Understanding epithelial to mesenchymal transition in human breast cancer

Sylvie Dubois-Marshall

Breakthrough Breast Cancer Research Unit, Western General Hospital, Edinburgh

MD Thesis

The University of Edinburgh

2012

Contents

Contents- Figures and Tables

List of Abbreviations

Acute myeloid leukaemia	AML
Adenocarcinoma	AC
Adenomatous polyposis coli	APC
Adherens junction	AJ
Anaplastic carcinoma	AnCar
Aldehyde dehydrogenase	ALDH
Aromatase inhibitor	AI
Automated quantitative analysis	AQUA
Basal B	BaB
Caspase 8	CASP8
Cancer associated fibroblast	CAF
Cancer stem cell	CSC
Casein kinase 1	CK1
Cyclophosphamide, methotrexate, fluorouracil	CMF
4',6-diamidino-2-phenylindole	DAPI
Ductal carcinoma *in situ*	DCIS
Epidermal growth factor receptor	EGFR
Epithelial repressors	EpR
Epithelial to mesenchymal transition	EMT
Endothelial growth factor	EGF
Embryonic stem cell	ESC
EMT promoting Smad complexes	EPSC
Foetal calf serum	FCS
Fibroblast growth factor receptor 2	FGFR2
Fibrocystic disease	F
Glycogen synthase kinase 3 beta	GSK3β
Haematoxylin and eosin	H&E

High-motility group A protein 2	HMGA2
Human epidermal growth factor receptor 2	HER2
Human mammary epithelial cells	HMECs
Hydrocortisone	H
Immunofluorescence	IF
Immunohistochemistry	IHC
Insulin	I
Invasive ductal carcinoma	IDC
Invasive ductal carcinoma of no special type	IDC-NST
Invasive lobular carcinoma	ILC
Ladybird homeobox1	LBX1
Low-density lipoprotein	LDL
Low-density lipoprotein receptor-related protein 5/6	LRP5/6
Lymph node	LN
Lymphocyte-specific protein 1	LSP1
Mammosphere forming efficiency	MSFE
Mammosphere	MS
Matrix metalloproteinase	MMP
Mesenchymal activators	MeA
Mesenchymal to epithelial transition	MET
Metaplastic breast cancer	MBC
MicroRNA	miRNA
Mitogen-activate protein kinase	MAPK
Normal	N
Not known	NK
Oestrogen receptor	ER
Papillary	Pap
Polymerase chain reaction	PCR

Progesterone receptor	PR
Quantitative real time polymerase chain reaction	qRT-PCR
RNA integrity number	RIN
Significance analysis of micro-arrays	SAM
Sodium hydroxide	NaOH
Somatic (adult) stem cell	SCC
T-cell-specific transcription factor/lymphoid enhancer binding factor	TCF/LEF
Terminal duct lobular unit	TDLU
Three-dimensional	3D
Tight junction	TJ
Tissue micro-array	TMA
Transforming growth factor β	TGF- β
Two-dimensional	2D
Western Blot	WB
Wilms' tumour protein	WT1

Abstract

Background and aims: Increasing evidence suggests that epithelial to mesenchymal transition (EMT) has a key role in breast cancer progression, underlying invasion, metastatic dissemination and acquisition of therapeutic resistance. However, this role is predominantly inferred from *in vitro* and animal studies and controversy regarding EMT in human cancer remains. This thesis has two principle aims. Firstly, to clarify the role of EMT in human breast cancer at the protein level. Secondly, to develop a three-dimensional *in vitro* assay to investigate cell invasion.

Experimental Design: Two independent patient cohorts of high-grade, invasive ductal breast cancer were interrogated for their expression of key EMT proteins using quantitative immunofluorescence. This analysis was extended to paired lymph node metastases for a subset of cases. EMT-related cell lines were selected based on gene and protein expression data. These lines were investigated using light-microscopy, immunohistochemistry and immunofluorescence in a three-dimensional assay that models invasion across the basement membrane.

Results: Two transcriptionally-driven EMT programmes were identified. One comprises vimentin, Snail and Slug and is uncoupled from E-cadherin down-regulation. A second is characterised by up-regulation of WT1, Snail and Slug and down-regulation of E-cadherin. Importantly, acquisition of this phenotype in lymph node metastases predicts poor outcome. Some aspects of these programmes were recapitulated *in vitro*.

Conclusions: These results suggest that EMT does occur in human breast cancer but in a manner distinct to that seen *in vitro*. The examination of primary tumours with their paired lymph node metastases may significantly contribute to understanding EMT. Lastly, *in vitro* models can reflect aspects of tumour biology and may prove invaluable in identifying clinically relevant, targetable pathways.

Chapter 1: Introduction

Clinical challenges in breast cancer

Breast cancer is the leading cause of cancer-related deaths in women world-wide [1]. Over 40,000 new cases are diagnosed and over 12,000 women die each year in England and Wales alone [2]. Furthermore, the incidence of breast cancer in the UK is rising, reflecting changes in population demographics, environmental factors and increased diagnosis as a result of screening [3]. Nonetheless, breast cancer related deaths have fallen by over 25% in the past two decades [4, 5], reflecting significant improvements in management. Breast cancer related deaths are mainly due to the currently incurable nature of metastatic spread of the disease. It is estimated that ~6% of women have metastatic disease at the time of their diagnosis whilst up to 50% will develop metastatic disease with time [6, 7]. The prognosis associated with metastases is poor. Median survival is 18-24 months and 20% remain alive after 5 years [8]. Preceding metastatic spread, the development of treatment resistance or recurrence is an important step in disease progression.

Systemic therapies in breast cancer

Surgery is the mainstay of treatment in early breast cancer across all age groups, in combination with adjuvant radiotherapy for those patients who have had breast-conserving surgery or those at high risk of local recurrence following mastectomy [9]. However, the use of systemic therapies, including endocrine, targeted and chemotherapies, is also critically important and advances in the use of these regimens is thought to be a major contributing factor to the observed decline in breast cancer recurrence and associated mortality [10]. The main aim of systemic adjuvant treatment is to target micrometastatic disease thus reducing recurrence rates and improving long-term overall survival. Some of these therapies are also used in the neoadjuvant setting in order to downstage disease prior to breast conserving surgery. In breast cancer, specific markers are routinely used in combination with information regarding tumour size, grade, lympho-vascular invasion and nodal stage to categorize patients into prognostic groups. Importantly, these markers are also used to predict response to therapy and guide treatment planning [11].

Oestrogen receptor status and endocrine treatments in breast cancer

Oestrogen receptor (ER) status is probably the most powerful predictor examined in breast cancer. The majority of breast cancers, corresponding to ~60% in women under 50 years and ~80% in women over 50 years, are ER positive [12]. Therefore, inhibition of the ER receptor either directly (using weak oestrogen agonists such as tamoxifen) or indirectly (by blocking the conversion of androgens to oestrogen using aromatase inhibitors) is an important treatment option in both early and advanced disease [13, 14].

The role of tamoxifen in reducing tumour recurrence and mortality in ER positive tumours has been well established. Meta-analysis of women with ER positive disease treated with 5 years of tamoxifen has shown that annual recurrence rates are almost halved (2.9% vs. 4.8%) whilst breast cancer mortality rates are reduced by a third (19.3% vs. 27.1%). However, even with 5 years of treatment (which is superior to 1-2 years), the 15-year recurrence and breast cancer mortality rates are still 33.2% and 25.6% respectively [15]. More recently several randomized clinical trials, including the ATAC, BIG 1-98, TEAM, MA-17, NSABP B-33 and ABSCG-6 studies, have examined the role of aromatase inhibitors (AI) in ER positive disease [16, 17]. These studies have examined the effects of either AI versus tamoxifen or various sequential approaches combining both agents. Although the most effective strategy remains to be determined, these studies have found that incorporation of AIs improves disease free survival, especially in the poorer prognostic subgroups. Consequently, postmenopausal women with ER positive disease are now treated with an AI for 5 years whilst patients treated with tamoxifen are either switched to an AI or receive a period of extended treatment with an AI on completion of treatment with tamoxifen. Tamoxifen remains the recommended treatment in pre- and perimenopausal women.

Progesterone receptor status

Progesterone (PR) is an oestrogen-regulated gene and consequently expression is thought to indicate a functioning ER pathway and may assist in predicting response to hormone therapy [11]. It has been shown that PR positive tumours are more likely

to respond to tamoxifen in both early and advanced disease [18] but the significance of PR expression in the absence of ER expression remains controversial [19, 20].

Human epidermal growth factor receptor 2 (HER2) status and targeted therapies in breast cancer

The human epidermal growth factor receptor 2 oncogene (HER2, also known as ERBB2) is a member of the epidermal growth factor receptor (EGFR) family of tyrosine kinases. Amplification and overexpression of the HER2 glycoprotein is seen in ~20-30% of breast cancers [21] and is associated with high tumour grade, lymph node involvement and increased rates of disease recurrence and mortality [22, 23]. HER2 status is predictive of response to targeted therapies, in particular trastuzumab (Herceptin; Genentech, South San Francisco, CA, USA), a monoclonal antibody that targets the HER2 extracellular domain. A key phase-III trial that compared trastuzumab plus chemotherapy versus chemotherapy alone in HER2 positive metastatic breast cancer demonstrated significant improvements in median time to progression (7.4 vs. 4.6 months) and median overall survival (25 vs. 20 months) [24]. More recently, a series of prospective randomized clinical trials have shown that trastuzumab also reduces recurrence and mortality rates in patients with early stage disease [25-27]. HER2 status may also be predictive of response to anthracycline- and taxane-containing chemotherapy regimens [28, 29].

Novel targeted agents for use in HER positive breast cancer are currently being developed with encouraging results [30]. New antibody-based approaches against HER2 include pertuzumab (Omnitarg, 2C4; Genentech), which targets a different epitope on the HER2 extracellular domain [31]. Preclinical studies with pertuzumab alone and in combination with trastuzumab demonstrate antitumour activity with down-regulation of PI3K/Akt and MAPK signalling pathways [32]. In addition, a recent single-arm phase II trial of pertuzumab plus trastuzumab in 66 patients with trastuzumab-refractory HER2 positive metastatic breast cancer demonstrated a complete response in 5 patients (7.6%), partial response in 11 patients (16.7%) and stable disease for at least 6 months in 17 patients (25.8%) [33]. Lapatinib (Tykerb; GlaxoSmithKline, London, UK) is a small molecule inhibitor of HER2 and EGFR/HER1 (an additional member of the EGFR family of tyrosine kinases) and is

currently approved for use in metastatic breast cancer. A key phase-III trial compared lapatinib plus capecitabine versus capecitabine alone in patients with locally advanced or metastatic HER2 positive disease and reported a significant improvement in median time to progression (8.4 vs. 4.4 months) [34]. Updated efficacy analyses have confirmed these findings, although no statistical difference in overall survival is seen [35, 36]. There is also emerging clinical evidence to support the use of trastuzumab in combination with lapatinib in HER2 positive breast cancer at various stages of disease [37, 38]. Other strategies against HER2 are also being developed and include anti-angiogenic therapies, heat shock protein 90 inhibitors, PI3K and mammalian target of rapamycin inhibitors and IGF-1R inhibitors [30].

Chemotherapy in breast cancer

The combination of cyclophosphamide, methotrexate and fluorouracil (or CMF) has been considered standard therapy for early breast cancer since the late 1970s [10]. Subsequently, a number of randomized clinical trials have examined the role of anthracyclines and compared them with CMF regimens. A collaborative meta-analysis of many of these trials has shown that 6 months of anthracycline-based polychemotherapy reduces annual breast cancer mortality by ~38% in women less than 50 years old and by ~20% in women aged 50-69 years, irrespective of ER status or other tumour characteristics. In addition, these regimens were significantly more effective that CMF regimens [15]. The NEAT/SCTBG Br9601 trial is another important study that demonstrated large and significant reductions in breast cancer recurrence (30%) and mortality (36%) when CMF plus epirubicin was compared to CMF alone [39]. As a result of these and other studies [40, 41], anthracycline-based regimens have replaced CMF as the mainstay of chemotherapy in women without pre-existing heart disease.

Chemotherapy is of particular importance in ER negative disease where the option of endocrine therapy is not available. A collaborative meta-analysis of ~6000 patients with ER-poor breast cancer treated with non-taxane-based polychemotherapy versus no chemotherapy reported significant reductions in 10-year recurrence (33% vs. 45% in women less than 50 years; 42% vs. 52% in women aged 50-69 years) and mortality rates (24% vs. 32% in women less than 50 years; 36% vs. 42% in women

aged 50-69 years) [42]. Despite substantial risk reductions, both recurrence and mortality rates remain high in these patients and, as expected, tamoxifen had little effect on recurrence or mortality in these patients.

Taxanes have also emerged as important chemotherapy agents in breast cancer. Randomized trials comparing a taxane-containing regimen with a non-taxane-containing regimen have shown significant improvements in disease free and overall survival [43] and the use of anthracycline-taxane combinations is now an accepted strategy, especially in high-risk patients [44]. Furthermore, concerns regarding cardiotoxicity and an increased incidence of leukaemia have led to the development of regimens featuring a taxane without an anthracycline. Results from the US Oncology 9735 and BCIRG 006 trials suggest equivalent efficiencies in taxane- and anthracycline-comparator arms, with reduced side-effects seen in the taxane-based regimens [45].

Whilst significant advances in chemotherapy treatments have been made, this is a complex area with ongoing controversies. The use of anthracyclines and taxanes in the adjuvant setting is also likely to pose a new challenge as an increasing number of women present with metastatic disease having already been exposed to these agents.

Despite clear advances in the management of breast cancer, several important clinical and scientific challenges remain. In part, these relate to understanding the molecular and cellular basis of cancer progression, in particular the development of resistance and ultimately metastatic spread [7, 46]. However, this is complicated by the fact that breast cancer is not a single disease but is highly heterogeneous at both the molecular and clinical level.

Breast cancer is a heterogeneous disease defined by molecular subtypes

Global gene expression analyses of human breast cancers have identified distinct tumour subgroups [47-49]. Two subtypes are ER negative and include a subgroup defined by increased expression of HER2 and a subgroup with characteristics of basal/myoepithelial cells. These tumours are associated with poor clinical outcomes. A third subtype is ER positive and termed 'luminal' and can be genetically divided into good outcome 'luminal A' tumours and poor outcome 'luminal B'. The fourth

subgroup resembles normal breast. These studies indicated that breast cancer is not a single disease with variable morphological features and biomarkers but rather a group of molecularly distinct neoplastic disorders. It is not surprising that diagnosis based on subtype adds significant prognostic and predictive information to standard parameters for patients with breast cancer [50].

Importantly, comparisons between mouse mammary carcinoma models and human breast tumours has led to the identification of an additional human molecular subtype, termed claudin-low. These cancers are characterised by low expression of genes involved in tight junctions and cell-cell adhesions, including claudins and E-cadherin, and relatively high expression of mesenchymal markers including vimentin. In addition, these are moderate-high grade invasive ductal carcinomas and appear distinct to lobular carcinomas despite their low expression of E-cadherin [51].

Metaplastic breast cancers (MBC) are aggressive, triple-negative (ER, progesterone (PR) and HER2 receptor negative) tumours [52]. These tumours express some markers associated with basal-like cancers and have been proposed to represent a form of basal-like cancer [53]. However, distinct clinical features such as chemoresistance suggest that MBCs may represent a unique subtype [54]. More recently, transcriptional profiling has shown that claudin-low cancers are related to metaplastic breast cancers [55]. Both of these triple-negative subtypes are characterised by low expression of genes responsible for cell-cell adhesion and high expression of stem cell and epithelial to mesenchymal transition (EMT) related genes.

Genetic and epigenetic events drive the initiation and progression of breast cancer

The exact aetiology of breast cancer is unclear. However, a family history of breast cancer remains one of the strongest predictors of risk. Mutations in high-penetrance genes such as *BRCA1*, *BRCA2* and *TP53* are thought to underlie ~25% of inherited susceptibility [56] whilst mutations in moderate- and low-penetrance genes such as *FGFR2* (fibroblast growth factor receptor 2), *CASP8* (caspase 8) and *LSP1* (lymphocyte-specific protein 1) underlie the majority of cases [57-59]. Nonetheless,

the mechanisms by which these genetic abnormalities influence breast cancer development remain to be fully understood.

In addition to genetic events, epigenetic changes are thought to contribute significantly to cancer progression. These can be defined as stable molecular alterations in the cellular phenotype of a cell which are heritable during somatic cell divisions but do not involve changes in DNA sequence. In addition, epigenetic changes occur more frequently than genetic changes, are potentially reversible and occur at defined regions within a gene [60].

DNA methylation is an established epigenetic mechanism and hypermethylation of gene promoter regions has been identified as a frequent mechanism of loss of gene function in cancer [61]. DNA methylation is a post-replication modification that occurs at the pyrimidine ring of cytosines that are located 5' to a guanosine [62]. During evolution, the number of CpG dinucleotides in the genome has been selectively reduced due to the inherent mutagenicity of methylated cytosines [63]. However, CpG dinucleotides occur with a much higher frequency within CpG islands, small stretches of DNA (~500-2000 bp) located in the 5'-regions of most genes [64, 65]. Whilst CpGs throughout the genome are generally methylated, CpGs within CpG islands and in particular those associated with gene promoters are usually unmethylated, allowing transcriptional activity to occur. In human cancers including breast, CpG islands may become methylated resulting in the silencing of gene expression [66]. DNA methylation may be of particular relevance to a claudin-low signature that comprises 9 down-regulated genes, previously known to be de novo methylated in breast and other cancers (see Table 5.1). Figure 1.1 illustrates how accumulated genetic and epigenetic alterations, combined with clonal expansion and selection, result in the initiation and progression of breast cancer. Corresponding histological changes are shown in Figure 1.2.

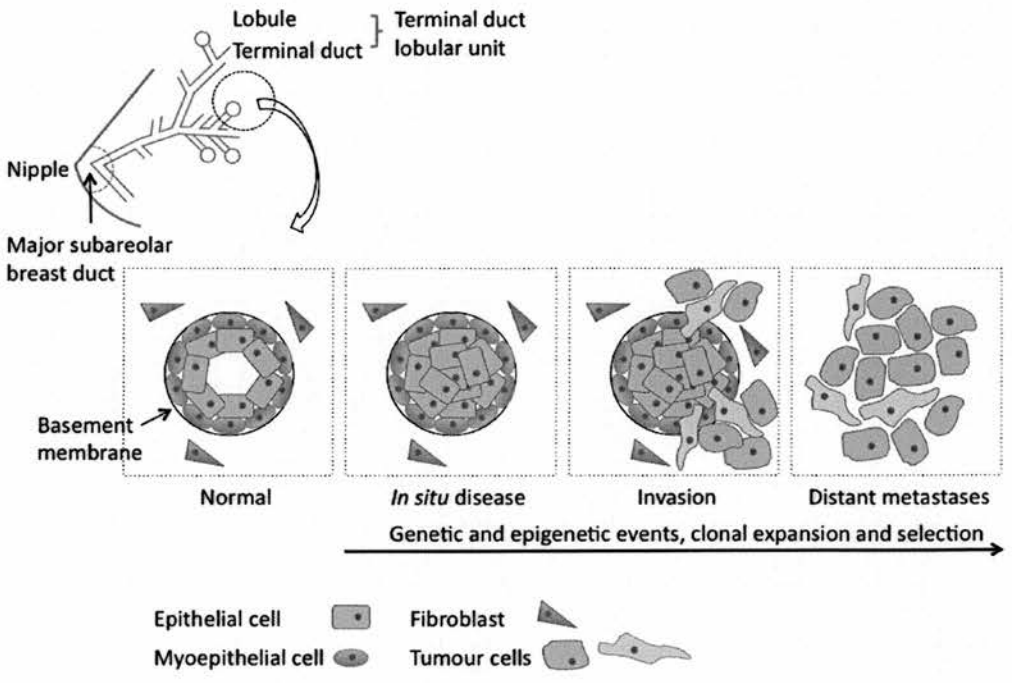


Figure 1.1. Breast cancer initiation and progression. A schematic illustration of normal, *in situ*, invasive and metastatic carcinoma is shown. Normal breast ducts are composed of a layer of luminal epithelial cells, myoepithelial cells and basement membrane. A number of cells exist within the surrounding stroma including fibroblasts, leukocytes, myofibroblasts and endothelial cells. As genetic and epigenetic alterations accumulate, epithelial cells acquire the ability to invade the basement membrane and ultimately invade to distant sites within the body.

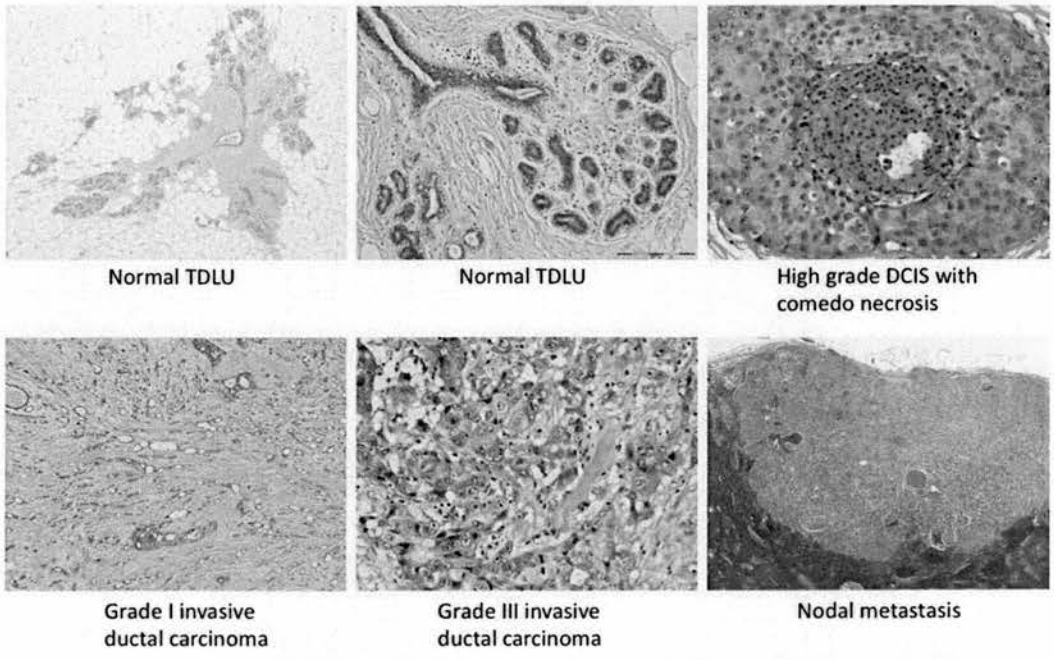


Figure 1.2. Histological changes in cancer progression. The normal architecture of the terminal duct lobular unit (TDLU) is shown. Luminal epithelial cells surround a hollow lumen and in turn are surrounded by a layer of myoepithelial cells that lie in direct contact with the basement membrane. Ductal carcinoma *in situ* (DCIS), where changes are contained by basement membrane, is a pre-invasive lesion with several architectural subtypes. Solid sheets of highly atypical, pleomorphic cells with intraluminal necrosis characterize comedo DCIS. Invasive ductal carcinoma is the commonest type of human breast cancer, accounting for ~70% of cases. The images shown are 4x, 10x and 20x magnification (*top panel, left to right*) and 10x, 20x and 4x magnification (*bottom panel, left to right*). Reproduced with kind permission from Dr. J. Thomas, Consultant Pathologist, Western General Hospital, Edinburgh.

Epithelial to mesenchymal transition, cell-cell contacts and mesenchymal markers

The trans-differentiation of cells from an epithelial to a mesenchymal phenotype, defined as EMT, is essential in embryogenesis and normal development [67]. Increasing evidence supports a role for EMT in the progression of many cancer types including breast, with critical roles in invasion and metastatic dissemination [68, 69] (Figure 1.3). EMT involves loss of cell-cell contacts and re-organisation of the actin cytoskeleton, resulting in loss of apical-basal polarity and acquisition of a spindle-like mesenchymal morphology [70]. EMT is associated with decreased expression of epithelial-specific proteins, including E-cadherin. This is possibly one of the most important consequences of EMT that results in the changed behaviour of tumour cells [71, 72]. Various mechanisms regulate the expression of E-cadherin in breast cancer. Somatic mutations of E-cadherin have been reported, although exclusively in the lobular subtype [73]. Hypermethylation of the CpG islands in the promoter region of *CDH1* has also been reported, occurring in all breast cancer types and associated with poorer clinical outcome [74, 75]. However, the most frequently described mechanism of E-cadherin loss is transcriptional repression [71]. Specific repressors which belong to a family of zinc finger binding proteins have been identified and include Snail, Slug, Zeb1, Zeb2 and Twist [76].

Cell-cell contacts are important in determining the normal structure and function of many epithelial tissues. These consist of three main adhesive structures: tight junctions (TJs), adherens junctions (AJs) and desmosomes [77]. TJs have two principle roles. Firstly, they regulate paracellular permeability, a function which relies on the ability of claudin proteins to form size and charge-selective pores [78]. Secondly, TJs form a physical barrier that prevents the intramembranous movement of lipids and proteins and contributes to the development of apico-basal polarity [79]. Apico-basal polarity underlies terminal differentiation of epithelial structures by orientating the trans-Golgi apparatus. This in turn sorts proteins to either the apical or basolateral membrane, ultimately allowing appropriate electrochemical gradients to be established [80]. E-cadherin is the main transmembrane protein of the AJ and initiates intercellular contacts by binding to cadherins on opposing cells (Figure 3.1) [81]. Cadherins also bind directly and indirectly to many cytoplasmic proteins, in

particular to members of the catenin family. The catenins in turn regulate organisation of the actin cytoskeleton, cadherin stability and intracellular signalling pathways that control gene transcription [82]. The functions and protein components of TJs and AJs make these particularly relevant to EMT and dysregulated cell-cell contact are increasingly implicated in the progression of various cancer types [83, 84].

In many epithelial cancers, loss of E-cadherin is accompanied by increased expression of N-cadherin and other mesenchymal cadherins, and is referred to as a 'cadherin switch' [85, 86]. This is thought to result in a significant change in the adhesive properties of cancer cells, so that they lose their affinity for epithelial neighbours and gain affinity for stromal cells [87]. In addition, the cadherin switch is associated with increased cell motility, matrix metalloproteinase (MMP) secretion, invasiveness, and poor prognosis [88, 89]. This is thought to result, at least in part, from the interaction of N-cadherin with fibroblast growth factor receptors, which results in sustained mitogen-activated protein kinase (MAPK) pathway activity [85, 90]. Importantly, transcriptional repressors of E-cadherin, including Snail, Slug and Zeb2 can also induce N-cadherin, suggesting that the cadherin switch is part of a transcriptionally driven reprogramming of cancer cells [91]. Vimentin and the extracellular matrix component fibronectin are two other mesenchymal proteins commonly expressed by cancer cells following EMT. Expression of vimentin is associated with changes in cell shape and cytoskeletal arrangement as well as changes in motility and adhesion [92-94]. The changes in extracellular matrix function and composition associated with cancer development are in part due to increased production of fibronectin [95], by both stromal and tumour cells. This is thought to generate a microenvironment conducive to tumour cell migration [96].

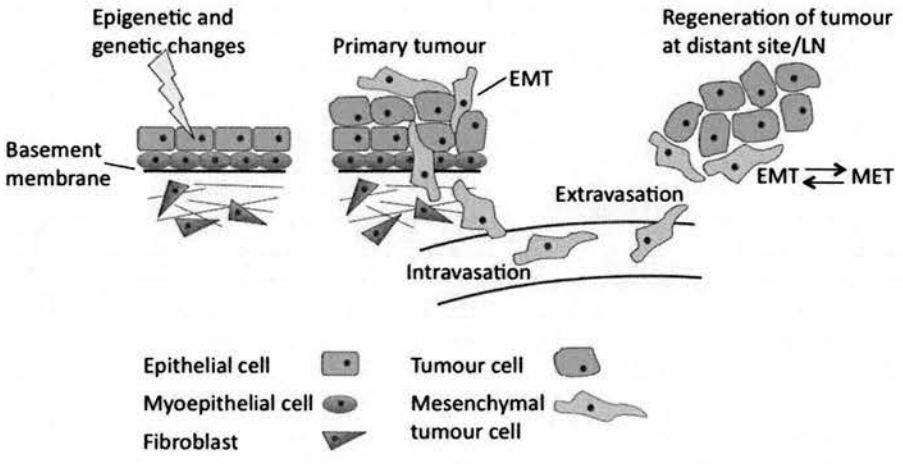


Figure 1.3. EMT is a dynamic process that reflects epithelial plasticity. Cells undergoing EMT during tumour progression are characterised by loss of cell-cell adhesion and polarity, cytoskeletal rearrangements, increased motility and increased invasive capacity. This is accompanied by a change in morphology. Cells that undergo EMT may revert to an epithelial phenotype by mesenchymal to epithelial transition (MET), and stable MET may occur at established distant metastases. Adapted from [76].

Important signalling pathways in EMT

Transforming growth factor-β (TGF-β) is an important inducer of EMT during development [97] and is also over expressed in many cancers where it is thought to drive EMT [98, 99]. However, tumour suppressor roles have also been described such that the role of TGF-β in cancer and its ability to induce EMT appears to be context dependent [100, 101]. TGF-β is known to co-operate with a number of signalling pathways including Wnt [102, 103], Ras/receptor tyrosine kinases [104, 105], Hedgehog [106] and Notch [107] to induce a complete EMT. These pathways also have roles in the renewal and proliferation of stem cells [108, 109]. Consequently, it is hypothesized that co-activation of these pathways may modify the cellular response to TGF-β, either promoting or inhibiting EMT [110]. This hypothesis is also in keeping with the emerging links between EMT and the acquisition of stem cell characteristics [111].

Binding of TGF-β to its receptors results in the formation of a heterotetrameric signalling complex involving type I and type II receptors, resulting in the activation of Smads- intracellular transcription factors and transducers of TGF-β signalling [112]. The receptor-activated Smad2 and Smad3 proteins associate with cytoplasmic Smad4 and translocate to the nucleus to regulate gene transcription [100]. However, Smads have a low affinity for DNA and interact with transcriptional cofactors to improve their affinity for target genes. In addition, the availability and activity of cofactors in the nucleus determines the cellular response to TGF-β and whether target genes are activated or repressed [100].

Recent studies have shown that many EMT promoting transcriptional factors including Snail, Zeb1/2, Twist and β-catentin can act as Smad cofactors [113-116]. Interaction between these transcription factors and Smads results in the formation of EMT promoting Smad complexes (EPSC) which drive EMT by repressing epithelial genes including E-cadherin and up-regulating mesenchymal genes including vimentin [110]. This is an important finding as EMT promoting transcriptional factors are of significant and current interest as drivers of EMT (discussed further in chapter 3) and the formation of EPSC represents a point of convergence between a number of signalling pathways (Figure 1.4). In addition, Wnt, Ras and TGF-β

signalling pathways can cooperate to activate and stabilize Snail, Zeb1/2, β-catentin and other EMT promoting transcriptional factors. For example, the up-regulation of Snail by TGF-β is mediated by transcriptional cooperation between Smads and the high-mobility group A protein 2 (HMGA2) [117], itself a transcriptional target of Ras signalling [118]. Growth factor mediated activation of Ras also cooperates with TGF-β to regulate Snail expression [119]. Snail is also directly targeted by Wnt signalling, and indirectly, through glycogen synthase kinase-3beta (GSK-3β) mediated phosphorylation and degradation [120, 121].

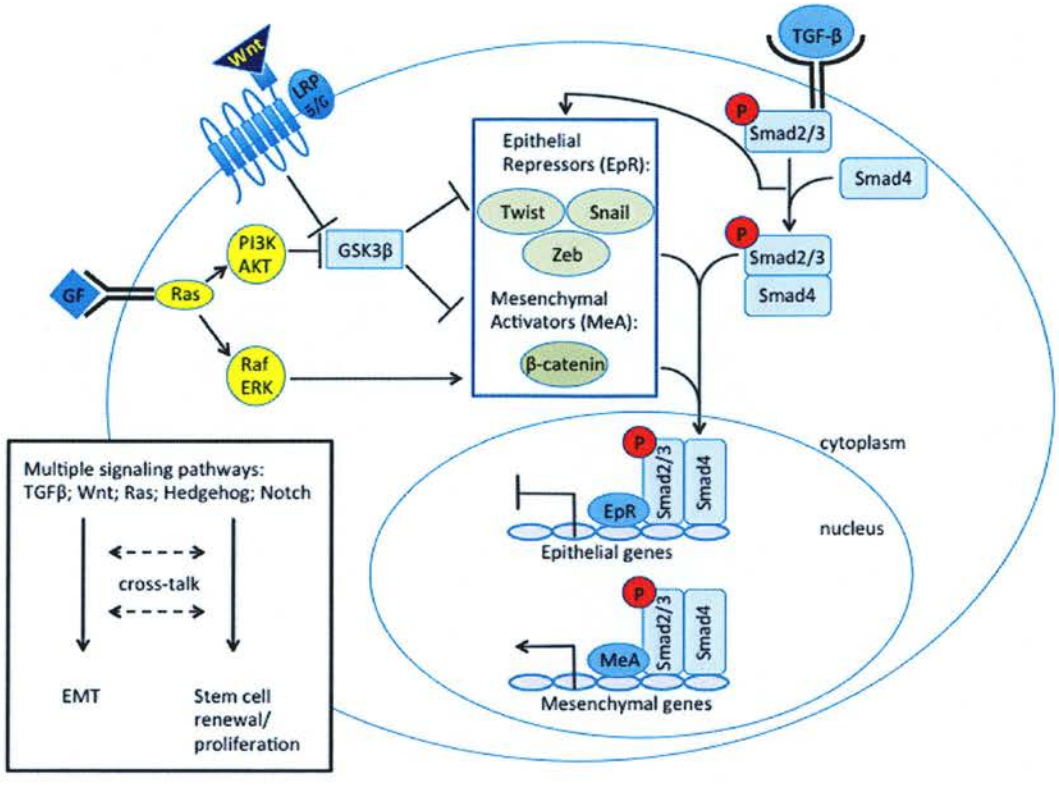


Figure 1.4. Transcriptional crosstalk between key signalling pathways in EMT. The binding of TGF-β to its receptor results in the phosphorylation and nuclear translocation of Smad proteins where they interact with cofactors to regulate gene transcription. Cofactors include EMT promoting transcription factors (including epithelial repressors (EpR) such as Snail, Zeb and Twist, and mesenchymal activators (MeA) such as β-catenin). Interaction between Smads and these cofactors results in the formation of EMT promoting Smad complexes (EPSC) which drive EMT. The Wnt, Ras and TGF-β signalling pathways also cooperate to activate EMT promoting transcription factors such that the formation of EPSC represents an important point of convergence between them. The availability of Smad cofactors, as influenced by the above pathways, may determine whether EPSC are formed and whether or not EMT results. GSK-3β is an important nodal protein whose activity is regulated by Wnt and Ras pathways and which in turn negatively regulates Snail and β-catentin stability. Adapted from [110].

Evidence for epithelial to mesenchymal transition in human breast cancer

Until recently, insufficient evidence for EMT in clinical samples has contributed to the ongoing controversy regarding the relevance of EMT in human cancer [122-124]. Furthermore, the role of EMT in human cancer has predominantly been inferred from *in vitro* studies using cell type-specific markers [125-127]. This makes tracing the actual origin of tumour associated mesenchymal cells difficult as, having undergone EMT, tumour epithelial cells will be phenotypically similar to normal stromal cells.

The description of small aggregates of tumour cells detaching from the invasive front of colorectal carcinomas has provided morphological evidence for the existence of EMT in human cancer [128]. Importantly, these morphological changes have been associated with loss of E-cadherin expression and deregulated Wnt signalling [129]. In breast, a recent immunohistochemical analysis of invasive ductal carcinomas reported significant associations between nuclear expression of Slug and loss of membranous E-cadherin protein expression [130]. In a second immunohistochemical study comprising 479 invasive human breast carcinomas, unsupervised hierarchical clustering identified up-regulation of additional EMT markers in a subset of cancers. Similarly, this correlated with loss of E-cadherin protein expression [131]. However, whilst these studies examined clinical specimens for evidence of EMT, neither formally investigated the relationship between morphology and the observed changes in protein expression.

In order to directly visualize EMT in cancer progression *in vivo*, stromal- and epithelial-specific cre-transgenic mice have been developed [132]. This genetic system allows epithelial and stromal cells to be tracked independently, without relying on cell type-specific markers. Three oncogene-driven (*myc, neu* and *PyMT*) breast cancer models, reflecting different molecular and cellular aspects of breast cancer, were selected. Investigation of these models using stromal- and epithelial-specific cre-transgenic mice showed that EMT was common in *myc*-induced tumours but rare in *neu*- and *PyMT*-initiated tumours. Interestingly, EMT was not a pre-requisite for invasiveness and metastases in this model as mice with *neu*- and *PyMT*-induced cancers had significant amounts of lung metastases. Conservation of gene

expression between tumour epithelial and stromal cells has been used as a marker of the likelihood of a common progenitor between these populations. Using this approach, Trimboli and colleagues subsequently showed that the incidence of EMT in invasive human breast cancers is rare (although when it occurs it is associated with amplification of *MYC*) [132]. A similar study that evaluated nuclear polymorphisms between tumour epithelial and stromal populations also suggested that mesenchymal cells within the stroma infrequently originate from an epithelial lineage [133]. Taken together, these studies suggest that EMT is a relevant phenomenon in at least a proportion of human cancers.

EMT and chemoresistance

There is increasing evidence in a variety of cancer types to suggest a molecular and phenotypic link between EMT and chemoresistance. Oxaliplatin resistance in colorectal cancer cell lines was associated with spindle morphology and expression of EMT markers [134]. Similar changes were seen in paclitaxel resistant ovarian carcinoma cells [135]. Another study found that up-regulation of the transcription factor Twist in breast cancer cell lines is associated with EMT and resistance to paclitaxel [136]. A second study in breast cancer cell lines found that multidrug resistance following adriamycin-induced EMT was partially reversed by depletion of Twist. In addition, Twist RNA interference improved the efficiency of adriamycin treatment in relation to tumour volume, development of metastases and survival in a mouse model of breast cancer [137]. Similarly, a study in resistant pancreatic cancer cell lines has shown that silencing of an additional EMT-related transcriptional repressor, Zeb1, increases the expression of epithelial markers and restores drug sensitivity [138]. These results support the need to target specific cancer subpopulations including those undergoing EMT in the prevention of recurrence.

Cancer stem cells in breast cancer

The role of stem cells in embryogenesis and in the renewal and maintenance of adult tissues has led to the exciting concept that similar 'cancer' stem cells (CSC) might play an important role in cancer development and progression [139].

General aspects of embryonic and somatic stem cells

As the concept of CSCs is derived from the general model of stem cells, embryonic (ESC) and adult or somatic stem cells (SCC) will be briefly considered first. Regardless of their source, stem cells are characterised by general properties- i) self-renewal; ii) asymmetric division to produce either a copy of themselves or a hierarchy of progressively more differentiated daughter cells and iii) homeostatic control [140]. However, whilst ESC can generate all three germ layers and ultimately all differentiated cells in the body, SCC have a more restricted potential and typically generate cell types of the tissue in which they reside [141]. In addition, ESC and SCC utilize different signalling pathways for their maintenance [142].

Cancer stem cells- identification and uncertainties

The CSC hypothesis proposes that there is a small subpopulation of cells within a tumour capable of initiating and sustaining tumour growth. These cells have been interchangeably described as CSCs or 'tumourigenic' or 'tumour initiating cells' [143]. These cells may play a key role in resistance to treatment in breast cancer (see subsequent section). Whilst increasingly accepted it is important to note that the CSC hypothesis is an area of ongoing controversy in terms of how to define CSCs, their origin (including their relation to SCCs) and even in terms of the validity of the CSCs hypothesis [144, 145].

CSCs have been isolated based on the presence of cell surface markers (using flow cytometry), with further characterization using sphere culture assays and transplantation into immune-compromised mice [146]. Sphere culture assays make use of an *in vitro* culture system where single cells are grown in suspension. Some stem or progenitor cells cultured in this way will form spherical colonies termed spheroids (or mammospheres in the context of breast), the self-renewal capacity of which is then tested by evaluating the ability of spheroid-derived cells to form new spheres containing multipotent cells [147]. The first convincing evidence for the existence of CSCs was found in acute myloid leukaemia (AML) [148]. In this study a rare subpopulation (0.2-100 cells per 10^6) of cells expressing the cell surface markers $CD34^+CD38^-$ were isolated. 5000 of these cells were sufficient to induce

leukaemic transformation in immune-compromised mice whilst much greater numbers of $CD34^{+}CD38^{+}$ cells could not. Subsequently, Al-Hajj and colleagues demonstrated evidence for the existence of CSCs in breast cancer [149]. In this study, a subpopulation of highly tumourigenic $CD44^{+}CD24^{-/low}$ cells was identified from metastatic pleural effusions associated with breast cancer. These cells formed multi-lineage mammospheres *in vitro* and as few as 100 $CD44^{+}CD24^{-/low}$ cells were sufficient to induce tumour formation in immune-compromised mice whilst again much larger numbers of non-CSCs could not. More recently, other groups have identified putative CSC populations in a variety of solid cancers including brain [150], colon [151], hepatocellular [152] and prostate cancer [153]. CSCs and SCCs both demonstrate shared characteristics (i.e. a capacity for self-renewal and the generation of heterogeneous progeny) whilst others including asymmetry and rates of cell division, the necessity for a specific microenvironment to maintain an undifferentiated state [140] and homeostatic control are not so clearly shared [154]. Whilst distinct, the similarity of roles and the similarity of surface marker expression between some CSCs (for example leukaemia and brain CSCs [155]) and their normal counterparts supports the hypothesis that CSCs can originate form SCCs [156, 157]. There is also evidence for shared signalling pathways between some SCCs and CSSs [158, 159]. The origin of CSCs remains an area of debate and is discussed in further detail subsequently.

The role of cell surface markers in the identification of CSCs is well established. Several biomarkers have been identified, some of which are common to different cancer types [155]. If these biomarkers characterize a unique and homogenous population, then this population should demonstrate all the characteristics of CSCs without further refinement. However, this is generally not the case. For example, in pancreatic cancer, distinct CSC populations appear to confer tumour growth and metastatic activity [160]. In addition, the expression of biomarkers is often much more prevalent than would be expected given that they are postulated to represent a 'rare' subpopulation. For example, the $CD44^{+}CD24^{-/low}$ population identified by Al-Hajj and colleagues represents 11-35% of the total population examined [149]. Furthermore, subsequent examination of these cells suggests that only 10-20% have self-renewal capacity [161]. Whilst improved biomarker specificity is required it

seems unlikely that CSC populations represent a single, homogenous population. In support of this, single CSCs have not been shown capable of tumour initiation whilst a single stem cell can repopulate the normal mammary gland in mice [162]. The difficulties in identifying and defining CSC populations must be considered when interpreting data from these studies.

Cancer stem cells underlie intrinsic resistance

Patients may relapse with time for various reasons. Firstly, multiple subpopulations of cancer cells within a tumour may gradually acquire resistance. In this situation, the relative proportions of subpopulations of cancer cells in residual tumours should be unchanged before and after treatment. Alternatively, a specific subpopulation may be intrinsically resistant to treatment, in which case the relative proportion of this subpopulation should increase after treatment (Figure 1.4). In support of the second hypothesis, it has been shown that the gene expression pattern of residual tumour cells after docetaxel is different to that of the initial tumour [163]. These findings provide the beginning of an explanation for the relative failures of conventional chemotherapies and even the observed enhancement of tumour progression seen in some settings [164].

Gene expression signatures based on $CD44^{+}CD24^{-/low}$ surface marker expression have already been used as independent prognostic predictors in breast and other cancers [165]. More recently, the role of $CD44^{+}CD24^{-/low}$ cells in human breast cancer resistance has been examined [166]. Neoadjuvant chemotherapy in HER2-negative patients significantly increased the proportion of $CD44^{+}CD24^{-/low}$ cells and was associated with increased mammosphere forming efficiency (MSFE) *in vitro*. In addition, cells from residual tumours were associated with increased tumour outgrowth when implanted into immunocompromised mice. These data provide clinical evidence for the existence of a subpopulation of intrinsically resistant CSCs in breast cancer. In the same study, lapatinib treatment of patients with HER2-positive tumours was associated with a decrease in the proportion of $CD44^{+}CD24^{-/low}$ cells and their MSFE. Lapatinib is an EGFR (epidermal growth factor receptor)/HER2 tyrosine kinase inhibitor and EGFR signalling has been shown to play a role in self-renewal [167]. These data may partially explain the significant

survival benefit seen when trastuzumab (which also targets EGFR/HER2 signalling) is given in conjunction with chemotherapy in comparison with chemotherapy alone [168]. In addition, lapatinib given in combination with chemotherapy has been shown to improve progression-free survival in women with HER2 positive metastatic disease [34]. These data suggest that the use of specific pathway inhibitors in combination with conventional therapies may target CSC populations and reduce recurrence of disease.

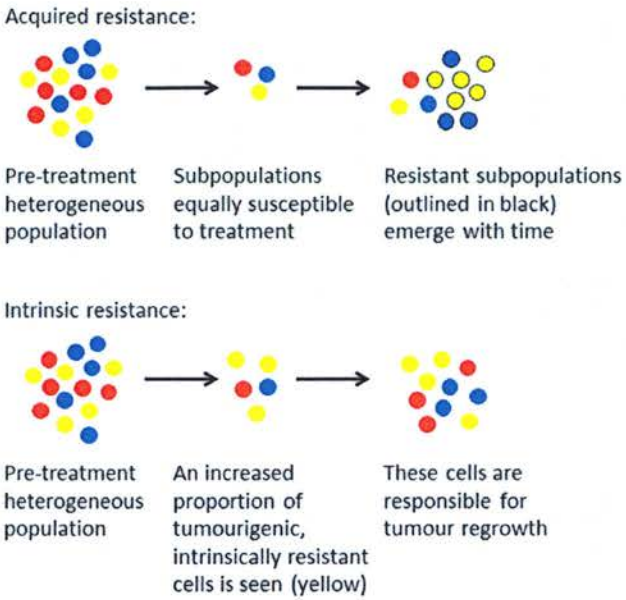


Figure 1.5. Acquired versus intrinsic resistance. There are two broad, although not exclusive, explanations for treatment resistance and clinical recurrence. Subpopulations of cancer cells may be equally susceptible to treatment but continued exposure results in the acquisition of resistance by some populations (*top panel*). Alternatively, a relatively rare, tumourigenic subpopulation may be intrinsically resistant to treatment. In this case, the relative proportions of cells in residual tumours with tumourigenic properties would be expected to increase with treatment (*bottom panel*).

Further attempts to understand the regulatory pathways underlying subpopulations of residual tumour cells after conventional treatments have been made [169]. A gene signature common to both $CD44^{+}CD24^{-/low}$ and mammosphere-forming (MS) cells was found to be predominantly expressed in tumours of the recently identified claudin-low subtype. In addition, both this signature and a previously defined claudin-low signature [51] were up-regulated in residual tumours after treatment with endocrine- (letrozole) or chemo-therapy (docetaxel). Importantly, the claudin-low subtype is characterised by the expression of many EMT associated genes. The increased expression of EMT markers in post-treatment tumours, including vimentin, Snail and MMP2, was confirmed. EMT has been increasingly implicated in cancer progression (discussed in detail below) and may contribute to treatment resistance. These data suggest that subpopulations of residual tumour cells after conventional treatments display not only CSC features but also evidence of EMT. Therefore, targeting EMT related pathways may provide an additional therapeutic strategy against recurrence. The full clinical significance of this work remains to be seen as there is no report on the long-term outcome of those patients with up-regulated expression of both the $CD44^{+}CD24^{-/low}$-MS-forming and claudin-low signatures. In addition, correlations between the claudin-low subtype and outcome have yet to be described by other groups. However, other studies support the link between EMT and CSCs. In one study, circulating tumour cells were isolated from breast cancer patients treated with chemotherapy and classified as responders versus non-responders. The non-responders were found to express both EMT and CSC signatures [170]. In addition, claudin-low tumours are related at the gene expression level to MBCs, which are associated with chemoresistance and poor outcome [55]. The association between EMT and chemoresistance has been briefly discussed. Taking these associations further, a study in ovarian cancer cells has shown that induction of EMT mediates radio- and chemo-resistance through the induction of stem cell promoting genes [171].

The clinical relevance of these studies is accumulating and the use of multiplex assays that identify markers of 'stemness' and EMT with prognostic and predictive value is likely to become increasingly apparent. In a recent study, co-expression of the stem cell markers ALDH1 and CD44 was found to significantly correlate with

poor clinical outcome in human breast cancer, independently of other established prognostic markers [172].

EMT and CSCs: two closely linked phenomena

In addition to the emerging link between EMT and chemoresistance, recent evidence suggests a direct link between EMT and the acquisition of stem cell-like properties. Mani and colleagues clearly demonstrated this in their series of recent experiments where they showed that immortalized human mammary epithelial cells (HMECs) that have undergone EMT also express stem-cell markers and increased MSFE [111]. In addition, stem cell-like cells isolated from human reduction mammoplasties and breast carcinomas were shown to express markers of EMT whilst transformed HMECs that had undergone EMT formed both mammospheres and tumours with greater efficiency. A subsequent study aimed to determine specifically the origin of tumourigenic $CD44^{+}CD24^{-/low}$ cells [173]. This study showed that $CD44^{+}CD24^{-/low}$ cells can originate from primary $CD44^{low}CD24^{+}$ HMECs following their neoplastic transformation. Similar findings were seen with the untransformed breast epithelial cell line MCF10A. Importantly, in both HMECs and MCF10A cells, acquisition of the $CD24^{-}$ phenotype was associated with a mesenchymal morphology and expression of EMT markers. This study supports the findings of Mani and colleagues and shows that CSCs can originate in the absence of normal stem cells. An additional study provides further evidence that EMT can generate CSCs, and found that these cells demonstrate enhanced resistance to drugs and radiation [174].

A recent study uses normal human mammary epithelial hierarchy as a framework for understanding the cellular origins of the molecular subtypes of breast cancer [175]. Gene signatures characterising the hierarchal subpopulations in normal breast were compared to those characterising the molecular subtypes of breast cancer. Interestingly, a 'stem cell signature' was enriched in the EMT-related claudin-low subtype, further supporting a link between stemness and EMT. Clearly, therapeutic strategies that target CSCs are required. In a key study, Gupta and colleagues exploit the ability of EMT to enrich populations for CSCs in order to implement high-throughput screening for selective CSC inhibitors [72]. They demonstrate for the first time that CSCs exhibiting features of EMT can be selectively targeted.

miRNA expression profiles suggest that CSCs can originate from their normal counterparts

An area of ongoing controversy surrounding CSCs regards their origin [176]. Mani and colleagues have shown a clear link between EMT and the generation of stem cell-like cells and Morel and colleagues have shown that these cells can originate from transformed, differentiated cells of $CD44^{low}CD24^{+}$ lineage [111, 173]. This implies that there is no prerequisite for transformation of normal stem cells. Nonetheless, normal stem cells have been implicated in the origin of CSCs. This is based on the following observations. Firstly, CSCs and normal stem cells share many properties. Secondly, the long lifespan and multiple mitoses undergone by normal stem cells would appear to make them potential targets for malignant transformation [177].

Comparison of microRNA (miRNA) expression in breast stem cell populations has led to the identification of shared regulators between normal and CSCs [178]. miRNAs are small noncoding RNAs that regulate the translation of messenger RNA (mRNA), either inhibiting or degrading the target mRNA [179]. In addition, miRNAs are critical regulators of self-renewal and differentiation [180, 181] and expression profiles correlate with clinical outcomes including tumour stage and prognosis [182]. Shimono and colleagues have identified three miRNA clusters, miR-200c-141, miR-200b-200a-429 and miR-183-96-182, which are down-regulated in both human breast CSCs and normal stem cells. Importantly, expression of miR-200c inhibited the formation of differentiated mammary structures by normal stem cells and tumour formation by CSCs *in vivo* [178]. This identification of molecular and functional links between normal and CSCs may be interpreted as evidence that CSCs can originate from their normal counterparts.

miRNA expression profiles strengthen the emerging links between EMT, CSCs and cancer progression

The findings of Shimono and colleagues [178] strengthen the emerging link between EMT and CSCs as the miR-200 family has recently been shown to prevent EMT by suppressing the expression of Zeb1 and Zeb2, two transcriptional repressors of E-

cadherin [183, 184]. In addition, re-expression of miR200c in ovarian cancer cell lines has been shown to restore E-cadherin expression and sensitivity to microtubule-targeting chemotherapy [185].

Other miRNAs with links to EMT pathways have been identified as promoters of metastases in breast cancer. miR-9 is up-regulated in human breast cancer and can directly target *CDH1*. In addition, overexpression of miR-9 promotes the formation of pulmonary metastases in mice and correlates with grade and metastatic status in human cancers [186]. Similarly, miR-10b, whose expression is induced by Twist, is highly expressed in metastatic human breast cancer [187]. Silencing miR-10b using specific antagomirs (modified anti-miRNA oligonucleotides) significantly suppresses the formation of metastases in a mouse model and demonstrates an additional potential therapeutic strategy for targeting metastases [188].

Aims

Increasingly, links are being established between EMT, CSCs and important clinical events, namely recurrence of disease and metastatic spread (Figure 1.5). In addition, therapeutic strategies that target EMT pathways and CSCs are beginning to emerge and precise, multiplex assays that identify markers related to these biological events with correlations to clinical outcomes are envisaged for the future. However, if EMT and its related processes are to be successfully targeted, their roles as well as their underlying mechanisms in human cancer need to be fully understood.

Therefore, the aims of this thesis are:

- To clarify the role of EMT in human breast cancer at the protein level. This will involve an analysis of the expression of key protein markers of EMT in primary tumours (Chapter 3), and, for a subset of these, in paired LN metastases (Chapter 4).
- To develop an *in vitro* assay for the investigation of a series of EMT-related breast cancer cell lines, with a particular focus on the mechanisms underlying invasion and EMT (Chapter 5).

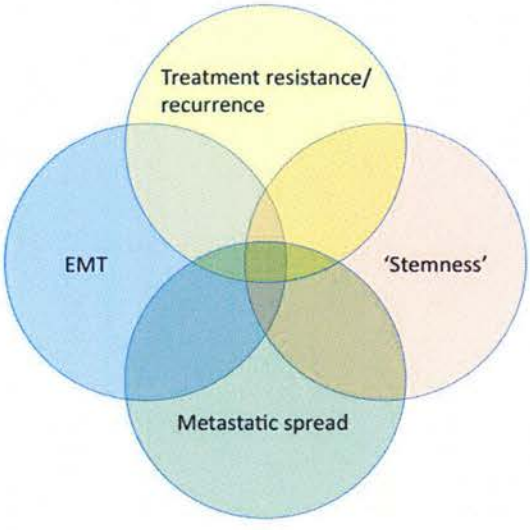


Figure 1.6. EMT and 'stemness' are two important biological processes that closely underlie clinical resistance and metastatic spread. The exact links between these processes remain to be fully clarified.

Chapter 2: Materials and Methods

Patient cohort selection for tissue microarray

Invasive ductal carcinomas of no special type (IDC-NST) were examined as these are the commonest form of breast cancer. The first patient cohort (set 1) was enriched for HER2 positive invasive breast cancers [189], therefore minimising the number of invasive lobular carcinomas (ILC) [190]. This is important as down-regulation of E-cadherin is much more common in ILC than IDC-NST [191-193]. In addition, hierarchical cluster analysis has clearly demonstrated that ILC and IDC-NST are molecularly distinct [194] and therefore down-regulation of E-cadherin is likely to occur through distinct mechanisms. Set 1 consisted of 122 patients with primary breast cancer who were subsequently treated with trastuzumab in the Edinburgh Breast Unit as previously described [189].

The second patient cohort (set 2) was enriched for large, high-grade cancers that had already given rise to lymph node (LN) metastases. The presence of these poor prognostic features [195] maximises the identification of EMT events as EMT is believed to contribute to invasion and metastatic dissemination [68]. This study population was derived from an original population of 521 patients with primary breast carcinomas treated in the Edinburgh Breast Unit from 1999 to 2002, previously described [196]. All received axillary LN dissections as part of surgery for large or high-grade invasive breast carcinomas, in the absence of known distant metastases. Out of these 521 cases, 156 had paired LN metastases and were selected for tissue micro-array (TMA) construction. Exclusion of 13 cases with insufficient tumour left 143 primary carcinomas with paired LNs for inclusion in the TMA.

Table 2.1. Available patient characteristics for set 1. Median follow-up was 21.6 months. The Nottingham Prognostic Index (NPI) is based on tumour size, LN status and histological grade. 3 prognostic categories may be derived accordingly: NPI $\leq$ 3.4 = low risk; NPI 3.4 – 5.4 = medium risk; NPI > 5.4 = high risk [197].

	Patients, *n* (%) *N*=122
Age (y)	
<50	49 (40.2)
>50	73 (59.8)
NK	0 (0.0)
Nottingham Prognostic Index	
<3.4	2 (1.6)
3.4-5.4	47 (38.5)
>5.4	62 (50.8)
NK	11 (9.0)
Tumour grade	
1	1 (0.8)
2	19 (15.6)
3	99 (81.1)
NK	3 (2.5)
Tumour stage	
I	35 (28.7)
II	64 (52.5)
III	12 (9.8)
IV	3 (2.5)
NK	8 (6.6)
Node stage at diagnosis	
Negative (N0)	26 (21.3)
Positive (N1)	87 (71.3)
NK	9 (7.4)
ER status	
>3	72 (59.0)
≤3	41 (33.6)
NK	9 (7.3)
HER2 status	
Positive	90 (73.7)
Negative	32 (26.3)
NK	0 (0.0)
Chemotherapy	
Anthracycline containing	66 (54.1)
Taxane containing	53 (43.4)
NK	3 (2.5)

Table 2.2. Available patient characteristics for set 2. Median follow-up was 90 months.

	Patients, *n* (%) ***N*=143**
Tumour grade	
1	4 (2.8)
2	72 (50.3)
3	67 (46.9)
NK	0 (0)
Tumour size (mm)	
≤ 20 (T1)	49 (34.3)
21-50 (T2)	84 (58.7)
> 50 (T3)	9 (6.3)
NK	1 (0.7)
Node stage at diagnosis	
Negative	1 (0.7)
Positive	142 (99.3)
NK	0 (0)
HER2 status	
Negative	109 (76.2)
Positive	18 (12.6)
NK	16 (11.2)
ER status	
≥ 3	90 (62.9)
<3	53 (37.1)
NK	0 (0)

Breast cancer tissue microarray construction

The characteristics of the patient cohorts used to construct the two TMA sets in this study are summarised in Tables 2.1 (set 1) and 2.2 (set 2). Construction of these sets was approved by the Lothian Research Ethics Committee (08/S1101/41). Tumour areas were selected for TMA construction on H&E slides and 0.6 mm^2 cores were placed into three separate TMA replicates for each sample, as previously described [198]. H&E slides for both TMA sets were subsequently re-examined and phenotyped by a senior pathologist (Dr. J. Thomas) to ensure that only IDC-NST were included in the analysis.

Immunofluorescence

AQUA (Automated QUantitative Analysis) methodology has been described elsewhere [189, 199, 200]. Briefly, antigen retrieval for all epitopes was carried out using heat treatment under pressure in a microwave oven for 5 min in citrate buffer (82ml 0.01M sodium citrate: 18ml 0.01M citric acid) pH 6.0. Slides were incubated with primary antibodies for 1 hour at room temperature. Details of primary antibodies are summarized in Table 2.3. Rabbit primary antibodies were incubated overnight with mouse anti-pancytokeratin (Invitrogen, #18-0059, 1:25) to visualise epithelial cells. Mouse primary antibodies were incubated overnight with rabbit anti-pancytokeratin (Dako, #Z0622, 1:150) and rabbit anti-pancadherin (Cell Signalling, #4068, 1:50). The epithelial compartment was then visualised with Cy3 (Invitrogen, anti-rabbit #A21422; anti-mouse #A21428, both used at 1:25). DAPI (4',6-diamidino-2-phenylindole) counterstain (Invitrogen, #P36931) was used to identify nuclei and Cy-5-tyramide (HistoRx, #AQ-EMM1-0001, 1:50) was used to detect protein 'targets'.

Table 2.3. Details of antibodies used for immunofluorescence and western blotting.

Target	Source	Catalogue No.	Host	Dilution (IF)	Dilution (WB)	Apparent Molecular Wt (KDa)
E-cadherin	BD	610181	Mouse	1:1500	1:2500	120 KDa
Claudin7	Abcam	Ab75347	Rabbit	n/a	1:1000	23 KDa
N-cadherin	BD	610921	Mouse	1:300	1:3000	130 KDa
Vimentin	Sigma	V 6630	Mouse	1:400	1:1000	58 KDa
Fibronectin	Abcam	ab2413	Rabbit	1:10000	1:5000	262 KDa
Zeb1	D. Darling	[201]	Rabbit	1:1000	n/a	n/a
Slug	LifeSpan Bio	LS-C30318	Rabbit	1:1000	1:4000	34 KDa
Snail	Abcam	ab17732	Rabbit	1:700	1:4000	29 KDa
β-catenin (total)	BD	610153	Mouse	1:500	1:2000	92 KDa
β-catenin (active)	Millipore	05-665	Mouse	n/a	1:1000	92 KDa
pan-WT1	Genetex	GTX15249	Rabbit	1:100	n/a	n/a
Tubulin	Abcam	Ab7291	Mouse	n/a	1:6000	50 KDa

AQUA automated image analysis

Monochromatic images of each TMA core were captured at 20x magnification using an Olympus AX-51 epifluorescence microscope. These high-resolution, digital images were individually, manually quality checked to crop aberrant imaging artefacts and exclude them from analysis. All areas of normal mammary duct and DCIS were also cropped to ensure that only invasive cancer was included in the analysis. Checked images were then analysed using AQUA software [199]. Briefly, a binary epithelial mask was created from the cytokeratin image of each TMA core. If the epithelium comprised <5% of total core area, the core was excluded from analysis. Similar binary masks were created for cytoplasmic and nuclear compartments based on DAPI staining of nuclei. Target protein expression was quantified by calculating the Cy5 fluorescent signal intensity on a scale of 0 to 255 within each image pixel. The AQUA score was generated by dividing the sum of Cy5 signal within the epithelial mask by the area of the cytoplasmic compartment. Cytokeratin, DAPI, 'protein target' and 'composite signal' AQUA images are illustrated in Figure 2.1. Quantitative analysis of ER, PR and HER2 expression using AQUA correlates well with traditional immunohistochemistry (IHC) techniques (Allred and HercepTest). Furthermore, AQUA demonstrates a particularly dynamic range of scores at higher IHC scores as shown in Figure 2.2 [196].

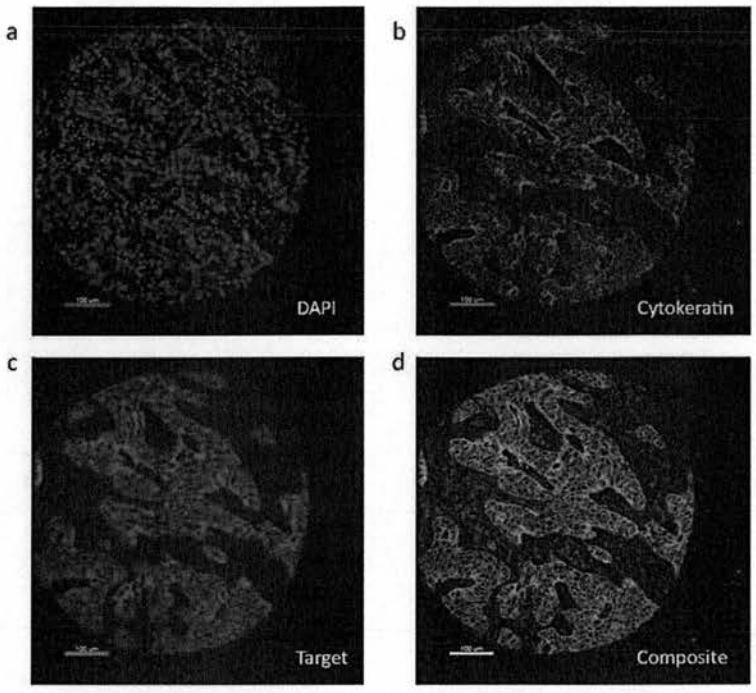


Figure 2.1. AQUA quantitative image analysis for compartmentalised analysis of tissue sections. (a) DAPI counterstain was used to identify nuclei, (b) anti-pancytokeratin was used to identify infiltrating tumour cells and normal epithelium and (c) Cy-5-tyramide was used to identify target proteins. Cytokeratin = green; DAPI = blue; E-cadherin = red. Bars = 100 μm.

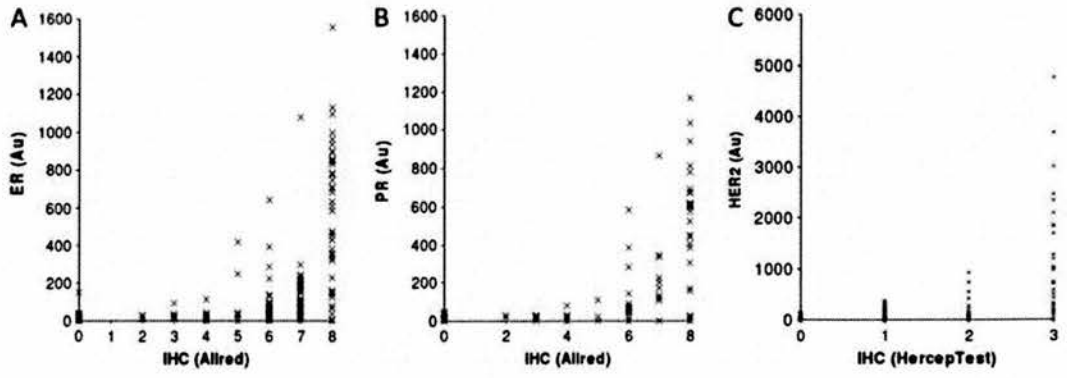


Figure 2.2. Quantitative receptor expression analysis using AQUA compared with IHC. (a) ER, (b) PR and (c) HER2. There is a continuous distribution of AQUA scores with a wide dynamic range of expression at high-intensity staining. AQUA scores for ER, PR and HER2 show good correlation with IHC (Pearson regression coefficients, $r = 0.66$, 0.68 and 0.50 respectively). Reproduced with kind permission from Dr. D. Faratian, Consultant Pathologist, Western General Hospital, Edinburgh [196].

Statistical analysis methods for AQUA data

An arbitrary cut-off of ≤ 50% available AQUA scores was selected for the exclusion of cases with insufficient data prior to analysis. Cases with special type carcinomas were also excluded from analysis. AQUAsition data was matched to TMA maps, facilitating the investigation of correlations between EMT markers and clinical parameters. AQUA scores for each array were $\log_2$ transformed and mean centred. The mean of the three replicates for each protein target was used for further analysis. As a consequence of this averaging process, zero no longer represents the mean of the data that is shown graphically.

Associations between variables were calculated using Pearson's correlation coefficients and differences in means with one-way ANOVA using SPSS software v14. Overall survival was subsequently assessed by Kaplan-Meier analysis with log-rank testing to determine statistical significance. Deriving cutpoints in Kaplan-Meier analysis using minimum *P* statistics can introduce a type I statistical error through multiple testing [202]. X-Tile, a bio-informatics tool that allows optimal cutpoints to be derived whilst correcting for the use of minimum *P* statistics [203], was used in order to reduce this type I error. The two statistical corrections used were the Monte-Carlo *P* value and the Miller-Siegmund minimal *P* correction [202]. Pairwise comparisons of change in phenotype between primary tumours and paired LN metastases were made using the Chi-squared test.

Retrospective power calculations were carried out using online formulas available at www.stattools.net/SSizSurvival_Pgm.php and www.dssresearch.com/KnowledgeCenter/toolkitcalculators/samplesizecalculators.aspx to determine the required population sizes to establish whether observed associations in Chapter 4 were real.

Cell lines and culture conditions for use in invasion assays

Table 2.4, amended from [204], gives details of all claudin-low/EMT-like cell lines as well as MCF10A cells.

Table 2.4. Details of cell lines and culture conditions used in invasion assays.

Cell Line	Gene cluster	ER	PR	Tumour type	Culture media	Culture conditions
HBL100	BaB	-	[-]	N	DMEM, 10% FCS	37°C, 5% CO_2
SUM1315	BaB	-	[-]	IDC (skin metastasis)	Ham's F12, 5%-IE	37°C, 5% CO_2
SUM159PT	BaB	[-]	[-]	AnCar	Ham's F12, 5%-IH	37°C, 5% CO_2
BT549	BaB	-	[-]	IDC, pap	RPMI, 10% FCS	37°C, 5% CO_2
MDAMB436	BaB	[-]	[-]	IDC	L15, 10% FCS	37°C, no CO_2
MDAMB157	BaB	-	[-]	MC	DMEM, 10% FCS	37°C, 5% CO_2
MDAMB231	BaB	-	[-]	AC	DMEM, 10% FCS	37°C, 5% CO_2
HS578T	BaB	-	[-]	IDC	DMEM, 10% FCS	37°C, 5% CO_2
MCF10A	BaB	-	[-]	F	DMEM/ F12*	37°C, 5% CO_2

AC, adenocarcinoma; AnCar, anaplastic carcinoma; BaB, Basal B (the basal-like cluster was found to comprise of two major subdivision termed 'A' and 'B'); F, fibrocystic disease; IDC, invasive ductal carcinoma; MC, metaplastic carcinoma; N, normal; pap, papillary. Expression data for ER and PR are derived from mRNA and protein levels, square brackets indicate that levels are inferred from mRNA levels alone. Media conditions: FCS, foetal calf serum (Harlan, #S-0001AE); I, insulin 0.01mg/ml (Sigma, #I9278); H, hydrocortisone 500ng/ml (Sigma, #H0888); E, EGF 20ng/ml (Sigma, #E9644); DMEM, Dulbecco's Modified Eagle medium, GIBCO #31885-023; RPMI, RPMI medium 1640, GIBCO #21875-034; Ham's F12, D-12 nutrient mixture (Ham), GIBCO #21765-029; DMEM/F12, Dulbecco's modified Eagle's medium: Nutrient mix F-12 (D-MEM/F-12), GIBCO #31330-038; L15, Leibovitz's L-15 medium, GIBCO #11415-049. For MCF10A, supplement DMEM/F12 media with 5% horse serum (Invitrogen, #16050-122), 20ng/ml EGF, 100ng/ml cholera toxin (Sigma, #C8052), 0.01mg/ml insulin and 500ng/ml hydrocortisone.

All cell lines obtained from ATCC (American Type Culture Collection; www.atcc.org) apart from SUM1315 and SUM159PT which were obtained from Dr. A. Orimo, University of Manchester.

Primary cell isolation for tissue culture

Normal fresh breast tissue cores were incubated for 1 hour at room temperature in tissue mix consisting of DMEM/F12, 1% fungizone (Invitrogen, #15290018), 1% penicillin/streptomycin (GIBCO, #15140, 10,000U/ml penicillin, 10,000ug/ml streptomycin), 10ug/ml insulin and 10% FCS. Tissue cores were then finely chopped (~1ml pieces) and put in a tissue mix/Collagenase I solution (GIBCO, #17100-017, made up with 200uL of 200U/ml Collagenase I to 20ml tissue mix) for digestion (2 hours at 37°C, 200 rpm). The digested tissue was then spun for 4 minutes at 60g. The resulting pellet was plated with fibroblast media (DMEM supplemented with 10% FCS, 50 U/ml penicillin and 50 mg/ml streptomycin) and the supernatant spun for a further 4 minutes at 600g, 4 times. The resulting second pellet was plated with HMEC media (CnT-22 (CELLNTEC, #CnT-22.BM) supplemented with 5% FCS). The same protocol was used to isolate epithelial cells and fibroblasts from fresh breast cancer specimens.

Rat tail collagen I preparation for use in invasion assay

Fresh rat tails were collected and frozen. Prior to harvesting these were placed in 70% ethanol. Tendons were stripped from the tails and returned to 70% ethanol to sterilise. The collected tendons were weighed and transferred to the appropriate volume of pre-cooled acetic acid (1g tendon to 250ml 0.5M acetic acid) and gently stirred for 48 hours at 4°C. The tendon/acetic acid mix was then centrifuged at 10,000g for 30 minutes and the pellet discarded. The remaining supernatant was measured and an equal volume of 10% (w/v) NaCl added. This mix was allowed to stand overnight at 4°C. The collagen-rich, insoluble 'bottom layer' was taken and collected by further centrifugation (10,000g for 30 minutes). The collagen-rich material was resuspended in 0.25M acetic acid at 4°C and dialysed against 1:1000 acetic acid at 4°C for 3 days, changing the dialysis buffer twice daily. The collagen solution was then sterilised by centrifugation (20,000g for 2 hours) and stored at 4°C. Collagen was diluted as required by the addition of sterile 1:1000 acetic acid to a stock concentration of 1.2mg/ml.

Collagen based 'on top' invasion assays

This assay was used to investigate invasion following the expression of C35 protein. Methodology for this assay has been previously described [205, 206]. 3ml of rat tail collagen I solution was prepared (25% Collagen stock, 55% sterile 1:1000 acetic acid, 10% DMEM, ~10% 0.22M NaOH (Sigma, #S2770), and 10% FCS to give a 10% excess) with 0.1 x 10^6 normal human breast fibroblasts per gel. This mix was transferred to a 35mm Petri dish and allowed to contract in fibroblast media over a period of 4-7 days. Sufficient contraction was judged as a ~4-fold reduction in size. When this was achieved, gels were carefully transferred to a 24 well plate and 0.3 x 10^6 H16N-2 cells from the desired lines seeded on top. H16N-2 cells were obtained as a kind gift from Dr. V. Band, Band and Sanger, 1991. These were then incubated as submerged cultures for 3-4 days in H16N-2 media (MEGM supplemented with 5% FCS (Lonza, #CC-3150). To induce invasion, the gels were raised to the air/liquid interface and incubated for a further 7 days. This was done by transferring the gels onto a raised mesh within a 50mm Petri dish. At this stage, the gels were fixed in 10% phosphate buffered formalin and wax embedded (Figure 2.3a). Details of transfections of H16N-2 with C35 protein are fully described elsewhere [206].

Matrigel based 'plug' invasion assays

This assay was used to investigate the invasion of cells across basement membrane. 200uL cell-collagen plugs and 75uL cell-Matrigel plugs were made in a u-shaped 96 well plate, with the aim of achieving comparable size after a 24hr incubation (day -1). A cell concentration of 1 x 10^6 was used for all plugs. Details of all lines used in this assay are shown in Table 2.4. Rat tail collagen I, for both plugs and subsequent embedding, was prepared as per the 'on top' assays. Growth factor reduced Matrigel was obtained from BD (#354230, 9.4mg/ml) and used at a 5mg/ml. Matrigel matrix is a soluble basement membrane extract of the Engelbreth-Holm-Swarm tumour that gels at room temperature to form a reconstituted basement membrane. The major components are laminin, collagen IV, entactin and heparin sulphate membrane. After the 24 hour incubation, cell plugs were carefully removed from their 96 well plate and embedded in 1ml of collagen in a 24 well plate (taken as day 0), with or without fibroblasts (used at 10,000/ml). These cultures were incubated for a further hour and

then carefully freed from the edges of the well (to allow contraction of the collagen) and supplemented with 0.5ml of cell-specific media. The cultures were then left to invade. Media was changed weekly. Gels were fixed at either 1 or 2 weeks in 10% phosphate buffered formalin and wax embedded (Figure 2.3b). This assay is similar to that of Sabeh and colleagues [207].

Haematoxylin and eosin staining

Slides were de-waxed and re-hydrated, stained and then de-hydrated and cover-slipped according to a standard protocol detailed below. The Haematoxylin used was Shandon's ready-made Harris Haematoxylin. This must be filtered before use and discarded after every 200 slides. Eosin is supplied by Shandon as EosinY.

De-waxing and Rehydration:

- De-wax- Xylene solution 1 — 5 minutes
- De-wax- Xylene solution 2 — 5 minutes
- De-wax- Xylene solution 3 — 5 minutes
- Rehydration- 100% Alcohol solution 1 — 2 minutes
- Rehydration- 100% Alcohol solution 2 — 2 minutes
- Rehydration- 80% Alcohol — 2 minutes
- Rehydration- 50% Alcohol — 2 minutes
- Wash in running water — 2 minutes

Staining:

- Haematoxylin — 5 minutes
- Wash in running water — 30 seconds to 10 minutes
- Scott's tap water substitute — 5 minutes
- EosinY — 5 minutes
- Wash in running water — briefly

Dehydration:

- Dehydration-50% Alcohol 30 seconds
- Dehydration-80% Alcohol 30 seconds
- Dehydration-100% Alcohol solution 1 2 minutes
- Dehydration-100% Alcohol solution 2 2 minutes
- Clearing-Xylene solution 1 5 minutes
- Clearing-Xylene solution 2 5 minutes
- Clearing-Xylene solution 3 5 minutes

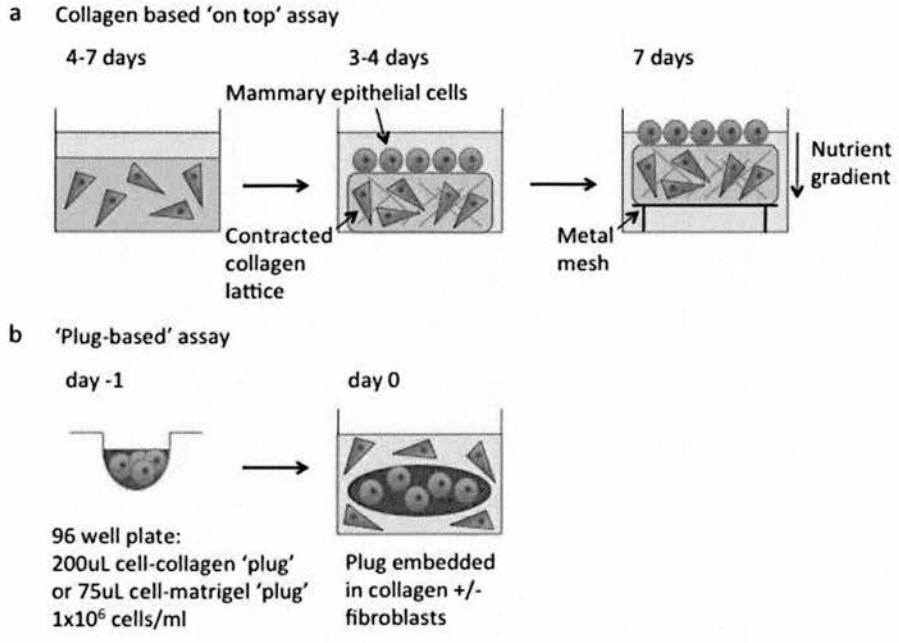


Figure 2.3. Schematic illustration of the 'on-top' and 'plug-based' invasion assays. (a) In the on-top assay, a contracted collagen gel containing fibroblasts is seeded with epithelial cells. These gels are then incubated as submerged cultures such that the epithelial cells invade down into the gel along a nutrient gradient. (b) In the plug-based assay, a collagen- or matrigel-based epithelial plug is generated in a 96 well plate. This plug is then embedded in additional collagen, which may or may not contain fibroblasts as required. With time, the epithelial cells invade out of the plug into surrounding collagen in a star-burst manner that is easily visualized by light microscopy.

SDS-PAGE

Protein lysates (50ug/well, as determined by MicroBCA protein assay) were resolved by SDS-PAGE after being denatured for 1 hour at 60°C. The resolving gel (7.5% w/v acrylamide, 0.37M TRIS pH8.85, 0.1% SDS, 0.02% AMPS, 0.25% TEMED) was set between glass plates using a Bio-Rad kit. Once the resolving gel had set, a stacking gel (3.6% w/v acrylamide, 0.12M TRIS pH6.8, 0.1% SDS, 0.03% AMPS, 0.33% TEMED) was layered and a comb used to create wells for sample loading. The loaded samples were electro-separated under constant current (100-200mA) using electrophoresis buffer (25mM Trizma Base, 0.19M Glycine, 10% SDS). Electro-transfer onto immobilon transfer membrane (Millipore, #IPVH304F0) was performed using transfer buffer (25mM Trizma Base, 0.19M Glycine) using a Bio-Rad kit, under constant electrical potential (~30mV for at least 2 hours). All chemicals used here and for western blotting came from Sigma unless otherwise stated.

Western Blotting

Nonspecific binding was blocked with Li-Cor Odyssey Blocking Buffer (Li-Cor, #927-40000), diluted 50:50 in PBS, for 1 hour at room temperature. Primary antibodies were diluted in Li-Cor Odyssey Blocking Buffer (see Table 2.3 for details), diluted 50:50 in 0.1% PBS-Tween20, and incubated with the blot overnight at 4°C. Blots were washed 3 times for 5 minutes with PBS-T before incubation with appropriate fluorescent secondary antibodies (Li-Cor, anti-rabbit 680nm, #926-32221; anti-rabbit 800nm, #926-32211; anti-mouse 680nm, #926-32220; anti-mouse 800nm, #926-32210), diluted 1:10,000 in Li-Cor Odyssey Blocking Buffer, diluted 50:50 in 0.1% PBS-Tween20, for 45 minutes at room temperature. Exposure to light was avoided. Subsequently, membranes were washed, dried and scanned on the Li-Cor Odyssey scanner. All washes/incubations were carried out under constant agitation.

RNA extraction, gene expression micro-array construction and analysis, and qRT-PCR (quantitative real time polymerase chain reaction)

RNA was extracted from the collagen invasion assays using an RNeasy Mini kit (Qiagen, #74104) and from cell lines cultured on plastic using an AllPrep DNA/RNA Mini kit (Qiagen, #80204). Quality and concentration was evaluated using an Agilent Bioanalyser and Agilent RNA 6000 Nano kit (Agilent Technologies, #5067-1511) and Nanodrop. An RNA Integrity Number (RIN) of >9 was required before proceeding as low quality, fragmented RNA may compromise qRT-PCR or micro-array analysis.

RNA from the collagen invasion assays was labelled using an Illumina TotalPrep RNA amplification kit (Ambion, #AMIL1791) according to manufacturer's instructions. Triplicate samples from invasion assays (1500ng cDNA per assay) were hybridised to Illumina BeadChips and whole genome gene expression analysis performed using the Illumina HumanRef-8 v3 Expression BeadChip and BeadArray Reader. Microarray data was analysed using packages within Bioconductor [208] (http://www.bioconductor.org) that implement R statistical programming. Gene expression data was normalised using quantile normalisation within the BeadArray package [209] and differential gene expression assessed using Significance Analysis of Microarrays (SAM) [210] and the siggenes package. The dataset from Hershkowitz and colleagues [51] was downloaded from the UNC Microarray Database (https://genome.unc.edu/). Microarrays were constructed by Dr Katz and Dr Larionov and analysed by Dr Sims (all are part of the Breakthrough Breast Cancer Research Unit, Western General Hospital).

RNA from cell lines cultured on plastic was converted to cDNA prior to PCR using a QuantiTect Reverse Transcription kit (Qiagen, #205311). Gene expression patterns for invasion assays (biological triplicates) and cell lines cultured on plastic (technical triplicates) were examined using the QuantiTect SYBR Green PCR kit (Qiagen, #204145) and a Corbett RotoGene 3000. Primers for *CDH1* were: forward 5'-CGGAGAAGAGGACCAGGACT-3', reverse 5'-GGTCAGTATCAGCCGCTTTC-3'; for *CLDN7*: forward 5'-AAAATGTACGACTCGGTGCTC-3', reverse 5'-AGACCTGCCACGATGAAAAT; for *TBP*: forward 5'-

GGGGAGCTGTGATGTGAAGT-3', reverse 5'-CCAGGAAATAACTCTGGCTCA-3'; for *ACTB*: forward 5'-CCTTCCTGGGCATGGAGTCCT-3', reverse 5'-GGAGCAATGATCTTGATCTT-3'. QuantiTect Primer Assays (Qiagen) were used for *KRT8, CRB3, MARVELD3, IRF6*, *MAL2*, *TACSTD1* and *SPINT2*. PCR programme was identical for all genes: 95°C, 15 minutes; (94°C, 15 seconds; 56°C, 30 seconds; 72°C, 30 seconds) x 50 cycles; 72°C, 5 minutes. Standard reference human cDNA was from Clontech (#639654), random primed. ~50ng RNA equiv/mL was used for quantification of mRNA expression. Final normalisation was performed against the geometrical mean of *ACTB* and *TBP* levels.

Gene promoter analysis

Using the presumptive promoter region for the 9 relevant genes (a 2kb region upstream of the presumptive transcription start site determined using Ensembl 52, Jan2009, based on NCBI 36 assembly), over-represented 6- and 7-mer oligos were identified using oligo-analysis [211] from the RSAT-tools package (http://rsat.scmbb.ulb.ac.be/rsat/) [212]. This program counts all oligonucleotide occurrences within the sequence set and estimates their statistical significance. A calibration is done using the entire genome promoter regions as a background model (Ensembl 52, Jan2009, based on the NCBI 36 assembly). The best 7-mer candidates were identified and their sequences compared to the entire collection of consensus binding sites available from Transfac professional [213] (release 2010.1) using the compare-pattern script (RSAT-tools) and the associated binding-factors identified. The analysis was performed by Dr. P. Gautier, MRC Human Genetics Unit Bioinformatics Service, Western General Hospital.

Chapter 3: The down-regulation of E-cadherin is uncoupled from an EMT programme in high-grade invasive ductal breast cancers

Introduction

Transcriptional repression is a key mediator of EMT

Increasing evidence suggests that EMT has a key role in cancer progression, underlying invasion, metastatic dissemination and acquisition of resistance. This role has predominantly been inferred from *in vitro* and animal studies and controversy regarding the precise role of EMT in human cancer remains. A decrease in E-cadherin expression is possibly one of the most important consequences of EMT resulting in the changed behaviour of tumour cells [72]. Various mechanisms may down-regulate E-cadherin expression in cancer but transcriptional repression is thought to be particularly important. A number of zinc finger-containing repressors, capable of interacting with E-boxes within the *CDH1* promoter, have been identified. These include Snail, Slug (Snail2), Zeb1, Zeb2 and Twist [76].

Following investigation in a series of human carcinoma cell lines, Snail was the first of these transcription factors to be characterised as a repressor of E-cadherin and inducer of EMT [214, 215]. Subsequently, Snail expression has been reported in various human tumours, including breast, where expression correlates with LN and distant metastases [126, 216, 217]. The expression of Slug is similarly associated with poor clinico-pathological outcomes in breast cancer [217, 218]. However, Slug expression has been found to correlate with a partially differentiated phenotype, suggesting that Snail and Slug may play different roles in the progression of breast cancer [219]. Zeb1 expression has predominantly been studied in colorectal and uterine cancers, where expression correlates with poor outcome [220, 221]. In the context of breast, Zeb2 expression has been reported in pleural effusions (although not predictive of outcome) [217], and both Zeb1 and Zeb2 are associated with repression of E-cadherin and EMT in transformed MCF10A cells [222]. The ability of Twist to induce EMT and the development of metastases was originally described in a mouse model. In addition, the same study reported an inverse relationship

between the expression of E-cadherin and Twist in lobular carcinomas of the breast [223]. However, further studies have shown expression of Twist to be a feature of high-grade breast carcinomas of various types [224]. Despite similarities between these transcription factors, the complexity of their individual roles in processes such as EMT remains to be fully understood.

Canonical Wnt signalling is an alternative EMT mechanism

β-catenin is a multifunctional protein. At the plasma membrane, it is an important component of the adherens junction, facilitating cell-cell adhesion by linking E-cadherin, in conjunction with α-catenin, to the actin cytoskeleton [82]. In addition, β-catenin is the main effector of canonical Wnt signalling. Normally, free cytoplasmic β-catenin is rapidly degraded by means of a degradation complex consisting of the serine and threonine kinases CK1 and GSK3β, the scaffold protein axin and the adenomatous polyposis coli protein APC. Binding of β-catenin to this complex results in β-catenin phosphorylation and subsequent proteosomal breakdown [225]. Binding of Wnt ligands to a coreceptor complex consisting of a seven, transmembrane domain Frizzled receptor and LDL receptor-related proteins LRP5 or LRP6 inhibits GSK3β [226]. Consequently, β-catenin accumulates in the cytoplasm and translocates to the nucleus. Here, it interacts with members of the T-cell-specific transcription factor/lymphoid enhancer binding factor (TCF/LEF) family of transcription factors to regulate gene expression [227] (Figure 3.1).

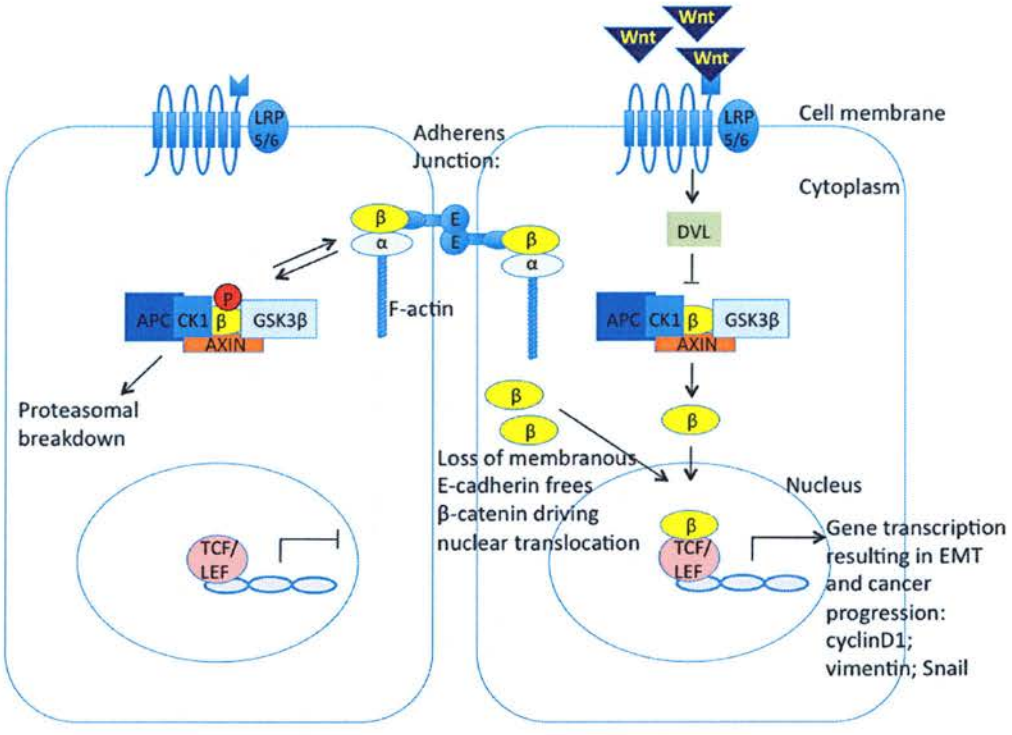


Figure 3.1. Canonical Wnt signalling. In the absence of Wnt signals (*left panel)*, β-catenin (β) is localized in adherens junctions at the plasma membrane, contributing to cell-cell adhesion in conjunction with E-cadherin (E) and α-catenin (α). Cytoplasmic β-catenin is targeted for proteasomal breakdown by a multiprotein degradation complex that includes APC, CK1, axin and GSK3β. Binding of Wnt molecules to frizzled receptors (*right panel*) inhibits the degradation complex and allows β-catenin to accumulate and translocate to the nucleus where specific transcriptional programmes are activated. Frizzled receptor mediated recruitment of the cytoplasmic protein dishevelled (DVL) is an important step in this signalling pathway [228]. Loss of membranous E-cadherin can also liberate β-catenin and drive its nuclear translocation. Adapted from [71].

The canonical Wnt signalling pathway is known to regulate EMT-programmes in the developing mammary gland and increasing evidence suggests that this pathway may be up-regulated in various cancers, including breast [229-231]. A study has shown that Wnt signalling, through a β-catenin-TCF/LEF complex, is capable of driving EMT by up-regulating Snail in human breast cancer cell lines [232]. In addition, loss of E-cadherin-β-catenin junctions at the cell membrane can independently drive EMT by allowing freed β-catenin to translocate to the nucleus [186, 233, 234]. These studies, carried out *in vitro* and in animal models, support the nuclear translocation of β-catenin as a key event driving EMT and poor outcome [235]. Additional studies indicate that there are multiple reciprocal interactions between E-cadherin, β-catenin and EMT-inducing transcriptional repressors [71]. Illustrating some of this complexity, Stemmer and colleagues [236] have demonstrated positive feedback stimulation of Wnt signalling by Snail that is independent of its repressor activity and independent of loss of E-cadherin.

A number of immunohistochemical studies have investigated β-catenin expression patterns and correlation with clinical outcome with mixed results [237-239]. However, Rimm and colleagues, who were the first to utilise an automated, quantitative immunohistochemical analysis of protein expression using the AQUA system, report a significant correlation between loss of cytoplasmic β-catenin and poor outcome [240]. However, this study does not support a translocation mechanism as almost no nuclear expression of β-catenin (<10/600 cases) was observed. Consequently, the role of β-catenin as a mechanism of EMT in human breast cancer warrants further investigation. A recent study using a TCF/LEF reporter has shown that canonical Wnt signalling activity is a marker of CSCs in colon cancer [241]. The evidence suggesting a link between EMT and stemness has already been discussed above. This additional evidence in the context of Wnt signalling makes the investigation of β-catenin in breast cancer particularly relevant.

Aims

To clarify the role of EMT in human breast cancer at the protein level we selected two patient cohorts that maximised the likelihood of identifying EMT-related events. As the most common form of the disease [242], only invasive ductal carcinomas of no special type (IDC-NST) were examined. The expression of key EMT-related proteins, including β-catenin and transcriptional repressors of E-cadherin, were examined using automated, quantitative, immunofluorescence analysis.

Results

Human breast cancer has a wide range of E-cadherin expression, highly correlated with β-catenin

To determine whether down-regulation of E-cadherin is evident in human IDC-NST, quantitative immunofluorescence analysis of E-cadherin protein expression was performed. Scores for immunofluorescence imaging were given by the AQUA system and were allocated as membrane/cytoplasmic and nuclear as appropriate. These scores are relative values and are based on expression within the whole tumour cohort.

A continuous range of E-cadherin expression was seen in both sets (Figure 3.2a). In set 1, an 18-fold change in the expression of E-cadherin was seen. A corresponding 11-fold change was seen in set 2. No correlation between E-cadherin expression and ER or HER2 receptor status was observed (Figure 3.2b). Qualitative assessment of immunofluorescence images showed that E-cadherin staining remains membranous across all breast tumours (Figure 3.2c). This is an important observation as tumours with non-functional, cytoplasmic E-cadherin (potentially indicative of EMT) would not be allocated distinctly low AQUA scores. Quantitative analysis of β-catenin expression was performed to determine the relationship with E-cadherin expression. Tumours with high cytoplasmic β-catenin expression were identified (Figure 3.3a). Nuclear β-catenin expression, although less distinct, was also observed and positively correlated with cytoplasmic expression (Table 3.1). Importantly, a positive correlation between E-cadherin and β-catenin expression was seen in both patient sets (Figure 3.3b and Table 3.1).

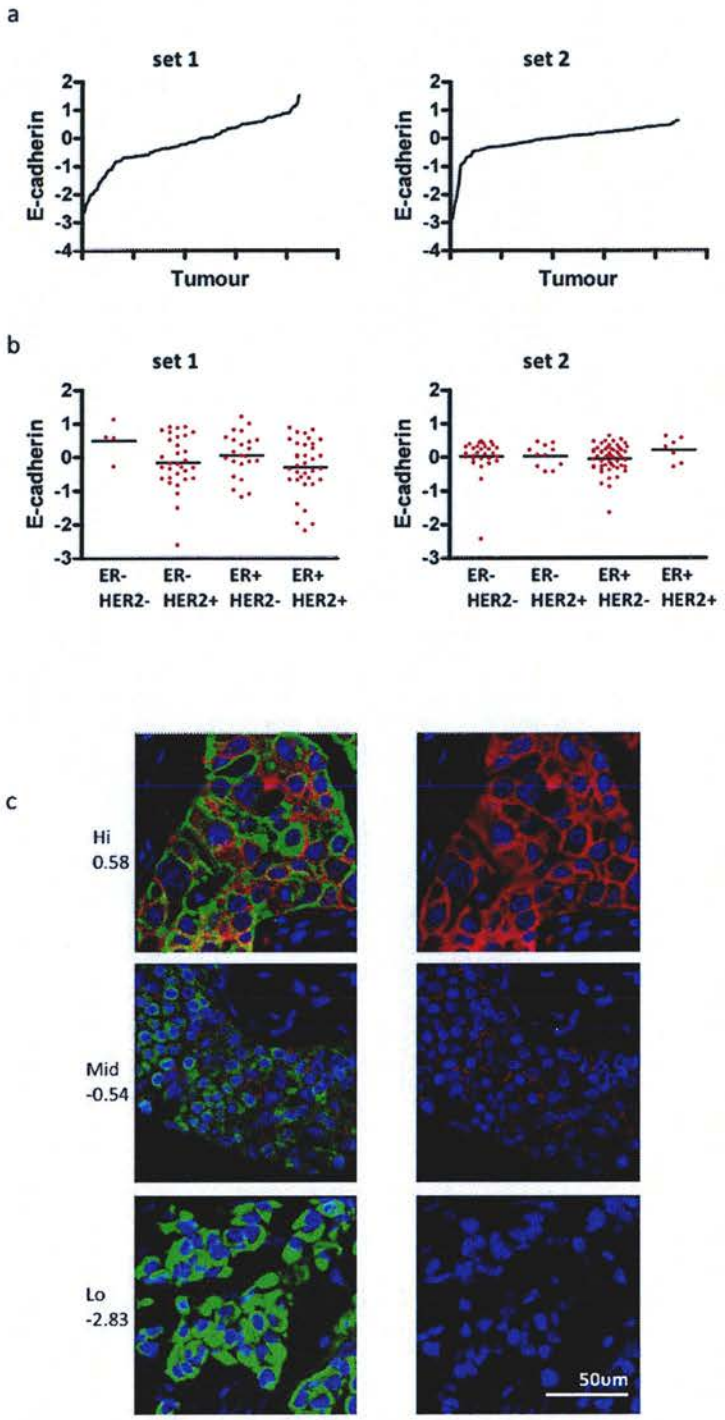


Figure 3.2. Human breast cancer has a wide range of E-cadherin expression. (a) Comparable, wide ranges of E-cadherin expression are seen in both TMA sets.

AQUA scores for each array were $\log_2$ transformed and mean centred (0 represents the adjusted mean for each individual TMA slide). The mean of three TMA replicates is shown. (b) E-cadherin protein expression distribution by ER and HER2 receptor status in both TMA sets. (c) Representative immunofluorescence images illustrating varying E-cadherin protein expression (*left panels*, with pan-cytokeratin mask; *right panels*, without pan-cytokeratin). Note the membranous staining of E-cadherin throughout. Normalised AQUA scores are given (arbitrary units). Bar = 50 μm. Cytokeratin = green; DAPI = blue; E-cadherin = red.

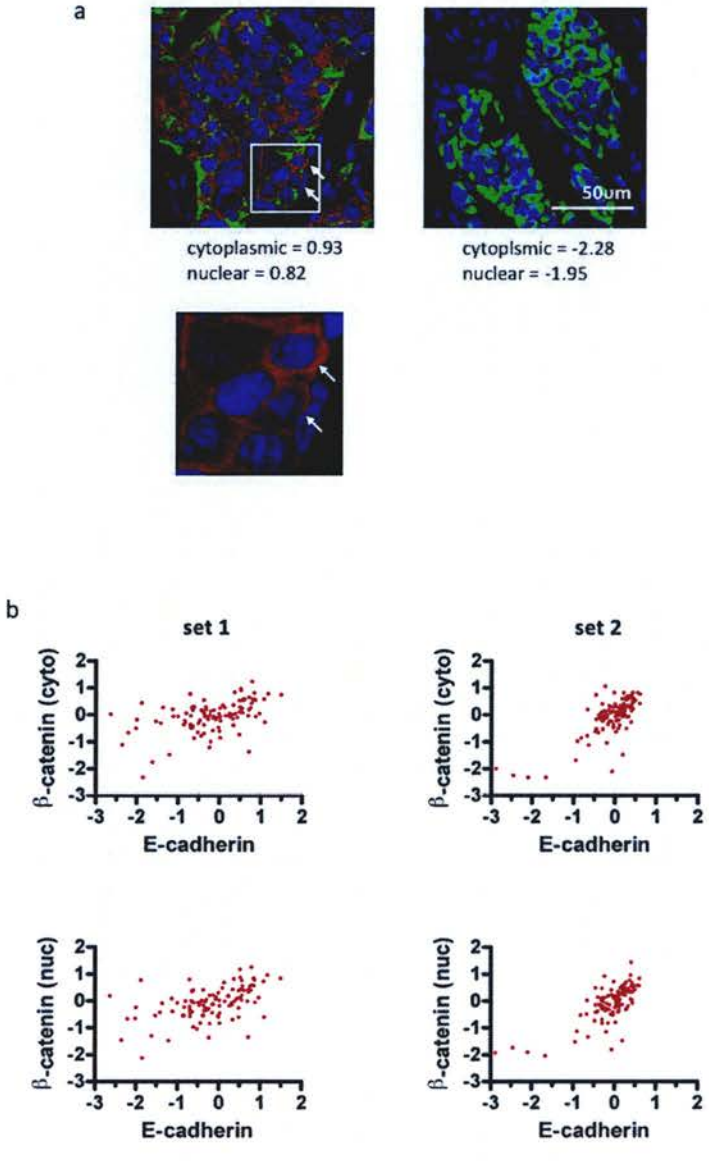


Figure 3.3. E-cadherin and β-catenin protein expression are highly correlated in human breast cancer. (a) Representative immunofluorescence images illustrating tumours with high (*left panel*) and low (*right panel*) expression of β-catenin. Inset shows nuclear staining in individual cells (arrows). Bar = 50 μm. Cytokeratin = green; DAPI = blue; β-catenin = red. (b) Plots comparing membrane/cytoplasmic and nuclear β-catenin expression levels to E-cadherin in both TMA sets (all $p<0.01$ by Pearson's correlation, see Table 3.1).

High expression of N-cadherin and other mesenchymal markers is found in human breast cancer

Having identified tumours with relatively low E-cadherin expression, we hypothesised that some of these might be undergoing EMT and should therefore show evidence of a 'cadherin switch'. Tumours with high N-cadherin expression (Figure 3.4a) were identified. A weak negative correlation with E-cadherin expression was observed in set 1 (Pearson's correlation, $r=-0.204$, $p<0.05$; Figure 3.5 and Table 3.2), the set enriched for HER2 expression. No such correlation was found in set 2 (Figure 3.5 and Table 3.2). Vimentin and fibronectin are two other established mesenchymal markers that are up-regulated in EMT in the context of breast and other cancers [243]. As with N-cadherin, tumours with relatively high vimentin and fibronectin expression were identified (Figure 3.4b-c), but no correlation with E-cadherin expression observed (Figure 3.5, Table 3.2). Immunofluorescence images show stromal and epithelial staining for these markers as expected. Interestingly, fibronectin was expressed focally and not evenly in high expressing tumours (Figure 3.4c). No correlation was observed between N-cadherin, vimentin and fibronectin expression in these tumours (Table 3.2). These findings suggest that up-regulation of mesenchymal markers is an uncommon event and that tumours that up-regulate one protein marker do not necessarily up-regulate others. Mesenchymal protein up-regulation is uncoupled from E-cadherin down-regulation in the breast cancers examined.

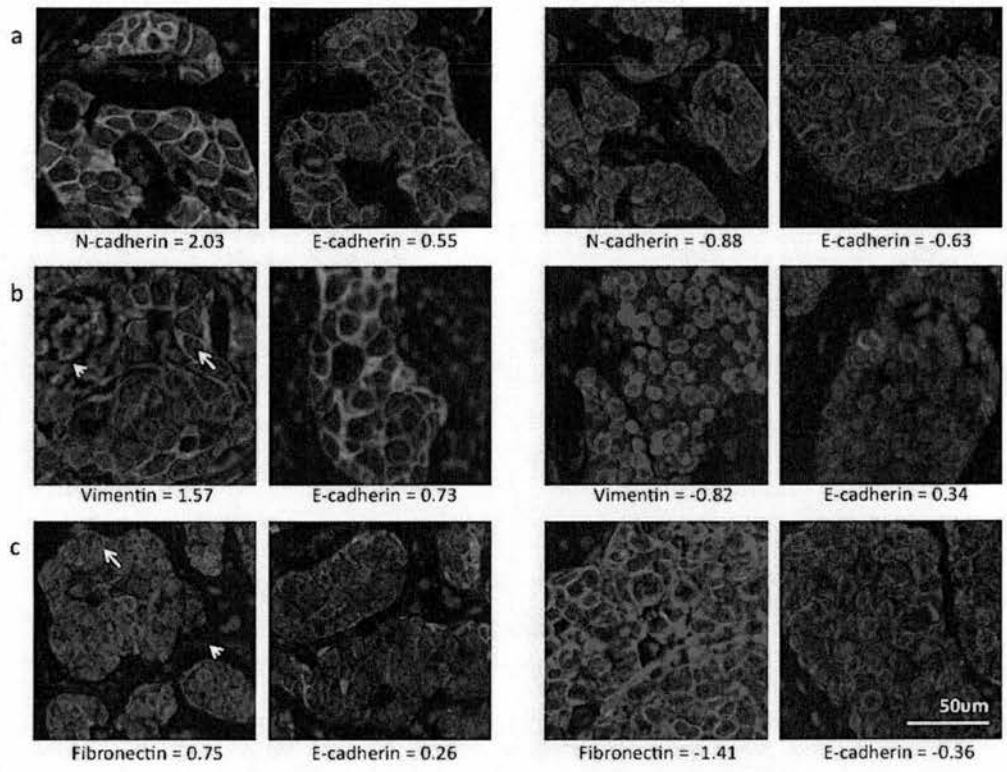


Figure 3.4. High expression of mesenchymal proteins is found in human breast cancer. Representative immunofluorescence images illustrating tumours with high (*left panels*) and low (*right panels*) expression of (a) N-cadherin, (b) vimentin and (c) fibronectin. Corresponding E-cadherin expression is shown in each case. Note the stromal (short arrow) as well as epithelial (long arrow) expression of vimentin and fibronectin. Bar = 50 μm. Cytokeratin = green; DAPI = blue; N-cadherin/vimentin/fibronectin = red.

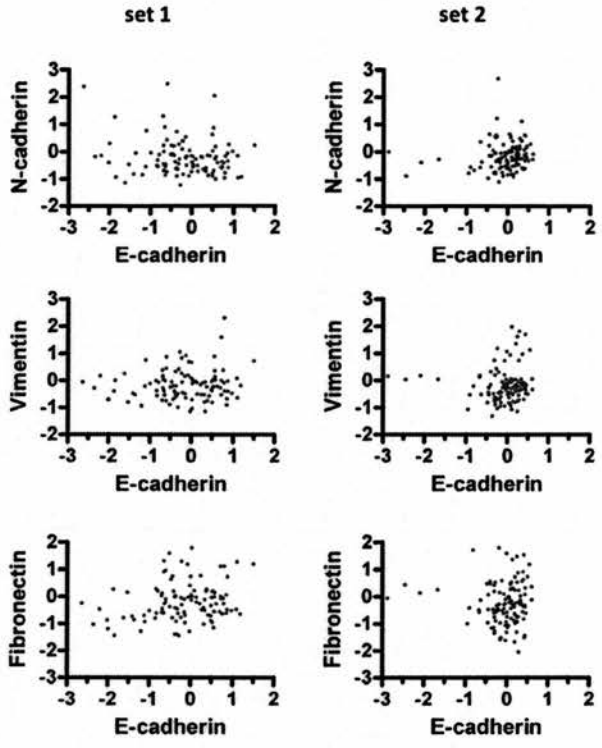


Figure 3.5. Correlations between mesenchymal proteins and E-cadherin. Plots comparing N-cadherin, vimentin and fibronectin protein expression levels to E-cadherin are shown for both sets.

High expression of Snail and Slug in human breast cancer

Quantitative analysis of transcription repressors was performed to determine their correlation with E-cadherin expression and other EMT markers. Tumours with relatively high Snail and Slug expression were identified (Figure 3.6), but no correlation with E-cadherin expression was observed (Figure 3.7, Table 3.2). In addition, protein expression of Slug and Snail did not differ significantly in relation to ER status (not shown), as previously reported using immunohistochemical methods [244]. Interestingly, immunofluorescence images show that Snail expression is predominantly nuclear, whilst Slug expression is predominantly cytoplasmic (Figure 3.6a-b). Cytoplasmic expression of EMT transcription repressors including Slug has been reported but the functional significance of this remains unknown [219]. Twist expression could not be reliably reported due to the poor specificity of commercially available antibodies. In addition, Zeb1 expression was purely stromal in set 1 and therefore not investigated in set 2 (Figure 3.8). These findings strengthen the observation that EMT events are relatively uncommon, do not occur uniformly and are uncoupled from E-cadherin loss.

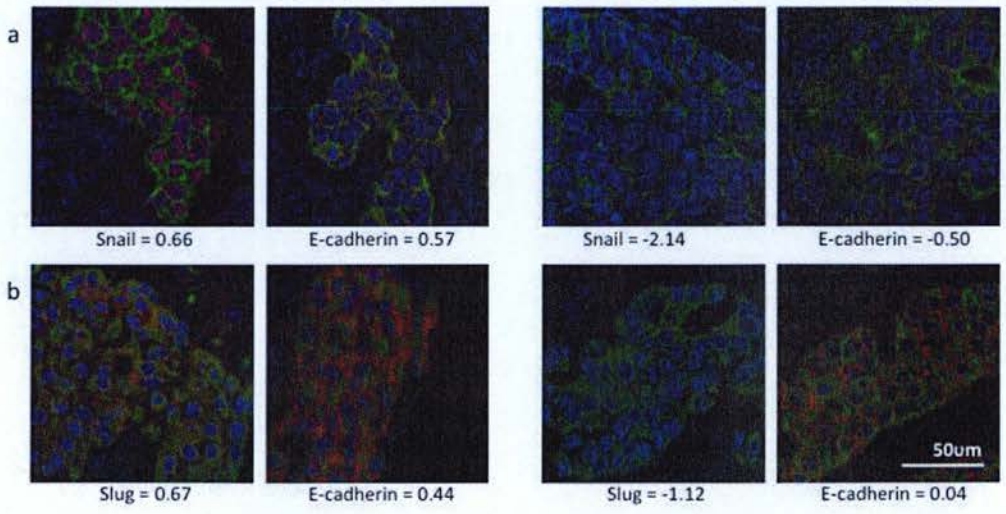


Figure 3.6. High expression of Snail and Slug in human breast cancer. Representative immunofluorescence images illustrating tumours with high (*left panels*) and low (*right panels*) expression of (a) Snail and (b) Slug. Corresponding E-cadherin expression is shown in each case. Note the clear nuclear expression of Snail whilst Slug is predominantly cytoplasmic. Bar = 50 μm. Cytokeratin = green; DAPI = blue; Snail/Slug = red

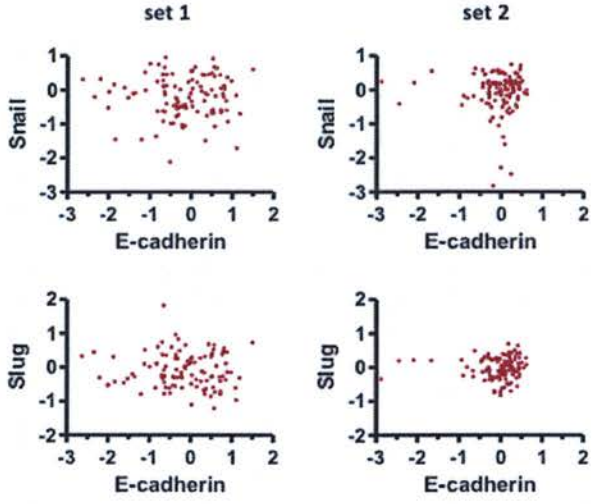


Figure 3.7. Correlations between transcriptional repressors and E-cadherin. Plots comparing Snail and Slug protein expression levels to E-cadherin are shown for both sets.

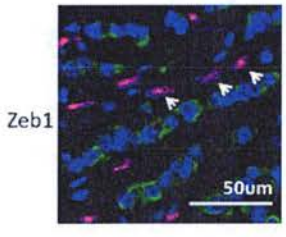


Figure 3.8. Zeb1 staining. Representative immunofluorescence image illustrating stromal staining of Zeb1 (arrows).

Correlations between markers of EMT support two distinct programmes in human breast cancer

Protein markers of EMT with reproducible, significant correlations across both patient sets are shown in Table 3.1. Univariate analysis showed no reproducibly significant correlations between the markers of EMT investigated in this study and patient survival (Table 3.3). The expression of E-cadherin significantly correlates with β-catenin, which in turn correlates with N-cadherin. The expression of vimentin correlates with Snail, which in turn correlates with Slug (Figure 3.9, Table 3.1). These correlations may represent two distinct programmes that are expressed in human breast cancers. These programmes appear related, but not identical, to EMT as described *in vitro*.

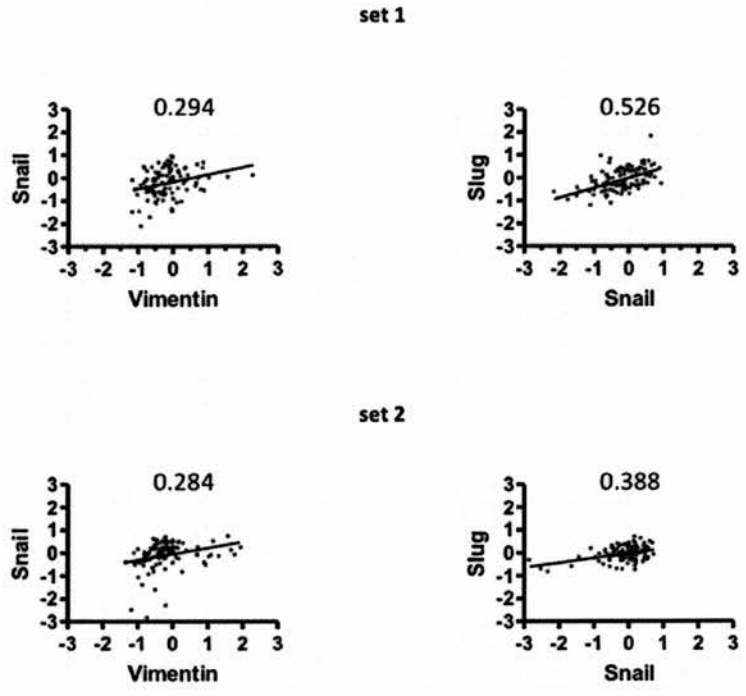


Figure 3.9. Correlations between protein markers of EMT suggest a transcriptionally driven programme in human breast cancers. Plots illustrating the reproducible correlations between vimentin, Snail and Slug as identified in Table 3.1. Correlation coefficients, significant at the 0.01 level, are shown alongside each graph. Both sets are shown with linear regression.

Table 3.1. Significant correlations between protein markers of EMT. Significant correlations, reproducible across both TMA sets, are seen between some protein markers of EMT. Upper row = set 1; bottom row = set 2. * Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed). Cyto = cytoplasmic; nuc = nuclear.

E-cadherin	0.476** 0.749**	β-catenin (cyto)
E-cadherin	0.501** 0.724**	β-catenin (nuc)
N-cadherin	0.274** 0.247**	β-catenin (cyto)
N-cadherin	0.262** 0.193*	β-catenin (nuc)
β-catenin (cyto)	0.944** 0.947**	β-catenin (nuc)
Vimentin	0.294** 0.284**	Snail
Snail	0.526** 0.388**	Slug

Table 3.2. Correlations between all examined markers of EMT in set 1 and set 2. Upper row = set1; bottom row = set 2. * Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed). Cyto = cytoplasmic; nuc = nuclear.

	β-catenin (nuc)	β-catenin (cyto)	Zeb1	Snail	Slug	E-cadherin	Vimentin	Fibronectin	N-cadherin
β-catenin (nuc)	- -	.944** .947**	.439** n/a	.376** .169	.251** .079	.501** .724**	.234* .158	.195* .070	.262** .193*
β-catenin (cyto)	.944** .947**	- -	.376** n/a	.355** .176	.240* .019	.476** .749**	.189 .194*	.162 -.014	.274** .247**
Zeb1	.439** n/a	.376** n/a	- -	.573** n/a	.586** n/a	.093 n/a	.348** n/a	.034 n/a	.328** n/a
Snail	.376** .169	.355** .176	.573** n/a	- -	.526** .388**	.000 -.018	.294** .284**	.003 -.092	.176 .108
Slug	.251** .079	.240* .019	.586** n/a	.526** .388**	- -	-.040 .009	.317** -.077	-.078 -.005	.311** .074
E-cadherin	.501** .724**	.476** .749**	.093 n/a	.000 -.018	-.040 .009	- -	.111 .064	.240* -.028	-.204* .126
Vimentin	.234* .158	.189 .194*	.348** n/a	.294** .284**	.317** -.077	.111 .064	- -	.015 .124	.071 .207*
Fibronectin	.195* .070	.162 -.014	.034 n/a	.003 -.092	-.078 -.005	.240* -.028	.015 .124	- -	.090 -.197*
N-cadherin	.262** .193**	.274** .247**	.328** n/a	.176 .108	.311** .074	-.204* .126	.071 .207*	.090 -.197*	- -

Table 3.3. Markers of EMT and patient survival. Univariate analysis showed no reproducible, significant correlations between markers of EMT and patient survival. Two statistical correction (Miller-Siegmund and Monte-Carlo) were used to correct type I statistical error that results from multiple testing for cutpoints in Kaplan-Meirer survival analysis. Upper row = set 1; bottom row = set 2.

Target	P, uncorrected	P, corrected	MonteCarlo corrected
E-cadherin	0.03 0.06	0.38 0.53	0.35 0.54
N-cadherin	0.08 0.002	0.63 0.057	0.49 0.05
β-catenin (nuc)	0.12 0.18	1 1	0.67 0.84
β-catenin (cyto)	0.12 0.18	1 1	0.73 0.79
Snail	0.02 0.01	0.23 0.21	0.18 0.22
Slug	0.13 0.18	1 1	0.79 0.81
Vimentin	0.05 0.07	0.46 0.61	0.40 0.64
Fibronectin	0.05 0.11	0.46 1	0.39 0.64
Zeb1	0.23 n/a	1 n/a	0.93 n/a

Chapter 4: The acquisition of a WT1-related EMT phenotype in LN metastases correlates with poor clinical outcome in human breast cancer

Introduction

WT1 in human breast carcinoma

WT1 (Wilms' tumour 1) encodes a transcriptional regulatory protein involved in the normal development of multiple tissues including the kidneys, spleen and heart [245]. Whilst several transcriptional modifications can occur, WT1 mRNA is subject to two alternative splicing events involving exon 5 (a 17 amino acid sequence) and KTS (a three amino acid sequence found between the third and fourth zinc finger domains). Alternative splicing at these two sites results in the expression of four predominant isoforms, including or excluding exon 5 and KTS respectively [246] (Figure 4.1). *WT1* was originally described as a tumour suppressor in the relatively rare paediatric kidney malignancy Wilms' tumour, following the identification of *WT1* mutations in 15% of sporadic cases [247]. However, overexpression of the wild-type WT1 has been reported in a variety of human cancers, including a proportion of Wilms' tumour cases, leading to the hypothesis that WT1 can function as a tumour suppressor or oncogene in a context dependent manner [248].

Current clinical trials have shown that WT1 is a potential molecular target for cancer immunotherapy [249]. A phase I trial reported clinical responses, defined as a decrease in tumour size or tumour marker expression, in 2/2 patients with advanced breast cancer [250]. Similar responses have been seen in lung and kidney [251, 252]. Therefore, understanding the role of WT1 and its expression patterns in cancer subtypes is clinically relevant. Whilst mutations in WT1 do not appear to be a feature of breast cancer [253], the role of wild-type WT1 remains to be clarified.

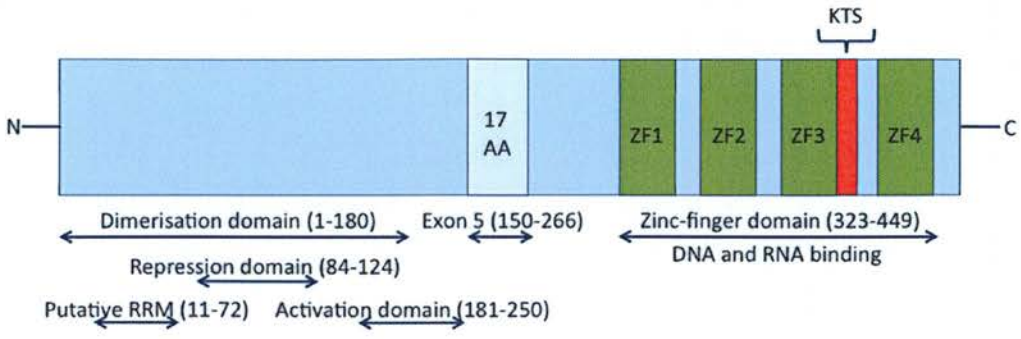


Figure 4.1. A schematic representation of the WT1 protein. The N-terminal comprises various domains: a dimerisation domain; a transcriptional activation and repression domain and a putative RNA recognition motif (RRM). Alternative splicing events involving exon 5 (a 17 amino acid sequence shown as 17AA) and KTS (a three amino acid sequence between the 3rd and 4th zinc finger domains ZF3 and ZF4 shown in red) generates the four predominant WT1 isoforms. Adapted from [254].

Using immunohistochemical methods, an early study examining WT1 expression in human breast tissues reported WT1 expression in normal breast. The complex expression patterns observed, with WT1 more strongly expressed in epithelial cells with characteristics of less differentiated cells, were consistent with a role for WT1 in developmental regulation. In addition, 15/21 breast tumours examined were classified as WT1 negative (although it is important to note that negativity was defined as less than 50% of cells staining for WT1) [255]. These findings were interpreted as indicating that WT1 has a tumour suppressor role in breast cancer and were supported by the subsequent finding that WT1 suppresses the growth of the breast cancer cell line MDA-MB-231 *in vitro* [256]. However, another group reported absent expression of WT1 in normal breast epithelium (19/20) and over expression in breast cancer (27/31) using RT-PCR [257]. Subsequently, WT1 expression was shown to correlate with poor clinical outcome in human breast cancer [258] and rates of proliferation in additional breast cancer cell lines [259]. WT1 expression according to breast carcinoma subtype has not been extensively examined. However, taken together, these studies support an oncogenic role for WT1 in breast cancer.

WT1 is a potential novel regulator of EMT

Multiple WT1 isoforms are variably expressed in individual cancers [257] and could explain, at least in part, the conflicting reports regarding the role of WT1. To examine this possibility, an additional study examined the effects of expressing distinct WT1 isoforms in a non-transformed mammary epithelial cell line [260]. Interestingly, expression of an isoform containing both exon 5 and the KTS insert (+Ex5/+KTS) induced morphological changes consistent with EMT. This was accompanied by a redistribution of E-cadherin from the cell membrane to the cytoplasm and up-regulation of vimentin. In contrast, an isoform lacking both these sequences (-Ex5/-KTS) was associated with impaired proliferation and induction of G2-cell cycle arrest. These results suggest that the ratio of expression of different isoforms may be critical in determining the role of WT1 in cancer progression. Importantly, this study also raises the possibility that WT1 may be a regulator of EMT. An association between WT1 and EMT has been described in additional

studies [261, 262]. The transcription factors Slug and Snail have an established role in EMT [76]. Genome-wide expression profiling has identified Slug as a direct target of WT1. In addition, Slug and WT1 co-express in the embryonic kidney [261]. A second study showed that WT1 directly regulates Snail and E-cadherin, and possibly Slug, in cardiovascular progenitor and embryonic stem cells [262].

Aims

Quantitative immunofluorescence was used to examine the expression of WT1 in a cohort of breast IDC-NST. The relationship of WT1 with the EMT-related proteins E-cadherin and β-catenin was examined. In addition, the hypothesis that WT1 might regulate an EMT programme in breast cancer involving the transcriptional repressors Snail and Slug was investigated. This analysis was extended to paired LN metastases.

Results

Human breast cancer has a wide range of WT1 expression

There is some controversy regarding WT1 expression patterns in human breast cancer. However, increasing evidence suggests that WT1 is over expressed in at least a proportion of these and correlates with poor clinical outcome [255, 257, 258]. Therefore, WT1 expression was examined in a cohort of IDC-NST by quantitative analysis. This analysis concentrated on the previously described set 2, a cohort enriched for large, high-grade cancers that are associated with LN metastases. An 11-fold change in WT1 expression was observed in primary tumours (Figure 4.2a). A similar range of expression was observed in LN metastases (Figure 4.2b). Interestingly, qualitative assessment of immunofluorescence images showed that WT1 staining is mostly uniform and not focal, as seen in normal development [263] (Figure 4.2c).

ER and HER2 are important clinical biomarkers in breast cancer and may correlate with WT1 expression. Higher expression of WT1 in ER negative cancers has been reported [255] and WT1 expression has been shown to induce ER-independent growth in ER positive breast cancer cell lines [264]. HER2 expression has also been associated with WT1 expression in breast cancer cell lines [265]. Therefore, the relationships between WT1, ER and HER2 were examined in primary tumours. No correlation was observed between WT1 expression and HER2 receptor status (Figure 4.2d). A trend demonstrating higher WT1 expression in ER positive tumours was observed (two-tailed t-test, $p=0.0528$; Figure 4.2e). However, a retrospective power calculation assuming 95% confidence intervals suggests that the sample size is sufficiently large to show significance if this association were real (www.dssresearch.com/KnowledgeCenter/toolkitcalculators/samplesizecalculators.aspx).

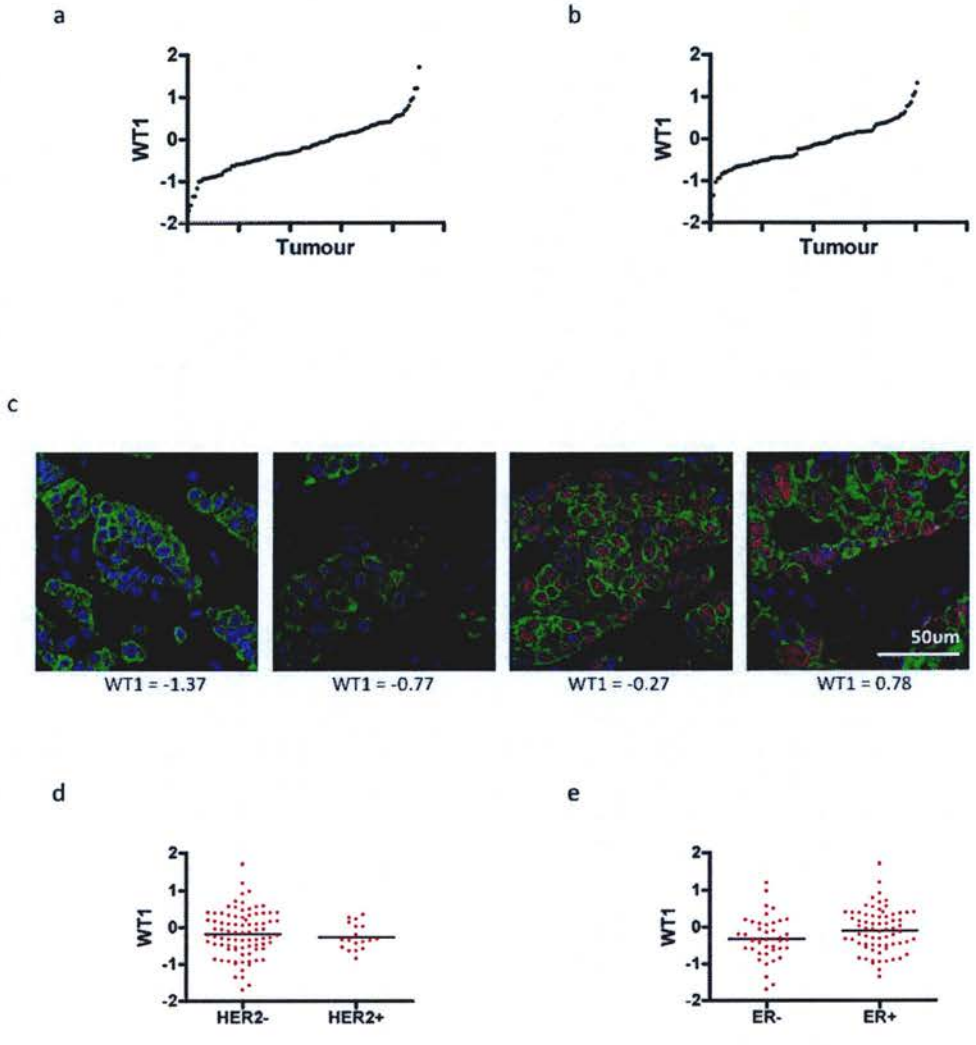


Figure 4.2. A wide range of WT1 expression is seen in both primary breast cancers and LN metastases. (a) A wide range of WT1 expression is seen in primary breast cancers. AQUA scores for each array were $\log_2$ transformed and mean centred (0 represents the adjusted mean for each individual TMA slide). The mean of three TMA replicates is shown. (b) A similar range of WT1 expression is seen in LN metastases. (c) Representative immunofluorescence images illustrating varying WT1 protein expression. Normalised AQUA scores are given (arbitrary units). Bar = 50 μm. Cytokeratin = green; DAPI = blue; WT1 = red. (d) WT1 protein expression distribution by HER2 and (e) ER receptor status.

WT1 expression is independent of E-cadherin and β-catenin in primary tumours

The relationship between WT1, E-cadherin and β-catenin is of potential interest. WT1 is emerging as a novel regulator of EMT [261, 262]. In addition, the cell junction protein E-cadherin, whose down-regulation is a key event in EMT, has been identified as a potential direct target of WT1 [266]. The nuclear translocation of β-catenin is an established, alternative regulator of EMT [71]. Importantly, WT1 has been shown to down-regulate β-catenin/TCF signalling in breast cancer cell lines [256]. Therefore, these relationships were examined in primary tumours. However, no significant correlation was seen between WT1 and E-cadherin expression (Figure 4.3a). Immunofluorescence images illustrate the co-expression of these markers in some tumours (Figure 4.3b). Similarly, no correlation was seen between WT1 and β-catenin expression (Figure 4.3c) and WT1-expressing tumours with maintained β-catenin expression at cellular junctions are seen (Figure 4.3d). Importantly, examination of correlations between WT1 and the EMT markers examined in Chapter 3 showed significant correlations with only fibronectin (Pearson's correlation = -0.319, $p<0.01$) and Slug (Pearson's correlation = 0.193, $p<0.05$).

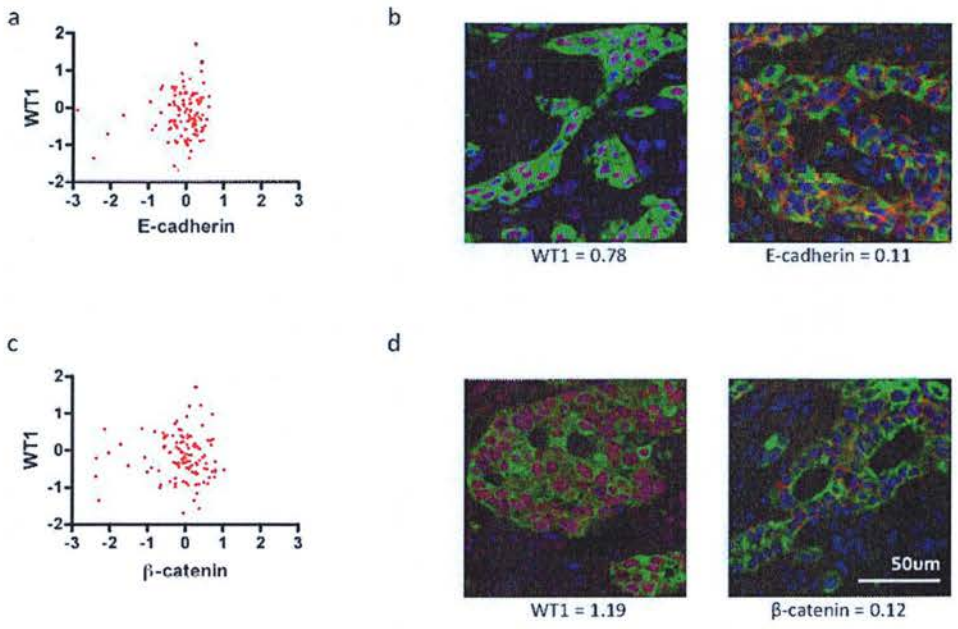


Figure 4.3. Correlations between WT1, E-cadherin and β-catenin protein expression in human breast cancer. Plots comparing E-cadherin and β-catenin protein expression to that of WT1 in primary tumours are shown in (a) and (c). Representative immunofluorescence images illustrating tumours with high expression of WT1 are shown in (b) and (d). Corresponding expression of E-cadherin and β-catenin is shown. Bar = 50 μm. Cytokeratin = green; DAPI = blue; WT1/E-cadherin/β-catenin = red.

A WT1-related EMT phenotype comprising Snail and Slug is seen in human breast cancer

Recent studies suggest that WT1 may regulate an EMT programme involving Snail and Slug [261, 262]. To test this hypothesis in human breast cancer, protein expression levels of WT1, E-cadherin, Snail and Slug were measured in primary tumours. Unsupervised hierarchical clustering of protein expression data from primary tumours revealed two principle clusters. Cluster 1 (shown in red) was characterised by down-regulated expression of E-cadherin and predominantly up-regulated expression of WT1, Snail and Slug. Cluster 2 (shown in blue) was characterised by high expression of E-cadherin and low/mixed expression of WT1, Snail and Slug (Figure 4.4a). The pattern of protein expression seen in cluster 1 is highly suggestive of an EMT phenotype whilst that seen in cluster 2 suggests an epithelial phenotype.

As the expression of protein markers is known to alter between primary and LN metastases in breast cancer with clinical and biological implications [196], paired LN metastases were similarly examined. Interestingly, similar pattern of protein expression was seen in paired LN metastases with an equivalent cluster 1 and cluster 2 (Figure 4.5a). However, in both primary and LN tissues, separation according to cluster showed no correlation with clinical outcome (Figure 4.4b and 4.5b). However, retrospective power calculations (assuming 95% confidence intervals) show that this data is significantly underpowered. Estimated total sample sizes of 1146 and 1019 cases would be required to determine whether the trends seen in Figures 4.4b and 4.5b are real (www.stattools.net/SSizSurvival_Pgm.php). Mean survival times by cluster for primary tumours and LN metastases are given in the respective figure legends.

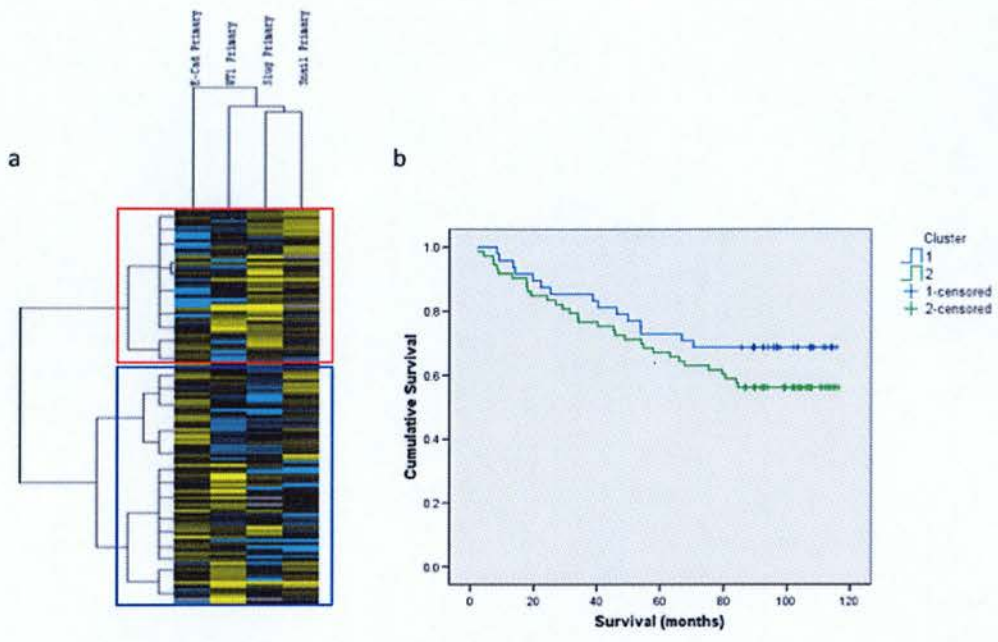


Figure 4.4. Unsupervised hierarchical clustering of protein expression data in primary breast tumours identifies a WT1-related EMT phenotype. (a) Two main clusters are identified. Down-regulated expression of E-cadherin and predominantly up-regulated expression of WT1, Snail and Slug characterises cluster 1 (outlined in red). High expression of E-cadherin and low/mixed expression of WT1, Snail and Slug characterises cluster 2 (outlined in blue). Cluster 1 has a mean survival of 90.7 months (confidence interval (79.7–101.7)) versus 82.5 months for cluster 2 (confidence interval (72.9–92.1)). Relative up-regulation of protein expression = yellow; relative down-regulation of protein expression = blue. (b) Separation according to cluster is not correlated with clinical outcome.

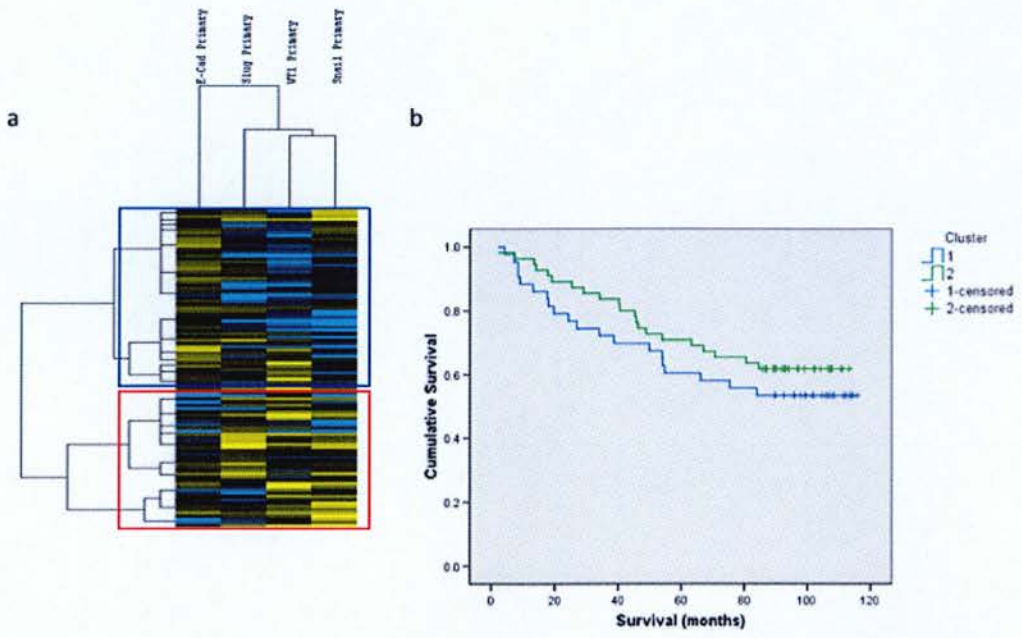


Figure 4.5. Unsupervised hierarchical clustering of protein expression data in paired LN metastases identifies a WT1-related EMT phenotype. (a) As in primary tumours, two main clusters are identified. Down-regulated expression of E-cadherin and predominantly up-regulated expression of WT1, Snail and Slug characterises cluster 1 (outlined in red). High expression of E-cadherin and low/mixed expression of WT1, Snail and Slug characterises cluster 2 (outlined in blue). Cluster 1 has a mean survival of 77.5 months (confidence interval (64.3-90.7)) versus 85.5 months for cluster 2 (confidence interval (75.4-95.6)). Relative up-regulation of protein expression = yellow; relative down-regulation of protein expression = blue. (b) Separation according to cluster is not correlated with clinical outcome.

Acquisition of a WT1-related EMT phenotype in LN metastases predicts poor clinical outcome

EMT is a dynamic process and reversion to an epithelial phenotype (mesenchymal to epithelial transition or MET) is a feature of both normal development and cancer progression [68]. The observation that distant metastases are mainly composed of cells with an epithelial phenotype closely resembling that of the primary tumour supports this [129, 144, 267]. Therefore, an ability to move between EMT and MET phenotypes depending on the microenvironment may play a role in determining the aggressiveness of tumour cells. Consequently, the change in cluster expression between primary tumours and paired LN metastases and correlation with outcome was examined.

Within the framework of the clusters identified here, four possibilities exist. Tumours may remain in the same cluster type, either maintaining an EMT phenotype (those that remain in cluster 1) or maintaining an epithelial phenotype (those that remain in cluster 2). Alternatively, tumours may change cluster as they invade locoregional LNs, either acquiring an EMT phenotype (those that change from cluster 2 to cluster 1) or reverting to an epithelial phenotype which can be equated with MET (those that change from cluster 1 to cluster 2). Importantly, tumours that acquired an EMT phenotype had the worst prognosis (Figure 4.6) and were close to statistical significance when compared with tumours that maintained the same phenotype, either EMT ($p = 0.058$) or epithelial ($p = 0.059$) (see Tables 4.1 and 4.2). Tumours that acquired an EMT phenotype also had a worse prognosis than those undergoing MET but this did not reach significance. It is important to note that patient numbers in this study are limited and retrospective power calculation (assuming 95% confidence intervals) suggests that a total sample size of at least 206 cases would be required to determine whether the trend shown in Figure 4.6 is real (www.stattools.net/SSizSurvival_Pgm.php). When those tumours that maintained the same phenotype are combined and then compared to those that acquired an EMT phenotype, a clearly significant difference in prognosis is seen ($p = 0.024$; Tables 4.3 and 4.4). These results suggest that acquisition of an EMT phenotype as tumour cells invade locoregional LNs confers the greatest aggressiveness to tumour cells.

Importantly, univariate analysis of individual markers expressed in LN metastases was not predictive of outcome (Table 4.5). In addition, change in cluster showed no relation to tumour grade, size, number of positive nodes and ER or HER2 receptor status (not shown).

Immunofluorescence images illustrating the change in expression in E-cadherin, WT1, Snail and Slug in a tumour that acquires this phenotype are shown (Figure 4.7). H&E stains for paired primary and LN tissues were examined by two pathologists (Dr. J. Thomas and Dr. D. Faratian) but no clear morphological subgroups in relation to change in cluster were identified (Figure 4.8).

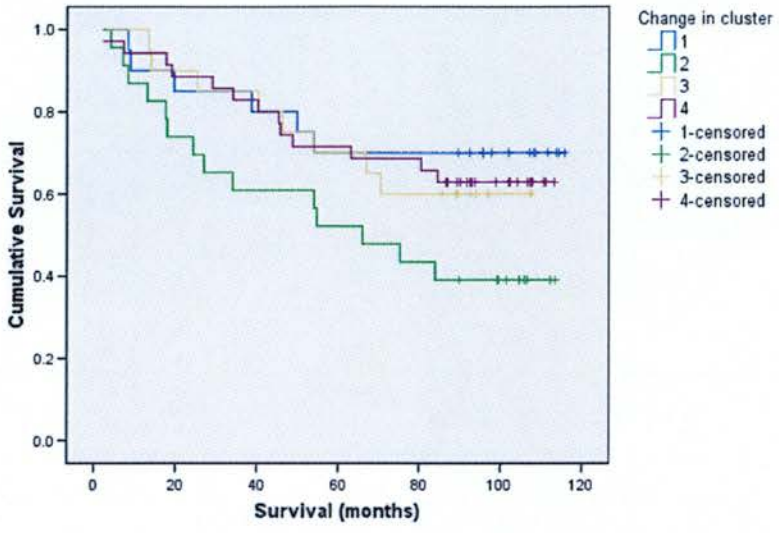


Figure 4.6. Change in cluster expression between primary tumours and paired LN metastases correlates with clinical outcome. The change in cluster expression between primary tumours and paired LN metastases is correlated with outcome. Cluster 1>1 = blue; cluster 2>2 = purple; cluster 1>2 = yellow; cluster 2>1 = green.

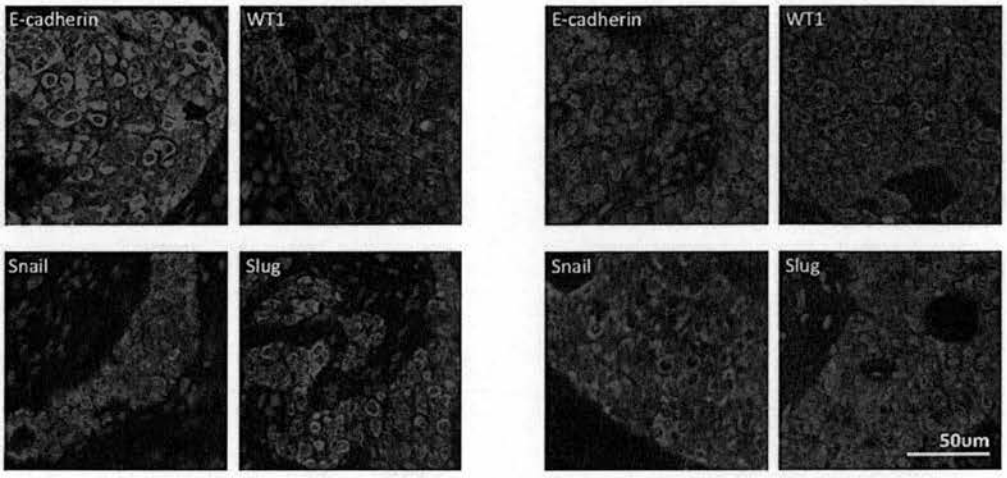


Figure 4.7. A subgroup of human breast tumours acquire an EMT phenotype as they invade locoregional LNs. Immunofluorescent staining is shown for E-cadherin, WT1, Snail and Slug protein expression (*left panels,* primary tumour; *right panels,* LN metastases). Note the down-regulation of E-cadherin expression and the up-regulation of WT1, Snail and Slug as this tumour invades locoregional LNs. Cell morphology appears unchanged. Bar = 50 μm. Cytokeratin = green; DAPI = blue; E-cadherin/WT1/Snail/Slug = red.

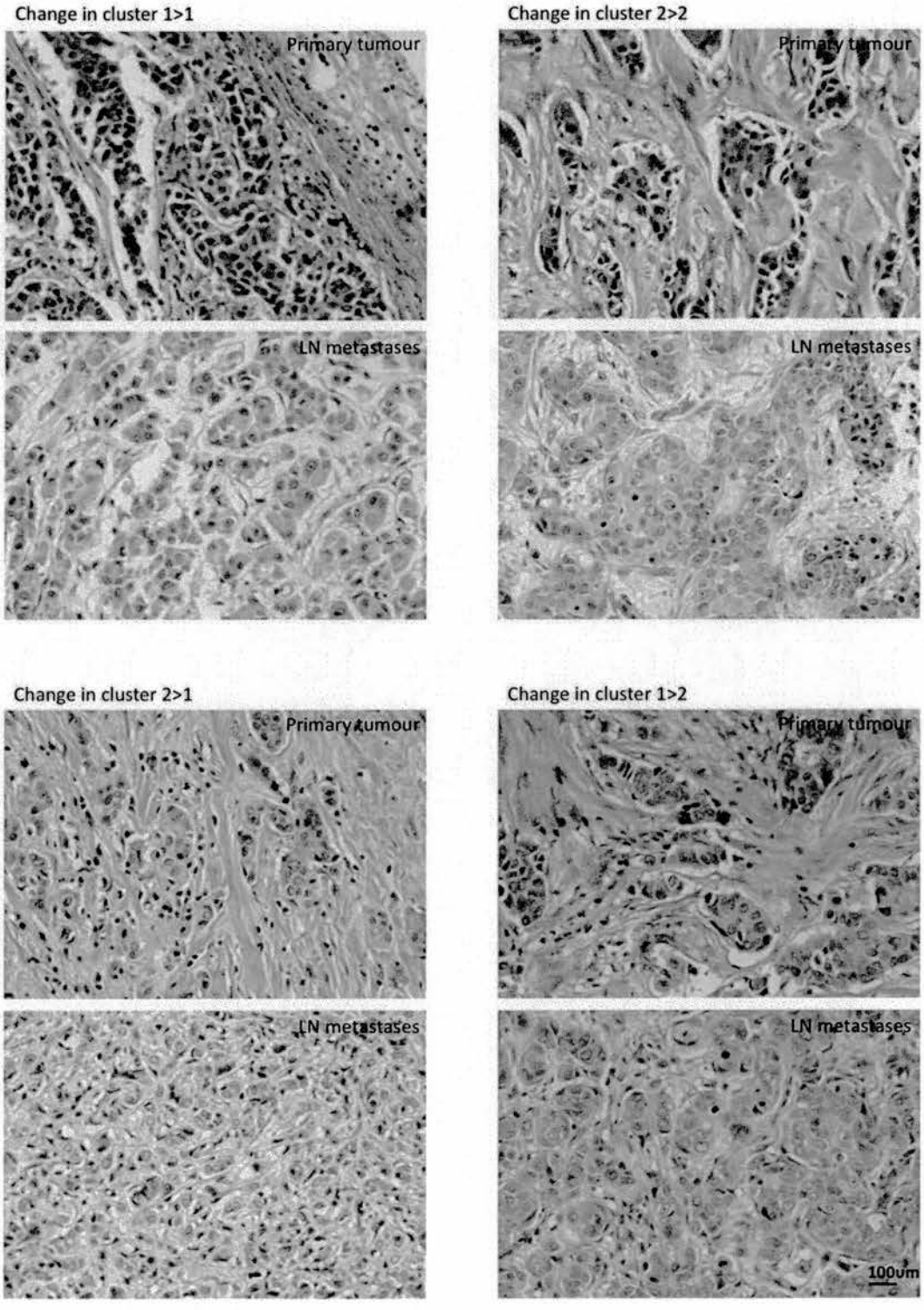


Figure 4.8. H&E stains according to change in cluster. A primary tumour (*top panel*) with its paired LN metastases (*bottom panel*) is shown for each change in cluster. Bars = 100 μm.

Table 4.1. Pairwise comparisons according to change in phenotype.

Change in cluster	2>1	
	Chi-Square	Significance
1>1	3.593	0.058
2>2	3.569	0.059
1>2	2.039	0.153

Table 4.2. Means and medians for survival time according to change in phenotype. (a) estimation is limited to the largest survival time if it is censored.

Change in cluster	Mean(a)				Median			
	Estimate	Std. Error	95% Conf. Int.		Estimate	Std. Error	95% Conf. Int.	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
1>1	90.07	9.09	72.26	107.88	-	-	-	-
2>2	86.00	6.51	73.24	98.76	-	-	-	-
1>2	81.19	7.84	65.80	96.54	-	-	-	-
2>1	65.64	9.05	47.90	83.37	66.10	17.01	32.76	99.44
Overall	82.85	4.23	74.57	91.14	-	-	-	-

Table 4.3. Pairwise comparisons according to change in phenotype when tumours maintaining the same phenotype are combined.

Change in cluster	2>1	
	Chi-Square	Significance
1>1 and 2>2	5.091	0.024
1>2	2.039	0.153

Table 4.4. Means and medians for survival time according to change in phenotype when tumours maintaining the same phenotype are combined. (a) Estimation is limited to the largest survival time if it is censored.

Change in cluster	Mean(a)				Median			
	Estimate	Std. Error	95% Conf. Int.		Estimate	Std. Error	95% Conf. Int.	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
1>1 and 2>2	88.51	5.40	77.93	99.09	-	-	-	-
1>2	81.19	7.84	65.80	96.54	-	-	-	-
2>1	65.64	9.05	47.90	83.37	66.10	17.01	32.76	99.44
Overall	82.85	4.23	74.57	91.14	-	-	-	-

Table 4.5. Markers of EMT in LN metastases and patient survival. Univariate analysis showed no significant correlations between markers of EMT in LN metastases and patient survival.

Target	P, uncorrected	P, corrected	MonteCarlo corrected
E-cadherin	0.023	0.303	0.300
Snail	0.054	0.517	0.450
Slug	0.0049	0.93	0.09
WT1	0.038	0.414	0.370

Chapter 5: Developing a three-dimensional (3D) model for the investigation of invasion

Introduction

The acquisition of an invasive phenotype and the ability to breach the basement membrane is a prerequisite for metastasis

Breast cancer related deaths are primarily due to metastatic progression [8]. Understanding the mechanisms that underlie this multistep process is essential to improving clinical outcome. The transformation of normal breast epithelial cells to metastatic cancer is the result of multiple epigenetic and genetic changes, in the context of deregulated interactions with the microenvironment. During this process, control of proliferation, cell survival, migration and differentiation is lost. The acquisition of an invasive phenotype, and in particular the ability to breach the basement membrane, is a critical event in cancer progression and a prerequisite for metastasis. In addition, the acquisition of an invasive phenotype is closely associated with EMT [68]. Having breached the basement membrane, cells may then access the circulation and lymphatics and attempt to establish distant tumour foci [268-270].

Cell culture models and cell lines for the investigation of invasion

Culture models of human breast cells provide an opportunity to investigate metastatic progression and in particular, migration and invasion. In addition, they allow new, targeted therapies to be tested in a pre-clinical setting. Although closely related to the tissue of origin, primary cells from human tissues have limited use *in vitro*. These cells do not divide indefinitely in culture and are not easily manipulated at the gene expression level. In addition, heterogeneity between samples from different individuals raises issues around interpretation of results and reproducibility [271]. Consequently, a large series of breast cancer cell lines have been developed for use in culture, resolving many of these issues. However, it is important to be aware that multiple variants of a single line can develop, making comparison between studies potentially difficult [272].

The value of cell culture models depends on a high degree of similarity between cell lines and human cancers. A number of studies have compared the genomic and transcriptional profiles of breast cancer cell lines with those of primary cancers [204, 273, 274]. In the largest of these studies Neve and colleagues compared a panel of 51 cell lines to 145 primary cancers [204]. Importantly, most genomic and transcriptional abnormalities characteristic of primary cancers were identified in cell lines. In addition, this study showed that a panel of cell lines can be used to identify molecular markers that predict response to targeted treatment. However, some differences between cell lines and primary cancers were also apparent. The genome copy number abnormalities seen between luminal and basal-subtype tumours were not present in corresponding cell lines and not all cancer subtypes were represented by the cell line panel. Importantly, luminal subtypes A and B, associated with distinct prognoses, were not distinguished. These studies suggest that cell lines are valuable tools in cell culture but results may not always be directly applicable to human cancers.

Normal breast architecture can be recapitulated in 3D culture

Many studies, including the important work of Neve and colleagues, have been carried out using cell lines cultured on plastic [204, 273, 275]. These two-dimensional (2D) cultures lack exposure to physiological substratum and fail to recapitulate 3D tissue architecture [276, 277]. In contrast, culture of normal breast epithelial cells in 3D laminin-rich extracellular matrix, also known as Matrigel, results in a well-characterised morphogenesis process that closely resembles normal breast. These cells undergo a limited number of divisions, after which they organise into polarised, growth-arrested colonies. Similar 3D culture of cancer cells results in unlimited proliferation, absence of tissue polarity and disorganised architecture (Figure 5.1) [278, 279]. In addition, signalling pathways are known to be differentially regulated in 2D versus 3D culture [280]. Nonetheless, comparison of the gene expression profiles of cell lines cultured in 2D versus 3D has shown that individual cell lines still cluster together, independently of culture conditions. Although this indicates that 3D culture does not induce a global change in gene

expression, a group of genes encoding signal transduction proteins significantly correlated with 3D culture [281].

Aims

Taken together, these studies suggest that 3D culture is closer to the *in vivo* situation. A significant amount of work has been done in this field and models that demonstrate accelerated invasion or make use of increasingly physiological matrices have been developed [282, 283]. In addition, cell lines with increased invasive potential have been developed in order to study the transition to an invasive phenotype [284]. The movement of cells across basement membrane has also been examined with important findings regarding the role of Snail and MMPs in cancer progression [285]. However, a model that examines cells before, during and after they cross the basement membrane has not been described. Therefore, we identified and characterised a series of potentially invasive, EMT-like cell lines and developed an *in vitro* model that allows cells to be followed and examined as they cross the basement membrane.

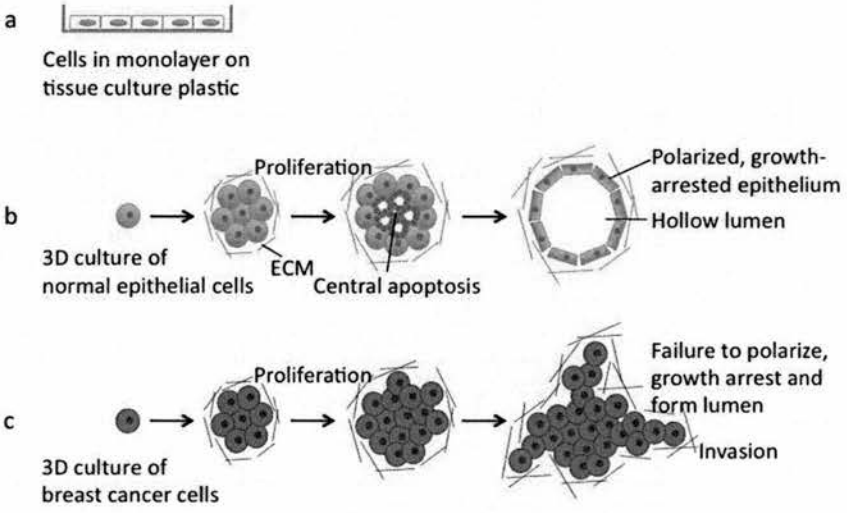


Figure 5.1. Three-dimensional culture recapitulates normal tissue architecture. (a) Two-dimensional (2D) cultures grown on plastic are exposed to a non-physiological substratum that lacks the normal components of the extracellular matrix *in vivo*. These cultures lack heterotypic cell-cell contacts and fail to recapitulate three-dimensional (3D) tissue architecture. (b) The morphology and behaviour of non-transformed epithelial cells in 3D culture more accurately mimics breast structure and function. These cells undergo a morphogenesis process similar to that seen *in vivo* in normal breast. (c) When breast cancer cell lines are grown in 3D culture, they fail to growth arrest, lack polarity, display abnormal architecture and may become invasive. Whilst this is similar to what is observed *in vivo*, not all changes occur at once in all cancer cell lines. Adapted from [271].

Results

Identification of a 9-gene claudin-low/ EMT signature derived from breast cancer models

Claudin-low breast cancers are a novel molecular subtype, identified following comparisons between mouse models of breast cancer and human cancers [51]. This subtype appears closely related to EMT, as well as to intrinsic resistance to common therapies [169]. Recent *in vitro* work has shown a clear overlap between the claudin-low phenotype and a C35-expressiong cell line system [206]. C35 (also known as C17orf37 or MGC14832) is a 12KDa membrane-anchored protein found on the HER2 amplicon and is overexpressed in ~11% of breast cancers [286]. Importantly, expression of C35 *in vitro* can induce mammary epithelial cell transformation associated with acquisition of an EMT phenotype. Increased expression of C35 was associated with increased invasion in 3D culture, down-regulation of E-cadherin expression and up-regulation of the transcription repressor Twist (Figure 5.2a-c) [206]. Corresponding AQUA scores for C35 and E-caderin protein expression are shown (Figure 5.2d). These results suggest that invading C35-expressing cells may consitute a good model for the investigation of EMT and the claudin-low phenotype.

Consequently, genes correlating with C35 expression (and consequently EMT) and those identifying the claudin-low phenotype were compared using gene expression micro-array data. The top 100 illumina probes most differentially expressed between C35 and parental cells were identified. Of these, 57 corresponded to genes examined in the work of Herschkowitz and colleagues [51]. These 57 genes were sufficient to cluster together the 13 claudin-low tumours identified and 9 of these 57 genes were shared with a 34 gene claudin-low cluster (Figure 5.3). In addition, all of the genes on the illumina array representing the 34 gene claudin-low cluster (of which there were 25) were significantly down-regulated in C35-expressing cells compared with parental cells ($p \leq 1x10^{-8}$).

The 9 genes commonly expressed between claudin-low tumours and C35-expressing cells can be grouped into functional clusters: polarity and cell-cell contact (*CDH1, CLDN7, CRB3, KRT8. TACSTD1*), protein trafficking (*MAL2, MARVELD3*) and

protease-related (*IRF6, SPINT2*) (Table 5.1). Some of these genes (*CDH1, CLDN7, TACSTD1, CRB3* and *KRT8)* have previously been reported to be down-regulated in either claudin-low cancers or in the context of EMT *in vitro* [287-291]. In addition, *SPINT2* (hepatocyte growth factor activator inhibitor-2 or HAI-2) has been shown to regulate the HGF-induced invasion of breast cancer cells *in vitro [292].*

Importantly, all 9 genes were down-regulated in both claudin-low tumours and C35-expressing cells, suggesting that a common repressive mechanism underlies this phenotype. Therefore, oligo-analysis was carried out to determine whether these genes share common regulatory elements in their promoter regions and identified a shared 7-mer (CAGGTGC/GCACCTG) [293]. This binding motif is targeted by E-box transcription repressors including members of the Snail and Zeb families. These findings raise the possibility that transcriptional repression regulates the expression of these EMT-related genes both *in vitro* and *in vivo*. Other regulatory mechanisms, such as DNA methylation, may also regulate this phenotype and evidence that these genes can be silenced by DNA methylation is present in the literature [294-298] (Table 5.1).

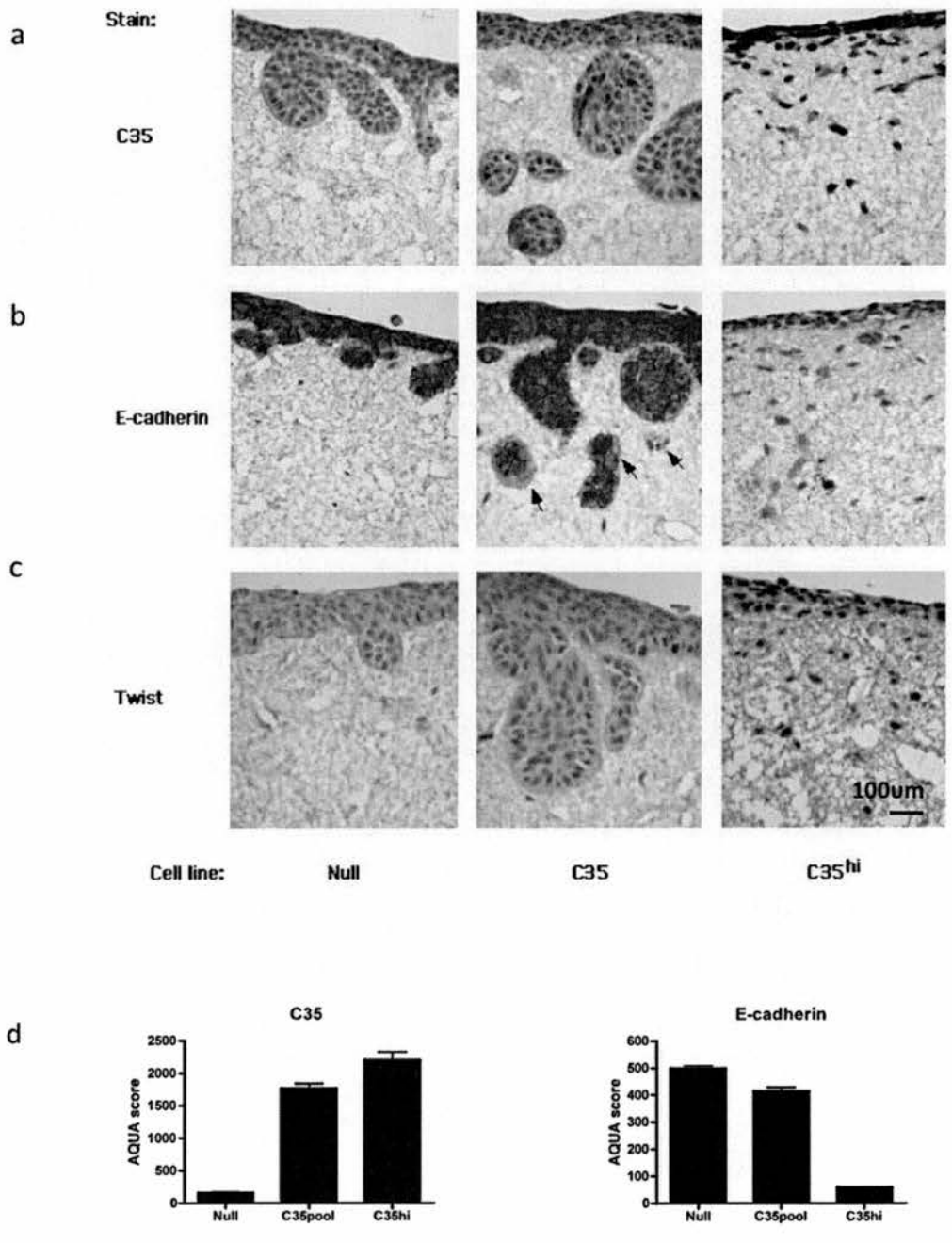


Figure 5.2. C35-expression leads to an invasive phenotype, associated with EMT. H16N-2 cells were retrovirally transfected with C35 protein. Colonies expressing i) empty vector (null, *left panels*), ii) variable levels of C35, termed C35-pool (C35, *middle panels*) and iii) high levels of C35 protein ($C35^{hi}$, *right panel*) were stained

by immunohistochemistry for C35 (a), E-cadherin (b) and Twist (c). Specific areas of E-cadherin loss are observed in the C35 pool (arrows) whilst a more general loss is observed in the $C35^{hi}$- expressing cells. Importantly, this is accompanied by up-regulation of Twist expression. Bar = 100 μm. (d) Quantification of C35 and E-cadherin by AQUA. Expression of C35 protein is inversely correlated to that of E-cadherin. The mean of three replicates is shown in all cases.

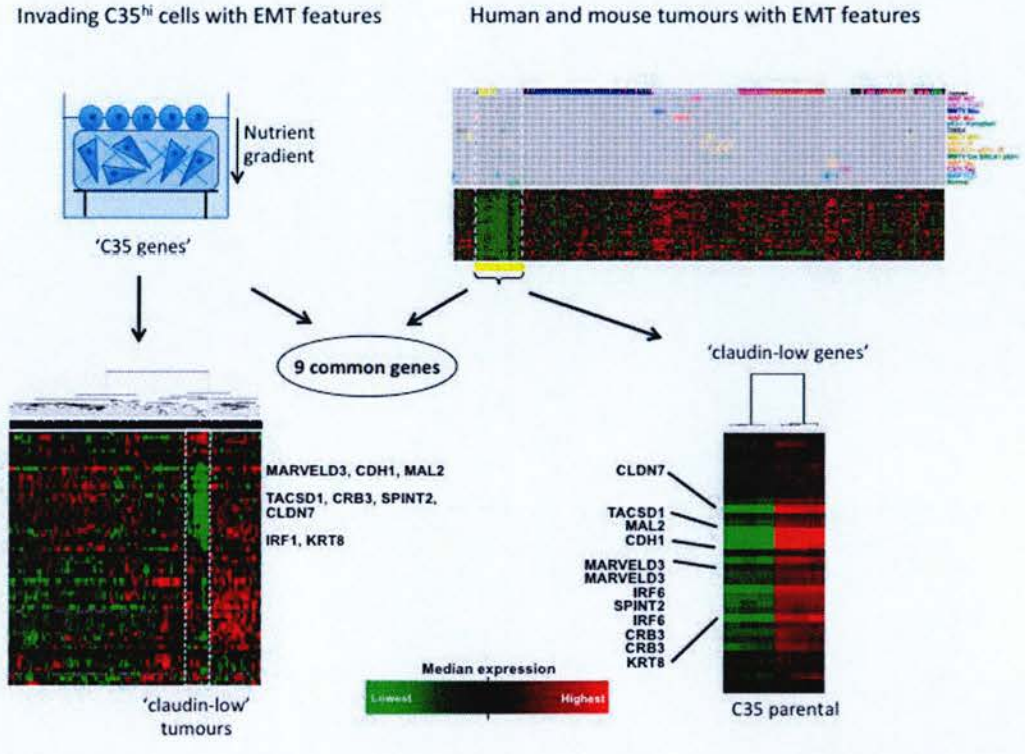
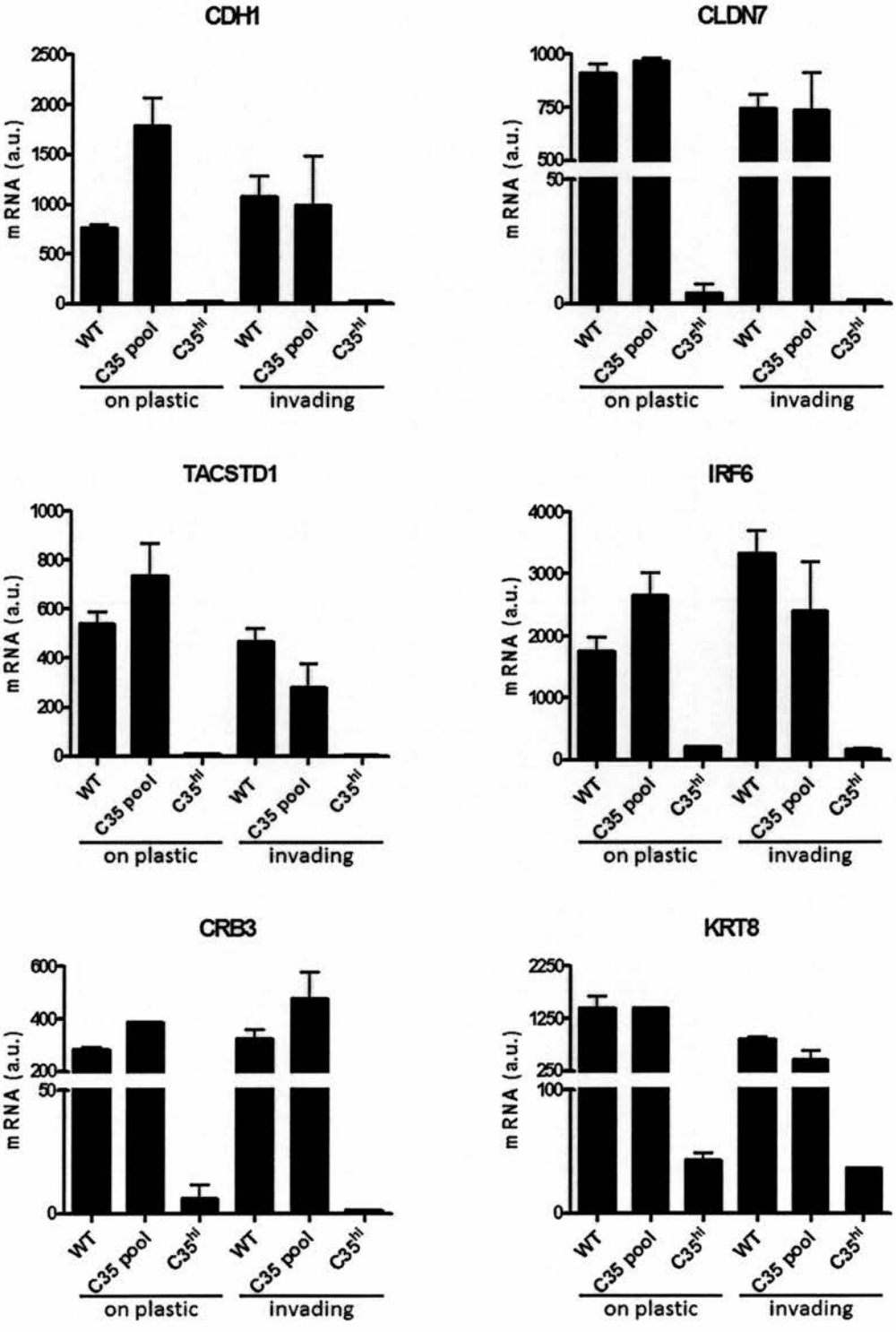


Figure 5.3. Comparison of genes correlating with C35 expression and those identifying the claudin-low phenotype identifies a 9-gene EMT signature. The top 100 illumina probes most differentially expressed between C35 and parental cells were identified, 57 of which correspond to genes examined in the Herschkowitz dataset [51]. These 57 genes were able to cluster together the 13 claudin-low tumours identified by Herschkowitz and colleagues and 9/57 are shared with a 34-gene claudin-low cluster. In addition, these 34 claudin-low genes are able to cluster together parental and C35-expressing cells. The 9 genes comprising this signature are shown in Table 5.1. For full gene lists refer to [293].

Validation of the 9-gene signature by qRT-PCR and identification of representative cell lines

The identification of a concise gene signature that reflects invasion and EMT provides a means of identifying candidate cell lines for the *in vitro* study of these processes. Quantitative RT-PCR was performed to validate the observed down-regulated expression of the 9-gene signature in the C35 model. High C35 expression was associated with the clear down-regulation of eight of these genes in comparison with parental cells (Figure 5.4). This is observed when epithelial cells are seeded on plastic and on contracted collagen lattices, on which they have the opportunity to invade. Marveld3 could not be assessed due to particularly low levels of expression. Varied expression levels are seen in the C35 pool (which contains clones with variable expression levels of C35) although expression tended to be similar to that of parental cells.

Having validated the 9-gene signature in the C35 model, a series of eight claudin-low/EMT cell lines were identified from the large set characterised by Neve and colleagues by means of gene expression micro-array data [204] (for details of these lines including culture conditions see Materials and Methods). The clear mesenchymal or spindle-shaped morphology of these cells when cultured on plastic is shown (Figure 5.5). These cells contact neighbouring cells only focally. This is in contrast to the cobblestone appearance of MCF10A cells which, in addition, maintain close contact with their neighbours. As with the C35 model, the down-regulation of the 9-gene signature in these cell lines, excluding Marveld3, was confirmed by quantitative RT-PCR (Figure 5.6). Normal HMECs were used for comparison. Culture conditions were limited to plastic for this series of PCRs.


CDH1
CLDN7
TACSTD1
IRF6
CRB3
KRT8
m RNA (a.u.)
WT
C35 pool
C35hi
on plastic
invading

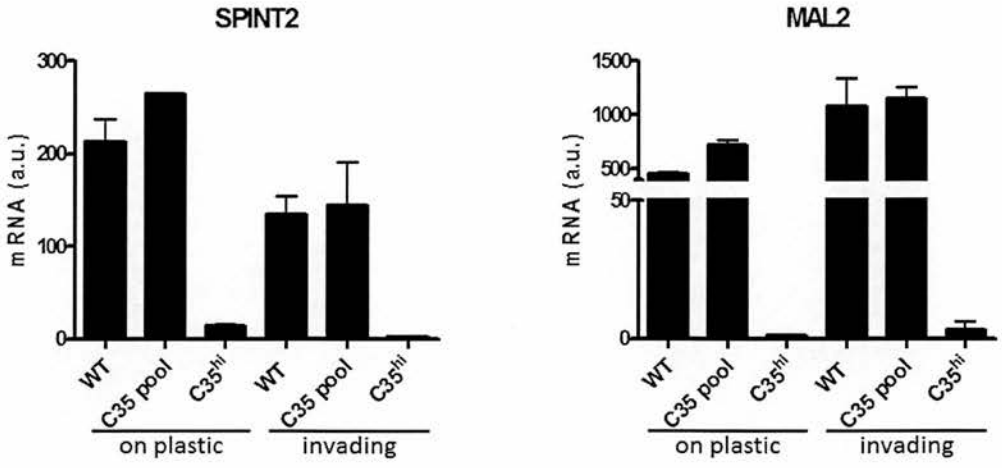


Figure 5.4. Genes down-regulated in the C35-induced, transformed phenotype. C35-induced down-regulation of CDH1, CLDN7, CRB3, KRT8, TACSTD1, IRF6, SPINT2 and MAL2 was confirmed by qRT-PCR. This is found both in cells grown on plastic and in cells grown on contracted collagen lattices with the opportunity to invade. MARVELD3 could not be assessed due to particularly low levels of expression. Biological triplicate mRNA expression data is shown for the empty vector, the C35 expressing pool and the $C35^{hi}$- expressing clone.

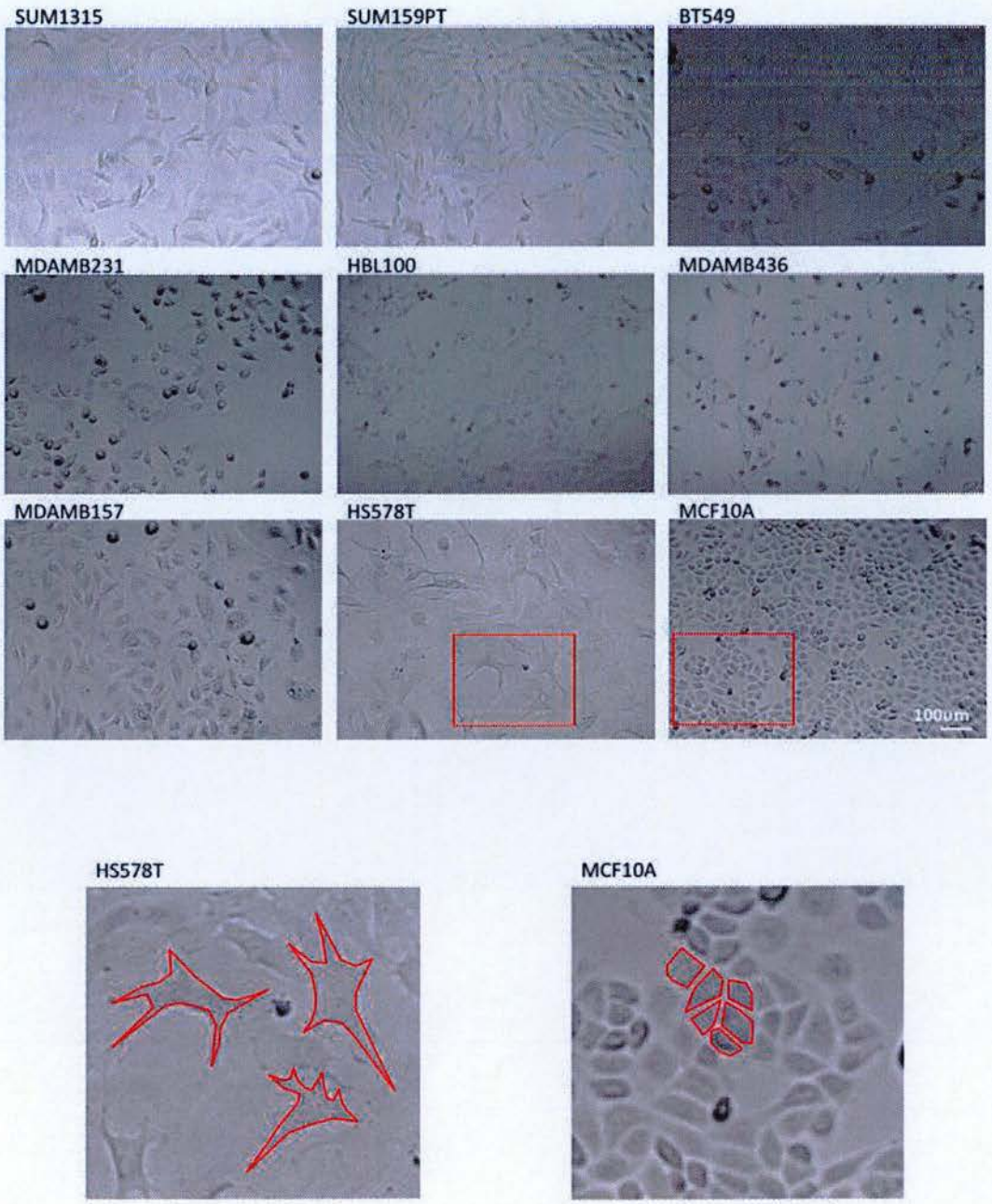
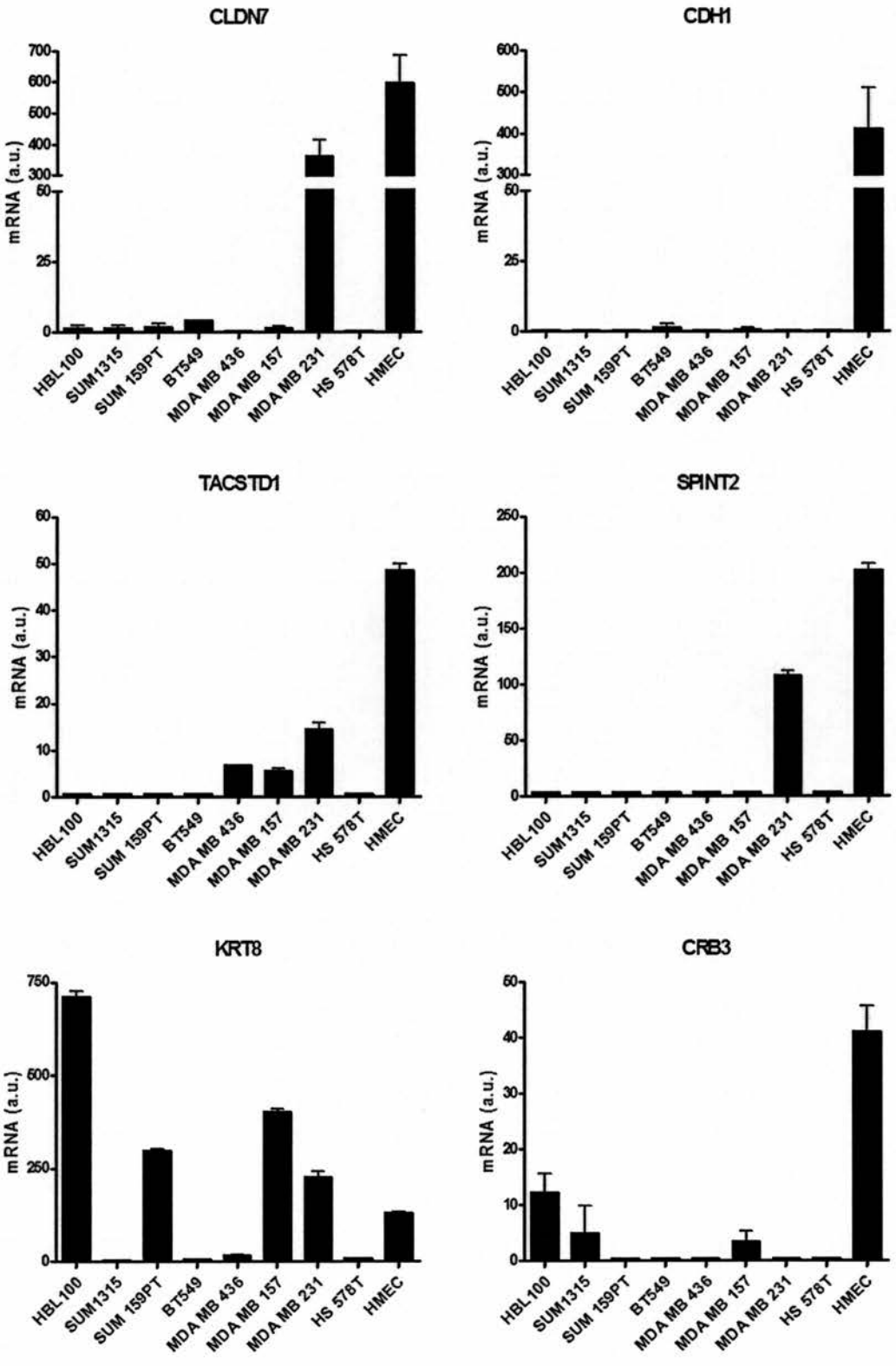


Figure 5.5. Claudin-low/EMT cell lines exhibit a mesenchymal morphology. Eight potential claudin-low/EMT cell lines were identified. Representative live microscopy images of these lines cultured on plastic are shown. Note the predominant mesenchymal morphology of these cells. The non-transformed cell line, MCF10A, shows a contrasting cobblestone morphology. Insets illustrate the contrasting morphologies and cell-cell contacts in HS578T and MCF10A cells. Bar = 100 μm.


CLDN7
CDH1
TACSTD1
SPINT2
KRT8
CRB3
mRNA (a.u.)
HBL100
SUM1315
SUM 159PT
BT549
MDA MB 436
MDA MB 157
MDA MB 231
HS 578T
HMEC

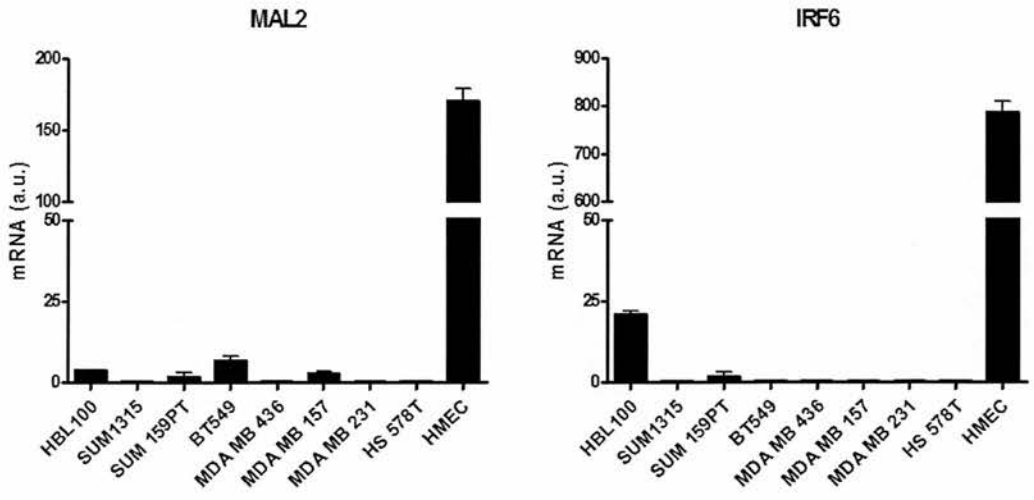


Figure 5.6. Genes down-regulated in the claudin-low/EMT cell lines. Low expression of CDH1, CLDN7, CRB3, KRT8, TACSTD1, IRF6, SPINT2 and MAL2 was confirmed by qRT-PCR in a panel of 8 cell lines. As before, MARVELD3 could not be assessed due to particularly low levels of expression. HMECs are shown for comparison and express significantly higher levels of mRNA for each gene examined. Technical triplicate mRNA expression data is shown for each line.

Claudin-low/EMT cell lines show up-regulation of key transcriptional drivers of EMT

Several transcription repressors have been identified as important drivers of EMT, both in development and in cancer progression. These include Snail, Slug, Zeb1, Zeb2 and Twist [76]. Importantly, examination of gene expression micro-array data suggests that the selected claudin-low/EMT cell lines have elevated expression of Slug, Zeb1 and Twist (Figure 5.7a). Western blots were carried out in order to investigate the expression of these and other transcription repressors at the protein level. The expression of other EMT-related proteins was also examined in order to further characterise the cell lines (Figure 5.7b). All claudin-low cell lines show low levels of E-cadherin and Claudin-7 in comparison to normal mammary epithelial cells. Most cell lines have detectable levels of Snail whilst Slug is absent only in MDA MB231 cells. Interestingly, strong expression of active β-catenin is seen in normal mammary epithelial cells. Whilst these blots are derived from cell lines cultured in 2D, this may reflect the positive correlation between β-catenin and E-cadherin seen in human tumours (Chapter 3). However, western blot gives no information about the cellular localization of these proteins, which is key to understanding their function. Most claudin-low lines variably express the mesenchymal markers N-cadherin, vimentin and fibronectin. Normal mammary epithelial cells express strong levels of fibronectin and the exact reason for this is not clear. Twist antibodies were non-specific for western blot and blots are therefore not shown. A suitable Zeb1 antibody for use in western blot also could not be found.

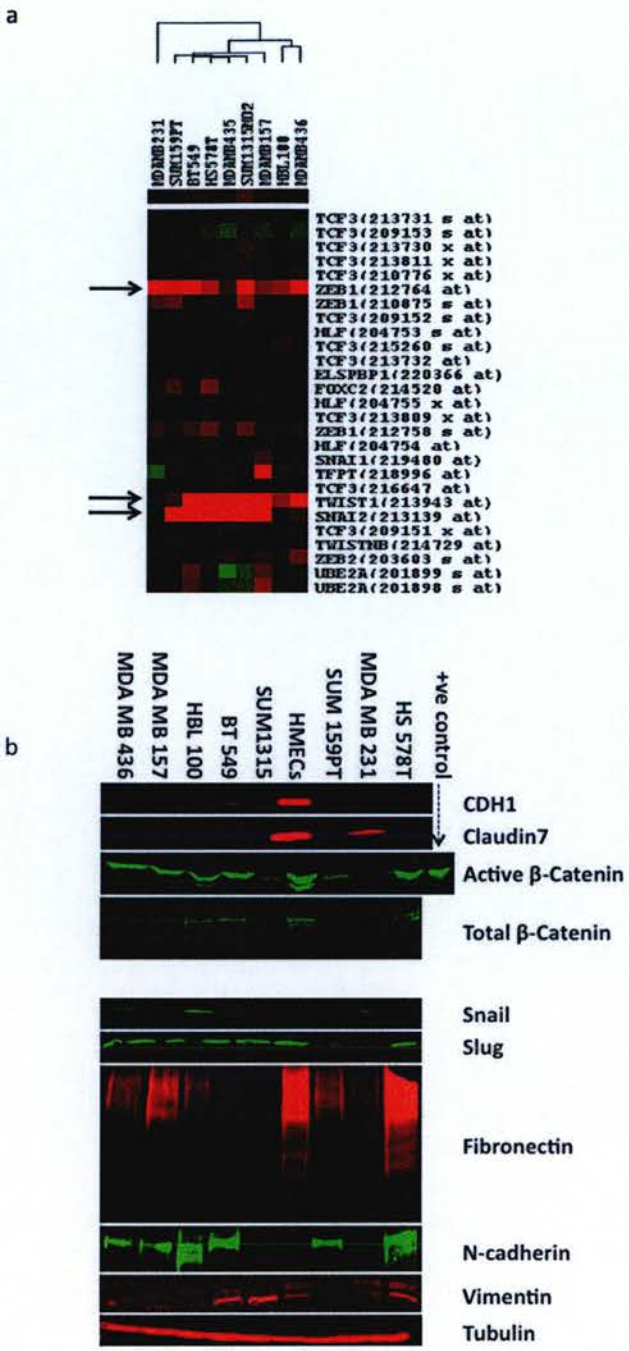


Figure 5.7. Claudin-low/EMT cell lines express key markers of EMT. (a) Examination of gene-expression micro-array data suggests that claudin-low/EMT cell lines have elevated expression of the transcriptional repressors Slug (Snail2), Zeb1 and Twist [204]. (b) Western blots examine the expression of these and other EMT-related markers at the protein level. HMECs are shown for comparison. As expected, these cells show high levels of E-cadherin and claudin-7.

A novel 3D invasion assay mimics invasion across the basement membrane

The acquisition of an invasive phenotype, and in particular the ability to breach the basement membrane, is a critical event in cancer progression. A 3D model that attempts to mimic this process was developed. Epithelial cells were embedded in laminin-rich, basement membrane-like Matrigel to generate a cell 'plug' which was subsequently embedded in collagen to mimic surrounding extracellular matrix (Figure 5.8). This model potentially generates a three-stage assay that allows investigation of cells i) contained by basement membrane, ii) as they invade across basement membrane and iii) as they invade more distally into surrounding collagen/ECM. In addition, the movement of cells in a horizontal plane can be easily followed by light microscopy, in contrast to the movement of cells in a vertical plane that occurs with the collagen-based 'on top' assay [206].

Three claudin-low/EMT lines (HBL100, HS578T and SUM 159PT) demonstrated clear and reproducible invasion in this novel assay. Importantly, all three lines adopt a round morphology when embedded in Matrigel (day 0), versus the predominantly elongated morphology that is seen in collagen. By day 7, the HBL100 and HS578T cells have reverted to an elongated morphology, indistinguishable from that seen in collagen, and are invading across basement membrane and into surrounding collagen. In contrast, many SUM159PT cells retain a round morphology, accompanied by delayed invasion. By day 14, SUM159PT cells appear to have overcome this inhibition and many elongated cells are now seen leaving the Matrigel plug (Figure 5.9). H&E staining confirms these morphological observations (Figure 5.10). MCF10A cells (an untransformed line) were also investigated in this assay and show no invasion. As expected, MCF10A cells appear to form polarised, growth arrested structures. These observations suggest that this model may allow the investigation of cells as they invade across the basement membrane. Importantly, SUM159PT cells are the most affected by Matrigel in terms of morphology and invasive capacity, and were therefore selected for further investigation.

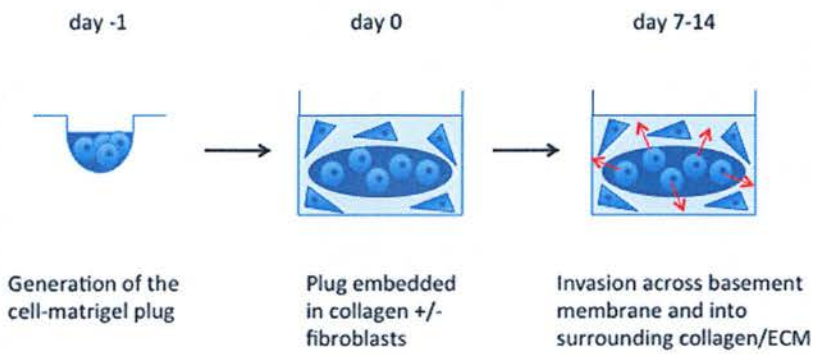
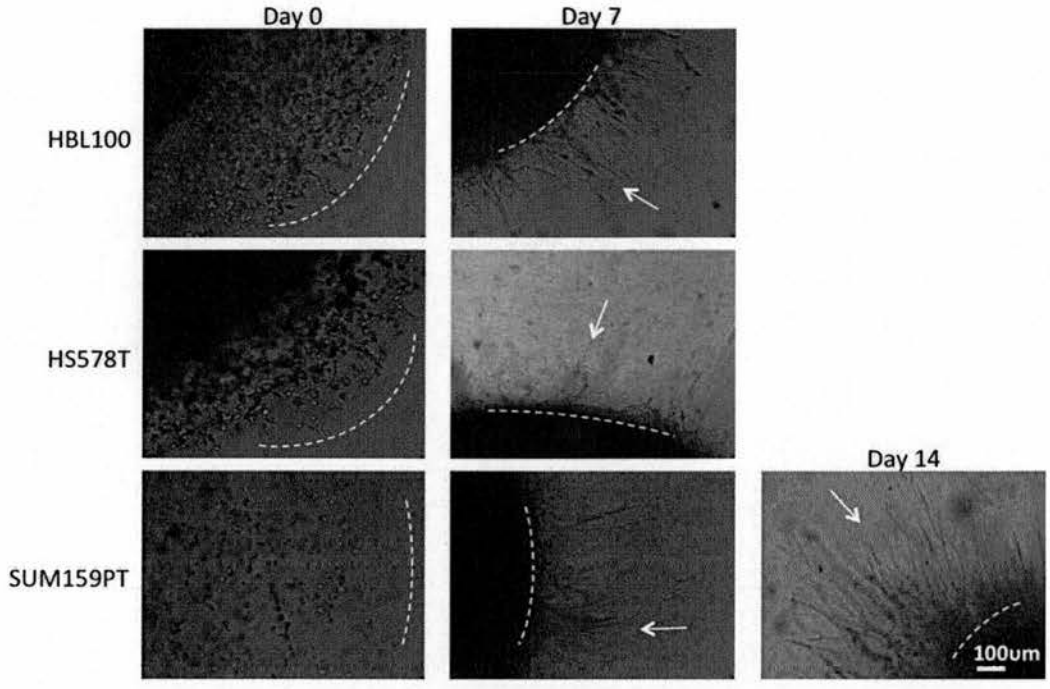


Figure 5.8. A novel 3D invasion assay. Epithelial cells are embedded in basement membrane-like Matrigel to generate a cell plug. This plug is subsequently embedded in collagen to mimic surrounding extracellular matrix. Epithelial cells are then observed as they invade across basement membrane and into extracellular matrix.

a Collagen-based plug



b Matrigel-based plug

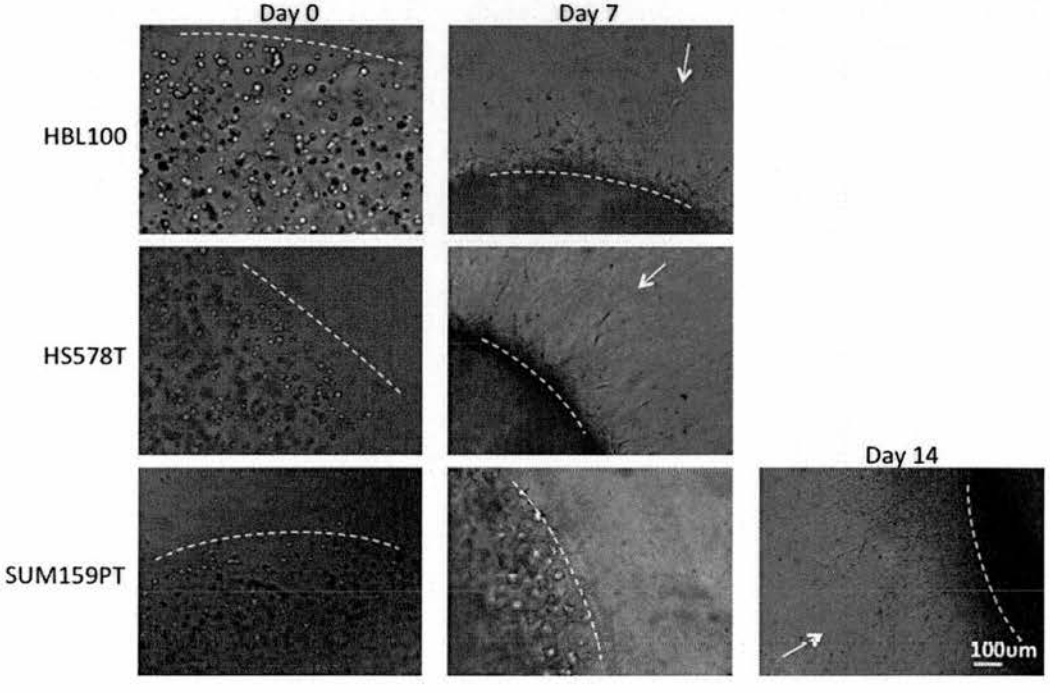


Figure 5.9. Morphological changes suggestive of spontaneous transitions between MET and EMT states are observed by light microscopy. (a) Cell-collagen plugs were made with HBL100, HS578T and SUM159PT cell lines. These exhibit a predominantly elongated morphology at day 0. Clear invasion into surrounding collagen is seen by day 7 in all three lines (arrows). (b) Cell-matrigel plugs made with the same lines exhibit a marked rounded morphology on day 0, contrasting with that seen with cell-collagen plugs. By day 7, HBL100 and HS578T cells have reverted to an elongated morphology and are invading into surrounding collagen. In contrast, SUM159PT cells retain a rounded morphology accompanied by delayed invasion (day 7) although this appears to be overcome by day 14. Dotted lines represent the original plug edge. Bar = 100 μm.

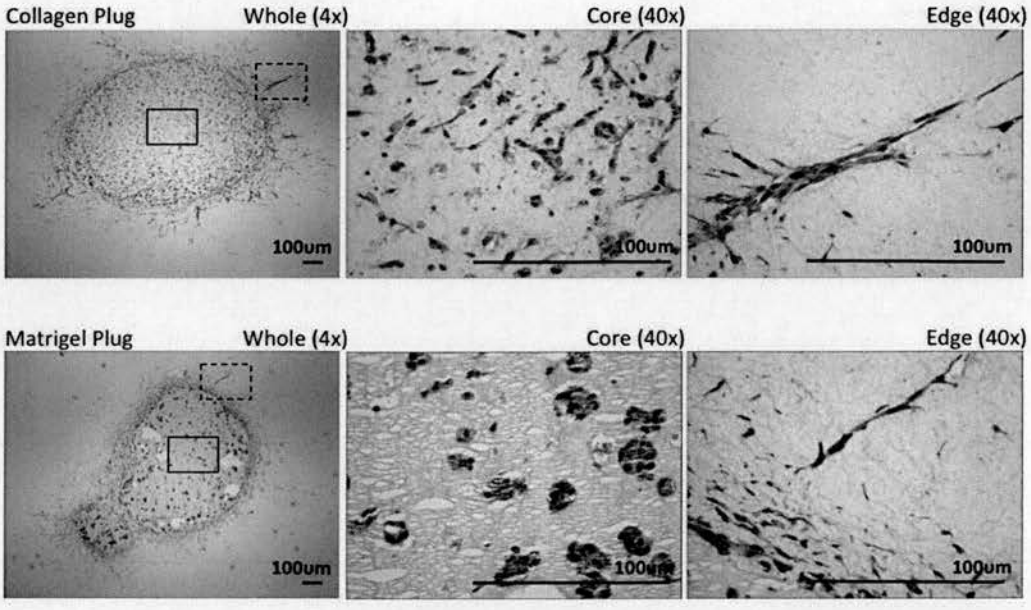
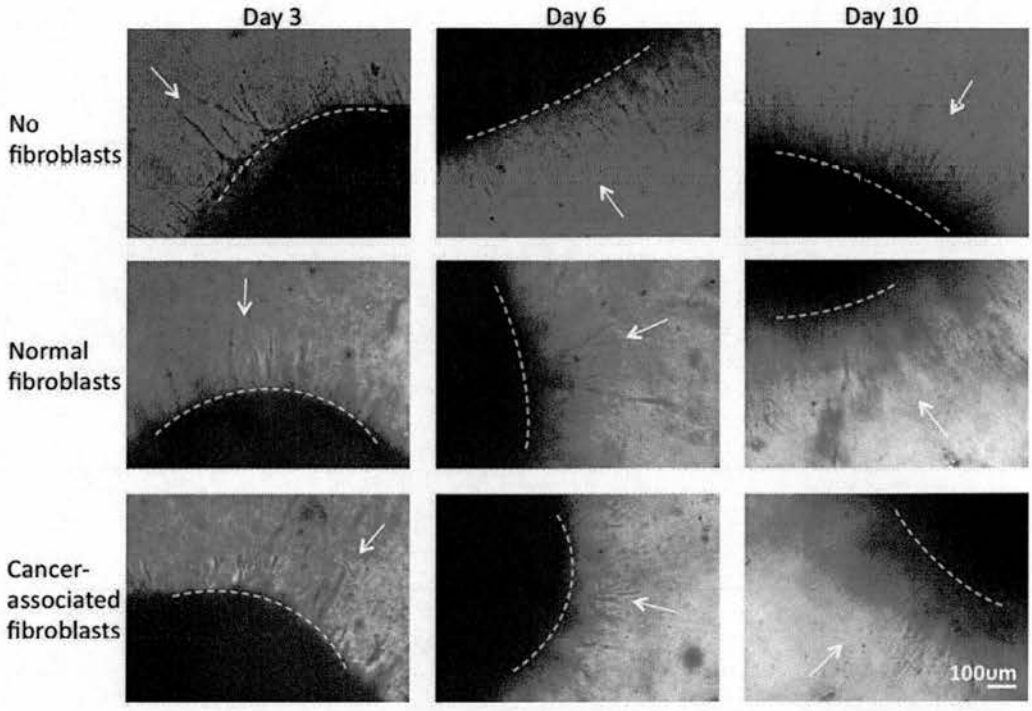


Figure 5.10. H&E staining confirms the observations made by light microscopy. Only assays using SUM159PT cells are shown here. Cell-collagen (*top panel*) and cell-matrigel (*bottom panel*) plugs were fixed at day 14 following a period of invasion. Images of the whole plugs (4x magnification; bars = 100 μm), plug cores and plug edges are shown (both 40x magnification; bars = 100 μm). Note the organised, rounded morphology in matrigel (*bottom panel*) in contrast to the elongated morphology as cells invade into surrounding collagen, suggestive of EMT.

Stromal fibroblasts have been shown to play critical roles in some models of invasion, remodelling the ECM and generating tracks along which epithelial cells can follow [299]. The role of normal and cancer-associated fibroblasts (CAFs) was therefore investigated here. No difference in invasion was evident with both normal fibroblasts and CAFs and this lack of effect on invasion was seen when epithelial cell plugs were made with both collagen and Matrigel (Figure 5.11). H&E staining at days 6 and 10 (Figure 5.12) confirms these observations.

Collagen-based plug



Matrigel-based plug

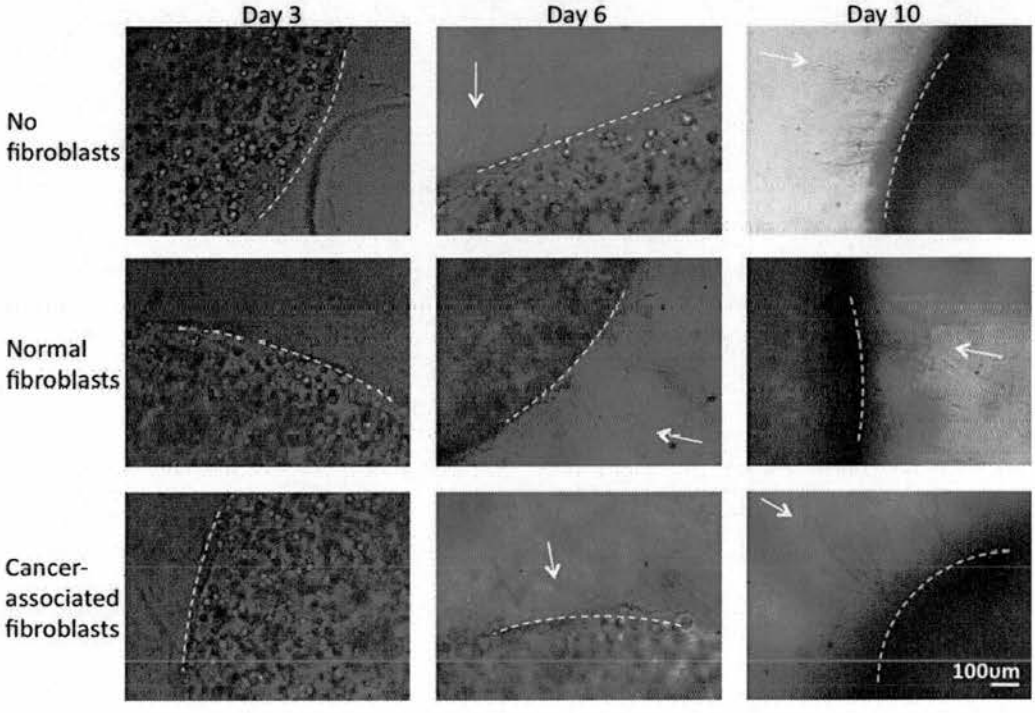


Figure 5.11. Fibroblasts have no obvious effect on the invasion of SUM159PT cells. Comparable invasion is seen by light microscopy with no, normal and cancer-associated fibroblasts at days 3, 6 and 10. This is observed both in collagen-based (*top panel*) and matrigel-based (*bottom panel*) cell plugs. As previously observed, SUM159PT cells exhibit a rounded morphology in matrigel and delayed invasion. Bar = 100 μm.

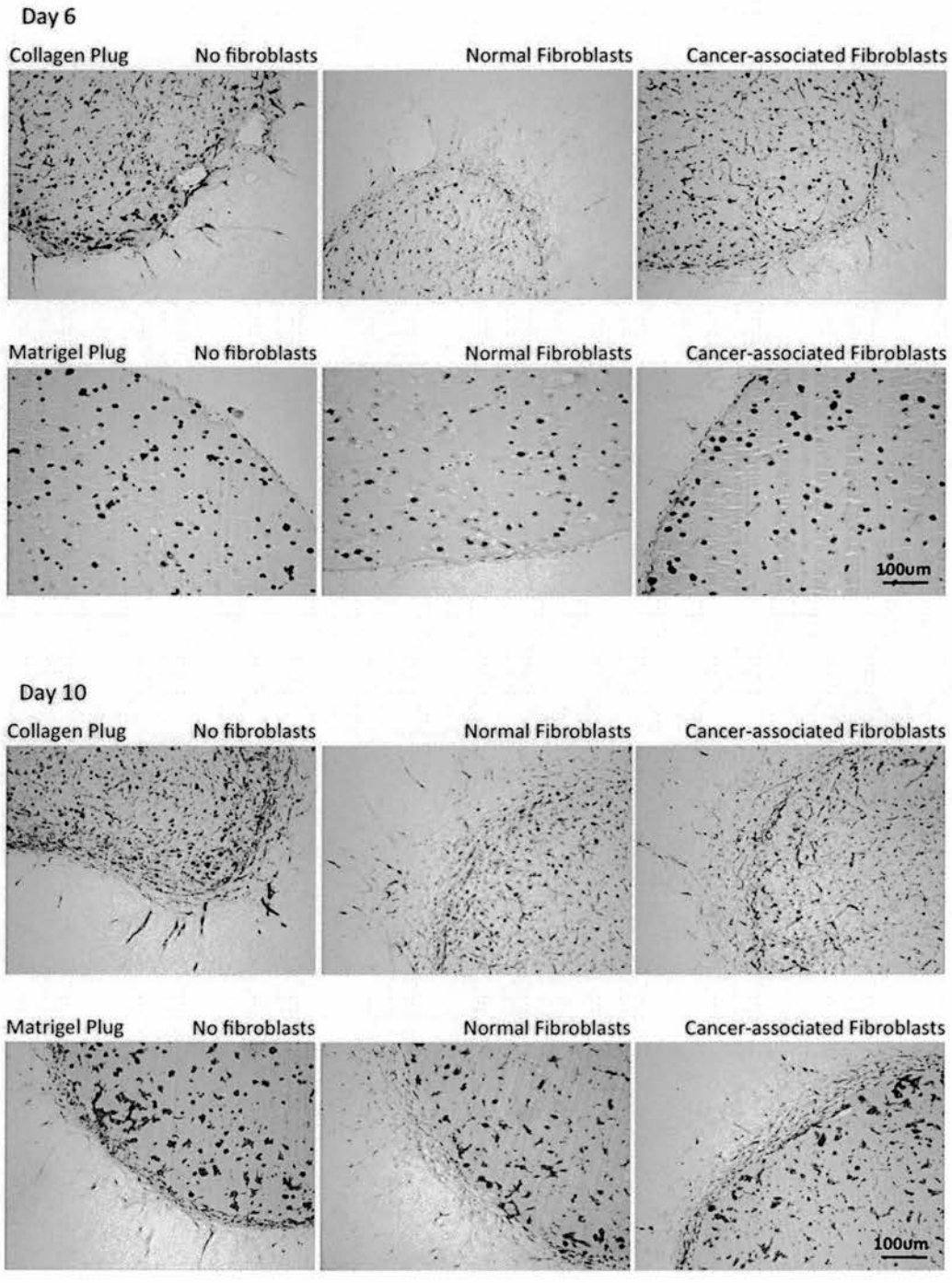


Figure 5.12. H&E staining confirms the comparable invasion of SUM159PT cells regardless of the presence or type of fibroblasts in the surrounding collagen. Comparable invasion of SUM159PT cells is seen with no, normal and cancer-associated fibroblasts. This is seen at both 6 (*top panel*) and 10 days (*bottom panel*) and is independent of the epithelial plug type. Bar = 100 μm.

The expression of EMT markers changes as cells invade across basement membrane

What mechanisms underlie the invasion of cells across the basement membrane as observed here? To begin to answer this question the expression of E-cadherin and N-cadherin, both important markers in EMT, was examined using immunofluorescence. As HBL100 and HS578T appear to rapidly overcome the influence of Matrigel, SUM159PT cells were again selected for comparison with MCF10A cells.

MCF10A cells show uniform, membranous expression of E-cadherin and no expression of N-cadherin. In contrast, SUM159PT cells show no membranal E-cadherin expression. Instead, E-cadherin expression appears to be nuclear, the significance of which is not clear but may reflect a non-functional E-cadherin. SUM159PT cells show membranal N-cadherin expression throughout the core of the plug whilst N-cadherin expression appears to be lost in the elongated, invading cells at the periphery (Figure 5.13 and 5.14). Therefore, two important observations are made: i) despite similar, round morphology within the plug, the pattern of E-cadherin expression in MCF10A cells is comparable to that of N-cadherin in SUM159PT cells; ii) perhaps surprisingly, SUM159PT cells demonstrate loss of N-cadherin as they invade. These findings may explain the lack of a cadherin switch that was observed in human tumours (Chapter 3).

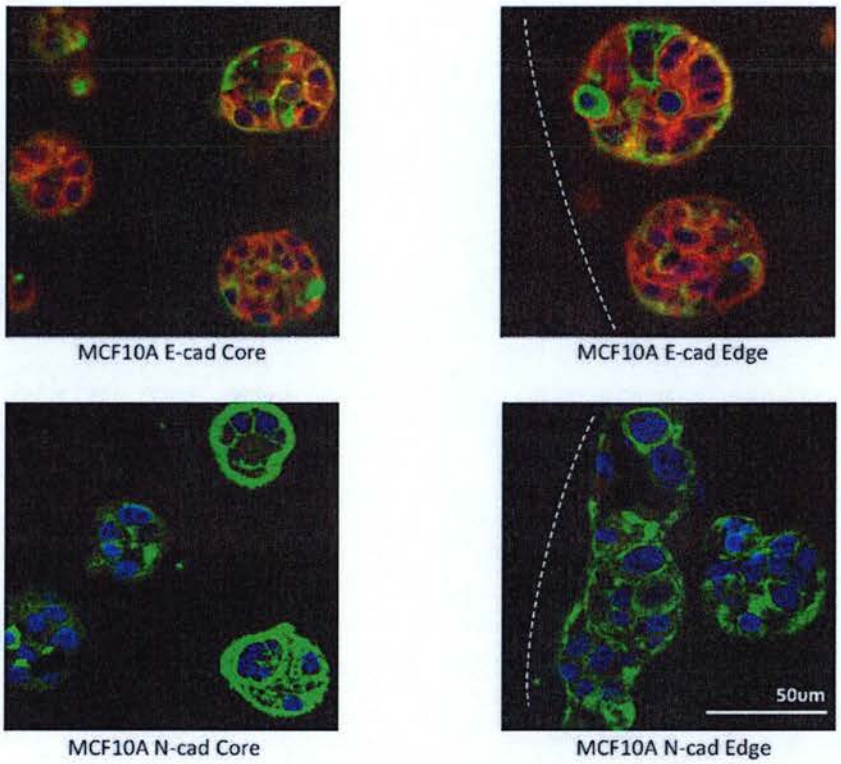


Figure 5.13. Immunofluorescent images illustrate expression patterns of E-cadherin and N-cadherin in MCF10A cells. MCF10A cells show membranous E-cadherin expression throughout the cell plug (*top 2 panels*). No expression of N-cadherin is seen (*bottom 2 panels*). In addition, MCF10A cells shown no evidence of invasion out of the plug. Dotted lines represent the plug edge. Bar = 50 μm.

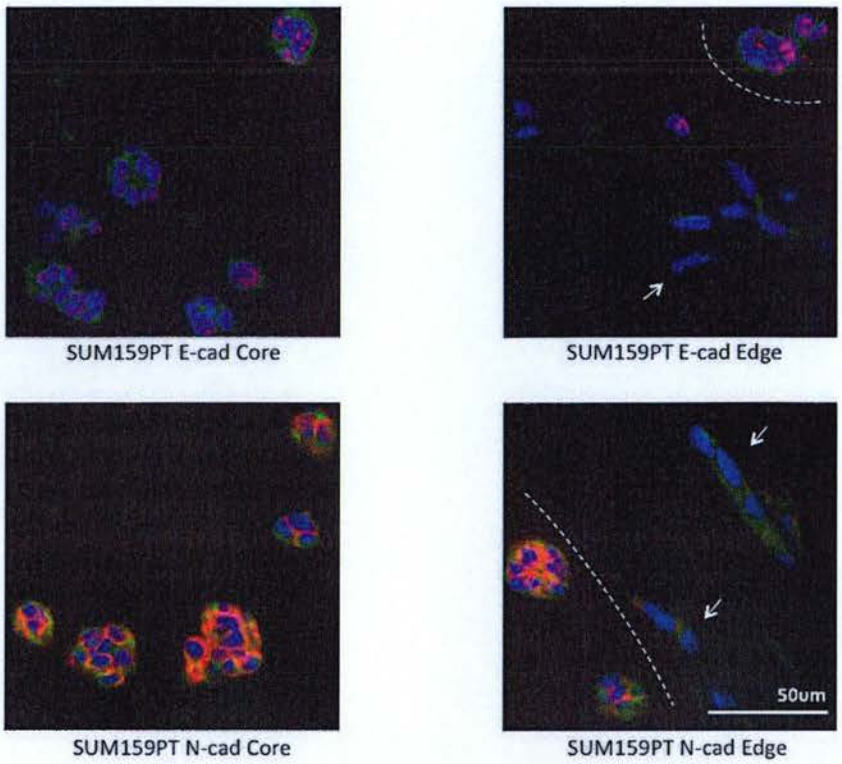


Figure 5.14. Immunofluorescent images illustrate expression patterns of E-cadherin and N-cadherin in SUM159PT cells. SUM159PT cells appear to express nuclear E-cadherin (*top 2 panels*). Despite the rounded morphology of cells within the plug, membranal N-cadherin expression is observed. In addition, N-cadherin expression is lost in the elongated, mesenchymal cells that invade out of the plug (*bottom 2 panels*). Dotted lines represent the plug edge. Bar = 50 μm.

Table 5.1. Comparison of genes correlating with C35 expression/EMT and claudin-low tumours identifies a 9-gene signature.

Gene	Protein Name	Function	Evidence of Promoter Methylation
CDH1	E-cadherin	Adherens junctions	Breast and other cancers
CLDN7	Claudin-7	Tight junctions	Breast
CRB3	Crumbs	Polarity complex	Not available
IRF6		Binds to maspin	Not available
KRT8	Keratin-8	Intermediate filaments	MCF7 cells
MAL2		Protein trafficking	Not available
MARVELD3		Related to MAL2	Not available
SPINT2	HAI-2	HGF activation	Liver cancer
TACSTD1	EpCAM	Cell-cell contact	Various cell lines

Chapter 6: Discussion

EMT is believed to be a mechanism that allows cells to acquire an invasive phenotype critical for cancer progression [68]. However, most of the evidence supporting this function in cancer is based on *in vitro* experimental systems using cell lines. This thesis had two broad aims. Firstly, to clarify the role of EMT in human breast cancer by examining the expression of key protein markers of EMT in primary tumours and, for a subset of these, in paired LN metastases. Secondly, this thesis aimed to develop a 3D *in vitro* model for the investigation of invasion and in particular the mechanisms underlying EMT as cells invade into collagen.

Evidence for transcriptionally driven EMT programmes is identified in human breast cancer

An EMT programme comprising vimentin, Snail and Slug

The expression of, and correlations between, a wide range of EMT markers was examined to determine whether relationships identified *in vitro* are maintained in human breast cancer. This led to the identification of significant correlations between vimentin, Snail and Slug, suggesting that a transcriptionally driven programme may be in place in human breast cancer. Although not predictive of clinical outcome, this programme was reproducible across two independent patient cohorts.

Snail and Slug are known transcriptional repressors of E-cadherin protein expression *in vitro* [76]. However, whilst strongly expressed in some tumours, these repressors showed no correlation with E-cadherin in both patient cohorts. Similar results were found for vimentin expression. These findings suggest that the programme identified here is uncoupled from down-regulation of E-cadherin and supports the hypothesis that EMT does occur *in vivo* but in a manner distinct to that suggested by *in vitro* studies. Up-regulation of some EMT markers with associated down-regulation of E-cadherin has been reported in a subset of human breast cancers [131]. However, this study used immunohistochemical methods which are less quantitative and objective than AQUA analysis. In addition, the arbitrary division of protein expression as positive or negative is unlikely to accurately reflect the biology of biomarker

expression [240]. Comparison between this study and the results presented here is further limited as only E-cadherin, N-cadherin and vimentin expression were examined in both cases.

A number of studies have reported expression of vimentin, Snail and Slug that is at least partially uncoupled from down-regulation of E-cadherin. Come and colleagues have examined the expression of Snail, Slug and E-cadherin mRNA in a series of transformed cell lines and breast carcinomas [219]. In accordance with other studies they found that Slug and Snail negatively correlated with E-cadherin in cell lines [300, 301]. However, in normal HMECs and in a large panel of breast tumour samples, E-cadherin was found to co-express with Snail and Slug. Cell line models have shown that Snail and Slug can target E-boxes in the E-cadherin promoter to repress E-cadherin transcription [300, 301] and the persistent co-expression of these markers in tumours suggests that other regulatory elements are involved *in vivo* [302, 303]. Another study examined the expression of E-cadherin, vimentin and Snail as well as migratory potential in a series of cell lines exposed to hypoxia [304]. Despite an increase in Snail expression across all lines examined the observed EMTs were described as partial as E-cadherin expression and migratory potential were variably effected. Chao and colleagues have examined the expression of epithelial and mesenchymal markers in paired breast primary and metastatic tumours [305]. They describe a significantly increased expression of E-cadherin in metastases but unchanged expression of mesenchymal markers including vimentin. This persistence of mesenchymal markers suggests that subsequent EMTs at the metastatic site might be important for disease progression (discussed subsequently) but also supports our finding that vimentin expression can be uncoupled from E-cadherin expression.

Candidate up-stream regulators of an EMT programme comprising vimentin, Snail and Slug in breast cancer may already have been described. Ladybird homeobox 1 (*LBX1*), a developmentally regulated homeobox gene, has been shown to regulate Snail in the context of breast cancer [306]. Down-stream targets of Snail and Slug, in the context of breast cancer, also need to be identified and may include members of the MMP gene family. These MMPs are up-regulated in many tumour types including breast [307]. One study identified a correlation between Slug mRNA

expression and MMP2 in ovarian carcinomas. This correlation was not seen in breast carcinomas although numbers in this part of the study were limited [217]. More recently, Weiss and colleagues reported that Snail can induce a basement membrane transmigration programme which is dependent on the membrane-anchored metalloproteinases, MT1-MMP and MT2-MMP. Silencing of these MMPs inhibited the ability of Snail to drive basement membrane invasion, angiogenesis and intravasation *in vitro* in breast cancer cell lines [285]. In addition, MMP1, MMP2 and MT1-MMP, are up-regulated by Snail in the context of hepatocellular carcinoma cell lines [308]. Both MMP1 and MMP2 are found in primary breast carcinomas and their expression is significantly higher in those primary tumours associated with distant metastases [309]. It has been recently shown *in vitro* that elongation of invadopodia, the specialized protrusions used by invasive cells to breach basement membrane, is dependent on an intact vimentin filament network. This is evidence that vimentin also has a functional role in MMP-mediated EMT [93].

A WT1-related EMT programme is identified, and acquisition in LN metastases predicts poor clinical outcome

As a result of recent studies, focus was directed to WT1 and its relationship with E-cadherin, Snail and Slug [261, 262]. In addition, due to the emerging importance of LNs as a source of biological information [196], analysis of these markers was extended to paired LN metastases. This modified approach resulted in the identification of a novel, WT1-related EMT programme in human breast cancer. This programme is evident in both primary tumours and LN metastases and is characterised by up-regulation of WT1, Snail and Slug and, importantly, down-regulation of E-cadherin. Expression of a WT1-related EMT programme is not predictive of poor clinical outcome in both primary tumours and LN metastases. However, the concept of examining marker expression in combination rather than independently is an attractive one. A recent study in gastric cancer found that combined altered expression of E-cadherin, vimentin, Snail and the stem cell marker CD44 (specifically loss of E-cadherin and gain of mesenchymal proteins or CD44) was significantly associated with aggressive clinical features and shortened disease free survival [310].

A key finding is that acquisition of this WT1-related EMT programme by cells invading locoregional LNs predicts worst clinical outcome. The effect on survival is significant in comparison with tumours that maintain their phenotype (i.e. maintain expression of either an EMT or a non-EMT programme) whilst tumours that appear to undergo MET form an intermediate group. These findings suggest that whilst an EMT programme may be advantageous for invasion, an ability to move between EMT and MET states in accordance with a changing microenvironment reflects greater adaptability and therefore aggressiveness [76]. These findings are thought to reflect the complexity of the metastatic process and the fact that a breast tumour must pass through cycles of invasion and tumour regeneration at new sites (i.e. LNs) in an effort to establish distant metastases. The role of an MET in enabling metastatic colonization has been proposed by a number of studies [68]. In addition, a recent study in prostate cancer has shown that E-cadherin expression in metastatic deposits is inversely correlated with the size of the metastasis [305]. This suggests that EMT states in metastatic deposits are associated with more aggressive disease and that sustained phenotypic plasticity beyond the primary tumour is important. Whilst reversion to an epithelial phenotype may facilitate the survival of tumour cells in the foreign environment of the metastatic site, it does not explain the generation of a macrometastasis or the generation of metastatic deposits at other sites. However, the increasingly supported concept that EMT confers many of the properties of stem cells including self-renewal [111, 173, 174] provides an attractive mechanism by which tumour cells can repopulate the metastatic site. Other studies support the role of Snail and Slug in disease progression in the context of breast cancer. Snail and Slug are significantly over expressed in tumours with associated LN metastases, although no comparison with paired primary tumours is reported [219]. Snail expression in metastatic pleural effusions has also been shown to be negatively predictive of disease free and overall survival in breast carcinoma [217].

We recently reported a similar, WT1-related programme in ovarian carcinoma (Faratian et al., unpublished observation). In contrast to the situation in breast, expression of this programme in primary tumours was associated with a significantly worse prognosis. In addition, multivariate analysis with stage and histological type showed that expression of this phenotype is an independent poor prognostic factor.

However, expression of the WT1-related programme was not enriched by and did not predict response to chemotherapy. This is in contrast to recent findings in breast [166, 169]. The differing predictive role of this EMT programme in ovarian and breast may reflect the features of ovarian cancer spread that distinguish it from other epithelial tumours. Ovarian carcinoma can spread by direct extension into adjacent organs and, whilst lymphatic dissemination occurs, spread through the vasculature is uncommon [311, 312]. In contrast, breast carcinoma cells must acquire an ability to invade locally, travel through the vasculature and lymphatics and finally establish distant metastases in a foreign microenvironment [268]. More recently, the protein expression of E-cadherin, Snail and Slug has been explored in renal cell carcinoma [313]. 14/61 patients were found to express an EMT signature comprising high E-cadherin and low Snail and Slug expression but this showed no correlation with clinical outcome.

The results presented here support a role for EMT-promoting transcription factors, in particular Snail and Slug, in cancer progression. As previously discussed, these factors can act as cofactors for Smads resulting in the formation of EMT promoting Smad complexes (EPSC) which represent a point of convergence between a number of signalling pathways [110]. In Chapter 5 a 9-gene claudin-low/EMT signature was identified [293]. Interestingly, transcriptional repressors are not part of this signature although they are broadly expressed in claudin-low cell lines. In addition, analysis of the promoter regions of the 9 genes shows that they have the potential to be targeted by E-box repressors including members of the Snail and Zeb families. Evidence that some of these genes are targeted by Snail and Slug is present in the literature [289, 301, 314-316], but further work is required to fully understand the role of these proteins in human breast cancer.

The results presented here also suggest that quantitative analysis of molecular markers between primary and nodal disease may hold important biological information. In support of this, a recent study showed that a significant number of patients show discordant expression of ER, PR or HER2 receptor status between primary and nodal disease [196]. These observations have clinical implications as adjuvant therapy decisions are founded on the molecular pathology of the diagnostic

core biopsy or primary resection specimen. Consequently, current treatment failures may, at least in part, reflect a changing biology as tumour cells leave the primary site.

EMT *in vivo* is distinct from the EMT observed *in vitro*

Examination of the expression of EMT markers, and in particular their relationship to E-cadherin, suggests that many of the relationships reported *in vitro* are not maintained in human breast cancer. The relationship between E-cadherin and transcriptional repressors has already been discussed. In addition, no reproducible, negative correlation with N-cadherin was observed, indicating that human breast cancers as a whole do not demonstrate a cadherin switch as observed *in vitro* [86].

The nuclear translocation of β-catenin, resulting from either Wnt signalling of loss of E-cadherin-β-catenin junctions, has been identified as a driver of EMT in *in vitro* and animal studies [186, 232]. Therefore, the relationship between E-cadherin and β-catenin was examined here. Importantly, a significant and reproducible correlation between E-cadherin and β-catenin protein expression was seen, suggesting that β-catenin does not drive EMT in human breast cancer. Rather, this suggests that the co-localization of E-cadherin and β-catenin at cellular junctions, as seen in normal epithelial tissues, is maintained [82]. In addition, univariate analysis showed no correlation between β-catenin expression and clinical outcome. These findings are in contrast to those of Rimm and colleagues who reported a significant correlation between loss of cytoplasmic β-catenin and poor clinical outcome using AQUA [240]. However, positive correlations between E-cadherin and β-catenin in human breast cancer, with no correlation to outcome, have recently been described using immunohistochemical methods [317]. Other immunohistochemical studies have investigated β-catenin expression patterns and correlation with clinical outcome with mixed results [238, 239].

Examination of EMT markers in primary tumours and their metastases may be key to understanding EMT in human breast cancer

Two, transcriptionally-driven, EMT programmes, both comprising Snail and Slug, have been identified here. Importantly, one programme appears uncoupled from E-cadherin down-regulation whilst the other includes it. In both cases, unsupervised clustering shows that these programmes are variably expressed in different tumours. These findings are not surprising given the accumulating evidence that EMT need not comprise a single, conserved programme. Partial or incomplete EMT phenotypes, where advanced carcinomas display some mesenchymal features whilst retaining characteristics of well-differentiated epithelial cells, have been described in several studies [318]. In a recent review, Klymkowsky and Savagner categorized various EMT-like phenotypes identified in a series of human carcinomas, based on features seen *in vitro*, into corresponding EMT-like 'stages' [122]. In addition, the examination of additional markers, in spite of the comprehensive analysis carried out here, may reveal more complete or consistently expressed EMT programmes.

Importantly, examination of LN metastases has shown that the acquisition of the WT1-related programme in the transition from primary tumour to LN metastases is predictive of clinical outcome. In contrast, no single marker examined here, both in primary tumour and in LN metastases, correlated with outcome. This suggests that understanding the role of EMT markers may depend on investigating how their expression changes as the tumour progresses.

A novel, 3D *in vitro* model of invasion demonstrates spontaneous EMT

A novel, 3D *in vitro* model that allows cells to be investigated as they migrate into collagen and undergo spontaneous EMT has been developed. A number of preliminary but key observations have been made. Claudin-low/EMT cell lines lose their mesenchymal morphology and acquire a rounded, organised morphology when contained within Matrigel. This morphology is very similar to that of the untransformed MCF10A cell line within the same assay and is consistent with a more benign phenotype [281, 319]. Importantly, claudin-low cell lines appear to undergo spontaneous EMT, re-acquiring a mesenchymal morphology, as they invade

into collagen. These morphological changes were particularly marked in SUM159PT cells which maintain a rounded morphology for longest and consequently exhibit the greatest delay in invasion. Importantly, immunofluorescent staining showed that despite a similar rounded morphology within Matrigel, MCF10A cells express membranal E-cadherin whilst SUM159PT cells express membranal N-cadherin and a presumed non-functional nuclear E-cadherin. A number of studies have shown that internalization of E-cadherin is a mechanism by which cells can disrupt E-cadherin function, supporting this hypothesis [320, 321]. In addition, SUM159PT cells loose N-cadherin expression as they undergo a supposed EMT and invade. These findings raise questions regarding the role of these proteins, and in particular the role of a cadherin switch, in regulating cell morphology and behaviour. Maeda and colleagues have made some interesting observations regarding the timescale of events when untransformed mouse mammary epithelial cells undergo EMT [322]. They found that E-cadherin expression was maintained until day 3 whilst morphological change and up-regulation of N-cadherin occurred within 1 day, suggesting that induction of N-cadherin rather than down-regulation of E-cadherin is important for morphological EMT. However, knock-down of N-cadherin did not prevent these cells from undergoing morphological EMT and N-cadherin over expression did not induce any loss of epithelial morphology. These findings support our suggestion that morphological change and N-cadherin up-regulation are independent events. The loss of N-cadherin by invading SUM159PT cells observed in our Matrigel based model is reminiscent of the absence of a cadherin switch observed in human tumours in Chapter 3 and raises further questions about the role of a cadherin switch *in vivo* [323]. Other studies report discrepancies between morphology and behaviour in EMT. For example, down-regulation of E-cadherin in HCT116 colorectal carcinoma cell lines is associated with up-regulation of vimentin and mesenchymal morphology but these cells fail to invade in a 3D matrix [324]. Equally, silencing of Snail in an ovarian cancer cell line impairs invasion although a mesenchymal morphology is retained [325]. In keeping with the observations made here, these findings emphasise the complexity of EMT and suggest that key features of EMT may be differentially regulated. This may have implications when relying on a single feature (i.e. morphological change) as indicative of EMT.

Our findings suggest that we have developed a model that may closely reflect some aspects of EMT biology as it occurs in human tumours. A next step will be to examine whether transcriptionally driven EMT programmes, as seen in human tumours, are reproduced in this model, and whether they can be manipulated to prevent or reverse EMT and invasion. Targetable pathways identified in this way may by clinically relevant. In addition, manipulation of EMT markers within this model may allow relationships between markers to be better understood and separated from simple associations.

Future directions

In addition to continuing the investigation of EMT markers in paired LN metastases and the mechanisms underlying EMT in our *in vitro* assay, other areas for future investigation are evident. Methylation and suppression by miRNAs are alternative mechanisms believed to regulate E-cadherin expression and EMT in breast cancer [74, 75, 186]. However, the use of historical, paraffin embedded tissues in our study, with potentially compromised preservation of nucleic acids, has limited investigation to the protein level. Teasing out the role of different mechanism of E-cadherin down-regulation is likely to be important in identifying and understanding EMT. For example, the down-regulation of E-cadherin by methylation but not mutation is related to EMT in cell lines [326]. Similarly, despite shared down-regulation of E-cadherin, ILCs are molecularly and clinically distinct to IDC-NST [194]. Future studies examining alternative mechanisms of E-cadherin loss in fresh frozen tissues, in parallel to immunofluorescence analysis, are envisaged. The combination of altered biomarker expression and morphological change as markers of EMT is also envisaged as a means of capturing more 'complete' EMTs.

This study has examined the expression of protein markers of EMT at the whole tumour level. Whilst many studies do not quantify changes in protein expression according to tumour edge and tumour core, a number of studies suggest that EMT primarily occurs at the tumour invading edge [327, 328]. Importantly, a study in human squamous cell carcinoma found that changes in the expression of two out of three EMT markers identified at the invasive edge where maintained at the whole tumour level [329]. Therefore, whilst examination at the whole tumour level is likely

to yield important biological information, a careful comparison to changes at the invading edge is needed.

Acknowledgements

I would like to thank my supervisors Dr Elad Katz, Professor David Harrison, Dr Richard Meehan and Professor Mike Dixon for their support and encouragement throughout my MD Thesis.

In addition I am particularly grateful to all members of the Breakthrough Breast Cancer research laboratory at the Western General Hospital, Edinburgh, for their patience and technical help. I would like to thank Ms Lynne Johnstone who constructed most of the TMAs; Ms Danielle Wilson and Ms InHwa Um for their technical assistance with immunofluorescent staining and the use of the AQUA system; Ms Helen Caldwell who contributed to all aspects of my work, in particular tissue culture and the sectioning and staining of specimens and Mr Peter Mullen for his help and knowledge regarding western blotting. Dr Dana Faratian and Dr Andy Sims provided invaluable assistance with statistical analyses. Dr Alexey Larionov, Dr Elad Katz and Dr Andy Sims were responsible for the micro-array construction and analysis that led to the identification of a claudin-low/EMT 9 gene signature and Dr Phillipe Gautier performed the promoter analysis of these genes. I am also extremely grateful to Dr Jeremy Thomas for his insight into cancer pathology and his help with obtaining, examining and marking H&Es.

Declaration

This thesis has been composed by myself, Sylvie Dubois-Marshall. The work presented here is largely my own apart and the contributions of other researchers/technical staff are clearly acknowledged. This work has not been submitted for any other degree or professional qualification.

S. Marshall

27/9/12

References

1. Kamangar, F., G.M. Dores, and W.F. Anderson, *Patterns of cancer incidence, mortality, and prevalence across five continents: defining priorities to reduce cancer disparities in different geographic regions of the world.* J Clin Oncol, 2006. **24**(14): p. 2137-50.
2. Harnett, A., et al., *Diagnosis and treatment of early breast cancer, including locally advanced disease--summary of NICE guidance.* BMJ, 2009. **338**: p. b438.
3. Turner, N.C. and A.L. Jones, *Management of breast cancer--part I.* BMJ, 2008. **337**: p. a421.
4. Mayor, S., *UK deaths from breast cancer fall to lowest figure for 40 years.* BMJ, 2009. **338**: p. b1710.
5. Peto, R., et al., *UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years.* Lancet, 2000. **355**(9217): p. 1822.
6. Jemal, A., et al., *Cancer statistics, 2007.* CA Cancer J Clin, 2007. **57**(1): p. 43-66.
7. Lu, J., et al., *Breast cancer metastasis: challenges and opportunities.* Cancer Res, 2009. **69**(12): p. 4951-3.
8. Fernandez, Y., et al., *Novel therapeutic approaches to the treatment of metastatic breast cancer.* Cancer Treat Rev, 2010. **36**(1): p. 33-42.
9. Ring, A., et al., *The treatment of early breast cancer in women over the age of 70.* British journal of cancer, 2011. **105**(2): p. 189-93.
10. Hassan, M.S., et al., *Chemotherapy for breast cancer (Review).* Oncology reports, 2010. **24**(5): p. 1121-31.
11. Payne, S.J., et al., *Predictive markers in breast cancer--the present.* Histopathology, 2008. **52**(1): p. 82-90.
12. Anderson, W.F., et al., *Estrogen receptor breast cancer phenotypes in the Surveillance, Epidemiology, and End Results database.* Breast cancer research and treatment, 2002. **76**(1): p. 27-36.
13. Anderson, W.F., et al., *Tumor variants by hormone receptor expression in white patients with node-negative breast cancer from the surveillance, epidemiology, and end results database.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2001. **19**(1): p. 18-27.
14. Crowe, J.P., Jr., et al., *Estrogen receptor determination and long term survival of patients with carcinoma of the breast.* Surgery, gynecology & obstetrics, 1991. **173**(4): p. 273-8.
15. *Effects of chemotherapy and hormonal therapy for early breast cancer on recurrence and 15-year survival: an overview of the randomised trials.* Lancet, 2005. **365**(9472): p. 1687-717.
16. Martinez Guisado, A., et al., *Initialization of adjuvant hormonal treatment for breast cancer.* Advances in therapy, 2011. **28 Suppl 6**: p. 66-84.
17. Lao Romera, J., et al., *Update on adjuvant hormonal treatment of early breast cancer.* Advances in therapy, 2011. **28 Suppl 6**: p. 1-18.
18. Ravdin, P.M., et al., *Prognostic significance of progesterone receptor levels in estrogen receptor-positive patients with metastatic breast cancer*

treated with tamoxifen: results of a prospective Southwest Oncology Group study. Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 1992. **10**(8): p. 1284-91.
19. Bardou, V.J., et al., *Progesterone receptor status significantly improves outcome prediction over estrogen receptor status alone for adjuvant endocrine therapy in two large breast cancer databases.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2003. **21**(10): p. 1973-9.
20. *Tamoxifen for early breast cancer: an overview of the randomised trials. Early Breast Cancer Trialists' Collaborative Group.* Lancet, 1998. **351**(9114): p. 1451-67.
21. Ross, J.S., et al., *The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine.* Oncologist, 2009. **14**(4): p. 320-68.
22. Burstein, H.J., *The distinctive nature of HER2-positive breast cancers.* The New England journal of medicine, 2005. **353**(16): p. 1652-4.
23. Yamauchi, H., V. Stearns, and D.F. Hayes, *When is a tumor marker ready for prime time? A case study of c-erbB-2 as a predictive factor in breast cancer.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2001. **19**(8): p. 2334-56.
24. Slamon DJ, L.-J.B., Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L., *Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2.* N Engl J Med, 2001. **344**(11): p. 783-92.
25. Joensuu, H., et al., *Adjuvant docetaxel or vinorelbine with or without trastuzumab for breast cancer.* The New England journal of medicine, 2006. **354**(8): p. 809-20.
26. Romond, E.H., et al., *Trastuzumab plus adjuvant chemotherapy for operable HER2-positive breast cancer.* The New England journal of medicine, 2005. **353**(16): p. 1673-84.
27. Smith, I., et al., *2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial.* Lancet, 2007. **369**(9555): p. 29-36.
28. Pritchard, K.I., et al., *HER2 and responsiveness of breast cancer to adjuvant chemotherapy.* The New England journal of medicine, 2006. **354**(20): p. 2103-11.
29. Konecny, G.E., et al., *Her-2/neu gene amplification and response to paclitaxel in patients with metastatic breast cancer.* Journal of the National Cancer Institute, 2004. **96**(15): p. 1141-51.
30. Tsang, R.Y. and R.S. Finn, *Beyond trastuzumab: novel therapeutic strategies in HER2-positive metastatic breast cancer.* British journal of cancer, 2012. **106**(1): p. 6-13.
31. Agus, D.B., et al., *Targeting ligand-activated ErbB2 signalling inhibits breast and prostate tumor growth.* Cancer Cell, 2002. **2**(2): p. 127-37.

32. Scheuer, W., et al., *Strongly enhanced antitumor activity of trastuzumab and pertuzumab combination treatment on HER2-positive human xenograft tumor models.* Cancer research, 2009. **69**(24): p. 9330-6.
33. Baselga, J., et al., *Phase II trial of pertuzumab and trastuzumab in patients with human epidermal growth factor receptor 2-positive metastatic breast cancer that progressed during prior trastuzumab therapy.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2010. **28**(7): p. 1138-44.
34. Geyer, C.E., et al., *Lapatinib plus capecitabine for HER2-positive advanced breast cancer.* N Engl J Med, 2006. **355**(26): p. 2733-43.
35. Cameron, D., et al., *Lapatinib plus capecitabine in women with HER-2-positive advanced breast cancer: final survival analysis of a phase III randomized trial.* The oncologist, 2010. **15**(9): p. 924-34.
36. Cameron, D., et al., *A phase III randomized comparison of lapatinib plus capecitabine versus capecitabine alone in women with advanced breast cancer that has progressed on trastuzumab: updated efficacy and biomarker analyses.* Breast cancer research and treatment, 2008. **112**(3): p. 533-43.
37. Bartsch, R., et al., *Impact of anti-HER2 therapy on overall survival in HER2-overexpressing breast cancer patients with brain metastases.* British journal of cancer, 2012. **106**(1): p. 25-31.
38. Baselga, J., et al., *Lapatinib with trastuzumab for HER2-positive early breast cancer (NeoALTTO): a randomised, open-label, multicentre, phase 3 trial.* Lancet, 2012. **379**(9816): p. 633-40.
39. Poole, C.J., et al., *Epirubicin and cyclophosphamide, methotrexate, and fluorouracil as adjuvant therapy for early breast cancer.* The New England journal of medicine, 2006. **355**(18): p. 1851-62.
40. Fumoleau, P., et al., *Randomized trial comparing six versus three cycles of epirubicin-based adjuvant chemotherapy in premenopausal, node-positive breast cancer patients: 10-year follow-up results of the French Adjuvant Study Group 01 trial.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2003. **21**(2): p. 298-305.
41. Goldhirsch, A., et al., *Meeting highlights: updated international expert consensus on the primary therapy of early breast cancer.* Journal of clinical oncology : official journal of the American Society of Clinical Oncology, 2003. **21**(17): p. 3357-65.
42. Clarke, M., et al., *Adjuvant chemotherapy in oestrogen-receptor-poor breast cancer: patient-level meta-analysis of randomised trials.* Lancet, 2008. **371**(9606): p. 29-40.
43. Nowak, A.K., et al., *Systematic review of taxane-containing versus non-taxane-containing regimens for adjuvant and neoadjuvant treatment of early breast cancer.* The lancet oncology, 2004. **5**(6): p. 372-80.
44. Palmieri, C., et al., *Rechallenging with anthracyclines and taxanes in metastatic breast cancer.* Nature reviews. Clinical oncology, 2010. **7**(10): p. 561-74.

45. Pal, S.K., B.H. Childs, and M. Pegram, *Emergence of nonanthracycline regimens in the adjuvant treatment of breast cancer.* Breast cancer research and treatment, 2010. **119**(1): p. 25-32.
46. Sleeman, J. and P.S. Steeg, *Cancer metastasis as a therapeutic target.* Eur J Cancer, 2010. **46**(7): p. 1177-80.
47. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.
48. 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.
49. Sorlie, T., et al., *Repeated observation of breast tumor subtypes in independent gene expression data sets.* Proc Natl Acad Sci U S A, 2003. **100**(14): p. 8418-23.
50. Parker, J.S., et al., *Supervised risk predictor of breast cancer based on intrinsic subtypes.* J Clin Oncol, 2009. **27**(8): p. 1160-7.
51. Herschkowitz, J.I., et al., *Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors.* Genome Biol, 2007. **8**(5): p. R76.
52. Reis-Filho, J.S., et al., *Metaplastic breast carcinomas are basal-like tumours.* Histopathology, 2006. **49**(1): p. 10-21.
53. Hennessy, B.T., et al., *Biphasic metaplastic sarcomatoid carcinoma of the breast.* Ann Oncol, 2006. **17**(4): p. 605-13.
54. Hennessy, B.T., et al., *Squamous cell carcinoma of the breast.* J Clin Oncol, 2005. **23**(31): p. 7827-35.
55. Hennessy, B.T., et al., *Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics.* Cancer Res, 2009. **69**(10): p. 4116-24.
56. Polyak, K., *Breast cancer: origins and evolution.* J Clin Invest, 2007. **117**(11): p. 3155-63.
57. Cox, A., et al., *A common coding variant in CASP8 is associated with breast cancer risk.* Nat Genet, 2007. **39**(3): p. 352-8.
58. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**(7148): p. 1087-93.
59. Hunter, D.J., et al., *A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer.* Nat Genet, 2007. **39**(7): p. 870-4.
60. Jovanovic, J., et al., *The epigenetics of breast cancer.* Mol Oncol, 2010. **4**(3): p. 242-54.
61. Widschwendter, M. and P.A. Jones, *DNA methylation and breast carcinogenesis.* Oncogene, 2002. **21**(35): p. 5462-82.
62. Bird, A., *DNA methylation patterns and epigenetic memory.* Genes Dev, 2002. **16**(1): p. 6-21.
63. Pfeifer, G.P. and A. Besaratinia, *Mutational spectra of human cancer.* Hum Genet, 2009. **125**(5-6): p. 493-506.
64. Jones, P.A. and D. Takai, *The role of DNA methylation in mammalian epigenetics.* Science, 2001. **293**(5532): p. 1068-70.

65. Takai, D. and P.A. Jones, *Comprehensive analysis of CpG islands in human chromosomes 21 and 22.* Proc Natl Acad Sci U S A, 2002. **99**(6): p. 3740-5.
66. Yang, X., L. Yan, and N.E. Davidson, *DNA methylation in breast cancer.* Endocr Relat Cancer, 2001. **8**(2): p. 115-27.
67. Thiery, J.P., *Epithelial-mesenchymal transitions in development and pathologies.* Curr Opin Cell Biol, 2003. **15**(6): p. 740-6.
68. Polyak, K. and R.A. Weinberg, *Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits.* Nat Rev Cancer, 2009. **9**(4): p. 265-73.
69. Thiery, J.P. and J.P. Sleeman, *Complex networks orchestrate epithelial-mesenchymal transitions.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 131-42.
70. Huber, M.A., N. Kraut, and H. Beug, *Molecular requirements for epithelial-mesenchymal transition during tumor progression.* Curr Opin Cell Biol, 2005. **17**(5): p. 548-58.
71. Schmalhofer, O., S. Brabletz, and T. Brabletz, *E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 151-66.
72. Gupta, P.B., et al., *Identification of selective inhibitors of cancer stem cells by high-throughput screening.* Cell, 2009. **138**(4): p. 645-59.
73. Berx, G., et al., *E-cadherin is inactivated in a majority of invasive human lobular breast cancers by truncation mutations throughout its extracellular domain.* Oncogene, 1996. **13**(9): p. 1919-25.
74. Caldeira, J.R., et al., *CDH1 promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer.* BMC Cancer, 2006. **6**: p. 48.
75. Nass, S.J., et al., *Aberrant methylation of the estrogen receptor and E-cadherin 5' CpG islands increases with malignant progression in human breast cancer.* Cancer Res, 2000. **60**(16): p. 4346-8.
76. Peinado, H., D. Olmeda, and A. Cano, *Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype?* Nat Rev Cancer, 2007. **7**(6): p. 415-28.
77. Brennan, K., et al., *Tight junctions: a barrier to the initiation and progression of breast cancer?* J Biomed Biotechnol, 2010. **2010**: p. 460607.
78. Van Itallie, C.M. and J.M. Anderson, *Claudins and epithelial paracellular transport.* Annu Rev Physiol, 2006. **68**: p. 403-29.
79. Tsukita, S., M. Furuse, and M. Itoh, *Multifunctional strands in tight junctions.* Nat Rev Mol Cell Biol, 2001. **2**(4): p. 285-93.
80. Mellman, I. and W.J. Nelson, *Coordinated protein sorting, targeting and distribution in polarized cells.* Nat Rev Mol Cell Biol, 2008. **9**(11): p. 833-45.
81. Hartsock, A. and W.J. Nelson, *Adherens and tight junctions: structure, function and connections to the actin cytoskeleton.* Biochim Biophys Acta, 2008. **1778**(3): p. 660-9.
82. Perez-Moreno, M. and E. Fuchs, *Catenins: keeping cells from getting their signals crossed.* Dev Cell, 2006. **11**(5): p. 601-12.

83. Cereijido, M., et al., *New diseases derived or associated with the tight junction.* Arch Med Res, 2007. **38**(5): p. 465-78.
84. Martin, T.A. and W.G. Jiang, *Loss of tight junction barrier function and its role in cancer metastasis.* Biochim Biophys Acta, 2009. **1788**(4): p. 872-91.
85. Hazan, R.B., et al., *Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis.* J Cell Biol, 2000. **148**(4): p. 779-90.
86. Hazan, R.B., et al., *Cadherin switch in tumor progression.* Ann N Y Acad Sci, 2004. **1014**: p. 155-63.
87. Yilmaz, M. and G. Christofori, *Mechanisms of motility in metastasizing cells.* Mol Cancer Res, 2010. **8**(5): p. 629-42.
88. Hulit, J., et al., *N-cadherin signalling potentiates mammary tumor metastasis via enhanced extracellular signal-regulated kinase activation.* Cancer Res, 2007. **67**(7): p. 3106-16.
89. Nieman, M.T., et al., *N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression.* J Cell Biol, 1999. **147**(3): p. 631-44.
90. Suyama, K., et al., *A signalling pathway leading to metastasis is controlled by N-cadherin and the FGF receptor.* Cancer Cell, 2002. **2**(4): p. 301-14.
91. Berx, G. and F. van Roy, *Involvement of members of the cadherin superfamily in cancer.* Cold Spring Harb Perspect Biol, 2009. **1**(6): p. a003129.
92. Mendez, M.G., S. Kojima, and R.D. Goldman, *Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition.* FASEB J, 2010. **24**(6): p. 1838-51.
93. Schoumacher, M., et al., *Actin, microtubules, and vimentin intermediate filaments cooperate for elongation of invadopodia.* J Cell Biol, 2010.
94. Yilmaz, M. and G. Christofori, *EMT, the cytoskeleton, and cancer cell invasion.* Cancer Metastasis Rev, 2009. **28**(1-2): p. 15-33.
95. Blaustein, M., et al., *Mammary epithelial-mesenchymal interaction regulates fibronectin alternative splicing via phosphatidylinositol 3-kinase.* J Biol Chem, 2004. **279**(20): p. 21029-37.
96. Hattar, R., et al., *Tamoxifen induces pleiotrophic changes in mammary stroma resulting in extracellular matrix that suppresses transformed phenotypes.* Breast Cancer Res, 2009. **11**(1): p. R5.
97. Nawshad, A., et al., *Transforming growth factor-beta signalling during epithelial-mesenchymal transformation: implications for embryogenesis and tumor metastasis.* Cells, tissues, organs, 2005. **179**(1-2): p. 11-23.
98. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.
99. Xu, J., S. Lamouille, and R. Derynck, *TGF-beta-induced epithelial to mesenchymal transition.* Cell research, 2009. **19**(2): p. 156-72.
100. Massague, J., *How cells read TGF-beta signals.* Nature reviews. Molecular cell biology, 2000. **1**(3): p. 169-78.
101. Pardali, K. and A. Moustakas, *Actions of TGF-beta as tumor suppressor and pro-metastatic factor in human cancer.* Biochimica et biophysica acta, 2007. **1775**(1): p. 21-62.

102. Eger, A., et al., *beta-Catenin and TGFbeta signalling cooperate to maintain a mesenchymal phenotype after FosER-induced epithelial to mesenchymal transition.* Oncogene, 2004. **23**(15): p. 2672-2680.
103. Nelson, W.J. and R. Nusse, *Convergence of Wnt, beta-catenin, and cadherin pathways.* Science, 2004. **303**(5663): p. 1483-7.
104. Seton-Rogers, S.E., et al., *Cooperation of the ErbB2 receptor and transforming growth factor beta in induction of migration and invasion in mammary epithelial cells.* Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(5): p. 1257-62.
105. Xie, L., et al., *Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro.* Neoplasia, 2004. **6**(5): p. 603-10.
106. Karhadkar, S.S., et al., *Hedgehog signalling in prostate regeneration, neoplasia and metastasis.* Nature, 2004. **431**(7009): p. 707-12.
107. Timmerman, L.A., et al., *Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation.* Genes & development, 2004. **18**(1): p. 99-115.
108. Molofsky, A.V., R. Pardal, and S.J. Morrison, *Diverse mechanisms regulate stem cell self-renewal.* Current opinion in cell biology, 2004. **16**(6): p. 700-7.
109. Gotoh, N., *Control of stemness by fibroblast growth factor signalling in stem cells and cancer stem cells.* Current stem cell research & therapy, 2009. **4**(1): p. 9-15.
110. Fuxe, J., T. Vincent, and A. Garcia de Herreros, *Transcriptional crosstalk between TGF-beta and stem cell pathways in tumor cell invasion: role of EMT promoting Smad complexes.* Cell Cycle, 2010. **9**(12): p. 2363-74.
111. Mani, S.A., et al., *The epithelial-mesenchymal transition generates cells with properties of stem cells.* Cell, 2008. **133**(4): p. 704-15.
112. Massague, J., J. Seoane, and D. Wotton, *Smad transcription factors.* Genes & development, 2005. **19**(23): p. 2783-810.
113. Kim, Y., et al., *Integrin alpha3beta1-dependent beta-catenin phosphorylation links epithelial Smad signalling to cell contacts.* The Journal of cell biology, 2009. **184**(2): p. 309-22.
114. Verschueren, K., et al., *SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5'-CACCT sequences in candidate target genes.* The Journal of biological chemistry, 1999. **274**(29): p. 20489-98.
115. Vincent, T., et al., *A SNAIL1-SMAD3/4 transcriptional repressor complex promotes TGF-beta mediated epithelial-mesenchymal transition.* Nature cell biology, 2009. **11**(8): p. 943-50.
116. Postigo, A.A., *Opposing functions of ZEB proteins in the regulation of the TGFbeta/BMP signalling pathway.* The EMBO journal, 2003. **22**(10): p. 2443-52.
117. Thuault, S., et al., *HMGA2 and Smads co-regulate SNAIL1 expression during induction of epithelial-to-mesenchymal transition.* The Journal of biological chemistry, 2008. **283**(48): p. 33437-46.

118. Watanabe, S., et al., *HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells.* The American journal of pathology, 2009. **174**(3): p. 854-68.
119. Horiguchi, K., et al., *Role of Ras signalling in the induction of snail by transforming growth factor-beta.* The Journal of biological chemistry, 2009. **284**(1): p. 245-53.
120. Zhou, B.P., et al., *Dual regulation of Snail by GSK-3beta-mediated phosphorylation in control of epithelial-mesenchymal transition.* Nature cell biology, 2004. **6**(10): p. 931-40.
121. Yook, J.I., et al., *A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells.* Nature cell biology, 2006. **8**(12): p. 1398-406.
122. Klymkowsky, M.W. and P. Savagner, *Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe.* Am J Pathol, 2009. **174**(5): p. 1588-93.
123. Tarin, D., E.W. Thompson, and D.F. Newgreen, *The fallacy of epithelial mesenchymal transition in neoplasia.* Cancer Res, 2005. **65**(14): p. 5996-6000; discussion 6000-1.
124. Thompson, E.W., D.F. Newgreen, and D. Tarin, *Carcinoma invasion and metastasis: a role for epithelial-mesenchymal transition?* Cancer Res, 2005. **65**(14): p. 5991-5; discussion 5995.
125. Mikaelian, I., et al., *Proteotypic classification of spontaneous and transgenic mammary neoplasms.* Breast Cancer Res, 2004. **6**(6): p. R668-79.
126. Moody, S.E., et al., *The transcriptional repressor Snail promotes mammary tumor recurrence.* Cancer Cell, 2005. **8**(3): p. 197-209.
127. Xue, C., et al., *The gatekeeper effect of epithelial-mesenchymal transition regulates the frequency of breast cancer metastasis.* Cancer Res, 2003. **63**(12): p. 3386-94.
128. Prall, F., *Tumour budding in colorectal carcinoma.* Histopathology, 2007. **50**(1): p. 151-62.
129. Brabletz, T., et al., *Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment.* Proc Natl Acad Sci U S A, 2001. **98**(18): p. 10356-61.
130. Prasad, C.P., et al., *Expression analysis of E-cadherin, Slug and GSK3beta in invasive ductal carcinoma of breast.* BMC Cancer, 2009. **9**: p. 325.
131. Sarrio, D., et al., *Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype.* Cancer Res, 2008. **68**(4): p. 989-97.
132. Trimboli, A.J., et al., *Direct evidence for epithelial-mesenchymal transitions in breast cancer.* Cancer Res, 2008. **68**(3): p. 937-45.
133. Allinen, M., et al., *Molecular characterization of the tumor microenvironment in breast cancer.* Cancer Cell, 2004. **6**(1): p. 17-32.
134. Yang, A.D., et al., *Chronic oxaliplatin resistance induces epithelial-to-mesenchymal transition in colorectal cancer cell lines.* Clin Cancer Res, 2006. **12**(14 Pt 1): p. 4147-53.
135. Kajiyama, H., et al., *Chemoresistance to paclitaxel induces epithelial-mesenchymal transition and enhances metastatic potential for epithelial ovarian carcinoma cells.* Int J Oncol, 2007. **31**(2): p. 277-83.

136. Cheng, G.Z., et al., *Twist transcriptionally up-regulates AKT2 in breast cancer cells leading to increased migration, invasion, and resistance to paclitaxel.* Cancer Res, 2007. **67**(5): p. 1979-87.
137. Li, Q.Q., et al., *Twist1-mediated adriamycin-induced epithelial-mesenchymal transition relates to multidrug resistance and invasive potential in breast cancer cells.* Clin Cancer Res, 2009. **15**(8): p. 2657-65.
138. Arumugam, T., et al., *Epithelial to mesenchymal transition contributes to drug resistance in pancreatic cancer.* Cancer Res, 2009. **69**(14): p. 5820-8.
139. Sell, S., *Stem cell origin of cancer and differentiation therapy.* Critical reviews in oncology/hematology, 2004. **51**(1): p. 1-28.
140. Morrison, S.J. and A.C. Spradling, *Stem cells and niches: mechanisms that promote stem cell maintenance throughout life.* Cell, 2008. **132**(4): p. 598-611.
141. Krtolica, A., *Stem cell: balancing aging and cancer.* The international journal of biochemistry & cell biology, 2005. **37**(5): p. 935-41.
142. Sanchez Alvarado, A., *Stem cells: time to check our premises.* Cell Stem Cell, 2008. **3**(1): p. 25-9.
143. Pardal, R., M.F. Clarke, and S.J. Morrison, *Applying the principles of stem-cell biology to cancer.* Nat Rev Cancer, 2003. **3**(12): p. 895-902.
144. Shipitsin, M., et al., *Molecular definition of breast tumor heterogeneity.* Cancer Cell, 2007. **11**(3): p. 259-73.
145. Marotta, L.L. and K. Polyak, *Cancer stem cells: a model in the making.* Current opinion in genetics & development, 2009. **19**(1): p. 44-50.
146. Badve, S. and H. Nakshatri, *Breast-cancer stem cells-beyond semantics.* The lancet oncology, 2012. **13**(1): p. e43-8.
147. Dontu, G., et al., *In vitro propagation and transcriptional profiling of human mammary stem/progenitor cells.* Genes Dev, 2003. **17**(10): p. 1253-70.
148. Lapidot, T., et al., *A cell initiating human acute myeloid leukaemia after transplantation into SCID mice.* Nature, 1994. **367**(6464): p. 645-8.
149. 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.
150. Yuan, X., et al., *Isolation of cancer stem cells from adult glioblastoma multiforme.* Oncogene, 2004. **23**(58): p. 9392-400.
151. O'Brien, C.A., et al., *A human colon cancer cell capable of initiating tumour growth in immunodeficient mice.* Nature, 2007. **445**(7123): p. 106-10.
152. Ma, S., et al., *Identification and characterization of tumorigenic liver cancer stem/progenitor cells.* Gastroenterology, 2007. **132**(7): p. 2542-56.
153. Collins, A.T., et al., *Prospective identification of tumorigenic prostate cancer stem cells.* Cancer research, 2005. **65**(23): p. 10946-51.
154. Maenhaut, C., et al., *Cancer stem cells: a reality, a myth, a fuzzy concept or a misnomer? An analysis.* Carcinogenesis, 2010. **31**(2): p. 149-58.
155. Zhou, J. and Y. Zhang, *Cancer stem cells: Models, mechanisms and implications for improved treatment.* Cell Cycle, 2008. **7**(10): p. 1360-70.

156. Cozzio, A., et al., *Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors.* Genes & development, 2003. **17**(24): p. 3029-35.
157. Krivtsov, A.V., et al., *Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9.* Nature, 2006. **442**(7104): p. 818-22.
158. Fan, X., et al., *Notch pathway inhibition depletes stem-like cells and blocks engraftment in embryonal brain tumors.* Cancer research, 2006. **66**(15): p. 7445-52.
159. Jamieson, C.H., et al., *Granulocyte-macrophage progenitors as candidate leukemic stem cells in blast-crisis CML.* The New England journal of medicine, 2004. **351**(7): p. 657-67.
160. Hermann, P.C., et al., *Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer.* Cell Stem Cell, 2007. **1**(3): p. 313-23.
161. Ponti, D., et al., *Isolation and in vitro propagation of tumorigenic breast cancer cells with stem/progenitor cell properties.* Cancer research, 2005. **65**(13): p. 5506-11.
162. Stingl, J., et al., *Purification and unique properties of mammary epithelial stem cells.* Nature, 2006. **439**(7079): p. 993-7.
163. Chang, J.C., et al., *Patterns of resistance and incomplete response to docetaxel by gene expression profiling in breast cancer patients.* J Clin Oncol, 2005. **23**(6): p. 1169-77.
164. Yamauchi, K., et al., *Induction of cancer metastasis by cyclophosphamide pretreatment of host mice: an opposite effect of chemotherapy.* Cancer Res, 2008. **68**(2): p. 516-20.
165. Liu, R., et al., *The prognostic role of a gene signature from tumorigenic breast-cancer cells.* N Engl J Med, 2007. **356**(3): p. 217-26.
166. Li, X., et al., *Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy.* J Natl Cancer Inst, 2008. **100**(9): p. 672-9.
167. Farnie, G., et al., *Novel cell culture technique for primary ductal carcinoma in situ: role of Notch and epidermal growth factor receptor signalling pathways.* J Natl Cancer Inst, 2007. **99**(8): p. 616-27.
168. Piccart-Gebhart, M.J., et al., *Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer.* N Engl J Med, 2005. **353**(16): p. 1659-72.
169. Creighton, C.J., et al., *Residual breast cancers after conventional therapy display mesenchymal as well as tumor-initiating features.* Proc Natl Acad Sci U S A, 2009. **106**(33): p. 13820-5.
170. Aktas, B., et al., *Stem cell and epithelial-mesenchymal transition markers are frequently overexpressed in circulating tumor cells of metastatic breast cancer patients.* Breast cancer research : BCR, 2009. **11**(4): p. R46.
171. Kurrey, N.K., et al., *Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells.* Stem Cells, 2009. **27**(9): p. 2059-68.

172. Neumeister, V., et al., *In situ identification of putative cancer stem cells by multiplexing ALDH1, CD44, and cytokeratin identifies breast cancer patients with poor prognosis.* Am J Pathol, 2010. **176**(5): p. 2131-8.
173. Morel, A.P., et al., *Generation of breast cancer stem cells through epithelial-mesenchymal transition.* PLoS One, 2008. **3**(8): p. e2888.
174. 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.
175. Lim, E., et al., *Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers.* Nat Med, 2009. **15**(8): p. 907-13.
176. Gupta, P.B., C.L. Chaffer, and R.A. Weinberg, *Cancer stem cells: mirage or reality?* Nat Med, 2009. **15**(9): p. 1010-2.
177. Vermeulen, L., et al., *Cancer stem cells--old concepts, new insights.* Cell Death Differ, 2008. **15**(6): p. 947-58.
178. Shimono, Y., et al., *Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells.* Cell, 2009. **138**(3): p. 592-603.
179. Filipowicz, W., S.N. Bhattacharyya, and N. Sonenberg, *Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight?* Nat Rev Genet, 2008. **9**(2): p. 102-14.
180. Koralov, S.B., et al., *Dicer ablation affects antibody diversity and cell survival in the B lymphocyte lineage.* Cell, 2008. **132**(5): p. 860-74.
181. Wang, Y., et al., *DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal.* Nat Genet, 2007. **39**(3): p. 380-5.
182. Calin, G.A., et al., *A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia.* N Engl J Med, 2005. **353**(17): p. 1793-801.
183. 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.
184. Wellner, U., et al., *The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs.* Nat Cell Biol, 2009. **11**(12): p. 1487-95.
185. Cochrane, D.R., et al., *MicroRNA-200c mitigates invasiveness and restores sensitivity to microtubule-targeting chemotherapeutic agents.* Mol Cancer Ther, 2009.
186. Ma, L., et al., *miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis.* Nat Cell Biol, 2010. **12**(3): p. 247-56.
187. 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.
188. Ma, L., et al., *Therapeutic silencing of miR-10b inhibits metastasis in a mouse mammary tumor model.* Nat Biotechnol, 2010. **28**(4): p. 341-7.
189. Faratian, D., et al., *Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab.* Cancer Res, 2009. **69**(16): p. 6713-20.

190. Arpino, G., et al., *Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome.* Breast Cancer Res, 2004. **6**(3): p. R149-56.
191. Berx, G., et al., *E-cadherin is a tumour/invasion suppressor gene mutated in human lobular breast cancers.* EMBO J, 1995. **14**(24): p. 6107-15.
192. Droufakou, S., et al., *Multiple ways of silencing E-cadherin gene expression in lobular carcinoma of the breast.* Int J Cancer, 2001. **92**(3): p. 404-8.
193. Mahler-Araujo, B., et al., *Reduction of E-cadherin expression is associated with non-lobular breast carcinomas of basal-like and triple negative phenotype.* J Clin Pathol, 2008. **61**(5): p. 615-20.
194. Weigelt, B., et al., *The molecular underpinning of lobular histological growth pattern: a genome-wide transcriptomic analysis of invasive lobular carcinomas and grade- and molecular subtype-matched invasive ductal carcinomas of no special type.* J Pathol, 2010. **220**(1): p. 45-57.
195. Megha, T., et al., *Traditional and new prognosticators in breast cancer: Nottingham index, Mib-1 and estrogen receptor signalling remain the best predictors of relapse and survival in a series of 289 cases.* Cancer Biol Ther, 2010. **9**(4).
196. Aitken, S.J., et al., *Quantitative analysis of changes in ER, PR and HER2 expression in primary breast cancer and paired nodal metastases.* Ann Oncol, 2009.
197. Sidoni, A., et al., *Prognostic indexes in breast cancer: comparison of the Nottingham and Adelaide indexes.* Breast, 2004. **13**(1): p. 23-7.
198. Kononen, J., et al., *Tissue microarrays for high-throughput molecular profiling of tumor specimens.* Nat Med, 1998. **4**(7): p. 844-7.
199. Camp, R.L., G.G. Chung, and D.L. Rimm, *Automated subcellular localization and quantification of protein expression in tissue microarrays.* Nat Med, 2002. **8**(11): p. 1323-7.
200. Giltnane, J.M., et al., *Comparison of quantitative immunofluorescence with conventional methods for HER2/neu testing with respect to response to trastuzumab therapy in metastatic breast cancer.* Arch Pathol Lab Med, 2008. **132**(10): p. 1635-47.
201. Darling, D.S., et al., *Expression of Zfhep/deltaEF1 protein in palate, neural progenitors, and differentiated neurons.* Gene Expr Patterns, 2003. **3**(6): p. 709-17.
202. Altman, D.G., et al., *Dangers of using "optimal" cutpoints in the evaluation of prognostic factors.* J Natl Cancer Inst, 1994. **86**(11): p. 829-35.
203. Camp, R.L., M. Dolled-Filhart, and D.L. Rimm, *X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization.* Clin Cancer Res, 2004. **10**(21): p. 7252-9.
204. 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.
205. Amjad, S.B., R. Carachi, and M. Edward, *Keratinocyte regulation of TGF-beta and connective tissue growth factor expression: a role in suppression of scar tissue formation.* Wound Repair Regen, 2007. **15**(5): p. 748-55.
206. Katz, E., et al., *A gene on the HER2 amplicon, C35, is an oncogene in breast cancer whose actions are prevented by inhibition of Syk.* Br J Cancer, 2010.

207. Sabeh, F., et al., *Tumor cell traffic through the extracellular matrix is controlled by the membrane-anchored collagenase MT1-MMP.* J Cell Biol, 2004. **167**(4): p. 769-81.
208. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics.* Genome Biol, 2004. **5**(10): p. R80.
209. Dunning, M.J., et al., *beadarray: R classes and methods for Illumina bead-based data.* Bioinformatics, 2007. **23**(16): p. 2183-4.
210. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response.* Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.
211. van Helden, J., B. Andre, and J. Collado-Vides, *Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies.* J Mol Biol, 1998. **281**(5): p. 827-42.
212. Thomas-Chollier, M., et al., *RSAT: regulatory sequence analysis tools.* Nucleic Acids Res, 2008. **36**(Web Server issue): p. W119-27.
213. Matys, V., et al., *TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes.* Nucleic Acids Res, 2006. **34**(Database issue): p. D108-10.
214. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nat Cell Biol, 2000. **2**(2): p. 84-9.
215. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nat Cell Biol, 2000. **2**(2): p. 76-83.
216. Blanco, M.J., et al., *Correlation of Snail expression with histological grade and lymph node status in breast carcinomas.* Oncogene, 2002. **21**(20): p. 3241-6.
217. Elloul, S., et al., *Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma.* Cancer, 2005. **103**(8): p. 1631-43.
218. Martin, T.A., et al., *Expression of the transcription factors snail, slug, and twist and their clinical significance in human breast cancer.* Ann Surg Oncol, 2005. **12**(6): p. 488-96.
219. Come, C., et al., *Snail and slug play distinct roles during breast carcinoma progression.* Clin Cancer Res, 2006. **12**(18): p. 5395-402.
220. Pena, C., et al., *E-cadherin and vitamin D receptor regulation by SNAIL and ZEB1 in colon cancer: clinicopathological correlations.* Hum Mol Genet, 2005. **14**(22): p. 3361-70.
221. Spoelstra, N.S., et al., *The transcription factor ZEB1 is aberrantly expressed in aggressive uterine cancers.* Cancer Res, 2006. **66**(7): p. 3893-902.
222. Chua, H.L., et al., *NF-kappaB represses E-cadherin expression and enhances epithelial to mesenchymal transition of mammary epithelial cells: potential involvement of ZEB-1 and ZEB-2.* Oncogene, 2007. **26**(5): p. 711-24.
223. Yang, J., et al., *Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis.* Cell, 2004. **117**(7): p. 927-39.

224. 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.
225. Fodde, R., R. Smits, and H. Clevers, *APC, signal transduction and genetic instability in colorectal cancer.* Nat Rev Cancer, 2001. **1**(1): p. 55-67.
226. Pinson, K.I., et al., *An LDL-receptor-related protein mediates Wnt signalling in mice.* Nature, 2000. **407**(6803): p. 535-8.
227. Korinek, V., et al., *Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma.* Science, 1997. **275**(5307): p. 1784-7.
228. Tauriello, D.V., et al., *Wnt/beta-catenin signalling requires interaction of the Dishevelled DEP domain and C terminus with a discontinuous motif in Frizzled.* Proceedings of the National Academy of Sciences of the United States of America, 2012.
229. Ayyanan, A., et al., *Increased Wnt signalling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism.* Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3799-804.
230. Chu, E.Y., et al., *Canonical WNT signalling promotes mammary placode development and is essential for initiation of mammary gland morphogenesis.* Development, 2004. **131**(19): p. 4819-29.
231. Reya, T. and H. Clevers, *Wnt signalling in stem cells and cancer.* Nature, 2005. **434**(7035): p. 843-50.
232. Yook, J.I., et al., *A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells.* Nat Cell Biol, 2006. **8**(12): p. 1398-406.
233. Gottardi, C.J., E. Wong, and B.M. Gumbiner, *E-cadherin suppresses cellular transformation by inhibiting beta-catenin signalling in an adhesion-independent manner.* J Cell Biol, 2001. **153**(5): p. 1049-60.
234. Onder, T.T., et al., *Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways.* Cancer Res, 2008. **68**(10): p. 3645-54.
235. Collu, G.M., O. Meurette, and K. Brennan, *Is there more to Wnt signalling in breast cancer than stabilisation of beta-catenin?* Breast Cancer Res, 2009. **11**(4): p. 105.
236. Stemmer, V., et al., *Snail promotes Wnt target gene expression and interacts with beta-catenin.* Oncogene, 2008. **27**(37): p. 5075-80.
237. Karayiannakis, A.J., et al., *Expression patterns of beta-catenin in in situ and invasive breast cancer.* Eur J Surg Oncol, 2001. **27**(1): p. 31-6.
238. Lin, S.Y., et al., *Beta-catenin, a novel prognostic marker for breast cancer: its roles in cyclin D1 expression and cancer progression.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4262-6.
239. Nakopoulou, L., et al., *c-met tyrosine kinase receptor expression is associated with abnormal beta-catenin expression and favourable prognostic factors in invasive breast carcinoma.* Histopathology, 2000. **36**(4): p. 313-25.
240. Dolled-Filhart, M., et al., *Quantitative in situ analysis of beta-catenin expression in breast cancer shows decreased expression is associated with poor outcome.* Cancer Res, 2006. **66**(10): p. 5487-94.

241. Vermeulen, L., et al., *Wnt activity defines colon cancer stem cells and is regulated by the microenvironment.* Nat Cell Biol, 2010. **12**(5): p. 468-76.
242. Ellis, P., et al., *Invasive breast carcinoma. In WHO Classification of Tumours. Pathology and Genetics of Tumours of the Breast and Female Genital Organs.*, F.A. Tavassoli and P. Devilee, Editors. 2003, Lyon Press. p. 13-59.
243. Thiery, J.P., et al., *Epithelial-mesenchymal transitions in development and disease.* Cell, 2009. **139**(5): p. 871-90.
244. Ye, Y., et al., *ERalpha signalling through slug regulates E-cadherin and EMT.* Oncogene, 2010. **29**(10): p. 1451-62.
245. Hohenstein, P. and N.D. Hastie, *The many facets of the Wilms' tumour gene, WT1.* Hum Mol Genet, 2006. **15 Spec No 2**: p. R196-201.
246. Haber, D.A., et al., *Alternative splicing and genomic structure of the Wilms tumor gene WT1.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9618-22.
247. Little, M. and C. Wells, *A clinical overview of WT1 gene mutations.* Hum Mutat, 1997. **9**(3): p. 209-25.
248. Yang, L., et al., *A tumor suppressor and oncogene: the WT1 story.* Leukemia, 2007. **21**(5): p. 868-76.
249. Oka, Y., et al., *WT1 peptide vaccine for the treatment of cancer.* Curr Opin Immunol, 2008. **20**(2): p. 211-20.
250. Oka, Y., et al., *Induction of WT1 (Wilms' tumor gene)-specific cytotoxic T lymphocytes by WT1 peptide vaccine and the resultant cancer regression.* Proc Natl Acad Sci U S A, 2004. **101**(38): p. 13885-90.
251. Iiyama, T., et al., *WT1 (Wilms' tumor 1) peptide immunotherapy for renal cell carcinoma.* Microbiol Immunol, 2007. **51**(5): p. 519-30.
252. Tsuboi, A., et al., *WT1 peptide-based immunotherapy for patients with lung cancer: report of two cases.* Microbiol Immunol, 2004. **48**(3): p. 175-84.
253. Oji, Y., et al., *Absence of mutations in the Wilms' tumor gene WT1 in primary breast cancer.* Jpn J Clin Oncol, 2004. **34**(2): p. 74-7.
254. Morrison, A.A., R.L. Viney, and M.R. Ladomery, *The post-transcriptional roles of WT1, a multifunctional zinc-finger protein.* Biochim Biophys Acta, 2008. **1785**(1): p. 55-62.
255. Silberstein, G.B., et al., *Altered expression of the WT1 wilms tumor suppressor gene in human breast cancer.* Proc Natl Acad Sci U S A, 1997. **94**(15): p. 8132-7.
256. Zhang, T.F., et al., *Inhibition of breast cancer cell growth by the Wilms' tumor suppressor WT1 is associated with a destabilization of beta-catenin.* Anticancer Res, 2003. **23**(5A): p. 3575-84.
257. Loeb, D.M., et al., *Wilms' tumor suppressor gene (WT1) is expressed in primary breast tumors despite tumor-specific promoter methylation.* Cancer Res, 2001. **61**(3): p. 921-5.
258. Miyoshi, Y., et al., *High expression of Wilms' tumor suppressor gene predicts poor prognosis in breast cancer patients.* Clin Cancer Res, 2002. **8**(5): p. 1167-71.

259. Zapata-Benavides, P., et al., *Downregulation of Wilms' tumor 1 protein inhibits breast cancer proliferation.* Biochem Biophys Res Commun, 2002. **295**(4): p. 784-90.
260. Burwell, E.A., et al., *Isoforms of Wilms' tumor suppressor gene (WT1) have distinct effects on mammary epithelial cells.* Oncogene, 2007. **26**(23): p. 3423-30.
261. Kim, H.S., et al., *Identification of novel Wilms' tumor suppressor gene target genes implicated in kidney development.* J Biol Chem, 2007. **282**(22): p. 16278-87.
262. Martinez-Estrada, O.M., et al., *Wt1 is required for cardiovascular progenitor cell formation through transcriptional control of Snail and E-cadherin.* Nat Genet, 2010. **42**(1): p. 89-93.
263. Carmona, R., et al., *Localization of the Wilm's tumour protein WT1 in avian embryos.* Cell Tissue Res, 2001. **303**(2): p. 173-86.
264. Wang, L. and Z.Y. Wang, *The Wilms' tumor suppressor WT1 induces estrogen-independent growth and anti-estrogen insensitivity in ER-positive breast cancer MCF7 cells.* Oncol Rep, 2010. **23**(4): p. 1109-17.
265. Tuna, M., A. Chavez-Reyes, and A.M. Tari, *HER2/neu increases the expression of Wilms' Tumor 1 (WT1) protein to stimulate S-phase proliferation and inhibit apoptosis in breast cancer cells.* Oncogene, 2005. **24**(9): p. 1648-52.
266. Hosono, S., et al., *E-cadherin is a WT1 target gene.* J Biol Chem, 2000. **275**(15): p. 10943-53.
267. Bloushtain-Qimron, N., et al., *Cell type-specific DNA methylation patterns in the human breast.* Proc Natl Acad Sci U S A, 2008. **105**(37): p. 14076-81.
268. Nguyen, D.X. and J. Massague, *Genetic determinants of cancer metastasis.* Nat Rev Genet, 2007. **8**(5): p. 341-52.
269. Van't Veer, L.J. and B. Weigelt, *Road map to metastasis.* Nat Med, 2003. **9**(8): p. 999-1000.
270. Weigelt, B., J.L. Peterse, and L.J. van 't Veer, *Breast cancer metastasis: markers and models.* Nat Rev Cancer, 2005. **5**(8): p. 591-602.
271. Vargo-Gogola, T. and J.M. Rosen, *Modelling breast cancer: one size does not fit all.* Nat Rev Cancer, 2007. **7**(9): p. 659-72.
272. Burdall, S.E., et al., *Breast cancer cell lines: friend or foe?* Breast Cancer Res, 2003. **5**(2): p. 89-95.
273. Charafe-Jauffret, E., et al., *Gene expression profiling of breast cell lines identifies potential new basal markers.* Oncogene, 2006. **25**(15): p. 2273-84.
274. Ross, D.T., et al., *Systematic variation in gene expression patterns in human cancer cell lines.* Nat Genet, 2000. **24**(3): p. 227-35.
275. Ross, D.T. and C.M. Perou, *A comparison of gene expression signatures from breast tumors and breast tissue derived cell lines.* Dis Markers, 2001. **17**(2): p. 99-109.
276. Bissell, M.J., et al., *The organizing principle: microenvironmental influences in the normal and malignant breast.* Differentiation, 2002. **70**(9-10): p. 537-46.

277. Bissell, M.J., A. Rizki, and I.S. Mian, *Tissue architecture: the ultimate regulator of breast epithelial function.* Curr Opin Cell Biol, 2003. **15**(6): p. 753-62.
278. Debnath, J., S.K. Muthuswamy, and J.S. Brugge, *Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures.* Methods, 2003. **30**(3): p. 256-68.
279. Park, C.C., et al., *Beta1 integrin inhibitory antibody induces apoptosis of breast cancer cells, inhibits growth, and distinguishes malignant from normal phenotype in three dimensional cultures and in vivo.* Cancer Res, 2006. **66**(3): p. 1526-35.
280. Inman, J.L. and M.J. Bissell, *Apical polarity in three-dimensional culture systems: where to now?* J Biol, 2010. **9**(1): p. 2.
281. Kenny, P.A., et al., *The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression.* Mol Oncol, 2007. **1**(1): p. 84-96.
282. Brusevold, I.J., et al., *Induction of invasion in an organotypic oral cancer model by CoCl2, a hypoxia mimetic.* Eur J Oral Sci, 2010. **118**(2): p. 168-76.
283. Nurmenniemi, S., et al., *A novel organotypic model mimics the tumor microenvironment.* Am J Pathol, 2009. **175**(3): p. 1281-91.
284. Rizki, A., et al., *A human breast cell model of preinvasive to invasive transition.* Cancer Res, 2008. **68**(5): p. 1378-87.
285. Ota, I., et al., *Induction of a MT1-MMP and MT2-MMP-dependent basement membrane transmigration program in cancer cells by Snail1.* Proc Natl Acad Sci U S A, 2009. **106**(48): p. 20318-23.
286. Evans, E.E., et al., *C35 (C17orf37) is a novel tumor biomarker abundantly expressed in breast cancer.* Mol Cancer Ther, 2006. **5**(11): p. 2919-30.
287. Lioni, M., et al., *Dysregulation of claudin-7 leads to loss of E-cadherin expression and the increased invasion of esophageal squamous cell carcinoma cells.* Am J Pathol, 2007. **170**(2): p. 709-21.
288. Prat, A., et al., *Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer.* Breast Cancer Res, 2010. **12**(5): p. R68.
289. Whiteman, E.L., et al., *The transcription factor snail represses Crumbs3 expression and disrupts apico-basal polarity complexes.* Oncogene, 2008. **27**(27): p. 3875-9.
290. Zhou, C., et al., *Proteomic analysis of tumor necrosis factor-alpha resistant human breast cancer cells reveals a MEK5/Erk5-mediated epithelial-mesenchymal transition phenotype.* Breast Cancer Res, 2008. **10**(6): p. R105.
291. Zhu, Y., M. Nilsson, and K. Sundfeldt, *Phenotypic plasticity of the ovarian surface epithelium: TGF-beta 1 induction of epithelial to mesenchymal transition (EMT) in vitro.* Endocrinology, 2010. **151**(11): p. 5497-505.
292. Parr, C. and W.G. Jiang, *Hepatocyte growth factor activation inhibitors (HAI-1 and HAI-2) regulate HGF-induced invasion of human breast cancer cells.* Int J Cancer, 2006. **119**(5): p. 1176-83.

293. Katz, E., et al., *An in vitro model that recapitulates the epithelial to mesenchymal transition (EMT) in human breast cancer.* PLoS One, 2011. **6**(2): p. e17083.
294. Chen, L., et al., *Widespread, exceptionally high levels of histone H3 lysine 4 trimethylation largely mediate "privileged" gene expression.* Gene Expr, 2007. **13**(4-5): p. 271-82.
295. Fukai, K., et al., *Hepatocyte growth factor activator inhibitor 2/placental bikunin (HAI-2/PB) gene is frequently hypermethylated in human hepatocellular carcinoma.* Cancer Res, 2003. **63**(24): p. 8674-9.
296. Kominsky, S.L., et al., *Loss of the tight junction protein claudin-7 correlates with histological grade in both ductal carcinoma in situ and invasive ductal carcinoma of the breast.* Oncogene, 2003. **22**(13): p. 2021-33.
297. Prasad, C.P., et al., *Epigenetic alterations of CDH1 and APC genes: relationship with activation of Wnt/beta-catenin pathway in invasive ductal carcinoma of breast.* Life Sci, 2008. **83**(9-10): p. 318-25.
298. van der Gun, B.T., et al., *Persistent downregulation of the pancarcinoma-associated epithelial cell adhesion molecule via active intranuclear methylation.* Int J Cancer, 2008. **123**(2): p. 484-9.
299. Gaggioli, C., et al., *Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells.* Nat Cell Biol, 2007. **9**(12): p. 1392-400.
300. Batlle, E., et al., *The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells.* Nature cell biology, 2000. **2**(2): p. 84-9.
301. Cano, A., et al., *The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression.* Nature cell biology, 2000. **2**(2): p. 76-83.
302. Batsche, E., et al., *RB and c-Myc activate expression of the E-cadherin gene in epithelial cells through interaction with transcription factor AP-2.* Molecular and cellular biology, 1998. **18**(7): p. 3647-58.
303. Stemmler, M.P., A. Hecht, and R. Kemler, *E-cadherin intron 2 contains cis-regulatory elements essential for gene expression.* Development, 2005. **132**(5): p. 965-76.
304. Lundgren, K., B. Nordenskjold, and G. Landberg, *Hypoxia, Snail and incomplete epithelial-mesenchymal transition in breast cancer.* Br J Cancer, 2009. **101**(10): p. 1769-81.
305. Chao, Y., et al., *Partial Mesenchymal to Epithelial Reverting Transition in Breast and Prostate Cancer Metastases.* Cancer microenvironment : official journal of the International Cancer Microenvironment Society, 2011.
306. Yu, M., et al., *A developmentally regulated inducer of EMT, LBX1, contributes to breast cancer progression.* Genes Dev, 2009. **23**(15): p. 1737-42.
307. Stallings-Mann, M. and D. Radisky, *Matrix metalloproteinase-induced malignancy in mammary epithelial cells.* Cells Tissues Organs, 2007. **185**(1-3): p. 104-10.

308. Miyoshi, A., et al., *Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells.* Br J Cancer, 2004. **90**(6): p. 1265-73.
309. Del Casar, J.M., et al., *Expression of metalloproteases and their inhibitors in primary tumors and in local recurrences after mastectomy for breast cancer.* J Cancer Res Clin Oncol, 2009.
310. Ryu, H.S., et al., *Combination of epithelial-mesenchymal transition and cancer stem cell-like phenotypes has independent prognostic value in gastric cancer.* Human pathology, 2012. **43**(4): p. 520-8.
311. Naora, H. and D.J. Montell, *Ovarian cancer metastasis: integrating insights from disparate model organisms.* Nat Rev Cancer, 2005. **5**(5): p. 355-66.
312. Vergara, D., et al., *Epithelial-mesenchymal transition in ovarian cancer.* Cancer Lett, 2010. **291**(1): p. 59-66.
313. O'Mahony, F.C., et al., *The Use of Automated Quantitative Analysis to Evaluate Epithelial-to-Mesenchymal Transition Associated Proteins in Clear Cell Renal Cell Carcinoma.* PLoS One, 2012. **7**(2): p. e31557.
314. Mittal, M.K., et al., *Mode of action of the retrogene product SNAI1P, a SNAIL homolog, in human breast cancer cells.* Molecular biology reports, 2010. **37**(3): p. 1221-7.
315. Peinado, H., et al., *Snail mediates E-cadherin repression by the recruitment of the Sin3A/histone deacetylase 1 (HDAC1)/HDAC2 complex.* Molecular and cellular biology, 2004. **24**(1): p. 306-19.
316. Tripathi, M.K., S. Misra, and G. Chaudhuri, *Negative regulation of the expressions of cytokeratins 8 and 19 by SLUG repressor protein in human breast cells.* Biochemical and biophysical research communications, 2005. **329**(2): p. 508-15.
317. Logullo, A.F., et al., *Concomitant expression of epithelial-mesenchymal transition biomarkers in breast ductal carcinoma: association with progression.* Oncol Rep, 2010. **23**(2): p. 313-20.
318. Christiansen, J.J. and A.K. Rajasekaran, *Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis.* Cancer Res, 2006. **66**(17): p. 8319-26.
319. Han, J., et al., *Molecular predictors of 3D morphogenesis by breast cancer cell lines in 3D culture.* PLoS Comput Biol, 2010. **6**(2): p. e1000684.
320. Ivanov, A.I., A. Nusrat, and C.A. Parkos, *Endocytosis of epithelial apical junctional proteins by a clathrin-mediated pathway into a unique storage compartment.* Molecular biology of the cell, 2004. **15**(1): p. 176-88.
321. Paterson, A.D., et al., *Characterization of E-cadherin endocytosis in isolated MCF-7 and chinese hamster ovary cells: the initial fate of unbound E-cadherin.* The Journal of biological chemistry, 2003. **278**(23): p. 21050-7.
322. Maeda, M., K.R. Johnson, and M.J. Wheelock, *Cadherin switching: essential for behavioral but not morphological changes during an epithelium-to-mesenchyme transition.* J Cell Sci, 2005. **118**(Pt 5): p. 873-87.
323. Dubois-Marshall, S., et al., *Two possible mechanisms of epithelial to mesenchymal transition in invasive ductal breast cancer.* Clinical & experimental metastasis, 2011.

324. Roger, L., et al., *Gain of oncogenic function of p53 mutants regulates E-cadherin expression uncoupled from cell invasion in colon cancer cells.* Journal of cell science, 2010. **123**(Pt 8): p. 1295-305.
325. Elloul, S., et al., *Mesenchymal-to-epithelial transition determinants as characteristics of ovarian carcinoma effusions.* Clinical & experimental metastasis, 2010. **27**(3): p. 161-72.
326. Lombaerts, M., et al., *E-cadherin transcriptional downregulation by promoter methylation but not mutation is related to epithelial-to-mesenchymal transition in breast cancer cell lines.* Br J Cancer, 2006. **94**(5): p. 661-71.
327. Huszar, M., et al., *Up-regulation of L1CAM is linked to loss of hormone receptors and E-cadherin in aggressive subtypes of endometrial carcinomas.* J Pathol, 2010. **220**(5): p. 551-61.
328. Vidal, M., et al., *A Role for the Epithelial Microenvironment at Tumor Boundaries. Evidence from Drosophila and Human Squamous Cell Carcinomas.* Am J Pathol, 2010.
329. Liu, L.K., et al., *Upregulation of vimentin and aberrant expression of E-cadherin/beta-catenin complex in oral squamous cell carcinomas: correlation with the clinicopathological features and patient outcome.* Mod Pathol, 2010. **23**(2): p. 213-24.

British Journal of Cancer (2010) 103, 401 – 410

www.bjcancer.com

A gene on the *HER2* amplicon, C35, is an oncogene in breast cancer whose actions are prevented by inhibition of Syk

E Katz*[,1], S Dubois-Marshall[1], AH Sims[1], D Faratian[1], J Li[2], ES Smith[2], JA Quinn[3], M Edward[3], RR Meehan[1,4], EE Evans[2], SP Langdon[1] and DJ Harrison[1]

[1]Breakthrough Research Unit and Division of Pathology, Institute of Genetics and Molecular Medicine, University of Edinburgh, Crewe Road, Edinburgh EH4 2XU, UK; [2]Vaccinex Inc., 1895 Mt Hope Avenue, Rochester, NY, USA; [3]Section of Dermatology, Division of Cancer Sciences, Faculty of Medicine, University of Glasgow, Glasgow, UK; [4]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Western General Hospital, Edinburgh EH4 2XU, UK

BACKGROUND: C35 is a 12 kDa membrane-anchored protein endogenously over-expressed in many invasive breast cancers. C35 (*C17orf37*) is located on the *HER2* amplicon, between *HER2* and *GRB7*. The function of over-expressed C35 in invasive breast cancer is unknown.

METHODS: Tissue microarrays containing 122 primary human breast cancer specimens were used to examine the association of C35 with HER2 expression. Cell lines over-expressing C35 were generated and tested for evidence of cell transformation *in vitro*.

RESULTS: In primary breast cancers high levels of C35 mRNA expression were associated with *HER2* gene amplification. High levels of C35 protein expression were associated with hallmarks of transformation, such as, colony growth in soft agar, invasion into collagen matrix and formation of large acinar structures in three-dimensional (3D) cell cultures. The transformed phenotype was also associated with characteristics of epithelial to mesenchymal transition, such as adoption of spindle cell morphology and down-regulation of epithelial markers, such as E-cadherin and keratin-8. Furthermore, C35-induced transformation in 3D cell cultures was dependent on Syk kinase, a downstream mediator of signalling from the immunoreceptor tyrosine-based activation motif, which is present in C35.

CONCLUSION: C35 functions as an oncogene in breast cancer cell lines. Drug targeting of C35 or Syk kinase might be helpful in treating a subset of patients with *HER2*-amplified breast cancers.

British Journal of Cancer (2010) **103,** 401 – 410. doi:10.1038/sj.bjc.6605763 www.bjcancer.com
Published online 13 July 2010

Keywords: breast cancer; C35; HER2; epithelial to mesenchymal transition; ITAM; Syk kinase

The gene C35 (*C17orf37*) is located within the smallest region of amplification of the *HER2* amplicon, between *HER2* and *GRB7*. It is a 12 kDa membrane-anchored protein over-expressed in 40 – 50% of invasive breast cancers (Evans *et al*, 2006). C35 has recently been implicated in conferring invasive potential in prostate cancer cell lines (Dasgupta *et al*, 2009). It contains a canonical immunoreceptor tyrosine-based activation motif (ITAM; Evans *et al*, 2006), a motif common in receptors of the immune system (Underhill and Goodridge, 2007), which has been associated with cell transformation through the activation of downstream Syk signalling. This raises the possibility that C35 can function as a transforming oncogene. The ability of ITAM-containing proteins to transform non-haematopoietic cells was previously shown using viral glycoproteins, such as the murine mammary tumour virus envelope protein (MMTV Env; Katz *et al*, 2005). Other examples of non-haematopoietic transformation by ITAM-containing proteins include latent membrane protein 2A of Epstein – Barr virus in skin keratinocytes (Lu *et al*, 2006) and K1 protein of Kaposi's sarcoma-associated herpes virus in endothelial cells (Wang *et al*, 2006). Particularly relevant were the observations that ITAM-containing proteins contribute to mammary epithelial cell (MEC) transformation and development of mammary carcinomas (Katz *et al*, 2005; Grande *et al*, 2006; Ross *et al*, 2006).

Using the ITAM-containing envelope protein of MMTV Env and a chimeric B-cell receptor protein, many researchers have made several key observations (Katz *et al*, 2005; Grande *et al*, 2006; Ross *et al*, 2006): (1) ITAM-containing protein expression can transform immortalised normal MECs in three-dimensional (3D) culture; (2) ITAM-induced transformation is dependent on its tyrosine phosphorylation and is associated with downstream Src and Syk kinase activation and (3) mutation of the ITAM tyrosines reduces tumour induction markedly by MMTV *in vivo* and influences its genomic integration. Therefore, ITAM-containing protein expression can switch on an intrinsic transformation programme in MECs. This programme is closely associated with epithelial to mesenchymal transition (EMT). Whereas epithelial markers such as E-cadherin and keratin-18 are down-regulated, mesenchymal markers such as N-cadherin and vimentin are up-regulated (Katz *et al*, 2005; Grande *et al*, 2006).

In this study, we determined the co-expression of C35 and HER2 proteins in human breast cancers. High levels of C35 expression were shown to induce invasion mediated by EMT *in vitro* 3D

*Correspondence: Dr E Katz; E-mail: elad.katz@ed.ac.uk
Received 13 April 2010; revised 7 June 2010; accepted 9 June 2010; published online 13 July 2010

cultures using cell lines. Mutation of ITAM of C35 (or downstream Syk inhibition) was sufficient for the reversal of C35-induced transformation. Syk inhibition in combination with anti-HER2 therapy was shown to be effective in BT474 cell line model, offering a possible therapeutic approach to treat HER2$^+$ tumours.

MATERIALS AND METHODS

Tissue microarray construction and AQUA analysis

The population characteristics of the trastuzumab-treated cohort are summarised in Supplementary Table S1. *HER2* gene amplification status was confirmed by fluorescence *in situ* hybridisation (FISH) according to the manufacturer's recommendations (*HER2* FISH PharmDx; Dako, Ely, Cambridge, UK). The use of this cohort was approved by the Lothian Research Ethics Committee (08/S1101/41). After H&E sectioning of representative tumour blocks, tumour areas were marked for TMA construction and 0.6 mm^2 cores were placed into three separate TMA replicates for each sample, as previously described (Kononen *et al*, 1998).

Immunofluorescence was carried out using methods previously described (Camp *et al*, 2002). Pan-cytokeratin antibody was used to identify infiltrating tumour cells and normal epithelial cells, DAPI counterstain to identify nuclei and Cy-5-tyramide detection for target (C35, 1 : 500 dilution; Vaccinex, Rochester, NY, USA) for compartmentalised (tissue and subcellular) analysis of tissue sections. Monochromatic images of each TMA core were captured at $\times 20$ objective using an Olympus AX-51 epifluorescence microscope (Olympus, Southend-on-Sea, UK), and high-resolution digital images analysed by the AQUAnalysis software (HistoRx Ltd., Branford, CT, USA). Briefly, a binary epithelial mask was created from the cytokeratin image of each TMA core. Similar binary masks were created for cytoplasmic and nuclear compartments on the basis of DAPI staining of nuclei. C35 expression was quantified by calculating the Cy5 fluorescent signal intensity on a scale of 0–255 within each image pixel, and the AQUA score was computed by dividing the sum of Cy5 signal within the epithelial mask by the area of the cytoplasmic compartment.

AQUA scores were averaged from replicate cores. If the tumour epithelium comprised $<5\%$ of total core area, the core was excluded from analysis. To determine the cut-point value for C35 expression in Kaplan–Meier analysis (Altman *et al*, 1994), we used X-Tile (Yale University New Haven, CT, USA), which allows determination of an optimal cut point while correcting for the use of minimum *P* statistic (Camp *et al*, 2004). Overall survival was subsequently assessed by Kaplan–Meier analysis with log-rank for determining statistical significance. Comparison of differences in means of C35 according to HER2 status was carried out using the Student's *t*-test. All calculations and analyses were two tailed where appropriate using SPSS 14.0 for Windows (SPSS Inc., Chicago, IL, USA).

Immunohistochemistry

The following antibodies were used: C35, an affinity-purified rabbit polyclonal antibody 78.2 (Vaccinex) at 0.42 μg ml^{-1}; cytokeratins 5/6 (CK5/6), rabbit polyclonal antibody (Dako) at 1 : 50 dilution; E-cadherin, mouse monoclonal (BD Biosciences, Oxford, UK) at 1 : 450 dilution; Twist, mouse monoclonal (Abcam, Cambridge, UK) at 1 : 100 dilution and claudin-7, rabbit polyclonal (Abcam) at 1 : 100 dilution.

Antigen retrieval for C35, E-cadherin and claudin-7 was carried out using sodium citrate buffer (18 μM citric acid, 82 μM sodium citrate, pH 6.0). Antigen retrieval for Twist was carried out using Tris/EDTA buffer (1 mM EDTA, 10 mM Tris-HCl base, pH 8.0). Standard immunohistochemistry protocol was carried out using the REAL EnVision mouse/rabbit kit (Dako), according to the manufacturer's instructions. For C35, comparative staining showed that automated AQUA immunofluorescence and manual immunohistochemistry scores correlated as follows: $<100:0$; $100-200:1+$; $201-300:2+$ and $>300:3+$.

HER2 immunohistochemistry was carried out using HercepTest (Dako), according to the manufacturer's instructions; with antigen retrieval at 96°C for 40 min. Staining was carried out on Autostainer (Dako). HER2 assessment was carried out according to the ASCO/CAP guidelines (Wolff *et al*, 2007). HER2 tumours were defined as positive only when the immunohistological score was $3+$ and *HER2* amplification was subsequently verified by FISH.

Cell lines, transfection and foci formation

The BT474, T47D, MBA-MD-231 and SKBr3 cell lines were obtained from the American Type Culture Collection. BT474, MBA-MD-231 and SKBr3 cells were cultured in RPMI 1640 (Invitrogen, Paisley, UK) supplemented with 10% donor bovine serum, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin. T47D cells were cultured in DMEM (Invitrogen) supplemented with 10% donor bovine serum, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin.

H16N-2 is an immortalised cell line derived from normal breast epithelium that does not over-express C35 (a kind gift from Dr V Band; Band and Sager, 1991). H16N-2 cells were cultured in DFCI media (Evans *et al*, 2006) or commercial MEGM (Lonza, Slough, UK) supplemented with 5% serum. The culture media were supplemented with 0.5 mg ml^{-1} G418 for vector selection. For detection of foci formation, we stained confluent monolayers with crystal violet (0.1% crystal violet, 20% ethanol) for 5 min, followed by de-stain rinse with water.

C35 and ITAM mutants through transfection

The coding region for human C35 protein was cloned into plasmid vector pIRESneo3 (Clontech, Mountain View, CA) at *Bsi*WI and *Bam*HI restriction sites. Plasmid DNA encoding wild-type (wt), Y39F/Y50F ITAM mutant or empty vector was transfected into host cells using Lipofectamine 2000 (Invitrogen) in OptiMem transfection medium following the manufacturer's protocol. Transfection medium was replaced with growth medium after 6 h. Transfectants were selected on G418, 48 h after transfection. Bulk transfected lines were cloned using cloning discs.

C35 recombinant cells by retroviral transduction

The coding region for human C35 protein (Evans *et al*, 2006) was cloned into retroviral vector pLXSN. To make a stable retrovirus producing line, we transfected pLXSN encoding wt C35 or empty vector into PA317 cells. Viral supernatants were collected, filter sterilised (0.45 μM) and titrated in the range of approximately 10^5 PFU per ml. H16N-2 were seeded at 3- to 5×10^6 cells in a T75 flask and incubated with 3 ml of viral supernatant and 2 μg ml^{-1} polybrene at 37°C for 6 h. Infection media were replaced with DFCI growth media and 0.5 μg ml^{-1} G418 was added at 48 h after infection. Bulk transduced lines were cloned by limiting dilution. Cell lines were assessed for C35 expression by western blot and/or immunofluorescence staining with C35 mouse monoclonal antibody (clone 1F2.4.1; Vaccinex) on fixed and permeabilised cells.

Soft agar colony formation assays

Triplicate wells of a six-well plate were seeded with uniform H16N-2 or MDA-MB-231 cell suspension diluted in DFCI, 0.33% agar (4×10^3 cells per well), which was layered over a bottom layer containing 0.625% agar. Plates were incubated up to 5 weeks at 37°C, fresh media were added to each well every week to replenish

nutrients and moisture. Presence of colonies was detected under light microscope and visual inspection, at which point colonies were stained with *P*-iodonitrotetrazolium violet dye (Sigma-Aldrich, Gillingham, UK). Iodonitrotetrazolium violet stock (dissolved in 95% ethanol at 20 mg ml^{-1}) was diluted to 1 mg ml^{-1} in PBS and 0.25 ml was added to each well. After overnight incubation at 37°C, visible colonies were counted in each well; counts from three wells were averaged. The number of colonies was normalised by multiplying the average number of soft agar colonies by the ratio of attached growth colonies normal to attached growth colonies transfectant. The attached growth assay was carried out at the same time as the soft agar assay, where 1/100 of each soft agar dilution was seeded into 100 mm dish. At 10–12 days after seeding, the dishes were stained with crystal violet and colonies were counted (Foos *et al*, 1998). Similar results were obtained in three independent experiments.

Collagen invasion assays

To characterise the mode of invasion of C35-expressing cells, we carried out collagen invasion assays essentially as previously described (Amjad *et al*, 2007). These assays are different from traditional Boyden chambers in several aspects: (1) the material used is a mixture of collagen and fibroblasts, generating a lattice of stroma-like substance; (2) the presence of live fibroblasts allows for continuing interaction with the epithelial cells; (3) importantly, the cells are examined as they invade the collagen lattice, not only measuring the number that have invaded right through the material.

Briefly, rat collagen I solution was mixed with 10^5 human breast fibroblasts (obtained from reduction mammoplasty, referenced in Amjad *et al*, 2007) per lattice and left to contract in fibroblast media (DMEM (Invitrogen) supplemented with 10% serum, 50 U ml^{-1} penicillin and 50 mg ml^{-1} streptomycin) for 4–7 days. When the lattices were of the required size (approximately four-fold contraction), 3×10^5 H16N-2 cells from the desired lines were seeded on top of the lattices and incubated as submerged cultures for 3–4 days in H16N-2 media. To induce invasion, we raised the lattices to the air/liquid interface and incubated for further 7 days before the they were fixed in 10% phosphate-buffered formalin and embedded in wax.

RNA extraction and RT–PCR

RNA was extracted by RNeasy Mini kit (Qiagen, Crawley, UK), evaluated on Agilent (South Queensferry, UK) Bioanalyzer (RIN > 9.5) and labelled using Illumina TotalPrep RNA amplification kit (Applied Biosystems/Ambion, Austin, TX, USA) according to the manufacturers' instructions. Triplicate samples from whole invasion assays (1500 ng cRNA each) were hybridised to Illumina BeadChips, according to the manufacturer's instructions. Whole-genome gene expression analysis was performed using Illumina HumanRef-8 v3 Expression BeadChip and BeadArray Reader. Microarray data were analysed using packages within Bioconductor (Gentleman *et al*, 2004; http://www.bioconductor.org) implemented in the R statistical programming language (http://www.r-project.org/). The gene expression data were normalised using quantile normalisation within the bead array package (Dunning *et al*, 2007) and differential gene expression was assessed using significance analysis of microarrays (SAM; Tusher *et al*, 2001) using the siggenes package. The data set of Herschkowitz *et al* (2007) was downloaded from the UNC Microarray Database (https://www.genome.unc.edu/).

Confirmation of gene expression patterns from biological triplicates of invasion assays was carried out using the QuantiTect SYBR Green kit (Corbett/Qiagen, Crawley, UK) on a Corbett Rotor-Gene 3000. Primers for *CDH1* were: forward 5′-CGGAGAA GAGGACCAGGACT-3′, reverse 5′-GGTCAGTATCAGCCGCTTTC-3′; for *CLDN7*: forward 5′-AAAATGTACGACTCGGTGCTC-3′, reverse 5′-AGACCTGCCACGATGAAAAT; for TBP: forward 5′-GGGGA GCTGTGATGTGAAGT-3′, reverse 5′-CCAGGAAATAACTCTGG CTCA-3′; for *ACTB*: forward 5′-CCTTCCTGGGCATGGAGTCCT-3′, reverse 5′-GGAGCAATGATCTTGATCTT-3′. QuantiTect Primer Assays (Qiagen) were used for *KRT8*, *MAL2*, *TACSTD1* and *SPINT2*. PCR programme was identical for all genes: 95°C, 15 min (94°C, 15 s; 56°C, 30 s; 72°C, 30 s) × 50 cycles; 72°C, 5 min. Standard reference human cDNA was from Clontech (catalogue no. 639654), random primed, ~50 ng RNA equivalent per μl was used for quantification of mRNA expression. Final normalisation as shown in Figure 4 was performed against the geometrical mean of *ACTB* and *TBP* levels.

Flow cytometry

shRNA constructs were cloned into Open Biosystems/Thermo-Fisher, Huntsville, AL) lentiviral inducible system; cell lines generated using non-silencing and shRNA-598 (agagagacactctc catgaaca) were evaluated for both C35 and Her2 expression. FACS analysis: cells were cultured in complete medium in the presence or absence of 0.5 μg ml^{-1} doxycycline for at least 7 days, collected with trypsin and re-suspended in FACS buffer (PBS (pH 7.2), 1% BSA). For HER2 staining, cells were incubated with 2 μg ml^{-1} biotinylated Herceptin or Remicade as human IgG1 isotype control, for 20 min on ice, followed by washing and incubation with 2 μg ml^{-1} streptavidin–APC. For C35 staining, cells were fixed and permeabilised according to the manufacturer's instruction using Invitrogen Fixation and Permeabilization kit GAS-004, and stained with 0.5 μg ml^{-1} C35 monoclonal antibody 1F2 or mouse IgG (BD Biosciences, catalogue no. 557732) conjugated to Alexa 647 for 45 min at room temperature. Cells were washed in FACS buffer and analysed on FACSCalibur. Samples were run in triplicate and averaged; ratio of median fluorescence intensity was plotted.

Three-dimensional cultures

Three-dimensional cultures have been used to study the behaviour of MECs in the presence of reconstituted basement membrane (Debnath and Brugge, 2005). This assay is particularly useful in observing oncogenic potential, by measuring morphological changes of the acinar structures formed in the culture. Such changes include enlarged acinar structures, local invasion and lack of lumen formation (Debnath and Brugge, 2005). We previously studied the effects of ITAM-containing proteins using 3D cultures (Katz *et al*, 2005; Grande *et al*, 2006), accurately predicting their contribution to tumour formation *in vivo* (Ross *et al*, 2006).

Cells (5×10^3 cells per chamber) were cultured on Matrigel (BD Biosciences) cushions following the precise protocol published previously (Debnath *et al*, 2003) using the usual cell culture media with the addition of 2% Matrigel. The structures were analysed, at a magnification of × 20, on a Leitz (Microscope Co., Glasgow, UK) Dialux 20 equipped with an Insight 4 video camera and SPOT software (Diagnostic Instruments, Sterling Heights, MI, USA). Quantification of structure size was carried out using a 10 × 50 μm grid reticule (Fisher Scientific, ThermoFisher, Huntsville, AL, USA), with 20–50 structures counted from each chamber. The inhibitors BAY61-3606 and piceatannol (Merck, Nottingham, UK) and trastuzumab/Herceptin (Roche Diagnostics, Penzberg, Germany) were added as follows (Figures 5 and 6):

(5a) T47D cells treated for 14 days with the Syk inhibitors BAY61-3606 (100 nM) or piceatannol (1 μg ml^{-1}) (added twice: at days 8 and 11).

(5b) BT474 cells treated for 13 days with trastuzumab (20 μg ml^{-1}) and/or BAY61-3606 (50 nM) (twice: at days 7 and 10).

(6b) Y39F/Y50F ITAM mutant or wt C35-expressing cell lines were treated for 14 days with the Syk inhibitor BAY61-3606 (50 nM, twice: at days 8 and 11).

For siRNA experiments, 10^5 cells were plated 48 h before the 3D culture. After 24 h, the cells were transfected with 100 nM non-targeted or Syk siRNA SmartPOOLs (Dharmacon, Cramlington, UK), using Lipofectamine (Invitrogen). At 24 h after the transfection cells were collected, counted and seeded on Matrigel as described above. The effectiveness of the SmartPOOLs *vs* single siRNA was measured by qPCR (all reagents from Dharmacon) after 48 h on plastic (Supplementary Figure S3).

Statistical analysis

For comparisons of means of structure diameters, two-tailed unpaired *t*-test was used. *P*-values were as follows:

(3c) C35, null *vs* C35pool: <0.0001; C35pool *vs* $C35^{hi}$: 0.0041. E-cadherin, C35: null *vs* C35pool: <0.0001; C35pool *vs* $C35^{hi}$: <0.0001.

(5a) None *vs* piceatannol: 0.0343; none *vs* BAY61-3606: 0.0119.

(5b) None *vs* trastuzumab/Herceptin: not significant; none *vs* BAY61-3606: 0.0356; none *vs* trastuzumab + BAY61-3606: <0.0001; BAY61-3606 *vs* trastuzumab + BAY61-3606: 0.0005; trastuzumab *vs* trastuzumab + BAY61-3606: <0.0001.

(5d) Non-targeted siRNA *vs* C35 siRNA: <0.0001; non-targeted siRNA *vs* HER2 siRNA: 0.0329; non-targeted siRNA *vs* Syk siRNA: 0.0104.

(6a) Neo *vs* Y39F/Y50F C35: not significant; neo *vs* wt C35: 0.0053; Y39F/Y50F C35 *vs* wt C35: 0.0026.

(6b) Y39F/Y50F C35 *vs* wt C35: 0.0111; Y39F/Y50F C35, none *vs* BAY61-3606: not significant; wt C35, none *vs* BAY61-3606: 0.0111.

(6c) Neo non-targeted *vs* Syk siRNA: not significant; neo non-targeted *vs* C35 non-targeted: 0.0011; C35 non-targeted *vs* Syk: 0.0018.

RESULTS

C35 protein is co-expressed with HER2 in human breast cancer cells

C35 protein expression was analysed by quantitative immunofluorescence using the HistoRx AQUA image analysis system (Camp *et al*, 2002) (1) to determine whether it is co-expressed with HER2 in the same cancer cells and (2) to investigate whether level of expression of protein was associated with therapeutic response to trastuzumab (Herceptin) in a retrospective clinical cohort of 122 treated patients, 32 of which were found later to be HER2 negative (Faratian *et al*, 2009). Pre-treatment C35 protein levels measured by immunofluorescence were significantly associated with *HER2* copy number amplification assessed by FISH (Figure 1A and B; mean AQUA score HER2 not amplified = 47.8 (s.d. 55.2; range 17.4–327.7), mean AQUA score HER2 amplified = 255.2 (s.d. 170.9; range 40.1–1014.9); Mann–Whitney *U*-test, $P<0.0001$). In cancers with no *HER2* amplification expression of C35 was uniformly low in all but two cases, with AQUA scores of less than 100.

We next sought to establish whether quantitative C35 expression was associated with response to trastuzumab as measured by the

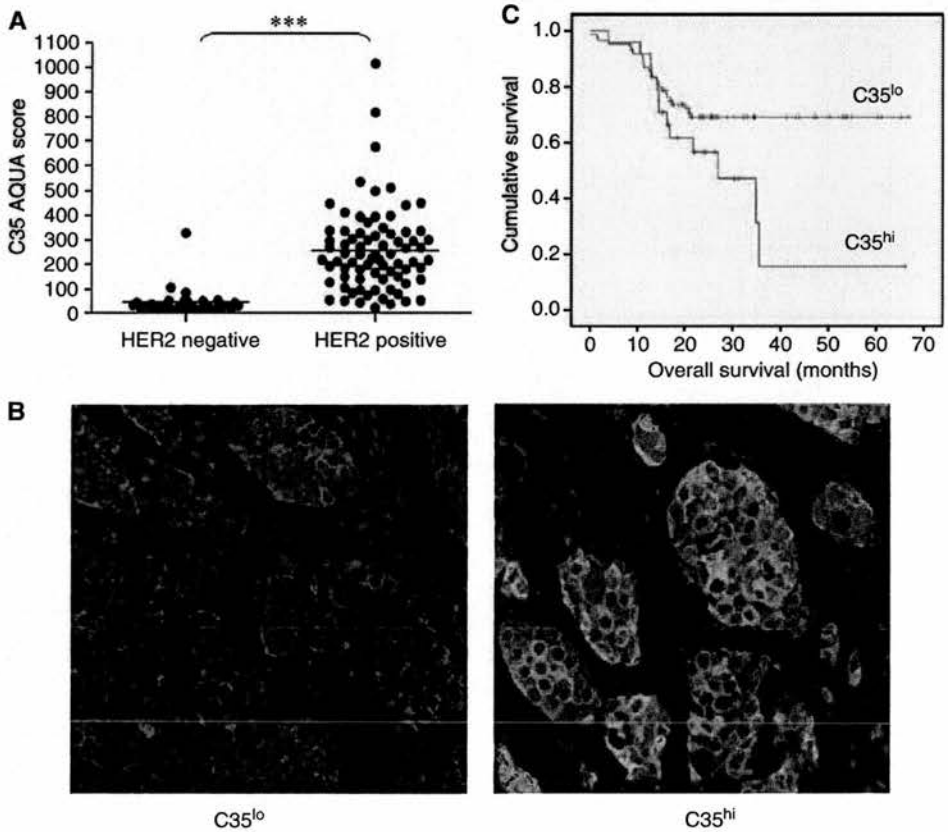


Figure 1 Clinical profile of C35 expression in human breast cancer. (**A**) Distribution of C35 immunofluorescence according to HER2 amplification status, as determined by fluorescence *in situ* hybridisation (***$P<0.001$). (**B**) representative examples of $C35^{hi}$ and $C35^{lo}$ immunofluorescence. Green: epithelial cell mask (pan-keratin); red: C35. Immunohistochemistry of C35 in primary breast cancers is shown in Supplementary Figure S1. (**C**) Kaplan–Meier survival curves according to optimal C35 cutpoint determined by minimum *P*-value method (log-rank test, $P=0.028$).

overall survival time. In univariate analysis high C35 expression (cut point AQUA score 304; minimum *P*-value method) was associated with worse overall survival (Figure 1C; log-rank test $P = 0.0285$), along with stage, ER status and chemotherapy regimen (Supplementary Table S1). However, only stage was associated with overall survival in a Cox regression multivariate analysis. Analysis of C35 expression in *HER2*-amplified tumours similarly did not yield a significant association with survival (Figure 1C).

Over-expression of C35 leads to EMT-mediated cell invasion

We carried out colony formation assays in soft agar to test whether C35 can induce MEC transformation. For this purpose, the normal MEC line H16N-2, which has been used previously for cell transformation assays (Burwell *et al*, 2007; Rhodes *et al*, 2009), was retrovirally transduced with wt C35. Colonies expressing high levels of C35 consistently formed enlarged structures in soft agar, whereas empty vector-expressing controls did not (Figure 2A). In contrast to the C35 transfectant pool, two of H16N-2 clones expressing high levels of wt C35 protein also showed foci formation when grown on plastic (Figure 2B; data not shown). In the breast cancer cell line MDA-MB-231, which normally expresses very low levels of C35, similar to those in the H16N-2 parental line (Evans *et al*, 2006), C35 expression was able to transform the MDA-MB-231 cell line at levels exceeding the transforming potential observed in the H16N-2 cell line (Figure 2C).

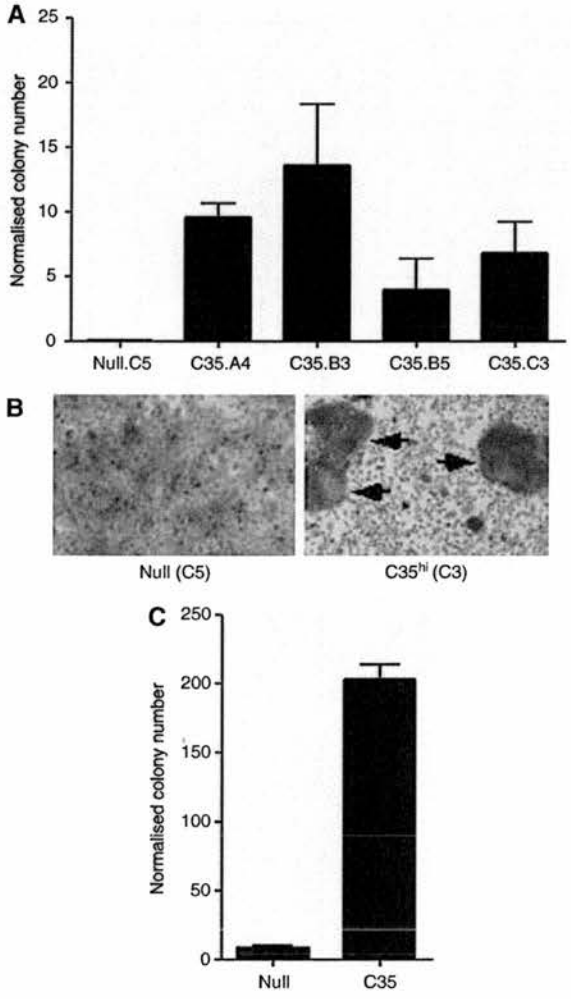


Figure 2 C35 expression in mammary epithelial cells leads to colony formation in soft agar. Quantification of soft agar assays in (**A**) H16N-2 cells expressing empty retroviral construct (null clone) or C35 (clones A4, B3, B5 and C3) and in (**C**) MDA-MB-231 cells (null or C35 transfectant pools). (**B**) Foci formation (marked by arrows) assay results are shown for null.C5 and the C35.C3 (high expressing) clones.

We previously reported that ITAM-containing proteins such as MMTV Env can induce an invasive phenotype in human MECs (Katz *et al*, 2005), likely to be caused by an EMT (Katz *et al*, 2005; Grande *et al*, 2006). It has also been shown that C35 promotes migration and invasion in prostate cancer cell lines (Dasgupta *et al*, 2009). To determine whether C35 expression also results in a similar behaviour in MEC, we used an invasion assay that used collagen lattices closely resembling breast stroma *in vivo* (Amjad *et al*, 2007; data not shown). The stroma-like lattices were generated by rat collagen I, contracted by seeding breast fibroblasts into the collagen gel. After the lattices contracted, MECs were seeded on top and invasion was induced by a nutrient gradient (Figure 3A). Although vector only (null) cells did not significantly invade the lattice, expression of C35 induced invasion. The C35 transfectant pool, which had variable levels of C35 expression, invaded mostly in large clusters of cells (Figure 3B). Three high-expressing clones showed complete transition to spindle cell phenotype, with single cells invading deep into the lattice (Figure 3B; Supplementary Figure S2). We chose one high-expressing clone, C35.C3, for further molecular characterisation (Figure 3C). Gradual loss of E-cadherin was apparent, occasionally in the C35-expressing pool and entirely within the $C35^{hi}$ clone (Figure 3B and C). Finally, all three major transcription factors known to be involved in EMT were examined. Slug expression was not detected and the level of Snail expression did not change in any C35-expressing cells. In contrast, Twist protein expression correlated positively with C35 expression (data not shown).

We carried out whole-genome expression array analysis to examine which transcripts correlate with C35 expression in the collagen invasion assays (raw gene expression files are publicly available from the caBIG-supported Edinburgh Clinical Research Facility Data Repository: https://www.catissuesuite.ecmc.ed.ac.uk/caarray/). Of the top 100 ranked differentially expressed genes by SAM (Tusher *et al*, 2001), the majority of the genes were down-regulated (62 of 98 probes, 63%, excluding a duplicate and a discontinued probe). First, we examined using KEGG analysis pathways activated or deactivated by C35 expression. The KEGG pathway that was most significantly over-represented by the most consistently differentially expressed genes by SAM analysis was cell communication ($P = 4.33E{-}08$, FDR $= 5.43E{-}05$). The genes responsible were *KRT15*, *GJB2*, *COL17A1*, *DSG3*, *KRT13*, *KRT6A*, *KRT6B*, *KRT14*, *KRT16*, *KRT8* and *LAMA3*. Using the DAVID Bioinformatics database (Huang *et al*, 2009), we found that the processes highlighted by this pathway are cell–cell contact (adherens junctions, tight junctions, desmosomes) and ECM–receptor interactions, including focal adhesions. Interestingly, gene expression of *PLAU* (uPA), *MMP9*, *VEGFA* and *VEGFB* did not correlate with C35 levels. This suggests involvement of a different set of activated signalling pathways in MECs compared with prostate cancer cells (Dasgupta *et al*, 2009).

When the most consistently differentially expressed genes were compared with those identified in two molecular subtypes of breast cancers linked recently to EMT, claudin low (Herschkowitz *et al*, 2007) and metaplastic breast cancers (Hennessy *et al*, 2009), a number of interesting results were discovered. Of the 23 commonly changed genes, 5 (22%) were among most changed by C35 expression: E-cadherin (*CDH1*), claudin-7 (*CLDN7*), MAL2, EpCAM (*TACSTD1*) and HAI-2 (*SPINT2*). Validation by quantitative PCR confirmed that all five genes were down-regulated by high expression of C35 in the invasion assays (Figure 4). Cells

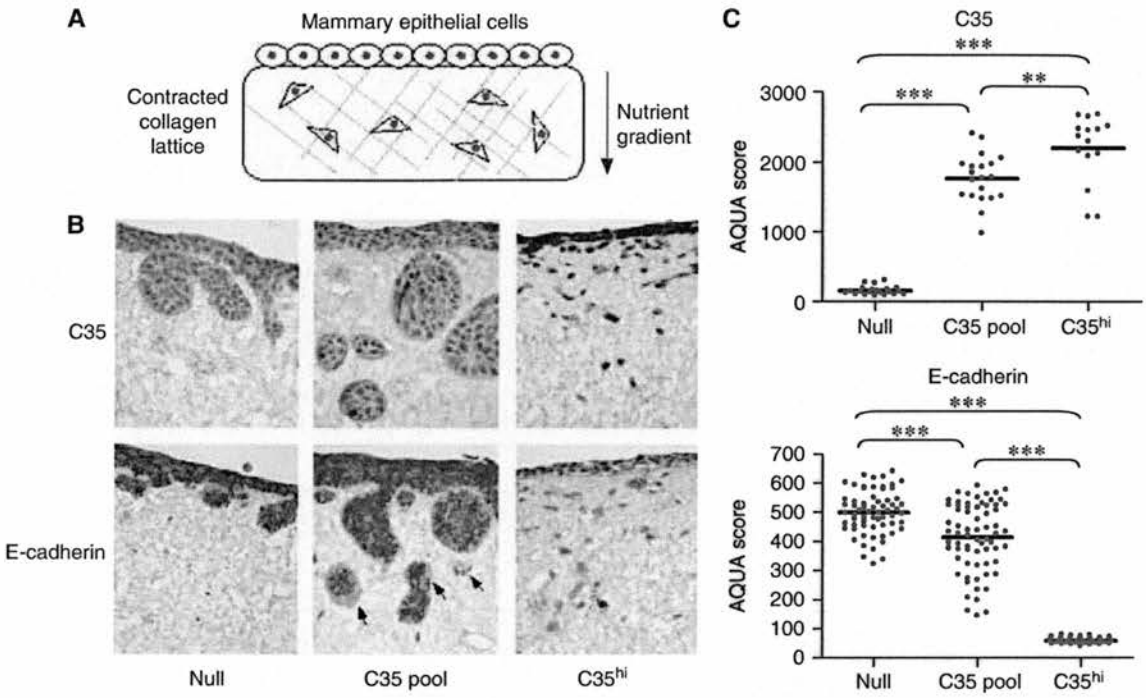


Figure 3 C35-expression leads to invasive phenotype, associated with epithelial to mesenchymal transition. (**A**) Schematic illustration of the invasion assay set-up in a collagen gel containing fibroblasts overlaid with epithelial cells. (**B**) H16N-2 cells expressing empty vector (null, left panels), C35-expressing pool (middle panels) and C35hi-expressing cells were stained by immunohistochemistry for C35 and E-cadherin. Note specific areas of E-cadherin loss in C35 pool (arrows). (**C**) Quantification of C35 and E-cadherin by AQUA is shown. Bar indicates mean of measurements (*P*-value indicators: * <0.05; ** <0.01; *** <0.001).

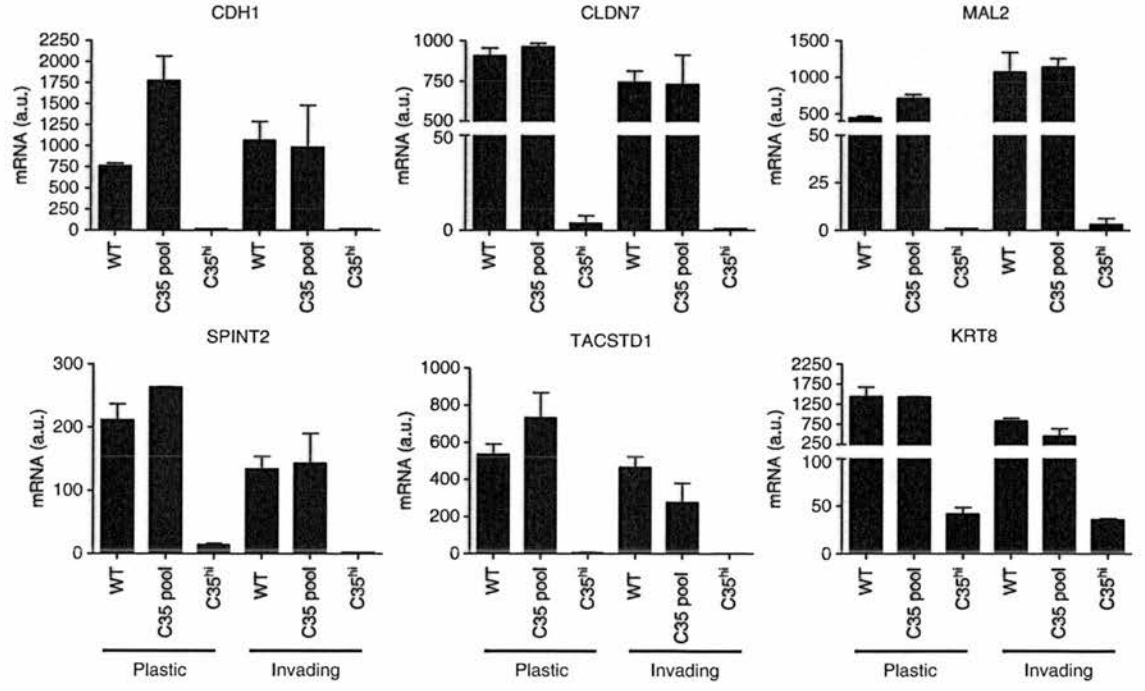


Figure 4 Genes down-regulated in C35-induced transformed phenotype. C35-induced down-regulation of *CDH1*, *CLDN7*, *KRT8*, *MAL2*, *TACSTD1* and *SPINT2* was observed in cells grown on plastic and in the invasion assays. Biological triplicate mRNA expression data are shown for empty vector, C35 expressing pool and C35^{hi}-expressing clone (C35.C3).

expressing high levels of C35 also down-regulated eight cytokeratin genes (out of 98 top ranked, 8%), also consistent with loss of epithelial phenotype. These included keratin-8 (*KRT8*, Figure 4), which is often down-regulated in EMT-like breast tumours (Herschkowitz *et al*, 2007; Hennessy *et al*, 2009). Both loss of cell–cell contact and down-regulation of cytokeratins have been linked with EMT and are thought to enable cancer cell invasion (Levayer and Lecuit, 2008).

C35-induced cell transformation is dependent on the function of its ITAM

Previous studies have used 3D cell culture, in which MECs are grown on reconstituted basement membrane (Matrigel) and form spherical structures resembling the terminal ductal lobular units in the breast. These cell cultures show many *in vivo* properties of MECs. This model has been extensively used to study oncogenic phenotypes (Debnath and Brugge, 2005). C35 contains an ITAM, a motif found in glycoproteins of oncogenic retroviruses, that is linked to epithelial cell transformation through the protein tyrosine kinase Syk (Katz *et al*, 2005; Lu *et al*, 2006; Wang *et al*, 2006). Syk binds to the ITAM through its tandem SH2 domains and activates multiple growth signalling pathways, including PI3K, PLCγ, Ras/MAPK and NFκB, among others (Underhill and Goodridge, 2007).

We determined the C35 and HER2 status of three breast cancer cell lines, as well as Syk expression. BT474 and SKBr3 lines harbour *HER2* and *C35* gene amplification and show high levels of mRNA expression of these genes (Supplementary Figure S4). T47D cells have no *HER2* gene amplification and they express moderate levels of C35 (21-fold less than SKBr3 cells, 4-fold more than MCF10A cells). T47D cells are sensitive to Syk inhibition, by piceatannol or BAY61-3606 treatment (Yamamoto *et al*, 2003; Figure 5A), similar to that of the H16N-2 C35-expressing line in 3D culture (Figure 6B). Therefore, the response of *HER2*-amplified cells to Syk inhibition was determined. BT474 cells were chosen as they form non-polarised but well-defined 'mass' 3D structures (Kenny *et al*, 2007), similar to those generated by T47D cells. Treatment with Syk inhibitors, or Syk siRNA, reduced the size of BT474 3D structures (Figure 5D). This effect was unlikely due to changes in HER2 expression (Figure 5C). Syk inhibition combined with Herceptin (trastuzumab) resulted in even smaller structures, similar in size to those seen with immortalised, but non-transformed, cell lines (Figure 5B).

We generated H16N-2 cell transfectant pools expressing the wt C35 protein or its Y39F/Y50F ITAM mutant. When grown in reconstituted basement membrane (3D culture), MECs expressing ITAM-containing proteins showed a transformed phenotype. This phenotype included enlargement of the acinar structures and was dependent on functional ITAM in these proteins (Katz *et al*, 2005; Grande *et al*, 2006). Consistent with these previous observations, when cultured in 3D, C35-expressing cells formed enlarged structures in comparison to empty vector-expressing cells (*t*-test, $P=0.0053$). Immunoreceptor tyrosine-based activation motif mutant C35-expressing cells formed similar structures to those of vector-expressing cells (Figure 6A).

Growth of C35-expressing H16N-2 cells was sensitive to Syk inhibition in 3D culture (Figure 6B) similar to other cell lines expressing ITAM-containing proteins (Katz *et al*, 2005; Grande *et al*, 2006). This was confirmed by siRNA knockdown for both Syk

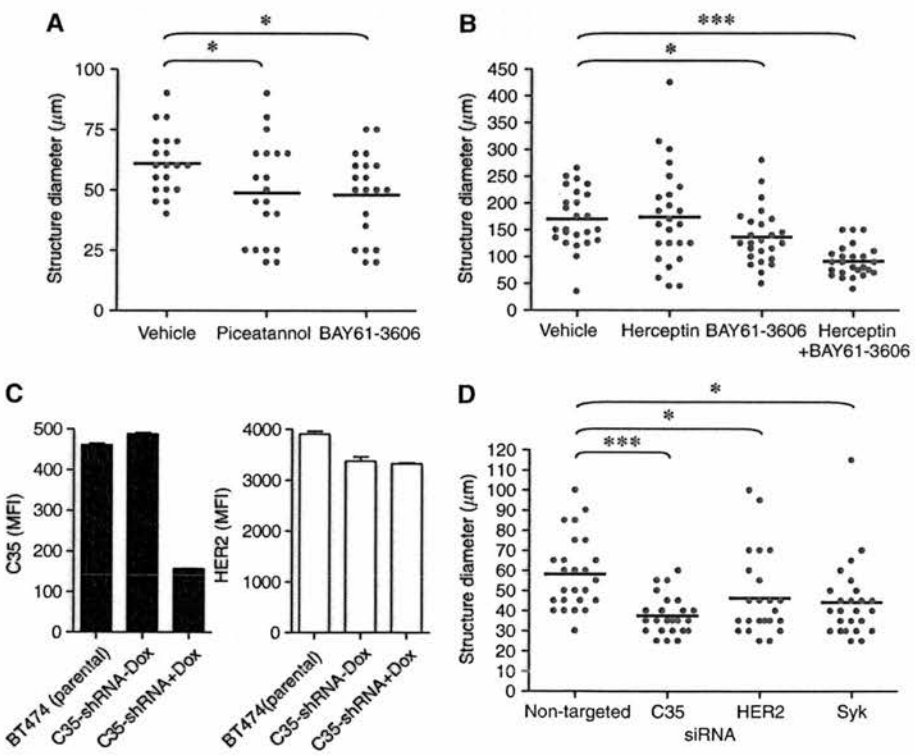


Figure 5 Inhibition of C35 and Syk reduces mammary epithelial cells acinar structure size. (**A**) Quantification of 3D structure size in T47D cells at day 14 after treatment with the Syk inhibitors BAY61-3606 or piceatannol. (**B**) Quantification of structure size in BT474 cells at day 13 after treatment with Herceptin (trastuzumab) and/or BAY61-3606. (**C**) Knockdown of C35 by siRNA in BT474 cells (left panel) has no effect on HER2 surface expression (right panel) as determined by flow cytometry. (**D**) Quantification of structure size of BT474 cells treated with non-targeted, C35, HER2 or Syk siRNA, at day 6 of 3D culture.

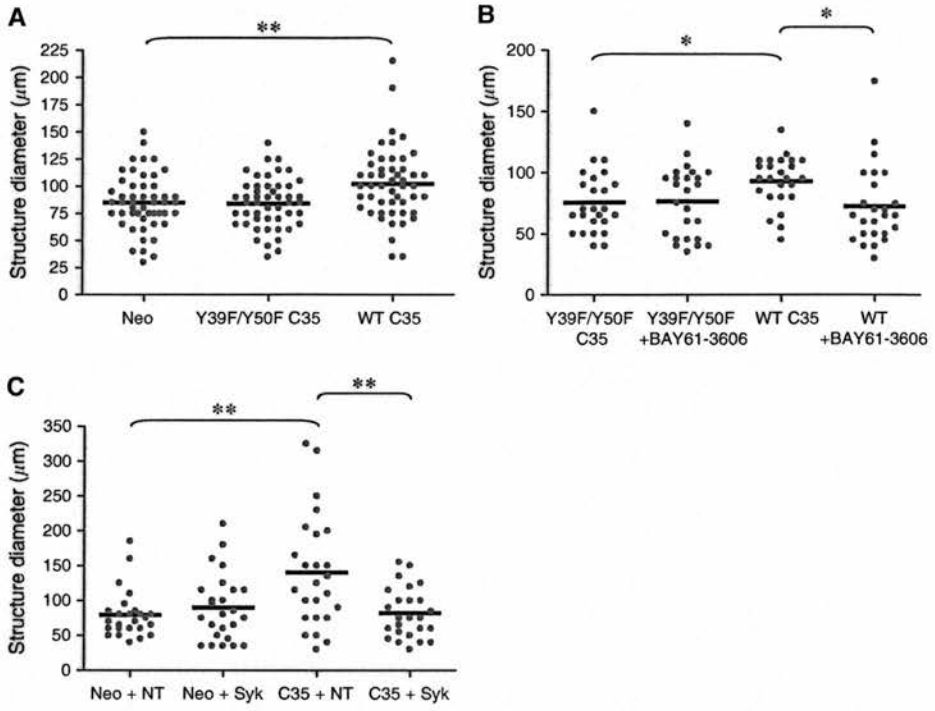


Figure 6 C35 expression in normal mammary epithelial cells leads to cell transformation in 3D cultures. (**A**) H16N-2 acinar structures (at day 11) expressing empty vector (Neo), Y39F/Y50F ITAM mutant or wild-type (wt) C35 protein. (**B**) Quantification of structure diameter in Y39F/Y50F ITAM mutant or wt C35-expressing cell lines at day 14 after treatment with the Syk inhibitor BAY61-3606 (50 nM, twice: at days 8 and 11). (**C**) Quantification of structure size of Neo and wt C35-expressing H16N-2 cells treated with non-targeted, Syk siRNA, at day 6 of 3D culture.

and C35 (Figure 6C). We also observed down-regulation of Syk mRNA in H16N-2 expressing the Y39F/Y50F ITAM mutant, compared with those expressing wt C35 (data not shown). This observation supports the view that Syk interacts with functional ITAM-containing C35.

DISCUSSION

HER2/ErbB2 amplification is a frequent and well-studied event in breast and other cancers. The genetic fragment being amplified is commonly known as the *HER2* amplicon. The smallest region of amplification of the *HER2/ERBB2* amplicon on human chromosome 17q12 contains 14 core genes, of which *STARD3*, *TCAP*, *PNMT*, *PERLD1*, *ERBB2*, *GRB7*, *GSDML* and *C17orf37*/C35 are over-expressed when amplified (Kauraniemi and Kallioniemi, 2006; Marchio *et al*, 2008). The function of HER2, in breast cancer in particular, has been subject to intense research efforts, culminating in the design of both small molecule inhibitors and monoclonal antibodies in treatment of $HER2^+$ patients (Bublil and Yarden, 2007). Recent efforts have concentrated on understanding how co-amplification of *HER2* with the non-core amplicon gene Topoisomerase II (*TOP2A*) may affect response to chemotherapy (Pritchard *et al*, 2008). Much less is known about the functional importance of the core genes co-amplified with HER2. One of the best studied of these core genes is *GRB7*. Co-expression of Grb7 and HER2 facilitates HER2 signal transduction and functions synergistically for tumour formation (Stein *et al*, 1994; Bai and Luoh, 2008). Tumours co-expressing high levels of Grb7 and HER2 have a worse outcome than those with only higher levels of HER2 (Nadler *et al*, 2010), in line with the clinical data presented here. Both *GRB7* and another core gene, *STARD3*, contribute to the growth of HER2-amplified cell lines *in vitro* (Kao and Pollack, 2006).

Here, we show that primary breast cancers have high levels of C35 protein expression when harbouring *HER2* gene amplification, and that over-expression of C35 and HER2 protein is correlated in both breast cancer cell lines and primary tumours, in agreement with previous findings (Evans *et al*, 2006). It is estimated that tumours can express 70–100 times the normal breast tissue C35 transcript level (Evans *et al*, 2006). Cell lines expressing high levels of C35 showed high invasive behaviour *in vitro*. The overall phenotype is consistent with EMT, including down-regulation of E-cadherin and up-regulation of Twist. Interestingly, more gene transcripts were down-regulated than up-regulated among the 100 most changed transcripts. This raises the possibility of common suppression mechanism of transcription, downstream of C35 expression. A study in a pancreatic cancer cell line has suggested that the protein inhibitor specific for HGF activator-1 (HAI-1), an HAI-2 homologue, may activate an EMT programme in these cells by up-regulating the transcription factor SIP-1/ZEB-2 and consequently repressing E-cadherin (Cheng *et al*, 2009).

We found that tyrosine mutation in the ITAM of C35, or Syk kinase inhibition, is sufficient to abolish the potential of C35 protein to cause enlargement of acinar structures in 3D cell culture. Studies in DLBCL lines have shown that some, but not all tumours, expressing ITAM-containing proteins may respond to Syk inhibition (Chen *et al*, 2008). Evidence in this study using C35-expressing MEC lines has supported this strategy *in vitro*. Syk expression and activation are also modulated by extracellular matrix, through integrin signalling (Zhang *et al*, 2009). Syk promotes cell–cell contact on plastic (Zhang *et al*, 2009) and its genetic knock-down promotes cell mobility and invasion (Sung *et al*, 2009; Zhang *et al*, 2009). Syk may also have a tumour suppressor function in breast cancer through its kinase activity in the nucleus (Coopman *et al*, 2000; Sung *et al*, 2009). A plausible mechanism is that the interaction of C35 with Syk mimics global knock-down of Syk by changing its localisation away from

integrins (Zhang *et al*, 2009) or the nucleus (Coopman *et al*, 2000). When activated in the cytoplasm, Syk functions as a promoter of cell growth (Zhou and Geahlen, 2009), consistent with the function postulated in this study. Our study results indicate that recently described Syk inhibitors (Braselmann *et al*, 2006; Chen *et al*, 2008) may be effective in C35 over-expressing breast cancer cells and thus have therapeutic value.

Other therapeutic approaches may be developed to take advantage of these findings in the treatment of human breast cancer, including the development of inhibitors of C35 interaction with proteins other than Syk, such as the novel ITAM-interacting protein Shb (Matskova *et al*, 2007). The Src kinase Lyn is unlikely to be involved in C35-induced EMT, because Lyn mRNA levels are reduced by approximately five-fold in $C35^{hi}$ cells, in comparison with both null and C35 transfectant pool cells.

In conclusion, we show here that the *HER2* amplicon contains a second oncogene, C35, in the context of breast cancer. Our observations suggest that targeting C35 as well as HER2 may be beneficial for patients with *HER2*-amplified breast cancers. C35/C17orf37 has recently been included in an expression signature predicting metastatic risk in node-negative breast cancer after chemotherapy (Jezequel *et al*, 2009). This signature does not include HER2, therefore suggesting a possible autonomous role for C35, and warrants further investigation.

ACKNOWLEDGEMENTS

We thank Alexey Larionov (Edinburgh Breakthrough Research Unit), Amy Sutton and Alan Howell (Vaccinex) for their technical assistance. We thank Jorge Reis-Filho for critically reading the manuscript. Supported by Scottish Funding Council and Breakthrough Breast Cancer. Clinical materials were obtained through the auspices of the Edinburgh Experimental Cancer Medicine Centre. For YMK.

Conflict of interest

ESS and EEE are employed by Vaccinex Inc., which identified C35 as a biomarker in human breast cancer. All other authors declare that they have no competing interests.

Supplementary Information accompanies the paper on British Journal of Cancer website (http://www.nature.com/bjc)

REFERENCES

Altman DG, Lausen B, Sauerbrei W, Schumacher M (1994) Dangers of using 'optimal' cutpoints in the evaluation of prognostic factors. *J Natl Cancer Inst* **86:** 829–835

Amjad SB, Carachi R, Edward M (2007) Keratinocyte regulation of TGF-beta and connective tissue growth factor expression: a role in suppression of scar tissue formation. *Wound Repair Regen* **15:** 748–755

Bai T, Luoh SW (2008) GRB-7 facilitates HER-2/Neu-mediated signal transduction and tumor formation. *Carcinogenesis* **29:** 473–479

Band V, Sager R (1991) Tumor progression in breast cancer. In *Neoplastic Transformation in Human Cell Culture*. Rhim JS, Dritschilo A (eds), pp 169–178. The Human Press: Totowa, New Jersey

Braselmann S, Taylor V, Zhao H, Wang S, Sylvain C, Balloum M, Qu K, Herlaar E, Lau A, Young C, Wong BR, Lovell S, Sun T, Park G, Argade A, Jurcevic S, Pine P, Singh R, Grossbard EB, Payan DG, Masuda ES (2006) R406, an orally available Syk kinase inhibitor blocks Fc receptor signaling and reduces immune complex-mediated inflammation. *J Pharmacol Exp Ther* **319**(3): 998–1008

Bublil EM, Yarden Y (2007) The EGF receptor family: spearheading a merger of signaling and therapeutics. *Curr Opin Cell Biol* **19:** 124–134

Burwell EA, McCarty GP, Simpson LA, Thompson KA, Loeb DM (2007) Isoforms of Wilms' tumor suppressor gene (WT1) have distinct effects on mammary epithelial cells. *Oncogene* **26:** 3423–3430

Camp RL, Chung GG, Rimm DL (2002) Automated subcellular localization and quantification of protein expression in tissue microarrays. *Nat Med* **8:** 1323–1327

Camp RL, Dolled-Filhart M, Rimm DL (2004) X-tile: a new bio-informatics tool for biomarker assessment and outcome-based cut-point optimization. *Clin Cancer Res* **10:** 7252–7259

Chen L, Monti S, Juszczynski P, Daley J, Chen W, Witzig TE, Habermann TM, Kutok JL, Shipp MA (2008) SYK-dependent tonic B-cell receptor signaling is a rational treatment target in diffuse large B-cell lymphoma. *Blood* **111:** 2230–2237

Cheng H, Fukushima T, Takahashi N, Tanaka H, Kataoka H (2009) Hepatocyte growth factor activator inhibitor type 1 regulates epithelial to mesenchymal transition through membrane-bound serine proteinases. *Cancer Res* **69:** 1828–1835

Coopman PJ, Do MT, Barth M, Bowden ET, Hayes AJ, Basyuk E, Blancato JK, Vezza PR, McLeskey SW, Mangeat PH, Mueller SC (2000) The Syk tyrosine kinase suppresses malignant growth of human breast cancer cells. *Nature* **406:** 742–747

Dasgupta S, Wasson LM, Rauniyar N, Prokai L, Borejdo J, Vishwanatha JK (2009) Novel gene C17orf37 in 17q12 amplicon promotes migration and invasion of prostate cancer cells. *Oncogene* **28:** 2860–2872

Debnath J, Brugge JS (2005) Modelling glandular epithelial cancers in three-dimensional cultures. *Nat Rev Cancer* **5:** 675–688

Debnath J, Muthuswamy SK, Brugge JS (2003) Morphogenesis and oncogenesis of MCF-10A mammary epithelial acini grown in three-dimensional basement membrane cultures. *Methods* **30:** 256–268

Dunning MJ, Smith ML, Ritchie ME, Tavare S (2007) Beadarray: R classes and methods for Illumina bead-based data. *Bioinformatics* **23:** 2183–2184

Evans EE, Henn AD, Jonason A, Paris MJ, Schiffhauer LM, Borrello MA, Smith ES, Sahasrabudhe DM, Zauderer M (2006) C35 (C17orf37) is a novel tumor biomarker abundantly expressed in breast cancer. *Mol Cancer Ther* **5:** 2919–2930

Faratian D, Goltsov A, Lebedeva G, Sorokin A, Moodie S, Mullen P, Kay C, Um IH, Langdon S, Goryanin I, Harrison DJ (2009) Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab. *Cancer Res* **69:** 6713–6720

Foos G, Garcia-Ramirez JJ, Galang CK, Hauser CA (1998) Elevated expression of Ets2 or distinct portions of Ets2 can reverse Ras-mediated cellular transformation. *J Biol Chem* **273:** 18871–18880

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* **5:** R80

Grande SM, Katz E, Crowley JE, Bernardini MS, Ross SR, Monroe JG (2006) Cellular ITAM-containing proteins are oncoproteins in nonhematopoietic cells. *Oncogene* **25:** 2748–2757

Hennessy BT, Gonzalez-Angulo AM, Stemke-Hale K, Gilcrease MZ, Krishnamurthy S, Lee JS, Fridlyand J, Sahin A, Agarwal R, Joy C, Liu W, Stivers D, Baggerly K, Carey M, Lluch A, Monteagudo C, He X, Weigman V, Fan C, Palazzo J, Hortobagyi GN, Nolden LK, Wang NJ, Valero V, Gray JW, Perou CM, Mills GB (2009) Characterization of a naturally occurring breast cancer subset enriched in epithelial-to-mesenchymal transition and stem cell characteristics. *Cancer Res* **69:** 4116–4124

Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, Hu Z, Rasmussen KE, Jones LP, Assefnia S, Chandrasekharan S, Backlund MG, Yin Y, Khramtsov AI, Bastein R, Quackenbush J, Glazer RI, Brown PH, Green JE, Kopelovich L, Furth PA, Palazzo JP, Olopade OI, Bernard PS, Churchill GA, Van Dyke T, Perou CM (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. *Genome Biol* **8:** R76

Huang DW, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4:** 44–57

Jezequel P, Campone M, Roche H, Gouraud W, Charbonnel C, Ricolleau G, Magrangeas F, Minvielle S, Geneve J, Martin AL, Bataille R, Campion L

(2009) A 38-gene expression signature to predict metastasis risk in node-positive breast cancer after systemic adjuvant chemotherapy: a genomic substudy of PACS01 clinical trial. *Breast Cancer Res Treat* **116:** 509–520

Kao J, Pollack JR (2006) RNA interference-based functional dissection of the 17q12 amplicon in breast cancer reveals contribution of coamplified genes. *Genes Chromosomes Cancer* **45:** 761–769

Katz E, Lareef MH, Rassa JC, Grande SM, King LB, Russo J, Ross SR, Monroe JG (2005) MMTV Env encodes an ITAM responsible for transformation of mammary epithelial cells in three-dimensional culture. *J Exp Med* **201:** 431–439

Kauraniemi P, Kallioniemi A (2006) Activation of multiple cancer-associated genes at the ERBB2 amplicon in breast cancer. *Endocr Relat Cancer* **13:** 39–49

Kenny PA, Lee GY, Myers CA, Neve RM, Semeiks JR, Spellman PT, Lorenz K, Lee EH, Barcellos-Hoff MH, Petersen OW, Gray JW, Bissell MJ (2007) The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression. *Mol Oncol* **1:** 84–96

Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S, Torhorst J, Mihatsch MJ, Sauter G, Kallioniemi OP (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* **4:** 844–847

Levayer R, Lecuit T (2008) Breaking down EMT. *Nat Cell Biol* **10:** 757–759

Lu J, Lin WH, Chen SY, Longnecker R, Tsai SC, Chen CL, Tsai CH (2006) Syk tyrosine kinase mediates Epstein–Barr virus latent membrane protein 2A-induced cell migration in epithelial cells. *J Biol Chem* **281:** 8806–8814

Marchio C, Natrajan R, Shiu KK, Lambros MB, Rodriguez-Pinilla SM, Tan DS, Lord CJ, Hungermann D, Fenwick K, Tamber N, Mackay A, Palacios J, Sapino A, Buerger H, Ashworth A, Reis-Filho JS (2008) The genomic profile of HER2-amplified breast cancers: the influence of ER status. *J Pathol* **216:** 399–407

Matskova LV, Helmstetter C, Ingham RJ, Gish G, Lindholm CK, Ernberg I, Pawson T, Winberg G (2007) The Shb signalling scaffold binds to and regulates constitutive signals from the Epstein–Barr virus LMP2A membrane protein. *Oncogene* **26:** 4908–4917

Nadler Y, Gonzalez AM, Camp RL, Rimm DL, Kluger HM, Kluger Y (2010) Growth factor receptor-bound protein-7 (Grb7) as a prognostic marker and therapeutic target in breast cancer. *Ann Oncol* **21:** 466–473

Pritchard KI, Messersmith H, Elavathil L, Trudeau M, O'Malley F, Dhesy-Thind B (2008) HER-2 and topoisomerase II as predictors of response to chemotherapy. *J Clin Oncol* **26:** 736–744

Rhodes DR, Ateeq B, Cao Q, Tomlins SA, Mehra R, Laxman B, Kalyana-Sundaram S, Lonigro RJ, Helgeson BE, Bhojani MS, Rehemtulla A, Kleer CG, Hayes DF, Lucas PC, Varambally S, Chinnaiyan AM (2009) AGTR1 overexpression defines a subset of breast cancer and confers sensitivity to losartan, an AGTR1 antagonist. *Proc Natl Acad Sci USA* **106:** 10284–10289

Ross SR, Schmidt JW, Katz E, Cappelli L, Hultine S, Gimotty P, Monroe JG (2006) An immunoreceptor tyrosine activation motif in the mouse mammary tumor virus envelope protein plays a role in virus-induced mammary tumors. *J Virol* **80:** 9000–9008

Stein D, Wu J, Fuqua SA, Roonprapunt C, Yajnik V, D'Eustachio P, Moskow JJ, Buchberg AM, Osborne CK, Margolis B (1994) The SH2 domain protein GRB-7 is co-amplified, overexpressed and in a tight complex with HER2 in breast cancer. *EMBO J* **13:** 1331–1340

Sung YM, Xu X, Sun J, Mueller D, Sentissi K, Johnson P, Urbach E, Seillier-Moiseiwitsch F, Johnson MD, Mueller SC (2009) Tumor suppressor function of Syk in human MCF10A in vitro and normal mouse mammary epithelium in vivo. *PLoS One* **4**(10): e7445

Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* **98:** 5116–5121

Underhill DM, Goodridge HS (2007) The many faces of ITAMs. *Trends Immunol* **28:** 66–73

Wang L, Dittmer DP, Tomlinson CC, Fakhari FD, Damania B (2006) Immortalization of primary endothelial cells by the K1 protein of Kaposi's sarcoma-associated herpesvirus. *Cancer Res* **66:** 3658–3666

Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* **25:** 118–145

Yamamoto N, Takeshita K, Shichijo M, Kokubo T, Sato M, Nakashima K, Ishimori M, Nagai H, Li YF, Yura T, Bacon KB (2003) The orally available spleen tyrosine kinase inhibitor 2-[7-(3,4-dimethoxyphenyl)-imidazo [1,2-c]pyrimidin-5-ylamino]nicotinamide dihydrochloride (BAY61-3606) blocks antigen-induced airway inflammation in rodents. *J Pharmacol Exp Ther* **306:** 1174–1181

Zhang X, Shrikhande U, Alicie BM, Zhou Q, Geahlen RL (2009) Role of the protein tyrosine kinase Syk in regulating cell-cell adhesion and motility in breast cancer cells. *Mol Cancer Res* **7**(5): 634–644

Zhou Q, Geahlen RL (2009) The protein-tyrosine kinase Syk interacts with TRAF-interacting protein TRIP in breast epithelial cells. *Oncogene* **28:** 1348–1356

OPEN ACCESS Freely available online

PLoS one

An *In Vitro* Model That Recapitulates the Epithelial to Mesenchymal Transition (EMT) in Human Breast Cancer

Elad Katz[1,2]*, Sylvie Dubois-Marshall[1,2], Andrew H. Sims[1], Philippe Gautier[3], Helen Caldwell[1,2], Richard R. Meehan[1,3], David J. Harrison[1,2]

1 Breakthrough Breast Cancer Research Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom, **2** Division of Pathology, Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, United Kingdom, **3** MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, Edinburgh, United Kingdom

Abstract

The epithelial to mesenchymal transition (EMT) is a developmental program in which epithelial cells down-regulate their cell-cell junctions, acquire spindle cell morphology and exhibit cellular motility. In human breast cancer, invasion into surrounding tissue is the first step in metastatic progression. Here, we devised an *in vitro* model using selected cell lines, which recapitulates many features of EMT as observed in human breast cancer. By comparing the gene expression profiles of claudin-low breast cancers with the experimental model, we identified a 9-gene signature characteristic of EMT. This signature was found to distinguish a series of breast cancer cell lines that have demonstrable, classical EMT hallmarks, including loss of E-cadherin protein and acquisition of N-cadherin and vimentin expression. We subsequently developed a three-dimensional model to recapitulate the process of EMT with these cell lines. The cells maintain epithelial morphology when encapsulated in a reconstituted basement membrane, but undergo spontaneous EMT and invade into surrounding collagen in the absence of exogenous cues. Collectively, this model of EMT *in vitro* reveals the behaviour of breast cancer cells beyond the basement membrane breach and recapitulates the *in vivo* context for further investigation into EMT and drugs that may interfere with it.

Citation: Katz E, Dubois-Marshall S, Sims AH, Gautier P, Caldwell H, et al. (2011) An *In Vitro* Model That Recapitulates the Epithelial to Mesenchymal Transition (EMT) in Human Breast Cancer. PLoS ONE 6(2): e17083. doi:10.1371/journal.pone.0017083

Editor: Syed Aziz, Health Canada, Canada

Received November 9, 2010; **Accepted** January 14, 2011; **Published** February 15, 2011

Funding: Clinical materials were obtained through the auspices of the Edinburgh Experimental Cancer Medicine Centre. This work was supported by Breakthrough Breast Cancer and the Scottish Funding Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: elad.katz@ed.ac.uk

Introduction

Breast cancer related deaths are primarily due to metastatic progression [1]. Understanding the mechanisms that underlie this multistep process is essential to improving clinical outcome. The transformation of normal breast epithelial cells to metastatic cancer is the result of multiple epigenetic and genetic changes, leading to deregulated interactions with the microenvironment [2]. During this process, inhibition of proliferation, cell survival, migration and differentiation is lost leading to the acquisition of an invasive phenotype. The ability to breach the basement membrane (BM) is a critical event in cancer progression and a prerequisite for metastasis. Having breached the BM, cells may then enter the lymphatic system, spread and attempt to establish themselves as distant tumor foci [3].

The trans-differentiation of cells from an epithelial to a mesenchymal phenotype is an essential part of normal embryogenesis and development [4]. Increasing evidence also supports a role for epithelial to mesenchymal transition (EMT) in the progression of many cancer types including breast, with critical roles in invasion and metastatic dissemination [5,6]. EMT involves loss of cell-cell junctions and re-organization of the actin cytoskeleton, resulting in loss of apical-basal polarity and acquisition of a spindle-like mesenchymal morphology [7]. At the same time, there is also decreased expression of epithelial-specific proteins, including E-cadherin, which may account at least in part for the altered properties of migrating tumor cells [8,9]. An important event in EMT is switching in expression from E-cadherin to N-cadherin [10]. In most cases this is associated with transcriptional repression of E-cadherin [9]. Several specific repressor factors have been identified including Snail, Slug, Zeb1, Zeb2 and Twist [11], all of which are zinc finger containing proteins that can bind with so called E-boxes within the *CDH1* gene promoter. N-cadherin is believed to promote cellular invasion by binding to and enhancing signalling by growth factors and is over-expressed in many invasive and metastatic human breast cancer cell lines and tumors [10,12,13].

Comparative analysis of mouse mammary carcinoma models and human breast tumors identified a novel human molecular subtype, termed 'claudin-low' cancers. These cancers are characterised by low to absent expression of genes involved in tight junctions and cell-cell adhesions, including claudins, occludins and E-cadherin [14,15]. In addition, these moderate-high grade invasive ductal carcinomas are morphologically distinct from lobular carcinomas despite their low expression of E-cadherin [14]. Similarities between claudin-low tumors and EMT *in vitro* have been documented, however these features have not previously been compared and analysed directly. Furthermore, while the contribution of the extra-cellular matrix to the

promotion of tumor progression is now appreciated [2], most current *in vitro* models do not take into account the contribution of stromal collagen into which cells undergoing EMT invade. The predisposition of tumours to undergo EMT can be enhanced by genetic alterations. For example, C35 is a 12KDa membrane-anchored protein found on the HER2 amplicon that is over-expressed in around 11% of breast cancers [16]. Cellular transformation associated with acquisition of an EMT phenotype can be induced in mammary epithelial cells transfected with a C35 expression construct resulting in increased invasion into stromal collagen, down regulation of E-cadherin and up regulation of the transcription repressor Twist [17]. This implies that collagen-invading C35-expressing cells can be used to model aspects of EMT in cancer cells.

Testing new treatments that may prevent EMT or tumor spread is challenging: conventional clinical trials may have difficulty in addressing the issues because of the ethical problems of leaving tumor *in situ*, or the limitation of a study to only very late stage disease. Robust models that can identify possible predictive biomarkers are essential. In this report, we describe a unique invasion assay, in which cell lines with known molecular pathology undergo spontaneous EMT when invading away from the basement membrane into collagen. We propose that this *in vitro* model of defined breast cancer cell lines can provide an improved representation of invasive breast cancer *in vivo*, compared to existing EMT models.

Materials and Methods

Gene expression analysis, RNA extraction and qRT-PCR

Microarray data was analysed using packages within Bioconductor [18] (http://www.bioconductor.org) that implement R statistical programming. Gene expression data was normalised using quantile normalisation within the BeadArray package [19] and differential gene expression assessed using Significance Analysis of Microarrays (SAM) [20] within the siggenes package. The dataset from Hershkowitz and colleagues [14] was downloaded from the UNC Microarray Database (https://genome.unc.edu/). RNA from the collagen invasion assays was labelled using a Illumina TotalPrep RNA amplification kit (Ambion) according to manufacturer's instructions. Triplicate samples from invasion assays (1500 ng cDNA per assay) were hybridised to Illumina BeadChips and whole genome gene expression analysis performed using the Illumina HumanRef-8 v3 Expression BeadChip and BeadArray Reader.

RNA from cell lines cultured on plastic was converted to cDNA prior to PCR using a QuantiTect Reverse Transcription kit (Qiagen). Gene expression patterns for invasion assays (biological triplicates) and cell lines cultured on plastic (technical triplicates) were examined using the QuantiTect SYBR Green PCR kit (Qiagen) and a Corbett RotoGene 3000. Primers for *CDH1* were: forward 5′-CGGAGAAGAGGACCAGGACT-3′, reverse 5′-GGTCAGTATCAGCCGCTTTC-3′; for *CLDN7*: forward 5′-AAAATGTACGACTCGGTGCTC-3′, reverse 5′-AGACCTGC-CACGATGAAAAT; for TBP: forward 5′-GGGGAGCTGTGA-TGTGAAGT-3′, reverse 5′-CCAGGAAATAACTCTGGCTCA-3′; for *ACTB*: forward 5′-CCTTCCTGGGCATGGAGTCCT-3′, reverse 5′-GGAGCAATGATCTTGATCTT-3′. QuantiTect Primer Assays (Qiagen) were used for *KRT8, CRB3, MARVELD3, IRF6, MAL2, TACSTD1* and *SPINT2*. PCR program was identical for all genes: 95°C, 15 min; (94°C, 15 s; 56°C, 30 s; 72°C, 30 s)×50 cycles; 72°C, 5 min. Standard reference human cDNA was from Clontech, random primed. ~50 ng RNA equiv/mL was used for quantification of mRNA expression. Final normalisation was performed against the geometrical mean of *ACTB* and *TBP* levels.

Gene promoter analysis

Using the presumptive promoter region for the 9 genes (a 2 kb region upstream of the presumptive transcription start site using Ensembl 52, Jan2009, based on NCBI 36 assembly), we looked for over-represented 6- and 7-mers oligos using oligo-analysis [21] from the RSAT-tools package (http://rsat.scmbb.ulb.ac.be/rsat/) [22]. The program counts all oligonucleotide occurrences within the sequence set, and estimates their statistical significance. A calibration is done using the entire genome promoter regions as a background model (Ensembl 52, Jan2009, based on NCBI 36 assembly). For the best 7-mers candidates, we compared the obtained oligo sequences to the entire collection of consensus binding sites available in Transfac professional [23] (release 2010.1) using the compare-pattern script (RSAT-tools) and listed the associated binding factor name.

E-value for best hit 7-mer CAGGTGC/GCACCTG (2.6×10^{-8}) represents the expected number of patterns which would be returned at random for a given probability. The weights in Table 1 reflect the number of matching positions, with a lower weight for matches between partially specified nucleotides (the weight for a perfect match to a 7-mer is 7). Both E-value and weights are calculated by RSAT-tools.

Cell lines

MCF10A, Hs578T, HBL100, BT549, MDA-MB157, MDA-MB231 and MDA-MB436 cell lines were obtained from American Type Culture Collection. SUM159PT and SUM1315MO2 cells were a kind gift from Akira Orimo (University of Manchester). The cells were cultured as previous described [24] at 37 deg C, 5% CO2: MCF10A in DMEM/F12 media (Invitrogen) with 5% horse serum (Invitrogen), 20 ng/ml EGF, 100 ng/ml cholera toxin, 0.01 mg/ml insulin and 500 ng/ml hydrocortisone (all from Sigma); MDA-MB157, MDA-MB231, HBL100 and HS578T in DMEM, 10% bovine serum (both from Invitrogen); SUM159PT in Ham's F12 (Invitrogen), 5% bovine serum, insulin, hydrocortisone; MDA-MB436 in L15 (Invitrogen), 10% bovine serum; BT549: RPMI-1640, 10% bovine serum; SUM1315MO2 in Ham's F12, 5% bovine serum, insulin, EGF.

Primary cell isolation for tissue culture

Fresh normal breast tissue and breast tumor materials were incubated for 1 hour at room temperature in tissue mix consisting of DMEM/F12, 1% fungizone, 1000 U/ml penicillin, 1000 µg/ml streptomycin, 10 µg/ml insulin and 10% bovine serum (all from Invitrogen). Tissue cores were then finely chopped (~1 ml pieces) and put in a tissue mix/Collagenase I solution (Invitrogen; made up with 200 µL of 200 U/ml Collagenase I to 20 ml tissue mix) for digestion (2 hours at 37 deg C, 200 rpm). The digested tissue was then spun for 4 mins at 60 g. The resulting pellet was plated with fibroblast media (DMEM supplemented with 10% bovine serum, 50 U/ml penicillin and 50 mg/ml streptomycin) and the supernatant spun for a further 4 mins at 600 g, 4 times. The resulting second pellet (mammary epithelial cells) was plated with HMEC media (CnT-22 (Cellntec) supplemented with 5% FCS).

Ethics Statement

The use of primary breast cells was approved by the Lothian Research Ethics Committee (08/S1101/41). Materials were obtained with written informed consent from all participants involved in this study.

Rat tails obtained from animals at the University of Edinburgh animal facilities scarified for other scientific purposes and did not require ethical approval.

Table 1. Common transcription factor binding sites in the 9-gene signature.

Best hit	Weight	Matrix consensus	Transfac ID	Factor name
GCACCTG	6.5	ASCACCTGTTNNCA	M00044	Snail*
CAGGTGC	6.5	RACAGGTGYA	M00060	Snail*
GCACCTG	6.5	VNRCACCTGKNC	M00414	AREB6/ZEB1*
CAGGTGC	6.21	CNNCAGGTGB	M00277	LMO2 complex*
CAGGTGC	6	RRCAGGTGNCV	M00693	E12/ELSPBP1*
CAGGTGC	6	CNGNRNCAGGTGNNGNA	M00929	MyoD*
GCACCTG	5.5	YNYACCTGWVT	M00412	AREB6/ZEB1*
GCACCTG	4	RRTGNMCYTNNTGAMCCNYNT	M00966	VDR, CAR, PXR
GCACCTG	3.5	GCTGGNTNGNNCYNG	M00947	CP2/LBP-1c/LSF
GCACCTG	3.5	RGNACNNKNTGTTCT	M00957	PR/Progesterone receptor
GCACCTG	3.5	TGGCASNNNGCCAA	M01196	CTF1

A list of binding sites matching to best 7-mer found in promoters of the common EMT gene signature. Three muscle initiator sequences with no further information were excluded.

*E-box binding transcription factors. E12 is part of the LMO2 complex.

doi:10.1371/journal.pone.0017083.t001

SDS-PAGE

Protein lysates (50 µg/well, as determined by MicroBCA protein assay) were resolved by SDS-PAGE after being denatured for 1 hour at 60 deg C. The resolving gel (7.5% w/v acrylamide, 0.37 M TRIS pH 8.85, 0.1% SDS, 0.02% AMPS, 0.25% TEMED; all from Sigma) was set between glass plates using a Bio-Rad kit. Once the resolving gel had set, a stacking gel (3.6% w/v acrylamide, 0.12 M TRIS pH 6.8, 0.1% SDS, 0.03% AMPS, 0.33% TEMED) was layered and a comb used to create wells for sample loading. The loaded samples were electro-separated under constant current (100–200 mA) using electrophoresis buffer (25 mM Trizma Base, 0.19 M Glycine, 10% SDS). Electro-transfer onto immobilon transfer membrane (Millipore) was performed using transfer buffer (25 mM Trizma Base, 0.19 M Glycine) using a Bio-Rad kit, under constant electrical potential (~30 mV for at least 2 hours).

Western Blotting

Nonspecific binding was blocked with Li-Cor Odyssey Blocking Buffer (Li-Cor), diluted 50:50 in PBS, for 1 hour at room temperature. Primary antibodies were diluted in Li-Cor Odyssey Blocking Buffer, diluted 50:50 in 0.1% PBS-Tween20, and incubated with the blot overnight at 4 deg C. Blots were washed 3 times for 5 mins with PBS-T before incubation with appropriate fluorescent secondary antibodies (Li-Cor), diluted 1:10,000 in Li-Cor Odyssey Blocking Buffer, diluted 50:50 in 0.1% PBS-Tween20, for 45 mins at room temperature. Exposure to light was avoided. Subsequently, membranes were washed, dried and scanned on the Li-Cor Odyssey scanner. All washes/incubations were carried out under constant agitation. Primary antibodies used as follows: E-cadherin, BD, 610181, Mouse, 1:2500; Claudin7, Abcam, Ab75347, Rabbit, 1:1000; N-cadherin, BD, 610921, Mouse, 1:3000; Vimentin, Sigma, V 6630, Mouse, 1:1000; Zeb2, BD, 611256, Mouse, 1:250; Slug, LifeSpan Bio, LS-C30318, Rabbit, 1:4000; Snail, Abcam, ab17732, Rabbit, 1:4000; Tubulin, Abcam, Ab7291, Mouse, 1:6000.

Rat tail collagen I preparation

Fresh rat tails were collected and frozen. Prior to harvesting these were placed in 70% ethanol. Tendons were stripped from the tails and returned to 70% ethanol to sterilise. The collected tendons were weighed and transferred to the appropriate volume of pre-cooled acetic acid (1 g tendon to 250 ml 0.5 M acetic acid) and gently stirred for 48 hours at 4 deg C. The tendon/acetic acid mix was then centrifuged at 10,000 g for 30 mins and the pellet discarded. An equal volume of 10% (w/v) NaCl was added to the supernatant and the mix allowed to stand overnight at 4 deg C. The collagen-rich, insoluble 'bottom layer' was taken and collected by further centrifugation (10,000 g for 30 mins). The collagen-rich material was resuspended in 0.25 M acetic acid at 4 deg C and dialysed against 1:1000 acetic acid at 4 deg C for 3 days, changing the dialysis buffer twice daily. The collagen solution was then sterilised by centrifugation (20,000 g for 2 hours) and stored at 4 deg C. Collagen was diluted as required by the addition of sterile 1:1000 acetic acid to a stock concentration of 1.2 mg/ml.

Establishment of 3D invasion assays

200 µL cell-collagen plugs and 75 µL cell-Matrigel plugs were made in a u-shaped 96 well plate, with the aim of achieving comparable size after a 24 hr incubation (day -1). A cell concentration of 1×10^6 was used for all plugs. Rat tail collagen I, for both plugs and subsequent embedding, was prepared as per the 'on top' assays. Growth factor reduced Matrigel was obtained from BD and used at a 5 mg/ml. Matrigel matrix is a soluble basement membrane extract of the Engelbreth-Holm-Swarm tumor that gels at room temperature to form a reconstituted basement membrane. The major components are laminin, collagen IV, entactin and heparin sulphate membrane. After the 24 hr incubation, cell plugs were carefully removed from their 96 well plate and embedded in 1 ml of collagen in a 24 well plate (taken as day 0), with or without fibroblasts (used at 10,000/ml). These cultures were incubated for a further hour and then carefully freed from the edges of the well (to allow contraction of the collagen) and supplemented with 0.5 ml of cell-specific media. The cultures were then left to invade. Media was changed weekly. Gels were fixed at either 1 or 2 weeks in 10% phosphate buffered formalin and wax embedded.

Immunofluorescence

Immunofluorescence was preformed as described previously [17]. Briefly, antigen retrieval for all epitopes was carried out using heat treatment under pressure in a microwave oven for 5 min in

citrate buffer (82 ml 0.01 M sodium citrate: 18 ml 0.01 M citric acid) pH 6.0. Slides were incubated with primary antibodies for 1 hr at room temperature. Primary antibodies were as follows: E-cadherin, BD, 610181, Mouse, 1:1500; N-cadherin, BD, 610921, Mouse, 1:300; Zeb2, BD, 611256, Mouse, 1:50; Slug, LifeSpan Bio, LS-C30318, Rabbit, 1:1000; Snail, Abcam, ab17732, Rabbit, 1:700. For Snail staining, mouse anti-pancytokeratin (Invitrogen, 1:25) was added to visualise epithelial cells. Mouse primary antibodies were incubated overnight with rabbit anti-pancytokeratin (Dako, 1:150). The epithelial compartment was then visualised with Cy3 (Invitrogen, anti-rabbit; anti-mouse, both used at 1:25). DAPI (4′,6-diamidino-2-phenylindole) counterstain (Invitrogen) was used to identify nuclei and Cy-5-tyramide (HistoRx, 1:50) was used to detect protein 'targets'. Monochromatic images of each TMA core were captured at 20× objective using an Olympus AX-51 epifluorescence microscope, and high-resolution digital images were analyzed by the AQUAnalysis software [17].

Results

Identification of a common EMT signature in the breast

In order to establish an *in vitro* EMT signature, we identified a set of 57 genes that strongly correlated with C35-induced EMT *in vitro* using significance analysis of microarrays (SAM, [20]). These 'C35 genes' were subsequently found to be sufficient to cluster claudin-low tumors together in a breast cancer dataset [14] (Figs. 1 and S1). In addition, a 34 gene 'claudin-low' signature identified in murine mammary carcinoma and human breast tumors [14], was significantly down-regulated in collagen-invading C35-expressing cells in comparison to parental cells (range $p = 0.048$ to $p = 1 \times 10^{-8}$; Figs. 1 and S1). Nine genes were common between the 'C35 genes' and 'claudin-low genes' signatures (Fig. 1): *CDH1*, *CLDN7*, *CRB3*, *KRT8*, *TACSTD1*, *IRF6*, *SPINT2*, *MAL2* and *MARVELD3*. Five of these, *CDH1* (E-cadherin), *CLDN7* (Claudin-7), *TACSTD1* (EpCAM), *IRF6* and *KRT8* (Keratin-8) have been previously implicated by their low expression in claudin-low cancers and/or in EMT *in vitro* [25,26,27]. *SPINT2* (Hepatocyte growth factor activation inhibitor-2, HAI-2) is capable of regulating a HGF-induced invasion of human breast cancer cells [28]. Two novel genes found to be down-regulated: the apical sorting protein *MAL2* [29] and its tight-junction-associated homologue *MARVELD3* [30].

We determined whether the nine EMT genes share common regulatory elements in their promoters and identified a shared 7-mer: CAGGTGC/GCACCTG. This binding motif is targeted by E-box transcription repressors, including Snail and ZEB families (Table 1) raising the possibility that these transcription factors repress all nine genes in the EMT pathway both *in vitro* and *in vivo*.

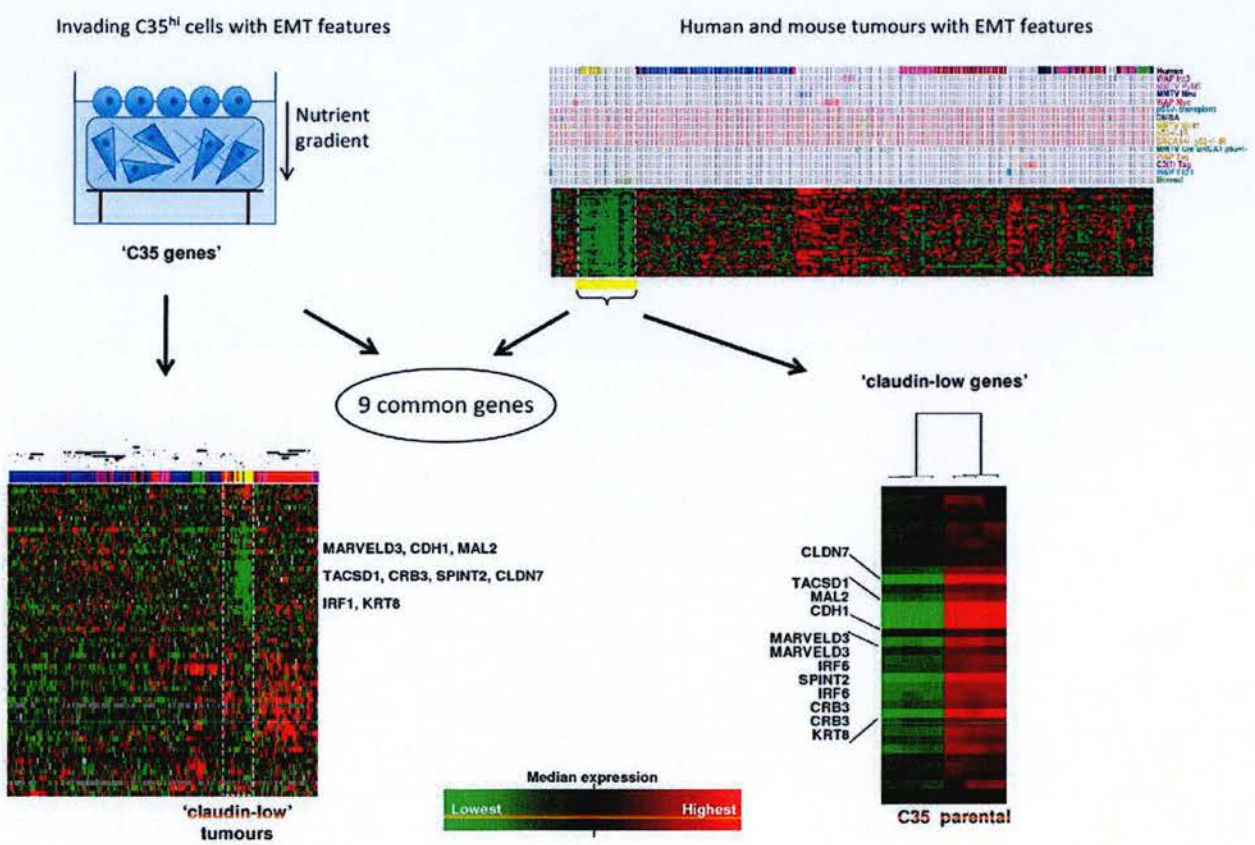


Figure 1. Comparison of genes correlating with C35 expression and those identifying the claudin-low phenotype identifies a 9-gene EMT signature. The 100 illumina probes most significantly differentially expressed between collagen-invading C35 and parental cells were represented by 57 genes that were able to cluster together the 13 claudin-low tumors identified by Herschkowitz and colleagues (*left panels*). A set of 34 claudin-low genes from the Herschkowitz were all significantly down-regulated in C35-expressing cells compared to parental cells (*right panels*). A signature of nine EMT-related genes is shared between the C35 and claudin-low gene lists (full lists in Fig. S1).
doi:10.1371/journal.pone.0017083.g001

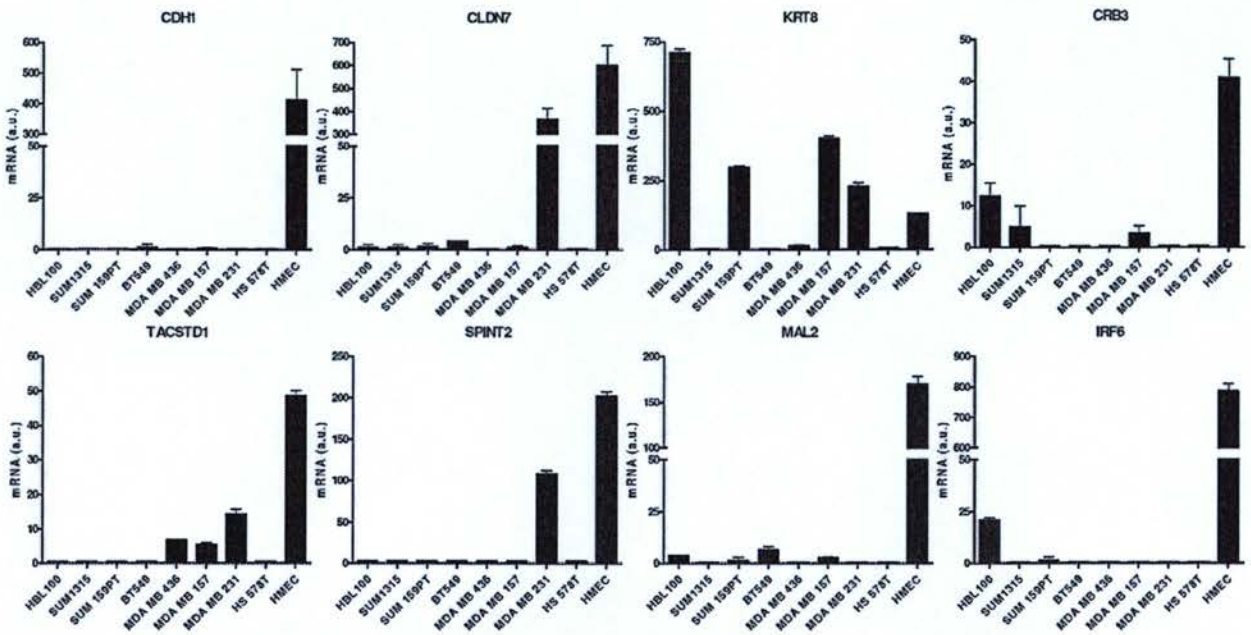


Figure 2. The 9-gene C35/claudin-low signature is down-regulated in a subset of human breast cell lines. Eight cell lines exhibit low expression of *CDH1, CLDN7, CRB3, KRT8, TACSTD1, IRF6, SPINT2* and *MAL2* when cultured on plastic. *MARVELD3* could not be assessed due to particularly low levels of expression. Technical triplicate mRNA expression data is shown for each line. HMEC cDNA is shown for comparison.
doi:10.1371/journal.pone.0017083.g002

Identification of cell lines with 'claudin-low' features

The 9-gene signature was identified in nine breast cancer cell lines from a previously published gene expression dataset [24] that all expressed low levels of the EMT genes: cell lines BT549, Hs578T, HBL100, MDA-MB157, MDA-MB231, MDA-MB435, MDA-MB436, SUM1315MO2 and SUM159PT respectively. We excluded the MDA-MB435 line from this cohort of cell lines due to doubts as to its tissue of origin [31]. The remaining eight cell lines show clear mesenchymal morphology when cultured on plastic (Fig. S2). We confirmed down-regulation of eight of the nine EMT genes by quantitative RT-PCR (Fig. 2) using normal human mammary epithelial cells (HMECs) as a positive control. We also validated low expression of these genes in the C35 model ([17] and data not shown).

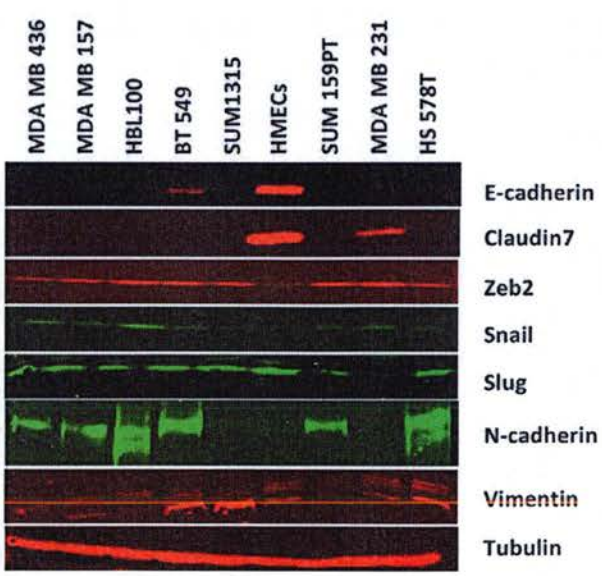


Figure 3. Claudin-low-like cell lines express key markers of EMT. Western blots demonstrate the expression of EMT-related markers at the protein level. HMEC lysates are shown for comparison.
doi:10.1371/journal.pone.0017083.g003

Western blotting was used to investigate the expression patterns of EMT-related proteins, including transcription repressors. All lines exhibit low levels of E-cadherin and Claudin-7 in comparison to normal mammary epithelial cells (Fig. 3), whereas ZEB2 (SIP-1), an E-box transcription factor that can induce EMT, is expressed in all the cell lines with claudin-low features. Most of the cell lines also have detectable expression of Snail, whereas Slug is absent in only one (MDA-MB231). Lastly, all of the cell lines express the mesenchymal marker vimentin and seven of the cell lines have detectable expression of N-cadherin.

A 3D invasion assay that mimics invasion into stromal collagen

A critical event in cancer progression is the acquisition of an invasive phenotype, and in particular the ability to breach the basement membrane (BM) into the stromal collagen. We developed a 3D model that attempts to mimic this process. Histologically normal breast epithelial cells are first embedded in a laminin-rich, BM-like Matrigel to generate a cell 'plug' which was subsequently embedded in collagen to mimic the surrounding extracellular matrix (Fig. 4a). This model potentially generates a three-stage assay that allows investigation of cells: i) contained by BM; ii) as they invade across BM; iii) as they invade more distally into surrounding collagen. In addition, the movement of cells in a horizontal plane can easily be followed by light microscopy, in

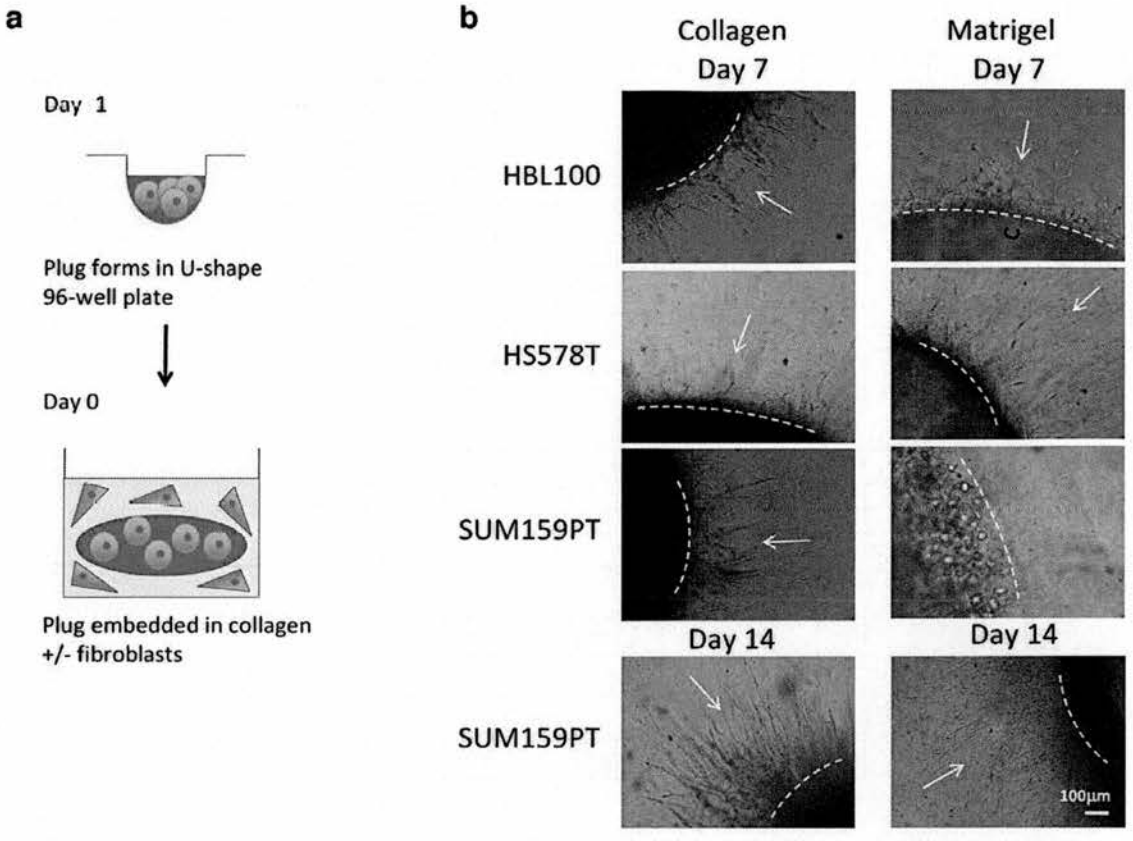


Figure 4. A novel invasion assay mimics EMT. (a) Schematic illustration of the 'plug' invasion assays. A collagen- or Matrigel-based epithelial plug is embedded in additional collagen, with or without fibroblasts. Epithelial cells then invade in a star-burst manner into surrounding collagen. (b) Morphological changes suggestive of spontaneous transitions between MET and EMT states are observed by light microscopy. Cell-collagen plugs were made with HBL100, HS578T and SUM159PT cell lines. These exhibit a predominantly elongated morphology at day 0. Clear invasion into surrounding collagen is seen by day 7 (*left panels*, arrows). Cell-Matrigel plugs with the same lines exhibit a rounded morphology on day 0. By day 7, HBL100 and HS578T cells have reverted to an elongated morphology and are invading into surrounding collagen. In contrast, SUM159PT cells retain a rounded morphology accompanied by delayed invasion (day 7) although this appears to be overcome by day 14 (*right panels*). Dotted lines represent the original plug edge. Bar = 100 μm.
doi:10.1371/journal.pone.0017083.g004

contrast to the movement of cells in a vertical plane that occurs with the collagen-based 'on top' assay [17].

Three different cell lines with low expression of the 9-gene EMT signature (HBL100, HS578T and SUM 159PT) demonstrated clear and reproducible invasion in this novel assay. Importantly, all three cell lines adopt a round morphology when embedded in Matrigel (day 0), versus the predominantly elongated morphology that is seen in collagen (Fig. S3). By day 7, the HBL100 and HS578T cells have reverted to an elongated morphology, indistinguishable from that seen in collagen, and are invading across BM and into surrounding collagen. In contrast, many SUM159PT cells retain a round morphology, accompanied by delayed invasion (Fig. 4b). By day 14, SUM159PT cells appear to have overcome this inhibition and many elongated cells are now seen leaving the Matrigel plug. Those cells that remain in the Matrigel plug still retain a more round morphology (Fig. 5a). MCF10A cells (a non-transformed line) were also tested in this assay and do exhibit an invasion phenotype. As expected, MCF10A cells appear to form polarised, growth arrested structures [32]. These observations suggest that this model may allow the investigation of cells as they invade across the BM. Importantly, SUM159PT cells are the most affected by Matrigel in terms of morphology and invasive capacity, and were therefore selected for further investigation.

SUM159PT cells, which are an excellent metastasis model *in vivo* [33,34], were selected for further EMT analysis with MCF10A cells serving as a control, as they show uniform, membranous expression of E-cadherin and no expression of N-cadherin. In contrast, SUM159PT cells show no membrane-specific E-cadherin expression but do show membranous N-cadherin expression throughout the core of the plug (Fig. 5b). In the elongated invading cells at the periphery N-cadherin expression appears to be down-regulated.

Stromal fibroblasts have been shown to play critical roles in some models of invasion, remodelling the ECM and generating tracks along which epithelial cells can follow [35]. The role of normal and cancer-associated fibroblasts (CAFs) was therefore also investigated. No difference in invasion was evident with both normal fibroblasts and CAFs. This lack of effect on invasion was seen when epithelial cells were embedded in both collagen and Matrigel (Fig. S4).

Discussion

This study identifies 9 key genes shared by breast cells undergoing EMT *in vitro* and EMT enriched claudin-low tumors.

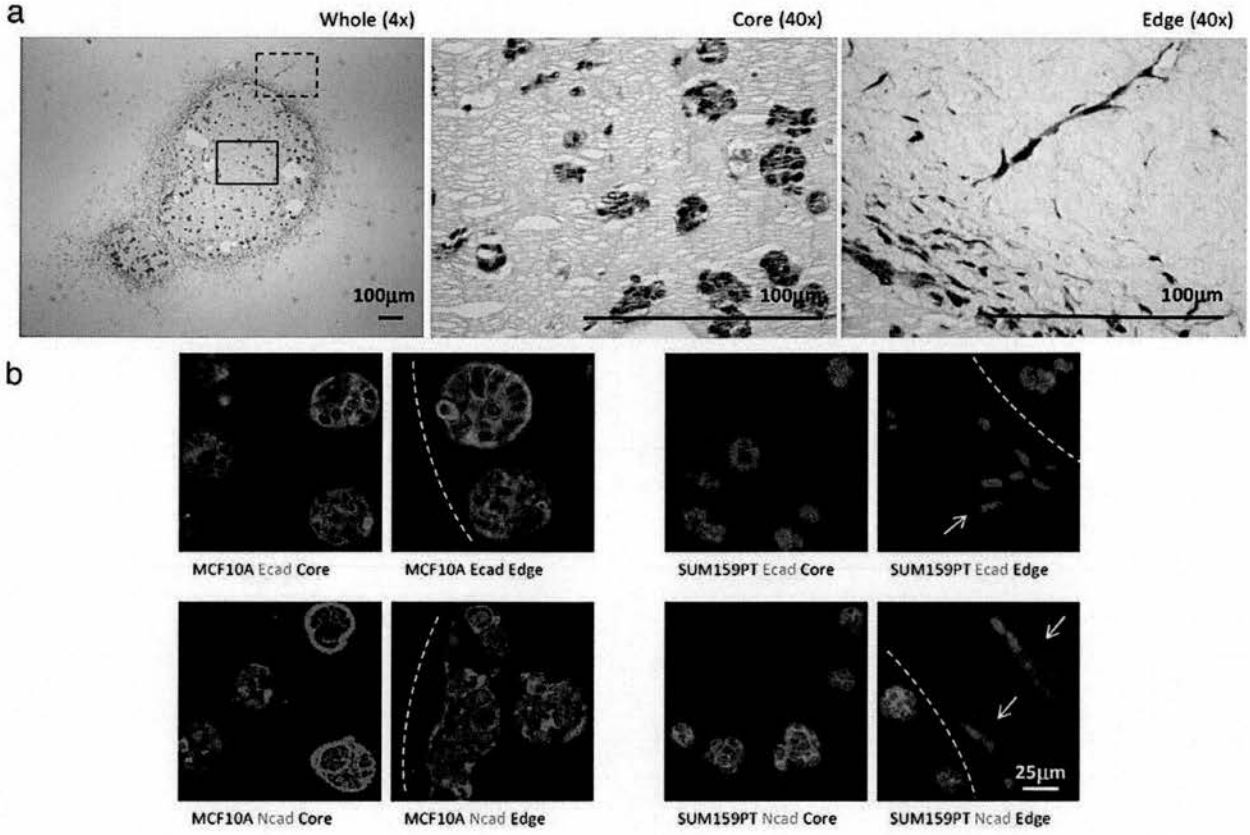


Figure 5. Changes in cells undergoing EMT while invading collagen stroma *in vitro*. (a) SUM159PT cell-Matrigel plugs were fixed at day 14 to monitor morphological changes during collagen invasion. Images of the whole plugs (4× magnification, *left panel*), core (*middle panel*) and plug edge (*right panel*) are shown (both 40× magnification). Note the organised, rounded morphology in Matrigel (*middle panel*) in contrast to the elongated morphology as cells invade into surrounding collagen (*right panel*), indicative of EMT. (b) E-cadherin expression in MCF10A cells is comparable to N-cadherin expression in SUM159PT cells. Representative immunofluorescence images of E-cadherin and N-cadherin protein expression in MCF10A (*left panel*) and SUM159PT cells (*right panel*) are shown. Expression within the plug (core) and at the edge where cells are seen to invade surrounding collagen (arrows) is compared. Note the change in morphology as cells invade. Bar = 50 μm.
doi:10.1371/journal.pone.0017083.g005

This signature in turn was used to identify breast cancer cell lines that are potentially useful in studying EMT *in vitro*. A 3D invasion model was developed that specifically addresses the link between EMT and invasion into stromal collagen in these cell lines, which may be representative of a general behaviour. This novel model was used to examine the expression patterns of cadherins in the EMT cell lines when invading from the basement membrane context to a collagen-rich environment.

The association of claudin-low breast cancer and epithelial to mesenchymal transition is now well established [14,27,36] and cell lines can be identified with gene expression profiles similar to those of claudin-low tumors [27]. The low expression signature is also found in these Basal B/mesenchymal/claudin-low cell lines, identified elsewhere [27]. Importantly, our results do not single out a particular EMT inducing transcriptional repressor, although these are broadly expressed (ZEB2, Snail and Slug) in the cell lines. This suggests that the induction of EMT may result from a combination of factors, resulting in repression of common downstream molecules. From a functional point of view, this is consistent with loss of cell-cell contact as a prerequisite for the detachment of invading cells from the tumor mass and their penetration of surrounding stroma [37].

Previously published invasion models have used either pure collagen environment [38] or non-physiological methylcellulose [39]. More physiologically relevant basement membrane-containing models, such as the chick chorioallantoic membrane [40] or peritoneal basement membrane [41], are inflexible, difficult to scale up and often have a very low yield. Our *in vitro* invasion model potentially offers a deeper investigation of the nature of EMT. The combination of basement membrane environment and surrounding collagen stroma maintains and mimics aspects of EMT *in vivo*.

The 3D model demonstrated here exemplifies how using the same cell line simultaneously in both basement membrane environment and in tissue-like collagen matrix may enable a better understating of EMT. Two novel observations were made using this model: within the basement membrane plug, N-cadherin expression in cells with EMT signature can phenocopy E-cadherin expression in normal mammary epithelial cells, maintaining a tight round morphology; and surprisingly, N-cadherin is lost as cells with EMT signature invade.

Claudin-low breast cancers are likely to represent the most acute EMT phenotype *in vivo*, but other subtypes may also present some EMT features [15,42]. The current study has extended our

understanding of common mechanisms of EMT in breast cancer. This study showed that the down-regulation of cell-cell contact molecules in claudin-low cancers is accompanied by changes in HGF signalling and apical sorting molecules. Furthermore, the 3D model has questioned the concept of a 'cadherin switch' *in vivo*. We have also observed elsewhere that in invasive ductal breast carcinomas there is no inverse correlation between E-cadherin and N-cadherin protein expression levels (S. Dubois-Marshall and E. Katz, unpublished observations). This raises the possibility that single cell invasion is cadherin-independent. This will to be verified in future experiments examining other cadherin molecules involved in cell motility, such as cadherin-11 [13]. Taken together, the 3D model presented here gives an opportunity to explore these possibilities relating to EMT as it may occur *in vivo* in claudin-low breast cancers and beyond.

Supporting Information

Figure S1 Comparison of genes correlating with C35 expression and those identifying the claudin-low phenotype. Full details of C35 and claudin-low signatures shown in Fig. 1.
(TIF)

Figure S2 Claudin-low cell lines exhibit a mesenchymal morphology. Eight claudin-low cell lines were identified. Representative live microscopy images of these lines cultured on plastic are shown. The non-transformed cell line, MCF10A, is shown for comparison. Bar = 100 μm.
(TIF)

Figure S3 Morphology of cell-collagen assays. SUM159PT cell-collagen plugs were fixed at day 14 following a period of invasion Images of the whole plugs (4× magnification, *left panel*), core (*middle panel*) and plug edge (*right panel*) are shown (both 40× magnification). Note the consistently elongated cell morphology unlike cell-Matrigel assays (Figure 5a).
(TIF)

Figure S4 Comparable invasion of SUM159PT cells regardless of the presence or type of fibroblasts in the surrounding collagen. Comparable invasion of SUM159PT cells is seen with no, normal and cancer-associated fibroblasts. This is seen with both cell-collagen (*top panel*) and cell-Matrigel (*bottom panel*) plugs. H&E staining relating to fixation at day 6 is shown here. Bar = 100 μm.
(TIF)

Acknowledgments

The authors would like to thank Alexey Larionov and Danielle Wilson for technical support.

Author Contributions

Conceived and designed the experiments: EK SDM DJH. Performed the experiments: EK SDM HC. Analyzed the data: EK SDM AHS PG RRM. Contributed reagents/materials/analysis tools: AHS PG. Wrote the paper: EK RRM DJH.

References

1. Fernandez Y, Cueva J, Palomo AG, Ramos M, de Juan A, et al. (2010) Novel therapeutic approaches to the treatment of metastatic breast cancer. Cancer Treat Rev 36: 33–42.
2. Arendt LM, Rudnick JA, Keller PJ, Kuperwasser C (2010) Stroma in breast development and disease. Semin Cell Dev Biol 21: 11–18.
3. Weigelt B, Peterse JL, van 't Veer LJ (2005) Breast cancer metastasis: markers and models. Nat Rev Cancer 5: 591–602.
4. Thiery JP (2003) Epithelial-mesenchymal transitions in development and pathologies. Curr Opin Cell Biol 15: 740–746.
5. Thiery JP, Sleeman JP (2006) Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol 7: 131–142.
6. Polyak K, Weinberg RA (2009) Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer 9: 265–273.
7. Huber MA, Kraut N, Beug H (2005) Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol 17: 548–558.
8. Gupta PB, Onder TT, Jiang G, Tao K, Kuperwasser C, et al. (2009) Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 138: 645–659.
9. Schmalhofer O, Brabletz S, Brabletz T (2009) E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. Cancer Metastasis Rev 28: 151–166.
10. Hazan RB, Qiao R, Keren R, Badano I, Suyama K (2004) Cadherin switch in tumor progression. Ann N Y Acad Sci 1014: 155–163.
11. Peinado H, Olmeda D, Cano A (2007) Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 7: 415–428.
12. Hazan RB, Phillips GR, Qiao RF, Norton L, Aaronson SA (2000) Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis. J Cell Biol 148: 779–790.
13. Nieman MT, Prudoff RS, Johnson KR, Wheelock MJ (1999) N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression. J Cell Biol 147: 631–644.
14. Herschkowitz JI, Simin K, Weigman VJ, Mikaelian I, Usary J, et al. (2007) Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol 8: R76.
15. Tomaskovic-Crook E, Thompson EW, Thiery JP (2009) Epithelial to mesenchymal transition and breast cancer. Breast Cancer Res 11: 213.
16. Evans EE, Henn AD, Jonason A, Paris MJ, Schiffhauer LM, et al. (2006) C35 (C17orf37) is a novel tumor biomarker abundantly expressed in breast cancer. Mol Cancer Ther 5: 2919–2930.
17. Katz E, Dubois-Marshall S, Sims AH, Faratian D, Li J, et al. (2010) A gene on the HER2 amplicon, C35, is an oncogene in breast cancer whose actions are prevented by inhibition of Syk. Br J Cancer 103: 401–410.
18. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, et al. (2004) Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 5: R80.
19. Dunning MJ, Smith ML, Ritchie ME, Tavare S (2007) beadarray: R classes and methods for Illumina bead-based data. Bioinformatics 23: 2183–2184.
20. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 98: 5116–5121.
21. van Helden J, Andre B, Collado-Vides J (1998) Extracting regulatory sites from the upstream region of yeast genes by computational analysis of oligonucleotide frequencies. J Mol Biol 281: 827–842.
22. Thomas-Chollier M, Sand O, Turatsinze JV, Janky R, Defrance M, et al. (2008) RSAT: regulatory sequence analysis tools. Nucleic Acids Res 36: W119–127.
23. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, et al. (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 34: D108–110.
24. Neve RM, Chin K, Fridlyand J, Yeh J, Baehner FL, et al. (2006) A collection of breast cancer cell lines for the study of functionally distinct cancer subtypes. Cancer Cell 10: 515–527.
25. Bailey CM, Khalkhali-Ellis Z, Kondo S, Margaryan NV, Seftor RE, et al. (2005) Mammary Serine Protease Inhibitor (Maspin) Binds Directly to Interferon Regulatory Factor 6: Identification of a Novel Serpin Partnership. J Biol Chem 280: 34210–34217.
26. Zhou C, Nitschke AM, Xiong W, Zhang Q, Tang Y, et al. (2008) Proteomic analysis of tumor necrosis factor-alpha resistant human breast cancer cells reveals a MEK5/Erk5-mediated epithelial-mesenchymal transition phenotype. Breast Cancer Res 10: R105.
27. Prat A, Parker JS, Karginova O, Fan C, Livasy C, et al. (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 12: R68.
28. Parr C, Jiang WG (2006) Hepatocyte growth factor activation inhibitors (HAI-1 and HAI-2) regulate HGF-induced invasion of human breast cancer cells. Int J Cancer 119: 1176–1183.
29. Fanayan S, Shehata M, Agterof AP, McGuckin MA, Alonso MA, et al. (2009) Mucin 1 (MUC1) is a novel partner for MAL2 in breast carcinoma cells. BMC Cell Biol 10: 7.
30. Steed E, Rodrigues NT, Balda MS, Matter K (2009) Identification of MarvelD3 as a tight junction-associated transmembrane protein of the occludin family. BMC Cell Biol 10: 95.

31. Ellison G, Klinowska T, Westwood RF, Docter E, French T, et al. (2002) Further evidence to support the melanocytic origin of MDA-MB-435. Mol Pathol 55: 294–299.
32. Debnath J, Brugge JS (2005) Modelling glandular epithelial cancers in three-dimensional cultures. Nat Rev Cancer 5: 675–688.
33. Kuperwasser C, Dessain S, Bierbaum BE, Garnet D, Sperandio K, et al. (2005) A mouse model of human breast cancer metastasis to human bone. Cancer Res 65: 6130–6138.
34. Flanagan L, Van Weelden K, Ammerman C, Ethier SP, Welsh J (2000) SUM-159PT cells: a novel estrogen independent human breast cancer model system. Breast Cancer Res Treat 58: 193–204.
35. Gaggioli C, Hooper S, Hidalgo-Carcedo C, Grosse R, Marshall JF, et al. (2007) Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. Nat Cell Biol 9: 1392–1400.
36. Taube JH, Herschkowitz JI, Komurov K, Zhou AY, Gupta S, et al. (2010) 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 107: 15449–15454.
37. Yilmaz M, Christofori G (2010) Mechanisms of motility in metastasizing cells. Mol Cancer Res 8: 629–642.
38. Sabeh F, Shimizu-Hirota R, Weiss SJ (2009) Protease-dependent versus - independent cancer cell invasion programs: three-dimensional amoeboid movement revisited. J Cell Biol 185: 11–19.
39. Wiercinska E, Naber HP, Pardali E, van der Pluijm G, van Dam H, et al. (2010) The TGF-beta/Smad pathway induces breast cancer cell invasion through the up-regulation of matrix metalloproteinase 2 and 9 in a spheroid invasion model system. Breast Cancer Res Treat [Epub ahead of print].
40. Yook JI, Li XY, Ota I, Hu C, Kim HS, et al. (2006) A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. Nat Cell Biol 8: 1398–1406.
41. Hotary K, Li XY, Allen E, Stevens SL, Weiss SJ (2006) A cancer cell metalloprotease triad regulates the basement membrane transmigration program. Genes Dev 20: 2673–2686.
42. Klymkowsky MW, Savagner P (2009) Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. Am J Pathol 174: 1588–1593.

Clin Exp Metastasis
DOI 10.1007/s10585-011-9412-x

RESEARCH PAPER

Two possible mechanisms of epithelial to mesenchymal transition in invasive ductal breast cancer

Sylvie Dubois-Marshall · Jeremy S. Thomas · Dana Faratian · David J. Harrison · Elad Katz

Received: 10 March 2011 / Accepted: 12 July 2011

Abstract Epithelial to mesenchymal transition (EMT) occurs in embryogenesis and normal development. It has been predominantly described in vitro and in animal studies, but EMT is also implicated in the progression of many cancers with proposed roles in invasion, metastasis and resistance to treatment. It is closely associated with loss of epithelial-specific protein expression and up-regulation of mesenchymal proteins, but several pathways are implicated in its execution. We explored what are the expression patterns of EMT proteins in human breast cancer. We interrogated two independent cohorts enriched for high-grade, invasive, ductal breast cancers. We used quantitative immunofluorescence to study the expression of key EMT proteins. Statistical associations to define protein profiles were based on Pearson's correlations. E-cadherin down-regulation in breast cancer was associated with β-catenin down-regulation, but not with up-regulation of mesenchymal markers. While EMT-related transcription repressors were expressed in some breast cancers, their expression did not negatively correlate with E-cadherin. Instead, an additional EMT profile was identified, composing Snail and Slug. In conclusion, EMT occurs in human breast cancer in a manner distinct to that seen in vitro. Certain EMT events are uncoupled from E-cadherin down-regulation and may constitute a novel EMT profile, which warrants further exploration.

Keywords Epithelial-mesenchymal transition · E-cadherin · Breast cancer · Neoplasm invasion · Beta-catenin

Introduction

The trans-differentiation of cells from an epithelial to a mesenchymal phenotype or epithelial to mesenchymal transition (EMT), occurs in embryogenesis and normal development [1]. Increasing evidence supports a role for EMT in the progression of many cancer types including breast, with critical roles in invasion, metastatic dissemination and acquisition of therapeutic resistance [1, 2]. EMT involves loss of cell–cell junctions and re-organisation of the actin cytoskeleton, resulting in a spindle-like mesenchymal morphology [3]. EMT is associated with decreased expression of epithelial-specific proteins [4]. Down-regulation of E-cadherin expression is possibly the most important consequence of EMT that leads to the changed behaviour of tumour cells [5, 6]. An important event in EMT is the switching of expression from E-cadherin, which becomes down regulated, to N-cadherin, which is up-regulated [7]. Other mesenchymal proteins, such as Vimentin, are also up-regulated during EMT [8, 9].

E-cadherin down-regulation is usually caused by transcriptional repression [5]. Several zinc finger containing repressors have been indentified, including Snail, Slug (Snail2) and Twist [10]. These transcriptional repressors interact with E-boxes within the *CDH1* gene promoter. Increased expression of Snail, Slug, and Twist has been

Electronic supplementary material The online version of this article (doi:10.1007/s10585-011-9412-x) contains supplementary material, which is available to authorized users.

S. Dubois-Marshall · J. S. Thomas · D. Faratian · D. J. Harrison · E. Katz (✉)
Breakthrough Breast Cancer Research Unit and Division of Pathology, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK
e-mail: elad.katz@ed.ac.uk

Published online: 26 July 2011

associated with poor prognostic clinico-pathological features and outcome in breast and other cancers [11, 12].

We recently compared claudin-low cell lines, which are enriched for EMT genetic events, to normal human mammary epithelial cells. Claudin-low cancer cell lines display uniform down-regulation of junction proteins (E-cadherin and claudin-7), but more variable up-regulation of Snail, N-cadherin and Vimentin [4].

Recent evidence indicates that there are reciprocal interactions between E-cadherin, β-catenin and EMT-inducing transcriptional repressors [5]. Wnt signalling, through a nuclear β-catenin-TCF complex, is capable of driving EMT by up-regulating Snail in human breast cancer cell lines [13]. In addition, loss of E-cadherin-β-catenin junctions at the cell membrane can independently drive EMT by allowing freed β-catenin to relocalise to the nucleus [14, 15].

The role of EMT in human cancer has predominantly been inferred from in vitro studies using cell type-specific markers [16–18]. Its importance in human cancer remains controversial and data are sometimes conflicting [19, 20]. While EMT in human breast cancer shares common traits with in vitro models, it may not display all the hallmarks predicted by them. We have identified recently a genetic EMT signature common to breast cancer in vitro models and in vivo [4]. This signature includes the down-regulation of E-cadherin alongside other cell–cell contact molecules, but it does not include increased expression of mesenchymal proteins or EMT repressors.

In order to gain insight into EMT in human breast cancer at the protein level, we selected two patient cohorts that maximised the likelihood of identifying EMT-events. We examined invasive ductal carcinomas of no special type (IDC-NST). We found that E-cadherin protein expression was associated closely with β-catenin expression level. However, E-cadherin down-regulation in breast cancer was not associated with the up-regulation of other EMT markers, such as N-cadherin. A distinct EMT profile containing Snail and Slug was identified in both patient cohorts.

Materials and methods

Patient cohort selection, exclusion of special type cancers and tissue microarray construction

We quantitatively determined expression of EMT-related proteins in the common type of human breast cancer, invasive ductal carcinoma/no special type breast cancers (IDC-NST). Down-regulation of E-cadherin is more common in invasive lobular cancer (ILC) than in IDC-NST [21]. ILC and IDC-NST are molecularly distinct [22] and therefore down-regulation of E-cadherin is likely to occur through distinct mechanisms in each type.

The first patient cohort (set 1) was enriched for HER2+ breast cancers [18]. The HER2+ sub-group usually contain a minimal number of invasive lobular cancers (ILCs) [23]. This study population consisted of 122 patients with primary breast cancer who, subsequently to the analysis specimens, were treated with trastuzumab in the Edinburgh Breast Unit, as previously described [18].

The second patient cohort (set 2) was enriched for large, high-grade breast cancers that had already given rise to lymph node metastases. The presence of these poor prognostic features maximised the chance of identifying EMT events [24]. This study population was derived from an original population of 521 patients with primary breast carcinomas treated in the Edinburgh Breast Unit from 1999 to 2002 [25, 26]. Patients with paired lymph node metastases were selected (143 patients) for tissue microarray (TMA) construction.

Construction of these sets and acquisition of linked clinical data was approved by the Lothian Research Ethics Committee (08/S1101/41): all cases were anonymised once data linkage had occurred. Tumour areas were selected for TMA construction on H&E slides and 0.6 mm^2 cores were placed into three separate TMA replicates for each sample, as previously described [27].

The characteristics of the patient cohorts used to construct the two TMA sets in this study are summarised in Tables 1 and 2. HER2 and ER scores were reported by us previously for all of these cases [18, 25]. The scores were given during the clinical process using Allred classification scoring.

Immunofluorescence

AQUA methodology has been described elsewhere [18]. Briefly, antigen retrieval for all epitopes was carried out using heat treatment under pressure in a microwave oven for 5 min in citrate buffer pH 6.0. Slides were incubated with primary antibodies for 1 h at room temperature. Primary antibodies used: E-cadherin (intracellular), BD, 610181, Mouse, 1:1500; N-cadherin, BD, 610921, Mouse, 1:300; Vimentin, Sigma, V6630, Mouse, 1;400; Fibronectin, Abcam, ab2413, Rabbit, 1:10,000; Slug, LifeSpan Bio, LS-C30318, Rabbit, 1:1000; Snail, Abcam, ab17732, Rabbit, 1:700; β-catenin, BD, 610153, Mouse, 1:500. The specificity of all antibodies was verified previously by Western blotting [4].

Rabbit primary antibodies were incubated overnight with mouse anti-pancytokeratin (Invitrogen, 1:25) to visualise epithelial cells. Mouse primary antibodies were incubated overnight with rabbit anti-pancytokeratin (Dako, 1:150) and rabbit anti-pancadherin (Cell Signalling, 1:50). The epithelial compartment was then visualised with Cy3 (Invitrogen, 1:25). DAPI (4′,6-diamidino-2-phenylindole)

Table 1 Patient characteristics of TMA set 1

	Patients, *n* (%) Original data set ($n = 122$)	Patients, *n* (%) After exclusions ($n = 106$)
Tumour grade		
1	1 (0.8)	1 (0.9)
2	18 (14.8)	11 (10.4)
3	99 (81.1)	93 (87.7)
NK	4 (3.3)	1 (0.9)
Prognostic index		
>5.4	59 (48.4)	53 (50.0)
3.4–5.4	46 (37.7)	40 (37.7)
<3.4	2 (1.6)	1 (0.9)
NK	15 (12.3)	12 (11.3)
Node stage		
Negative	26 (21.3)	24 (22.6)
Positive	85 (69.7)	75 (70.8)
NK	11 (9.0)	7 (6.6)
HER2 status		
Positive	82 (67.2)	73 (68.9)
Negative	32 (26.2)	30 (28.3)
NK	8 (6.6)	3 (2.8)
ER status		
≥3	71 (58.2)	60 (56.6)
<3	39 (32.0)	36 (34.0)
NK	12 (9.8)	10 (9.4)

Nine cases were excluded due to ≤50% available data. A further 7 cases were excluded (ILC 5, no/insufficient tumour 1, no E-cadherin data available 1), leaving 106 cases for statistical analysis

Table 2 Patient characteristics of TMA set 2

	Patients, *n* (%) Original data set ($n = 143$)	Patients, *n* (%) After exclusions ($n = 111$)
Tumour grade		
1	4 (2.8)	3 (2.7)
2	72 (50.3)	49 (44.1)
3	67 (46.9)	59 (53.2)
NK	0 (0)	0 (0)
Tumour size (mm)		
≤20 (T1)	49 (34.3)	44 (39.6)
21–50 (T2)	84 (58.7)	62 (55.9)
>50 (T3)	9 (6.3)	4 (3.6)
NK	1 (0.7)	1 (0.9)
Nodes		
Negative	1 (0.7)	1 (0.9)
Positive	142 (99.3)	110 (99.1)
NK	0 (0)	0 (0)
HER2 status		
Positive	18 (12.6)	18 (16.2)
Negative	109 (76.2)	83 (74.8)
NK	16 (11.2)	10 (9.0)
ER status		
≥3	90 (62.9)	69 (62.2)
<3	53 (37.1)	42 (37.8)
NK	0 (0)	0 (0)

Seventeen cases were excluded due to ≤50% available data. A further 15 cases were excluded (mixed NST/ILC 3, ILC 9, other special types 3), leaving 111 cases for statistical analysis

counterstain (Invitrogen) was used to identify nuclei and Cy-5-tyramide (HistoRx, 1:50) was used to detect protein 'targets'.

AQUA automated image analysis

Monochromatic images of each TMA core were captured at 20× magnification using an Olympus AX-51 epifluorescence microscope. These high-resolution, digital images were individually, manually quality checked to exclude aberrant imaging artefacts from analysis. All areas of normal mammary duct and ductal carcinoma in situ (DCIS) were excluded to ensure that only invasive cancer was included in the analysis. Images were then analysed using AQUAnalysis software, as previously described [18].

Statistical analysis methods

Cases with ≤50% of all available AQUA scores were excluded from further analysis. Cases lacking in TMA cores stained and scored for E-cadherin were also excluded. AQUA scores for each array were $\log_2$ transformed and mean centred. The mean of the three replicates for each protein target were used for further analysis. For each replicate, zero is the mean after transformation. After averaging across the three replicates, zero is not necessarily the mean of the entire dataset. Associations between variables were calculated using Pearson's correlation coefficients and differences in means with one-way ANOVA using SPSS software (v14). The resulting *P*-values were adjusted for multiple-testing using the conservative Bonferroni correction (uncorrected $P < 0.0014$).

Results

Human breast cancer has a wide range of E-cadherin expression, highly correlated with β-catenin

Two patient cohorts containing large numbers of high-grade invasive breast cancers [18, 25] were used to gain insight into possible EMT events and establish whether they are present in independent datasets. A continuous

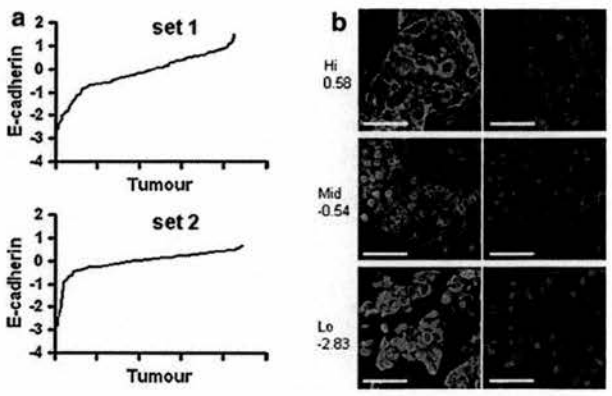


Fig. 1 Human breast cancer has a wide range of E-cadherin protein expression. **a** Comparable, dynamic ranges of E-cadherin expression are seen in both TMA sets. AQUA scores for each array were $\log_2$ transformed and mean centred (0 represents the adjusted mean for each individual TMA slide). The mean of three TMA replicates is shown. **b** Representative immunofluorescence images (from set 2) illustrating varying E-cadherin protein expression (*left panels*, with pan-cytokeratin mask; *right panels*, without pan-cytokeratin). Note the membranous staining of E-cadherin. Cytokeratin *green*; DAPI - *blue*; E-cadherin *red*. Normalised AQUA scores are given (arbitrary units). *Bars* 50 μm in all images

range of membrane/cytoplasmic E-cadherin expression was seen in both sets (Fig. 1a). In set 1, an 18-fold range of expression of E-cadherin was seen. A corresponding 11-fold range was seen in set 2. No correlation between E-cadherin expression and estrogen (ER) or HER2 receptor status was observed (data not shown). Qualitative assessment of immunofluorescence images showed that E-cadherin staining remains membranous across all breast tumours (Fig. 1b), similarly to its distribution in the normal breast (Supp. Fig 1a). In all cases, diffused E-cadherin staining could be detected also in the nucleus, although its functional importance is unknown (not shown).

Tumours with high cytoplasmic β-catenin expression were identified (Fig. 2a). Nuclear β-catenin expression, although less distinct, was also observed and positively correlated with cytoplasmic expression (Table 3; Supplementary Tables 1–2 contain all correlations in this study, corrected for multiple testing). A positive correlation between E-cadherin and β-catenin protein expression was seen in both patient sets (Fig. 2b and Table 3). This established that in human breast cancer there is an EMT-related profile in which concomitant down-regulation of E-cadherin and β-catenin occurs.

High expression of N-cadherin and other mesenchymal markers in human breast cancers

Having identified tumours with relatively low E-cadherin expression, we examined whether these tumours show evidence of a 'cadherin switch'. Tumours with high cytoplasmic N-cadherin expression (Fig. 3a) were identified. Nonetheless, no correlation between E-cadherin and N-cadherin expression levels was found in either set (Supplementary Tables 1–2). The lack of a consistent cadherin switch in human breast cancers is in agreement with our findings in cell line 3-dimensional cultures [4].

Vimentin and fibronectin are mesenchymal markers that are up-regulated in EMT in the context of breast and other cancers [1]. As with N-cadherin, tumours with relatively high cytoplasmic Vimentin and fibronectin expression were identified (Fig. 3b–c), but no correlation with E-cadherin expression observed (Supplementary Tables 1–2). Immunofluorescence images show stromal and epithelial staining

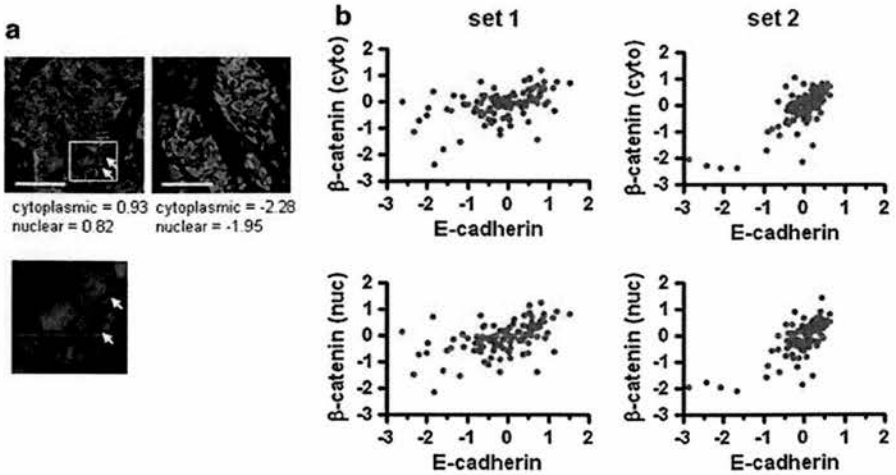


Fig. 2 E-cadherin and β-catenin protein expression are highly correlated in human breast cancer. **a** Representative immunofluorescence images (from set 2) illustrating tumours with high (*left panel*) and low (*right panel*) expression of β-catenin. Inset shows nuclear staining in individual cells (*arrows*). Cytokeratin *green*; DAPI *blue*; β-catenin *red*. **b** Plots comparing membrane/cytoplasmic and nuclear β-catenin expression levels to E-cadherin in both TMA sets (all $P < 0.01$ by Pearson's correlation, see Table 3)

Table 3 Significant correlations between protein markers of EMT

Protein marker	*r* value	Protein marker
E-cadherin	0.476 (set 1)	β-catenin (cytoplasmic)
	0.749 (set 2)	
E-cadherin	0.501	β-catenin (nuclear)
	0.724	
β-catenin (cytoplasmic)	0.944	β-catenin (nuclear)
	0.947	
Snail	0.526	Slug
	0.388	

All reproducible correlations (significant after Bonferroni correction) observed between protein markers of EMT

for these markers as expected. Epithelial fibronectin expression was focal in high expressing tumours (Fig. 3c). No correlation was observed between N-cadherin, Vimentin and fibronectin expression in these tumours (Supplementary Tables 1–2). These findings suggest that up-regulation of mesenchymal markers is unlikely to be a common event and that tumours that up-regulate one protein marker do not necessarily up-regulate other markers.

High expression of Snail and Slug in human breast cancers

Quantitative analysis of transcription repressors was performed to determine their correlation with E-cadherin expression and other EMT markers. Tumours with high nuclear Snail and Slug expression were identified (Fig. 4). Snail expression is predominantly nuclear, whilst Slug expression is mostly cytoplasmic (Fig. 4a–b). Cytoplasmic expression of EMT transcription repressors, including Slug, has been reported anecdotally in previous studies, although the functional importance of this is unknown [28]. This expression pattern of Slug is not seen in ovarian cancers (Supplementary Fig 1).

No correlation of E-cadherin expression with either Slug or Snail was observed (Supplementary Tables 1–2). In addition, protein expression of Slug and Snail did not differ significantly between ER− and ER+ tumours (not shown), as suggested using immunohistochemical methods [29].

Correlations between markers of EMT identify two distinct profiles in human breast cancer

Protein markers of EMT with reproducible, significant correlations across both patient sets are shown in Table 3. Univariate analysis showed no significant correlations between the markers of EMT investigated in this study and patient survival (not shown). Down-regulation of E-cadherin significantly correlates with down-regulation of β-catenin (Fig. 2b). In a second EMT event profile, the expression of Snail correlates with Slug (Fig. 5). These observations were reproducible in both sets (Supplementary Tables 1–2). Interestingly, in Set 1, Snail correlated

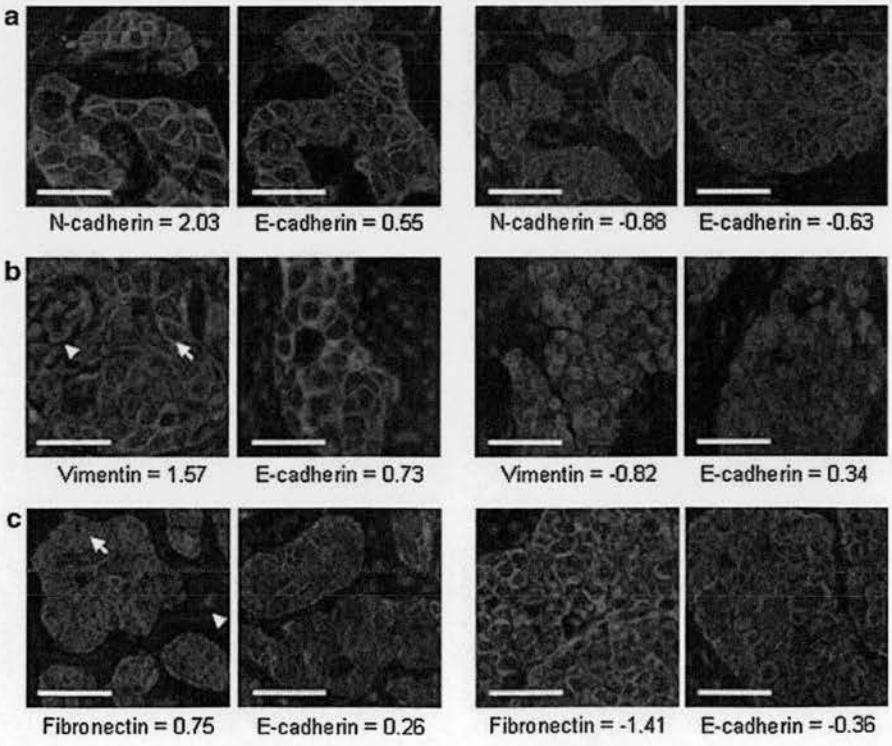


Fig. 3 High expression of mesenchymal proteins is found in human breast cancer. Representative immunofluorescence images (from set 1) illustrating tumours with high (*left panels*) and low (*right panels*) expression of **a** N-cadherin, **b** Vimentin and **c** fibronectin. Corresponding E-cadherin expression is shown in each case. Note the stromal (*short arrow*) as well as epithelial (long arrow) expression of Vimentin and fibronectin. Cytokeratin *green*; DAPI *blue*; E-cadherin/N-cadherin/Vimentin/fibronectin *red*

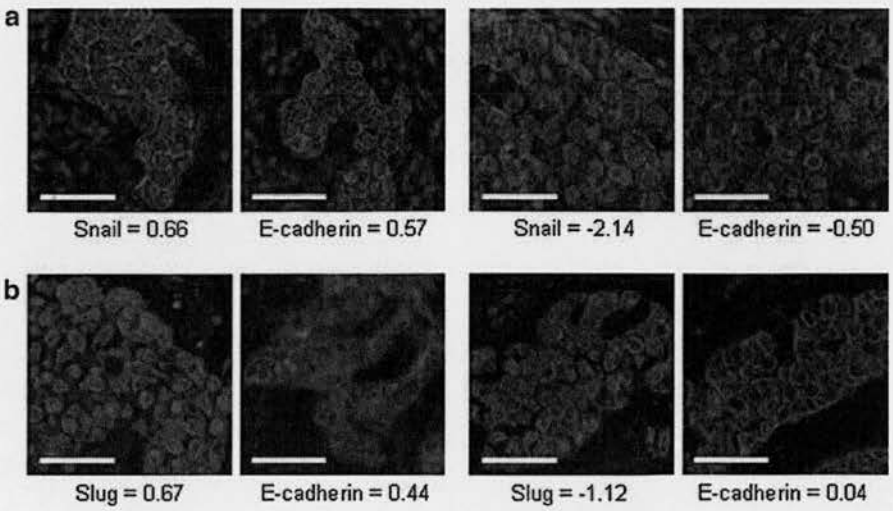


Fig. 4 High expression of Snail and Slug in human breast cancer. Representative immunofluorescence images (from set 1) illustrating tumours with high (*left panels*) and low (*right panels*) expression of **a** Snail and **b** Slug. Corresponding E-cadherin expression is shown in each case. Note the clear nuclear expression of Snail whilst Slug is predominantly cytoplasmic. Cytokeratin *green*; DAPI *blue*; Snail/Slug *red*

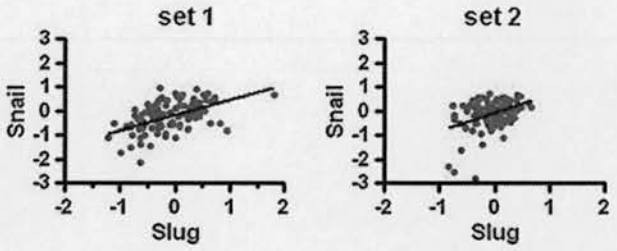


Fig. 5 Correlations between protein markers suggest a second EMT profile in human breast cancers. Plots illustrating the correlations between Snail and Slug as identified in Table 3

with β-catenin (cytoplasmic and nuclear) expression. Slug correlated with N-cadherin and Vimentin protein expression (Supplementary Table 1). This may reflect unique roles for these EMT transcription factors in HER2+ breast cancers.

Discussion

EMT is believed to be a mechanism that allows cells to acquire an invasive phenotype [2]. There is accumulating evidence that EMT may not be a single, conserved programme [19]. Partial or incomplete EMT phenotypes, where advanced carcinomas display some mesenchymal features whilst retaining characteristics of well-differentiated epithelial cells, have been described in several studies [30]. A recent systems analysis has suggested that in partial EMT in vitro, loss of epithelial components is more uniform than the gain of mesenchymal components [31]. Until now, EMT features were found mainly in ER- subtypes of breast cancers and in some mouse models [9]. Basal-like cancers contain higher levels of N-cadherin and Vimentin, but only some down-regulate E-cadherin. Notably, β-catenin expression is more prevalent in this sub-type [32]. Two other sub-types, namely claudin-low and metaplastic breast cancers display strong down-regulation of E-cadherin and other junction proteins, but more variable up-regulation of markers such as EMT transcription factors [4, 8, 33]. The current study is a more global analysis of EMT proteins in invasive human breast cancers, building on previous studies showing that cytokeratin-expressing breast epithelial cells may co-express mesenchymal proteins such as Vimentin [8, 9].

We examined EMT in human breast cancer by investigating the expression of key protein markers and the associations between them. We tested two hypotheses: (i) IDC-NSTs expressing low levels of E-cadherin undergo cadherin switch to N-cadherin and (ii) that this is driven by known transcriptional repressors of E-cadherin. We also investigated β-catenin nuclear translocation in these cancers as an alternative mode of EMT.

Two EMT profiles were identified in invasive breast cancers. Each of these profiles presents some of the EMT events as seen in vitro. Using quantitative immunofluorescence analysis, we identified a dynamic range of E-cadherin protein expression across two patient cohorts. Our findings suggest that co-expression of E-cadherin and β-catenin at cellular junctions is maintained in the majority of human breast cancers [34]. A significant, positive correlation between E-cadherin and β-catenin was seen, but without an apparent patient outcome association. Positive correlations between E-cadherin and β-catenin in human breast cancer with no correlation to outcome have been described using immunohistochemical methods [35]. These and our findings are in contrast to a different report in which a significant correlation between loss of β-catenin and poor clinical outcome [36]. The lack of cadherin switch in this profile is consistent with our previous observations in cell line models [4].

The correlation between Snail and Slug suggests that a transcriptionally driven EMT mechanism may also be in place in human breast cancer. This profile is uncoupled from down-regulation of E-cadherin, unlike what is found in cell line models [4]. This suggests that EMT does occur in vivo but in a manner distinct from that suggested by in vitro studies. A number of studies support our findings. In a series of human breast invasive ductal cancers with associated lymph node metastases, both Snail and Slug were significantly over-expressed at the mRNA level. Importantly, expression of Snail or Slug did not preclude significant E-cadherin expression [28]. Twist expression could not be reported here due to the poor specificity of commercially available antibodies. In cell lines, Twist expression induced up-regulation of both Snail [37] and Slug [38], as well as down-regulation of E-cadherin and β-catenin protein expression [38].

In conclusion, our observations of EMT events in breast cancer manifest in two distinct profiles. These validated profiles give first comprehensive insight into EMT in human breast cancer. This should enable more fitting applications of EMT knowledge gained in vitro in the clinical setting.

Acknowledgments The authors would like to thank InHwa Um, Danielle Wilson and Helen Caldwell for their technical support throughout this study. We would also like to thank Andy Sims and Richard Meehan for their advice. Clinical materials were obtained through the auspices of the Edinburgh Experimental Cancer Medicine Centre. This work was supported by Breakthrough Breast Cancer and the Scottish Funding Council.

Conflict of interest The authors declare no conflict of interest.

References

1. Thiery JP, Acloque H, Huang RY et al (2009) Epithelial-mesenchymal transitions in development and disease. Cell 139(5):871–890
2. Polyak K, Weinberg RA (2009) Transitions between epithelial and mesenchymal states: acquisition of malignant and stem cell traits. Nat Rev Cancer 9(4):265–273
3. Huber MA, Kraut N, Beug H (2005) Molecular requirements for epithelial-mesenchymal transition during tumor progression. Curr Opin Cell Biol 17(5):548–558
4. Katz E, Dubois-Marshall S, Sims AH et al (2011) An in vitro model that recapitulates the epithelial to mesenchymal transition (EMT) in human breast cancer. PLoS ONE 6(2):e17083
5. Schmalhofer O, Brabletz S, Brabletz T (2009) E-cadherin, beta-catenin, and ZEB1 in malignant progression of cancer. Cancer Metastasis Rev 28(1–2):151–166
6. Gupta PB, Onder TT, Jiang G et al (2009) Identification of selective inhibitors of cancer stem cells by high-throughput screening. Cell 138(4):645–659
7. Hazan RB, Qiao R, Keren R et al (2004) Cadherin switch in tumor progression. Ann N Y Acad Sci 1014:155–163
8. Prat A, Parker JS, Karginova O et al (2010) Phenotypic and molecular characterization of the claudin-low intrinsic subtype of breast cancer. Breast Cancer Res 12(5):R68
9. Damonte P, Gregg JP, Borowsky AD et al (2007) EMT tumorigenesis in the mouse mammary gland. Lab Invest 87(12): 1218–1226
10. Peinado H, Olmeda D, Cano A (2007) Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? Nat Rev Cancer 7(6):415–428
11. Elloul S, Elstrand MB, Nesland JM et al (2005) Snail, Slug, and Smad-interacting protein 1 as novel parameters of disease aggressiveness in metastatic ovarian and breast carcinoma. Cancer 103(8):1631–1643
12. Martin TA, Goyal A, Watkins G et al (2005) Expression of the transcription factors snail, slug, and twist and their clinical significance in human breast cancer. Ann Surg Oncol 12(6): 488–496
13. Yook JI, Li XY, Ota I et al (2006) A Wnt-Axin2-GSK3beta cascade regulates Snail1 activity in breast cancer cells. Nat Cell Biol 8(12):1398–1406
14. Ma L, Young J, Prabhala H et al (2010) miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. Nat Cell Biol 12(3):247–256
15. Onder TT, Gupta PB, Mani SA et al (2008) Loss of E-cadherin promotes metastasis via multiple downstream transcriptional pathways. Cancer Res 68(10):3645–3654
16. Mikaelian I, Blades N, Churchill GA et al (2004) Proteotypic classification of spontaneous and transgenic mammary neoplasms. Breast Cancer Res 6(6):R668–R679
17. Moody SE, Perez D, Pan TC et al (2005) The transcriptional repressor Snail promotes mammary tumor recurrence. Cancer Cell 8(3):197–209
18. Katz E, Dubois-Marshall S, Sims AH et al (2010) A gene on the HER2 amplicon, C35, is an oncogene in breast cancer whose actions are prevented by inhibition of Syk. Br J Cancer 103(3):401–410
19. Klymkowsky MW, Savagner P (2009) Epithelial-mesenchymal transition: a cancer researcher's conceptual friend and foe. Am J Pathol 174(5):1588–1593
20. Tarin D, Thompson EW, Newgreen DF (2005) The fallacy of epithelial mesenchymal transition in neoplasia. Cancer Res 65(14):5996–6000
21. Droufakou S, Deshmane V, Roylance R et al (2001) Multiple ways of silencing E-cadherin gene expression in lobular carcinoma of the breast. Int J Cancer 92(3):404–408
22. Weigelt B, Geyer FC, Natrajan R et al (2010) The molecular underpinning of lobular histological growth pattern: a genome-wide transcriptomic analysis of invasive lobular carcinomas and grade- and molecular subtype-matched invasive ductal carcinomas of no special type. J Pathol 220(1):45–57
23. Arpino G, Bardou VJ, Clark GM et al (2004) Infiltrating lobular carcinoma of the breast: tumor characteristics and clinical outcome. Breast Cancer Res 6(3):R149–R156
24. Megha T, Neri A, Malagnino V et al (2010) Traditional and new prognosticators in breast cancer: Nottingham index, Mib-1 and estrogen receptor signaling remain the best predictors of relapse and survival in a series of 289 cases. Cancer Biol Ther 9(4). doi: 10.4161/cbt.9.4.10659
25. Aitken SJ, Thomas JS, Langdon SP et al (2010) Quantitative analysis of changes in ER, PR and HER2 expression in primary breast cancer and paired nodal metastases. Ann Oncol 21(6): 1254–1261
26. Somner JE, Dixon JM, Thomas JS (2004) Node retrieval in axillary lymph node dissections: recommendations for minimum numbers to be confident about node negative status. J Clin Pathol 57(8):845–848

27. Kononen J, Bubendorf L, Kallioniemi A et al (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nat Med 4(7):844–847
28. Come C, Magnino F, Bibeau F et al (2006) Snail and slug play distinct roles during breast carcinoma progression. Clin Cancer Res 12(18):5395–5402
29. Ye Y, Xiao Y, Wang W et al (2010) ERalpha signaling through slug regulates E-cadherin and EMT. Oncogene 29(10): 1451–1462
30. Christiansen JJ, Rajasekaran AK (2006) Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. Cancer Res 66(17):8319–8326
31. Thomson S, Petti F, Sujka-Kwok I et al (2011) A systems view of epithelial-mesenchymal transition signaling states. Clin Exp Metastasis 28(2):137–155
32. Sarrio D, Rodriguez-Pinilla SM, Hardisson D et al (2008) Epithelial-mesenchymal transition in breast cancer relates to the basal-like phenotype. Cancer Res 68(4):989–997
33. Taube JH, Herschkowitz JI, Komurov K et al (2010) Core epithelial-to-mesenchymal transition interactome gene-expression signature is associated with claudin-low and metaplastic breast cancer subtypes. Proc Natl Acad Sci USA 107(35):15449–15454
34. Perez-Moreno M, Fuchs E (2006) Catenins: keeping cells from getting their signals crossed. Dev Cell 11(5):601–612
35. Logullo AF, Nonogaki S, Pasini FS et al (2010) Concomitant expression of epithelial-mesenchymal transition biomarkers in breast ductal carcinoma: association with progression. Oncol Rep 23(2):313–320
36. Dolled-Filhart M, McCabe A, Giltnane J et al (2006) Quantitative in situ analysis of beta-catenin expression in breast cancer shows decreased expression is associated with poor outcome. Cancer Res 66(10):5487–5494
37. Mironchik Y, Winnard PT Jr, Vesuna F et al (2005) Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer. Cancer Res 65(23): 10801–10809
38. Casas E, Kim J, Bendesky A et al (2011) Snail2 is an essential mediator of Twist1-induced epithelial mesenchymal transition and metastasis. Cancer Res 71(1):245–254