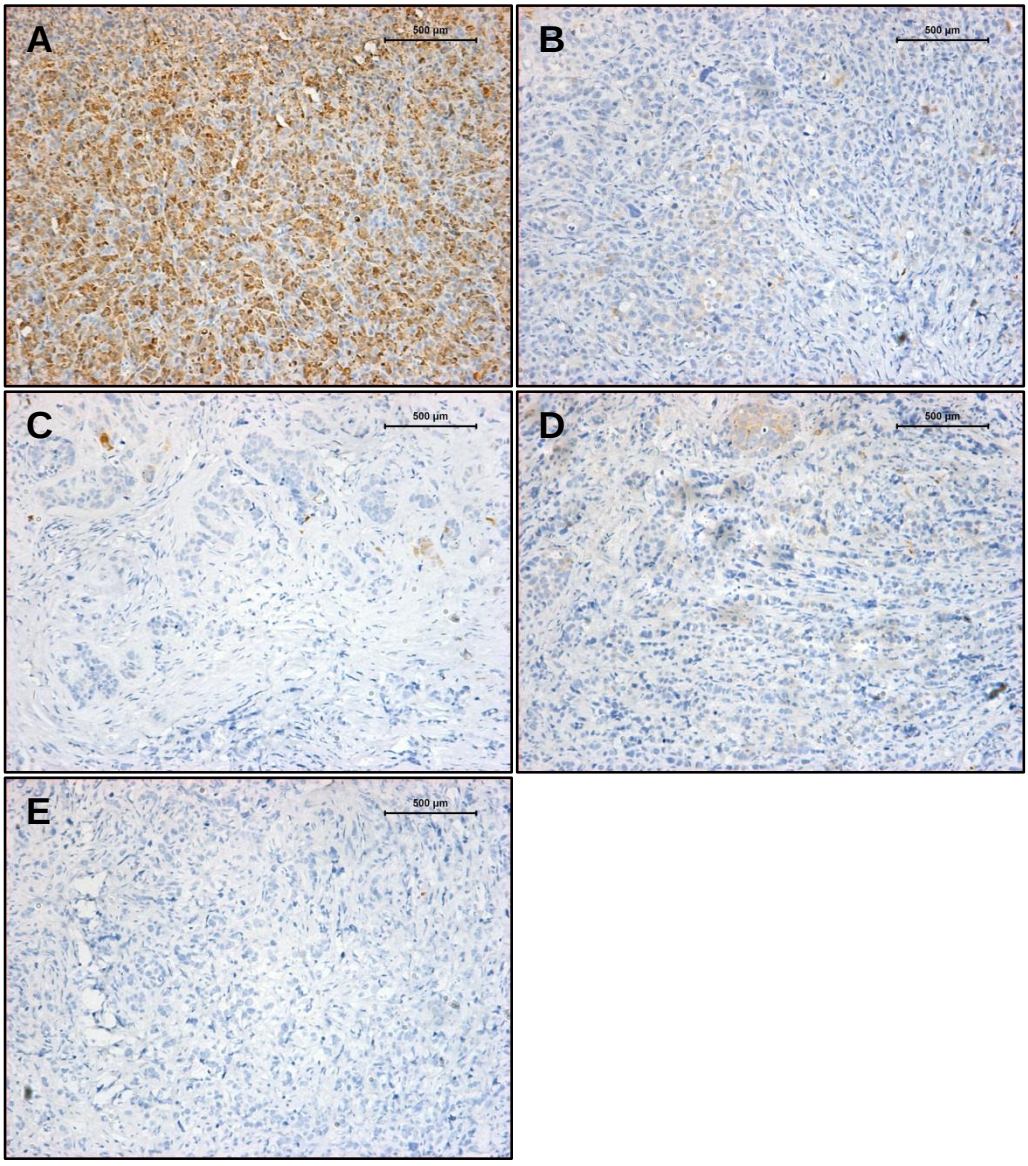


Figure 83. IHC analysis of S100A4 in mouse xenografts. A. Positive control (MDA-MB-231 xenograft from a study performed in house). **B-C.** Xenografts from mice 1-2 (Group 1; MCF-7 alone). **D-E.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs).

5.2.4 Assessing the effect of "MSCs boost" in a bioluminescent MCF-7 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model – Rag2$^{-/-}$ γc$^{-/-}$ mice

This study aimed to overcome possible depletion of MSCs early in the study by injecting MSCs both at initiation and at later stages. Here, bioluminescent imaging was performed immediately after initiation of the study (rather than a week later as previously) to better determine the timing of the effect of MSCs.

5.2.4.1 Tumour growth and metastasis

Bioluminescent MCF-7 cells were co-implanted in the MFP of Rag2 mice as before. An equal amount of MSCs, as implanted initially, was injected in the mice at day 21. Tumour establishment and tumour growth was monitored by measuring the tumour volumes and bioluminescent signal. The tumours established early in the study and then reached a plateau, as indicated by the bioluminescent imaging (Figure 83A&C). Apart from day 96, where the signal was unexpectedly higher in the "MCF-7/fLuc only group), there was significant different between the two groups, as assessed by a two-way Anova test (Figure 83A). Furthermore, there was no significant difference in the tumour volumes between the two groups (Figure 83B). Looking at the individual tumour volumes (Figure 83D) only the tumour in mouse 6 from the "MCF-7/fLuc + MSCs" group was grew substantially throughout the study. It is worth noting, there is a variation in the tumour volumes in this group, as opposed to the "MCF-7/fLuc" group, which might suggest a difference in the effect of MSCs in the different mice. At the end of the study, a post-mortem image was taken in order to detect any metastasis, but no metastasis was observed. Also, the tumours were removed, imaged and weighed. The bioluminescent signal was proportional to the tumour weight (Pearson r=0.8247, p=0.0118), but no significant difference was observed in either the bioluminescent signals or the tumour weights between the two groups

(data not shown). In summary, it was shown that the MSCs and the "MSCs boost" did not have a significant effect on tumour growth.

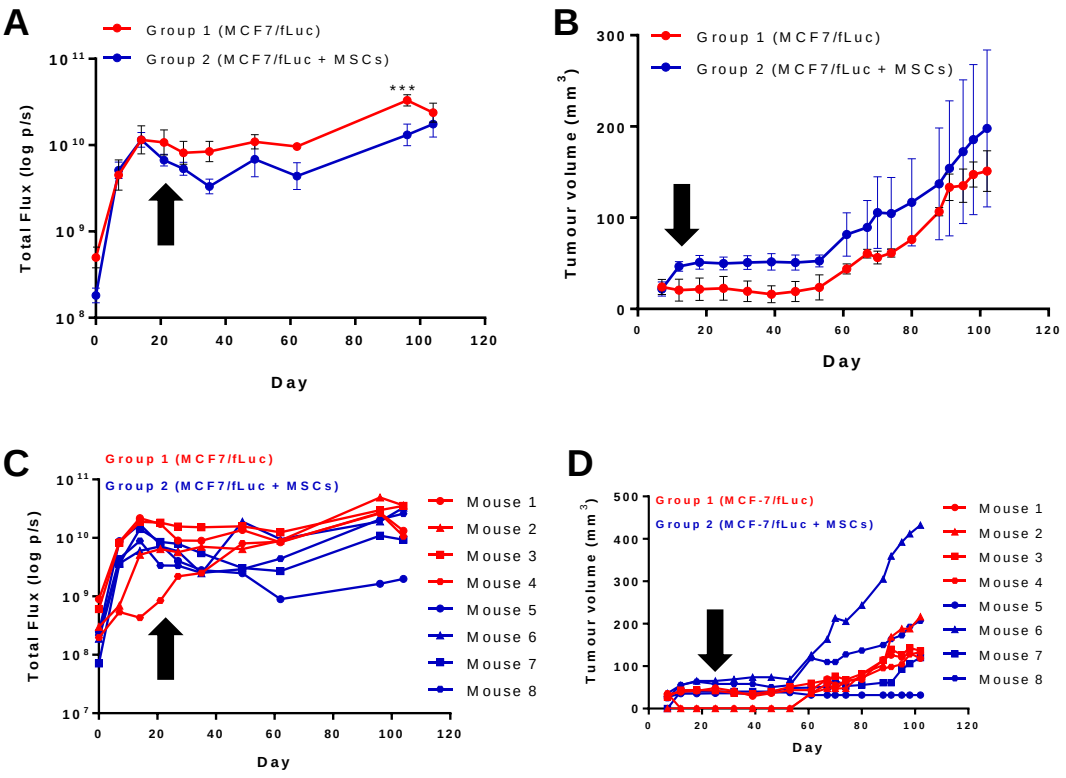


Figure 84. Tumour formation and tumour growth of bioluminescent MCF-7 cells in MFP mouse models. A. Bioluminescent measuments (total flux; photons/second) as measured after administration of D-luciferin. **B.** Tumour volumes (mm^3). Each bar represents the average total flux (A) or volume (B) ± SEM. N=4 per group. Statistical analysis was performed with two-way Anova. *P≤0.05, **P≤0.01 and ***P≤0.001. **C.** Bioluminescence of individual mice. **D.** Tumour volumes of individual mice. The arrow indicates the day of "MSCs boost".

5.2.4.2.1 E-cadherin

E-cadherin expression was detected in all the xenografts from both groups (Figure 84). While all the cells within the tumours stained positive for E-cadherin, the density of the staining varied within the same xenograft i.e. both strong and weak signals were detected in the same sample. Furthermore, while in certain xenografts the staining was mostly strong (xenografts 1, 2, 3 and 5; Figures 84A, B, C and E), in others the signal was mostly weak (xenografts 4, 6, 7 and 8; Figures 384E, F, G, and H). No particular difference in pattern/intensity of staining was observed between the two groups, showing that MSCs did not a substantial effect on E-cadherin expression. Strong staining was also observed in xenografts 6; the only one that had the effect of MSCs on growth. For isotype controls see appendix 20.

5.2.4.2.2 Vimentin

The xenografts were stained for Vimentin. None of the xenografts expressed Vimentin (Figure 85), indicating that MSCs did not increase the expression of MSCs, as it was detected in a previous study (Figure 75). A positive control was included, which was successfully stained, indicating that the technique worked. For isotype controls see appendix 21.

5.2.4.2.3 S100A4

The xenografts were also stained for S100A4. For isotype controls see appendix 22. Expression of S100A4 was detected only in xenograft 6 (Figure 86F); it was the only xenografts where an effect of MSCs was seen (both in tumour growth promotion and S100A4 expression).

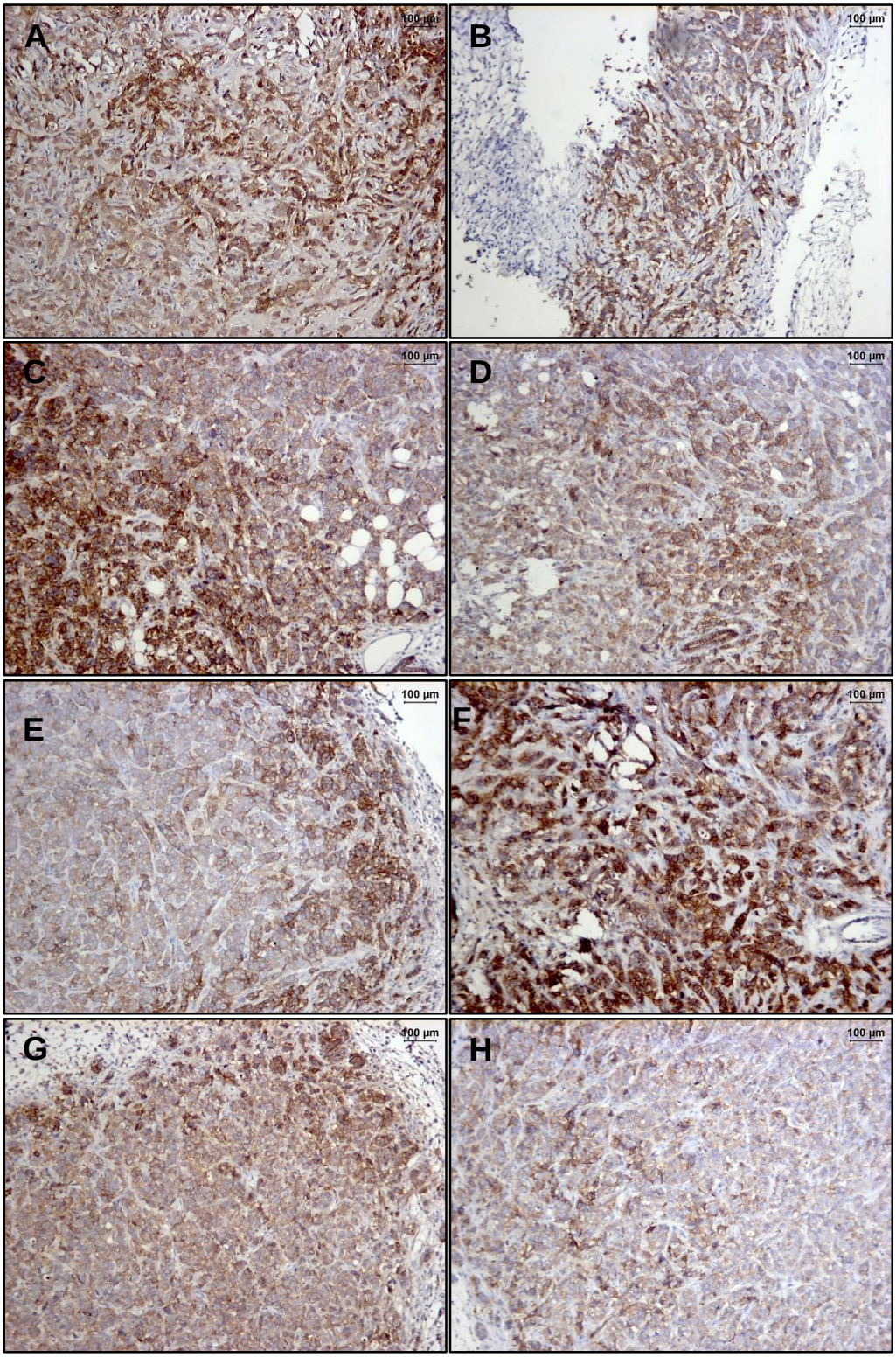


Figure 85. IHC staining of E-cadherin in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/fLuc). **E-H** Mice 5-8 from group 2 (MCF-7/fLuc + MSCs).

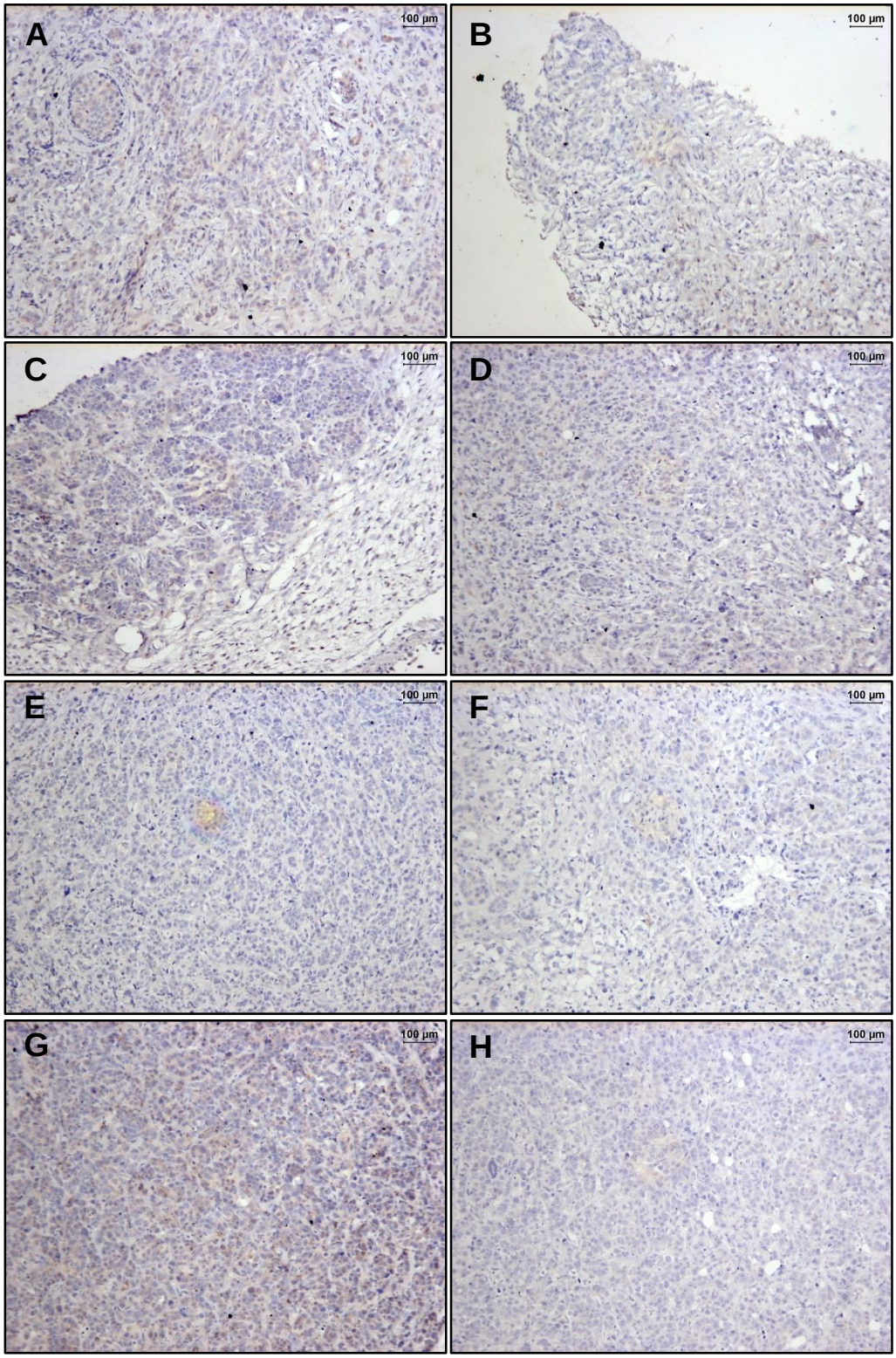


Figure 86. IHC staining of Vimentin in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/fLuc). **E-H** Mice 5-8 from group 2 (MCF-7/fLuc + MSCs).

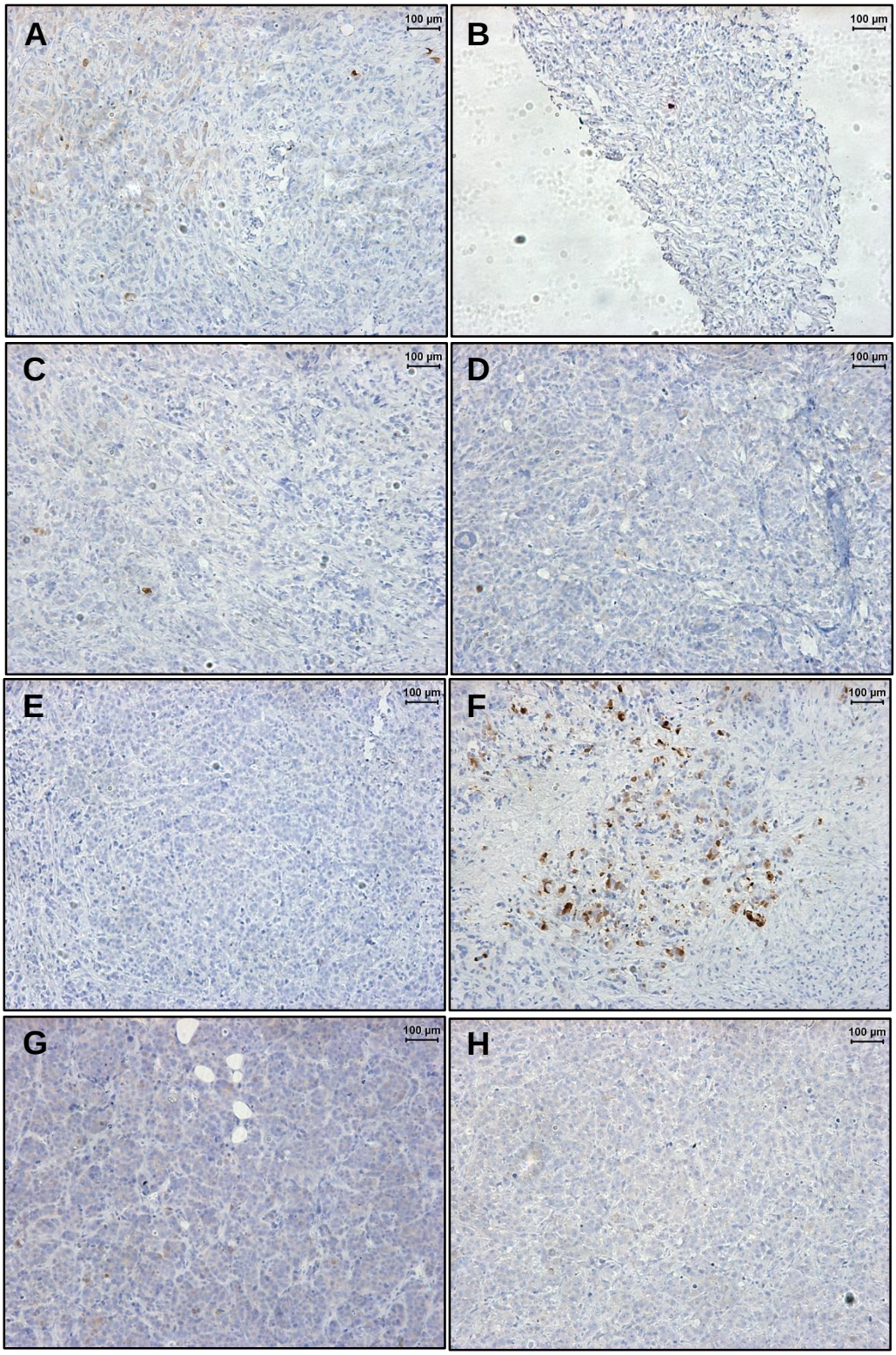


Figure 87. IHC staining of S100A4 in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/fLuc). **E-H** Mice 5-8 from group 2 (MCF-7/fLuc + MSCs).

Summary and discussion

In this chapter, the different *in vivo* studies were described, which were performed in order to assess the tumorigenicity of the BC cell lines, and generate a model of EMT and/or metastasis, to assess the activity of the S100A4 reporter.

Initially, the tumorigenicity of the MCF-7 and BT59 cells was tested in different implantation sites. There was no significant difference in the growth between the two different sites for the MCF-7 cells and the BT549 cells did not grow. Furthermore, no metastasis was observed. When MSCs were co-implanted with MCF-7 in an orthotopic nude model, the tumour growth was significantly upregulated and EMT-like traits were induced. When BT549 were co-implanted with MSCs, no significant tumour growth promotion was observed. In the Rag2 model, neither promotion of tumour growth was observed or EMT-like traits. It was hypothesised that the MSCs were depleted early in the study; therefore, a different study in which a “MSCs boost” was included was performed in the Rag2 mice. Here, no growth promotion was observed, but S100A4 expression was detected in a particular xenograft, which was the biggest and belonged to the “MCF-7/fLuc + MSCs” group.

In the different studies, variation in the tumour growth was observed both between and within the different groups. In some cases, growth was only observed in a particular mouse. One possible explanation for this is the difference in the amount of oestrogen uptake by the mice. As explained earlier, oestrogen supports the tumour growth of MCF-cells [180]. In our studies the oestrogen (estradiol) was included in the diet; therefore, difference in the amount of food consumed by each mouse would result in differences in the oestrogen uptake. The reason for choosing to give the oestrogen in the diet rather than in the form of a pellet was based on previous experiments performed in house that showed that administration of estradiol resulted in well-documented side effects such as bladder calculi and urine scald. Another difference that could explain the variation in the results, concerns the technicality of implanting the cells in the MFP site. In our studies, the cells were implanted into the 4th

mammary fad pad located in the abdomen, which is the most common technique used. This approach is based on the rationale that the 4th mammary pad is the largest and relatively easy to access, and is thus being assumed that the cells will be successfully implanted in the MFP. This technique is not guided by a microscope as opposed to an orthotopic implantation under direct vision (ODV), and can miss the targeted MFP entirely, resulting in intradermal, subcutaneous or intramuscular implantation [101].

While MSCs promoted the tumour growth of MCF-7 cells in MF1 nude mice, they did not promote tumour growth in the Rag2 mice. The role of MSCs in tumour progression per se is controversial (reviewed in [187]). Many studies have shown that MSCs promote the tumour growth and metastasis of BC [72, 74, 184, 188], but inhibition of tumour growth has also been reported [189]. MSCs also induced EMT-like traits in the orthotopic MCF-7 models, including downregulation/loss of E-cadherin, upregulation of Vimentin, and upregulation of S100A4. MSC-induced EMT in BC associated with loss of E-cadherin and upregulation of Vimentin have been demonstrated *in vitro* [73]. To our knowledge, this is the first study in which hMSCs were co-implanted with MCF-7 cells and loss of E-cadherin and upregulation of Vimentin was observed. Furthermore, we showed that the MSCs increased the expression of S100A4, which plays an early role in EMT [121], and this was accompanied (in the case of nude mice) with the loss of E-cadherin and Vimentin. In the Rag2 mice, upregulation of S100A4 was only observed in one mouse that belonged to the "+ MSCs group", which again indicated MSC-induced expression of S100A4. In this case, no other EMT-like traits were observed, suggesting that the cells were maybe in an early stage of undergoing EMT.

In summary, it was shown that MSCs have the ability to promote the tumour growth of MCF-7 cells and/or induce EMT-like like traits (loss of E-cadherin, upregulation of Vimentin and upregulation of S100A4), when co-implanted in the MFP of nude mice. MSCs can act as a progenitor for CAFs which are characterised by the expression of S100A4 [34].

Induction of EMT in MCF-7 cells by CAFs associated with loss of E-cadherin and acquisition of Vimentin has been previously reported [167, 190]. Therefore, based on the findings, we propose that MSCs generated CAFs expressing S100A4 (and hence the S100A4 expression), which in turn induced EMT-like traits (E-cadherin downregulation/loss and upregulation of Vimentin).

The S100A4 reporter will be tested in an orthotopic MCF-7 model, in which MSCs will be co-implanted and later injected in the study.

CHAPTER 6: RESULTS 4 – Assessing the activity of the EMT reporter *in vivo*

This chapter describes the assessment of the *in vivo* activity of the S100A4 and its refinement.

The activity of the reporter had been validated in MCF-7 and BT549 cells in *in vitro* models, under hypoxic conditions (Chapter 3). It was shown that *S100A4* expression was upregulated under hypoxia and that the bioluminescent signal was also upregulated. In chapter 5, a number of *in vivo* studies were performed, in order to identify a model of EMT and/or metastasis or a model where changes in the S100A4 expression would be detected. It was shown that co-implanting the MSCs with the MCF-7 cells in an orthotopic nude mouse model resulted in tumour growth promotion and induction of EMT-related traits, including upregulation of S100A4 expression. In Rag2 mice, expression of S100A4 was observed only in one xenograft, which belonged to the group in which MSCs were co-implanted and injected later in the study. The particular xenograft was also the biggest. Since the nude mice were not available to use, the study here was performed using the Rag2 mice, where MSCs were co-implanted with MSCs and an “MSCs boost” was received as before.

The model was designed in way that the tumour growth and the activity of the S100A4 reporter was validated in the same mouse at different times, which avoids the need to sacrifice multiple groups of mice at different end-points; hence, the total number of animals needed is reduced. The tumour growth was monitored by fluorescent readings and the activity of the reporter was monitored by bioluminescent readings. The model was assessed by correlating the fluorescent and bioluminescent signal with the tumour volume. The activity of the reporter was assessed by measuring the bioluminescent signal and comparing with the *ex vivo* analysis of the S100A4 expression. To assess whether the reporter detected changes in EMT, *ex vivo* analysis of EMT markers was performed and this was correlated with reporter activity.

6.1 Design of the model

The MCF-7 cells expressing the S100A4 reporter were transduced with a lentiviral vector constitutively expressing the fluorescent marker mCherry (pLVX-CMV-mCherry) and stable transductants were selected by Flow Cytometric Cell Sorting. The resulting cell line will be referred as "MCF7/S100A4-fLuc/mCherry". The constitutive expression of mCherry allowed the tumour growth to be monitored. The activity of the S100A4 reporter was monitored by bioluminescent imaging. The cell line ($2x10^6$ cells /mouse) was co-implanted in the MFP of Rag2 mice with or without MSCs ($1x10^6$ cells/ mouse) and the mice received a "MSCs boost" ($1x10^6$ cells/ mouse), as before.

For more accurate readings, the images were adjusted to remove excess autofluorescence and the background fluorescent signal was subtracted from the signal in the tumours (Figure 87A). Occasionally, the resulting signal had a negative value. This was due to the fact that the machine generated artifacts when the values of the tumour and background signal were similar. The negative values were included in the graphs as zero. Following fluorescent imaging, the mice received D-luciferin and the bioluminescence signal was measured, to assess the activity of the reporter. A representative image showing fluorescent and bioluminescent images in the same mice is shown in Figure 87.

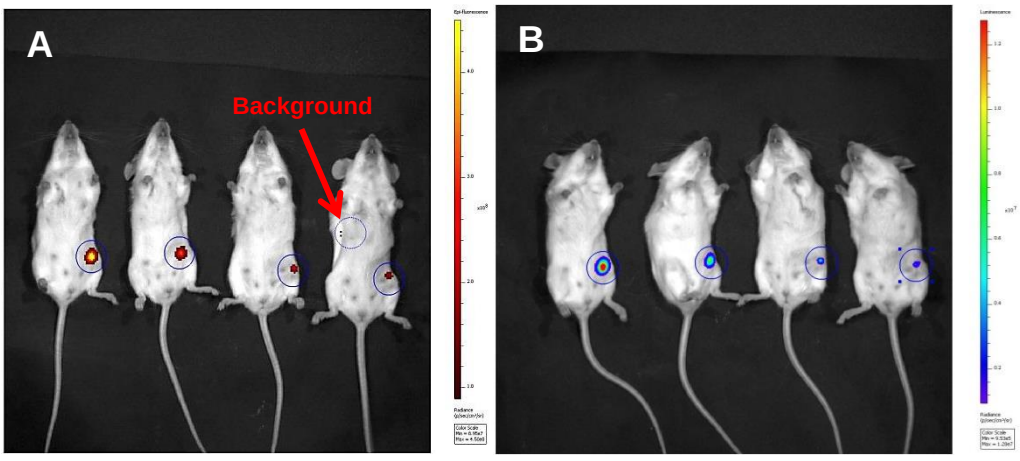


Figure 88. Fluorescent and bioluminescent readings. A. Fluorescent measurements. Each circle represents the regions for measuring tumour signal and the dotted circle shows the background signal. **B.** Bioluminescent readings of the same group. Each circle represents the regions for measuring signal. Scale=total flux (p/s)

6.2 Tumour growth

Tumour establishment and growth was monitored by fluorescent imaging and calliper measurements. At the beginning of the study, a discrepancy in the signal was observed between the two groups but no significant difference was observed between the two groups throughout the study (Figure 88A). An increase in the fluorescent signal was observed at the beginning of the study (from day 7 to day 14) in both groups, then reached a plateau, increased at day 62 and then remained the same until the end of the study (Figure 88A). The average fold increase in the signal at the end of the study compared to day 7 was 11.7 (16.5 compared to day 0). Similar to the fluorescent signal, the tumour volumes increased at the beginning of the study, and then reached a plateau, then started to increase again (Figure 88B). Unlike the fluorescent signal, the tumour volumes kept increasing until the end of the study. The average fold increase in the volume at the end of the study compared to day 7 was 13.5. No significant difference was observed between the two groups as assessed by a two-way Anova test. Looking at the individual mice, the fluorescent signal was similar between and within the groups (Figure 89A). The tumour volumes of individual mice in group 1 ranged from 13.5mm^3 (mouse 3) to 167.4mm^3 (mouse 1) (Figure 89B). A bigger variation was observed in group 2, where extremely low volumes were observed in mice 5 and 6, but big tumour volumes were observed in mice 7 and 8. It is worth noting that in that the higher fluorescent signal and higher tumour volumes were observed in these mice (7 and 8), both of which belonged to the group in which MSCs were co-implanted (Figure 89). The tumour volumes were measured weekly, but for direct comparisons between the volume and the signal, only specific points are shown in the graphs.

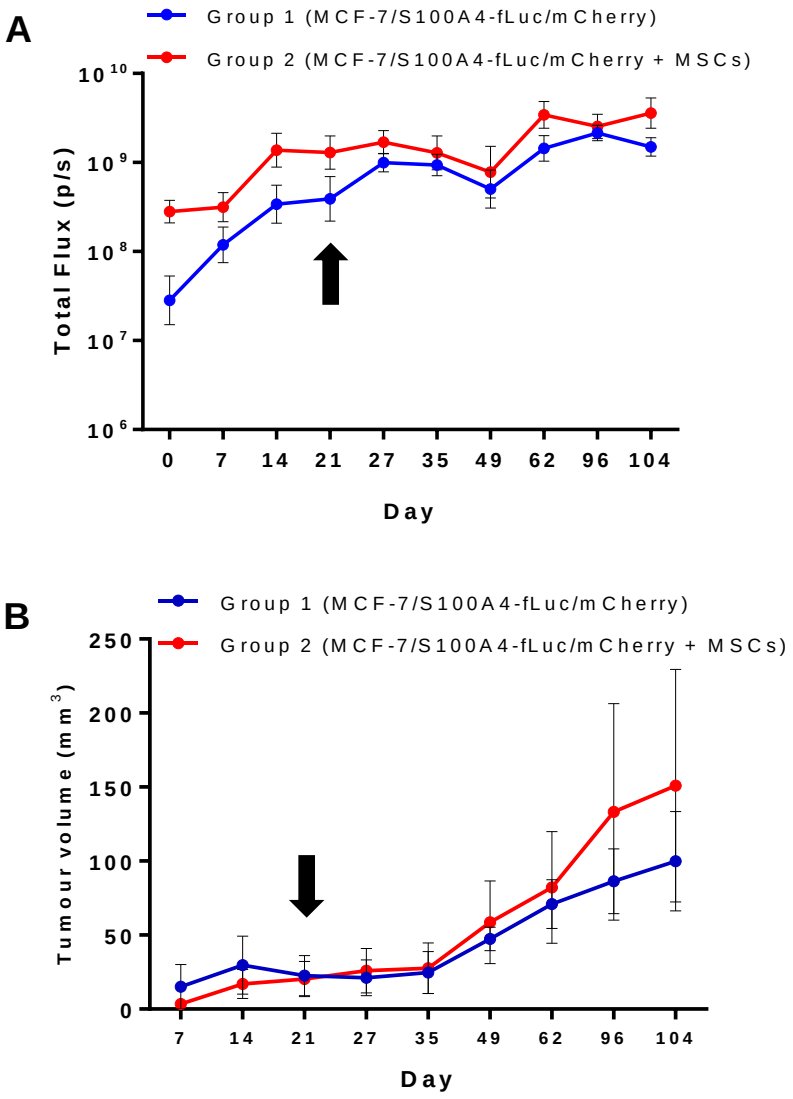


Figure 89. Tumour formation and tumour growth of the reporter line in MFP mouse models. A. Fluorescent measuments (total flux; photons/second) **B.** Tumour volumes (mm^3). Each bar represents the total flux (A) or volume (B) ± SEM. N=4 per group. Statistical analysis was performed with two-way Anova. *P≤0.05, **P≤0.01 and ***P≤0.001. The arrow indicates the day of "MSCs boost".

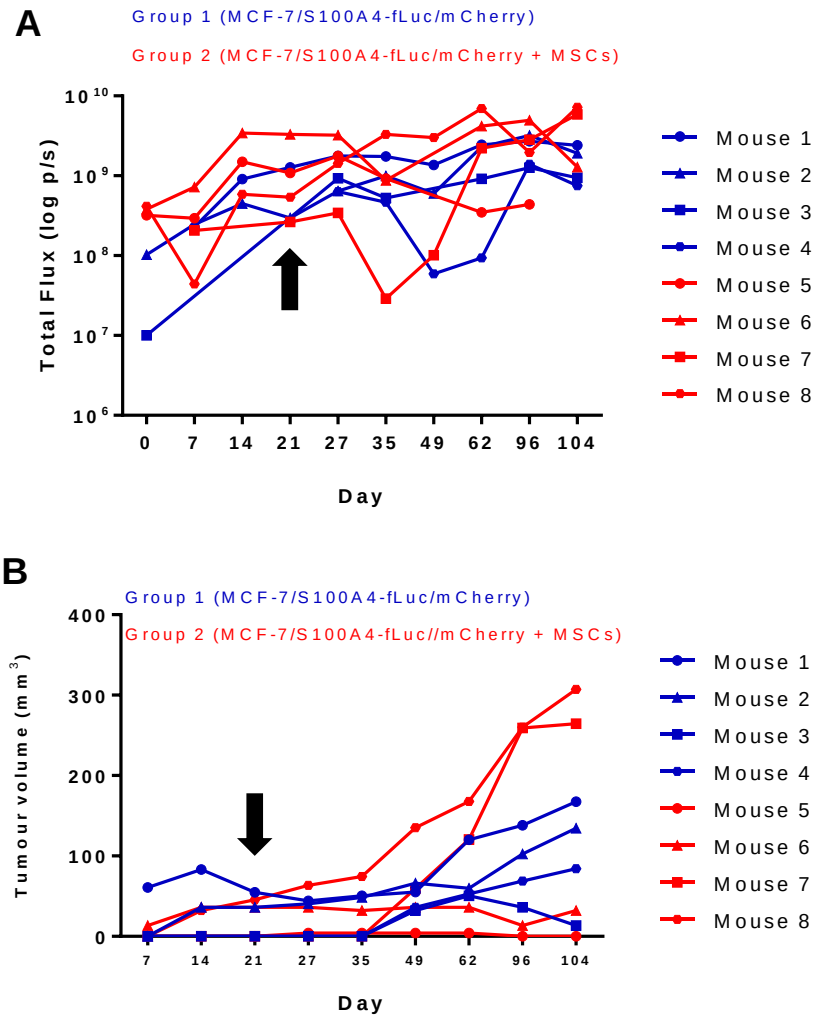


Figure 90. Tumour formation and tumour growth of the reporter line in MFP mouse models in individual mice. A. Fluorescent measurements (total flux; photons/second). **B.** Tumour volumes (mm^3) in individual mice. Each point represents the total flux (A) or volume (B) The arrow indicates the day of "MSCs boost".

6.3 Measuring reporter activity

The bioluminescent signal increased in the beginning of the study and then reached a plateau (Figure 90). Significant difference in the bioluminescent signal was observed only at day 96 (Figure 90A). The average fold increase in the signal at the end of the study compared to day 7 was 8.6 (71.3 compared to day 0). Looking at the readings from individual mice (Figure 90B), there was a peak in the signal in all the mice from group 2 at day 96, which dropped by the next imaging.

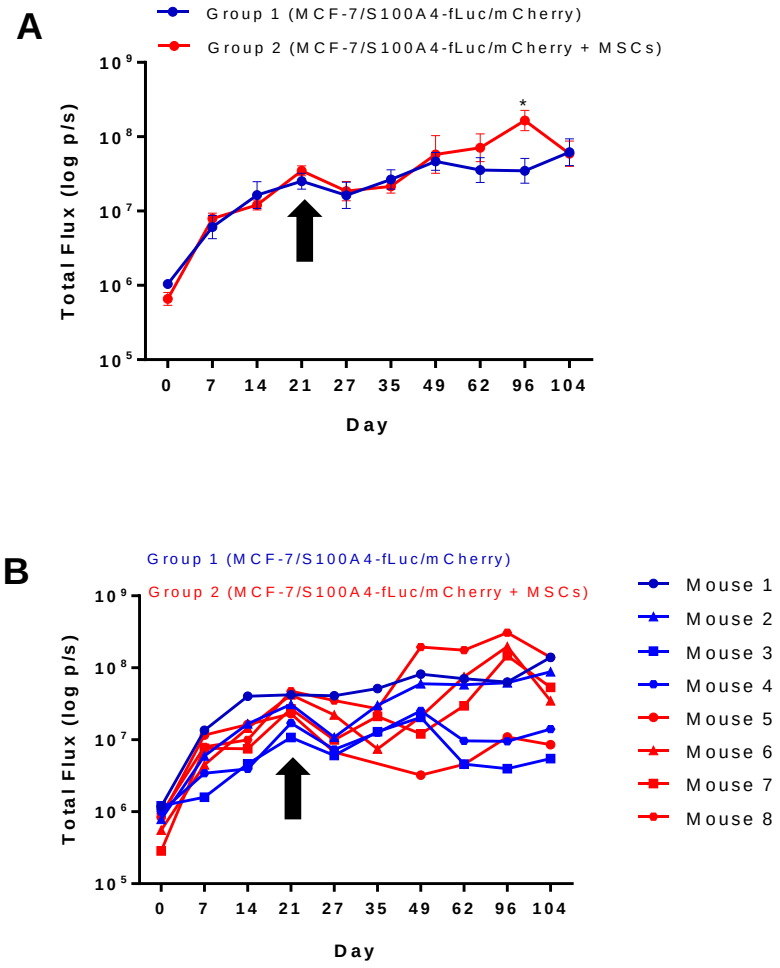


Figure 91. Bioluminescent readings measuring reporter activity. A. Average readings in group 1 and group 2. Each point represents average total flux ± SEM. N=4 per group. Statistical analysis was performed with two-way Anova. *P≤0.05, **P≤0.01 and ***P≤0.001. **B.** Signal in individual mice. The arrow indicates the day of "MSCs boost".

6.4 Assessing/refining the model

6.4.1 Correlation between tumour volumes and fluorescent signal

As all the cells constitutively express mCherry, the fluorescent signal should serve as an indicator of tumour growth. To validate this, the relationship between the signal and the tumour volume was examined. A significant positive correlation was observed in mice 1, 2, 3, 7 and 8, as assessed by the pearson's correlation test (Figure 91A-C, G & H).

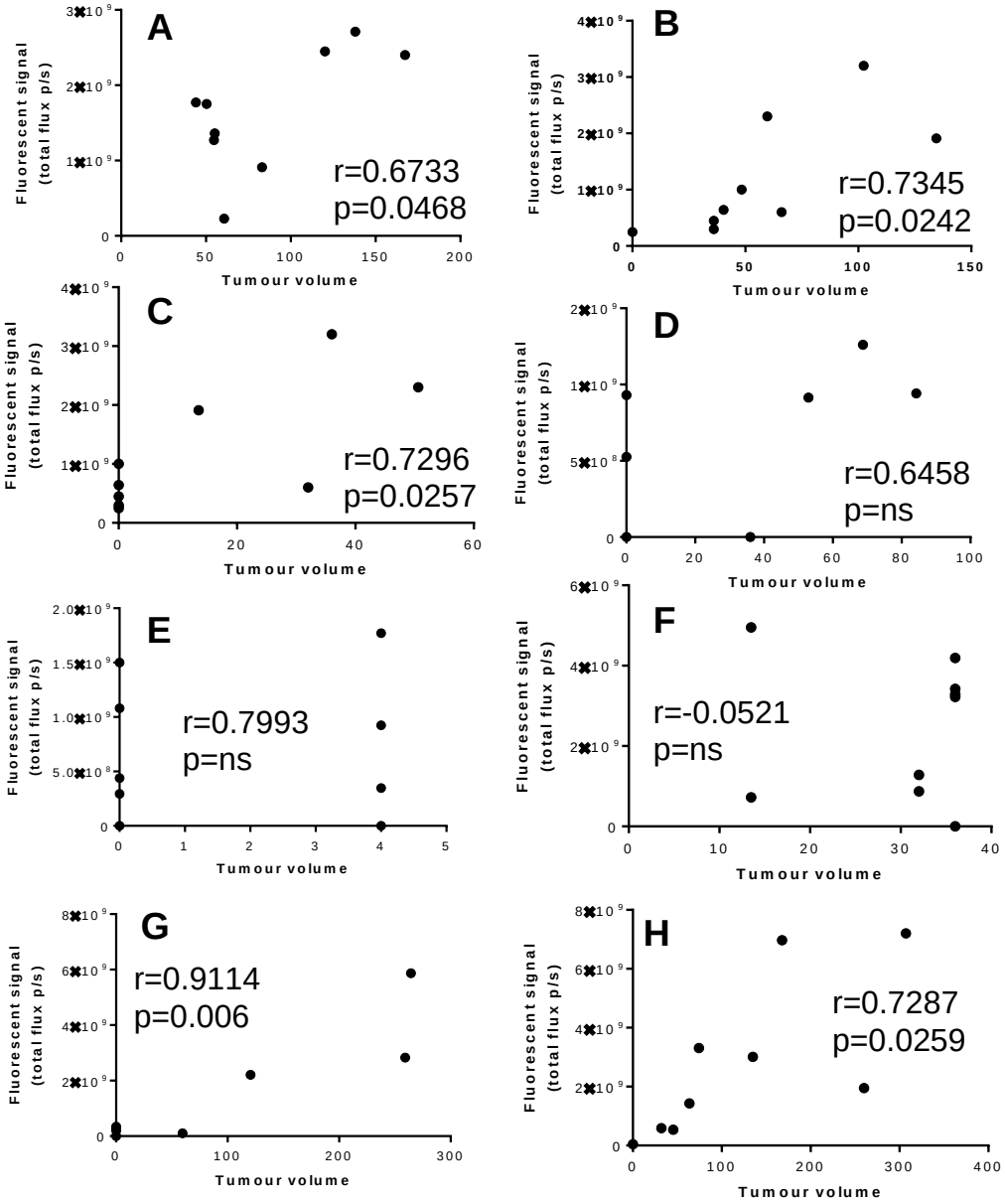


Figure 92. Correlation of fluorescent signal and tumour volume. Each point represents the tumour volume and fluorescent signal at different time-points throughout the study. **A-H.** Mice 1-8. r=pearson's correlation coefficient. N=9 (overlapping values appear as single dot). ns=not significant.

When assessing this relationship in all mice from both groups at different time points a positive correlation between the fluorescent signal and the tumour volumes was only observed at on day 49 (Figure 92A) and day 104 (Figure 92B). No significant correlation was detected in the other time points (data not shown). At day 49, a drop in the signal (Figure 92C) and an increase in the tumour volumes (Figure 92D) was observed. At day 104, there was no change in the signal or tumour volumes compared to previous measurements (Figure 92C-D).

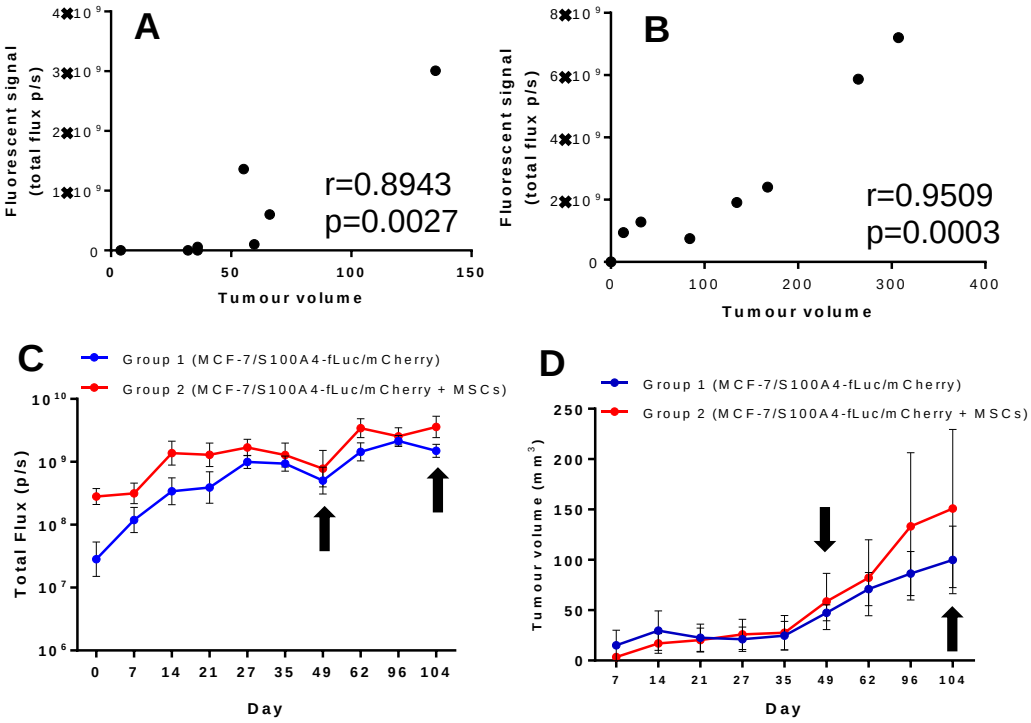


Figure 93. Correlation between fluorescent signal and tumour volume at different time points. A. Correlation analysis at day 49. **B.** Correlation analysis at day 104. Each point represents a different mouse. N=8. r=pearson correlation coefficient. p=P value.The days are pointed by the black arrows in the fluorescent **(C)** and tumour volume **(D)** graphs.

6.4.2 Correlation between tumour volumes and bioluminescent signal

In previous chapters, it was shown that the bioluminescent signal of the cells expressing the S100A4 reporter was proportional to the cell number. Therefore, we wanted to assess whether the signal would change proportional to tumour growth. It was anticipated that the bioluminescent signal would not correlate with tumour volume upon activation of the reporter, since this would be independent of cell number and hence tumour volume (if tumour volume is an indicator of cell number). A significant positive correlation was observed between the signal and the tumour volumes in mice 1, 2, 8 and 9 (Figure 93A, B, G & H). No correlation was observed in mice 3, 4 and 6 (Figure 93C, D and F), while a negative correlation was observed in mouse 5 (Figure 93E). When assessing this relationship in all mice from both groups at different time points, a positive correlation between the bioluminescent signal and the tumour volumes observed in most days throughout the study, apart from day 7, day 35 and day 96 (Figure 94). At day 96, the signal in the group where MSCs were co-implanted was significantly higher than the group where the reporter cells were implanted without MSCs. This could imply that the reporter was activated, and therefore a relationship between the tumour volume and the signal would not be expected, as the reporter activity is not dependent on cell number.

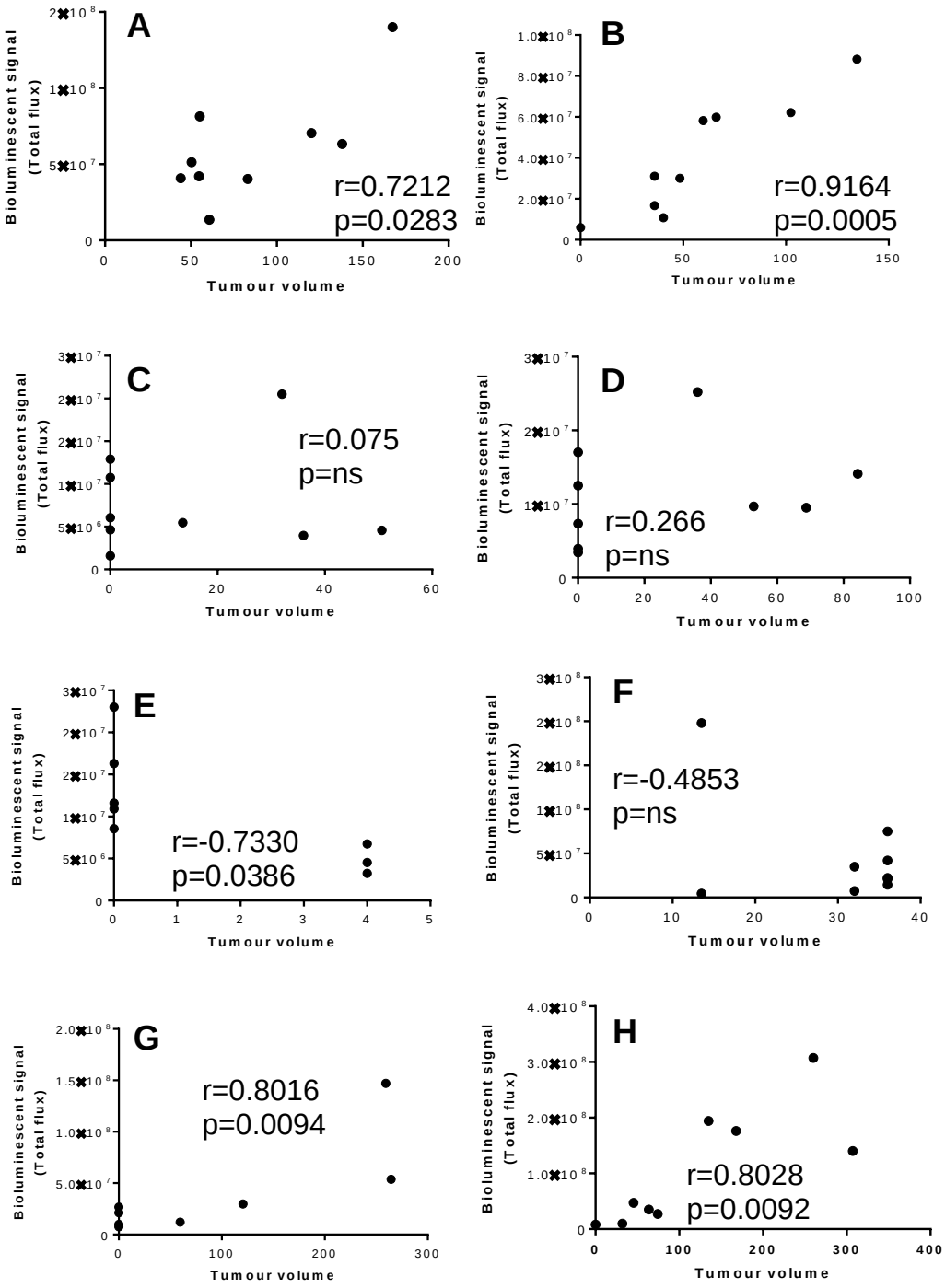


Figure 94. Correlation of bioluminescent signal and tumour volume in individual mice. A-H. Mice 1-8. Each point represents the tumour volume and fluorescent signal taken on the same day throughout the study. N=9. r=pearson correlation coefficient. p=P value. ns=not significant.

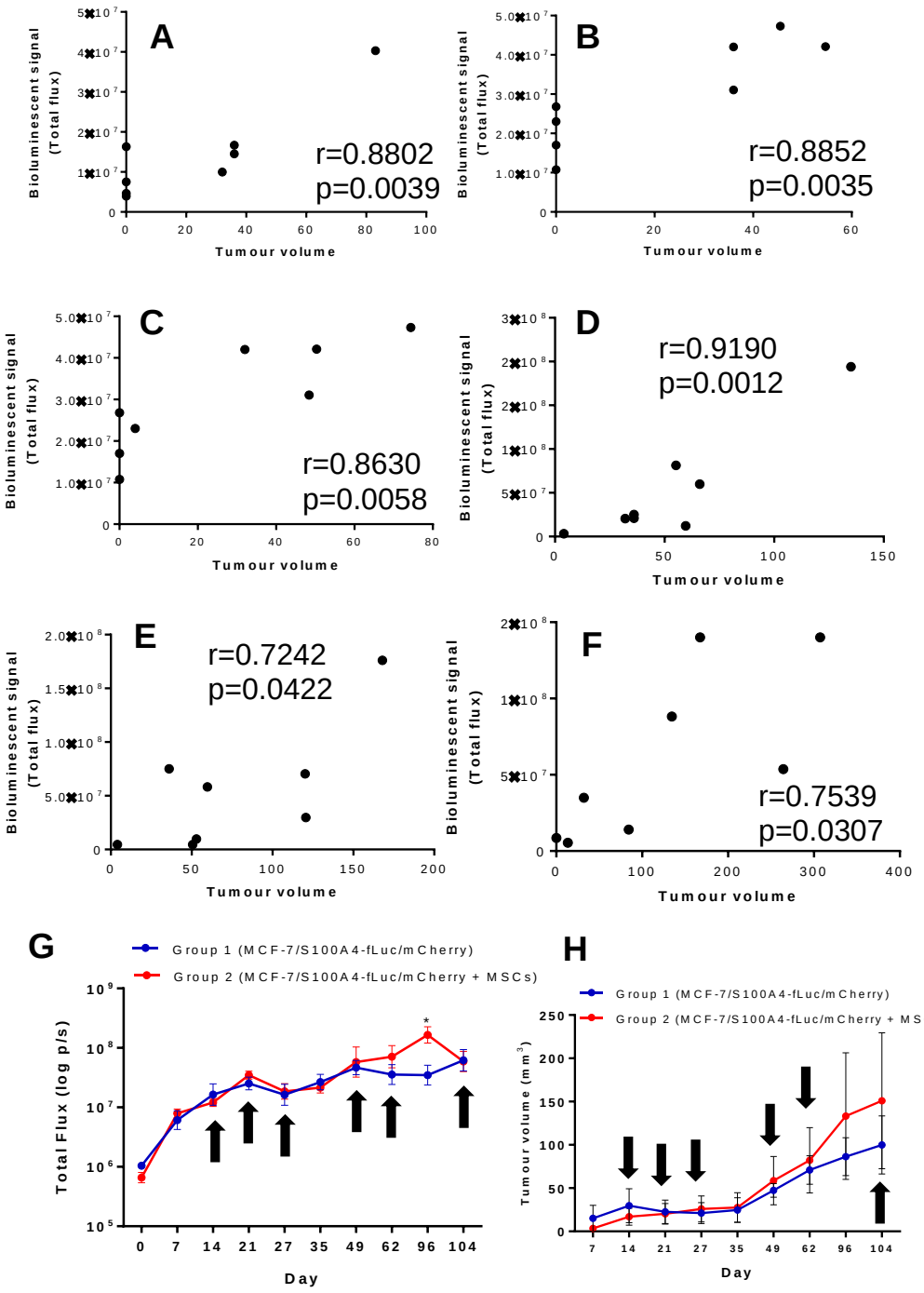


Figure 95. Correlation between bioluminescent signal and tumour volume at different time points. Correlation analysis at day 14 **(A)**, day 21 **(B)**, day 27 **(C)**, day 49 **(D)**, day 62 **(E)** and day 104 **(F)**. Each point represents a different mouse. N=8. r=pearson correlation coefficient. p=P value. The days are pointed by the black arrows in the bioluminescent **(G)** and tumour volume **(H)** graphs.

6.4.3 Correlation of reporter activity and S100A4 expression

As shown by the bioluminescent readings, the bioluminescent signal was significantly higher in the “MCF-7/S100A4-fLuc/mCherry + MSCs” group at day 96 (Figure 90A), which was not proportional to the tumour volume (Figure 94). An increase in the signal was observed in all the mice from this group on that day (Figure 90B). It was hypothesised that the reporter/ S100A4 gene was activated at this point. Although the signal was not increased in the “MCF-7/S100A4-fLuc/mCherry” group on that day, high signals were detected in mouse 1 and 2 throughout the study. To determine whether the S100A4-generated signal correlated with the expression of S100A4, the protein levels of S100A4 in the xenografts were analysed by IHC and correlated with the latest bioluminescent readings. Staining was observed in all xenografts apart from xenografts 5 and 7 (Figure 95). For the isotype controls see appendix 23. While substantial staining was observed in some xenografts (e.g. xenograft 9; Figure 95A) only a few cells were stained in others (e.g. xenograft 15; Figure 95G). On the other hand, expression of S100A4 was not restricted to the group where MSCs were co-implanted, implying that the expression of S100A4 was not due to MSCs (or MSC-induced EMT).

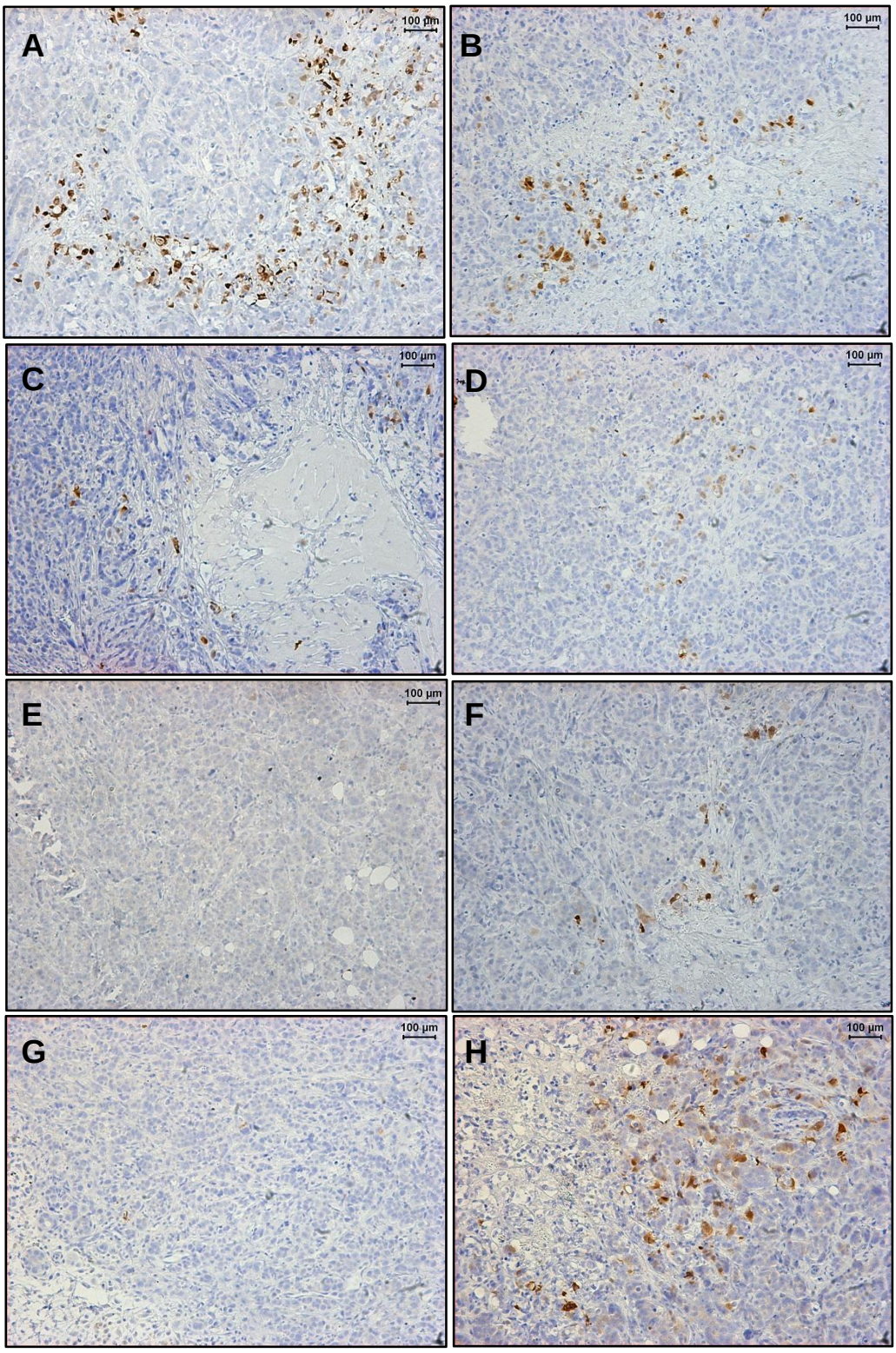


Figure 96. IHC staining of S100A4 in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).

The areas where staining was detected were called “S100A4 hotspots” and they were used for scoring the expression of S100A4. The expression of S100A4 was determined by measuring the number of stained cells (number of features), using semi-automated quantitative computer assisted image analysis. The hotspots were detected using the LEICA DMLB microscope and the image analysis was performed using the Qwin programme (version 3.5.1). The picture was taken, the negative and positive areas (blue and brown respectively) were manually selected (example shown in Figure 96A-C), and the programme generated an output for the number of positive cells and other information (e.g. area of cells). During the selection of positive and negative areas, false positive data (e.g. dark blue cells and dirt) were manually excluded from the analysis and false negative data (brown cells not selected by the programme) were manually included in the analysis. For more accurate results the size (area) of the features was calculated (example shown in Figure 96D). We estimated that anything with a size less than 20μm^2, was not a cell but an artifact, and excluded from the analysis.

The total number of stained cells in the “S100A4 hotspots” was calculated for each xenograft (Table 11) and correlated to the bioluminescent signal detected in the last day of the study (day 104). A strong positive correlation was observed between the bioluminescent signal and the number of cells stained positive for S100A4, as assessed by the pearson’s correlation test (Figure 97). This indicated that the reporter could detect the levels of S100A4 expression at protein levels.

The number of hotspots varied between the xenografts (Table 11) and the proportion of stained to unstained cells within the hotspots was extremely low (varied from 0.0002% to 0.05% as estimated using the ImageScope programme). This would suggest that the reporter can detect small number of cells expressing S100A4..

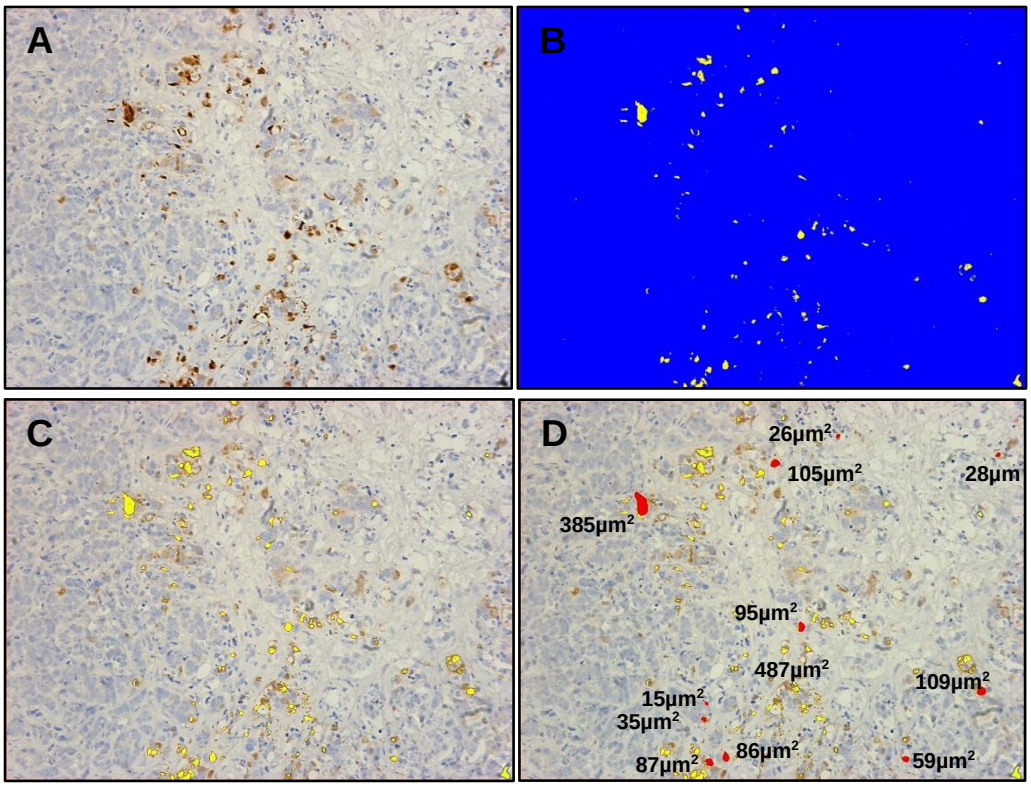


Table 11. Expression of S100A4. Table shows the total number (sum) of stained cells in the "EMT hotspots" for each mouse xenograft, as calculated by a semi-automated quantitative computer assisted image analysis. The number of "EMT hotspots" varied between the xenografts from 1 to 5.

Xenograft	Total number of stained cells	Number of hotspots
1	114	3
2	90	2
3	5	1
4	20	1
5	0	0
6	27	3
7	0	0
8	128	5

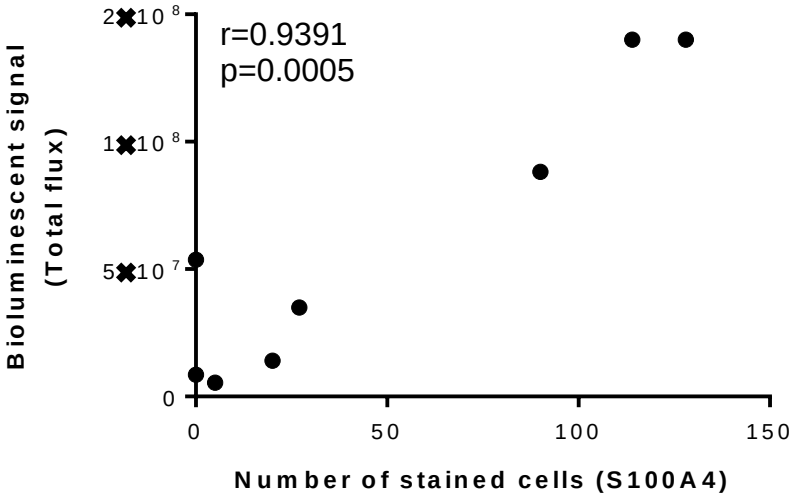


Figure 98. Correlation of bioluminescent signal from the S100A4 reporter and S100A4 expression. Each point represents the signal at day 104 or number of positive cells for a particular mouse/xenograft (Table 11). N=8 (4 per group). r=pearson's correlation coefficient. p=P value.

6.4.6 Assessing EMT in the model

In order to assess whether the reporter could detect EMT-related traits, the xenografts were stained for E-cadherin and Vimentin. For the isotype controls see appendices 24-25. Expression of E-cadherin was detected in all xenografts (Figure 98). The H-scores were calculated as described in chapter 5. No significant difference was observed between the two groups (Figure 99), as assessed by an unpaired t-test. Furthermore, there was no correlation between the S100A4 expression and the E-cadherin expression, as assessed by a pearson's correlation test (Figure 100). Finally, no Vimentin expression was detected (Figure 101). In summary, no difference in the expression of the EMT markers (S100A4, E-cadherin and Vimentin) was observed between the two groups, indicating that MSC-induced EMT was not present at termination.

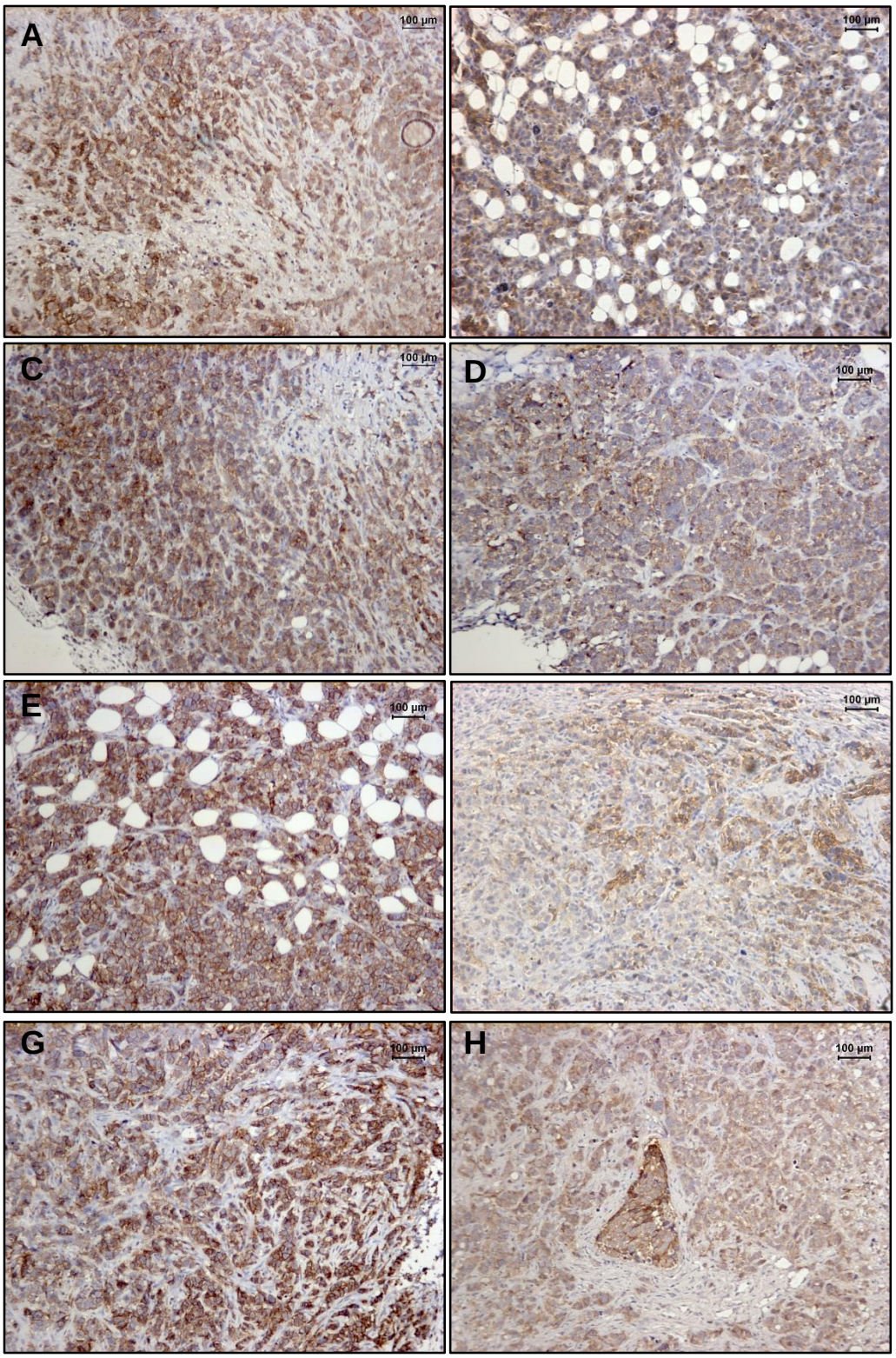


Figure 99. IHC staining of E-cadherin in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).

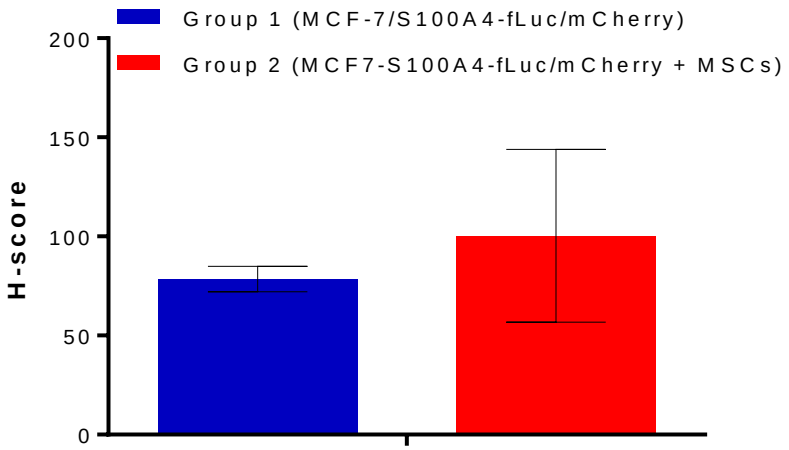


Figure 100. E-cadherin H-score. Each bar represents the average H-score for E-cadherin as detected by IHC and calculated using ImagScore for each group ± SEM. N=4 per group.

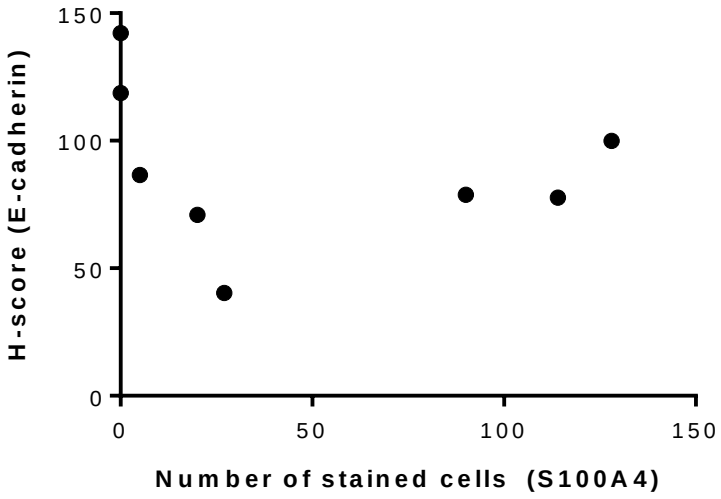


Figure 101. Correlation between E-cadherin and S100A4 expression. Each point represents the H-score for E-cadherin and the corresponding number of stained cells for S1004 (Table 1) in each mouse xenografts. N=8 (4 per group).

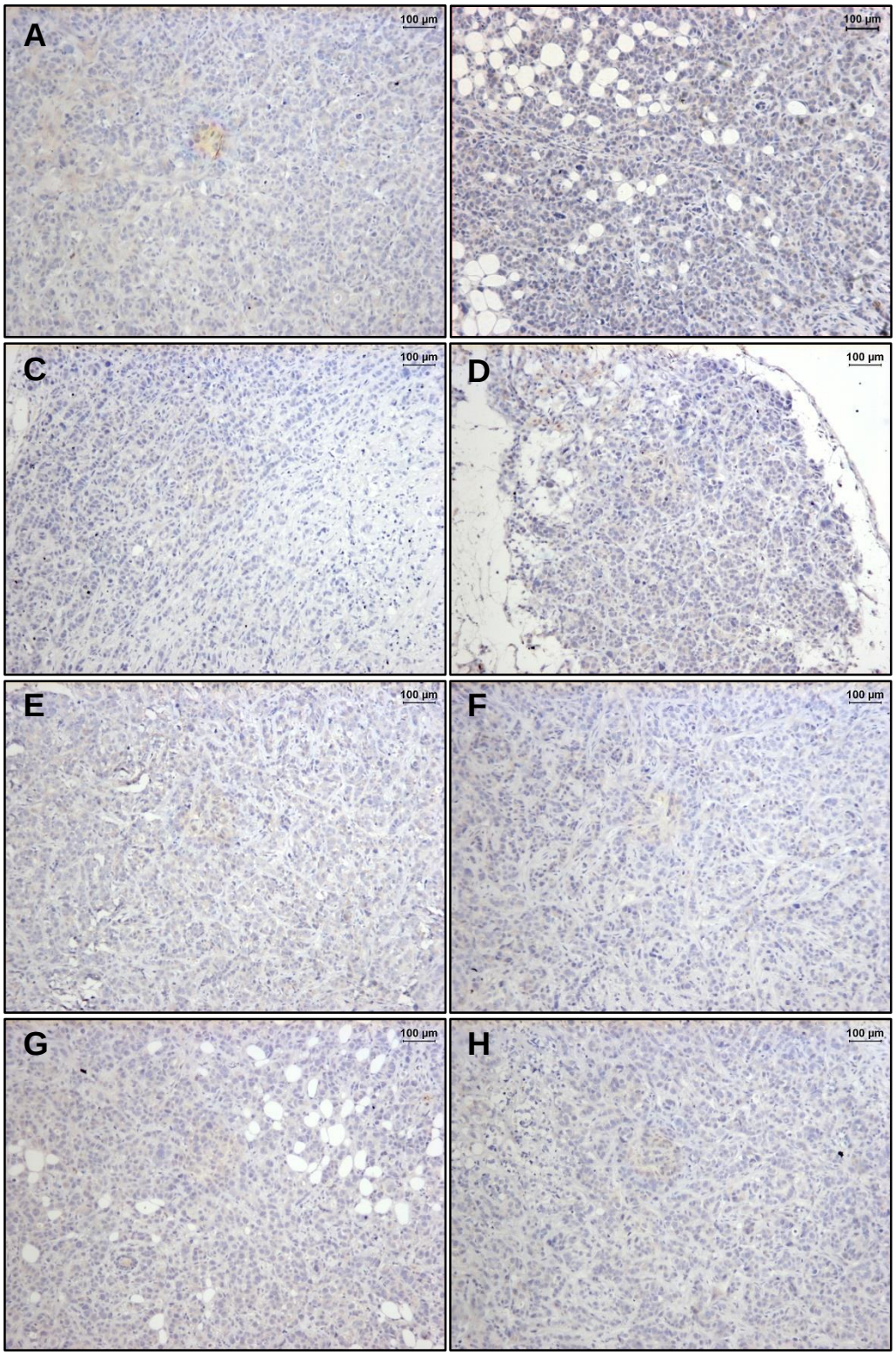


Figure 102. IHC staining of Vimentin in mouse xenografts. A-D Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).

Summary and discussion

In this chapter the activity of the S100A4 reporter was validated in an orthotopic MCF-7 mouse model, and the model itself was assessed.

The model was designed to assess whether a fluorescent reporter could be used to measure tumour growth and to monitor the activity of the reporter by bioluminescent readings. It was shown that the fluorescent readings did not correlate with the tumour volume throughout the study, indicating that the fluorescent readings might not be very reliable for measuring tumour growth. The most possible explanation for this is the decrease sensitivity of the fluorescent measurements due to high autofluorescence. Autofluorescence is an intrinsic property of the cells; fluorescence emission arises from endogenous fluorochromes, such as NAPDH and Flavin [191]. In tissues, the autofluorescence originates mostly from the extracellular matrix. The molecules that contribute to autofluorescence emit fluorescence between 450 and 650nm [192], which covers the emission of mCherry used in this study, 620nm. Another possible explanation is the fact that dead cells tend to be really bright; a phenomenon that has been repeatedly observed *in vitro* (i.e. the apoptotic/rounded/floating cells appear very bright under the microscope). While the fluorescent signal relies on excitation by an external light source, the bioluminescent signal does not require excitation and hence, the autofluorescence is minimised [116]. The bioluminescent signal is therefore a more reliable way to measure tumour growth. To minimise the problem of autofluorescence, it is suggested that a far-red fluorochrome (mPlum, emission 655nm) is used in future studies [193]. Furthermore, a different way of illuminating fluorescence is suggested. In this study the fluorescence was illuminated from the top (epi-illumination), but the IVIS Spectrum can also illuminate from the bottom (trans-illumination). While in epi-illumination the excitation and emission reside on the same side of the tissue within the animal, trans-illumination involves illuminating the tissue from one side and detecting the emitted light that has passed [194]. In the latter, the measurements come from deep inside the tissue, through which each measurement is

an average of the tissue volume sampled by the bulk of tissue through which the light has passed. Furthermore, the sensitivity of epi-illumination is limited in the ability to probe deep tissues due to scattered light and autofluorescence; autofluorescence is mostly absent on trans-illumination images. The trans-illumination technique became available since the study was carried out and therefore could not have been used throughout the study. Another way to refine the model, is to transduce the cells with a dual reporter (or two different reporters) expressing Renilla and Firefly luciferase; suitable as they have different substrate and light output requirements [195].

Even though the reporter measures the gene levels of S100A4, It was shown that the activity of the reporter (as measured by bioluminescent readings) correlated with the protein expression of S100A4. Due to the high homology between the mouse and human *S100A4*, it was not possible to analyse the gene expression levels. It could be implied that the gene was activated when the peak in the signal (at day 96) was observed, followed by translation of the protein. On the other hand, S100A4 expression (protein levels detected by IHC) was detected in xenografts in which this peak was not observed. There was an increase in the signal at the beginning of the study, which can imply that S100A4 was expressed early in the study. Taken together with the data from the *in vitro* studies, it can be concluded with confidence that the S100A4 reporter can detect the baseline gene expression of S100A4, the changes in the gene expression and the changes in protein expression both *in vitro* and *in vivo*. For more accurate and powerful results, the studies should be repeated using more animals, and the *ex vivo* analysis of S100A4 should be performed at early and late stages to capture the moment at which S100A4 was expressed. Furthermore, a bigger study will allow determining a cut off value for the bioluminescent signal for better refining the reporter/model.

MSCs did not promote the tumour growth and did not alter the expression of the EMT markers, E-cadherin and Vimentin, suggesting that EMT was not induced. Furthermore, S100A4 expression was observed in

xenografts from both groups (“MCF-7 alone” and “MCF-7 co-implanted with MSCs”), suggesting that the expression was not induced by hMSCs. So, is S100A4 a good marker of EMT and is the S100A4 reporter an EMT reporter? S100A4 is expressed by tumour cells, immune cells and fibroblasts, and has different functions in cancer progression including cell regulation of cell motility, angiogenesis, invasion and metastasis (reviewed in [119]), and has been used as an early marker of EMT [121]. The expression of S100A4 can be induced by different factors involved in EMT (e.g. TGFβ1) and at the transcriptional level it can be regulated by different transcription factors (e.g. β-catenin) [120]. As previously discussed, S100A4 is also activated under hypoxic conditions [131, 132], and this was also supported by our data that showed upregulation of the *S100A4* in hypoxic cultures. Therefore, it is clearly shown that the expression of S100A4 is regulated by different factors and in order to answer the question, further investigation is needed. For example in order to investigate whether the S100A4 expression related to hypoxia, hypoxia should be assessed in the xenografts and relate to the localisation of S100A4 expression. A common way to do this, is by injecting pimonidazole hydrochloride to the mice prior to termination, which is chemical compound that specifically binds to proteins in hypoxic cells and can be detected by IHC [196]. Hypoxia can also be detected by staining from HIF-1α or glucose transporter 1 (Glu-1) [197]. Expression of S100A4 in the invasive areas of the tumour in our experiments can support the theory that S100A4 detected the EMT-like/early metastatic cells but again, this needs further investigation. If indeed the S100A4 expression is involved in the early stages of EMT, this could explain why only a few cells expressed S100A4 and no changes in E-cadherin and Vimentin were observed. *In vitro* and *in vivo* models of EMT should be established and the expression of S100A4 should be correlated with the expression of EMT markers and functional assays of EMT (e.g. invasion assays) to further refine the model.

In summary, the activity of the reporter was validated in vitro and in vivo, where the S100A4-generated bioluminescent corresponded/ correlated

with changes in the gene and protein levels of S100A4. The reporter could be potentially used for detecting early stages of EMT or hypoxic regions etc., upon further validation/ refinement of the model.

CHAPTER 7: DISCUSSION

Summary of findings

EMT is thought to be involved in the early stages of metastatic BC, where cells in the invasive border of the tumour undergo EMT which allows them to metastasise. It is not clear though when EMT is happening, which cells are involved and what happens to those cells after they metastasis, and to our knowledge this has not been shown in real-time. This project aimed to generate inducible bioluminescent EMT reporters that would allow a real-time read-out of EMT in *in vivo* models, and to improve orthotopic and metastatic BC model by incorporating key elements of the tumour microenvironment. A bioluminescent EMT reporter could provide mechanistic insights to the EMT process and be used to identify potential targets against EMT, as well as be used to test drugs targeted at the process of EMT and metastasis.

In summary, an inducible bioluminescent reporter, that measures the transcription of S100A4, was constructed and validated in MCF-7 cells both *in vitro* and *in vivo*, where it was shown that S100A4-generated bioluminescence associated/ correlated with S100A4 expression. Upon refinement of the *in vivo* model, the reporter/model could potentially be used for imaging the process of EMT and/or metastasis in real-time and for testing novel drugs targeted at these processes. To the author's knowledge no drugs have been developed to target EMT in BC. Therefore, the model can be used for testing and monitoring drug response in real-time.

Cell lines used: their significance and limitations

Three main human BC cell lines were used (MCF-7, MDMAB231 and BT549) during this study. They were characterised according to the expression of EMT markers and CSC population, and, in agreement with the literature, MCF-7 cells were defined as "epithelial" with low CSC population and MDAMB231 and BT549 were defined as "mesenchymal" with high CSC population. These cell lines not only have different EMT phenotypes, but they represent different subtypes of BC; MCF-7 belong to Luminal and MDAMB231 and BT549 belong to Basal B subtypes [81].

This allowed us to explore the EMT potential (or expression of the different EMT and CSC markers) in different *in vitro* and *in vivo* experiments that will be useful for future studies. For example, it was shown that the transcription of the S100A4 reporter was activated in MCF-7 and BT549 cells under hypoxic conditions *in vitro* and in *in vivo* models, but not in MDAMB231. As S100A4 is involved in the early stages of EMT, targeting S100A4 in these two cell lines might have therapeutic potential. Using cell lines that belong to different subtypes would be useful for testing the effects of new drugs targeted at the process of EMT and/or metastasis in different stages/subtypes of BC. There are of course many other available cell lines that could be used that represent other BC subtypes, that have not been explored here, including Basal A subtype [81]. The ultimate goal of the study was to develop models that could be used for patient-relevant drug testing of novel drugs targeted at the process of EMT. But, when using human BC cell lines the question of whether established cell lines represent the heterogeneity of human BCs still remains controversial. The pros and cons of using cell lines have been previously highlighted, and it has been suggested that primary cultures are more likely to provide more useful data since they are isolated directly from patients and represent the heterogeneity and genotype of the human tumours [198]. Indeed, studies have shown that when primary tumours are used as an orthotopic xenograft, there is a stronger predictive response value compared to cell lines [199, 200]. On the other hand, primary cultures are hard to get, are challenging to establish due to the high possibility of contamination, they have other limitations such as finite lifespan, slow population time, and their genetic manipulation is difficult [198]. Therefore, it might prove difficult to transduce primary cells with the S100A4 reporter, induce EMT, and generate a large amount of these cells for testing them *in vivo*.

Inducing EMT *in vitro* and *in vivo*

Different methods were attempted to induce EMT in MCF-7 *in vitro* but none of them were successful. Methods that alter a single factor (e.g.

overexpression of Twist) were tried, and then experiments that involve components of the tumour microenvironment (e.g. 3D cultures) were assessed. The techniques used here were an adaptation of the techniques used literature where induction of EMT in MCF-7 cells have been reported [69, 141, 155, 162, 164, 165, 168], but they were unsuccessful most likely because of differences in the cell density, length of the assay, dimensionality, and the origin, batch and passage number of the cell lines used. The fact that none of the methods worked, even those that included components of the tumour microenvironment, stresses the need for better models of EMT. EMT-related traits were only observed *in vivo* (co-implantation of MCF-7 and MSCs) in a certain setting but not *in vitro* (co-cultures of MCF-7 and MSCs). While MSCs induced tumour growth promotion and EMT-like traits in MF1 nude mice, they did not have an effect on MCF-7 cells implanted in the Rag2 mice or in MCF-7 cells in *in vitro* co-cultures. The role of MSCs in tumour growth and EMT is controversial. For instance, induction of EMT associated with decreased E-cadherin expression and increased expression of Vimentin, N-cadherin, Twist and Snail, has been reported in T47D, SK-Br3 and MDAMB231 cells co-cultured with MSCs [73], but not in MCF-7 cells [201]. To the author's knowledge MSC-induced EMT has not been demonstrated in *in vivo* BC models.

Xenograft models: their significance and limitations

For the *in vivo* studies we chose orthotopic xenograft models, as orthotopic xenografts have a main advantage over other models (e.g. GEM models): they are useful for testing drugs targeting human molecules as they use human cell lines [87, 88]. On the other hand, they use human cell lines that are have been growing in plastic, which may have undergone selection and genetic drift such that they are genotypically and phenotypically very distant to real patient tumours. Furthermore, the xenografts are performed in immunocompromised mice that don't provide a species matched TME, and the genetics and histology of the mouse xenografts do not recapitulate those of the human

tumours. More realistic pre-clinical models are thought to be achieved by grafting fresh cancer tissue (PDX) in immunodeficient mice, as well as by using humanised chimeric animals [202]. Such models preserve the heterogeneity of the human cancers, and more accurately represent the interactions between the cancer cells the TME.

The role of hMSCs and related issues

In our attempt to include a more relevant TME and humanise the xenograft models, we co-implanted human MSCs, but this did not entirely solved the issues of the xenograft models discussed above. Regarding the effect of MSCs on tumour growth, while certain authors reported that MSCs increased the tumorigenicity of BC *cells in vitro* [74, 188], others reported that MSCs inhibited the proliferation of MCF-7 cells [189]. This shows that the effects of MSCs on tumour growth are complex and may be context-dependent, which is supported by the studies performed in this project.

It has been reported that MSCs promoted tumour growth and metastasis through generating an immunosuppressive environment [184], through paracrine signalling [72], through cellular interactions [188] and through generation of cytokine networks that regulate CSCs [74]. In this thesis, MSC-mediated growth and EMT was only observed in nude mouse models but not in Rag2 mouse models. A main difference between the two mouse strains lies in their immune system; the nude mice lack T cells, while the Rag2 mice lack T, B and NK cells and reduced macrophage and dendritic cell function [176-178]. As the Rag2 mice are more immunocompromised and lack a functional immune system, it would be expected that tumour formation (with or without MSCs) would be greater in these mice; which was not the case. In order to better understand why the MSCs did not have an effect on tumour growth in the certain models it is important to elucidate the fate and lifespan of the MSCs; MSCs are pluripotent stem cells that have the ability to proliferate and differentiate to other cells types when recruited to tumours (e.g. endothelial cells and CAFs) (reviewed in [187]). One way to explore this is to use fluorescently

labelled MSCs *in vivo* to image them and detect their location and lifespan, and then isolate the cells for further characterisation (e.g. isolate the cells with Flow Cytometric Cell Sorting and stain them for different cell-specific markers).

Another point worth considering is the appropriate number/ratio of MSCs to use in. In this study $1x10^6$ cells have been implanted *in vivo* with MCF-7 cells (2 MCF-7: 1 MSCs; same ratio *in vitro*), but others have used either more MSCs than MCF-7 (1:3 BC cells: 2 MSCs in co-cultures [73] and 1 BC cells: 20 MSCs *in vivo* [184]) or the same number of MSCs and BC cells (1 BC cells: 1 MSCs *in vivo* [74]). The number of stromal cells recruited to tumours in the clinical scenario is not known; therefore, it is not clear what should be the correct number/ ration of cells to use in studies of this nature. Studies performed here and by others, demonstrate the need for better models to better understand and monitor the role, fate and lifespan of MSCs in BC progression.

Using inducible bioluminescence reporters

The use of optical reporters in this project, allowed imaging of the same animal repeatedly at different times, which avoided the need to sacrifice multiple groups of mice at different end-points; hence, the total number of animals needed was reduced. All in all, this study contributed to 1 of the 3Rs; “Reduction” (http://www.nc3rs.org.uk/).

In order to visualise the process of EMT in real-time, inducible bioluminescent EMT reporters were constructed. While the activity of the S100A4 reporter was validated in MCF-7 and BT549 cells in hypoxia, it was not possible to validate this reporter in MDAMB231 cells or the E-cadherin and N-cadherin reporters. As previously discussed, when designing such reporters and choosing the setting in which the reporters will be validated, it is important to consider the different TFs that bind to the promoters and their gene regulation in general. For validating the activity of the E-cadherin and N-cadherin reporters, alternative models should also be established. Hypoxia was used to validate their activity as it regulates their expression through induction of Twist1 [133, 134]. It has

been shown though that the methyltransferase SET8 is important for Twist-1 induced EMT as it acts as a dual epigenetic modifier on the promoters of the E-cadherin and N-cadherin genes via its H4K20 monomethylation activity [141]; therefore, overexpression of Twist1 and SET8 could be used to induce changes in the genes and for testing these reporters.

A "good" EMT reporter

Upon validation of the S100A4 reporter *in vitro*, the reporter was tested *in vivo*, where the S100A4-generated bioluminescent correlated with the protein expression of S100A4. While the activity of the reporter was validated, the use of a bioluminescent reporter did not allow to localisation of the expression of S100A4 (suggested early EMT) *in vivo*. As stated earlier, a "good" EMT reporter should allow identification of a small population of cells undergoing EMT *in vitro* and *in vivo* and/or metastasising *in vivo*. To overcome this, a fluorescent reporter could be developed, which would allow the use of high resolution live-cell imaging applications, such as intravital and confocal microscopy and implantation of tissue window devices [108]. On the other hand, low levels of EMT/ EMT reporter-generated fluorescent and drug efficacy might be difficult to detect using these systems; bioluminescence is more sensitive that fluorescencet as discussed before.

None of the reporters generated was validated upon induction of EMT either *in vitro* or *in vivo*, which stresses the need of determining whether the reporters are "good" EMT reporters. A "good" EMT reporter should detect EMT-related traits both *in vitro* and *in vivo*, and the subset of cells that are undergoing EMT and metastasise. Upon establishment of EMT models, the activity of the reporters should be tested upon induction of EMT, by comparing/correlating the EMT/reporter-generated bioluminescent with reporter gene expression. As our results showed upregulation of Vimentin in *in vivo* models (and some *in vitro* models) a Vimentin reporter might also be useful for future studies.

Future directions

Future studies will involve repeating the *in vivo* study, in which MCF-7 cells were co-implanted with MSCs in MF1 nude mice, as soon as the mice become available. The study will aim to repeat/ replicate the data; MSC-induced tumour growth and induction of EMT-related traits. Should this model of EMT be established, the S100A4 reporter (existing and new fluorescent version of the reporter) will be assessed. The activity of the reporter will be studied at different time points and be correlated with the expression of EMT and CSC markers, to provide a real-time read out of S100A4/ early EMT.

For validation of the reporters *in vitro* upon induction of EMT, variations of the methods used will be attempted, including direct co-cultures of epithelial cells with BrCAFs, 3D co-cultures with MSCs or CAFs, and co-cultures with both MSCs and BrCAFs. In these models different ratios of epithelial: stromal cells, and different batches/ sources of stromal cells will be explored. Alternative methods will also be attempted for the induction of EMT *in vitro*, such co-cultures with inflammatory cells (e.g. TAMs), and inhibition or downregulation of miRNAs (e.g. inhibition of the miRNA-200 family); the role of inflammatory cells and miRNAs on EMT has been discussed in the introductory chapter. The induction of EMT will be tested by assessing the changes in the morphology of the cells, the changes in the expression of EMT markers, enrichment of CSCs and functional assays (invasion and migration assays). For achieving induction of EMT and/or metastasis *in vivo*, different ratios of epithelial: MSCs will be explored and co-implantation with different numbers of BrCAFs will be attempted. EMT will be assessed *ex vivo* by staining for EMT and CSC markers (Vimentin, E-cadherin, N-cadherin, Twist, CD44, ALDH1 etc.). Furthermore, more BC cell lines will be included in future studies covering different clinical BC subtypes, as well as primary cells as soon as they become available. Similarly, apart from orthotopic xenograft models, the reporters will be tested in PDXs upon availability. As mentioned earlier, a fluorescent version of the reporters will be constructed and be used for visualising EMT *in vitro* (by confocal

microscope) and *in vivo* (using the IVIS system initially to test sensitivity and moving to intravital microscopy in collaborators' lab). Apart from the S100A4, E-cadherin and N-cadherin reporters, a Vimentin reporter will be constructed and tested and possibly others (e.g. Snail reporter). Other than the promoter-based reporters, other types of reporters will be tested that are based on post-transcriptional regulatory mechanisms such as splicing, miRNA-regulated gene regulation or translational regulation of EMT-related genes. For example, a bichromatic fluorescent reporter which has been used to detect epithelial to mesenchymal transitions [203, 204], which is available for use through a collaboration with Dr Sebastian Oltean, can be used to identify EMT in *in vitro* and *in vivo* models. Briefly, the reporter is based on the alternative splicing of the fibroblast growth factor receptor 2 (FGFR2); in epithelial cells alternative cells leads to the expression of FGFR2(IIIb) isoform, while in mesenchymal cells it leads to the synthesis of FGFR2(IIIc) isoform. The dual reporter (EGFP and RFP) contains the exon IIIc and an intronic activator and repressor (ISAR) which activates IIIb and represses exon IIIc, in an epithelial-manner; inclusion of IIIc results in a fusion protein in frame with EGFP, while skipping this exon results in RFP expression. This reporter will be used to visualise and time EMT changes in real-time.

Significance and potential use of S100A4 reporter

The study aimed to use the S100A4 reporter for imaging the early stages of EMT in real-time. The reporter can also be used to provide mechanistic insights regarding the pathways and molecules involved. It is suggested in the literature that EMT occurs in a subset of tumour cells that acquire invasive abilities, enter the circulation and metastasise to distant parts, where they revert back through MET. To our knowledge this has not been shown in real-time. *In vivo* studies can be performed to image changes in the reporter in real-time, and in which mice will be sacrificed at different time points (i.e. when change in reporter activity is observed) for *ex vivo* analysis of EMT, CSC and metastatic markers. This could provide new insights as to what is involved in the different stages of the "EMT-

metastasis-MET" process. Building up on this, the reporter could be used to identify new prognostic and predictive markers to stratify patients.

Upon further validation and refinement, the S100A4 reporter can be used to identify drugs targeted at the EMT process. Once it is proven that the reporter is a "good" EMT reporter and *in vitro* and *in vivo* models of EMT are established, the S100A4 reporter can be used in high-throughput screening (HTS) drug discovery to identify compounds that target EMT, by for example using a small compound library as previously reported [83]. It can also be used for testing drugs targeted at EMT. For example, the reporter gene and other key genes involved can be knocked down (e.g. by shRNA) and effect on EMT and metastasis can be evaluated in *in vivo* models in real-time. Clinically relevant drugs could then be developed to target early EMT and therefore metastasis.

Conclusive remarks

In summary, in this study a potential EMT reporter was constructed and validated in an epithelial BC cell line. The different *in vitro* and *in vivo* experiments performed stressed the need for improved models of EMT, and will guide future studies for improving pre-clinical BC models. Through the use of inducible reporters such as this EMT reporter, more detailed information on the activity of specific biological pathways during tumour development drug treatment can be gained, allowing selection of the most relevant models to be selected for specific studies.

APPENDICES

Appendix 1: Study protocol for “Bioluminescent MCF-7 tumour EMT / metastasis pilot in an orthotopic mammary fatpad (MFP) model”

1. STUDY OBJECTIVE

Comparison of ability to generate bioluminescent MCF-7 cell primary and metastatic growth via different initiation routes with hormone supplementation;

2. REGULATORY GUIDELINES

This study will be conducted under the UK Home Office Licence number PPL 40/3559, 19b (3). NCRI guidelines for the welfare and use of animals in cancer research, LASA good practice guidelines and FELASA working group on pain and distress guidelines will also be followed.

3. TEST SUBSTANCE AND VEHICLE DETAILS

N/A

4. FORMULATION

As above

5. REASON FOR THE CHOICE OF SPECIES, ROUTE OF ADMINISTRATION AND DOSE LEVELS

Immunodeficient mice have been used extensively to generate human tumour xenografts and remain the experimental method of choice for testing anti-tumour efficacy of new compounds prior to administration in man and hence are appropriate for the development and refinement of orthotopic models. Female MF-1 nude mice are appropriate due to the use of breast cancer cells and bioluminescent reporters.

6. ANIMALS (species, strain, sex, number, source)

8 MF-1 female nude mice at 4 to 6 weeks of age will be obtained from Harlan (or PRECOS via Harlan).

6.1 Location of the Study

The study will be located in the CSU Level 2 containment facility, BSU, University of Nottingham.

6.2 Housing and Environment

Mice will be maintained in IVCs (Tecniplast UK) within a barriered unit illuminated by fluorescent lights set to give a 12 hour light-dark cycle (on 07.00, off 19.00), as recommended in the United Kingdom Home Office Animals (Scientific Procedures) Act 1986. The room will be air-conditioned by a system designed to maintain an air temperature range of 21 ± 2ºC and a humidity of 55% ± 10%.

Mice will be housed in social groups during the procedure with irradiated bedding and provided with autoclaved nesting materials and environmental enrichment.

6.3 Diet and Water

Sterile irradiated 2919 rodent diet (Harlan Teklad UK, product code Q219DJ1R2) and autoclaved water will be offered *ad libitum.*

6.4 Animal welfare

An experienced technician will check the condition of the mice daily. Adverse effects will be noted and reported to the Named Animal Care and Welfare Officer (NACWO) and Named Veterinary Surgeon (NVS). Animals may be terminated at any time during the study if any unexpected adverse effects are noted according to Home Office Project Licence PPL 40/3559. See also ***7.4.6***.

All efforts shall be made to image mice prior to termination and on Post Mortem to assess excised liver tumour burden by BLI.

7. EXPERIMENTAL DESIGN

7.1 Animal Identification

Each animal will be earmarked and assigned a cage number which will appear on the data sheets. Ear markings will be as follows; pl (plain or no marking), 1r (1 notch in right ear), 1l (1 notch in left ear), 1r1l, 2r etc

7.2 Group Sizes, Doses and Identification Numbers

4 mice per group will be initiated. Once allotted to their treatment/initiation groups the mice will be given ID numbers from 1 to 8 that relate to the cage and earmarking in order to simplify optical imaging and data analysis.

B Pilot study

There will be 4 groups with 4 mice per group initiated with **MCF-7 Fluc cells** as supplied by **NN** as follows;

Group 1 Standard mammary fatpad (MFP) initiation plus dietary estradiol

Group 2 Subcutaneous initiation with dietary estradiol

7.3 Body Weights

Animals will be weighed weekly from initiation for the duration of the study.

7.4 Experimental Procedure

7.4.1 Cell maintenance

The MCF-7 Fluc cells are maintained *in vitro* in RPMI culture medium (Sigma, UK?) containing 10% (v/v) heat inactivated foetal bovine serum (Sigma, Poole, UK) & 2 mM L-glutamine (Sigma, UK) at 37^0C in 5% CO_2 and humidified conditions. Cells from sub-confluent monolayers are harvested with 0.025% EDTA, washed in culture medium and counted. Cells with viability of >90% will be re-suspended and will be seeded at $2x10^6$ cells per T150 flask and incubated for 48 hours. On day of initiation cells are harvested from semi-confluent monolayers with 0.025% EDTA, washed twice in the culture medium and counted 3 times as above. Cells with viability of >90% will be re-suspended, for *in vivo* administration, in sterile standard formulation matrigel at **1 to $2x10^6$ cells/100ul (1 to 2 $x10^7$ cells per ml).** Cells will be aliquoted into sterile tubes and transported on ice to the F Floor facility for initiation within 30mins of arrival

7.4.2 Tumour Initiation

Mice will be initiated with 100ul of 1 to 2 $x10^7$/ml MCF-7 Fluc cells into the appropriate site.

7.4.3 Monitoring/tumour measurement

Tumour establishment and growth will be monitored during the experiment by caliper measurements twice weekly and 2D and 3D optical imaging, carried out under anaesthesia in an IVIS Spectrum, once weekly and at termination to provide an optical post mortem. 3D image

reconstruction if required and tumour measurements will be made using Living Image (4.3.1 build) software.

Mice will be evaluated daily by an experienced technician until termination.

7.4.3b Treatment

Animals will be checked daily. The study will continue until tumours reach maximum permitted volume/mass as specified in Home Office Project Licence PPL 40/3559 and following NCRI guidelines.

7.4.4 Termination

Each mouse will continue on the study as described in **7.2**, or until the tumour size (as measured bioluminescently and specified in the Home Office licence) or other clinical signs necessitate removal of that mouse from the study. Animals may be terminated at any time during the study if the tumour size becomes excessive (tumour burden should not exceed 5% of the animal's normal bodyweight or estimated body mass) or any adverse effects, as described below, are noted according to Home Office Project Licence PPL 40/3559. Mice will be weighed weekly. Prior to termination the mice will be given D-Luciferin substrate to enable the excised liver and spleen to be assessed for tumour burden via bioluminescent imaging. The mice will be terminated by an approved S1 method.

Terminal imaging will be carried out on the cadaver with the skin and muscle layers open to expose the primary site (sc or MFP) and adjacent lymphatics followed by the body cavity and then all dissected organs in dorsal and ventral orientations to assess tumour burden by bioluminescence.

Lymph nodes with identifiable light emission will be preserved in cold media for dissagregation to extract cells that may have a metastatic phenotype. These cells will be re-passaged to develop a reliable metastatic model. If there is enough material some will be fixed for IHC analysis

Other metastatic sites might be of interest – a decision may be made on the day about their importance otherwise see below.

After imaging other organs of interest will be weighed and tumours or organ thought to contain tumour cells will be preserved half formalin fixed and half frozen. If tumour is too small then fix only.

7.4.5 Adverse effects and humane end points

Dosing of substances	Adverse effects specific to the substance under test	% Unknown	Where supporting data cannot be provided in a similar model (mouse strain and tumour model) or dose level/frequency; short (14 days) pilot tolerability studies will be carried out (n=2 per level) to assess novel preparations. Duration and dose of the substance will be selected to minimise frequency and adverse effects without affecting scientific outcome. Animals exhibiting adverse effects (e.g. loss of bodyweight 20%, subdued behaviour patterns etc) will be referred to the named persons. Any which are fail to improve within a reasonable time will be killed by a schedule 1 method advice
Subcutaneous initiation of tumour	Although care is taken to position the tumour graft so that subsequent growth will take place on the flank to minimise disruption to movement, in <1% animals the	<0.1%	Where the experimental results will not be compromised the animal will be killed by a schedule 1 method. Where the loss of the animal would compromise the data with the resulting waste of the other animals in the study, the Named Veterinary surgeon, and if the adverse effects exceed or are likely to exceed moderate the Home Office Inspector, will be consulted.

	tumour will develop in such a way that this occurs.		
	Tumours may grow rapidly and suddenly exceed permitted size.	<1% except during pilot studies when may approach 10%	NCRI guidelines with regard to maximum permissible size and weight will be observed, if the tumour size suddenly (rather than gradually) approaches this, the animal will be humanely killed. Pilot studies will always be performed with new tumour lines to determine growth rates.
	Ulceration	<5%	Injected cells grown subcutaneously have a tendency to ulcerate whilst the tumour is quite small. The animals will therefore be observed regularly as soon as growth is established, and the mouse will be Schedule 1 sacrificed and tumour passaged as soon as the skin covering the tumour appears to be stretched.
	Primary xenografts, particularly those from luminal sites e.g. oesophageal, gastric and colorectal, may contain residual bacterial contamination.	<5%	The Primary tissue is washed and incubated with antibiotic/antimycotic media overnight prior to implantation. Animals implanted with these tissue-types are treated prophylactically with an antibiotic under advice of the NVS. Any animals exhibiting symptoms of local (reddening of the implant site, infected wound) or systemic (subdued behaviour patterns, piloerection, intermittent abnormal breathing patterns, weight loss of

			20%) infection will be immediately terminated by a Schedule 1 method.
	Slow growing xenografts or those exhibiting a pronounced metastatic phenotype may result in metastatic tumour spread to a distant organ.	<5%	Symptoms may be specific to the metastasised organ e.g. lung (intermittent abnormal breathing patterns); in addition to general symptoms: subdued behaviour patterns, piloerection, weight loss of 20%. Abdomen will be palpated to detect the presence of tumours as per NCRI Guidelines. Each specific experiment protocol will include the specific organs likely to be affected and associated clinical signs and specific humane endpoints Animals exhibiting such symptoms will terminated by a schedule 1 method.
Injection into the mammary fat pad	Due to the location of the tumour (lower left mammary pad) subsequent growth may cause disruption to movement.	up to 100%	NCRI guidelines with regard to maximum permissible size will be observed. If the tumour size and position affects normal animal behaviour, the animal will be humanely killed. Where the loss of the animal would compromise the data with the resulting waste of the other animals in the study, the Named Persons will be consulted.
	Injected cells grown subcutaneously (the fat pad is in effect a subcutaneous site) may ulcerate	<5%	Animals will be terminated by a Schedule 1 method.

	independent of tumour size. The animals will therefore be observed regularly as soon as growth is established.		
	Pronounced metastatic phenotype may result in metastatic tumour spread to a distant organ.	% Unknown To be determined by initial pilot studies	Symptoms may be specific to the metastasised organ e.g. lung (intermittent abnormal breathing patterns); in addition to general symptoms: subdued behaviour patterns, piloerection weight loss of 20%. Animals exhibiting such symptoms will terminated by a schedule 1 method.
General and Terminal Anaesthesia	Animals undergoing repeated anaesthesia events over a single day are at increased risk of dehydration and a drop in body temperature	<1%	Veterinary advice will be sought in order to ensure that animals are able to be re-hydrated between anaesthetic events. Body temperature will be maintained by the use of appropriate warming, with monitoring to confirm that this is achieved. Repeated anaesthetic events will require the Named Persons or their nominees to be notified in advance in order to ensure animals' welfare is not unduly compromised, and will only be carried out when animals have returned to normal eating, drinking and mobility. No more than two sessions will be carried out in any one day, such daily sessions

			being no less than two days apart for two weeks or weekly for five weeks.

8. DATA & STATISTICAL ANALYSIS

Body weight and tumour measurement data will be recorded and reported in spreadsheet and graphical format. Images and videos will be provided as required on the appropriate medium. Statistical analysis will be performed if appropriate using either SPSS/Minitab/Prism programmes for the PC. Data will be provided to the sponsor as soon as possible, the final report will follow when completed.

9. FINAL REPORT

A final report will be completed within 4-weeks of study completion, all data generated will be included as part of the therapy study report.

10. STORAGE OF DATA

All documents relating to the study, including raw data, will be stored at PCO for 3 years after which time it will become the responsibility of the Sponsor.

11. Consumables excluding tech time and overheads

20.00 to 45.00 per mouse depending on source

175.00 Shipping cages, transportation and delivery

€30.00/kg Mouse Chow – estradiol supplemented

12. CHANGES TO PROTOCOL (to be signed off by PPL Holder)

Appendix 2: Sequence of pLVX-S100A4promoter-fLuc

Bold = Restriction enzymes (ClaI, XhoI) / Cloning sites

Underlined = S100A4 promoter

Box = Change from the original sequence (T to A)

Double underlined = Firefly Luciferase

…TTACAGGGACAGCAGAGATCCAGTTTATCGAT**AT^CGAT**AACTC
CCCTATTTCATGCCATTTCACCTCTAACTCTCCACCCCAACCTGGA
TTCTTCATTCCTGACACTCATCCCAACTTTAAATGGCCCCTCCTGAT
ACCCTCTCCGAACCTGAGATCTATCCGTGAGCCCCCACGCCTCAC
TGCCACTCCACTCCATCACTACCTCACCCAGGACCTTTCCCACTGA
CGTTCCTGAGGTGGTCCCAGAGCCTCCTTTGGGTGTGAGCCTGTT
CCCCTCCAGATCCCCCCGCCCCGACCCTGAGCCTTACTTGGCATG
GCAGACAGTACCGGGCATGGGGATCCCCACCCCAGTTTTTGTTTC
TGAATCTTTATTTTTTTAAGAGACAAGGTCCTCTGTGTTGCTCAGGC
TGGAGAGCAGTGGCTTGAGCATAGCCAACTGCAGTCTCGAACTCC
TGGGCTCAAATGATCCTCCTGTCTCAGCTTCCTGACTAGCTGGGA
CTACAGGCTACAGCCATGCTGCCCAGCTAATTAAAAAAAAAAATTG
TTTTTCCTTTTTATAGAGACAGAAGTCTCTCTATGTTGCCTAGGCTG
GTCTTGAACTCCTGGCCTCAGGCGATCCTCCCATCTCCCCCCTAG
CTTTTGTGTCACCACATTTCCAGGGCAATCTCCCACCTGTCACCCA
CCACCCCCTGCATCTCCTTTCCTAGGTCCCCATGGGACTACTCCCT
GTCCCCCATGCTCCAGGCACAGGCTGCCCCTTCCTCCACCTCTCT
AAAACTCAGGCTGAGCTATGTACACTGGGTGGTGCCCATCTCATC
CAGTCCCCTGCTAGTAACCGCTAGGGCTTACCCGTTACCCACGGG
TGCCCACCTGGGAACAGGAGGCTTGGTTCCACGGCTGGGCTGGT
GGAGGGTGCTGTGGCACTTACCGCATCAGCCCACAGCAGGAAGG
CAGTATCCGCTCTCCCCTGTCCCCTGCTATGGGCAGGGCCTGGCT
GGGGTATAAATAGGTCAGACCTCTGGGCCGTCCCCATTCTTCCCC
TCTCTACAACCCTCTCTCCTCAGCGCTTCTTCTTTCTTGGATTGGT
GAGTTGTGTTG**C^TCGAG**CTCAAGCTTCGAATTCTGCAGTCGACG
GTACCGCGGGCCCGGGATCTGGCCTCGGCGGCCAAGCTTGGCAA

TCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAAACATTA
AGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGA…

Appendix 3: Sequence of pLVX-Ncadherin promoter-fLuc

Underlined = N-cadherin promoter
Box = Change from original sequence (T to C)
Bold = Restriction enzymes (ClaI, XhoI) / Cloning sites
Double underlined = Firefly Luciferase

…ATAGAAGAAGAAGGTGGAGAGAGAGACAGAGACAGATCCATTC
GATTAGTGAACGGATCTCGACGGTATCGCCTTTAAAAGAAAAGGG
GGGATTGGGGGGTACAGTGCAGGGGAAAGAATAGTAGACATAATA
GCAACAGACATACAAACTAAAGAACTACAAAAACAAATTACAAAAAT
TCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCAGTTT**AT**
^CGATGGCTCTAGGGGCTGGATTGCGGGTCCCCTCTGCTCCTTCT
TCGGGGCCGAGCAAGCTCGCGGGGGGTGTCCCCAGGCAGCAGG
AGGGTGAGGCAGGATAACTCCCAGGTTTTGCAGTTTTCTGTGCCT
TTCGCCCCTCGCATAGTACTTTGCTCCAATTTTTTTGTTTATTTCTG
CAGGGATGGGTCGCGAGTTTACAGTAGGAATTCTCACTCTGGTGT
CTTAGTTACGGGACTTGCCTCCACCTTACCCCACCCTGCACCCGT
GCGCCTCCCGGTGCTCAAGGGCTCTACTGTCCTATTGTCTTTATCC
CAGTCTGGAGTTGCCTCCTTTTCTTCTTCTCTTGCCCTCACAATATT
CTGTGTTCGTCCCACCTGGAAAGAAAATCTGGCCCTCTTGCCCGG
AGGAGTCTCTGCCGGTCAGCGCCTGCCTGTAGAACGAGTCTGGC
CTCTCACGCCTAGTATCTCCCCCGGGCGAGAAGTTCCGTTACGCG
GCTGCTCCTTTCTCTGTCCGGCTGCCTGTGCCCTATATGCGCCCC
GGGGGCGCCTTTGCCTGGAGCCTTCCGTACTTTTCATACCCTCGG
GAGAGAGCCAGCGAGCCTTTTCACTTAATCTCTACTAAGAGGCTC
CTCCGCGGCAGGAAGCACCTGCTCCAGTCCGGGCGCAGGTTCCC
TCGCCACTGCACGCTGCGGGTTCCCAGCCCGGGGGCAGCCCAAC
TTTCCGCCGCCTCCAACCCGAGGATGCCCCCAGGCTCCTCCCGA

CCAGCCAGCCCAGGCCTGGAGGCCACTTCACACAGCCTCGGCCG
GGACCCTCTAGCTGGGCGGGTGCAGTGCCCTGGTCCCAAGTCGG
TCTCCGCCTGCCGTGGGGTCCCTCGAAGGTTGCGGGAAAGGGGC
GTGCCCACCCTCGCCGCGGGGCAGGAAAATCCACGGAGGGTCCC
CCTGCCTGTCGCCCTCGGGGCAAGGGGAGAGGTGCTCTACTGCC
AGCTGGCAGCCCCGAAAAGTGCGGACGCAGCTCGGAAGTGCCTG
CTCCGGGGGTGGAGACAGGTGGTGTGGATTACCCCGCTCCGTGG
TCTTCACGGGTCAGGGCCTGGCGCGCTCCGGCCCGCCGGGGGG
TGAGTGCTGCTGGGGCGGGGGCAGGGGCTGGACGGGGCTGGCC
TCCGCAGTCTCTCGCGCTTTCTCCCGCAACCCCTAGACGGTCGTG
GGCCGCCGCCGCGCGGAGCCCTGTCCATGGTGCGCCATGGTGC
GGCACGCCTGTCCCTGCCCGGCTCCACGCTCTCCTCCCGGGTCG
TGCTCTCTCACTTTGTCTCTCCCCACCCCTGCCTCCCAGGCCAGG
GTCCCAGCGTTCCTTTTCCCTTTCTTGGATTTCTTCTTTTGTTCTTT
AGTCTCCTACCGGTCTTTCTCCCTTCCCCCACTCGCCATCCGCCAC
TCGGCTCTCTCCTGACTCGGGGATTTCAGAACTATTAGACTAGACA
TTTTTGGCCTTTGCAGTTTTGGAGAGTGGATGGAATTCCTTCTAAC
TTGGGGGAGGAGACCTGCCAGTCACTTGCTAACAAAAGTACCTCC
TGTTATAAAACCGCCCTTTATTCTCCCAGCGCACACTCCTCATTCTT
TGACCTCCTGTCACTTCATTTATTGTTTCTTCTTTAAAAATGGGGTT
TATTAAAGAGGTTAGGCTGAAGGAGGAAGAAGGAAAAAAGTCCTG
CTTCTAAAGTATTTGAAGTTAAATGGAAGGCAAGACAAACTTAAAG
ACTCCTGTGTATTAGGATTTTTTTTTAACAAAGGAGATGTGATACAT
TTTTTACAAATCTTACTTTAAAAACACTAATTTTTGGACAGCATTGG
GTCGTCTTATTTACAGACAGTAAATCCTCAGTGTGAATCCAGCAGC
AGATGGTTATTTTCACAAACTTTTTCATCTAAAAGTGAGTTTTAATT
GAATACTGTGCAGCGAAGTTAGGACACGAAGAGACATTTGTTGGT
CATGAATACTGTGGGAAATCTATGTGTGTTACAAGTGGAGCACTCT
CATTTTCTGTTTTAGTCATCTGCCTTGAAGTCAGGTGTGAATTTTCT
GTCAGTTTTTTCTTAAGTGTTTCTAAAAATATCGTTTATAAGCATTAA
ATAATGTGCTTTTCCATTGTATGTTCTTTCTCAGTTTCTTCTATACCA
ATGTTGACTTTTTTAAACCTTCAGGAAAGGAGAAATTATAACTTTCT
CATCAAGTTATTTGGCAGTCCTCTAATTCATGAGTACAAGTCAAGC

AGTAACATTGAATGATTTGAAAGGTCATTAGGAGTGGAAGCAGAGC
AGTTTACGCAACATGAGTTTTAGGGTTAAGTGCACCATGTGGATTG
TACAACTGAGATTAGAAATGACTTTGTGCTAATGGTCTGCTTTACA
CGCCCGAACAACA**C^TCGAG**CTCAAGCTTCGAATTCTGCAGTCGA
CGGTACCGCGGGCCCGGGATCTGGCCTCGGCGGCCAAGCTTGG
CAATCCGGTACTGTTGGTAAAGCCACCATGGAAGATGCCAAAAAC
ATTAAGAAGGGCCCAGCGCCATTCTACCCACTCGAAGACGGGACC
GCCGGCGAGCAGCTGCACAAAGCCATGAAGCGCTACGCCCTGGT
GCCCGGCACCATCGCCTTTACCGACGCACATATCGAGGTGGACAT
TACCTACGCCGAGTACTTCG…

Appendix 4: Sequence of pLVX-E-cadherin promoter-fLuc

Underlined = E-cadherin promoter
Bold = Restriction enzymes (ClaI, XhoI) / Cloning sites
Box = Change from original sequence (C to T)
Double underlined = Firefly Luciferase

…AAATTCAAAATTTTCGGGTTTATTACAGGGACAGCAGAGATCCAG
TTT**AT^CGAT**AAATTAGGCTGCTAGCTCAGTGGCTCATGGCTCACA
CCTGAAATCCTAGCACTTTGGGAGGCCAAGGCAGGAGGATCGCTT
CAGCCCAGGAGTTCGAGACCAGGCTGGGCAATACAGGGAGACAC
AGCGCCCCCACTGCCCCTGTCCGCCCCGACTTGTCTCTCTACAAA
AAGGCAAAAGAAAAAAAAAATTAGCCTGGCGTGGTGGTGTGCACC
TGTACTCCCAGCTACTAGAGAGGCTGGGGCCAGAGGACCGCTTG
AGCCCAGGAGTTCGAGGCTGCAGTGAGCTGTGATCGCACCACTG
CACTCCAGCTTGGGTGAAAGAGTGAGCCCCATCTCCAAAACGAAC
AAACAAAAAATCCCAAAAAACAAAAGAACTCAGCCAAGTGTAAAAG
CCCTTTCTGATCCCAGGTCTTAGTGAGCCACCGGCGGGGCTGGG
ATTCGAACCCAGTGGAATCAGAACCGTGCAGGTCCCATAACCCAC
CTAGACCCTAGCAACTCCAGGCTAGAGGGTCACCGCGTCTATGCG
AGGCCGGGTGGGCGGGCCGTCAGCTCCGCCCTGGGGAGGGGTC

CGCGCTGCTGATTGGCTGTGGCCGGCAGGTGAACCCTCAGCCAA
TCAGCGGTACGGGGGGCGGTGCCTCCGGGGCTCACCTGGCTGCA
GCCACGCACCCCCTCTCAGTGGCGTCGGAACTGCAAAGCACCTGT
GAGCTTGCGGAAGTCAGTTCAGACTCCAGCCCGCTCCAGCCCGG
CCCGACCCGACCGCACCCGGCGCCTGCCCTCGCTCGGCGTCCCC
GGCCAGCCATGGGCCCTTGGAGCCGCAGCCTCTCGGCGCTGCTG
CTGCTGCTGCAGGTACCTCGGATCCCCTGACTTGCGAGGGACGC
ATT**C^TCGAG**CTCAAGCTTCGAATTCTGCAGTCGACGGTACCGCG
GGCCCGGGATCTGGCCTCGGCGGCCAAGCTTGGCAATCCGGTAC
TGTTGGTAAAGCCACCATGGAAGATGCCAAAAACATTAAGAAGGG
CCCAGCGCCATTC…

Appendix 5: DNA sequence of Twist1 expressing vector

Double underlined = CMV promoter

Green = *TWIST1* (Underlined = Exon 1, Bold = Part of Exon 2)

Bold = Restriction enzymes (*EcoRI* & *BamHI*) / Cloning sites

Italic = IRES

Italic underlined = tdTomato

…GCTTGGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCC
TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC
GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA
TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTA
AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACT
TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG
GTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTT
GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG
AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTA

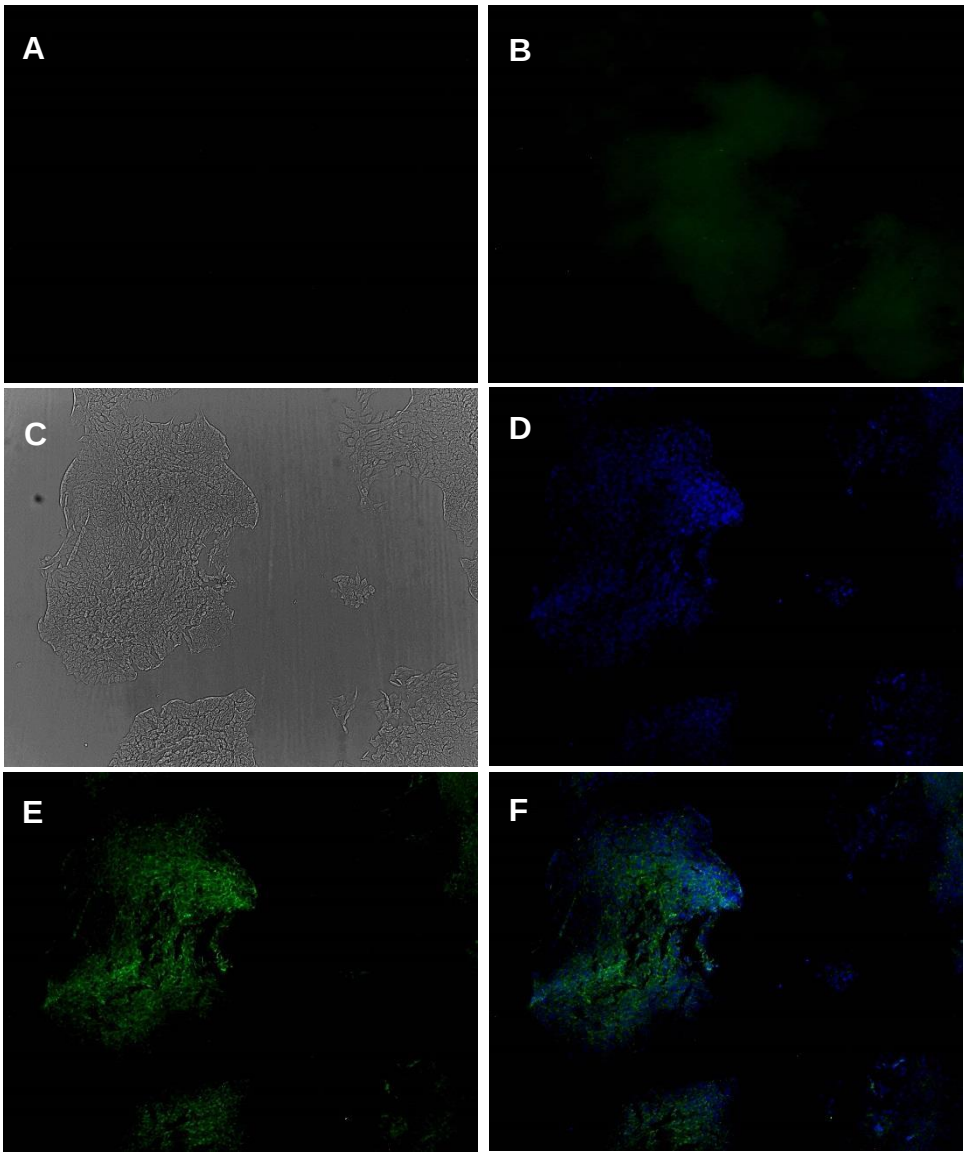
ACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGT
GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCG
CCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACC
GACTCTACTAGAGGATCTATTTCCGGT**G^AATTC**GAGGTATAAGAG
CCTCCAAGTCTGCAGCTCTCGCCCAACTCCCAGACACCTCGCGGG
CTCTGCAGCACCGGCACCGTTTCCAGGAGGCCTGGCGGGGTGTG
CGTCCAGCCGTTGGGCGCTTTCTTTTTGGACCTCGGGGCCATCCA
CACCGTCCCCTCCCCCTCCCGCCTCCCTCCCCGCCTCCCCCGCG
CGCCCTCCCCGCGGAGGTCCCTCCCGTCCGTCCTCCTGCTCTCTC
CTCCGCGGGCCGCATCGCCCGGGCCGGCGCCGCGCGCGGGGGA
AGCTGGCGGGCTGAGGCGCCCCGCTCTTCTCCTCTGCCCCGGGC
CCGCGAGGCCACGCGTCGCCGCTCGAGAGATGATGCAGGACGTG
TCCAGCTCGCCAGTCTCGCCGGCCGACGACAGCCTGAGCAACAG
CGAGGAAGAGCCAGACCGGCAGCAGCCGCCGAGCGGCAAGCGC
GGGGGACGCAAGCGGCGCAGCAGCAGGCGCAGCGCGGGCGGC
GGCGCGGGGCCCGGCGGAGCCGCGGGTGGGGGCGTCGGAGGC
GGCGACGAGCCGGGCAGCCCGGCCCAGGGCAAGCGCGGCAAGA
AGTCTGCGGGCTGTGGCGGCGGCGGCGGCGCGGGCGGCGGCG
GCGGCAGCAGCAGCGGCGGCGGGAGTCCGCAGTCTTACGAGGA
GCTGCAGACGCAGCGGGTCATGGCCAACGTGCGGGAGCGCCAG
CGCACCCAGTCGCTGAACGAGGCGTTCGCCGCGCTGCGGAAGAT
CATCCCCACGCTGCCCTCGGACAAGCTGAGCAAGATTCAGACCCT
CAAGCTGGCGGCCAGGTACATCGACTTCCTCTACCAGGTCCTCCA
GAGCGACGAGCTGGACTCCAAGATGGCAAGCTGCAGCTATGTGG
CTCACGAGCGGCTCAGCTACGCCTTCTCGGTCTGGAGGATGGAG
GGGGCCTGGTCCATGTCCGCGTCCCACTAGCAGGCGGAGCCCCC
CACCCCCTCAGCAGGGCCGGAGACCTAG**ATGTCATTGTTTCCAGA**
GAAGGAGAAAATGGACAGTCTAGAGACTCTGGAGCTGGATAACT
AAAAATAAAAATATATGCCAAAGATTTTCTTGGAAATTAGAAGAG
CAAAATCCAAATTCAAAGAAACAGGGCGTGGGGCGCACTTTTAA
AAGAGAAAGCGAGACAGGCCCGTGGACAGTGATTCCCAGACGG
GCAGCGGCACCATCCTCACACCTCTGCATTCTGATAGAAGTCTGA
ACAGTTGTTTGTGTTTTTTTTTTTTTTTTTGACGAAGAATGTTTTTA

TTTTTATTTTTTTCATGCATGCATTCTCAAGAGGTCGTGCCAATCA
GCCACTGAAAGGAAAGGCATCACTATGGACTTTCTCTATTTTAAA
ATGGTAACAATCAGAGGAACTATAAGAACACCTTTAGAAATAAAA
ATACTGGGATCAAACTGGCCTGCAAAACCATAGTCAGTTAATTCT
TTTTTTCATCCTTCCTCTGAGGGGG^GATCC*CGCCCCTCTCCCTCC*
CCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCG
GTGTGCGTTTGTCTATATGTTATTTTCCACCATATTGCCGTCTTTTG
GCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCA
TTCCTAGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTT
GAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGACA
AACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACC
TGGCGACAGGTGCCTCTGCGGCCAAAAGCCACGTGTATAAGATAC
ACCTGCAAAGGCGGCACAACCCCAGTGCCACGTTGTGAGTTGGAT
AGTTGTGGAAAGAGTCAAATGGCTCACCTCAAGCGTATTCAACAAG
GGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGATCTGAT
CTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTA
AAAAACGTCTAGGCCCCCCGAACCACGGGGACGTGGTTTTCCTTT
GAAAAACACGATGATAAT<u>ATGGTGAGCAAGGGCGAGGAGGTCATC</u>
<u>AAAGAGTTCATGCGCTTCAAGGTGCGCATGGAGGGCTCCATGAAC</u>
<u>GGCCACGAGTTCGAGATCGAGGGCGAGGGCGAGGGCCGCCCCT</u>
<u>ACGAGGGCACCCAGACCGCCAAGCTGAAGGTGACCAAGGGCGGC</u>
<u>CCCCTGCCCTTCGCCTGGGACATCCTGTCCCCCCCAGTTCATGTAC</u>
<u>GGCTCCAAGGCGTACGTGAAGCACCCCGCCGACATCCCCGATTA</u>
<u>CAAGAAGCTGTCCTTCCCCGAGGGCTTCAAGTGGGAGCGCGTGAT</u>
<u>GAACTTCGAGGACGGCGGTCTGGTGACCGTGACCCAGGACTCCT</u>
<u>CCCTGCAGGACGGCACGCTGATCTACAAGGTGAAGATGCGCGGC</u>
<u>ACCAACTTCCCCCCCGACGGCCCCGTAATGCAGAAGAAGACCATG</u>
<u>GGCTGGGAGGCCTCCACCGAGCGCCTGTACCCCCGCGACGGCGT</u>
<u>GCTGAAGGGCGAGATCCACCAGGCCCTGAAGCTGAAGGACGGCG</u>
<u>GCCACTACCTGGTGGAGTTCAAGACCATCTACATGGCCAAGAAGC</u>
<u>CCGTGCAACTGCCCGGCTACTACTACGTGGACACCAAGCTGGACA</u>
<u>TCACCTCCCACAACGAGGACTACACCATCGTGGAACAGTACGAGC</u>
<u>GCTCCGAGGGCCGCCACCACCTGTTCCTGGGGCATGGCACCGGC</u>

AGCACCGGCAGCGGCAGCTCCGGCACCGCCTCCTCCGAGGACAA
CAACATGGCCGTCATCAAAGAGTTCATGCGCTTCAAGGTGCGCAT
GGAGGGCTCCATGAACGGCCACGAGTTCGAGATCGAGGGCGAGG
GCGAGGGCCGCCCCTACGAGGGCACCCAGACCGCCAAGCTGAAG
GTGACCAAGGGCGGCCCCCTGCCCTTCGCCTGGGACATCCTGTC
CCCCCAGTTCATGTACGGCTCCAAGGCGTACGTGAAGCACCCCGC
CGACATCCCCGATTACAAGAAGCTGTCCTTCCCCGAGGGCTTCAA
GTGGGAGCGCGTGATGAACTTCGAGGACGGCGGTCTGGTGACCG
TGACCCAGGACTCCTCCCTGCAGGACGGCACGCTGATCTACAAGG
TGAAGATGCGCGGCACCAACTTCCCCCCCGACGGCCCCGTAATG
CAGAAGAAGACCATGGGCTGGGAGGCCTCCACCGAGCGCCTGTA
CCCCCGCGACGGCGTGCTGAAGGGCGAGATCCACCAGGCCCTGA
AGCTGAAGGACGGCGGCCACTACCTGGTGGAGTTCAAGACCATCT
ACATGGCCAAGAAGCCCGTGCAACTGCCCGGCTACTACTACGTGG
ACACCAAGCTGGACATCACCTCCCACAACGAGGACTACACCATCG
TGGAACAGTACGAGCGCTCCGAGGGCCGCCACCACCTGTTCCTG
*TACGGCATGGACGAGCTGTACAAG*TGAACGCGTCTGGAACAATCA
ACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACT
ATGTTGCTCCTTTTACGCTATGTGGATA…

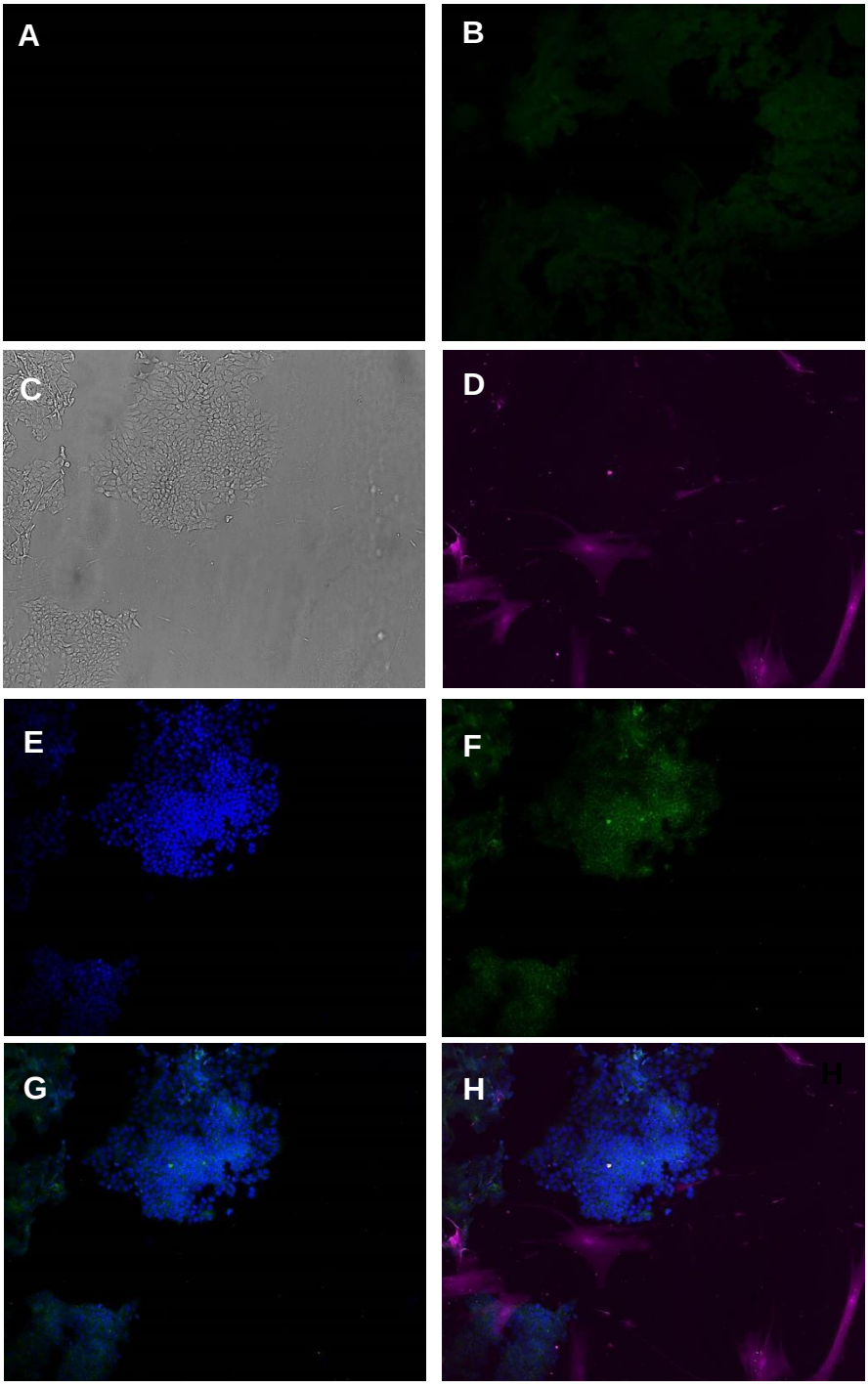
Appendix 6: Immunofluorescence staining of mono-cultures for E-cadherin.

A. Image of cells stained with the secondary antibody only taken with the FITC channel. B. Image of cells stained with the isotype control taken with the FITC channel. Image of cells stained with the E-cadherin antibody taken with the brightfield channel (C), the Dapi channel (D) and the FITC channel (E). F. Merge picture of cells stained with E-cadherin of the Dapi and FITC channels. Dapi shows the nuclei of the cells and FITC shows the expression of the secondary antibody (AF 488) bound to the isotype control or E-cadherin.



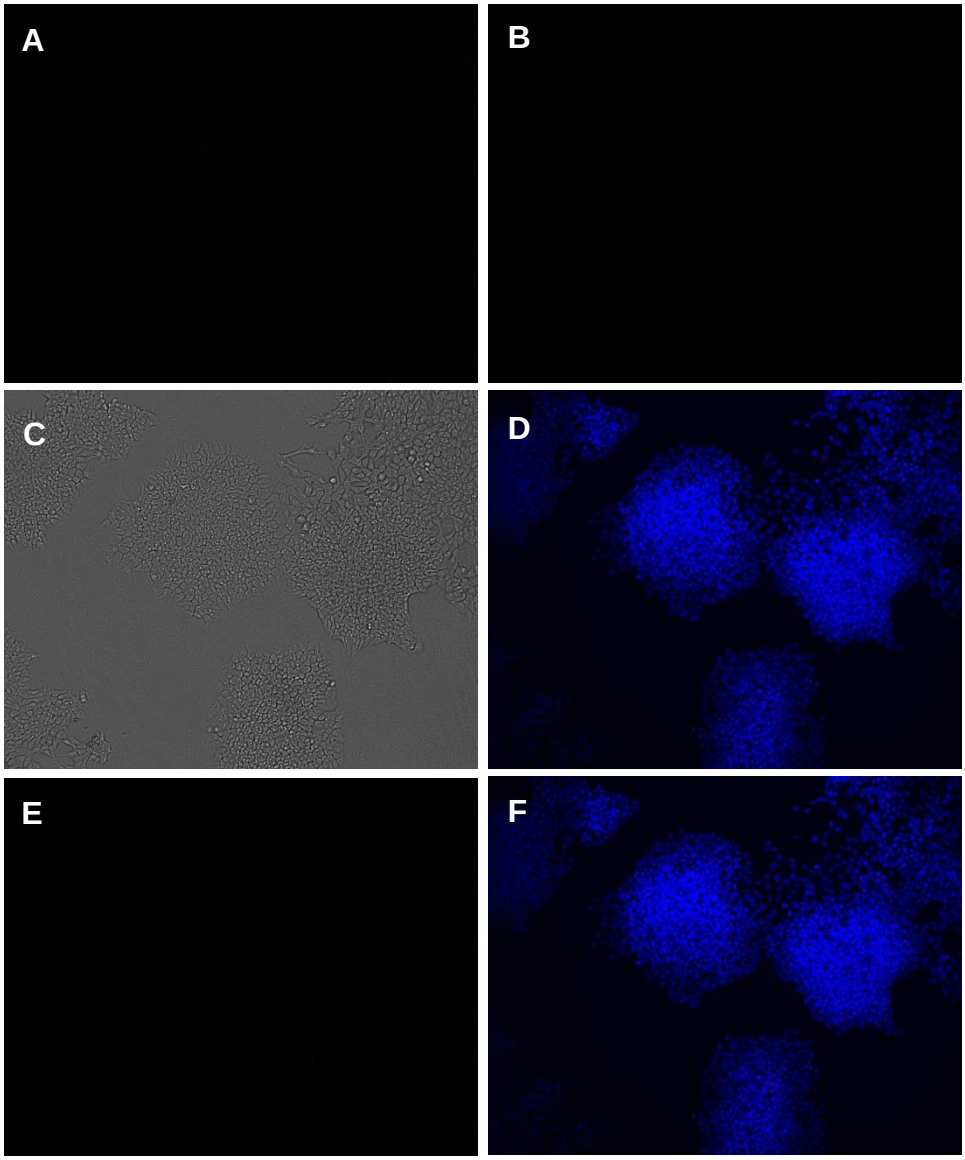
Brightfield/Dapi/FITC (Secondary antibody/E-cadherin)

Appendix 7: Immunofluorescence staining of co-cultures for E-cadherin. A. Image of cells stained with the secondary antibody only taken with the FITC channel. B. Image of cells stained with the isotype control taken with the FITC channel. Image of cells stained with the E-cadherin antibody taken with the brightfield channel (C), the HcRed channel (mCherry MSCs), the Dapi channel (E) and the FITC channel (F). Merge picture of cells stained with E-cadherin of the Dapi and FITC channels (G) or the Dapi, FITC and HcRed (H). Dapi shows the nuclei of the cells and FITC shows the expression of the secondary antibody (AF 488) bound to the isotype control or E-cadherin.



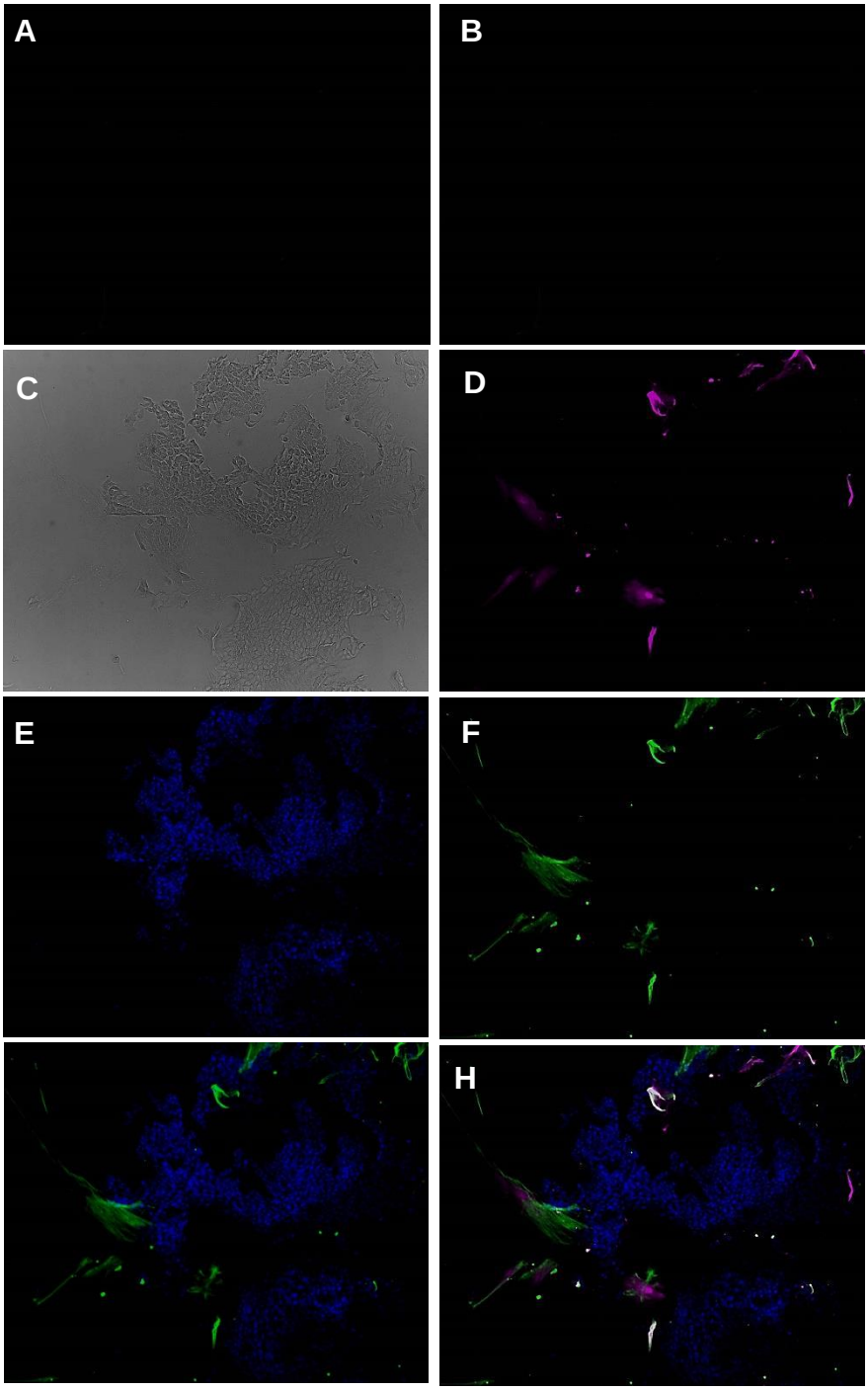
Brightfield/Dapi/FITC (Secondary antibody/E-cadherin)/ HcRed (mCherry MSCs)

Appendix 8: IF staining of mono-cultures for Vimentin. A. Image of cells stained with the secondary antibody only taken with the FITC channel. B. Image of cells stained with the isotype control taken with the FITC channel. Image of cells stained with the Vimentin antibody taken with the brightfield channel (C), the Dapi channel (D) and the FITC channel (E). F. Merge picture of cells stained with E-cadherin of the Dapi and FITC channels. Dapi shows the nuclei of the cells and FITC shows the expression of the secondary antibody (AF 488) bound to the isotype control or Vimentin.



Brightfield/Dapi/FITC (Secondary antibody/Vimentin)

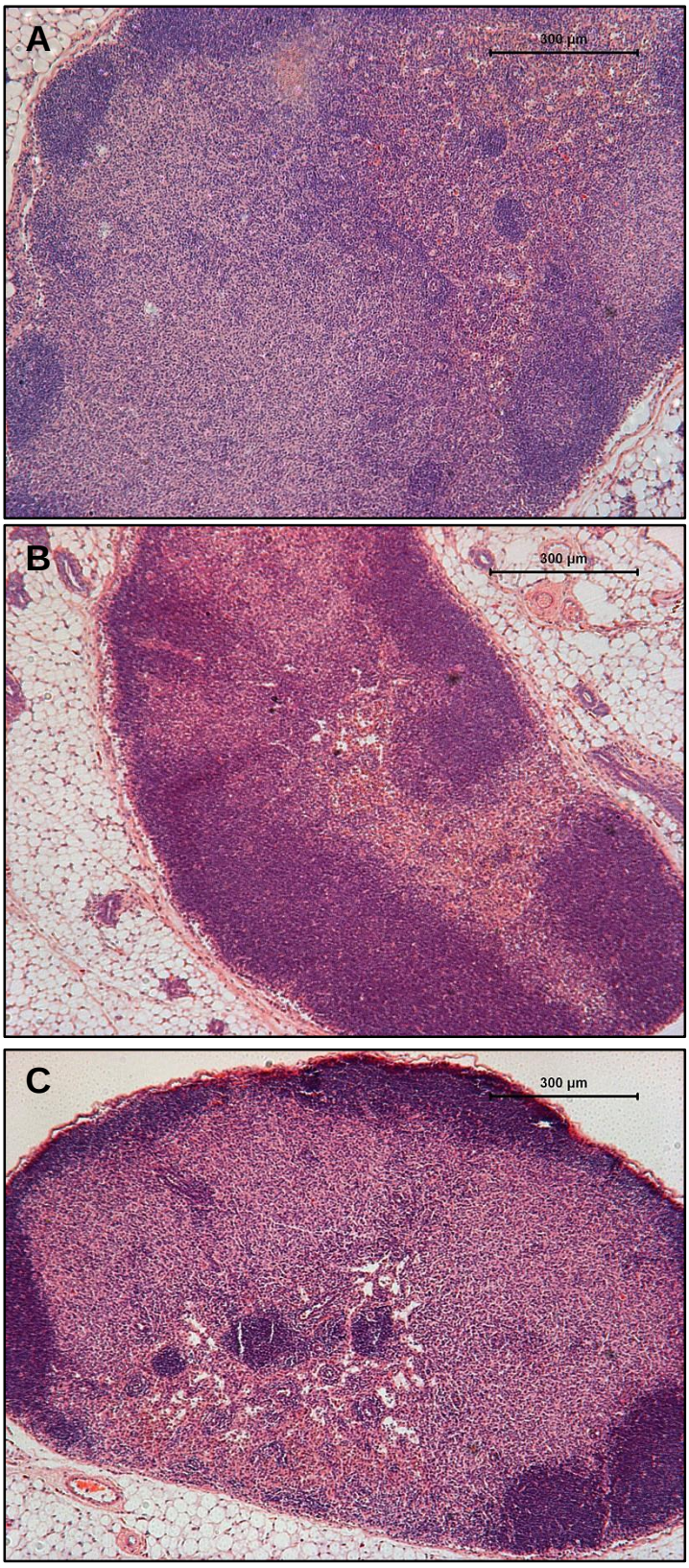
Appendix 9. IF staining of co-cultures for Vimentin.
A. Image of cells stained with the secondary antibody only taken with the FITC channel. B. Image of cells stained with the isotype control taken with the FITC channel. Image of cells stained with the Vimentin antibody taken with the brightfield channel (C), the HcRed channel (mCherry MSCs), the Dapi channel (E) and the FITC channel (F). Merge picture of cells stained with Vimentin of the Dapi and FITC channels (G) or the Dapi, FITC and HcRed (H). Dapi shows the nuclei of the cells and FITC shows the expression of the secondary antibody (AF 488) bound to the isotype control or Vimentin.



Brightfield/Dapi/FITC (Secondary antibody/E-cadherin)/ HcRed (mCherry MSCs)

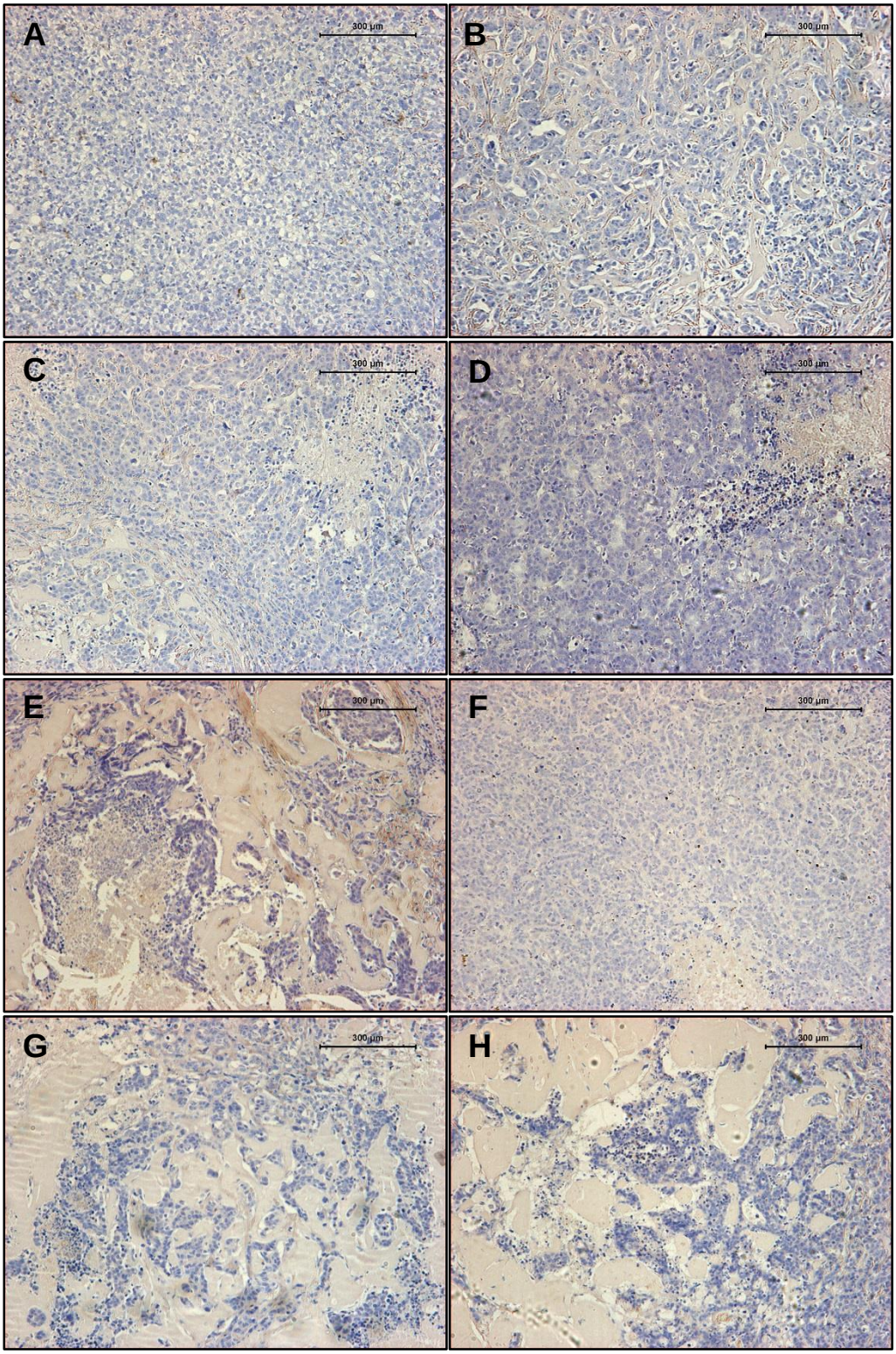
Appendix 10: H&E staining of LNs

A. Lymph node 1 (LHS). **B.** Lymph node 3 (LHS). **C.** Lymph node 4 (LHS).



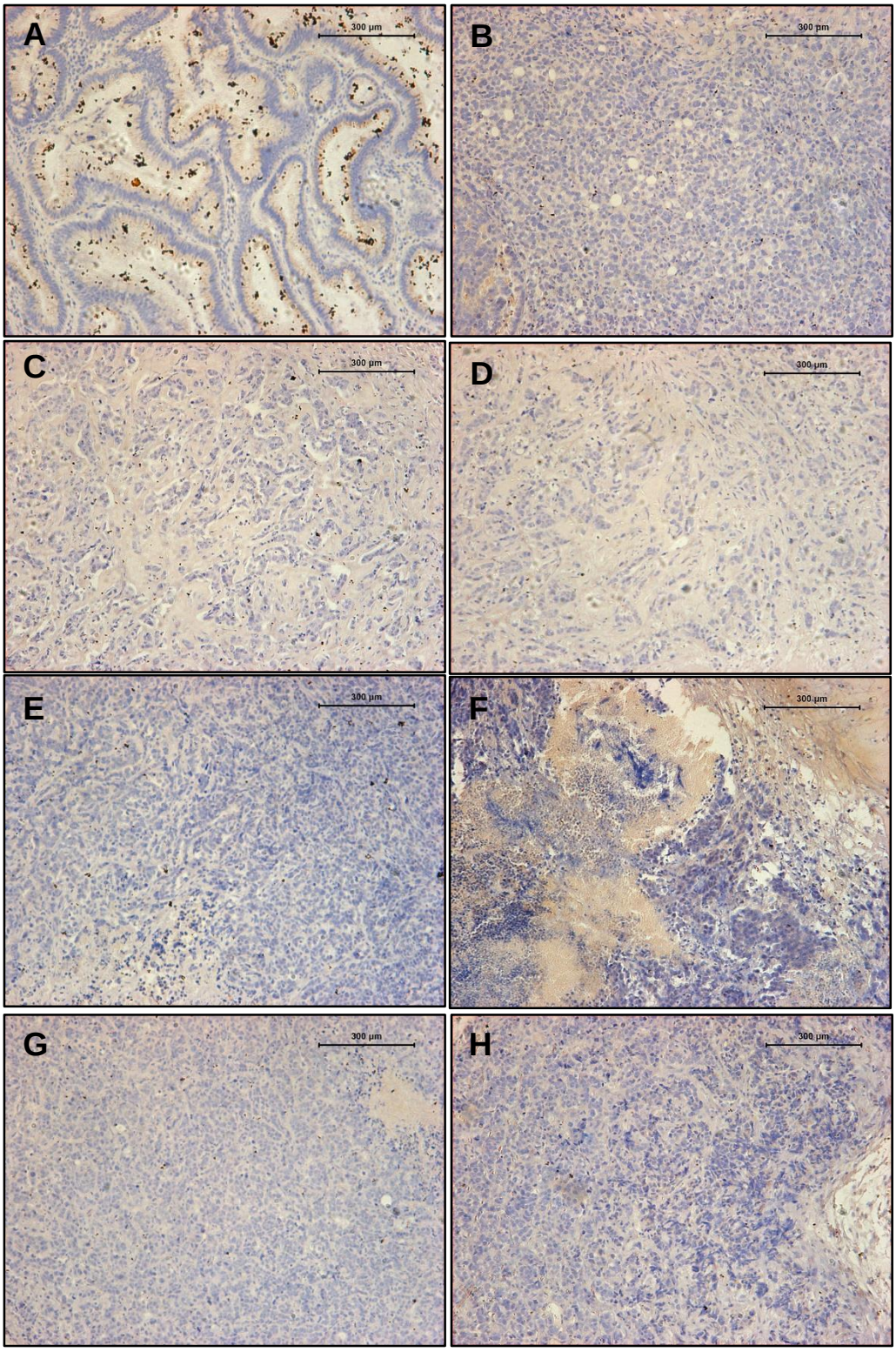
Appendix 11: Isotype controls for E-cadherin staining

A-D. Mice 1-4 from group 1 (MCF-7/fLuc cells in the MFP). **C-H** Mice 5-8 from group 2 (MCF-7/fLuc cells SC). Control antibody = Mouse IgG1



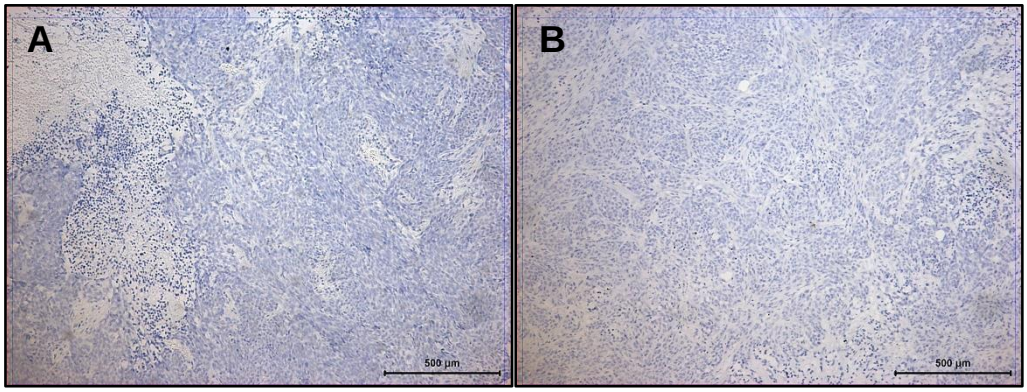
Appendix 12. Isotype controls for Vimentin staining

A. Positive control (human colorectal tumour). **B-E** xenografts 1-4 from group 1. **F-H.** Xenografts 5-7 from group 2. Isotype control = Mouse IgG1



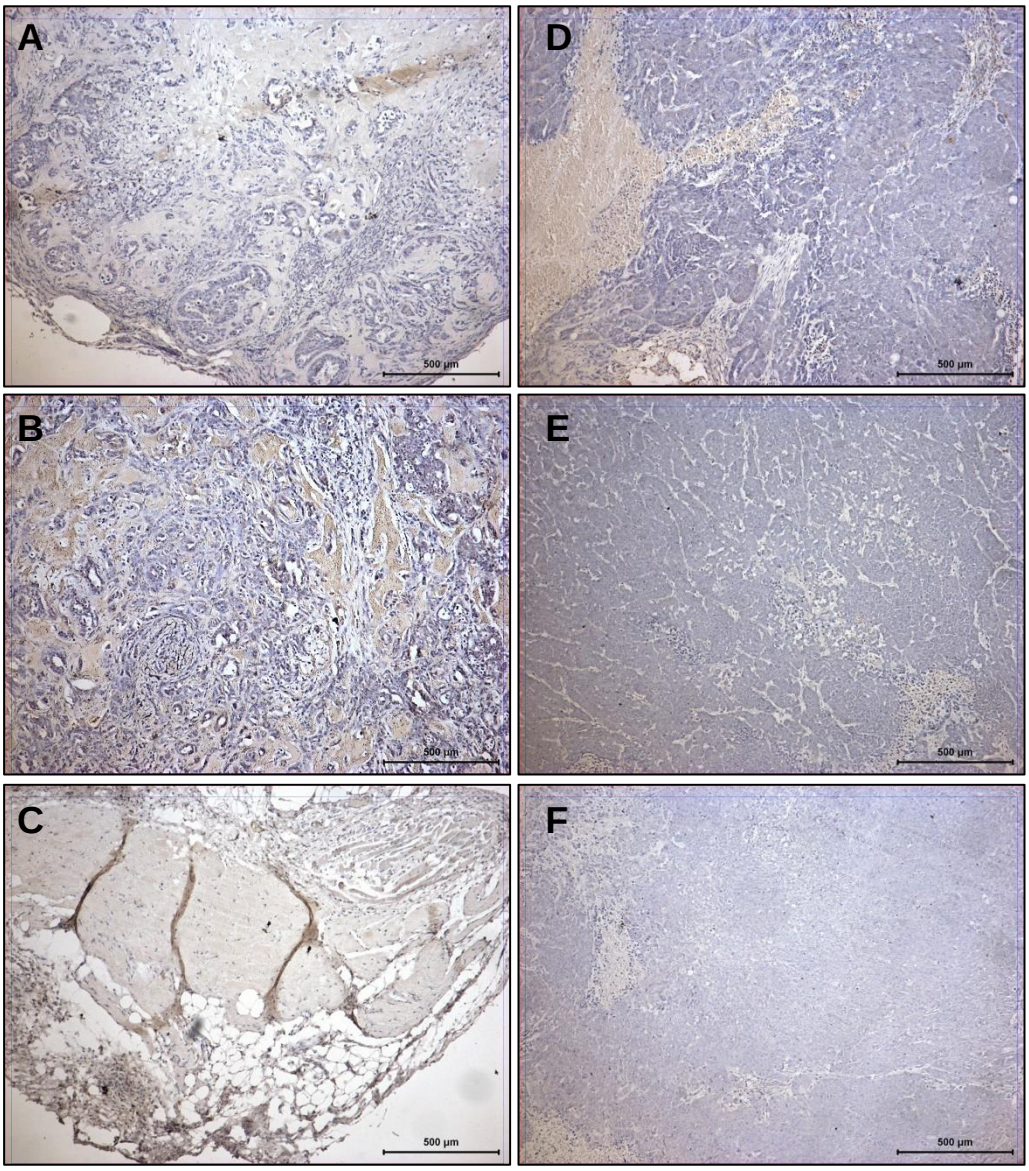
Appendix 13: Isotype controls for S100A4

A. Xenograft from mouse 1. **B.** Xenograft from mouse 2. Control = Rabbit serum.



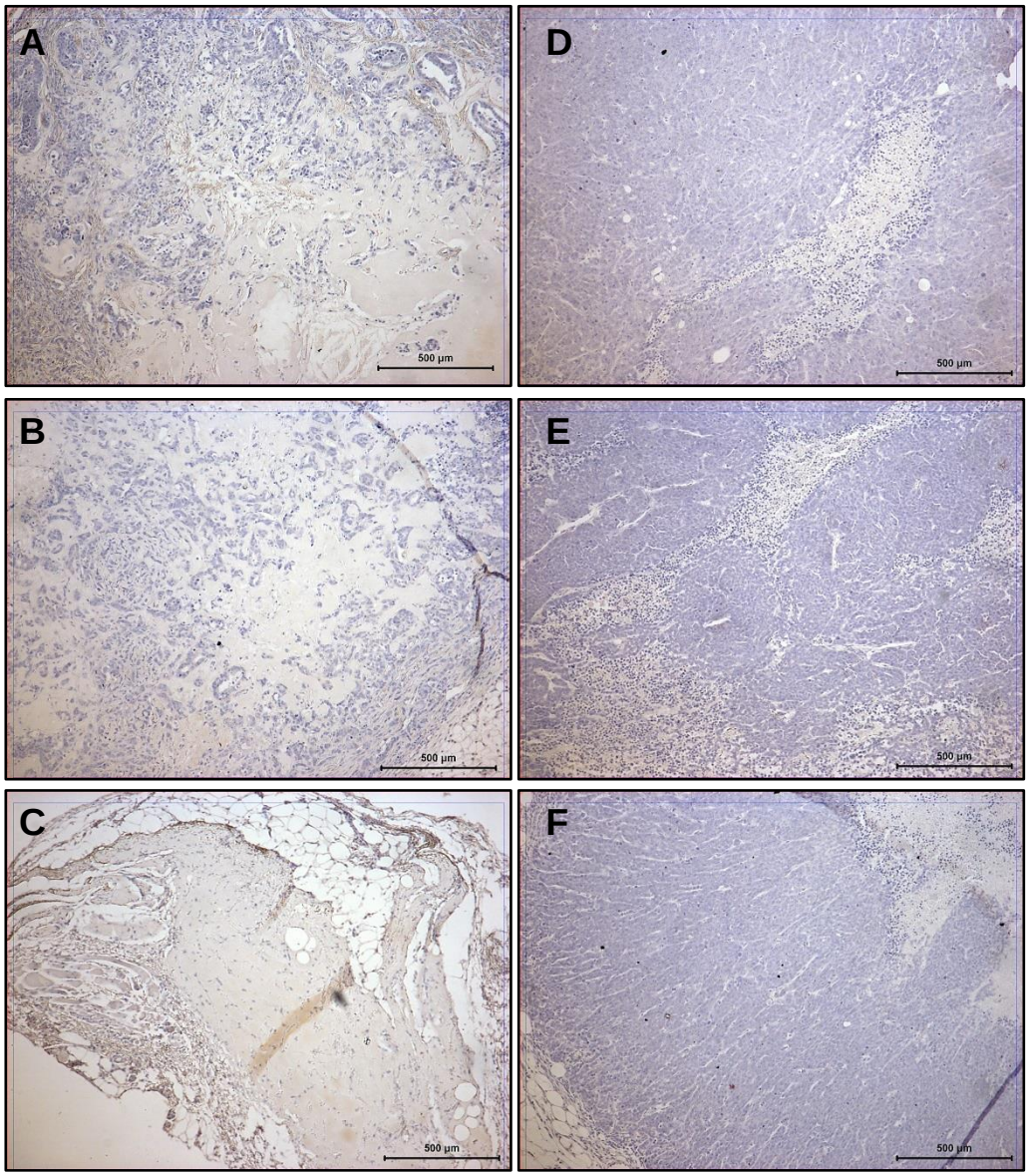
Appendix 14: Isotype controls for E-cadherin staining

A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-6 (Group 2; MCF-7 + MSCs). Isotype control = Mouse IgG1



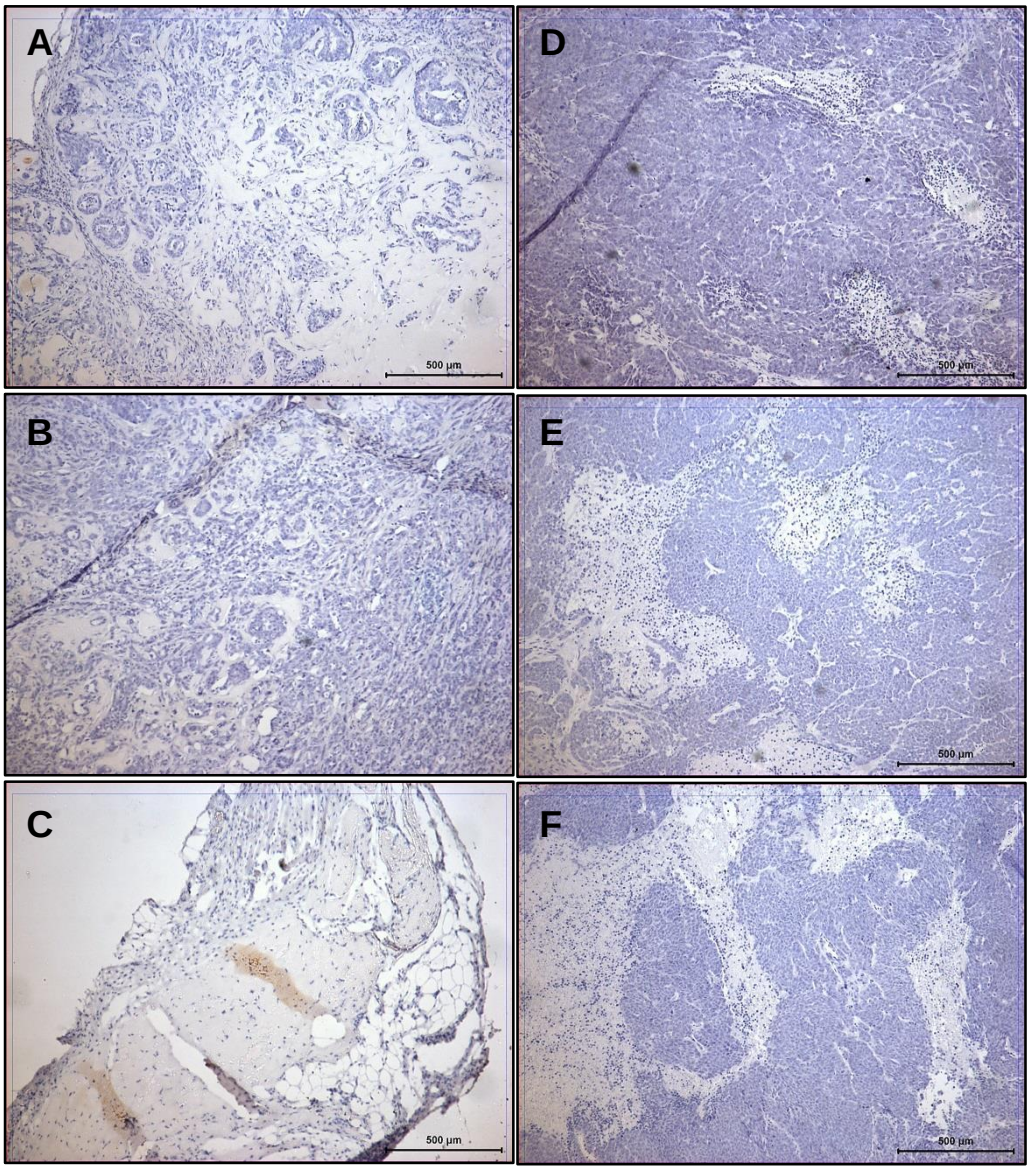
Appendix 15: Isotype controls for Vimentin staining

A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-5 (Group 2;MCF-7 + MSCs). Isotype control = Mouse IgG1.



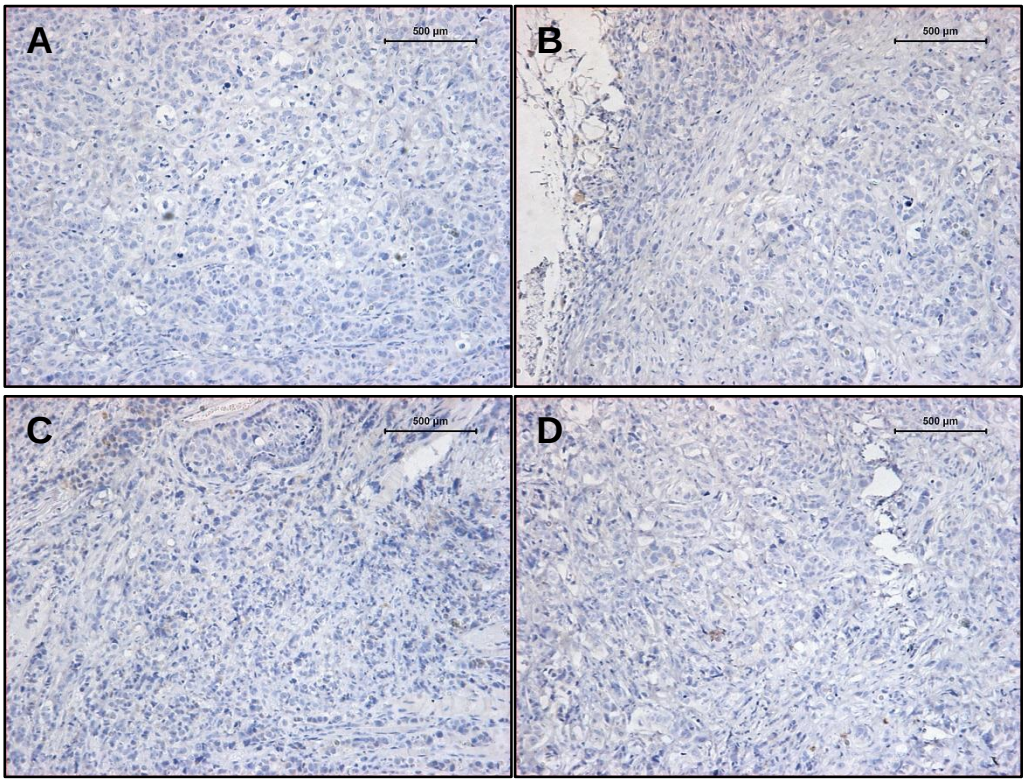
Appendix 16: Isotype controls for S100A4 staining

A-C. Xenografts from mice 1-3 (Group 1; MCF-7 alone). **D-F.** Xenografts from mice 4-6 (Group 2; MCF-7 + MSCs). Control = Rabbit serum



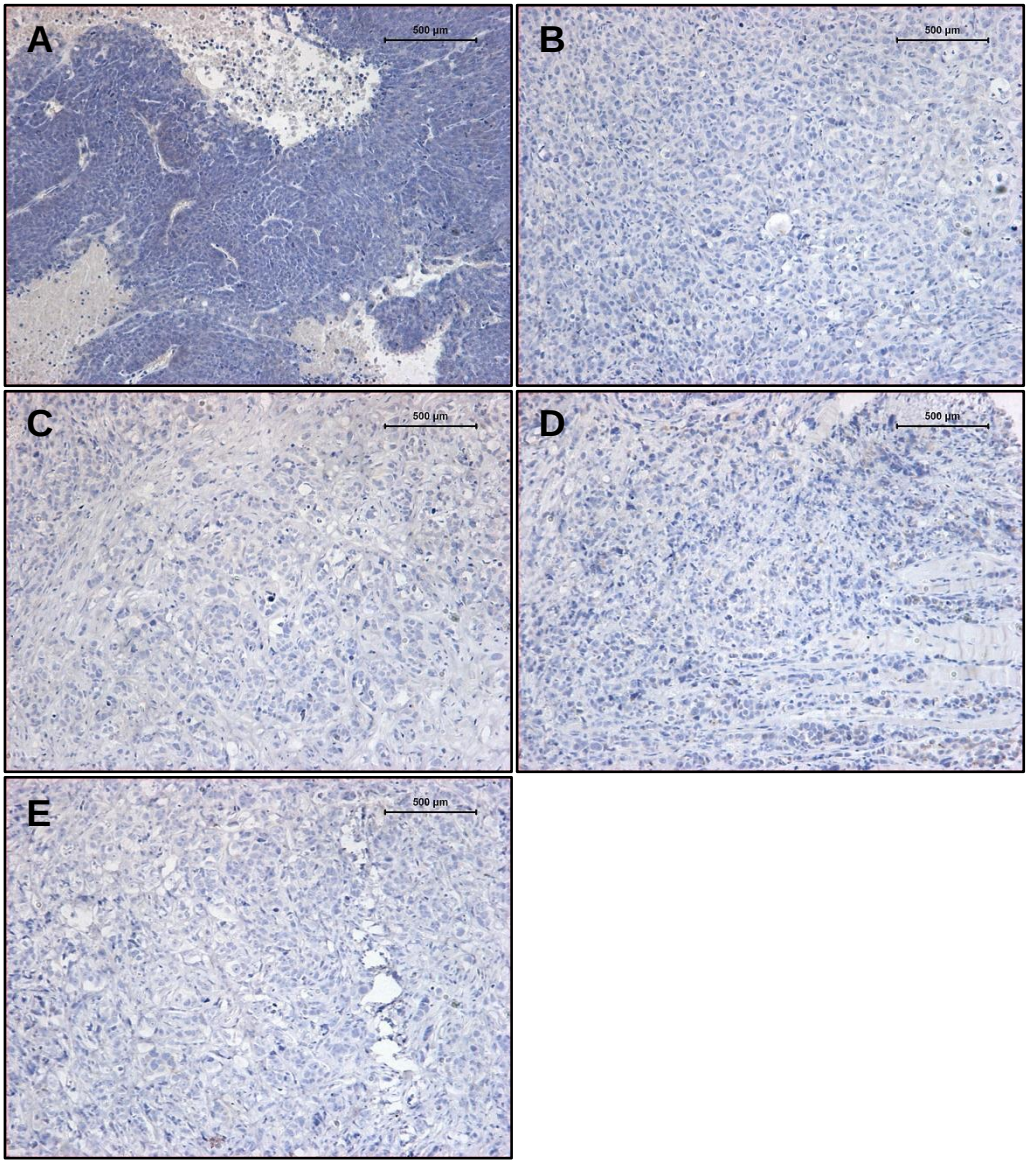
Appendix 17: Isotype controls for E-cadherin staining

A-B. Xenografts from mice 1-2 (Group 1; MCF-7 alone). **C-D.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs). Isotype control = Mouse IgG1.



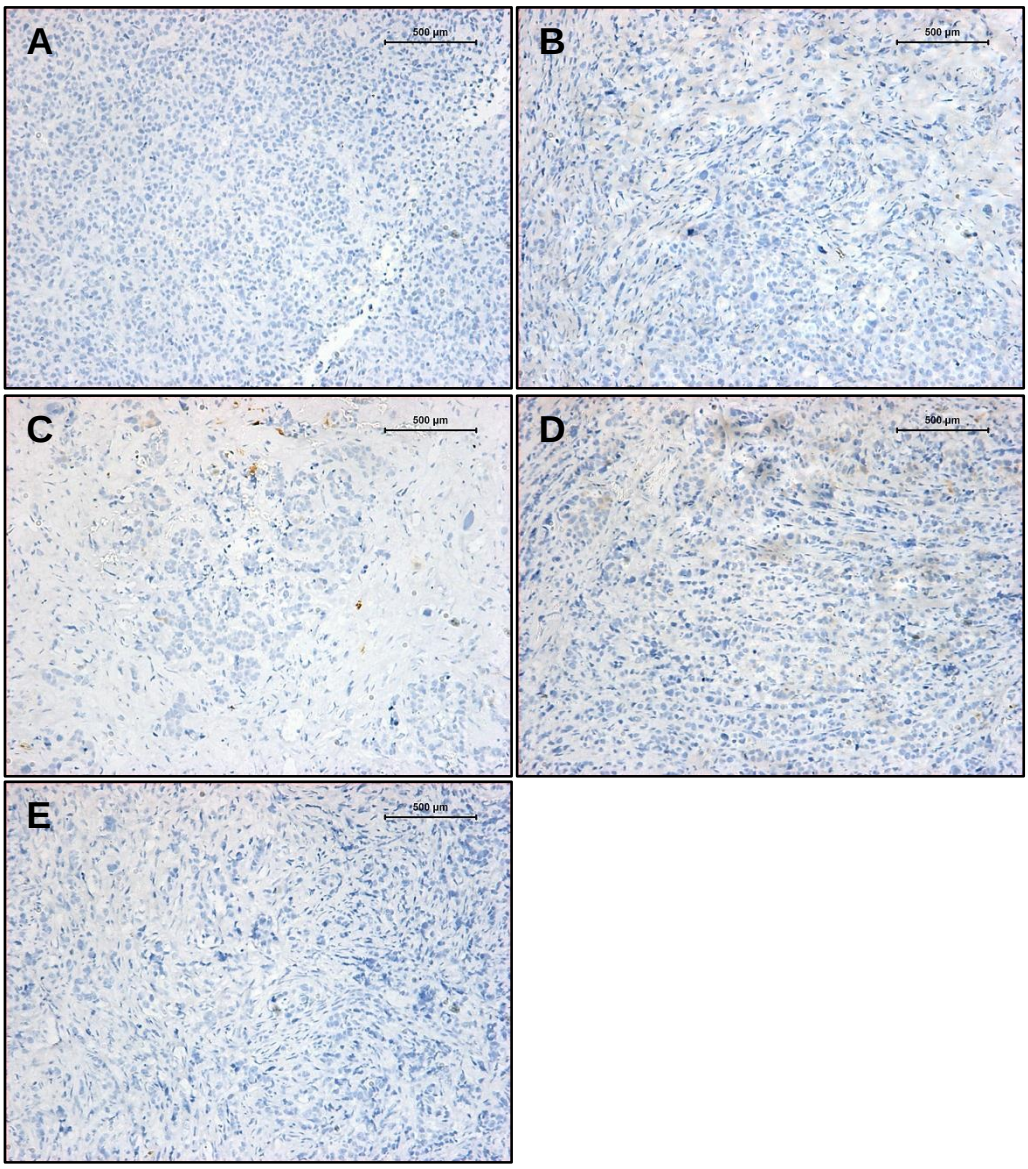
Appendix 18: Isotype controls for Vimentin staining

A. Positive control (xenograft 6, Figure 81A). **B-C.** Xenografts from mice 1-2 (Group 1; MCF-7 alone). **D-E.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs). Isotype control = Mouse IgG1.



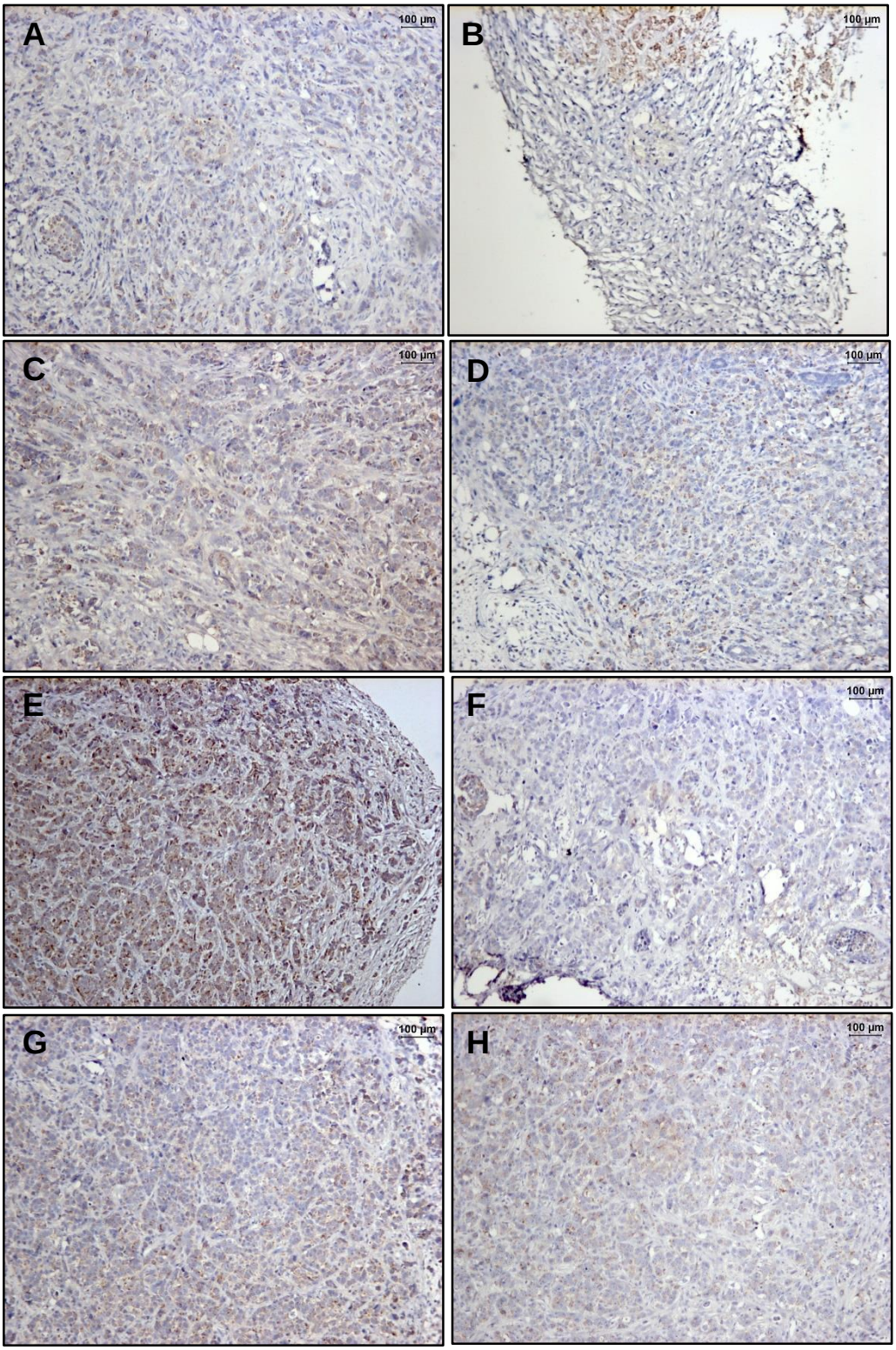
Appendix 19: Isotype controls for S100A4 staining

A. Positive control (MDA-MB-231 xenograft from a study performed in house). **B-C.** Xenografts from mice 1-2 (Group 1; MCF-7 alone). **D-E.** Xenografts from mice 3-4 (Group 2; MCF-7 + MSCs). Control = Rabbit serum.



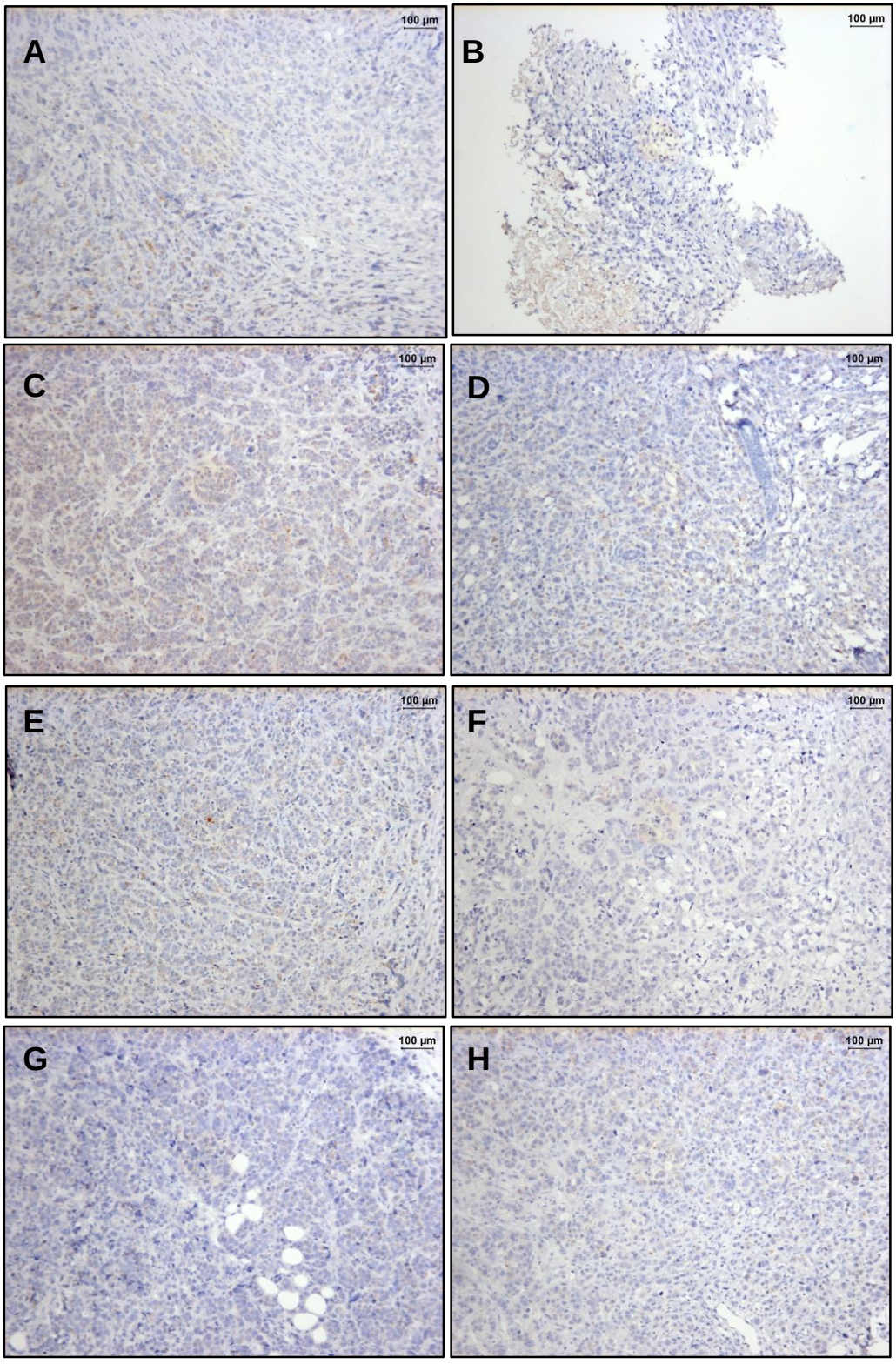
Appendix 20: Isotype controls for E-cadherin staining

A-D. Xenografts from mice 1-4 (Group 1; MCF-7/fLuc). **E-H.** Xenografts from mice 5-8 (Group 2; MCF-7/fLuc + MSCs). Isotype control = Mouse IgG1



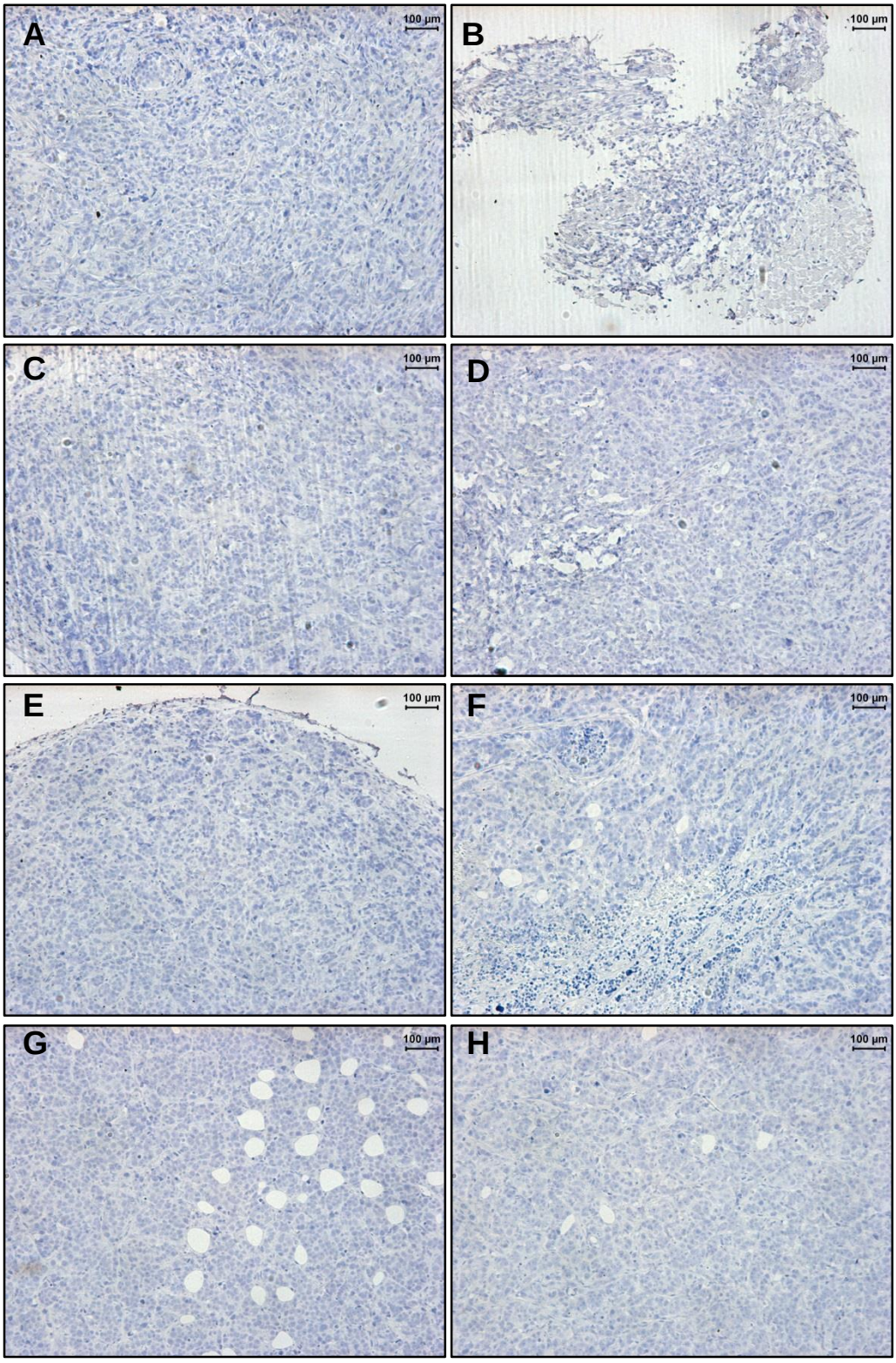
Appendix 21: Isotype controls for Vimentin staining

A-D. Xenografts from mice 1-4 (Group 1; MCF-7/fLuc). **E-H.** Xenografts from mice 5-8 (Group 2; MCF-7/fLuc + MSCs). Isotype control = Mouse IgG1



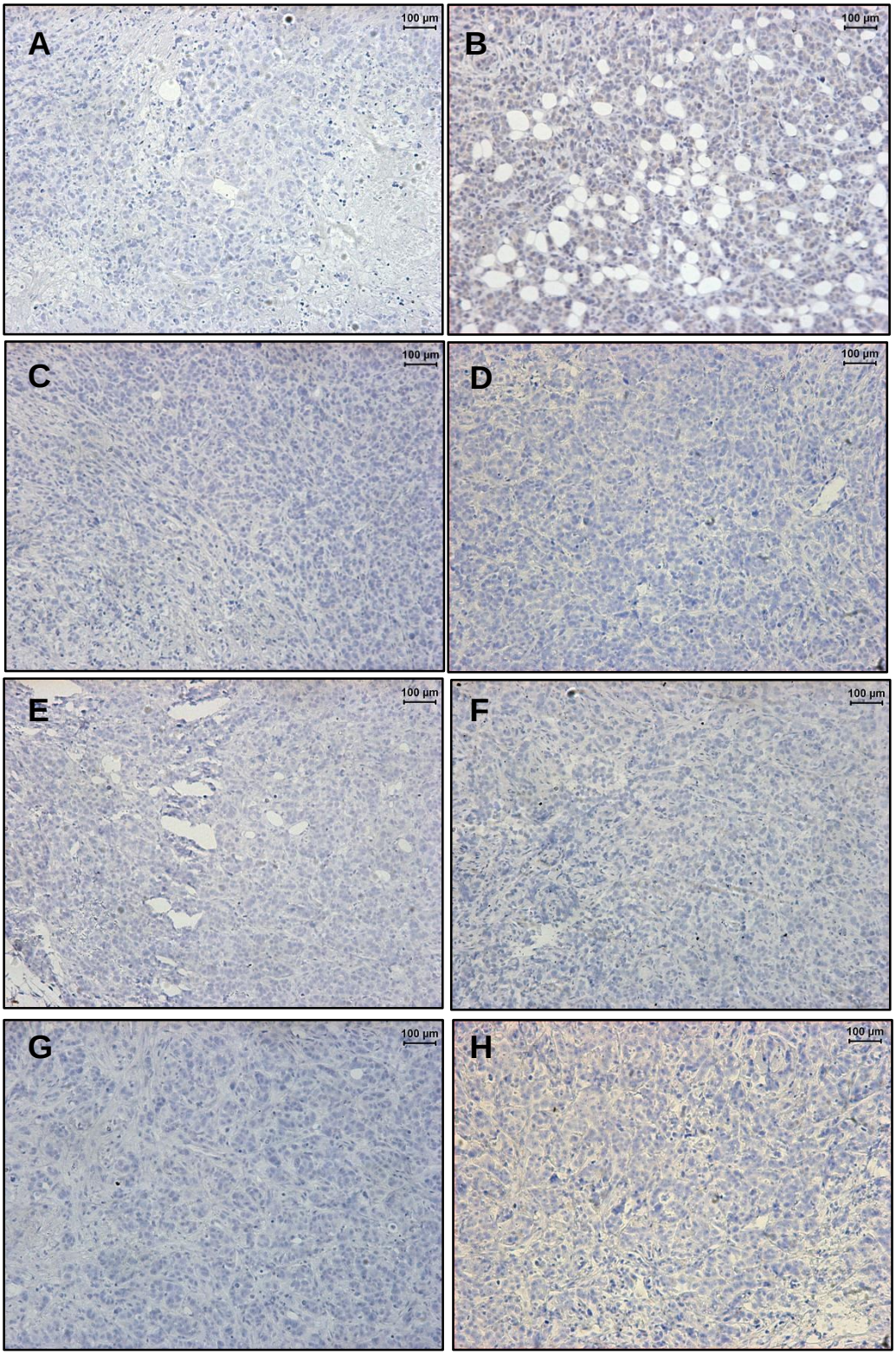
Appendix 22: Isotype controls for S100A4 staining

A-D. Xenografts from mice 1-4 (Group 1; MCF-7/fLuc). **E-H.** Xenografts from mice 5-8 (Group 2; MCF-7/fLuc + MSCs). Isotype control = Rabbit serum



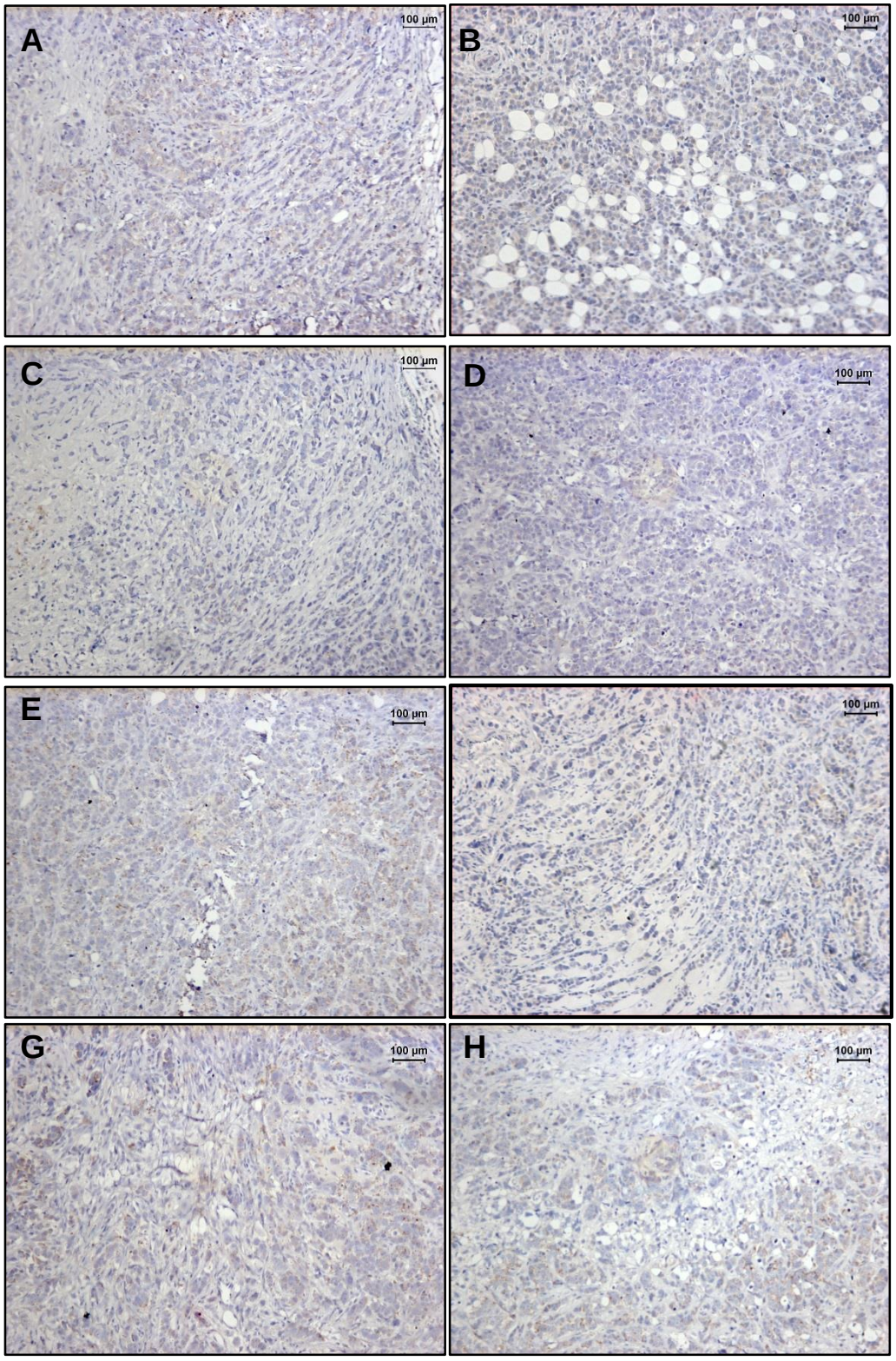
Appendix 23: Isotype controls for S100A4 staining

A-D. Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).. Control antibody = Rabbit serum



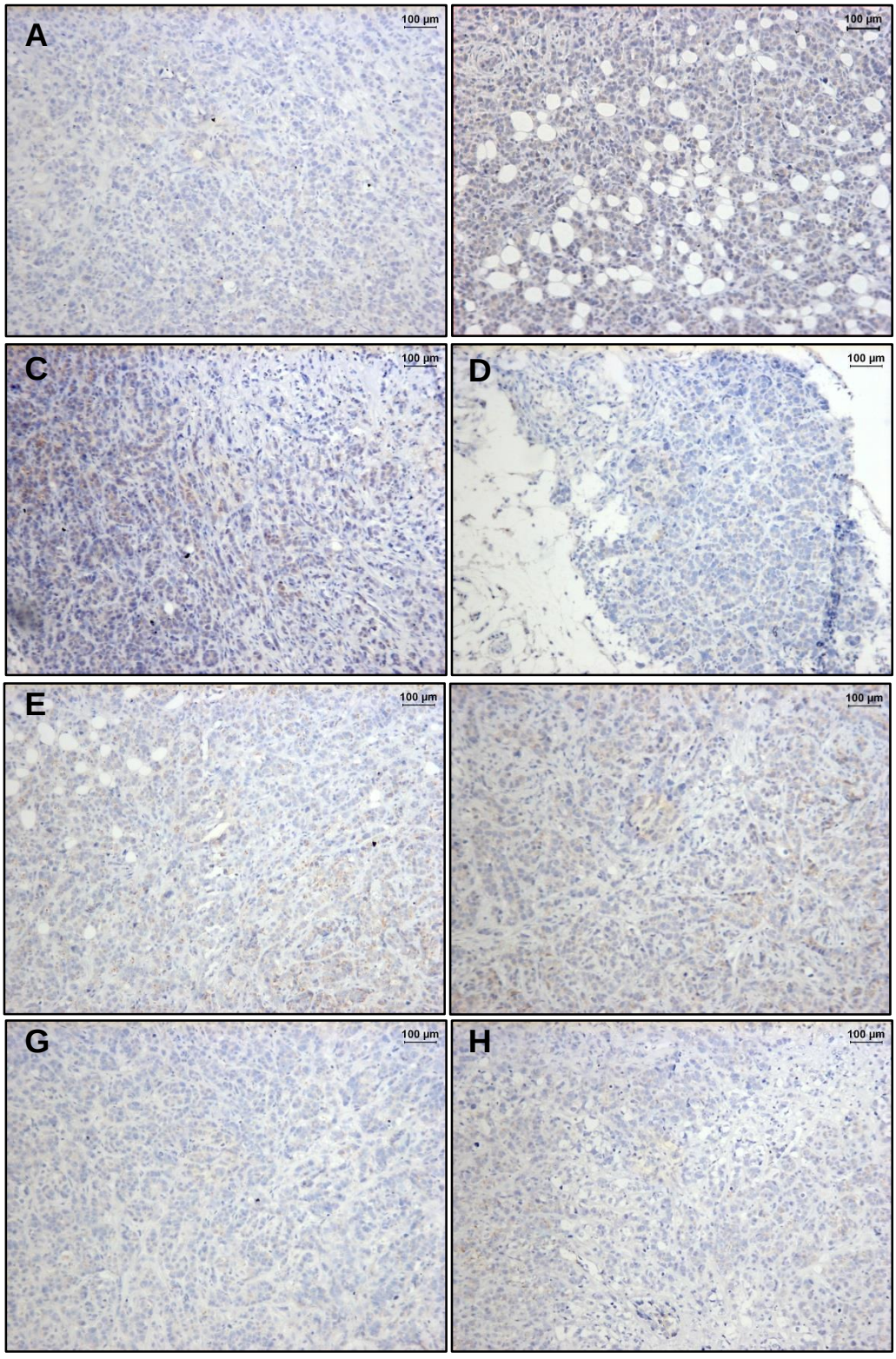
Appendix 24: Isotype controls for E-cadherin staining

A-D. Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).. Control antibody = Mouse IgG1



Appendix 25: Isotype controls for Vimentin staining

A-D. Mice 1-4 from group 1 (MCF-7/S100A4-fLuc/mCherry). **E-H** Mice 5-8 from group 2 (MCF-7/S100A4-fLuc/mCherry + MSCs).. Control antibody = Mouse IgG1



REFERENCES

1. Donald A. Berry, K.A.C., Sylvia K. Plevritis, Dennis G. Fryback, Lauren Clarke, M.S., Marvin Zelen, Jeanne S. Mandelblatt, Andrei Y. Yakovlev, J. Dik F. Habbema, and Eric J. Feuer, *Effect of Screening and Adjuvant Therapy on Mortality from Breast Cancer.* N Engl J Med, 2005. **353**: p. 1784-1792.
2. Overgaard, M., et al., *Postoperative radiotherapy in high-risk postmenopausal breast-cancer patients given adjuvant tamoxifen: Danish Breast Cancer Cooperative Group DBCG 82c randomised trial.* Lancet, 1999. **353**(9165): p. 1641-8.
3. Ahmedin Jemal, F.B., Melissa M. Center, Jacques Ferlay, Elizabeth Ward, David Forman, *Global Cancer Statistics.* CA Cancer J Clin, 2011. **61**: p. 69-90.
4. Webb, M.L., et al., *A failure analysis of invasive breast cancer: Most deaths from disease occur in women not regularly screened.* Cancer, 2013. **9**(10): p. 28199.
5. Chambers A.F, G.A.C., MacDonald I.C, *Dissemination and growth of cancer cells in metastatic sites.* Nature Reviews Cancer, 2002. **2**(8): p. 563-572.
6. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
7. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proc Natl Acad Sci U S A, 2001. **98**(19): p. 10869-74.
8. Sotiriou S., et al., *Breast cancer classification and prognosis based on gene expression profiles from a population-based study.* Proc Natl Acad Sci U S A., 2003. **100**(18): p. 10393-8.
9. Shah, M. and C. Allegrucci, *Keeping an open mind: highlights and controversies of the breast cancer stem cell theory.* Breast Cancer: Targets and Therapy, 2012. **4**: p. 156-166.
10. Hogan, N.M., et al., *Mesenchymal stem cells in the colorectal tumor microenvironment: recent progress and implications.* Int J Cancer, 2012. **131**(1): p. 1-7.
11. Cleator, S., W. Heller, and R.C. Coombes, *Triple-negative breast cancer: therapeutic options.* Lancet Oncol, 2007. **8**(3): p. 235-44.
12. van de Rijn, M., et al., *Expression of cytokeratins 17 and 5 identifies a group of breast carcinomas with poor clinical outcome.* Am J Pathol, 2002. **161**(6): p. 1991-6.

13. Nielsen, T.O., et al., *Immunohistochemical and clinical characterization of the basal-like subtype of invasive breast carcinoma.* Clin Cancer Res, 2004. **10**(16): p. 5367-74.
14. Brenton, J.D., et al., *Molecular classification and molecular forecasting of breast cancer: ready for clinical application?* J Clin Oncol, 2005. **23**(29): p. 7350-60.
15. Foulkes, W.D., et al., *The prognostic implication of the basal-like (cyclin E high/p27 low/p53 +/-glomeruloid-microvascular-proliferation+) phenotype of BRCA-1 related breast cancer.* Cancer Res, 2004. **64**(3): p. 830-5.
16. Wright, M.H., et al., *Brca1 breast tumors contain distinct CD44+/CD24- and CD133+ cells with cancer stem cell characteristics.* Breast Cancer Res, 2008. **10**(1): p. R10.
17. Marija Balic, H.L., Lillian Young, et al., *Most Early Disseminated Cancer Cells Detected in Bone Marrow of Breast Cancer Patients Have a Putative Breast Cancer Stem Cell Phenotype.* Clin Cancer Res, 2006. **12**(5615-5621).
18. Guarino M., Rubino B., and Ballabio G., *The role of epithelial-mesenchymal transition in cancer pathology.* Pathology., 2007. **39**(3): p. 305-18.
19. Heppner, G.H. and B.E. Miller, *Tumor heterogeneity: biological implications and therapeutic consequences.* Cancer Metastasis Rev, 1983. **2**(1): p. 5-23.
20. Reya, T., et al., *Stem cells, cancer, and cancer stem cells.* Nature, 2001. **414**(6859): p. 105-11.
21. Bonnet, D. and J.E. Dick, *Human acute myeloid leukemia is organized as a hierarchy that originates from a primitive hematopoietic cell.* Nat Med, 1997. **3**(7): p. 730-7.
22. Nowell, P.C., *The clonal evolution of tumor cell populations.* Science, 1976. **194**(4260): p. 23-8.
23. Campbell, L.L. and K. Polyak, *Breast tumor heterogeneity: cancer stem cells or clonal evolution?* Cell Cycle, 2007. **6**(19): p. 2332-8.
24. Liu, S. and M.S. Wicha, *Targeting breast cancer stem cells.* J Clin Oncol, 2010. **28**(25): p. 4006-12.
25. Visvader, J.E. and G.J. Lindeman, *Cancer stem cells in solid tumours: accumulating evidence and unresolved questions.* Nat Rev Cancer, 2008. **8**(10): p. 755-68.
26. Clevers, H., *The cancer stem cell: premises, promises and challenges.* Nat Med, 2011. **17**(3): p. 313-9.
27. Al-Hajj, M., et al., *Prospective identification of tumorigenic breast cancer cells.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3983-8.

28. Phillips, T.M., W.H. McBride, and F. Pajonk, *The response of CD24(-/low)/CD44+ breast cancer-initiating cells to radiation.* J Natl Cancer Inst, 2006. **98**(24): p. 1777-85.
29. Sheridan, C., et al., *CD44+/CD24- breast cancer cells exhibit enhanced invasive properties: an early step necessary for metastasis.* Breast Cancer Res, 2006. **8**(5): p. R59.
30. Fillmore, C.M. and C. Kuperwasser, *Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy.* Breast Cancer Res, 2008. **10**(2): p. R25.
31. Ricardo, S., et al., *Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype.* J Clin Pathol, 2011. **64**(11): p. 937-46.
32. Honeth, G., et al., *The CD44+/CD24- phenotype is enriched in basal-like breast tumors.* Breast Cancer Res, 2008. **10**(3): p. R53.
33. Ginestier, C., et al., *ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome.* Cell Stem Cell, 2007. **1**(5): p. 555-67.
34. Servais, C. and N. Erez, *From sentinel cells to inflammatory culprits: cancer-associated fibroblasts in tumour-related inflammation.* J Pathol, 2013. **229**(2): p. 198-207.
35. Santisteban, M., et al., *Immune-induced epithelial to mesenchymal transition in vivo generates breast cancer stem cells.* Cancer Res, 2009. **69**(7): p. 2887-95.
36. Blick, T., et al., *Epithelial mesenchymal transition traits in human breast cancer cell lines parallel the CD44(hi/)CD24 (lo/-) stem cell phenotype in human breast cancer.* J Mammary Gland Biol Neoplasia, 2010. **15**(2): p. 235-52.
37. Morel, A.P., et al., *Generation of breast cancer stem cells through epithelial-mesenchymal transition.* PLoS One, 2008. **3**(8): p. e2888.
38. Kalluri, R. and E.G. Neilson, *Epithelial-mesenchymal transition and its implications for fibrosis.* J Clin Invest, 2003. **112**(12): p. 1776-84.
39. Acloque, H., et al., *Epithelial-mesenchymal transitions: the importance of changing cell state in development and disease.* J Clin Invest, 2009. **119**(6): p. 1438-49.
40. Wang, Y. and B.P. Zhou, *Epithelial-mesenchymal transition in breast cancer progression and metastasis.* Chin J Cancer, 2011 **30**(9): p. 603-11.
41. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.

42. Kalluri, R. and R.A. Weinberg, *The basics of epithelial-mesenchymal transition.* J Clin Invest, 2009. **119**(6): p. 1420-8.
43. Thiery, J.P. and J.P. Sleeman, *Complex networks orchestrate epithelial-mesenchymal transitions.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 131-42.
44. Foroni, C., et al., *Epithelial-mesenchymal transition and breast cancer: Role, molecular mechanisms and clinical impact.* Cancer Treat Rev, 2012. **38**(6): p. 689-697.
45. Tomaskovic-Crook, E., E.W. Thompson, and J.P. Thiery, *Epithelial to mesenchymal transition and breast cancer.* Breast Cancer Res, 2009. **11**(6): p. 213.
46. Lamouille Samy., Jian Xu., and Derynck R., *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol., 2014. **15**(3): p. 178-96. doi: 10.1038/nrm3758.
47. Teschendorff, A.E., et al., *Elucidating the altered transcriptional programs in breast cancer using independent component analysis.* PLoS Comput Biol, 2007. **3**(8): p. e161.
48. Willipinski-Stapelfeldt, B., et al., *Changes in cytoskeletal protein composition indicative of an epithelial-mesenchymal transition in human micrometastatic and primary breast carcinoma cells.* Clin Cancer Res, 2005. **11**(22): p. 8006-14.
49. Mironchik, Y., et al., *Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer.* Cancer Res, 2005. **65**(23): p. 10801-9.
50. Moody, S.E., et al., *The transcriptional repressor Snail promotes mammary tumor recurrence.* Cancer Cell, 2005. **8**(3): p. 197-209.
51. Sarrio, D., et al., *Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype.* Cancer Res, 2008. **68**(4): p. 989-97.
52. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
53. Gregory, P.A., et al., *The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1.* Nat Cell Biol, 2008. **10**(5): p. 593-601.
54. Ma, L., J. Teruya-Feldstein, and R.A. Weinberg, *Tumour invasion and metastasis initiated by microRNA-10b in breast cancer.* Nature, 2007. **449**(7163): p. 682-8.
55. Cai J., et al., *MicroRNA-374a activates Wnt/beta-catenin signaling to promote breast cancer metastasis.* J Clin Invest., 2013. **123**(2): p. 566-79. doi: 10.1172/JCI65871. Epub 2013 Jan 16.

56. Gwak JM., et al., *MicroRNA-9 is associated with epithelial-mesenchymal transition, breast cancer stem cell phenotype, and tumor progression in breast cancer.* Breast Cancer Res Treat., 2014. **147**(1): p. 39-49. doi: 10.1007/s10549-014-3069-5. Epub 2014 Aug 3.
57. Radisky, D.C. and M.A. LaBarge, *Epithelial-mesenchymal transition and the stem cell phenotype.* Cell Stem Cell, 2008. **2**(6): p. 511-2.
58. Shimono, Y., et al., *Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells.* Cell, 2009. **138**(3): p. 592-603.
59. Taube, J.H., et al., *Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes.* Proc Natl Acad Sci U S A, 2010. **107**(35): p. 15449-54.
60. Zeng, Q., et al., *CD146, an epithelial-mesenchymal transition inducer, is associated with triple-negative breast cancer.* Proc Natl Acad Sci U S A, 2011. **109**(4): p. 1127-32.
61. Liotta, L.A. and E.C. Kohn, *The microenvironment of the tumour-host interface.* Nature, 2001. **411**(6835): p. 375-9.
62. Quail DF., and Joyce JA, *Microenvironmental regulation of tumor progression and metastasis.* Nat Med., 2013. **19**(11): p. 1423-37. doi: 10.1038/nm.3394.
63. Korkaya, H., S. Liu, and M.S. Wicha, *Breast cancer stem cells, cytokine networks, and the tumor microenvironment.* J Clin Invest, 2011. **121**(10): p. 3804-9.
64. Mao Y., et al., *Stromal cells in tumor microenvironment and breast cancer.* Cancer Metastasis Rev., 2013. **32**(1-2): p. 303-15. doi: 10.1007/s10555-012-9415-3.
65. Tse, J.C. and R. Kalluri, *Mechanisms of metastasis: epithelial-to-mesenchymal transition and contribution of tumor microenvironment.* J Cell Biochem, 2007. **101**(4): p. 816-29.
66. Jing, Y., et al., *Epithelial-Mesenchymal Transition in tumor microenvironment.* Cell Biosci, 2011. **1**: p. 29.
67. Solinas, G., et al., *Inflammation-mediated promotion of invasion and metastasis.* Cancer Metastasis Rev, 2010. **29**(2): p. 243-8.
68. Dong, R., et al., *Role of nuclear factor kappa B and reactive oxygen species in the tumor necrosis factor-alpha-induced epithelial-mesenchymal transition of MCF-7 cells.* Braz J Med Biol Res, 2007. **40**(8): p. 1071-8.

69. Cannito S., et al., *Redox mechanisms switch on hypoxia-dependent epithelial-mesenchymal transition in cancer cells.* Carcinogenesis., 2008. **29**(12): p. 2267-78. doi: 10.1093/carcin/bgn216. Epub 2008 Sep 12.
70. Lester RD., et al., *uPAR induces epithelial-mesenchymal transition in hypoxic breast cancer cells.* J Cell Biol., 2007. **178**(3): p. 425-36.
71. Louie, E., et al., *Identification of a stem-like cell population by exposing metastatic breast cancer cell lines to repetitive cycles of hypoxia and reoxygenation.* Breast Cancer Res, 2010. **12**(6): p. R94.
72. Karnoub, A.E., et al., *Mesenchymal stem cells within tumour stroma promote breast cancer metastasis.* Nature, 2007. **449**(7162): p. 557-63.
73. Martin, F.T., et al., *Potential role of mesenchymal stem cells (MSCs) in the breast tumour microenvironment: stimulation of epithelial to mesenchymal transition (EMT).* Breast Cancer Res Treat, 2010. **124**(2): p. 317-26.
74. Liu, S., et al., *Breast cancer stem cells are regulated by mesenchymal stem cells through cytokine networks.* Cancer research, 2011. **71**(2): p. 614-24.
75. Yao, C., et al., *Interleukin-8 modulates growth and invasiveness of estrogen receptor-negative breast cancer cells.* International journal of cancer. Journal international du cancer, 2007. **121**(9): p. 1949-57.
76. Benoy, I.H., et al., *Increased serum interleukin-8 in patients with early and metastatic breast cancer correlates with early dissemination and survival.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2004. **10**(21): p. 7157-62.
77. Burger, J.A. and T.J. Kipps, *CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment.* Blood, 2006. **107**(5): p. 1761-7.
78. Chu, Q.D., et al., *High chemokine receptor CXCR4 level in triple negative breast cancer specimens predicts poor clinical outcome.* The Journal of surgical research, 2010. **159**(2): p. 689-95.
79. Vargo-Gogola, T. and J.M. Rosen, *Modelling breast cancer: one size does not fit all.* Nature reviews. Cancer, 2007. **7**(9): p. 659-72.
80. Lacroix, M. and G. Leclercq, *Relevance of breast cancer cell lines as models for breast tumours: an update.* Breast cancer research and treatment, 2004. **83**(3): p. 249-89.
81. Neve, R.M., et al., *A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes.* Cancer Cell, 2006. **10**(6): p. 515-27.

82. E, K., et al., *Translation of a tumor microenvironment mimicking 3D tumor growth co-culture assay platform to high-content screening.* J Biomol Screen., 2013. **18**(1): p. 54-66. doi: 10.1177/1087057112456874. Epub 2012 Aug 24.
83. Li Q et al., *3D models of epithelial-mesenchymal transition in breast cancer metastasis: high-throughput screening assay development, validation, and pilot screen.* J Biomol Screen., 2011. **16**(2): p. 141-54. doi: 10.1177/1087057110392995.
84. Yoo YA et al., *Sustained co-cultivation with human placenta-derived MSCs enhances ALK5/Smad3 signaling in human breast epithelial cells, leading to EMT and differentiation.* Differentiation., 2009. **77**(5): p. 450-61.
85. Heppner, G.H., F.R. Miller, and P.M. Shekhar, *Nontransgenic models of breast cancer.* Breast cancer research : BCR, 2000. **2**(5): p. 331-4.
86. HM, K., et al., *Using a xenograft model of human breast cancer metastasis to find genes associated with clinically aggressive disease.* Cancer Res., 2005. **65**(13): p. 5578-87.
87. Richmond A. and Su Y., *Mouse xenograft models vs GEM models for human cancer therapeutics.* Dis Model Mech., 2008. **1**(2-3): p. 78-82. doi: 10.1242/dmm.000976.
88. Becher OJ. and Holland EC., *Genetically engineered models have advantages over xenografts for preclinical studies.* Cancer Res, 2006. **2006 Apr 1;66**(7): p. 3355-8.
89. Balkwill, F., K.A. Charles, and A. Mantovani, *Smoldering and polarized inflammation in the initiation and promotion of malignant disease.* Cancer Cell, 2005. **7**(3): p. 211-7.
90. Hovey, R.C., T.B. McFadden, and R.M. Akers, *Regulation of mammary gland growth and morphogenesis by the mammary fat pad: a species comparison.* Journal of mammary gland biology and neoplasia, 1999. **4**(1): p. 53-68.
91. Bos PD et al., *Genes that mediate breast cancer metastasis to the brain.* Nature., 2009. **459**(7249): p. 1005-9. doi: 10.1038/nature08021. Epub 2009 May 6.
92. Minn AJ et al., *Genes that mediate breast cancer metastasis to lung.* Nature., 2005. **436**(7050): p. 518-24.
93. Qi Q et al., *Involvement of matrix metalloproteinase 2 and 9 in gambogic acid induced suppression of MDA-MB-435 human breast carcinoma cell lung metastasis.* J Mol Med (Berl). 2008. **86**(12): p. 1367-77. doi: 10.1007/s00109-008-0398-z. Epub 2008 Sep 6.

94. Harrell JC et al., *Estrogen receptor positive breast cancer metastasis: altered hormonal sensitivity and tumor aggressiveness in lymphatic vessels and lymph nodes.* Cancer Res., 2006. **66**(18): p. 9308-15.
95. Iorns E et al., *A new mouse model for the study of human breast cancer metastasis.* PLoS One, 2012. **7**(10): p. e47995.
96. Weigelt, B., J.L. Peterse, and L.J. van 't Veer, *Breast cancer metastasis: markers and models.* Nature reviews. Cancer, 2005. **5**(8): p. 591-602.
97. Jonkers J. and Derksen PW., *Modeling metastatic breast cancer in mice.* J Mammary Gland Biol Neoplasia., 2007. **12**(2-3): p. 191-203.
98. Landis MD. et al., *Patient-derived breast tumor xenografts facilitating personalized cancer therapy.* Breast Cancer Res., 2013. **15**(1): p. 201. doi: 10.1186/bcr3355.
99. Proia DA. and Kuperwasser C., *Reconstruction of human mammary tissues in a mouse model.* Nat Protoc, 2006. **1**(1): p. 206-14.
100. Wu M. et al., *Dissecting genetic requirements of human breast tumorigenesis in a tissue transgenic model of human breast cancer in mice.* Proc Natl Acad Sci U S A., 2009. **106**(17): p. 7022-7.
101. Rashid OM. et al., *An improved syngeneic orthotopic murine model of human breast cancer progression.* Breast Cancer Res Treat., 2014. **147**(3): p. 501-12.
102. Tao K. et al., *Imagable 4T1 model for the study of late stage breast cancer.* BMC Cancer., 2008. **8:228.**
103. MM, Tomayko and Reynolds CP., *Determination of subcutaneous tumor size in athymic (nude) mice.* Cancer Chemother Pharmacol, 1989. **24**(3): p. 148-54.
104. Euhus DM. et al., *Tumor measurement in the nude mouse.* J Surg Oncol., 1986. **31**(4): p. 229-34.
105. Ayers GD. et al., *Volume of preclinical xenograft tumors is more accurately assessed by ultrasound imaging than manual caliper measurements.* J Ultrasound Med., 2010. **29**(6): p. 891-901.
106. Ishimori T., Tatsumi M., and Wahl RL., *Tumor response assessment is more robust with sequential CT scanning than external caliper measurements.* Acad Radiol., 2005. **12**(6): p. 776-81.
107. Dazai J. et al., *Multiple mouse biological loading and monitoring system for MRI.* Magn Reson Med., 2004. **52**(4): p. 709-15.
108. Isherwood B. et al., *Live cell in vitro and in vivo imaging applications: accelerating drug discovery.* Pharmaceutics., 2011. **3**(2): p. 141-70.

109. Haugwitz M. et al., *Multiplexing bioluminescent and fluorescent reporters to monitor live cells.* Curr Chem Genomics., 2008. **1:11-9.**
110. Ladunga, I., *Computational Biology of Transcription Factor Binding*, I. Ladunga, Editor 2010, Humana Press: New York. p. 1-42.
111. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods., 2001. **25**(4): p. 402-8.
112. Blick T. et al., *Epithelial mesenchymal transition traits in human breast cancer cell lines.* Clin Exp Metastasis, 2008. **25**(6): p. 629-42.
113. Minn AJ.et al., *Distinct organ-specific metastatic potential of individual breast cancer cells and primary tumors.* J Clin Invest., 2005. **115**(1): p. 44-55.
114. Zeisberg, M. and E.G. Neilson, *Biomarkers for epithelial-mesenchymal transitions.* J Clin Invest, 2009. **119**(6): p. 1429-37.
115. Thorne N., Inglese J. and Auld DS., *Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology.* Chem Biol., 2010. **17**(6): p. 646-57.
116. Choy G. et al., *Comparison of noninvasive fluorescent and bioluminescent small animal optical imaging.* Biotechniques, 2003. **2003 Nov;35**(5): p. 1022-6.
117. Negrin, R.S. and C.H. Contag, *In vivo imaging using bioluminescence: a tool for probing graft-versus-host disease.* Nat Rev Immunol., 2006. **6**(6): p. 484-90.
118. Pauwels K. et al., *State-of-the-art lentiviral vectors for research use: risk assessment and biosafety recommendations.* Curr Gene Ther., 2009. **9**(6): p. 459-74.
119. Garret SG. et al., *S100A4, a mediator of metastasis.* J Biol Chem., 2006. **281**(2): p. 677-80.
120. Schneider M., Hansen JL., and Sheikh SP., *S100A4: a common mediator of epithelial-mesenchymal transition, fibrosis and regeneration in diseases?* J Mol Med (Berl). 2008. **86**(5): p. 507-22.
121. Okada H. et al., *Early role of Fsp1 in epithelial-mesenchymal transformation.* Am J Physiol., 1997. **273**(4 Pt 2): p. F563-74.
122. Teng Y., Zeisberg M. and Kalluri R., *Transcriptional regulation of epithelial-mesenchymal transition.* J Clin Invest., 2007. **117**(2): p. 304-6.

123. Robertson H. et al., *Biliary epithelial-mesenchymal transition in posttransplantation recurrence of primary biliary cirrhosis.* Hepatology., 2007. **45**(4): p. 977-81.
124. Boye K. and Maelandsmo GM., *S100A4 and metastasis: a small actor playing many roles.* Am J Pathol., 2010. **176**(2): p. 528-35.
125. Maître JL. and Heisenberg CP., *Three functions of cadherins in cell adhesion.* Curr Biol., 2013. **23**(14): p. R626-33.
126. van Roy F. and Berx G., *The cell-cell adhesion molecule E-cadherin.* Cell Mol Life Sci., 2008. **65**(23): p. 3756-88.
127. Derycke LD. and Bracke ME., *N-cadherin in the spotlight of cell-cell adhesion, differentiation, embryogenesis, invasion and signalling.* Int J Dev Biol, 2004. **48**(5-6): p. 463-76.
128. Lamouille, S., J. Xu, and R. Derynck, *Molecular mechanisms of epithelial-mesenchymal transition.* Nat Rev Mol Cell Biol., 2014. **15**(3): p. 178-96. doi: 10.1038/nrm3758.
129. Ziello JE, Jovin IS. and Huang Y., *Hypoxia-Inducible Factor (HIF)-1 regulatory pathway and its potential for therapeutic intervention in malignancy and ischemia.* Yale J Biol Med., 2007. **80**(2): p. 51-60.
130. Ke Q. and Costa M., *Hypoxia-inducible factor-1 (HIF-1).* Mol Pharmacol., 2006. **70**(5): p. 1469-80.
131. Liao SH. et al., *Proteomics-based identification of two novel direct targets of hypoxia-inducible factor-1 and their potential roles in migration/invasion of cancer cells.* Proteomics., 2009. **9**(15): p. 3901-12.
132. Zhang R. et al., *Subcellular distribution of S100A4 and its transcriptional regulation under hypoxic conditions in gastric cancer cell line BGC823.* Cancer Sci., 2010. **101**(5): p. 1141-6.
133. Yang MH. and Wu KJ., *TWIST activation by hypoxia inducible factor-1 (HIF-1): implications in metastasis and development.* Cell Cycle., 2008. **7**(14): p. 2090-6.
134. Yang MH et al., *Direct regulation of TWIST by HIF-1alpha promotes metastasis.* Nat Cell Biol., 2008. **10**(3): p. 295-305.
135. Zhang L. et al., *Hypoxia induces epithelial-mesenchymal transition via activation of SNAI1 by hypoxia-inducible factor -1alpha in hepatocellular carcinoma.* BMC Cancer., 2013. **13:108.**
136. Moriyama EH. et al., *The influence of hypoxia on bioluminescence in luciferase-transfected gliosarcoma tumor cells in vitro.* Photochem Photobiol Sci., 2008. **7**(6): p. 675-80.

137. Khalil AA. et al., *The Influence of Hypoxia and pH on Bioluminescence Imaging of Luciferase-Transfected Tumor Cells and Xenografts.* Int J Mol Imaging, 2013. **2013:287697.**
138. Teng, Y., M. Zeisberg, and R. Kalluri, *Transcriptional regulation of epithelial-mesenchymal transition.* The Journal of Clinical Investigation, 2007. **117**(304-306).
139. Croker, A.K., et al., *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, 2009. **13**(8B): p. 2236-52.
140. Liu Y. et al., *Lack of correlation of stem cell markers in breast cancer stem cells.* Br J Cancer., 2014. **110**(8): p. 2063-71.
141. Yang, F., et al., *SET8 promotes epithelial-mesenchymal transition and confers TWIST dual transcriptional activities.* EMBO J, 2011. **31**(1): p. 110-23.
142. Pagès G.and Pouysségur J.,*Transcriptional regulation of the Vascular Endothelial Growth Factor gene--a concert of activating factors.* Cardiovasc Res., 2005. **65**(3): p. 564-73.
143. Fang U., Sullivan R. and Graham CH., *Confluence-dependent resistance to doxorubicin in human MDA-MB-231 breast carcinoma cells requires hypoxia-inducible factor-1 activity.* Exp Cell Res., 2007. **313**(5): p. 867-77.
144. Bendinelli P. et al., *NF-kappaB activation, dependent on acetylation/deacetylation, contributes to HIF-1 activity and migration of bone metastatic breast carcinoma cells.* Mol Cancer Res., 2009. **7**(8): p. 1328-41.
145. Lee JW. et al., *Hypoxia-inducible factor (HIF-1)alpha: its protein stability and biological functions.* Exp Mol Med., 2004. **36**(1): p. 1-12.
146. Matteucci E. et al., *Bone metastatic process of breast cancer involves methylation state affecting E-cadherin expression through TAZ and WWOX nuclear effectors.* Eur J Cancer., 2013. **49**(1): p. 231-44.
147. Liu YN. et al., *Regulatory mechanisms controlling human E-cadherin gene expression.* Oncogene., 2005. **24**(56): p. 8277-90.
148. Higashimura Y. et al., *Up-regulation of glyceraldehyde-3-phosphate dehydrogenase gene expression by HIF-1 activity depending on Sp1 in hypoxic breast cancer cells.* Arch Biochem Biophys., 2011. **509**(1): p. 1-8.
149. Mauro L. et al., *Evidences that leptin up-regulates E-cadherin expression in breast cancer: effects on tumor growth and progression.* Cancer Res., 2007. **67**(7): p. 3412-21.

150. Liu J. et al., *S100A4 is upregulated via the binding of c-Myb in methylation-free laryngeal cancer cells.* Oncol Rep., 2014. **31**(1): p. 442-9.
151. Yoshiura K. et al., *Silencing of the E-cadherin invasion-suppressor gene by CpG methylation in human carcinomas.* Proc Natl Acad Sci U S A., 1995. **92**(16): p. 7416-9.
152. Le Mée S., Fromigué O. and Marie PJ., *Sp1/Sp3 and the myeloid zinc finger gene MZF1 regulate the human N-cadherin promoter in osteoblasts.* Exp Cell Res., 2005. **302**(1): p. 129-42.
153. Knapinska, A.M., et al., *Molecular Mechanisms Regulating mRNA Stability: Physiological and Pathological Significance.* Current Genomics, 2005. **6**(6): p. 1-16.
154. Masuda K. et al., *RNA-binding proteins implicated in the hypoxic response.* J Cell Mol Med., 2009. **13**(9A): p. 2759-69.
155. Casas, E., et al., *Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis.* Cancer Res., 2011. **71**(1): p. 245-54.
156. Komar AA. and Hatzoglou M., *Cellular IRES-mediated translation: the war of ITAFs in pathophysiological states.* Cell Cycle., 2011. **10**(2): p. 229-40.
157. Polyak K. and Weinberg RA., *Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.* Nat Rev Cancer., 2009. **9**(4): p. 265-73.
158. Tang Y. et al., *Induction and analysis of epithelial to mesenchymal transition.* J Vis Exp., 2013. **(78).**
159. Scheel C. et al., *Paracrine and autocrine signals induce and maintain mesenchymal and stem cell states in the breast.* Cell., 2011. **145**(6): p. 926-40.
160. Suzuki H. et al., *Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer.* Nat Genet., 2004. **36**(4): p. 417-22.
161. Kuhnert F. et al., *Essential requirement for Wnt signaling in proliferation of adult small intestine and colon revealed by adenoviral expression of Dickkopf-1.* Proc Natl Acad Sci U S A., 2004. **101**(1): p. 266-71.
162. Foroni L. et al., *The role of 3D microenvironmental organization in MCF-7 epithelial-mesenchymal transition after 7 culture days.* Exp Cell Res., 2013. **319**(10): p. 1515-22.
163. Yagi H. and Kitagawa Y., *The role of mesenchymal stem cells in cancer development.* Front Genet., 2013. **4:261.**

164. Sullivan NJ. et al., *Interleukin-6 induces an epithelial-mesenchymal transition phenotype in human breast cancer cells.* Oncogene., 2009. **28**(33): p. 2940-7.
165. Xie G. et al., *IL-6-induced epithelial-mesenchymal transition promotes the generation of breast cancer stem-like cells analogous to mammosphere cultures.* Int J Oncol., 2012. **40**(4): p. 1171-9.
166. Räsänen, K. and M. Herlyn, *Paracrine signaling between carcinoma cells and mesenchymal stem cells generates cancer stem cell niche via epithelial-mesenchymal transition.* Cancer Discov., 2012. **2**(9): p. 775-7. doi: 10.1158/2159-8290.CD-12-0312.
167. Yu Y. et al., *Cancer-associated fibroblasts induce epithelial-mesenchymal transition of breast cancer cells through paracrine TGF-beta signalling.* Br J Cancer., 2014. **110**(3): p. 724-32.
168. Sun X. et al., *IL-6 secreted by cancer-associated fibroblasts induces tamoxifen resistance in luminal breast cancer.* Oncogene., 2014. **33**(35): p. 4450.
169. Hue G. and Hemmings BA., *Phosphorylation of basic helix-loop-helix transcription factor Twist in development and disease.* Biochem Soc Trans., 2012. **40**(1): p. 90-3.
170. Escobar-Cabrera Je. et al., *The Role of Transcription Factor TWIST in Cancer Cells.* Genet Syndr Gene Ther, 2013. **4**(1): p. 1-7.
171. Craene, B.D. and G. Berx, *Regulatory networks defining EMT during cancer initiation and progression.* Nat Rev Cancer., 2013. **13**(2): p. 97-110.
172. Meacham CE. and Morrison SJ., *Tumour heterogeneity and cancer cell plasticity.* Nature., 2013. **501**(7467): p. 328-37.
173. Micalizzi DS., Farabaugh SM. and Ford HL., *Epithelial-mesenchymal transition in cancer: parallels between normal development and tumor progression.* J Mammary Gland Biol Neoplasia., 2010. **15**(2): p. 117-34.
174. Frese KK. and Tuveson DA., *Maximizing mouse cancer models.* Nat Rev Cancer., 2007. **7**(9): p. 645-58.
175. Kim IS. and Baek SH., *Mouse models for breast cancer metastasis.* Biochem Biophys Res Commun., 2010. **394**(3): p. 443-7.
176. Pantelouris EM. and Hair J., *Thymus dysgenesis in nude (nu nu) mice.* J Embryol Exp Morphol., 1970. **24**(3): p. 615-23.
177. Shinkai Y. et al., *RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement.* Cell., 1992. **68**(5): p. 855-67.

178. Puck JM. et al., *Mutation analysis of IL2RG in human X-linked severe combined immunodeficiency.* Blood., 1997. **89**(6): p. 1968-77.
179. Kleinman HK. et al., *Isolation and characterization of type IV procollagen, laminin, and heparan sulfate proteoglycan from the EHS sarcoma.* Biochemistry., 1982. **21**(24): p. 6188-93.
180. Mullen P. et al., *Effect of Matrigel on the tumorigenicity of human breast and ovarian carcinoma cell lines.* Int J Cancer., 1996. **67**(6): p. 816-20.
181. Schuh JC., *Trials, tribulations, and trends in tumor modeling in mice.* Toxicol Pathol., 2004. **32**(Suppl 1): p. 53-66.
182. Ottewell PD., Coleman RE., and Holen I., *From genetic abnormality to metastases: murine models of breast cancer and their use in the development of anticancer therapies.* Breast Cancer Res Treat., 2006. **96**(2): p. 101-13.
183. Gravekamp C. et al., *Behavior of metastatic and nonmetastatic breast tumors in old mice.* Exp Biol Med (Maywood). 2004. **229**(7): p. 665-75.
184. Ljujic B et al., *Human mesenchymal stem cells creating an immunosuppressive environment and promote breast cancer in mice.* Sci Rep, 2013. **3:2298.**
185. Contag CH. et al., *Use of reporter genes for optical measurements of neoplastic disease in vivo.* Neoplasia., 2000. **2**(1-2): p. 41-52.
186. Abd El-Rehim DM et al., *Expression of luminal and basal cytokeratins in human breast carcinoma.* J Pathol., 2004. **203**(2): p. 661-71.
187. Klopp AH. et al., *Concise review: Dissecting a discrepancy in the literature: do mesenchymal stem cells support or suppress tumor growth?* Stem Cells., 2011. **29**(1): p. 11-9.
188. Mandel K. et al., *Mesenchymal stem cells directly interact with breast cancer cells and promote tumor cell growth in vitro and in vivo.* Stem Cells Dev., 2013. **22**(23): p. 3114-27.
189. Qiao L. et al., *Dkk-1 secreted by mesenchymal stem cells inhibits growth of breast cancer cells via depression of Wnt signalling.* Cancer Lett., 2008. **269**(1): p. 67-77.
190. Soon, P.S., et al., *Breast cancer-associated fibroblasts induce epithelial-to-mesenchymal transition in breast cancer cells.* Endocr Relat Cancer, 2013. **20**(1): p. 1-12.
191. Monici M et al., *Cell and tissue autofluorescence research and diagnostic applications.* Biotechnol Annu Rev, 2005. **11**: p. 227-56.

192. Viegas MS. et al., *An improved and cost-effective methodology for the reduction of autofluorescence in direct immunofluorescence studies on formalin-fixed paraffin-embedded tissues.* Eur J Histochem., 2007. **51**(1): p. 59-66.
193. Shaner NC et al., *A guide to choosing fluorescent proteins.* Nat Methods., 2005. **2**(12): p. 905-9.
194. Leblond F. et al., *Pre-clinical whole-body fluorescence imaging: Review of instruments, methods and applications.* J Photochem Photobiol B., 2010. **98**(1): p. 77-94.
195. Rowel L., Dikici E., and Daunert S., *Engineering bioluminescent proteins: expanding their analytical potential.* Anal Chem., 2009. **81**(21): p. 8662-8.
196. Asosingh K. et al., *Role of the hypoxic bone marrow microenvironment in 5T2MM murine myeloma tumor progression.* Haematologica., 2005. **90**(6): p. 810-7.
197. Young RJ., and Möller A., *Immunohistochemical detection of tumour hypoxia.* Methods Mol Biol, 2010. **611:151-9.**
198. Burdall SE. et al., *Breast cancer cell lines: friend or foe?* Breast Cancer Res, 2003. **5**(2): p. 89-95.
199. Johnson Jl. et al., *Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials.* Br J Cancer., 2001. **84**(10): p. 1424-31.
200. Kerbel RS., *Human tumor xenografts as predictive preclinical models for anticancer drug activity in humans: better than commonly perceived-but they can be improved.* Cancer Biol Ther., 2003. **2**(4 Suppl 1): p. S134-9.
201. Dittmer A. et al., *Human mesenchymal stem cells induce E-cadherin degradation in breast carcinoma spheroids by activating ADAM10.* Cell Mol Life Sci., 2009. **66**(18): p. 3053-65.
202. Choi Sy. et al., *Lessons from patient-derived xenografts for better in vitro modeling of human cancer.* Adv Drug Deliv Rev., 2014. **79-80:222-37.**
203. Oltean S. et al., *Alternative inclusion of fibroblast growth factor receptor 2 exon IIIc in Dunning prostate tumors reveals unexpected epithelial mesenchymal plasticity.* Proc Natl Acad Sci U S A., 2006. **103**(38): p. 14116-21.
204. Oltean S., Febbo PG., and Garcia-Blanco MA., *Dunning rat prostate adenocarcinomas and alternative splicing reporters: powerful tools to study epithelial plasticity in prostate tumors in vivo.* Clin Exp Metastasis, 2008. **25**(6): p. 611-9.