Epithelial stromal interaction in breast cancer

Eldo Thomas Verghese, MBBS, MRCP (UK)

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Medicine

The candidate confirms that the work submitted is his own and that appropriate credit has been given when reference has been made to the work of others

Submission for examination: October, 2013

ACKNOWLEDGEMENTS

I would like to acknowledge the patients whose tissues I have used in this study.

In addition I would like to thank the following people:

Supervisors: Associate Professor Thomas A Hughes, Professor Andrew M Hanby, Professor Mark A Hull

Collaborators: Professor Valerie Speirs, Ms Claire Nash, Dr Xiaomei Lu, Dr Alexandre Zougman, Dr Caroline Green, Dr Ruth Drury, Dr James L Thorne, Dr Nikki Ingram, Dr Mihaela Lorger, Dr Louise Colette, Prof David Westhead, Dr Helene H Thygesen, Mr Mike Shires and Ms Sarah Perry

Mentor: Professor Phil Quirke

I also want to thank the Pathological Society of Great Britain & Ireland and Breast Cancer Campaign for the pilot grants, National Institute of Health Research for the academic clinical fellowship and Medical Research Council for the clinical research fellowship.

And last but not the least, my family whose support and encouragement has been essential for me to complete this work.

ABSTRACT

In recent years there has been an increasing awareness of the role played by the microenvironment surrounding breast cancer epithelium in modulating the behaviour of breast tumours. A major component of this microenvironment in breast cancer is the fibroblasts. These are commonly referred to as carcinoma-associated fibroblasts (CAFs). CAFs are phenotypically different from their normal counterpart in normal breast tissue. MicroRNAs, a family of small non-coding RNAs that are key players in the posttranscriptional regulation of mRNAs, have been shown to play a role in controlling the behaviour of cancers. They have been extensively studied in breast cancer epithelial cells and have important roles in breast cancer initiation and progression.

My hypotheses were that miRNAs also play roles in controlling the behaviour of breast CAFs and in turn, impact on the behaviour of malignant breast epithelial cells.

I used a single patient sample and a co-culture model to examine the expression of miRNAs in breast CAFs and normal fibroblast (NFs). Six miRNAs were consistently down-regulated in CAFs as compared to NFs with a fold change > 10 in both the tissue co-culture model and in patient samples. Of these, miR-26b was significantly down-regulated in CAFs in a further 14 cases of microdissected matched NFs and CAFs from clinical FFPE (Formalin Fixed Paraffin Embedded) samples and consistently in a further 4 cases of matched primary NFs and CAFs in vitro.

Functional influences on fibroblasts were examined by overexpressing and ‘knocking down’ miR-26b. Although overexpression of miR-26 in immortalised breast fibroblasts was associated with decreased proliferation and increased apoptosis, this was felt to be not physiologically relevant since the degree of over-expression exceeded the differential expression seen in clinical samples.

Reducing the level of miR-26b in immortalised breast fibroblasts resulted in a small decrease in proliferation, and a very notable increase in migration and invasion. Next, the functional influences of fibroblasts with reduced miR-26b on breast cancer epithelial cells were examined. There was no difference in the growth of cancer epithelial cells when they were co-cultured with fibroblasts with low miR-26b as compared to controls. However, these fibroblasts increased the migration and invasion of breast cancer epithelial cells.

To identify the potential targets and pathways targeted by miR-26b, I used a combination of mass spectrometry and in-silico analysis. I identified multiple proteins that were differentially expressed between miR-26b knockdown and control fibroblasts. COL12-A1, TNKS1BP1 and CPSF7 were identified as potential targets of miR-26b. I used publically available databases to examine the clinical relevance of these targets in the stroma of breast cancers. Higher expression of these targets was associated with increased rates of recurrence of breast cancer. In addition, pathway analyses of the differentially expressed proteins were enriched for components of the “glycolytic” pathway and “cytoskeleton regulation by RHO GTPase” pathway.

In conclusion, I have shown that down regulation of miR-26b in breast cancer fibroblasts increases the migration and invasion of breast cancer epithelial cells, and I have identified molecular changes that are downstream of miR-26b.

GLOSSARY

AGO	ArGOnaute
ATCC	American Type Culture Collection
CAF	Carcinoma- Associated Fibroblast
cDNA	complementary Deoxyribo-Nucleic Acid
DMEM	Dulbecco's Modified Eagle Medium
DGCR8	DiGeorge Syndrome Critical Region 8
DMSO	DiMethyl SulfOxide
dNTP	deoxyriboNucleotide TriPhosphate
dTTP	deoxyThymidine TriPhosphate
ECM	Extra-Cellular Matrix
ER	Oestrogen Receptor
FACS	Fluorescence Activated Cell Sorting
FCS	Foetal Calf Serum
FFPE	Formalin Fixed Paraffin Embedded
GFP	Green Fluorescent Protein
H&E	Haematoxylin & Eosin
HER	Human Epidermal growth factor Receptor-2
hTERT	Human Telomerase Reverse Transcriptase
IDC-NST	Infiltrating Ductal Carcinoma-No Special Type
ISH	In-Situ Hybridisation
LCM	Laser Capture Microdissection
LOH	Loss Of Heterozygosity
MCF-7	Michigan Cancer Foundation-7

MIAME	Minimum Information About a Microarray Experiment
MiRNA	Micro RiboNucleicAcid
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium
NF	Normal Fibroblast
ORF	Open Reading Frame
PACT	Protein ACTivator of the interferon-induced protein kinase
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PR	Progesterone Receptor
PTEN	Phosphatase and TENsin homolog
qRT-PCR	quantitative Real Time- Polymerase Chain Reaction
Ran-GTP	RAs-related Nuclear protein-Guanosine-5'-TriPhosphate
REC	Research Ethics Committee
RISC	*RNA*-Induced Silencing Complex (RISC)
RNA	Ribo Nucleic Acid
RNase	Ribonuclease
RPMI	Roswell Park Memorial Institute
RT	Reverse Transcriptase
SDPP	Stromal Derived Prognostic Predictor
SILAC	Stable Isotope Labelling with Amino Acids in Cell culture
STR	Short Tandem Repeats
TAR	HIV-1 TransActivating Response
TARBP	TAR RNA-Binding Protein
TP53	Tumour Protein 53
WHO	World Health Organisation

TABLE OF CONTENTS

1 INTRODUCTION

1.1 BREAST CANCER

Breast cancer is the most common cancer in females in the United Kingdom (UK). In the UK there were 49,564 new cases of breast cancer diagnosed in 2010. The number of cases diagnosed per year has been steadily increasing over the past 35 years (1975-2010) and this trend is reflected in the incidence rate, which is currently at 126 per 100,000 population. However, survival rates have been increasing over the past 40 years with the advent of improved treatments strategies and newer modalities. Even so, in 2008, there were about 12,000 women who died from breast cancer **(CancerResearchUK)**.

1.2 BREAST CANCER HETEROGENEITY

Breast cancer is a heterogeneous disease and this is reflected in the various morphological subtypes of breast cancer that are recognised by the World Health Organisation (WHO) classification system **(Lakhani and Ellis, 2012)**. More recently, gene expression profiling has identified various molecular subtypes of breast cancer. **(Fan et al., 2006)**. This heterogeneity in breast cancer encompasses substantial variation in not only the morphology and molecular profiles, but also in clinical behaviour and response to treatment. This variation is most likely a reflection of the various biological processes that underpin the development of breast cancers **(Bertucci and Birnbaum, 2008; Stingl and Caldas, 2007)**. Therefore, it is valuable to have a better understanding of these processes to better predict clinical behaviour of breast cancers and develop further improved treatment strategies. This understanding is also useful when undertaking research into breast cancer, since it would help with better design of experiments by taking into consideration the heterogeneity of breast cancer. Therefore, I will first provide an overview of the classification systems of breast cancer and their key features.

1.2.1 HISTOLOGICAL CLASSIFICATION OF BREAST CANCER

1.2.1.1 MORPHOLOGICAL SUBTYPES

The WHO classification describes 18 different morphological subtypes of invasive breast cancer that are of epithelial origin **(Lakhani and Ellis, 2012)** (**Table 1.1**). Of these, the more common subtypes include infiltrating ductal carcinoma-no special type (IDC-NST), tubular carcinoma, invasive lobular carcinoma, mucinous carcinoma, invasive cribriform carcinoma, neuroendocrine carcinoma, metaplastic carcinoma, papillary carcinoma, apocrine carcinoma and medullary-like carcinoma. Each subtype has strict morphological criteria that are used to define it.

For the purposes of pathological reporting of breast cancer for clinical use, tubular carcinoma, invasive lobular carcinoma, mucinous carcinoma, invasive cribriform carcinoma, neuroendocrine carcinoma and medullary-like carcinoma fall into the category of “special types”. The common thread that combines these tumour subtypes is that they show the classical morphological characteristics (“special pattern”) that define the subtypes. For example mucinous carcinoma will contain islands of uniform cells that appear to sit in lakes of mucin. For a tumour to be classified as a special subtype, the special pattern will need to be present in greater than 90% of the tumour mass.

In contrast, tumours that fall into IDC-NST category do not show any unique morphological features. In addition, there is significant case-to-case variation in the morphological features. For a given tumour to be classified as IDC-NST, such nondescript areas must constitute at least 50% of the tumour mass as judged by a detailed morphological assessment of representative sections from the tumour. It should be noted here that the term “representative” is not very well defined but would generally mean at least one section from each tumour block that was sampled at the time of macroscopic examination of the breast cancer specimen. The remaining part of the tumour can be formed by any of the special types.

From the above discussion it is obvious that occasionally some tumours will show special type characteristics that forms greater than 50% but less that 90% of the tumour. This pattern is not infrequent and is generally referred to as having a mixed pattern.

Tumours of breast	
Invasive breast carcinoma	**Others**
Invasive ductal carcinoma, not otherwise specified (NOS) (40-70%)	Precursor lesions
Invasive lobular carcinoma (5-15%)	Benign epithelial lesions
Tubular carcinoma (2-7%)	Myoepithelial tumour
Invasive cribriform carcinoma (0.8-3.5%)	Mesenchymal tumours
Medullary carcinoma (1-7%)	Fibroepithelial tumour
Mucinous carcinoma (<2%)	Lymphoma
Neuroendocrine tumours (2-5%)	Metastatic carcinoma
Invasive papillary carcinoma (1-2%)	
Invasive micropapillary carcinoma (2%)	
Apocrine carcinoma (0.3-4%)	
Metaplastic carcinomas (<1%)	
Lipid-rich carcinoma (1-6%)	
Secretory carcinoma (0.15%)	
Oncocytic carcinoma *	
Adenoid cystic carcinoma (0.1%)	
Acinic cell carcinoma *	
Glycogen-rich clear cell carcinoma 1-3%)	
Sebaceous carcinoma *	

Table 1.1 2012 WHO classification of breast tumours. Invasive breast carcinoma includes only those that are of epithelial cell origin. The proportion of subtypes is indicated in brackets. * These subtypes are very rare.

The morphological subdivisions are clinically useful since many of these subtypes have distinct clinical presentations and associated prognostic implications **(Ellis et al., 1992)**. Most of the special subtypes, such as mucinous tumour, tubular carcinoma, medullary–like carcinoma and invasive cribriform carcinoma, have excellent overall survival rates. For example in one series the overall survival of tubular and mucinous carcinoma was equal to age matched women without breast cancer **(Diab et al., 1999)**. However, the majority of breast cancers belong to IDC-NST. The 10 year survival rate for IDC-NST ranges between 35 to 50% **(Lakhani and Ellis, 2012)**. Therefore, this morphological classification system is of limited value since it fails to stratify most of the breast cancers into subgroups with distinct clinical behaviour. In addition, typing of breast carcinoma in this way has been shown to relatively poorly reproducible (overall kappa 0.61) by the National Health Service breast cancer screening programme (NHSBSP).

1.2.1.2 HISTOLOGICAL GRADE, LYMPH NODE STATUS AND TUMOUR SIZE

In addition to the classification based on histological subtype, breast cancer can also be classified into subgroups based on a number of other pathological parameters that have a strong influence on the outcomes of breast cancer, such as tumour grade, lymph node status and tumour size **(Elston and Ellis, 1991)**.

Patients who have ≥ 1 positive lymph node in the axilla generally have a poorer prognosis. Similarly, patients who have a large tumour size have a worse outcome than those with small tumour size. The notion of tumour grade is more complex and will be explained in detail.

Tumour grade is based on the assessment of three key morphological parameters, namely tubule formation, nuclear pleomorphism and mitotic index. These parameters amount essentially to an assessment of the degrees of differentiation (tubule formation and nuclear pleomorphism) and proliferative activity (mitotic index), and give an indication of the aggressiveness of the tumour. The basis for the current grading system was put forward by Bloom and Richardson and is known as the Bloom-Richardson grading system. Since

this grading system is based on the subjective assessment of these morphological parameters it is prone to poor reproducibility. To improve this Elston and Ellis developed the assessment further and made it more quantitative **(Elston and Ellis, 1991)**.

The Elston and Ellis modification of the Bloom-Richardson scoring system consists of assigning a score between 1 and 3 to each of the morphological parameters. A final score is calculated by adding the scores for the individual parameters and this is used to determine the grade of the tumour (**Table 1.2**). This has helped improve the concordance between pathologists for the grading of breast cancers **(Longacre et al., 2006)**. Although the histological grade of the tumour was not included in the 2003 revised AJCC cancer staging manual, it has been incorporated into breast cancer treatment algorithms such as the Nottingham prognostic index and "Adjuvant online" **(Singletary et al., 2002)**.

Total scores of the three morphological parameters (each 1-3)	Grade of tumour
3, 4 or 5	1
6 or 7	2
8 or 9	3

Table 1.2. The scores for tubule formation, nuclear pleomorphism and mitotic index are added together to obtain a total score. The grade of the tumour is based on the total score.

The biological basis for the grading system has also been demonstrated in a study that identified a "gene expression grade index" **(Sotiriou et al., 2006)**. This index correlated very well with Grade 1 and Grade 3 tumours. Interestingly, the index did not correlate very well with the Grade 2 tumours and instead appeared to show overlapping features of Grade 1 and Grade 3. In addition, when recurrence data were examined, the index appeared to reclassify the Grade 2 tumours into those had a low risk of recurrence and a high rate of recurrence **(Sotiriou et al., 2006)**. Clearly this indicates that there is a greater degree of heterogeneity within the grade 2 tumours, with some tumours behaving like Grade 1 tumours and some like Grade 3 tumours. It is interesting to note that in the study that examined the concordance for the grading system, they noticed that although the

kappa values for Grade 1 and Grade 3 was very good, this was not the case for Grade 2 tumours **(Longacre et al., 2006)**. Therefore, it is reasonable to speculate that tumours that belong to the category of grade 2 tumours are likely to be an “artifactual” group that arises from the limitations inherent in the morphological assessment of tumours.

1.2.2 MOLECULAR CLASSIFICATION OF BREAST CANCER

Recent molecular studies have reinforced the idea of tumour heterogeneity. These studies have also led to a paradigm shift in the way that we classify breast cancers. One of the earliest studies that contributed to this molecular classification was from Botstein’s group **(Perou et al., 2000)**. The investigators examined the gene expression profiles of 36 cases of IDC-NST, 2 cases of invasive lobular carcinoma and one case of ductal carcinoma in-situ. They were initially able to identify 4 distinct subgroups that they defined as ER+/luminal-like, basal-like, Her-2-like and normal-like. They subsequently published a second study that confirmed these subtypes **(Sorlie et al., 2001)**. They also refined the ER+/luminal-like group by subdividing this into two groups named Luminal type A and Luminal type B, with distinct expression profiles. They were also able to show different outcomes for these subgroups, albeit using univariate analysis and a small sample size (n=49).

Although there is some debate regarding these intrinsic subtypes **(Gusterson, 2009; Moinfar, 2008)**, they are now generally accepted and the existence of the intrinsic subtypes has been recently validated by other investigators using a larger number of tumour samples **(Network, 2012)**. They used multiple analytical platforms including genomic DNA copy number arrays, DNA methylation arrays, messenger RNA arrays, microRNA sequencing and reverse-phase protein arrays to examine 348 breast tumours. They were able to confirm the existence of the four main intrinsic breast cancer subtypes i.e. luminal A, luminal B, basal like and Her-2. However this study also demonstrated significant heterogeneity of somatic mutations, DNA methylation, miRNA expression and protein expression even within these classes. For example, the most commonly mutated

genes were TP*53*, *PIK3CA* and *GATA3.* These mutations were present at an overall frequency of 37%, 36% and 11%, with similar diversity within the subtypes, highlighting the heterogeneity even within these subgroups.

In addition to this intrinsic classification of breast cancer, there have been numerous studies that have classified tumours using different gene sets to define prognostic subgroups **(Chang et al., 2005; Ma et al., 2004; Paik et al., 2004; Parker et al., 2009; van 't Veer et al., 2002; Wang et al., 2005)**. For example, in one of the studies, investigators examined 78 breast tumours from patients with lymph node negative breast cancer and identified a 70 gene signature that predicted recurrence **(van de Vijver et al., 2002)**. Using this signature they were able to identify correctly the outcome in 65 of the 78 patients.

The concordance of these various classification systems was examined in another study **(Fan et al., 2006)**. They concluded that although there is limited overlap between the gene sets of these gene predictors, they found that these classification systems had high rates of concordance in their outcome predictions for the individual samples **(Buyse et al., 2006; Chang et al., 2005; Fan et al., 2006; Parker et al., 2009; Wang et al., 2005)**. Interestingly, in all but one of the classification systems, grade remained an independent prognostic predictor of outcome even on multivariate analysis **(Fan et al., 2006)**. A number of these gene expression predictors are being marketed under various trade names and a few of them are being validated in clinical trials (**Table 1.3**).

Recently, it has been shown that by using a limited set of five immunohistochemical markers, one is able to subdivide breast cancer into the intrinsic subtypes with distinct clinical behaviour **(Blows et al., 2010)**. This suggests that use of these markers might be robust enough for routine clinical practice where treatment can be targeted to specific groups. This strategy has even been adopted in analysing data from clinical trials **(Hugh et al., 2009)**.

Trade name	Patient group tested	Current ongoing trial
Oncotype DX **(Paik et al., 2004)**	ER positive and Lymph node negative	TAILORx
Mammaprint® (NKI70) **(Buyse et al., 2006; van de Vijver et al., 2002)**	ER positive and Lymph node negative	MINDACT
Rotterdam 76-gene signature (ROT76) **(Wang et al., 2005)**	ER positive and Lymph node negative	-
PAM50 subtype assay **(Parker et al., 2009; Sorlie et al., 2001)**	ER positive and Lymph node negative	-

Table 1.3. Common gene expression predictors that are currently used as a research tools or being validated in clinical trial.

These data support the view that breast cancer is heterogeneous with different morphological appearances, genetic alterations, gene expression profiles, protein expression and clinical behaviours. This means when researchers are studying breast cancer, they should take into consideration this variability.

The source of heterogeneity that has been detailed above is thought to be predominantly due to differences within the epithelial component. However, there is emerging evidence to suggest that part of this heterogeneity might be contributed by the stroma (further discussed in section 1.3) **(Bergamaschi et al., 2008; Chang et al., 2005; Network, 2012)**. In addition to this, there is also evidence that there is significant heterogeneity even within the breast cancer stroma and this could potentially contribute to the intratumoural heterogeneity **(Finak et al., 2008)**. This is further discussed in section 1.3.2.

1.3 MICROENVIRONMENT IN BREAST CANCER

From the morphological appearance, it is obvious that breast cancer tissue consists of not only malignant epithelial cells but also a number of other cell types and the extracellular matrix. This milieu is commonly referred to as the tumour microenvironment. In addition some investigators use the term “tumour microenvironment” to refer to not only cells and extracellular matrix but also soluble factors such as signalling molecules, cytokines, etc. The different cells types mentioned earlier include vascular cells (endothelium and pericytes), the inflammatory immune cell (lymphocytes, macrophages and neutrophils), myoepithelial cells, fibroblasts and adipocytes **(Ronnov-Jessen et al., 1996)**. These non-epithelial cell types along with the extracellular matrix are commonly referred to as the tumour stroma and form the physical scaffolding within which the tumour cells sit.

The vast majority of the investigations described in section 1.2, and indeed the majority of on-going cancer studies, involves the study either of the epithelial component of carcinomas or of whole tumours including the tumour microenvironment. It is only in recent years that the tumour microenvironment itself has been a research focus with increasing interest in the role that this plays in the initiation and progression of breast cancer.

1.3.1 ROLE OF THE MICROENVIRONMENT

One of the earliest pieces of evidence for a role for the microenvironment in breast carcinogenesis came from a study that examined the degree of differentiation that occurred in murine mammary mouse carcinoma. Here, investigators grew mammary mouse carcinoma cell lines alone or in co-culture with various other embryonic tissues **(DeCosse et al., 1975)**. They were able to conclude that embryonic mammary mesenchyme was most effective at promoting differentiation of mammary mouse carcinoma. In another study when investigators irradiated the mammary gland stroma, they found that there was an increase in the incidence of tumour formation by non-irradiated epithelial cells that were injected with the irradiated mammary stroma

(Barcellos-Hoff and Ravani, 2000). In addition they also found that these tumours had a significantly larger size compared to the controls.

Although these experiments took place in the context of *in-vivo* or *in-vitro* models, there is compelling evidence from studies on human tissue that the microenvironment in human breast cancer is also important. For example, recently investigators have examined the gene expression profile of the stromal component of breast cancer in humans **(Finak et al., 2008)**. Here, investigators laser capture microdissected the stromal compartment from a series of breast cancers and examined the mRNA expression. They were able to identify three distinct classes of tumour that differed significantly in rates of recurrence. These predictors were independent of standard clinico-pathological parameters such as grade and tumour size, ER & Her-2 positivity, and lymph node status. In addition they also constructed a stromal based prognostic predictor (SDPP) that predicted outcomes using gene expression data from whole tumours. Furthermore, they demonstrated that the SDPP was an independent prognostic factor and added value by increasing the predictive accuracy of metastasis when combined with other gene expression predictors. An extracellular matrix signature has been defined based on the expression of ECM related genes using unsupervised hierarchical clustering **(Bergamaschi et al., 2008)**. Again, analysis of this signature allowed identification of breast cancers with different survival outcomes.

Investigators have also identified Loss of heterozygosity (LOH) at several loci within the stromal compartment **(Moinfar et al., 2000)**. For some of these, the LOH was present exclusively within the stromal compartment. The most frequent changes seen exclusively in the stromal compartment were on chromosomes 17q24, 16q23.1-24.2, 3p14.2, and 11q21-23.2. Another group identified somatic mutation in PTEN and TP53 genes within the stromal compartment **(Kurose et al., 2002)**. Even in benign condition such as radial scar, increased propensity to develop breast cancer in patients with this condition has been attributed, by some, in part to their abnormal stromal component **(Jacobs et al., 1999)**.

These studies indicate a clear role for the tumour stroma in influencing the behaviour of cancers. These studies have examined the stromal compartment as a whole and therefore the contribution of each cell type is not clear.

1.3.2 FIBROBLASTS WITHIN THE TUMOUR STROMA

The commonly described hallmarks of cancer are sustaining proliferation signalling, evading growth suppression, avoiding immune reaction, enabling replicative immortality, activating invasion and metastasis, inducing angiogenesis, resisting cells death and deregulation of cellular metabolism. For most of these hallmarks, there is evidence for a contribution from stromal cells of the tumour stroma to these characteristics **(Hanahan and Weinberg, 2011)**. The degree of contribution of one particular cell type will vary depending on the organ system and tumour type. In breast cancer, each of the individual stromal cell types, including fibroblasts **(Holliday et al., 2007; Orimo et al., 2005)**, myoepithelial cells **(Jones et al., 2003)**, inflammatory cells specifically macrophages **(Wyckoff et al., 2004; Wyckoff et al., 2007)**, endothelium **(Buchanan et al., 2012)** and adipocytes **(Dirat et al., 2011)**, have been investigated in the context of breast cancer **(Coussens and Werb, 2002)**. However, the main cellular component of the tumour stroma in invasive breast cancer is fibroblasts, and these are likely to have important roles to play in modulating the behaviour of cancer epithelial cells. Therefore, they will be my focus of investigation.

The fibroblasts within the tumour stroma are commonly referred to as carcinoma-associated fibroblasts (CAF) **(Ronnov-Jessen et al., 1995)**. There are various theories on the origin of CAFs. They could potentially arise from resident stromal fibroblasts, bone marrow-derived mesenchymal stem cells or from trans-differentiation of epithelial cells, adipocytes, endothelial cells or pericytes. It is also likely that the origins of CAFs from these sources are not mutually exclusive. Although the true origin of CAFs remains uncertain, the majority view is that CAFs arise from normal resident stromal fibroblasts.

It is now fairly well established that CAFs promote tumour initiation and promote tumour progression **(Place et al., 2011)**. There have been a number of studies that have examined how the CAFs exert this influence. In the majority of these studies investigators have concluded that the fibroblasts secrete tumour promoting factors such as SDF-1 **(Orimo et al., 2005)**, SDC1 **(Maeda et al., 2006)**, HGF **(Jedeszko et al., 2009)** and TGF-beta **(Kojima et al., 2010; Kuperwasser et al., 2004)**. These in turn act on either the

tumour epithelium or on the vascular endothelium to promote tumour growth or invasion. In addition to secreted factors that have direct influences on epithelial cells, there is evidence that CAFs can modify the structural characteristics of the tumour microenvironment, which might influence epithelial cell behaviour. For example, one of the protein families that have been extensively studied in the context of invasion and metastasis of breast cancer is the matrix metalloproteinases (MMPs). The mechanism by which MMPs exert their influence is through degradation of proteins in the extracellular matrix **(Duffy et al., 2000)**. There is also emerging evidence that abnormal physical characteristics such as stiffness and abnormal collagen cross-linking can contribute to breast cancer tumour progression **(Levental et al., 2009)**. Clearly the above mechanisms are not mutually exclusive and there is extensive interplay between them.

One of the concerns in the experiments described above is the “stability” of the “CAF phenotype”. This is particularly relevant in studies where investigators have isolated CAFs from primary breast cancers and subsequently used them either as primary fibroblasts or immortalised fibroblast cell lines. It is generally accepted that CAFs originate from normal resident stromal fibroblasts. These normal fibroblasts acquire the various properties (as described above) and transform into CAFs in the presence of adjacent malignant epithelial cells and the surrounding microenvironment. Therefore, it is reasonable to query whether these properties remain intact after the fibroblasts have been removed from this context. Although some degree of reprogramming is bound to occur when fibroblasts are isolated, for most part, the evidence supports the contention that the “CAF phenotype” is stable **(Orimo et al., 2005)**. There is further support for this line of thought, when one considers that the wide spread epigenetic changes (described below) that are largely responsible for the maintenance of the CAF phenotype are often stable changes **(Hu et al., 2005)**.

The underlying mechanism by which CAFs acquire a “tumour promoting” phenotype has also been investigated to some extent. Investigators have identified distinct epigenetic changes in fibroblasts that might play roles in defining the behaviour of breast cancers **(Fiegl et al., 2006; Hu et al., 2005)**. For example, the CDC42EP5 and HOXD4 genes have been shown to be differentially methylated between normal and tumour fibroblasts **(Hu et al., 2005)**. Other studies have specifically examined the gene expression profiles of CAFs and compared them to either other cells types **(Allinen et al., 2004)** or normal

fibroblasts **(Bauer et al., 2010)**. These studies have identified significant difference in the expression profiles of CAFs. Consistent with the findings from previous investigators, a number of the genes that were deregulated in CAFs were secreted proteins. In addition, gene expression of factors that are constituents of the extracellular matrix, such as Collagen type 1- alpha 1, were also found to be deregulated. Tumour suppressor genes such as PTEN have also been shown to play a role. PTEN inactivation in fibroblasts result in accelerated initiation, progression and malignant transformation of mammary epithelial tumours **(Trimboli et al., 2009)**. Interestingly somatic mutation or LOH could not be identified in fibroblasts **(Allinen et al., 2004; Qiu et al., 2008)**. This is in contrast to some of the aforementioned studies that identified LOH and somatic mutation within the stromal compartment of breast cancer. However, the role of miRNAs (described later), in the acquisition of “tumour promoting” phenotype by CAFs has not been examined.

1.4 MICRORNAS

MicroRNA (miRNA) are small single stranded RNA molecules, 21-23 nucleotides in length that have emerged as important players in the post-transcriptional regulation of gene expression. MiRNAs themselves do not code for proteins but instead regulate the amount of protein expressed by coding RNAs. They were first discovered in 1993, from the observation that lin-4 gene, which is essential in the development of nematode worm *Caenorhabditis elegans*, does not encode a protein **(Lee et al., 1993; Wightman et al., 1993)**. Instead, it regulates another important gene, lin-14, by the binding of the lin-4 RNA to the 3' untranslated region of the lin-14 transcript leading to translational inhibition. Since the discovery of this example, there have been a huge number of publications concerning the types, biogenesis and functions of miRNAs in many eukaryotic species. The Sanger database (http://www.mirbase.org/cgi-bin/mirna), which lists all miRNA sequences, currently contains 2578 individual miRNA sequences for humans (accessed on 1/7/2013).

1.4.1 BIOGENESIS OF MIRNAS

MiRNAs are formed by multi-stage processing of an initial transcript (**Figure 1.1**). The primary transcript is produced by RNA polymerase II, as a long transcript termed the pri-miRNA. Pri-miRNAs contain regions of intra-molecular base-pairing that allow formation of hairpin structures. These are recognised and processed within the nucleus by Drosha, an RNase III-type protein **(Lee et al., 2003)**, and its co-factor DGCR8 (DiGeorge syndrome critical region 8) to form pre-miRNAs, which comprise the isolated hairpin RNAs. The Drosha:DGCR8 complex is commonly termed the Microprocessor complex **(Gregory et al., 2004)**.

Pre-miRNAs are then transported to the cytoplasm by exportin 5 and Ran-GTP **(Kim, 2004)**. Within the cytoplasm the pre-miRNAs are further processed by Dicer **(Chendrimada et al., 2005; Doi et al., 2003)** to form ~22 nt miRNA duplexes. MiRNA

duplexes then associate with AGO proteins, TAR RNA-Binding proteins (TRBP) **(Chendrimada et al., 2005)** and/or PACT **(Lee et al., 2006)** to form RNA-Induced Silencing Complexes (RISC). This complex produces and contains the mature, single-stranded, miRNA species. The passenger strand, identified by the name of the miRNA followed by an asterisk (miRNA*), is usually degraded by the AGO proteins.

The mature miRNAs can then bind to the 3' end of target mRNAs, commonly through an imperfect complementarity that result in translational repression. The miRNA-mRNA pairing is usually mediated through the interaction between a 7-8 nucleotide region, that forms a contiguous “Watson-Crick pairing” **(Bartel, 2009)**. Here the “Watson-Crick pairing” specifically refers to A:U and G:C base pairs. Occasionally, the entire miRNA binds by perfect complementarity to the message, often within the open reading frame of the mRNA, and this result in degradation of the message (Rhoades et al., 2002). However, this is thought to be rare in humans, while it is common in plants **(Davis et al., 2005)**.

The levels of mature miRNAs in cells can be controlled at various stages. Various transcription factors (e.g. c-Myc, p53) and other proteins that are involved in epigenetic alterations of DNA (e.g. DNA Methyltransferase DNMT1 and DNMT3b) can positively or negatively regulate miRNA levels by controlling transcription. The existence of the miRNA gene within the exome of its target gene can give rise to an auto-regulatory feedback loop (e.g. miR-26b and CTDSP1, see section 7.1). In addition, control can also be exerted at (i) miRNA processing by 5’ terminal capping of transcripts, (ii) miRNA stability by post-transcriptional addition of nucleotides to the 3’ end of pre-miRNA or mature miRNA transcripts, and (iii) regulation of RISC complex activity by various RNA binding proteins (e.g. FMRP, Dnd1) **(Krol et al., 2010)**.

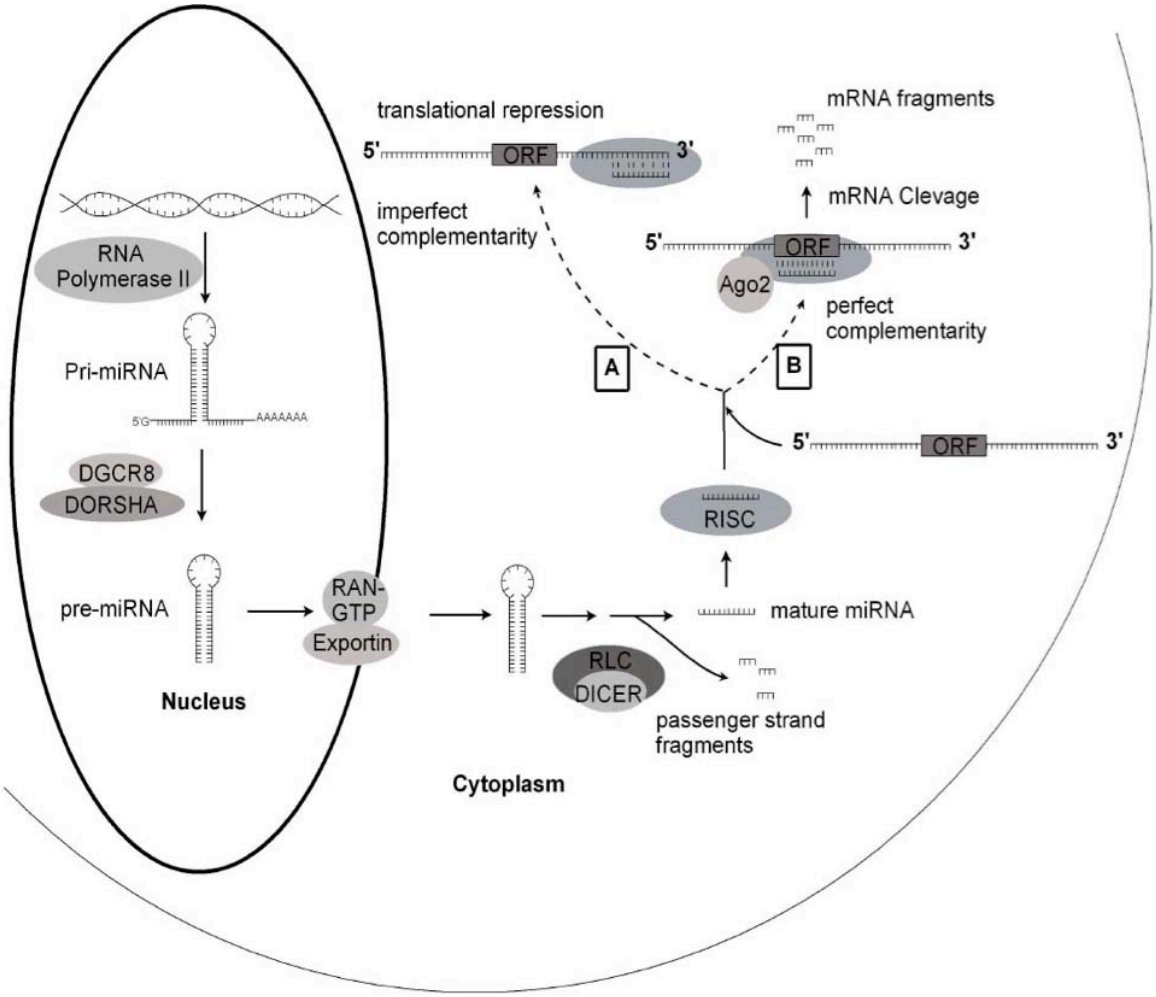

Figure 1.1. Mature miRNAs are produced by a multi-stage pathway, from initial transcription by RNA polymerase II (top left) to binding target mRNAs as part of the RISC complex. MiRNA are known to act by at least two alternative pathways (A and B). A: This is most common mechanism whereby the miRNA/RISC combines with the 3' untranslated region of the target mRNA and causes translational repression or occasionally cleavage and degradation of mRNA. This usually requires only an imperfect complementarity. B: The miRNA/RISC complex binds to the open reading frame (ORF) of target mRNAs. This requires a near perfect or, exceptionally, perfect complementarity and results in mainly cleavage and degradation of the mRNA by Argonaute2 (Ago2) (Verghese et al., 2008).

1.4.2 TARGETS OF MIRNAS

The initial strategy to identify potential targets of specific miRNAs involved computational approaches based on simple identification of mRNAs containing potential binding sites **(Lewis et al., 2003)**. Identification of the miRNA targets in plants using this approach was relatively straight-forward because of the extensive complementarity that is required between the miRNA and the target mRNA's ORF.

However, identifying the relevant targets in humans was difficult because of the non-requirement of this extensive complementarily between the 5' region of miRNA and the 3'UTR of the target mRNA **(Bartel, 2009)**. The initial algorithms performed genome-wide computational analysis by examining the 3'UTRs of mRNAs for miRNA matches. These algorithms took into consideration the degree of thermodynamic stability of the RNA:RNA duplex interaction **(Lewis et al., 2003)**. The critical component of successful target prediction was the identification of mRNAs that have a region within their 3' UTR that base pairs with the 5' region of miRNA. This region of the mRNA is called the seed region. The seed region of a miRNA is present in the 3' UTR of more than one transcript. Similarly, the 3' UTR of a single transcript will contain the seed region for more than one miRNA. Therefore each miRNA can have multiple targets and each target can have multiple miRNAs that control its translation.

Based on the degree of contiguous Watson-Crick pairing there are multiple site types with varying degree of efficacy (**Figure 1.2**). The canonical sites include 7mer-A1 site, 7mer-m8 site and 8mer site. These require base pairing to occur between the 2^{nd} to 7^{th} nucleotide of the 5' region of miRNA and the 3'UTR. The 7mer-A1 site has an additional requirement, which is to have an adenosine flanking the position 1 on the 3' UTR of the target mRNA. The 7mer-m8 sites requires an additional match at position 8. The 8mer site requires an additional match at position 8 and an adenosine flanking the position 1. In addition to the canonical sites there are marginal sites and atypical sites where the stringency for base pairing is less.

However, this approach predicted a large number of false negative results **(Lewis et al., 2003)**. Over the years the algorithms have been modified to included additional criteria

including site context, site number, site type and conservation of the target site across species **(Bartel, 2009)**. The main bioinformatics tools that rely on these algorithms include TargetScan, PicTar, miRanda, miRBase Targets and PITA. The performances of these tools were compared against the actual changes in protein output as measured by quantitative mass spectrometry when a miRNA was overexpressed or knocked down. TargetScan and PicTar emerged as the best performing tools **(Bartel, 2009)**. However, even using these tools two thirds of the predicted targets appeared to be non-responsive to manipulation of miRNA levels in cells, perhaps indicating that they were false positive predictions.

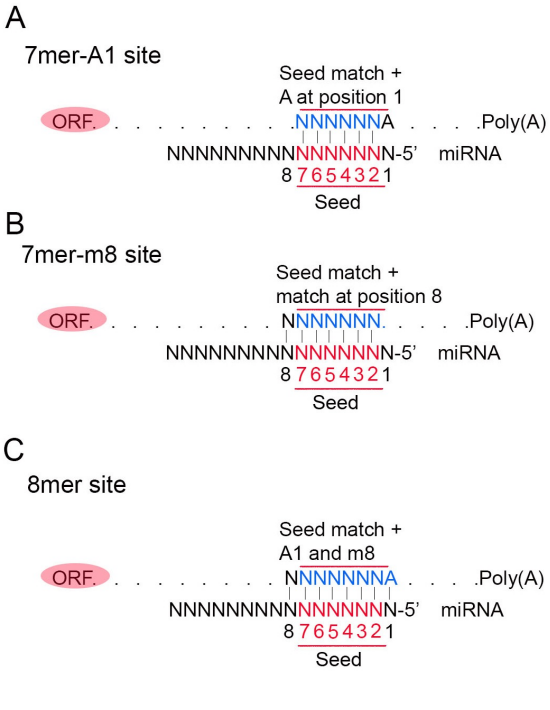

Figure 1.2. The common canonical miRNA binding site types. Six nucleotides within the 3'UTR (blue) of mRNA and the 6 nucleotides in the 5' region (red) of miRNA represents the 6 nucleotide "seed region" that undergoes base pairing. This is common for all the canonical binding sites. The pairing is represented by the solid lines between these nucleotides. In addition the 7mer-A1 site has an adenosine nucleotide at position 1 (A), the 7mer-m8 site has a matching pair at position 8 (B) and the 8mer site has both an adenosine nucleotide at position 1 and a matching pair at position 8 (C).

In addition it is also reasonable to assume that the true targets will depend on the cell type. Furthermore, it is important to bear in mind that nearly all the data that has been obtained are from a minority of miRNAs that are highly expressed. For miRNAs that have low expression levels the currently-understood rules that govern the mRNA targeting might not hold entirely true.

1.4.3 MIRNAS IN WHOLE TUMOUR OR EPITHELIUM OF BREAST CANCER

There have been a number of profiling and functional studies examining the expression and roles of miRNA in breast cancer. Iorio et al. undertook the first profiling study where they compared the miRNA profile of normal breast tissue with that of breast cancer tissue **(Iorio et al., 2005)**. The authors compared the expression profile of 76 individual cases of breast cancer with the expression profile obtained from 6 pooled samples each containing 5 cases of normal breast tissues, and 4 additional cases of breast tissues. They identified 29 miRNAs that were differentially expressed between normal and cancer tissues. There was consistent down-regulation of miR10b, miR-125b, miR-145 and up-regulation of miR-21 and miR-155. They also examined the variation in expression profile between groups defined by standard clinic-pathological parameters such as ER^+/ER^-, PR^+/PR^-, $HER2^+/HER2^-$, positive and negative lymph node status, presence and absence of vascular invasion, high and low proliferation index, and histopathological subtype. They found that in each of these comparisons there were sets of miRNAs that were differentially expressed between the groups.

A huge number of other profiling studies have followed, and have demonstrated differential expression of many individual miRNAs **(Verghese et al., 2008)**. Based on the data from these profiling studies, various functional studies have been performed where the researchers have identified a number of miRNAs that were de-regulated and appeared to function as either "Oncogenes" or "Tumour suppressor genes". For example miR-21, miR10b, miR27a have been shown to be up-regulated in breast cancer and have been shown to have functional roles in enhancing tumour growth, invasion and metastasis in cell

culture and/or xenograft carcinoma mouse models **(Ma et al., 2007; Scott et al., 2006; Si et al., 2007)**. Similarly, miR-335 and miR-126 have been shown to be down-regulated and identified as metastasis suppressor microRNAs in human breast cancer **(Tavazoie et al., 2008)**.

A key point to be noted here is that these profiling studies have used whole tumour tissue, in which the contributions from the various different cell types are combined. MicroRNAs expressed within the stroma have not been considered in these profiling studies and it has been assumed that the various reported deregulations of miRNA expression have taken place within the epithelial cells. Furthermore, the functional roles of these deregulations have been studied only within the epithelial cells.

1.4.4 ROLE OF MIRNAS IN CARCINOMA-ASSOCIATED FIBROBLASTS

There are very few published papers on the roles of miRNA in CAFs in solid tumours, although it is obvious that they are likely to contribute to the substantial change in phenotype that fibroblasts undergo during tumourigenesis. Below, I describe essentially all the published data.

1.4.4.1 BREAST CANCER

Zhao and colleagues have examined the miRNA expression profile of CAFs in breast cancer **(Zhao et al., 2012b)**. They examined the expression profile from established primary cultures of six CAFs and their paired NFs and identified a number of miRNAs that were differentially expressed. miR-221-5p, miR-31-3p, miR-221-3p were up-regulated in the CAFs compared to NFs, whereas miR-205, miR-200b, miR-200c, miR-141, miR-101, miR-342-3p, let-7g and miR-26b were down-regulated in the CAFs. However, there are a few limitations with this study. Firstly, it is not certain that the miRNA expression profiles of

fibroblasts are stable when they are removed from the context of the cancer epithelial cells and grown in tissue culture flasks. Secondly, they did not directly examine the functional roles of any of these deregulated miRNAs in CAFs.

The expression of miR-21 has been examined in a large number of breast cancers using in situ hybridisation (ISH) on tissue microarrays. It was shown that miRNA-21 was present not only in the epithelial compartment, but in many cases in both the epithelium and CAFs, and in some cancer cases solely in the CAFs **(Sempere et al., 2007)**. This was confirmed by another study where ISH was used to show that miR-21 was highly expressed in the stroma **(Rask et al., 2011)**. Increased expression of miR-21 correlated with increased cancer cell proliferation.

1.4.4.2 OTHER TUMOUR TYPES

Expression of miR-15 and miR-16 is reportedly down-regulated in prostate cancer CAFs as compared to normal fibroblasts. This down-regulation was associated with increased tumour growth and progression, apparently through reduced post-transcriptional repression of Fgf-2 (Fibroblast growth factor-2) and Fgfr-1 (Fibroblast growth factor recptor-1) **(Musumeci et al., 2011)**. Similarly, miR-31 was down-regulated in CAFs from endometrial cancer and re-expression of miR-31 impaired tumour cell migration and invasion **(Aprelikova et al., 2010)**. miR-31 and miR-214 were down-regulated in ovarian CAFs, whereas miR-155 was unregulated **(Mitra et al., 2012)**. Up-regulation of miR-31 and miR-214 and down-regulation of miR-155 simultaneously in ovarian CAFs reduced CAF migration and invasion as well as the invasion and growth of ovarian cancer epithelial cells. Although up-regulation of mir-21 has been shown to be important in colorectal carcinogenesis **(Schetter et al., 2008)**, this up-regulation apparently occurred predominantly within the stromal compartment of the tumours, in particular within the fibroblasts **(Nielsen et al., 2011)**.

2 HYPOTHESIS

My hypothesis is that changes in the levels of miRNAs regulate the behaviour of carcinoma-associated fibroblasts (CAFs) in breast cancer. These CAFs, in turn, regulate the behaviour of malignant epithelial cells. Therefore, my aims were to:

i) Identify miRNAs that are consistently differentially expressed between normal fibroblasts and CAFs;

ii) Examine the influence of these miRNAs within fibroblasts on cancer cell behaviour;

iii) Identify the molecular pathways targeted by these miRNAs in CAFs.

3 METHODS

3.1 ETHICAL ISSUES

Ethical permission for using archival breast tissues and primary breast cell cultures was obtained from Leeds (East) Research Ethics Committee, reference 06/Q1206/180 (See Appendix).

3.2 CELL CULTURE

MCF-7, HB2, T47D and BT474 cells were obtained from ATCC via Prof Valerie Speirs (LIMM, University of Leeds). The identity of these cell lines was confirmed annually by Short Tandem Repeats profiling by the Breast Research Group (BRG) in LIMM. MCF-7 and HB2 cells that stably expressed green fluorescent protein (GFP) were kindly donated by Prof Valerie Speirs. Cells that stably expressed firefly luciferase protein (MCF-7-Luc) were obtained from Cell Biolabs (Cat No. AKR-23). 293TN packaging cell line was obtained from Systems Biosciences (Cat no. LV900A-1). All cell were grown at 37^{o}C, with 5% CO_2 and were passaged when they were 70-80% confluent. When the cells lines were first obtained, multiple vials containing the cells were also frozen down as stock cells. Cells that were cultured beyond 25 passages from their first use were discarded and a new vial of frozen cells were thawed and used.

HB2 cell lines stably expressing GFP were routinely maintained in RPMI-1640, GlutaMAX™ I (Invitrogen, Cat No: 61870) with 5% FCS. Two weeks before setting up co-culture experiments with fibroblasts, the medium for these epithelial cell lines was

changed to DMEM (Invitrogen, Cat No: 31966) with 10% FCS. T47D, BT474, MCF-7-Luc and 293TN were routinely maintained in DMEM with 10% FCS.

Immortalised breast cancer fibroblasts (IF1) and normal fibroblasts (IF2) were kindly donated by Prof Valerie Speirs. These were created by viral transduction with hTERT as previously described **(Verghese et al., 2011)**. Fibroblasts that were used for the initial co-culture experiments were maintained in DMEM with 10% FCS, 1μg/ml penicillin, 1mg/ml streptomycin and 2.5μg/ml puromycin. Two weeks before the co-culture experiments the medium was changed to DMEM with 10% FCS without the antibiotics. For all other experiments immortalised fibroblasts were maintained in DMEM with 10% FCS. IF1 cells that were knocked down for miR-26b were designated as $26b^{k/d}$ whereas control knockdown cells were designated as $con^{k/d}$. IF2 that were knocked down for miR-26b were designated as $26b2^{k/d}$ where as the control knockdown cells were designated as $cont2^{k/d}$. $26b^{k/d}$, $con^{k/d}$, $26b2^{k/d}$ and $cont2^{k/d}$ were all maintained in DMEM with 10% FCS. Characteristics of all cell lines have been described in **Table 3.2**

Cell Line	Tissue of Origin	Cell type	Disease state	Specific characteristics
MCF-7	Breast	Epithelial	Metastatic adenocarcinoma cells from pleural fluid	ER+, PR+, Her-2 -
MCF-7-Luc	Breast	Epithelial	Metastatic adenocarcinoma cells from pleural fluid	ER+, PR+, Her-2 -, luciferase expressing cell line
HB2	Breast	Epithelial	Luminal epithelial cells isolated from milk, Immortalised using simian virus 40 (SV40) large T antigen	ER+, PR+, Her-2 -
T47D	Breast	Epithelial	Metastatic adenocarcinoma cells from pleural fluid	ER+, PR+, Her-2 -
BT474	Breast	Epithelial	Ductal adenocarcinoma	ER-, PR+, Her-2+,
293TN	Embryonic kidney	Unclear	Normal human embryonic kidney cells, transformed using adenovirus	SV40 large T Antigen & Neomycin resistance gene
IF1	Breast	Fibroblast	Breast cancer, immortalised by viral transduction with hTERT	-
IF2	Breast	Fibroblast	Normal breast tissue, immortalised by viral transduction with hTERT	-

Table 3.2 Characteristics of cell lines used **(ATCC, 2014; Subik et al., 2010)**.

3.3 FLUORESCENCE–ACTIVATED CELL SORTING

Fluorescence-activated cell sorting (FACS) was performed using a MoFlo Cell Sorter (Dako). Cells were trypsinised and washed twice with PBS. The samples were prepared as single cell suspension by passing them through a cell strainer (70 µm pore size) into 5 ml round bottom tubes (BD Falcon, Cat no. 352008) and resuspended in PBS with 2% FCS. For co-culture experiments, GFP positive epithelial cells (HB2 or MCF-7 cells) and fibroblasts were grown together (at appropriate ratios) in 6 well plates. GFP positive epithelial cells and fibroblasts that were grown alone were used as positive and negative controls to define appropriate flow-cytometry gates. These gates helped to separate majority of the GFP positive epithelial cells and fibroblast with minimal contamination from the second cell type. At least 10,000 cells were sorted into 5 ml round bottom tubes containing DMEM with 20% FCS. FACS was also used to enrich cultures for GFP positive transduced fibroblasts i.e. $26b^{k/d}$, $con^{k/d}$, $26b2^{k/d}$ and $cont2^{k/d}$ (See section 3.12). The percentage of cells that were GFP positive was monitored throughout the experiments that were performed with these cell lines. When percentages fell below 90%, cells that were growing in a $75cm^2$ flask were trypsinised and processed as above. Cells that were not transduced were used as negative control to define the flow-cytometry gates. The entire population was sorted into a 5 ml round bottom tube containing DMEM with 20% FCS.

3.4 SELECTION OF CLINICAL CASES

A single breast cancer case of grade 2 infiltrating ductal carcinoma-no special type (ER+, PR+, Her2-, Lymph node negative) was identified for use in the initial microarray work. Representative tumour and normal tissue blocks for this case were retrieved from Leeds Teaching Hospital Pathology Archive. Subsequently, 14 sequential cases with a similar clinical and immunohistochemical profile were chosen for validation using further laser

capture microdissection (LCM) and quantitative polymerase chain reaction (qPCR) analyses. In all instances, a Haematoxylin & Eosin (H&E) stained section of the tumour and normal tissue was examined to confirm the presence of an adequate area of stroma containing predominantly fibroblasts. Tumours that had substantial infiltration of lymphocytes in the stroma were excluded.

In addition, RNAs that were obtained from primary carcinoma–associated fibroblasts and matched normal fibroblasts from four cases were kindly donated by Prof Val Speirs' group. The pathological and immunohistochemical characteristics of all clinical samples are outlined in **Table 3.1**.

Usage	Type	Grade	LN status (+/-)	ER status (+/-)	Her2 status (+/-)
Microarrays	ductal-NST	2	-	+	-
qPCR 1	ductal-NST	2	-	+	-
qPCR 2	ductal-NST	2	-	+	-
qPCR 3	ductal-NST	2	-	+	-
qPCR 4	ductal-NST	2	-	+	-
qPCR 5	ductal-NST	2	-	+	-
qPCR 6	ductal-NST	2	-	+	-
qPCR 7	ductal-NST	2	-	+	-
qPCR 8	ductal-NST	2	-	+	-
qPCR 9	ductal-NST	2	-	+	-
qPCR 10	ductal-NST	2	-	+	-
qPCR 11	ductal-NST	2	-	+	-
qPCR 12	ductal-NST	2	-	+	-
qPCR 13	ductal-NST	2	-	+	-
qPCR 14	ductal-NST	2	-	+	-
1^{o} cultures 1	ductal-NST	2	-	+	-
1^{o} cultures 2	ductal-NST	3	-	+	-
1^{o} cultures 3	ductal-NST	1	-	+	-
1^{o} cultures 4	Mixed	1	unknown	+	+

Table 3.1. Pathological features of the breast cancer cases included in this study. NST: no special type; LN: lymph node; ER: oestrogen receptor alpha.

3.5 LASER CAPTURE OF EPITHELIAL AND STROMAL COMPONENTS

For each archival block up to four 10 µm sections were prepared on uncoated glass slides. These were then placed on a hot plate (55^0C) for 5 min to fix the tissue to the slides. The duration was increased to 8-10 min if there was considerable fat content within the section on visual inspection. The sections were allowed to cool to room temperature. They were then transferred into xylene (maintained at 60^oC), and 100% alcohol (room temperature) for 3 min each. They were then air dried for 1 min, stained with 1% toluidine blue for 30 sec and then washed with 100% alcohol.

Cell were visualised using the Zeiss/PALM Laser Capture Dissection Microscope using 100X and/or 200X magnification. This microscope was also used for laser capture. Epithelial and stromal compartments were identified from their morphology and digitally marked. Areas of increased vascularity, inflammation and necrosis were avoided. The energy setting for the laser varied between samples types and settings allowing successful dissection were obtained separately for every tissue by trial and error. For each compartment a total area 10 mm^2 was microdissected (**Figure 3.1**).

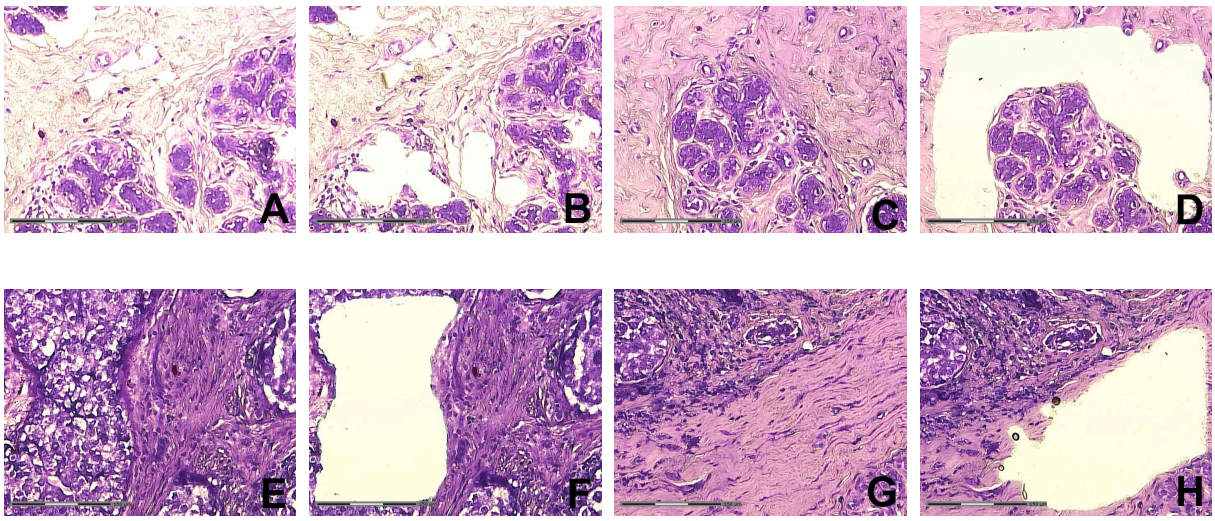

Figure 3.1. Laser-capture allows successful separation of epithelial and stromal compartments. Toluidine blue stained section of a case of breast cancer showing the normal breast tissue (A, C) and cancer tissue (E, G). The epithelial compartment of the normal (B) and cancer tissue (F) has been microdissected leaving the stromal compartment. Similarly, the stromal compartment of the normal (D) and cancer tissue (H) has been microdissected leaving behind the epithelial compartment. Scale bar: 150 µm

3.6 RNA EXTRACTION FROM CELL LINES AND FFPE SAMPLES

3.6.1 RNA EXTRACTION FROM CELL LINES FOR MIRNA

Total RNA was extracted using *mir*Vana miRNA Isolation Kit® (Life technologies, Cat No. AM1560) using the manufacturer's protocols. Briefly, the samples were first lysed using a denaturing lysis buffer. Samples were then subject to Acid-Phenol: Chloroform extraction and then passed through a column-based glass-fibre filter. RNA bound to this filter in the presence of a high salt concentration. Total RNA was finally eluted using 60 µl of nuclease free water.

3.6.2 RNA EXTRACTION FROM CELL LINES FOR MRNA

Total RNA was extracted using Qiagen’s RNA easy plus mini kit (Cat no. 74134) using the manufacturer's protocols. Briefly, the samples were lysed using a lysis buffer that contained guanidine-isothiocyanate. Lysates were then passed through gDNA eliminator columns, which allow removal of genomic DNA. The elute containing the RNA, was then passed through a membrane filter that binds to the RNA. Total RNA was eluted from this filter in 30 µl of nuclease free water.

3.6.3 RNA EXTRACTION FROM FFPE SAMPLES FOR MIRNA AND MRNA

Total RNA was extracted using Ambion's RecoverAll™ Total Nucleic Acid Isolation Kit for FFPE Tissues (Life technologies, Cat No. AM1975) using the manufacturer's protocols. Briefly, FFPE samples were deparaffinised using a series of xylene and ethanol washes. They then underwent an initial protease digestion and subsequent digestion with DNA nuclease. Total RNA was then purified on a column-based filter and eluted into 60 µl of nuclease free water.

3.7 MIRNA MICROARRAY

Microarray analyses were performed using Agilent's miRNA microarray system Version 1.7 (Cat No. G4170-90010) and the Human miRNA V2 Microarray slides (8 arrays/slide, (8*15K)), Design ID = 019118, using the manufacturer's protocols. The array work was done under guidance from Prof David Westhead (School of Molecular and Cell biology, Leeds University) and Ruth Burton (Agilent technologies).

Briefly, 100 ng of total RNA was initially dephosphorylated using calf intestinal phosphatase (TaKaRa Cat No. 2250A). Secondary structure was disrupted by addition of DMSO (Sigma, Cat No. D8418). 3' ends of the RNAs were labelled using a cyanine dye (Cy3) by ligation. This ligation was performed using T4 RNA Ligase (New England Biolabs, Cat No. M0204S). Labelled RNA was purified using Micro Bio-Spin® Columns with Bio-Gel® (Bio-Rad, Cat No. 732-6221). Agilent's blocking agent was used to prevent non-specific binding. Hybridisation was carried out at 55^{o}C for 20 hours. The slides were then washed using Agilent's Gene Expression Wash buffers (Agilent, Cat No. 5188-5327) and scanned using Agilent's microarray scanner. The settings used for the scanner were as follows (**Table A1**).

Agilent's features extraction software (Version 9.5.3) was used to probe the data. This gave an output in tabular format, which listed all the miRNAs, the signal intensity, and a final column stating whether the miRNA was detected or not. Further analysis of this data was done using the vendor's recommendations. Standardisation was done so that all signal intensity values for each feature below 0.01 (including negative values) were reassigned to 0.01. The signal values were transformed to log base 2 and 75-percentile normalization was performed. A list of all miRNAs with their corresponding fold changes was calculated.

The files required as part of the MIAME (Minimum information about a microarray experiment) are available in the folder "miRNA Array data" in the submitted CD and include the array layout, design file, raw data, data processed by Feature extraction and normalised data **(Brazma et al., 2001)**.

3.8 REAL TIME – QUANTITATIVE POLYMERASE CHAIN REACTION

3.8.1 MIRNA

MiRNAs were quantified using qRT-PCR. This was performed using Taqman assays (Life technologies, Cat no. 4427975). The quantity of miRNAs in microdissected samples was expected to be very small. Initial experiments using NCode™ miRNA First-Strand cDNA Synthesis and qRT-PCR Kit (Life technologies, Cat no. MIRQ-100) for quantifying miRNA showed that the levels of miRNAs in microdissected samples were close to the limit of detection. Therefore a pre-amplification step was necessary to quantify the levels of miRNAs in these samples reliably. A modification of the Applied Biosystems' protocol for miRNA expression profiling was used for this purpose. Expression profiles using this pre-

amplification step have been shown to have strong correlations to expression profiles obtained from non-amplified samples **(Mestdagh et al., 2008)**.

Briefly, for each sample, 0.8 µl of MegaPlex^{TM} RT Primers Human Pool A (Life technologies, Cat No. 4401090), 0.2 µl of dNTP with dTTP (100mM), 1.5 µl of MultiScribe Reverse transcriptase (50U/µl), 0.80 µl of 10X RT Buffer, 0.9 µl of $MgCl_2$, RNase inhibitor (20U/µl) and 0.20 µl of nuclease free water was combined to make a total of 4.5 µl of RT reaction mix. 3 µl of RNA, containing 100ng of total RNA was added to this master mix. This was incubated on ice for 5 min. It was then transferred to a thermal cycler and reverse transcription was performed.

2.5 µl of the reverse transcription product was then combined with 12.5 µl of the TaqMan® PreAmp Master Mix (Life technologies, Cat no. 4391128), 2.5 µl of MegaPlex^{TM} RT Primers Human Pool A (Life technologies, Cat no. 4399233) and 7.5 µl of nuclease free water in a PCR tube. This was transferred to a thermal cycler using the thermal cycling conditions outlined in Life technologies protocol for "MegaPlex Pools, for miRNA expression analysis" (Thermal cycling parameters: Hold: 95^0C, 10 min, Hold: 55^0C, 2 min Hold: 72^0C, 2 min, Cycle (12 cycles): 95^0C, 15 sec; 60^0C 4 min, Hold: 99.9^0C, 10 min).

The resultant product contained the amplified cDNA. This was then diluted with 75 µl of 0.1 x TE buffer, pH 8.0.

For each sample, 0.5 µl of TaqMan® Small RNA Assay (20X) (Life technologies, Cat no. 4427975), 0.5 µl of amplified and diluted RT product and 5 µl of TaqMan® Universal PCR Master Mix, No AmpErase® UNG (Life technologies, Cat no. 4324018) was combined in a 1.5 ml microcentrifuge tube. This was made up to 10 µl using nuclease free water and inverted several times to mix. 10 µl of this mix was transferred into a 96 well plate. This was then amplified on an Applied Biosystems 7900HT Fast Real-Time PCR System using the thermal cycling conditions outlined in Applied Biosystems protocol for "Taqman Small RNA Assay". All reactions were performed in triplicates and data was analysed using the same system.

The two reference genes were used: U6 snRNA and RNU48 (SNORD48). U6 is small nuclear RNA present in the ribonucleoprotein complex that forms the spliceosome **(Forne et al., 1996)**. RNU48 is a small nucleolar RNA that are required as part of rRNA synthesis **(Kiss-Laszlo et al., 1996)**. They were selected on the basis that they were the most abundant and least variable in a series of cell lines and normal human tissue that were tested by Life technologies **(Biosystems, 2007)**.

3.8.2 MRNA

mRNA was also quantified using qRT-PCR. This was done using Taqman gene expression assay (Life technologies, Cat No. 4331348). Briefly, total RNA was extracted using Qiagen's RNA easy+ mini kit. Reverse transcription to cDNA was performed using High Capacity cDNA Reverse Transcription Kit (Life technologies, Cat no. 4368814). For this at least 100ng of total RNA was used. For each sample 2 µl of RT Buffer, 0.8 µl dNTP, 2 µl of random primers and 1 µl of MultiScribe reverse transcriptase was added and made up to 10 µl with nuclease free water. This was then transferred to a thermal cycler and reverse transcription was performed (Thermal cycling parameters: Hold: 25^{o}C, 10 min, Hold: 37^{o}C. 120 min, Hold: 85^{o}C, 5 min).

2 µl of the cDNA template (reverse transcription product) was mixed with 0.5 µl of Taqman gene expression assay and 5 µl of TaqMan® Universal PCR Master Mix, No AmpErase® UNG and made up to 10 µl with nuclease free water. 10 µl of this mix was transferred to a 96 well plate. It was then amplified on an Applied Biosystems 7900HT Fast Real-Time PCR System (Thermal cycling parameters: Hold: 50^{o}C, 2 min, Hold: 95^{o}C, 10 min, Cycle (40 cycles): 95^{o}C, 15 sec; 60^{o}C, 1 min). All reactions were performed in triplicates and data was analysed using the same system.

The reference gene used was Glyceraldehyde 3-phosphate dehydrogenase (GAPDH). GAPDH is one of the most commonly used reference gene and it has been validated in large panel of human tissue **(Barber et al., 2005)**.

3.9 TRANSFECTION OF CELL LINES

3.9.1 TRANSFECTION OF FIBROBLASTS WITH PRE-MIR PRECURSOR MOLECULES OR ANTI-MIR INHIBITORS

Transfection complexes were initially prepared by adding 1 to 5 µl of 6.25 µM of Pre-miR miRNA-26b, miRNA-30b or miRNA-21 Precursor Molecules (Life technologies, Cat no. AM17100) to a final concentration of 10 to 50 nM, and 3 µl of HiPerFect (Qiagen, Cat no. 301704) to FCS free DMEM medium for experiments involving overexpression of miRNAs. The total volume was made up to 125 µl using FCS free DMEM medium. 25 µl of 6.25 µM of mirVana® miRNA-26b Inhibitors (Life technologies, Cat no.464084), for a final concentration of 250 nM, was used for transient knockdown of miR-26b. Equivalent amounts of Pre-miR™ miRNA Precursor Molecules-Negative Control #1 (Life Technologies, Cat no. AM17110) or, mirVana® miRNA Inhibitor-Negative control #1 (Life Technologies Cat no. 4464076) were used as controls. The transfection complexes were left at room temperature for at least 5 min. In the meantime, 0.5 ml of FCS free DMEM medium containing cells at a concentration of 10,000 to 30,000 cells/ml was added to each well of a 24 well plate. This was incubated at 37^{o} C for 5 min. 125 µl of the transfection complexes was added to each well for a final volume of 625 µl. The media was changed to DMEM medium with 10% FCS after 18 hours.

3.9.2 TRANSFECTION OF FIBROBLASTS WITH S*ILENCER*® SELECT SIRNAS

Transfection complexes were prepared by adding 0.125 to 1.25 µl of Silencer® Select siRNAs (Life Technologies Cat no. 4390843) for a final concentration of 1.25 to 12.5 nM, and 3 µl of HiPerFect to FCS free DMEM. The total volume was made up to 50 µl using FCS free DMEM. Silencer® Select-Negative Control #1 (Life Technologies Cat no. 4404021) was used as the negative control. Transfection complexes were kept at room temperature for 5 min. 450 µl of FCS free DMEM medium containing cells at a concentration of 3 x 10^4 cells/ml were added to each well of a 24 well plate. This was incubated at 37^o C for 5 min. 50 µl of the transfection complexes was added to each well for a final volume of 500 µl. The media was changed to DMEM medium with 10% FCS after 18 hours.

Knockdown efficiency for each siRNA was calculated for varying concentrations. The lowest final concentrations of siRNA that produced greater than 90% knockdown was 2.5 nM, 5 nM, 12.5 nM and 10nM for MTDH-1 (Metadherin-1), MTDH-2 (Metadherin-2), CDH-11 (Cadherin-11) and PDGFR-α (Platelet derived growth factor receptor) respectively. This was the concentration that was used for subsequent knockdown experiments.

3.10 FUNCTIONAL ASSAYS

3.10.1 CELL VIABILITY ASSAY

Cell viability assays were performed in 24 well plates using 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) (Sigma, Cat no. M-2128). Briefly, 10,000 to

15,000 cells were plated in each well of a 24 well plate. Transfections were performed if required. A working solution of MTT was prepared at a concentration of 1mg/ml using PBS. The media in each well of the 24 well plates was replaced with 250 µl of MTT working solution. It was then incubated in the dark at 37^{o}C for 3 hours. MTT solution was then removed and replaced with 250 µl of propan-2-ol. 50 µl of this mixture was then transferred on to a 96 well plate. The optical density was measured using a microplate reader (Opsys MR^{TM} Dynex technologies Ltd, UK) at 570nm.

3.10.2 APOPTOSIS ASSAY

Apoptosis assays were performed by quantifying the percentage of cells in early to mid apoptosis using the PE Annexin V Apoptosis Detection Kit (BD Pharmingen Cat no. 559763) at 24 and 48 hours. Cells were plated in 6 well plates. Transfections were performed if required. Cells were trypsinised and washed twice with PBS. The cells were then resuspended in 500 µl of 1x Annexin V Binding Buffer. 100 µl of the cell suspension was transferred into 5 ml round bottom tubes. Cells were stained with 5 µl of Annexin V, 5 µl of 7-amino-actinomycin (7-AAD) and incubated in the dark at 37^{o}C for 15 min. Following this, 400 µl of 1X Annexin binding buffer was added to each of the 4 tubes and mixed well. The percentage of Annexin V-positive and 7AAD-negative cells were analysed using a LSR II flow cytometer (BD Bioscience).

3.10.3 CELL CYCLE ASSAY

Cell cycle assay was performed by evaluating the percentage of cells in each stage of the cell cycle using Propidium iodide (PI) staining. Briefly, cells were grown in 6 well plates

until they were 70% confluent. The cells were trypsinised and washed twice with PBS. The cells were fixed in 2 ml of 70% ethanol maintained at 4^{o}C for at least 2 hours. The ethanol was removed and the cells were resuspended in 400ul of PBS. 100 µl of PI staining solution (0.25 mg/ml PI with 0.1% (v/v) Triton X100) was added to the cells. This mix was incubated for 10 min in the dark. Data were acquired using a LSR II flow cytometer (BD bioscience) and analysed using ModFit LT™ (Verity software).

3.10.4 MIGRATION ASSAY

3.10.4.1 TRANSWELL MIGRATION ASSAY

Transwell migration assay was performed using 24 well plate inserts with a pore size of 8 µm (BD bioscience Cat no. 353097). The number of cells that migrated through the membrane pores was quantified and compared to appropriate controls. Briefly, 10,000 to 15,000 cells were resuspended in 250 µl of serum free medium and were added to the inserts on day 1. DMEM containing 10% FCS was added to the bottom wells. Cells were allowed to migrate through the membrane for 12 to 48 hours. The inserts were removed and washed in PBS for 2 min. The cells remaining on the top of the inserts were removed using a cotton swab. The inserts were fixed in 4% formaldehyde for 2 min and then air-dried. They were then stained with 1% toluidine blue for 30 sec, washed in water and allowed to dry. Membranes were detached from the inserts and mounted onto glass slides and were covered with coverslips. The number of cells in four peripheral and one central field at 200x magnification was counted.

For experiments that required transfection with siRNAs or Anti-miR inhibitors, reverse transfection was performed as described earlier, in the inserts. Media was changed to

serum free DMEM media after 18 hours. DMEM containing 10% FCS was added to the bottom wells. Migration assays were performed as described before.

3.10.4.2 CELL SCRATCH ASSAY

Cell scratch assays were performed as described previously **(Liang et al., 2007)**. The percentage area of the scratch/"wound" that closed was determined after 18 hours. Briefly, cells were grown to a confluent monolayer in a 6 well plate. A scratch was made on the monolayer using P200 pipette tip. Reference points were marked on the bottom of the dish and images of the scratch were taken using E1000 Eclipse microscope (Nikon) at 200X magnification. The cells were incubated for 18 hours at 37^0C and a set of second images were acquired at the reference markings. The JPEG images were analysed using Photoshop® (Abode). The area of the scratch closed by cells was calculated as percentage of the initial area of the scratch.

3.10.5 INVASION ASSAY

3.10.5.1 TRANSWELL INVASION ASSAY

Transwell invasion assays were performed using 24 well plate inserts that were pre-coated with Matrigel (BD, Cat No. 354480). This assay was performed in the same manner as the migration assay described above. The number of cells that migrated through the membrane pores was quantified at 12, 24 or 48 hours.

3.10.6 CO-CULTURE ASSAYS

3.10.6.1 PROLIFERATION ASSAY

Proliferation of MCF-7-Luc cells in co-culture with fibroblasts was quantified using luciferase assays, while overall growth (i.e. that of fibroblast and MCF-7-Luc combined) was analysed using MTT assays. Briefly, MCF-7-Luc with either $con^{k/d}$ or $26b^{k/d}$ were plated together in each well of a 24 well plate. The initial epithelial to fibroblast ratio was 1:3. This was based on ratios that have been used in the literature **(Orimo et al., 2005)**. In addition, other ratios (3:1 & 1:1) were also tested. The experiments were set up in duplicate so as to perform both an MTT assay and a luciferase assay (see below) on separate wells. The MTT assay (as described before) was performed to assess the overall growth of both epithelial cells and fibroblasts at 24, 48 and 72 hours. To monitor the growth of MCF-7-Luc cells alone, firefly luciferase assay was performed at the same time points.

3.10.6.1.1 LUCIFERASE ASSSAY

Firefly luciferase reporter assay was performed using Luciferase Assay System (Promega, Cat no. E1500). Briefly, the medium was removed and cells were washed twice with PBS. 200 µl of 1X Passive lysis buffer was added to the each well of the 24 well plates and incubated for 15 min. 25 µl of this lysate was transferred to a 96 well plate. 10 µl of luciferase assay reagent II was added to this and agitated. Firefly luciferase activity was measured using Luminometer (Berthold Mithras LB 940) after 10 sec.

3.10.6.2 MIGRATION ASSAY

3.10.6.2.1 INDIRECT CO-CULTURE TRANSWELL MIGRATION ASSAY

Indirect transwell migration assays were performed using 24 well plate inserts with a pore size of 8 µm (BD bioscience Cat no. 353097). $con^{k/d}$ or $26b^{k/d}$ cells were grown in the lower chamber of a Boyden chamber and MCF-7-Luc cells were grown in the upper chamber. The number of epithelial cells that migrated through the membrane pores was quantified. Briefly, 50,000 $con^{k/d}$ or $26b^{k/d}$ were resuspended in 500 µl of DMEM medium with 10% FCS and were plated on to 24 well plates. The next day the medium was changed to serum free DMEM medium. 10,000 MCF-7-Luc that had been starved overnight by growing them in serum free DMEM medium were added to the inserts, and incubated for 2 hours. The inserts were then placed into the wells containing the fibroblasts. The number of MCF-7-Luc that migrated through the membrane pores was quantified by counting the cells as described (Section 3.10.4.1).

3.10.6.2.2 DIRECT TRANSWELL MIGRATION ASSAY

These were also performed using the 24 well plate inserts (BD bioscience, Cat no. 353097). Here, the fibroblasts and epithelial cells were grown together in the upper chamber of the Boyden chamber and the number of epithelial cells that migrated through the membrane pores was quantified. MCF-7-Luc and equal numbers of either $con^{k/d}$ or $26b^{k/d}$ were resuspended in 250 µl of serum-free DMEM media and were added to the inserts. DMEM media containing 10% FCS was added to the bottom wells. Cells that migrated through the membrane were quantified at 12 or 24 hours using the firefly luciferase assay. The inserts were removed and washed in PBS for 2 min. The cells

remaining on the top of the inserts were removed using a cotton swab. Individual inserts were then placed in 300 µl of passive lysis buffer and agitated gently for 15 min. 25 µl of this lysate was used to perform Luciferase assays as described (Section 3.10.6.1.1).

3.10.6.3 INVASION ASSAY

3.10.6.3.1 INDIRECT AND DIRECT TRANSWELL INVASION ASSAY

Indirect and direct transwell invasion assays were performed using 24 well plate inserts that were pre-coated with Matrigel (BD Cat No. 354480). The assays were performed in the same manner as the indirect and direct co-culture transwell migration assay described above. MCF-7-Luc cells that migrated through the membrane pores were quantified either by counting the cells or by using luciferase assays.

3.10.6.3.2 3D SPHEROIDAL INVASION ASSAY

3D spheroidal invasion assay were performed using spheroids consisting of fibroblasts and epithelial cells that were embedded in an extracellular matrix **(Nowicki et al., 2008)**. The invasion of epithelial cells was quantified.

For generating spheroids, equal numbers of epithelial cells and fibroblasts were suspended in DMEM medium containing 10% FCS at a final concentration of 1-2 x 10^6 /µl. 25 µl of this was placed in the centre of a 35 mm petri dish (Ibidi, Cat no. 81158) and inverted. 5 ml of DMEM with 10% FCS and added to the lid of the petri dish to prevent drying. This was incubated at 37^0C for 24 hours to allow formation of spheroids.

The next day 300 µl of a collagen-1/Matrigel mix (Appendix 9.2) was added to each well of a 24 hours well plate and incubated for 1 hour for the gel to solidify. Cells that had

formed tight spheroids were harvested using a 200 µl pipette and placed in the centre of the well containing the gel. An additional 300 µl of collagen-1/Matrigel mix was added on top of the spheroid and incubated for 1 hour at 37^0C. Finally 300 µl of DMEM was added to the well. This was then incubated for 48 to 96 hours. Phase-contrast images were obtained at 24, 48 and 96 hours using an E1000 Eclipse microscope. The invasive front of epithelial cells was quantified relative to the area of the central spheroidal area (**Figure 5.10**)

3.11 WESTERN BLOTTING

3.11.1 PREPARATION OF LYSATES

Western blots for protein of interest were performed using lysates prepared using Radioimmune precipitation assay (RIPA) buffer (Sigma Cat no. R0278). Briefly, the cells of interest were grown in 6 well or 24 well plates. Medium was removed and cells were washed twice with PBS. 1 ml or 200 µl of the working solution of RIPA buffer (Appendix 9.2) was added to each well and allowed to stand for 30 seconds. Cells were dislodged using a cell scrapper and the solution was transferred to a microfuge tube. The solution was pipetted up and down to break up any clumps. It was incubated on ice for 20 min and then spun at 10,000g for 15 min at 4^oC. The protein concentrations of the lysates were quantified (see below) and the remaining lysate was stored at -80^0C.

3.11.2 QUANTIFICATION OF PROTEINS

The protein concentration of the lysates was quantified using the DC protein assay (BIO-RAD Cat no. 500-0116). This assay is based on a colourimetric assay following detergent solubilisation. Briefly, 5 µl of either the sample or serial dilutions of the protein standard (BIO-RAD Cat no. 500-0112) was added to each well of a 96 well microtitre plate. 25 µl of a working solution of alkaline copper tartrate solution (Solution A + Solution S) was added to the wells. This was followed by 200 µl of dilute Folin's Reagent (Reagent B) into each well. The solutions were mixed by gently agitating the plate. After 15 min the optical density was measured using a microplate reader (Opsys MRTM Dynex technologies Ltd, UK) at 670nm. The protein concentration of the sample was estimated from the standard curve obtained from the serial dilutions of the protein standard.

3.11.3 GEL ELECTROPHORESIS OF PROTEINS

Gel electrophoresis of proteins was performed to allow separation of proteins based on molecular weight. Briefly, all protein samples were standardized to ensure equal volumes by adding deionised water. 5 µl of 4X NuPage® LDS sample buffer (Life technologies, Cat no. NP0007) and 1 µl beta-mercaptoethanol was added to each sample. Deionised water was added to make up to 24 µl in total. Samples were placed in hot block for 5 min at 105^0C, centrifuged, resuspended and placed on ice for 5 min. Precast NuPAGE® Novex® 4-12% Bis-Tris Gels (Invitrogen, Cat no. NP0321) were used with the Xcell SureLock™ Mini-Cell system and 1 x NuPAGE® MOPS SDS Running Buffer (Appendix 9.2). 20 µl of the samples and 6 µl of the Novex® sharp prestained protein standard (Life

technologies, Cat no. LC5800) were loaded on to the gels and electrophoresis was performed for 1 to 1.5 hours at 180V.

3.11.4 TRANSBLOTTING OF SDS GEL

Proteins that were separated using the gel were then transferred onto a Polyvinylidene difluoride (PVDF) (Thermo SCIENTIFIC, Cat No. 88518) membrane. The PVDF membrane was activated by soaking it in methanol for 30 sec. It was then rinsed for 5 min x 3 in running water and then soaked in 1 x transfer buffer (See buffer and reagents). Proteins were transferred in 1 x transfer buffer using the XCell II Blot Module (Life technologies, Cat no. EI9051) for 1.5 hours at 30V.

3.11.5 IMMUNOBLOTTING OF PROTEINS

The PVDF membrane with the transferred protein was blocked with 5% skimmed milk for 1 hour at room temperature. It was then transferred to 5% skimmed milk containing the primary antibody at an appropriate dilution and allowed to incubate either overnight at $4^{o}C$ or 1 hour at $37^{o}C$. The membrane was washed three times with PBST, with each wash lasting for at least 10 min. It was then transferred to 5% skimmed milk containing the secondary antibody at an appropriate dilution and allowed to incubate for one hour at room temperature. Three washes with PBST were repeated. The membrane was developed using either SuperSignal West Femto Chemiluminescent Substrate (Thermo SCIENTIFIC Cat no. 34095) or SuperSignal West Pico Chemiluminescent Substrate (Thermo SCIENTIFIC Cat no. 34077) and visualised using ChemiDoc® MP imaging system. Primary antibodies: Col12A1 (Sigma, Cat no. HPA009143; dilution 1:500),

TNKS1BP1 (Sigma, Cat no. HPA037929; dilution 1:250), CPSF7 (Sigma, Cat no. HPA041094; dilution 1:250), GAPDH (Sigma, Cat no. G9545; dilution 1:5000), β-actin (Sigma, Cat no. A5441; dilution 1:5000), α-SMA (DAKO, Cat no M851; dilution 1:1000). Secondary antibodies: Swine Anti-rabbit (Dako, Cat no. P0260, Rabbit Anti-mouse (Dako, Cat no. P0217. Ladder: Precision Plus Protein™ Dual Color Standards (BioRad, Cat no. 161-0374).

3.12 GENERATION OF STABLE CELL LINES WITH MIR-26B KNOCKDOWN

For transduction of cell lines the "Lentivirus expression system" from Systems Biosciences was used. This was performed in a Genetically Modified Category II cell culture facility that was approved for viral work under permission obtained by Dr Georgia Mavria (LIMM).

3.12.1 PRODUCTION OF LENTIVIRAL EXPRESSION CONSTRUCTS

Bacterial plates containing colonies expressing the plasmid that contained the lentivirus expression constructs for anti-miR-26b and a non-targeting control were obtained from Systems biosciences (Cat no. MZIP26b-PA-1 and Cat no. MZIP000-PA-1). Plasmid DNA was purified using Qiagen's plasmid Maxi kit (Cat no. 12162) following the manufacturer's protocols. Note that cultures were incubated at a lower temperature than is typical (30^{o}C rather than 37^{o}C) to reduce the chances of undesired recombination events, as recommended by Systems biosciences.

3.12.2 PREPARING PSEUDOVIRAL PARTICLES

Replication-incompetent pseudoviral particles containing the expression constructs were created using the 293TN packaging cell line. 3 x 10^6 293TN cells were plated on to 75 cm^2 cell culture flasks with 10 ml of DMEM medium with 10% FCS medium. Cells were cultured until 70% confluence. 2 µg of plasmids containing the lentivirus expression construct (either miR-26b or Control) and 20 µl of the packaging plasmid mix pPACKH1 (Systems Biosciences, Cat No. LV500A-1) was added to 400 µl of DMEM medium without serum. 20 µl of Plus reagent (Life technologies, Cat no. 11514-015) was added to the plasmid mix. It was mixed thoroughly and left for 15 min at room temperature. In the meantime, 30 µl of Lipofectamine was added to 400 µl of serum free DMEM and mixed gently. This mixture was added to the plasmid mix drop by drop. It was mixed gently and left at room temperature for 15 min. The above mix was added to the 293TN cells and cells were incubated for 48 hours at 37^{o}C. At the end of 48 hours the medium containing the pseudoviral particles was collected in a 15 ml sterile tube and centrifuged at 4000 g for 15 min. The supernatant was transferred to another 15 ml sterile tube and aliquoted into 1.5 ml sterile microfuge tubes.

3.12.3 TRANSDUCTION OF FIBROBLASTS

Cell lines with stable knockdown of miR-26b (or control) were made by transduction of the parent fibroblast cell lines using the pseudoviral particles. 1 x 10^5 fibroblasts were plated onto a 12 well plate in DMEM medium with 10% FCS and incubated overnight at 37^{o}C. The following day the medium was replaced with 1 ml of DMEM medium containing 10% FCS and polybrene (Sigma, Cat. No. H9268) at a final concentration of 5µg/ml. 100 to 400 µl of the pseudoviral particles containing the anti-miR-26b construct or non-

targeting construct was added to this. 100 to 400 µl of DMEM medium was added to control wells of cells as a mock transduced control. This was incubated overnight at 37^0C. The next day media was removed and replaced with DMEM with 10% FCS (without polybrene). Growth of cells was monitored daily and the cells were split when they were 90% confluent with subsequent transfer to 25 cm^2 and 75 cm^2 tissue culture flasks. The transduction efficiency was determined by assessing the percentage of GFP positive cells using flow-cytometry. Mock-transduced cells were used as negative controls to define the flow-cytometry gates. The cell populations were enriched using Fluorescence-activated cell sorting (FACS) with MoFlo Cell Sorter (Dako) as described before (Section 3.3).

3.13 LUCIFERASE REPORTER ASSAY

Luciferase reporter assay were performed using the pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega Cat no. E1330) for confirming the knockdown of miR-26b and the targets of miR-26b.

3.13.1 CONSTRUCTION OF REPORTER ASSAY FOR MIR-26B KNOCKDOWN

3.13.1.1 PREPARING INSERTS

For confirming the knockdown of miR-26b, oligonucleotide sequences containing the full complementary sequence of the mature miR-26b (top strand: 5'-AACCTATCCTGAATTACTTGAAG-3'; bottom strand: 5'-TCGACTTCAAGTAATTCAGGATAGGTTAGCT-3') were made to order (Life Technologies). The ends of these oligonucleotides were designed such that the annealed

pair would have restriction enzymes, Sal1 and Sac1 compatible overhangs at each end. These “sticky ends” enabled insertion of the sequence into the Sal1/Sac1 cloning sites of pmirGLO Dual-Luciferase miRNA Target Expression Vector, which is located at the 3′ of the firefly Luciferase gene. 1 Optical Density Unit (ODU) of each oligonucleotide was added to a final volume of 50 µl of TE buffer. This mix was heated to 80^0C for 3 min and was then cooled at 1 min/s to room temperature in order to allow the separate strands to anneal into double-stranded DNA.

3.13.1.2 PREPARING VECTOR AND CLONING

The vector was linearized by digesting with Sal1 (Promega, Cat no. R6051) and Sac1 (Promega, R6061) and subsequent gel purification. Different concentrations of the synthesised oligonucleotides were ligated to the linearized vector by incubating them with T4 ligase (New England Biolabs, Cat no. M0202S) at 16^0C overnight and transformations were performed with MAX Efficiency® DH5α™ Competent Cells (Invitrogen, Cat no. 18258-012). The bacteria containing plasmids from the different ligation condition were plated on to individual ampicillin containing agar plates and incubated overnight. Individual colonies were picked and screened for successful insertion of the appropriate insert on the basis that the insert recreated the Sal1 but not the Sac1 site. The final construct was verified by sequencing.

3.13.2 PLASMID REPORTERS FOR MIRNA TARGETS

Plasmids containing the predicted binding sites for miR-26b in the 3'UTRs of COL12A1, TNK1, CPSF7 was kindly provided by Dr James Thorne (LIMM). Briefly, pmirGLO Dual-Luciferase miRNA Target Expression Vector was digested with restriction enzymes Xba1

and Sac1. Target DNAs were amplified with primers (Appendix, **Table A8**) from MCF-7 genomic template to obtain regions of the 3' UTRs (Appendix, **Table A7**) that include the miR-26b binding sites. Primers included Xba1 and Sac1 sites for appropriate directional cloning. Ligations were performed overnight and were transformed into MAX Efficiency® DH5α™ Competent Cells E Coli. Presence of inserts was confirmed using colony PCR and final constructs were verified by sequencing.

3.13.3 REPORTER ASSAY

3.13.3.1 TRANSFECTION OF FIBROBLAST WITH PLASMIDS

Reporter plasmids were introduced into the cell lines by reverse transfection. Briefly, for each well of a 24 well plate, 500 ng of plasmid DNA was diluted with 25 µl of serum free DMEM, and 1 µl of Lipofectamine-2000 (Life technologies, Cat no. 11668-019) was diluted with 25 µl of SF DMEM. After 15 min the above two mixes were combined and left for 15 min at room temperature. In the meantime 25,000 to 50,000 cells suspended in 450µl of SF DMEM medium were added to each well of a 24 well plate and incubated at 37^0C for 5 minutes. Then 50 µl of DNA/Lipofectamine mix was added to the cells drop wise and incubated at 37^0C for 18 hours. Medium was changed to DMEM with 10% FCS. Luciferase reporter assays (Section 3.13.3.3) were performed after 6 hours.

3.13.3.2 CO-TRANSFECTION OF DNA PLASMID AND ANTI-MIR-26B INHIBITOR

Co-transfection of reporter plasmids and Anti-miR-26b inhibitors was performed by forward transfection. Briefly 25,000 cell were plated in each well of a 24 well plate using DMEM with 10% FCS. On the following day the transfection mix was prepared. For each well of a 24 well plate, 50 ng of plasmid DNA and 2.5 µl of Anti-miR-26b inhibitors were diluted with 50 µl of Optimem (Life technologies, Cat no. 31985-062), and separately 0.5 µl of Lipofectamine 2000 (Life technologies, Cat no. 11668-019) was diluted with 50 µl of Optimem. After 15 min the above two mixes was combined and left for 15 min at room temperature. In the mean time the growth medium was removed from the wells and the cells were washed with PBS. 100 µl of the plasmid DNA/Anti-miR-26-inhibitor/Lipofectamine mix was added to each well. This was replaced with DMEM with 10% FCS after 18 hours. Luciferase reporter assays (described below) were performed after 6 hours.

3.13.3.3 LUCIFERASE ASSAY

Dual luciferase reporter assays were performed using the Dual-Luciferase® Reporter Assay System (Promega, E1910). Briefly, medium was removed and the cells were washed twice with PBS. 200 ml of 1X Passive lysis buffer was added to the each well of the 24 well plates. 20 µl of this lysate was transferred to a 5 ml round bottom tube (Sarstedt, Cat no. 555.476), 100 µl of luciferase assay reagent II was added and mixed vigorously, and Firefly luciferase activity was measured after 10 sec using a Luminometer. Next, 100 µl of the Stop & Go® reagent was added, mixed vigorously and Renilla luciferase activity was measured after 10 seconds. The Firefly luciferase activity for each individual well was normalised to the Renilla luciferase activity.

3.14 IN-SITU HYBRIDISATION

In-situ hybridisation protocol was performed and optimised using Exiqon's miRCURY microRNA optimisation kit-2 (Cat no. 90002). 0.5 nM to 2 nM of 5' DIG-labelled LNA™ U6 snRNA probe (Exiqon, Cat no. 99002-15 Probe sequence: 5'-CACGAATTTGCGTGTCATCCTT-3', Tm: 84^{0}C) was used to optimise the protocol. 50 to 500 nM double-DIG labelled LNA™ miR-21 probe (Exiqon, Cat no. 38102-15, Probe sequence: TCAACATCAGTCTGATAAGCTA, Tm: 83 ^{0}C) for detecting miR-21 was used as a positive control, since this miRNA has been shown to be one of the most highly expressed in breast cancer tissue **(Iorio et al., 2005)**. 50 to 500 nM of double-DIG labelled LNA™ miR-26b probe was used to detect miR-26b (Exiqon, Cat no.38793-15, Probe sequence: GAGCCAAGTAATGGAGAACAGG, Tm: 81 ^{0}C). 5'-DIG labelled LNA™ scrambled microRNA probe, (Exiqon, Cat no. 99004-01, Probe sequence 5'-GTGTAACACGTCTATACGCCCA-3', Tm: 87 ^{0}C) was used as the negative control.

All reagents were prepared using nuclease free water (Appendix 9.2). All equipment was ensured to be RNAse free by cleaning with RNase Zap (Life technologies, Cat no. AM9780). A fresh microtome blade was used to cut 6 μm sections from tissue blocks onto superfrost plus glass slides. The slides were left to dry for 2 hours and then stored at 4^{0}C. The day before the experiment the slides were incubated at 60^{0}C for 45 minutes and stored at 4^{0}C. The next day the sections were deparaffinised using xylene, washed initially with decreasing concentrations of ethanol and finally with PBS. Next, the sections were treated with 300 µl of proteinase K reagent (See buffers and reagent) that had a proteinase K concentration in the range of 10 to 30 µg/ml. The slides were incubated at 37^{0}C for 10 to 30 min. The sections were washed with PBS and dehydrated using increasing concentrations of ethanol. 25 µl of the hybridization mix containing varying concentration of the LNA^{TM} probe was applied to the slides. A coverslip was placed on top of the sections, the edges sealed with rubber gum and it was hybridized for 1 hour at 50^{0}C to 60^{0}C (or approximately 30^{0}C below the Tm) in an oven. Then the coverslips were

removed and the sections were washed with decreasing concentration of SSC buffer (See buffers and reagents) maintained at the hybridisation temperature for 5 min each. A final wash with 0.2X SSC maintained at room temperature was also performed.

The slides were washed with PBS and incubated for 15 min with the blocking solution (See buffers and reagents). The blocking solution was removed and 200 µl of anti-DIG reagent (See buffers and reagents) was added and incubated for 60 min at room temperature. The slides were washed with PBS-T for 3 min x 3. Freshly prepared AP substrate (See buffers and reagents) was added to the sections and incubated for 2 hours at 30^0C in a humidifying chamber. At this point it was ensured that the slides were protected from light. The AP substrate was removed and washed with KTBT buffer (See buffers and reagents) for 5 min X 2 to stop the reaction. The slides were then washed with water for 1 min x 2. For the nuclear counter stain, 200 µl of Nuclear Fast Red^{TM} (Vector laboratories, Cat. No. H-3403) was applied for one minute. The slides were washed with water for 10 min. The sections were dehydrated with increasing concentration of ethanol, dried, mounted using Eukitt mounting medium (VWR, Cat no.3618940) and examined under a microscope.

The ability of ISH probes to bind to the target sequence and the subsequent intensity of signal obtained depends upon various parameters such the degree of proteinase K treatment, probe concentration, hybridisation temperature and stringency of washes. Therefore it was necessary to optimise these conditions. For each case, the proteinase K concentration and duration of proteinase K treatment was varied to optimise the conditions for proteinase K treatment. The intensity of nuclear staining using 0.5 nM of DIG labelled LNA^{TM} U6 snRNA probe was used as the readout. The ideal duration of proteinase K treatment in all cases was between 15 to 20 min by using a proteinase-K concentration of 20 µg/ml. The hybridisation temperature was set at approximately 30^0C below the melting temperature. Hybridisation temperature was set at 51^0C for U6, 53^0C for miR-21 and 57^0C for miR-26b.Probe concentrations and the stringency of washes were changed if there was no staining. The U6 probe concentration and stringency of washes were not varied for the optimisation step.

3.15 XENOGRAFTS

This was performed in collaboration with Dr Louise Coletta, Dr Nikki Ingram (Home office personal license no. 40/9070) and Dr Mihaela Lorger (Home office Personal license no. 40/10163) under home office project license number 40/3291. All procedures were carried out in accordance with the Home Office Animal Scientific Procedures Act 1986.

We used 8 athymic Ncr-nu/nu nude female mice (2 experimental groups of 4). Since we were examining the growth of MCF-7-Luc cells, exogenous oestrogen supplementation was required. This was provided in the form of 17β-estradiol pellets (0.36gm /pellet, 90-day release, Innovative research of America, Cat no. NE-121). The pellet insertion and the inoculation of cell/matrigel mix were performed by Dr Mihaela Lorger. A single pellet was inserted subcutaneously behind the ear of each mouse one week before the inoculation.

$26^{k/d}$, $con^{k/d}$ and MCF-7-Luc cells were grown in 150 cm^2 flask till they were confluent. On the day of the implant, the cells were washed with serum free DMEM. For each mouse 1 x 10^6 MCF-7-Luc cells were mixed with 3 x 10^6 fibroblasts (either $26^{k/d}$ or $con^{k/d}$) and suspended in 15 µl of Matrigel for a total volume of 30 µl. A master mix was made for each experimental group. This was maintained on ice until inoculation. For the inoculation, a small skin incision was made over the right lower mammary gland. The mammary gland was identified by gross dissection and 30 µl of the cell/Matrigel mix was injected into the mammary gland. Mice inoculated with $26^{k/d}$ and MCF-7-Luc was designated as $26^{k/d}$-mice and those inoculated with $con^{k/d}$ and MCF-7-Luc was designated as $con^{k/d}$-mice.

Since the lentiviral vector in $26^{k/d}$ and $con^{k/d}$ encoded a GFP gene reporter, their growth was monitored used fluorescence imaging. Whereas the growth of MCF-7-Luc was estimated by quantifying the Luciferase bioluminescence after injecting 100 µl (12 mg/ml) of luciferin (Promega, Cat no. E1605) into each mouse 10 min before imaging. Imaging of

fibroblasts and MCF-7-Luc cells were performed by Dr Nikki Ingram using the IVIS imaging system. I performed subsequent analysis using Living Image software (Caliper Life Sciences). The growth of fibroblasts was plotted as GFP mean fluorescence intensity after subtracting the background auto-fluorescence. The growth of MCF-7-Luc was plotted as the total luciferase flux (p/s = photons/second). This was calculated individually for each mouse and for the above days.

The mice were sacrificed on day 25. The tumour xenografts were dissected out and fixed in 4% formalin.

3.16 IMMUNOHISTOCHEMISTRY

3.16.1 FOR 3D-SPHEROIDS

The collagen gels containing collagen gel spheroids were removed from the wells and fixed in 4% formalin overnight. They were transferred to 70% alcohol and subsequently processed for paraffin embedding. Multiple 5 µm sections of the collagen gel spheroids were prepared on to Superfrost Plus slides (Menzel-Glaser, Braunschweig, Germany) by the Histology core facility in LIMM. Every 5th section from each spheroid was dewaxed in xylene and rehydrated with graded ethanol. They were then stained in Mayer's haematoxylin, Scott's tap water substitute, and eosin. Following this, sections were dehydrated in ethanol and xylene, and mounted in DPX. They were examined to identify the level at which the spheroids has the greatest diameter. The sections before and after this level were used to perform immunohistochemistry using rabbit anti-cytokeratin (Abcam, Cat no. Ab9377), a broad-spectrum anti-cytokeratin antibody. This was performed by the Histology core facility. The antibody was diluted at 1:100 in Zymed antibody diluent (Invitrogen; Paisley, UK). Antigen retrieval was performed using Access

Revelation Solution in the Access retrieval unit (Menarini; Florence, Italy) and staining using an automated stainer (IntelliPATH running a standard X-Cell plus protocol). Slides were counterstained in haematoxylin, mounted in DPX and images taken using the E1000 Eclipse microscope (Nikon; Tokyo, Japan).

3.16.2 IMMUNOHISTOCHEMISTRY FOR XENOGRAFT BLOCKS

5 µm FFPE sections were dewaxed in xylene and rehydrated with graded ethanol. This was followed by a 5 minutes wash with water. Endogenous peroxidase activity was blocked using 0.3% H_2O_2 (10min). Avidin/biotin blocking was performed using Avidin/biotin kit (Vector cat number SP-2001). 100 µl of vimentin antibody (Dako, Cat no. M0725), diluted in antibody diluent solution (Invitrogen, USA, Cat no. 00-3218; dilution 1:500), was added to each section and incubated for 30 min at room temperature. Slides were then rinsed twice with TBS-tween-20 (TBS-T), followed by a single wash with TBS. Each wash lasted 5 minutes. Slides were incubated in anti-mouse horseradish peroxidase (HRP) conjugated polymer secondary antibody (Dako, Cat No. K4001) for 30 minutes. The slides were then re-washed with TBS-T and TBS as before, and then 100ul of diaminobenzidine (DAB) was added to each slide for exactly 10 minutes, followed by a wash in water for 5 minutes. The slides were stained with haematoxylin for one minute, followed by a one minute wash in water. This was followed by a wash in Scott's tap water for one minute, and a subsequent one minute wash in water. The slides were dehydrated with graded ethanol and xylene, and mounted in DPX.

3.17 *IN-SILICO* ANALYSIS FOR MIRNA TARGETS

In-silico analysis of miR-26b targets was performed using publically available web based algorithms. The common algorithms include miRanda, miRBase targets, PicTar, PITA and TargetScan **(Baek et al., 2008)**. These algorithms use a number of different parameters to identify the targets. These parameters vary between the algorithms and therefore the lists of potential targets identified by each individual algorithm were likely to be different. I limited myself to using TargetScan and PicTar, since these have been shown have fewer false positives when compared to others at predicting the true targets **(Baek et al., 2008)**.

For each target that is identified, TargetScan allocates a rank based on the probability of it being a true target. This is reported as a context score. The context score is a value determined by a mathematical model that takes into consideration the site-type, local AU content, position of seed region, target site abundance and seed-pairing stability **(Grimson et al., 2007)**. The lower the context score the more likely it is to be target. PicTar initially identifies targets that have a seed match within their 3'UTR. Next it calculates the free energies of RNA:RNA duplexes, which it incorporates into the algorithm. Targets that are conserved across multiple species are considered more favourably. In addition PicTar also takes into consideration that multiple miRNAs (a few of which are possibly co-expressed) target a single mRNA. The probability of a target being a true target is reported as a PicTar score, similarly to the context score for TargetScan. However, the higher the PicTar score the greater the likelihood of it being a true target **(Krek et al., 2005)**.

3.18 LABEL-FREE QUANTIFICATION USING MASS SPECTROMETRY

75 cm^2 cell culture flasks were plated in triplicate with either $26b^{k/d}$ or $con^{k/d}$ and grown in DMEM medium containing 10% FCS, until they were 70% confluent. The medium was removed and the cells were washed twice with PBS. 2 ml of cell detachment buffer (Gibco, Cat no. Cat 13151014) was added and the cells were incubated at 37^0C for 5 min. The cells were transferred to a universal and washed again with PBS. Cell pellets were dried by pipetting out any residual fluid. They were then stored at -80^0C.

Label-free quantification using mass spectrometry was performed in collaboration with Dr Alexander Zougman of the Clinical and Biomedical Proteomics Group (LIMM). The samples were further processed to obtain individual peptides by digestion with trypsin. These were then separated by capillary liquid chromatography (LC) using RSLCnano system (Dionex/Thermo Scientific; Waltham, MA, USA). The liquid chromatography was connected to LTQ-Orbitrap Velos mass spectrometer (Thermo Scientific; Waltham, MA, USA), which was used for performing the tandem mass spectrometry (MS/MS). These data were compared with IPI Human 3.87 database to identify the proteins and label-free quantification (LFQ) was performed using MaxQuant 1.2.2.5. The final data that were provided to me contained the following parameters: Protein ID, protein name, gene name, number of peptides and number of unique peptides detected for each protein, and an intensity values indicating the amount of protein in each sample. All proteins, for which at least one unique peptides was detected, were included in subsequent analyses.

Further analysis of the data was performed under the guidance of biostatistician Dr Helene Thygesen (Statistics Group, LIMM). A median normalisation of the intensity values was performed for the individual samples. All values were then log transformed to base 2. For each protein, a t-test was performed comparing the log transformed intensity values of $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts. A 'false discovery rate' (FDR) was calculated (by Dr Helene Thygesen), since this type of multiple comparisons can generate false positive results. The FDR is the proportion of true positive results that are

likely to be false positive. Since the up-regulated proteins were likely to include both the direct miR-26b targets and those that are due to downstream effects, the MS data were compared to the list of potential targets generated by in-silico analyses. I used the list generated by TargetScan since its algorithm appeared to identify potential targets more comprehensively.

3.19 INTERROGATION OF PUBLICLY AVAILABLE DATABASES

Large numbers of datasets of gene expression profiles of breast cancer tissues compared to relevant controls are available online through various website. However, most of the data currently available are either from whole tumour tissue or the epithelial compartment. There is only a single study, by Finak et al. that specifically examined the gene expression profile of the stromal compartment **(Finak et al., 2008)**. Finak et al. laser capture microdissected the stromal compartment of 53 primary breast tumour and 31 matched normal stroma. They examined the gene expression profiles using Agilent's microarray platform (Agilent-012391 Whole Human Genome OligoMicroarray G4112A) and correlated this to the clinical outcomes, local or distant recurrence. Agilent's microarray platform allows measurement of a total of 19,189 transcripts. The data mining was performed using a web based bioinformatics tools "Oncomine" (https://www.oncomine.org/) **(Rhodes et al., 2007)**. This tool allows the user to access gene expression values of individual genes from various publically available datasets relatively conveniently. For Finak's dataset, expressions were available as the log2 median normalised values.

I obtained the gene expression values for COL12A1, TNK1BP1 and CPSF-7 for the 53 tumours. Although the study stated that matched normal stroma was available for 31 cases, gene expression data were publicly available for only 6 cases. I compared the median expression for all three genes in the tumour stroma to the normal stroma. Next I

dichotomised the gene expression values of the 53 tumours into those having low and high expression, using receiver operating characteristic analysis, and correlated this to key clinico-pathological parameters and clinical outcome. I performed a univariate and a multivariate analysis for each gene.

3.20 STATISTICAL ANALYSES

Statistical analysis was performed using Prism Graphpad (Version 6) or SPSS version (Version 20). Unless otherwise stated the means were compared using student's t test and error bars represent the standard error of mean.

4 MIR-26B IS DOWN-REGULATED IN BREAST CANCER CAFS

4.1 ABSTRACT

In order to identify microRNAs that may act in CAFs to modify breast cancer epithelial cell behaviour, I initially examined the expression of miRNAs from CAFs from a single archival breast cancer tissue and a tissue culture model. Firstly, I micro-dissected NFs and CAFs from a single breast cancer FFPE sample (grade-2, invasive-ductal-carcinoma, ER+ and PR+ and Her-2-) and determined expression levels of 723 individual microRNAs using microarrays. Secondly, I examined fibroblast microRNA expression in immortalised breast fibroblasts co-cultured with a breast carcinoma cell line, thereby representative of CAFs, and in immortalised breast fibroblasts co-cultured with a normal breast epithelial cell lines, thereby representative of NFs. In each case, expression in the CAFs was compared to the NFs and differentially expressed microRNAs were identified. 125 miRNAs showed consistent differences in expression between NFs and CAFs in both models. Of these 6 were selected for further analyses based on the magnitude of differential expression (>10 fold). Differential expression was examined in a further 14 cases of microdissected matched NFs and CAFs from clinical FFPE samples and 4 cases of matched primary NFs and CAFs using qRT-PCR. miR-26b was significantly down-regulated in CAFs as compared to NFs.

4.2 LASER CAPTURE MICRODISSECTION ALLOWS DISSECTION OF DIFFERENT TISSUE COMPONENTS

In order to examine the expression of miRNAs within the stroma of breast cancers I aimed to separate and examine independently the epithelial and stromal components of both normal breast and breast cancer tissue. I used laser microdissection to isolate these components from archival tissues of breast cancer resections. Initially, a single case of breast cancer was selected. The chosen case was representative of the most common histopathological subtype of breast cancer (infiltrating ductal carcinoma - NST, grade 2, ER +, PR +, her2 -, lymph node -). The epithelial and stromal compartments were microdissected from tissue sections and RNA was extracted. The areas selected for the stroma were highly enriched in fibroblasts, as areas with prevalent vascularity (endothelial cells), inflammation (lymphocytes) and fat (adipocytes) were specifically excluded; therefore this compartment has been termed fibroblasts below. RNA was also extracted from whole sections containing mainly epithelial cells and fibroblasts. Expression levels of a total of 723 miRNAs were examined using hybridisation microarrays. For each sample, the initial plan was to perform technical replicates - I found this was not feasible since the amount of RNA obtained from laser capture microdissection was limited. Therefore for each sample, only a single array was used.

In order to verify that laser capture microdissection had allowed separation of different compartments, miRNA expression profiles of the epithelial compartments and the fibroblast compartments were compared. This analysis was performed for both normal and tumour tissue. Of the 723 miRNAs that were analysed, I detected 124 miRNAs that were present in both the epithelium and the fibroblasts obtained from normal tissue. 3 miRNAs were detected only in epithelium and 86 were detected only in fibroblasts (**Figure 4.1A**). Similarly, 87 miRNAs were detected in both tumour epithelium and tumour fibroblasts while 7 miRNAs were unique to the tumour epithelium and 68 to the

tumour fibroblasts (**Figure 4.1B**). These differences confirmed that microdissection had purified different populations of cells.

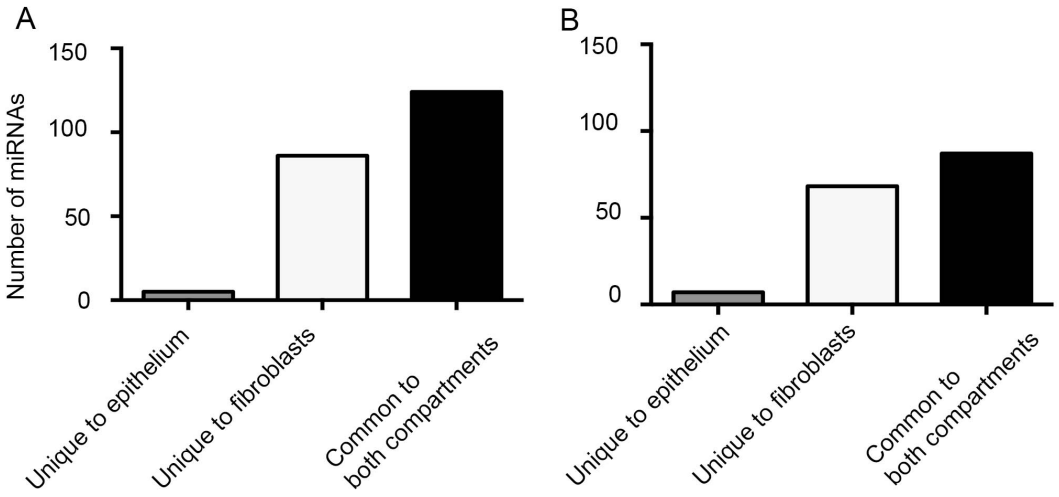

Figure 4.1. Laser microdissection allows identification of different miRNA expression profiles in epithelial and fibroblast component. Epithelium and fibroblasts were isolated by laser microdissection from normal breast tissue (A) and breast cancer tissue (B) from one breast cancer resection sample. Expression of 723 microRNA was examined using microarrays. The numbers of microRNAs detected in only the epithelial compartments, in only the fibroblast compartments, or in both compartments are shown.

I then compared the differences in miRNA expression that were determined between normal and tumour tissues when testing either the whole breast tissue sample, the purified epithelial compartment, or the fibroblast compartment. 68 miRNAs were

detected as being down-regulated in tumour in all three samples (**Figure 4.2A**), indicating some commonality between the changes occurring in the two separate compartments. Surprisingly, 8 miRNAs were found to be down-regulated in both the tumour epithelium and CAFs as compared to their normal equivalents, but not in the whole tissue samples; the changes in the isolated compartments were presumably masked in the context of the whole tissue by expression in other cell types, and this observation underlines the value of studying the different components in isolation. 52 miRNAs were detected as being down-regulated in tumour in whole tissue and in fibroblasts, but not in the epithelial component, suggesting that a substantial proportion of the changes seen in whole tissue are, in fact, limited to the stroma. Relatively few miRNAs (6) were found to be down-regulated in whole tissues and in epithelia, but not in fibroblasts, suggesting that analyses of whole tissues is fairly poor at detecting epithelial-specific changes. Similar findings were made for miRNAs that were up-regulated in tumours (**Figure 4.2B)**, with a similar bias for many of the changes seen in whole tissues being within the fibroblasts, rather than the epithelial components.

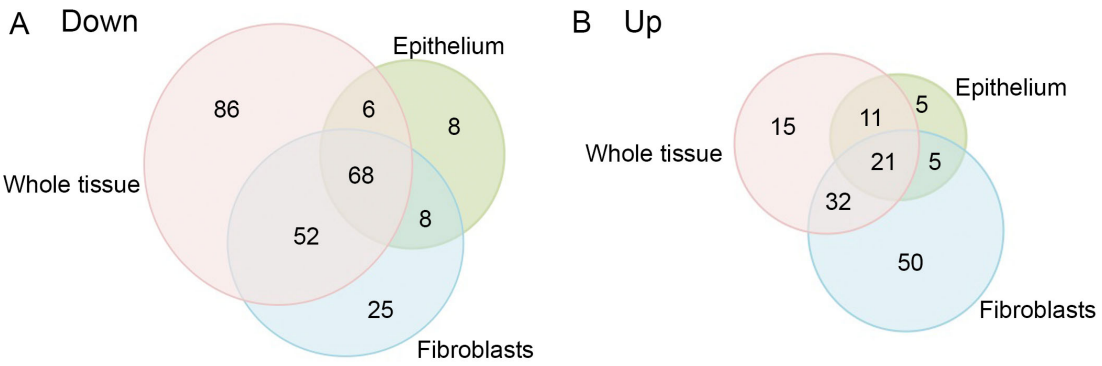

Figure 4.2. Deregulation of miRNAs as assessed using whole section differs from deregulations detected in laser microdissected epithelial or fibroblast compartments. miRNA expression of whole tumour tissue, tumour epithelium and CAFs was compared to whole normal breast tissue, normal epithelium and

normal fibroblasts respectively. The number of miRNA that were either down-regulated (A) or up-regulated (B) in whole tumour tissue, tumour epithelium and CAFs are shown. The values in the different areas represent the number of miRNA that were either down-regulated or up-regulated in the various subsets.

It is clear that deregulation of miRNAs as assessed using whole tissues is different from that determined in the epithelial or fibroblast compartments. These differences in expression between the samples representative of each compartment also demonstrate that microdissection has allowed separation of different populations of cells. However, a drawback of this type of analysis is that it did not allow me to confirm if the population of cells that I obtained by laser capture microdissection was pure. A more robust experiment to verify the purity of microdissected samples would have been to examine and quantify the mRNA expression of epithelial (e.g. Epithelial Membrane Antigen or Ber-EP4), mesenchymal (e.g. Vimentin or Smooth Muscle Antigen), endothelial (e.g. CD31 or CD34) and lymphocyte (e.g. CD45) markers in these samples. Unfortunately, again this was difficult because of the limited amount of sample obtained after microdissection and the quality of the mRNA obtained.

4.3 MICRORNA EXPRESSION IN NORMAL FIBROBLASTS IS DIFFERENT FROM CARCINOMA-ASSOCIATED FIBROBLASTS

Next, I wanted to examine how the miRNA expression profiles of carcinoma-associated fibroblasts (CAFs) differed from those of the normal fibroblasts (NFs). MiRNAs that were detected in NFs but not in CAFs were presumed to be down-regulated. Similarly miRNAs that were detected in CAFs but not in NFs were presumed to be up-regulated. If the miRNA was present in both NFs and CAFs a fold change was calculated. Only miRNAs that had fold change of greater than 2 were regarded as being deregulated.

Using this strategy 109 miRNAs were determined to be expressed at higher levels in CAFs compared to NFs, and 153 miRNAs were expressed at lower levels in CAFs compared to NFs. These are listed in the Appendix (**Table A1**).

4.4 A CO-CULTURE SYSTEM TO MODEL CAF BIOLOGY

Since a very large number of miRNAs were identified from the clinical samples as being differentially expressed between NFs and CAFs, an additional screening method was required to determine which miRNAs were likely to be the most functionally relevant and therefore those most suitable for further study.

In order to investigate the functional influence of fibroblasts on the behaviour of breast epithelial cells, colleagues within the Breast Research Group established a tissue co-culture system and examined the effect on growth of MCF-7 cells and HB2 cells when they were co-cultured with CAFs. Firstly, breast epithelial cells were stably-transfected with plasmids allowing overexpression of green fluorescent protein (GFP). Two alternative epithelial cell types were used: MCF-7 cells, representing breast cancer cells, or HB2 cells, representing non-transformed "normal" breast epithelial cells. Secondly, primary breast CAFs were immortalised by retroviral-transduction with human

telomerase **(Verghese et al., 2011)**. The immortalised breast fibroblasts and the GFP-labelled epithelial cells were cultured either alone or in the context of an epithelial: fibroblast co-culture (ratio of epithelial cells: fibroblasts 1:4). Proliferation of epithelial cells was monitored by counting using FACS on the basis of the presence of GFP. The immortalised fibroblasts had a potent and differential effect on the growth of the two breast epithelial cell types; growth of the MCF-7 breast cancer cell line was stimulated by the presence of fibroblasts, while growth of the benign breast cell line HB2 was not. (**Figure 4.4**). It is clear that the fibroblasts had a differential influence on the cancer epithelial cells (MCF- 7 cells, **Figure 4.4A**), but not the benign epithelial cells (HB2, **Figure 4.4B**), being stimulated by the fibroblasts. We concluded that the MCF-7: fibroblast interaction represented a functional interaction analogous of the interaction between breast epithelial cancer cells and CAFs, while HB2 and CAFs exhibited less functional cross-talk **(Orimo et al., 2005)**. It was clear that this co-culture model could potentially allow study of the changes that occur in fibroblasts as they functionally interact with tumour cells. Caroline Green (PhD student supervised by Prof Valerie Speirs) performed the above experiments.

I aimed to use this MCF-7: fibroblast co-culture as a model of functional CAFs. On the other hand, fibroblasts grown with HB2 cells did not show this functional ability to stimulate epithelial cell growth. I aimed to use this as a model of NFs. The difference in miRNA expression in fibroblasts in these two situations may provide a suitable second method of identifying miRNAs of functional importance in CAFs. Therefore, I used the *in-vitro* tissue culture to model epithelial: fibroblast interactions and to provide an alternative system for analysis of miRNA expression.

Fibroblasts were grown with or without GFP labelled epithelial cells (either HB2 cells or MCF-7 cells) for 9 days. The two different cell types in the co-cultures were separated using flow-cytometry. RNA was extracted from the fibroblasts and expression of miRNAs was examined using hybridisation arrays as above.

MiRNA expression in fibroblasts that were grown with MCF-7 cells (representing CAFs) was compared to that in fibroblasts that were grown with HB2 cells (representing NFs). 60 miRNAs were more highly expressed in the ‘CAF' model as compared to the ‘NFs', while 166 miRNAs were expressed at lower levels in the ‘CAF' model. These are listed in the Appendix **(Table A2).**

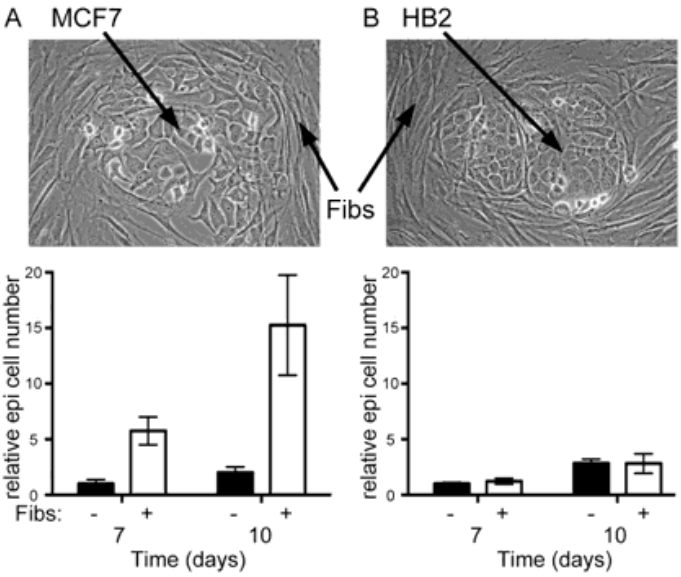

Figure 4.4. Growth of cancer cells, but not benign HB2 cells, is stimulated by fibroblasts. A and B) Phase-contrast images of co-cultures of fibroblasts with (A) or HB2 (B) cells. Fibroblast (Fibs) and epithelial cells (MCF-7 or HB2) are labelled. Equal starting numbers of either MCF-7 (left) or HB2 (right) cells were seeded alone (black bars), or in co-culture with fibroblasts (white bars), and epithelial cell number was measured after 7 or 10 days. Data points represent mean of triplicates. [Data provided by Prof Valerie Speirs, Breast Research Group, University of Leeds]

4.5 A SUBSET OF MIRNAS HAVE SIMILAR DEREGULATIONS IN-VIVO AND IN-VITRO

In order to identify the miRNAs that were consistently differentially expressed between NFs and CAFs, I compared data from the two experimental systems. MiRNAs that were differentially expressed in the same direction in both systems would represent the most promising candidates for functionally important miRNAs. There were 105 miRNAs that were down-regulated and 10 miRNAs that were up-regulated consistently in both systems (Appendix, **Table A3**). In order to reduce the number of candidate miRNAs to a number that could practically be analysed further, I applied an arbitrary cut-off of miRNAs with fold changes greater than 10 in both systems. Six microRNAs were selected using these criteria (**Table 4.1**).

miRNA	Regulation	Fold change (tissues)	Fold change (co-culture)	Geomean fold change
miR-26b	Down	21	49.3	32.2
miR-7f	Down	47.6	10.2	22
miR-30b	Down	12.3	30.2	19.3
let-7g	Down	15	11.4	13
miR-107	Down	13.1	11.5	12.3
miR-15b	Down	14.2	10.8	12.4

Table 4.1. Six miRNAs are consistently down-regulated more than 10 fold in CAFs as compared to NFs in both tissue and co-culture model. 'Geomean fold change' is the geomean of the fold change in tissues and the co-culture model.

4.6 VALIDATION OF MIRNAS USING QRT-PCR USING FURTHER CLINICAL SAMPLES

4.6.1 SCREENING STRATEGY FOR CANDIDATE MIRNAS

The above experiments allowed me to identify miRNAs for which expression differed similarly in the two experimental systems. Next, I wanted to examine whether this deregulation was consistent in more clinical samples. Therefore, I decided to laser capture microdissect the NFs and CAFs from further clinical cases that had a similar profile as the first case, extract total RNA, and examine the expression of candidate miRNAs using qPCR. For this, I initially selected 40 cases of breast cancer (infiltrating ductal carcinoma – NST grade 2, ER +, PR +, her2 -, lymph node -) from the Leeds Teaching Hospital pathology tissue archives. The H&E stained sections of these cases were examined. This allowed me to identify cases where it was feasible to microdissect a relatively pure population of fibroblasts. Of the 40 cases that I initially screened, I selected 25 cases for further microdissection.

One difficulty that I encountered was that the amounts of RNA obtained from laser capture microdissection for downstream work was variable on a case-by-case basis. Of the twenty five cases that I microdissected, I was able to extract sufficient RNA for further analysis from only fourteen cases. However, this gave me the samples required to detect a standardised difference of at least 1, 95% power and $p<0.05$. This sample size was obtained from the normogram produced by Altman **(Whitley and Ball, 2002)**.

Another difficulty that I encountered was choosing an appropriate gene for normalisation. There was conflicting evidence in the literature regarding the appropriate reference for miRNAs and therefore I decided to use two separate reference genes. RNU48 and U6 **(Gee et al., 2011)**. I first decided to determine if the levels of these two

potential reference genes were correlated in the samples. There was a strong positive correlation between the reference genes in the 28 RNA samples (14 cases of NF and 14 cases of CAF), indicating that potentially either could be used to normalise miRNA data without strongly influencing the conclusions (**Figure 4.4**).

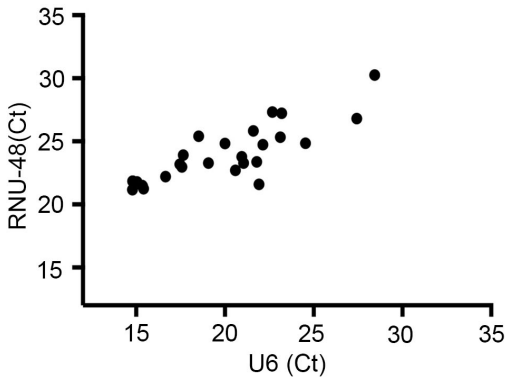

Figure 4.4. A strong positive correlation was seen between qPCR CT values for U6 and RNU48 in 28 clinical samples. The values on X-axis and Y-axis represent mean CT values of triplicate wells for U6 and RNU48 respectively. The Spearman's coefficient of correlation was 0.83 (P<0.0001).

4.6.2 MIR-26B IS CONSISTENTLY DOWN-REGULATED IN CAFS

I decided to prioritise examination of miR-26b, as this had shown the greatest consistent fold change (>20) in the two experimental systems tested using microarrays (**Table 6.4.1**). In addition, I decided to examine another two miRNA, one randomly selected from the remaining 5 miRNAs with fold change greater than 10 (miR-30b) and another randomly selected from the literature as being important in cancer biology (miR-10a). Expression levels of miR-26b, miR-30b and miR-10a were determined, and data were

initially normalised to U6 or RNU48. The geomean (see appendix for formula) of the relative expression was subsequently calculated for further analysis.

Eleven of fourteen cases showed reduced expression of miR-26b in CAFs compared to NFs (**Figure 4.5A**). The median fold change was 2.3 (IQR= 1.6, 5.7). Eleven of fourteen cases showed reduced expression of miR-30b in CAFs as compared to NFs (**Figure 4.5B**). The median fold change was 1.6. Ten of fourteen cases showed reduced expression of miR-10a in CAFs compared to NFs (**Figure 4.5C**). The median fold change was 1.5. However, only miR-26b showed a significant difference in the expression levels between NFs and CAFs (p=0.04, Wilcoxon matched-pairs signed rank test) (**Figure 4.5C**). The difference between NFs and CAFs for miR-30b and miR-10a were not statistically significant (p=0.06 and p=0.07 respectively, Wilcoxon signed rank test) (**Figure 4.5E & 4.5F**).

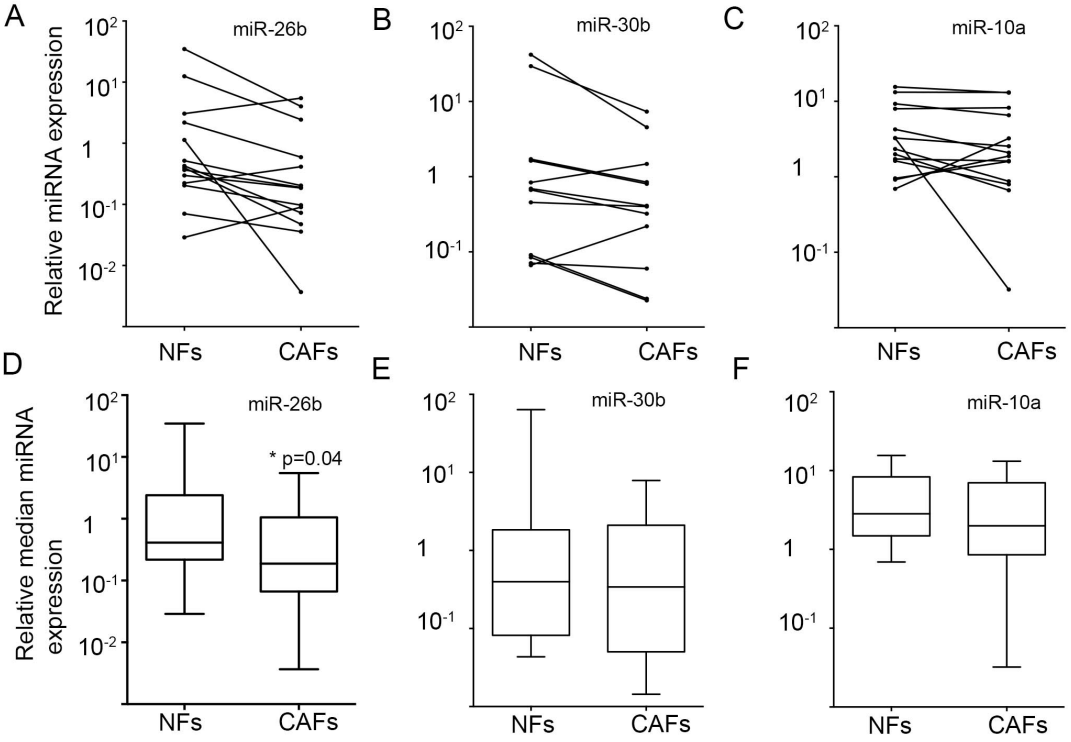

Figure 4.5. Relative expression of miR-26b, 30b and miR-10a in NFs and matched CAFs in 14 breast cancer cases using the geomean of expression levels of U6 and RNU48 as the reference value. 11 of 14 cases showed down-regulation of miR-26b (A), 11 of 14 cases showed down-regulation of miR-30b (B) and 10 of 14 cases showed down-regulation of miR-10a (C) in CAFs compared to NFs. Each dot represents the mean of experimental triplicates of an individual sample. The two dots that are joined by the straight line represent a single case. The median expression of miR-26b (D) was significantly reduced in CAFs compared to NFs (Wilcoxon signed rank test, p=0.04). There was no significant difference in the median expression of miR-30b (E) or miR-10a (F) between CAFs and NFs (Wilcoxon signed rank test, p=0.06, p=0.07 respectively).

4.7 VALIDATION OF MIR-26B DOWN-REGULATION IN CAFS

Although laser capture microdissection allowed me to enrich for fibroblasts and confirm the down-regulation of miR-26b in CAFs compared to NFs in a large number of samples, contamination from other cell types remained a theoretical possibility. Therefore to validate these findings, I decided to

1. Confirm the differential expression of miR-26b in primary fibroblasts AND
2. Confirm the differential expression of miR-26b in FFPE clinical samples using In-situ-hybridization (ISH).

4.7.1 MIR-26B IS DOWN-REGULATED IN PRIMARY CAFS

NFs and CAFs are isolated from fresh tissue and grown in tissue cell culture routinely by the Breast Research Group in LIMM. Examining the expression level of miR-26b in these primary fibroblasts provided another method to confirm the down-regulation of miR-26b in CAFs.

RNA from four cases of matched primary NFs and CAFs were kindly donated by Prof Valerie Speirs. MiR-26b expression in these RNA samples was determined using qRT-PCR. Once again, U6 and RNU48 were used as reference genes.

In all four cases miR-26b was down-regulated in CAFs compared to matched NFs (**Figure 4.6A**). However, the difference in degree of down-regulation was not significant (median fold change=2.3, p=0.125, Wilcoxon matched-pairs signed rank test) (**Figure 4.6B**).

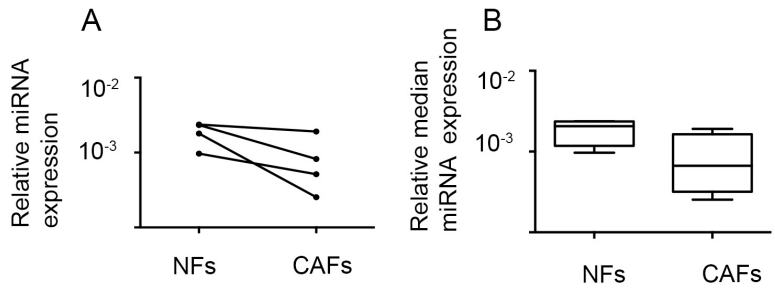

Figure 4.6. Relative expression of miR-26b in primary NFs and matched CAFs in 4 breast cancer cases using the geomean of expression levels of U6 and RNU48 as the reference value. All 4 cases showed down-regulation of miR-26b. Each dot represents the mean of experimental triplicates of an individual sample. The two dots that are joined by the straight line represent a single case. There was no difference in the median expression of miR-26b between CAFs and NFs (Wilcoxon signed rank test, p=0.125).

4.7.2 MIR-26B EXPRESSION CANNOT BE DETECTED BY ISH

ISH is a type of hybridisation technique that uses a labelled complementary nucleotide probe to determine the presence of specific DNA or RNA molecules within the tissue. Potentially this would allow the visualisation and localisation of miR-26b expression within the fibroblasts. I examined the expression of miR-26b in the same FFPE clinical samples that were used to perform the laser capture microdissection. For optimisation, I initially chose cases that had high levels of miR-26b expression as determined by qRT-PCR.

4.7.2.1 EXPRESSION OF MIR-26B

Expression of miR-26b was examined with miR-26b probe. ISH probes for U6 and miR-21 were used as positive controls and a scrambled miRNA probe was used as a negative control.

ISH using U6 probes showed strong nuclear staining in normal epithelial cells and normal fibroblasts (**Figure 4.7A**) and cancer epithelial cells and CAFs (**Figure 4.7C**). ISH using miR-21 probe showed strong staining, albeit only in CAFs (**Figure 4.7D**). There was no staining within the cancer epithelial cells (**Figure 4.7D**), normal epithelial cells or normal fibroblasts (**Figure 4.7B**). ISH using the scrambled miRNA probe did not show any staining.

However there was no staining using the probe for miR-26b. A number of different hybridisation conditions were tried but no staining was seen. It is likely that this could be due to the poor sensitivity of ISH compared to qRT-PCR and expression array.

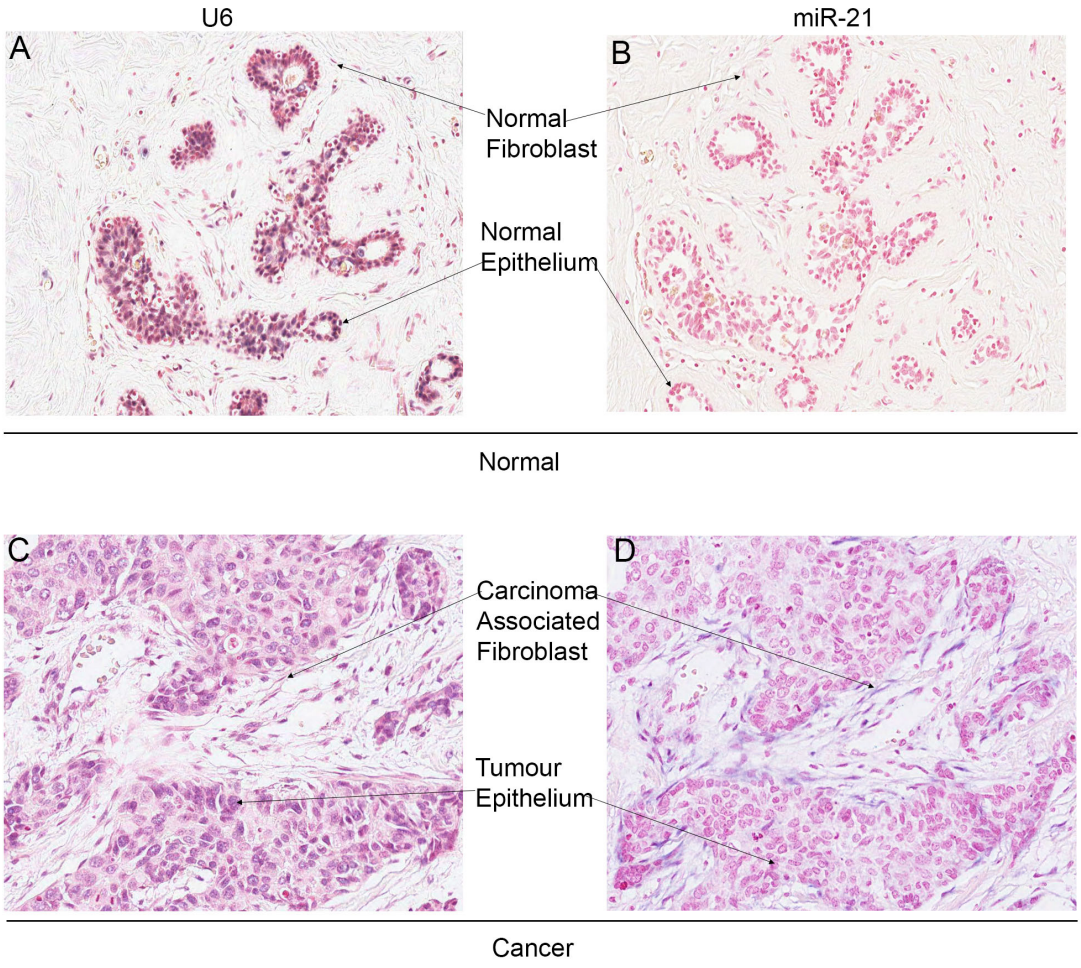

Figure 4.7. Representative examples of in-situ hybridisation show that miR-21 is expressed in CAFs rather than cancer epithelial cells. Breast cancer tissues were stained with nuclear fast red (a counter stain that labels the nucleus red) and ISH probe (blue) for the specific RNA. U6 was used as a positive control. U6 shows strong nuclear staining in both the epithelium and fibroblasts in normal (A) and breast cancer tissue (C). The miR-21 probe shows strong cytoplasmic staining in the CAFs but no staining in cancer epithelium (D), normal epithelium or normal fibroblasts (B). (*Image link: http://slides.virtualpathology.leeds.ac.uk/Research_4/Breast_Group/Eldo/2012-03-22/view.apml*)

4.7.3 MIR-21 IS FREQUENTLY UP-REGULATED IN CAFS

The predominant stromal/fibroblast expression of miR-21 in breast cancer tissue was a surprise. miR-21 has been reported to be up-regulated in breast cancer in a number of studies **(Iorio et al., 2005; Qi et al., 2009)**. However, this up-regulation was assumed to be within the epithelial compartment. Subsequently, a number of functional studies have been performed where investigators have manipulated the expression of miR-21 in epithelial cells to better understand epithelial cell biology **(Si et al., 2007; Yan et al., 2011)**. However, in a few of the recent studies there have been suggestions that miR-21 might be present within the stromal compartment **(Rask et al., 2011)**.

Therefore to investigate the unexpected localisation of miR-21 that was seen using ISH in more detail, I examined the expression of miR-21 using qRT-PCR in the microdissected samples that were obtained from my initial cohort of 14 samples.

Ten of the thirteen cases showed up-regulation of miR-21 in CAFs compared to NFs (**Figure 4.8A**). There was a significantly higher expression of miR-21 in CAFs compared to NFs (median fold change=8; Wilcoxon signed rank test, p=0.03; **Figure 4.8B**).

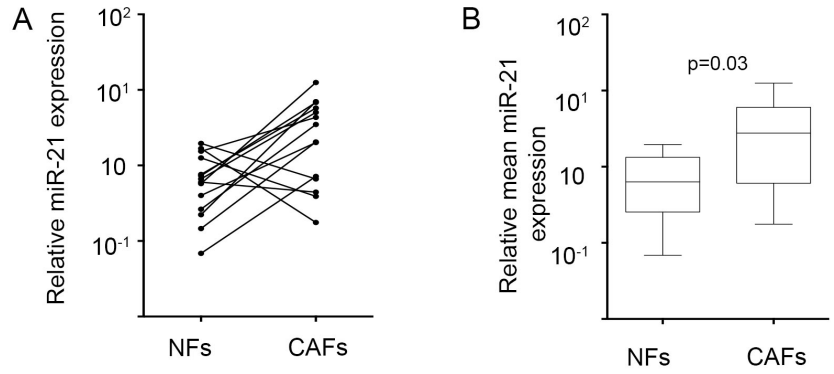

Figure 4.8. Relative expression of miR-21 in NFs and matched CAFs in 13 breast cancer cases using the geomean of expression levels U6 and RNU48 as the reference value. 10 of 13 cases showed up-regulation of miR-21 (A). Each dot represents the mean of experimental triplicates of an individual sample. The two dots that are joined by the straight line represent a single case. The median expression of miR-21 (B)) was significantly increased in CAFs compared to NFs (Wilcoxon signed rank test, p=0.03).

To confirm the up-regulation of miR-21 in fibroblasts, I performed ISH hybridisation on a further three cases that had high levels of miR-21 expression using qRT-PCR. This confirmed that the high miR-21 expression was in the fibroblasts present within the stromal compartment of tumour. Expression of miR-21 was absent in the cancer cells.

Interestingly, when I examined the dataset obtained from the miRNA microarrays of the initial case that was microdissected for the expression levels of miR-21, it was seen to be down-regulated in the cancer as compared to the normal tissue in both the epithelial and the fibroblast compartments. This highlights the limitation of using only a single case for the initial screen.

During this period, further evidence emerged that confirmed the overexpression of miR-21 in fibroblasts in breast cancer, also as assessed by ISH **(Rask et al., 2011)**. Therefore, based on my data and the published literature, our group decided to

investigate the role of miR-21 in fibroblasts; this work is currently being undertaken in our laboratory with funding from Cancer Research UK.

4.8 DISCUSSION

The initial screening strategy to identify miRNAs that were deregulated in CAFs helped me to identify miR-26b as one of the miRNAs that was down-regulated. Although, the use of only a single archival case has been described, I had initially planned to laser capture microdissect CAFs from at least 3 different cases. However, I found this to be extremely difficult since the laser capture microdissection was taking much longer than expected. This was mainly attributable to a lack of available expertise. However, the two-pronged approach of using the single clinical sample and combining it with a co-culture model allowed me to partly overcome these shortcomings. This is demonstrated by the subsequent data that showed miR-26b, identified using this screening strategy, was significantly down-regulated in further clinical samples.

I attempted to validate the down-regulation of miR-26b in primary fibroblasts and confirm the differential expression of miR-26b in FFPE clinical samples using in-situ-hybridization (ISH). I was able to obtain 4 matched samples of primary NFs and CAFs to assess the expression of miR-26b. Although, the down-regulation of miR-26b did not reach statistical significance, it was seen to be down-regulated in all 4 samples. The lack of statistical significance is most likely due to the small number of samples. Unfortunately, I was unable at the time to obtain further matched primary samples to expand the series. There appears to be variation in the degree of down-regulation of miR-26b in the primary CAFs compared to its matched NFs. This was similar to that seen in the FFPE samples. I attempted to examine if the there was any relation with the passage number of the primary fibroblasts and the degree of down-regulation. Unfortunately, I was unable to obtain this information.

ISH was successful in detecting miR-21. However, no signal could be obtained for miR-26b. Although, I reasoned that this could be due to poor sensitivity of ISH, another

possibility might be a defect in the miR-26b probe. Although this is unlikely, since these probes have been validated in the literature **(Wienholds et al., 2005)**, I attempted to validate the miR-26b probes in our lab using a positive control. Unfortunately, obtaining a positive control was difficult. The miR-26b probes had been validated using embryonic tissue of zebrafish. I did not have access to this tissue. Therefore, as an alternative I made cell blocks containing fibroblasts that had high levels of miR-26b by transfecting them with miRNA-26b precursor molecules (see section 5.2). These were then formalin fixed and paraffin embedded and ISH was performed. However, ISH using these were not successful since the cells did not adhere to the slides.

5 FUNCTIONAL ROLES OF MIR-26B IN BREAST FIBROBLASTS AND, INDIRECTLY ON EPITHELIAL CANCER CELLS

5.1 ABSTRACT

In this chapter, I have studied functional roles of miR-26b in breast CAFs and the potential influences of CAFs on breast cancer epithelial cells.

Initially, I examined roles of miR-26b in defining the behaviour of the fibroblasts themselves, by either overexpressing or reducing miR-26b levels. Overexpression of miR-26b in breast fibroblasts caused a decrease in proliferation and increase in apoptosis. These results were consistent with the results of most published studies in which investigators overexpressed miR-26b in other cell lines. However, as the degree of miR-26b overexpression in these experiments was far greater than that seen in clinical samples, these results may lack clinical relevance. Reducing miR-26b expression in breast fibroblasts caused a decrease in proliferation and a possible G2/M cell cycle defect, albeit only to a minor degree. However, there was a significant increase in the migration and invasion of these fibroblasts (designated '$26b^{k/d}$') in comparison to control fibroblasts (designated '$con^{k/d}$') in transwell assays.

In order to examine the influence of $26b^{k/d}$ fibroblasts on breast cancer epithelial cells, I performed co-culture assays. Although, secreted factors from fibroblasts are known to cause increase in the growth of cancer epithelial cells, there was no differential influence on the growth of cancer epithelial cells when they were grown along with $26b^{k/d}$ fibroblasts or $con^{k/d}$ fibroblasts. Next, I examined the influence of $26b^{k/d}$ fibroblasts on migration and invasion of breast cancer epithelial cells. There was no difference in the migration of cancer epithelial cells in indirect co-culture transwell migration assays

but there is a significant increase in epithelial migration in direct co-culture migration assays. In addition, there was increased invasion of cancer epithelial cells in 3D spheroidal invasion assays.

In conclusion, knock down of miR-26b increased the migration of fibroblasts, which, in turn, increased the migration and invasion of cancer epithelial cells, demonstrating a potential clinically-relevant role for the reduced expression of miR-26b in CAFs observed in breast cancers.

5.2 FUNCTIONAL INFLUENCE OF MIR-26B ON FIBROBLASTS

Having determined that miR-26b was frequently down-regulated in CAFs, I aimed to examine whether this down-regulation might be associated with altered fibroblast behaviour, i.e. differences in growth, migration or invasion. To this end, I transfected immortalised breast fibroblasts (IF1 cells) either to over-express or to knockdown miR-26b, and examined a number of different functional readouts.

5.2.1 OVEREXPRESSION OF MIR-26B IN FIBROBLASTS LEADS TO REDUCED GROWTH

IF1 cells were transfected with miRNA-26b precursor molecules or with miRNA precursor negative controls. The control molecules are double stranded oligonucleotide molecules that are processed by the cellular machinery (section 1.4.1) but that have no known mRNA targets. MTT assays were performed to investigate the growth and survival of the cells after 24, 48 and 72 hours (**Figure 5.1**) Overexpression of miR-26b caused a dramatic reduction in cell numbers after 48 and 72 hours compared to the control transfected cells.

Transfection with miRNA-26b precursor molecules caused a decrease in cell numbers from 48 to 72 hours. Apoptosis assays were performed using PE Annexin-V/7-AAD staining at 24 and 48 hours to examine whether apoptosis was responsible for this reduction in cell number. The percentage of cells in early to mid apoptotic phase was quantified (**Figure 5.1 A &D**). Overexpression of miR-26b resulted in 14.8% of the cells in the early apoptotic phase compared to 6.3 % of the control cells at 48 hours (**Figure 5.1 D**). This confirmed that the decrease in cell viability was in part due to apoptosis.

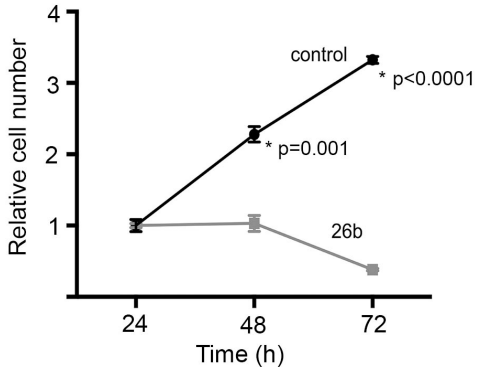

Figure 5.1. Growth/viability of immortalised breast CAFs was decreased when miR-26b was overexpressed. Immortalised fibroblasts were transfected with either pre-miR-26b (grey line) or control pre-miR (dark line). MTT assays were performed to quantify viable cells at the times shown. Data points represent the relative cell number of three biological triplicates and are representative of 3 independent experiments.

These data are compatible with a number of studies in which miR-26b has been overexpressed in various cancer types, including breast cancer (epithelial cells), nasopharyngeal carcinoma and gliomas **(Ji et al., 2010; Liu et al., 2011; Wu et al., 2011)**. These studies have variously shown both decreased proliferation and increased apoptosis.

However, I wondered whether the degree of overexpression *in-vitro* was comparable to that seen in the clinical samples. The median fold change between NFs and CAFs for miR-26b in the clinical FFPE samples was 2.3 (range 2 to 308). Therefore, I examined the fold change in miRNA-26b achieved by transfection *in-vitro* at 48 hours. This time point was chosen since the difference in cell viability was first noticed at 48 hours. The fold change was 976 after 48 hours (**Figure 5.3**). This was substantially more than the maximal degree of fold change seen in the clinical samples. This result raises the possibility that the effects seen when overexpressing miR-26 may not be physiologically relevant.

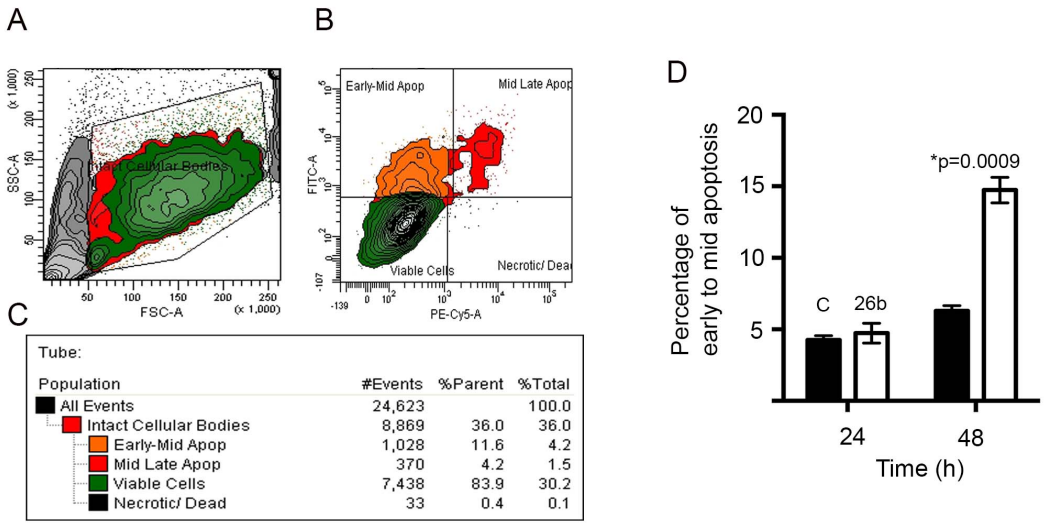

Figure 5.2 Overexpression of miR-26b increased the percentage of cells in early to mid apoptosis. Immortalised fibroblasts were transfected with pre-miR-26b (26b) or control. Apoptosis assays were performed at 24 and 48 hours with PE Annexin-V/7-AAD. Intact cells were initially gated using control transfected cells (A). This population was examined for Annexin-V and 7-AAD staining (B). Cells that were positive for Annexin-V (FITC-A) and negative for 7-AAD (PE-Cy5-A) were counted (C). Data points represents the mean percentage of cells that were positive for Annexin V and negative for 7-AAD of three biological replicates, black bars=control, open bar=miR-26b (D). This is representative of 2 independent experiments.

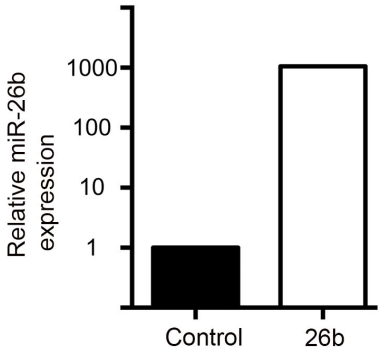

Figure 5.3. Transfection with miR-26b precursor molecules results in dramatic up-regulation of miR-26b levels. Immortalised fibroblasts were transfected with pre-miR-26b or control molecules. Expression of miR-26b relative to U6 at 48 hours is shown, as assessed by qRT-PCR. Data represents the mean of technical triplicates (error bars are too small to be visible).

5.2.2 KNOCKDOWN OF MIR-26B IN FIBROBLASTS

5.2.2.1 TRANSIENT KNOCKDOWN OF MIR-26B IN FIBROBLAST RESULTS IN INCREASED MIGRATION AND A DECREASE IN GROWTH

Next, I examined changes in breast fibroblast behaviour associated with knocking down miR-26b. For this I initially used a transient knockdown model where miR-26b function was reduced in IF1 cells. Since the degree of fold change would inform the physiological significance of the findings I initially examined the knockdown efficiency of this inhibitor molecule.

IF1 cells were transfected with Anti-miR-26b™ miRNA Inhibitors or a miRNA inhibitor negative control. The control molecule is a single-stranded sequence that does not target any miRNA and has no identifiable effects on cells. The relative expression of miR-26b miRNA was examined using qRT-PCR 48 hours after transfection. There was a 3.5 fold decrease in miR-26b levels (**Figure 5.4A**). This degree of fold change was similar to the differential expression seen in clinical samples (median fold difference, 2.3).

Next, I examined the proliferation of IF1 cells after transfection using MTT assays at 24, 48 and 72 hours. Surprisingly, the MTT assays showed that there was a small but significant decrease in the number of cells after miR-26b knockdown at 48 hours, compared to the cells with the control knockdown (**Figure 5.4B**). However this difference was not significant at 72 hours. The loss of this difference at the later time point could be due to loss of functional knockdown after 78 hours; although a trend for reduced cell number was still evident.

I also examined the migration of IF1 cells after transfection using transwell migration assays. The transfections were performed in the transwell inserts. The number of cells

that had migrated to the underside of the membrane was quantified by microscopy after 25 hours. There were a significantly greater number of IF1 cells with miR-26b knockdown that had migrated to the underside of the membrane compared to the control knockdown (**Figure 5.4C**).

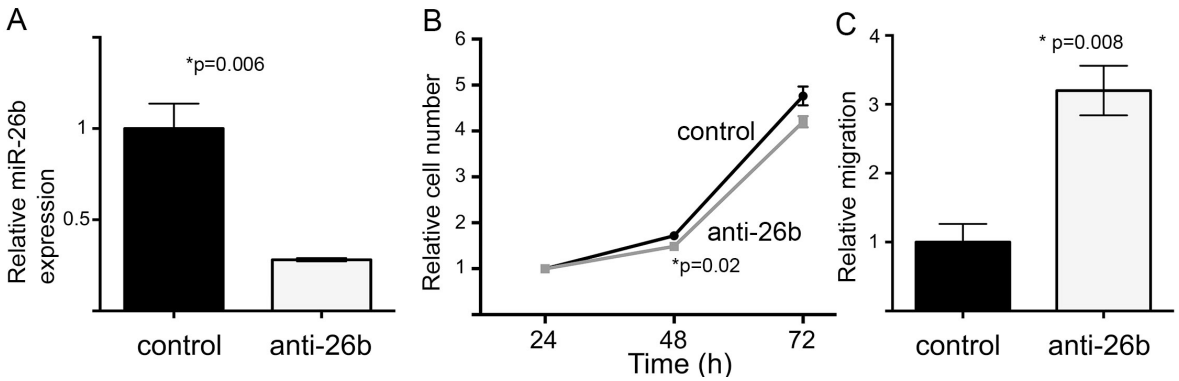

Figure 5.4. Transient down-regulation of miR-26b results in a small but significant decrease in growth and a 3 fold increase in migration. Immortalised fibroblasts were transfected with anti-26b molecules or control, and miR-26b expression relative to U6 was examined at 48 hours (A). MTT assays were performed at the time points shown (B). Transwell migration assays were performed 24 hours after transient transfection of fibroblasts (C). The data points for A are means of three technical replicates, whereas for B & C are means of biological triplicates.

5.2.2.2 STABLE KNOCKDOWN OF MIR-26B RESULTS IN INCREASED MIGRATION AND INVASION OF FIBROBLASTS

Modification of the extracellular matrix has been to shown to increase migration and invasion of malignant epithelial cells (see section 1.3). Furthermore, there is emerging evidence that increased motility of fibroblasts can influence the remodelling of the

extracellular matrix **(Shieh et al., 2011)**. To examine these potential influences requires longer term assays and co-culture studies with epithelial cells. Therefore, fibroblasts with stable knockdown of miR-26b were generated. These were generated by transducing IF1 cells with pseudoviral particles encoding anti-miR26b molecules (see section 3.12). IF1 cells with miR-26b knockdown were designated as "$26b^{k/d}$" whereas the control knockdown cells (virally transduced to express a control molecule) were designated as "$con^{k/d}$".

Again, the knockdown efficiency using this viral system was an important consideration. The knockdown efficiency was confirmed by examining the relative expression of miR-26b using qRT-PCR. This showed that there was 2 fold decrease in miR-26b level in $26b^{k/d}$ fibroblasts compared to $con^{k/d}$ fibroblasts (**Figure 5.5A**).

In addition, the degree of “functional knockdown” of miR-26b in these cell lines was also assessed. A full length complementary sequence for miR-26b (see section 3.13.1.1) was cloned into the 3’ UTR of the firefly luciferase cassette within a plasmid reporter vector. The vector also contained a control renilla luciferase expression cassette allowing constitutive expression; firefly luciferase activity was the normalised to renilla luciferase activity. $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts were transfected with the luciferase reporter plasmid and luciferase activities were measured after 24 hours. There was a 2 fold increase in the normalised luciferase activity in $26b^{k/d}$ fibroblasts compared to $con^{k/d}$ fibroblasts (**Figure 5.5B**) indicating a reduction in the inhibitory activity of miR-26b due to the presence of anti-miR-26b.

Finally, the cell viability and proliferation of $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts were assessed. Equal numbers of $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts were seeded in 24 well plates and MTT assays were performed at 24, 48 and 72 hours. This also showed a small but significant decrease in the growth of $26b^{k/d}$ fibroblasts compared to $con^{k/d}$ fibroblasts at 72 hours (Figure 5.5C). This was a similar result to that seen when miR-26b was transiently knocked down.

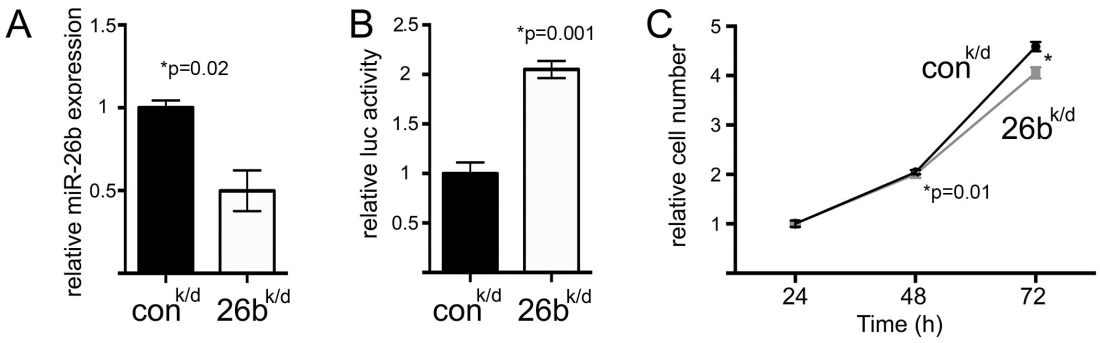

Figure 5.5. Stable knockdown of miR-26b was associated with decreased growth of immortalised fibroblasts. The level of miR-26b expression relative to U6 was examined in both con$^{k/d}$ and 26b$^{k/d}$ fibroblasts (A). con$^{k/d}$ and 26b$^{k/d}$ fibroblasts were transfected with a luciferase reporter plasmid expressing a full length complementary sequence for miR-26b in the 3'UTR of the luciferase transcript and normalised luciferase reporter activity was measured 24 hours after transfection (B). Equal numbers of con$^{k/d}$ and 26b$^{k/d}$ fibroblasts were plated and MTT assays were performed at the time points shown (C). Data points for A represent means of technical triplicates, whereas for B & C they represent the mean of biological triplicates. B & C are representative examples of at least 2 independent experiments.

It was postulated that the difference in growth could be a consequence of miR-26b's influence on the cell cycle. Therefore I performed cell cycle analysis of 26b$^{k/d}$ and con$^{k/d}$ cells using propidium iodide staining. Equal numbers of 26b$^{k/d}$ and con$^{k/d}$ cells were seeded in 6 well plates and the proportion of cells in each phase of the cell cycle were quantified after 48 hours (**Figure 5.6 A-C**). This demonstrated a small but significant decrease in the percentage of 26b$^{k/d}$ fibroblasts in G1 phase and a significant increase in the percentage of 26b$^{k/d}$ fibroblasts in the G2/M phase (**Figure 5.6D**). There was no difference in the proportion of cells in S phase. These data, in conjunction with the decrease in growth of 26b$^{k/d}$ fibroblasts, suggests a potential G2/M defect in these cells.

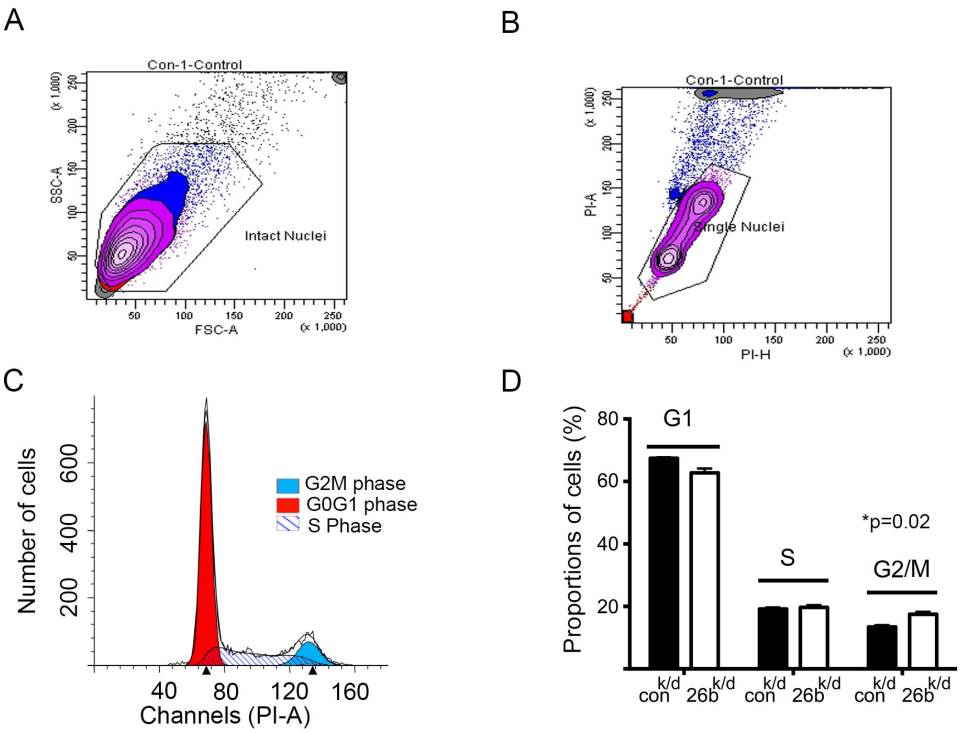

Figure 5.6. Cell cycle assays demonstrated a potential G2/M defect in 26b$^{k/d}$ fibroblasts. Equal number of 26b$^{k/d}$ and con$^{k/d}$ were seeded in 6 well plates. Cell cycle assays were performed using propidium iodide staining. Panels A and B represent flow-cytometry gating strategy: cells were selected on the basis of forward/side scatter (A); clumps and doublets were excluded from this population (B). A model was created to fit this DNA staining pattern and the number of cells in each phase of the cell cycle was estimated using this model (C). The proportion of cells in each phase of the cell cycle was quantified after 48 (D). The data points represent the mean of biological triplicates and the data are representative of 2 independent experiments.

To confirm the influence of miR-26b on motility of fibroblasts, I performed cell scratch assays and transwell migration assays. For the cell scratch assay, 26b$^{k/d}$ fibroblasts or

con$^{k/d}$ fibroblasts were grown until they were confluent and a wound was created. Images were taken at 0 and 18 hours (**Figure 5.7A**) and the area of the wound that had closed after 18 hours were quantified. There appeared to be an increase in the percentage area of wound closed by 26b$^{k/d}$ fibroblasts compared to con$^{k/d}$ fibroblasts (**Figure 5.7B**). However this difference fell short of being significant.

Next transwell migration assays were performed to assess the degree of migration at different time points. 26b$^{k/d}$ fibroblasts or con$^{k/d}$ fibroblasts were plated onto the inserts and the number of cells that migrated to the underside of the membranes at 12 hours was quantified by counting (**Figure 5.7C**). There were a significantly greater number of 26b$^{k/d}$ fibroblasts that migrated to the underside of the membrane compared to con$^{k/d}$ fibroblasts (**Figure 5.7D**).

The transwell assay was repeated using a second pair of cell lines that had either stable knockdown of miR-26b or control. This second pair was generated using a second fibroblast cell line (IF2). IF2s that was knocked down for miR-26b was designated as "26b$^{k/d}$2" and the control knockdown was designated as "con$^{k/d}$2". The degree of migration was assessed at 12 hours. Again there was a significantly greater number of 26b$^{k/d}$2 fibroblasts that migrated compared to the con$^{k/d}$2 fibroblasts at 12 hours (**Figure 5.7E**).

Next, I examined the ability of 26b$^{k/d}$ fibroblasts to invade through matrigel using transwell invasion assays. The experimental setup was similar to the migration assay performed above. The degree of invasion was examined at 24 hours. There were a significantly greater number of 26b$^{k/d}$ fibroblasts that migrated to the underside of the membrane compared to con$^{k/d}$ fibroblasts (**Figure 5.7F**). This was repeated using the 26b$^{k/d}$2 & con$^{k/d}$2 fibroblasts, with similar results (**Figure 5.7G**).

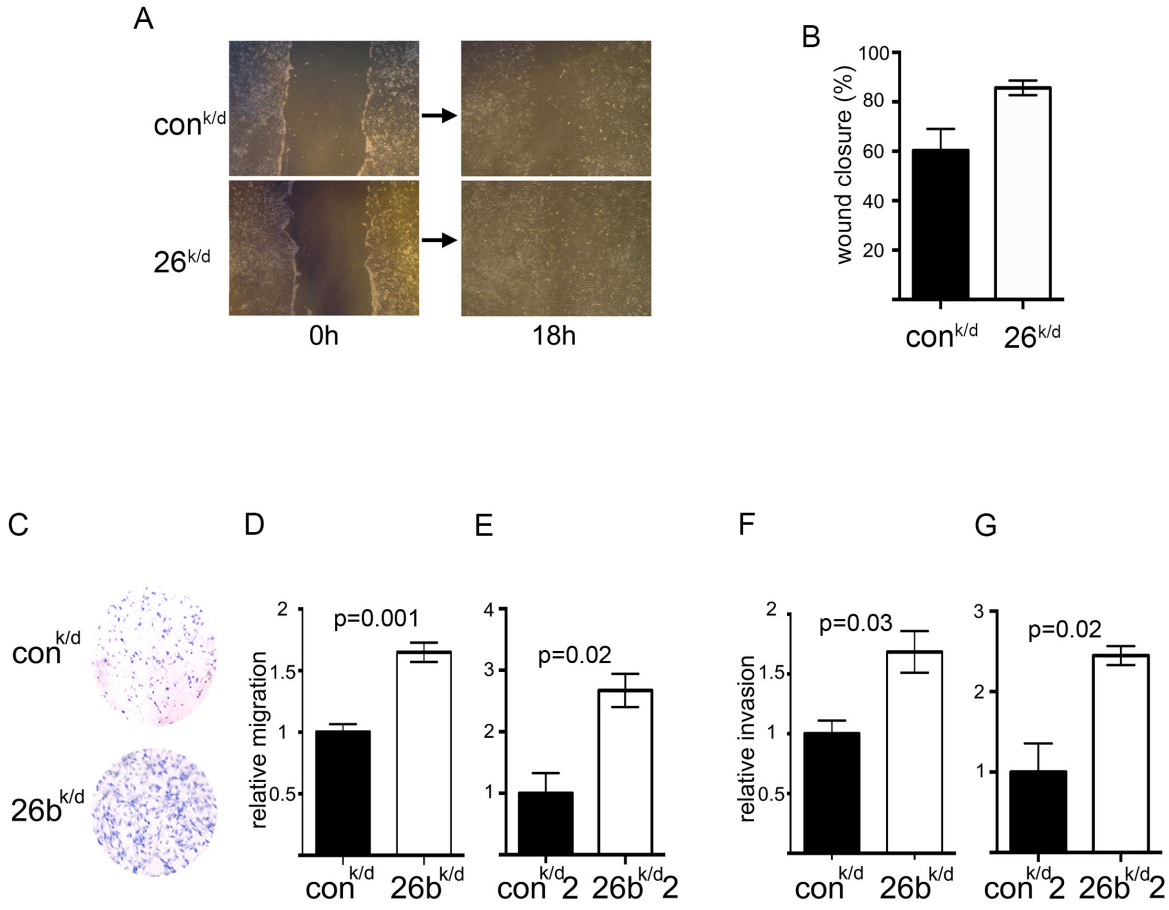

Figure 5.7. Stable knockdown of miR-26b in breast fibroblasts resulted in increased migration and invasion compared to control fibroblasts. Cell scratch assays were performed (A and B). Representative examples of a cell scratch on a confluent layer of 26b$^{k/d}$ and con$^{k/d}$ fibroblasts at 0 and 18 hours are shown (A). The percentage areas of the wound closed were calculated at 18 hours (B). Equal numbers of 26b$^{k/d}$ or con$^{k/d}$ and 26b$^{k/d}$2 or con$^{k/d}$2 cells were plated on the upper side of transwell inserts without (D & E) or with matrigel (F & G). The number of cells that migrated to the underside of the membrane was counted (C) at 12 hours

for the migration assay (D & E) and 24 hours for the invasion assay (F & G). Number of cells counted: 20-150 cells. A representative example of the trans-well migration result is shown. The data points represent the mean of biological triplicates.

5.3 FUNCTIONAL INFLUENCE ON EPITHELIAL CELLS DUE TO 26BK/D FIBROBLASTS

It was clear from the above results that there are changes in cellular behaviour within the fibroblasts due to down-regulation of miR-26b. However, a key question was whether these changes could potentially result in changes in the behaviour of epithelial cancer cells directly or indirectly. Therefore, I examined changes in epithelial cell behaviour when they were co-cultured with fibroblasts that were knocked down for miR-26b or control.

5.3.1 $26B^{K/D}$ FIBROBLASTS DO NOT RESULT IN DIFFERENTIAL GROWTH OF EPITHELIAL CELLS

The most commonly reported change in the epithelial phenotype due to the presence of CAFs, is an enhancement in the growth of epithelial cells. Therefore, I initially examined to see if there was any difference in the growth of MCF-7-Luc cells when they were cultured along with either $26b^{k/d}$ or $con^{k/d}$ fibroblasts. MCF-7 cells are a breast cancer cell line that is ER positive and Her2 negative. They represent the ER+/luminal type subtype of breast cancer (Section 1.2.2), as was used for the initial profiling of miRNA expression in fibroblasts. MCF-7-Luc cells are cells that stably express luciferase. MCF-

7-Luc cells lines were used so that epithelial cells could be specifically quantified when they were co-cultured with fibroblasts.

MCF-7-Luc cells were cultured along with $26b^{k/d}$ or $con^{k/d}$ fibroblasts in direct contact with each other. The initial ratio of epithelial cells to fibroblasts was 1: 3. This ratio was based on *in-vivo* studies that had shown a difference in the growth of epithelial cells using a similar ratio **(Orimo et al., 2005)**. The growth of MCF-7-Luc cells was quantified at 24, 48 and 72 hours using luciferase assays. There was no difference in the growth of MCF-7-Luc cells when they were grown with either $26b^{k/d}$ or $con^{k/d}$ cells (**Figure 5.8A**).

I varied the ratio of MCF-7-Luc to $26b^{k/d}$ or $con^{k/d}$ fibroblasts to 1:1 and 3:1 and examined the growth of MCF-7-Luc at 72 hours. There were no significant differences in the growth of MCF-7-Luc using any of these ratios (**Figure 5.8B**).

MCF-7-Luc cells are positive for oestrogen receptor. Fetal calf serum is known to contain multiple hormones, including oestrogen, that stimulate the growth of epithelial cells. The presence of these hormones might mask the influences of potential signalling molecules secreted from fibroblasts that could influence the growth of epithelial cells. Therefore I also modified cell culture conditions by using charcoal stripped fetal calf serum and varying the concentration of fetal calf serum in the DMEM medium. The above co-culture experiments were repeated using these modified conditions, but I could not detect a difference in the growth of MCF-7-Luc cells when cultured with the two fibroblast lines.

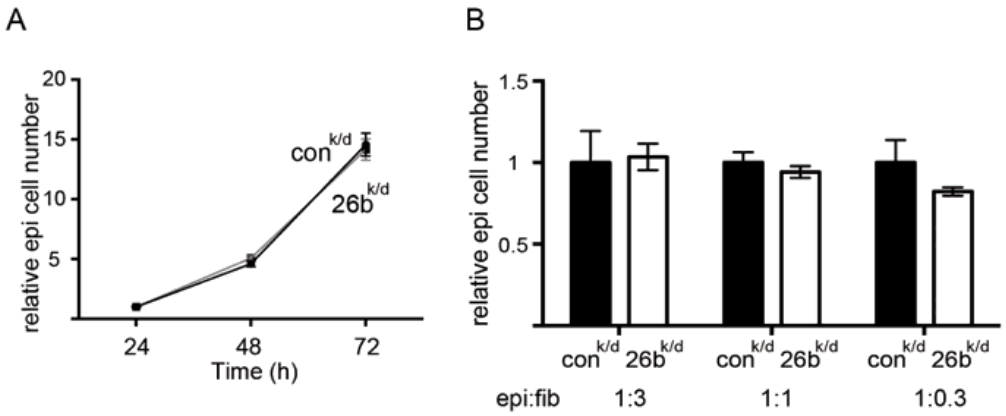

Figure 5.8. The was no difference in the growth of MCF-7 breast cancer epithelial cells when they were co-cultured with breast fibroblasts that were knocked down for miR-26b as compared to control breast fibroblasts. MCF-7 cells (luciferase positive) were grown with $26b^{k/d}$ or $con^{k/d}$ fibroblasts at an initial seeding ratio of 1:3. Epithelial cell growth was monitored using luciferase assays at the time points shown (A). Cells were also co-cultured with $26b^{k/d}$ or $con^{k/d}$ fibroblasts in the ratios shown (B). The growth of epithelial cells was examined at 72 hours using luciferase assays. The data points represent the mean of biological triplicates.

5.3.2 $26B^{K/D}$ FIBROBLASTS INCREASE MIGRATION OF EPITHELIAL CELLS IN DIRECT CO-CULTURE MIGRATION ASSAY

The presence of CAFs is also known to increase the migration of malignant epithelial cells (See section 1.3.2). This could be due to a number of reasons that include paracrine influences on epithelial cells from soluble factors secreted by fibroblasts, alterations in the physical interaction between the fibroblasts and epithelial cells, or from insoluble ECM component laid down by the fibroblasts (see section 7.3)

To explore these possibilities I initially performed indirect and direct transwell migration assays. I examined if the presence of $26b^{k/d}$ or $26b^{k/d}2$ fibroblasts, or appropriate control fibroblasts resulted in differences in migration of MCF-7-Luc cells.

5.3.2.1 INDIRECT MIGRATION ASSAY

I initially performed an indirect co-culture experiment to examine if there were any soluble factors secreted by $26b^{k/d}$ fibroblasts that could result in an increase migration of epithelial cells.

$26b^{k/d}$ or $con^{k/d}$ fibroblasts were plated onto each well of a 24 well plate. MCF-7-Luc cells were plated onto the upper well of transwell inserts. The transwell inserts were initially placed in empty wells of a 24 well plates. After 24 hours the inserts were placed in the wells containing fibroblasts. Migration of MCF-7-Luc cells was assessed after 24 hours by counting the number of cells that migrated to the underside of the membrane.

There was no difference in the number of MCF-7-Luc cells that migrated with either $26b^{k/d}$ or $con^{k/d}$ fibroblasts (**Figure 5.9A**). This demonstrated that there were no

secreted factors that resulted in an increase in migration at 24 hours of MCF-7-Luc cells with the above conditions.

5.3.2.2 DIRECT MIGRATION ASSAY

It is possible that migration might be also be influenced by direct physical interactions between the epithelial cells and surrounding stromal cells **(Camp et al., 2011; Sung et al., 2011)**. To examine if this might be the case with my cells, migration of epithelial cells was examined when they were grown in direct contact with $26b^{k/d}$ fibroblasts or control fibroblasts.

For this I plated MCF-7-Luc cells and $26b^{k/d}$ or $con^{k/d}$ fibroblasts in a ratio of 1:3 onto the upper side of transwell inserts. The number of MCF-7-Luc cells that had migrated to the underside of the membrane was assessed at 24 hours. Since only MCF-7-Luc cells had a luciferase gene, luciferase activity was measured as a surrogate marker to quantify the migration of only these cells through the membrane **(Coombe et al., 1998)**.

There were a significantly greater number of MCF-7-Luc cells that migrated through the membrane when they were grown along with $26b^{k/d}$ cells compared to MCF-7-Luc cells grown along with $con^{k/d}$ fibroblasts (**Figure 5.9B**). Combining the results from the direct and indirect co-culture experiments indicates that the stimulation of migration of epithelial cells may require direct cell-cell contact between $26b^{k/d}$ fibroblasts and MCF-7-Luc. However, this does not necessarily mean mechanical forces acting on epithelial cells by fibroblasts. The increased migration could still be a soluble factor secreted by fibroblast or changes in pro-migratory signalling that has occurred in the epithelial cells either of which is activated by physical contact with the fibroblasts.

5.3.2.3 DIRECT CO-CULTURE TRANSWELL INVASION ASSAY

In invasion assays using monocultures, the migrating cells initially digest the extracellular matrix, and then squeeze through pores of the membrane to reach the underside of the membrane. In a co-culture setting, this process might be aided by neighbouring cells. Neighbouring cells could either provide enhancement of the digestive capacity or directly contribute to the movement of epithelial cells through cell-to-cell interactions.

The transwell assays were performed as described before. But here transwell inserts coated with Matrigel were used. As before, luciferase assays were used to quantify the epithelial cells that had invaded through the Matrigel barrier at 24 hours.

There appeared to be an increased invasion of MCF-7-Luc cells when they were grown along with $26b^{k/d}$ fibroblasts compared to when they were grown with $con^{k/d}$ fibroblasts **(Figure 5.9C**); however, this did not reach statistical significance.

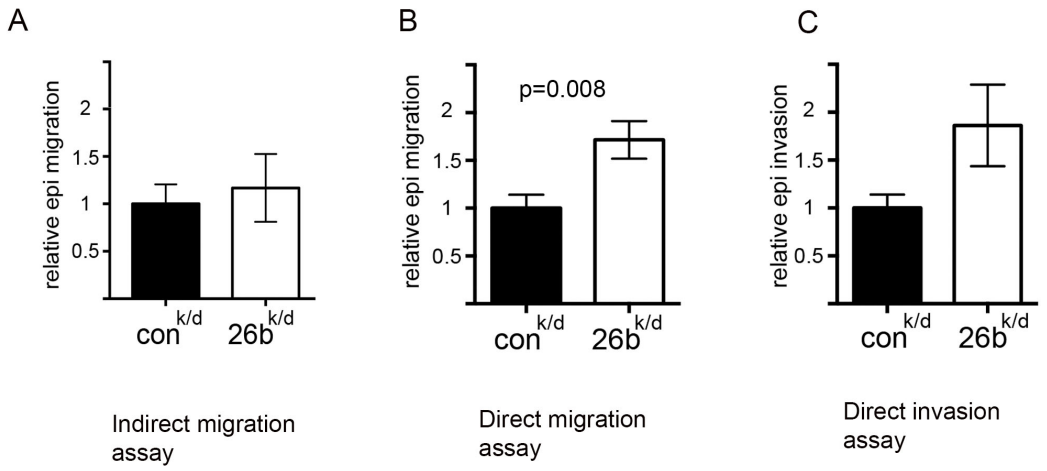

Figure 5.9. Migration of MCF-7 breast cancer epithelial cells increased when they were in direct co-culture with breast fibroblasts knocked down for miR-26b. For in-direct co-culture migration assays $26b^{k/d}$ or $con^{k/d}$ fibroblasts were plated in the lower Boyden chamber and migration of epithelial cells from the top of the transwell inserts was quantified by counting at 24 hours (A),. For direct co-culture migration and invasion assays $26b^{k/d}$ or $con^{k/d}$ fibroblasts and MCF-7-Luc cells (luciferase positive) were plated together on top of the transwell inserts without (B) or with a Matrigel barrier (C). Luciferase activity of migrated cells was measured at 24 hours. Data points for A and C represent the mean of biological triplicates. Data points for B represent the mean of nine experimental replicates.

5.3.3 THE PRESENCE OF $26B^{K/D}$ FIBROBLASTS INCREASED INVASION OF EPITHELIAL CELLS IN A 3D MODEL

Invasion of cancer cells through collagen is important not only in metastasis (intravasation and extravasation) but also in the context of growth of tumour cells into the surrounding tissue. Unlike the basement membrane that contain mainly collagen

type IV, the stroma of normal breast tissue consists mainly of collagen type 1 **(Jarocka-Cyrta and Bankowski, 1991)**. Therefore, the ability of epithelial cells to invade was assessed using a 3-D spheroidal assay containing mainly collagen type I.

5.3.3.1 3-D SPHEROIDAL INVASION ASSAY

The ability of epithelial cells to invade into the surrounding tissue was assessed by performing a 3D spheroidal invasion assay, using collagen gel that had collagen-I as one of its major component (appendix 9.2). This type of 3D invasion assay also helps to examine if cell-matrix interactions play a role in the invasion of cells.

Spheroids containing MCF-7-Luc cells and $26b^{k/d}$ or $con^{k/d}$ cells were generated. Here the ratio of epithelial cells to fibroblasts was 1:1. This was based on the previous work that used similar ratios to study the influence of CAFs on squamous cell carcinoma cell lines **(Gaggioli et al., 2007)**. The spheroids were embedded in a Matrigel/collagen-I gel. Invasion of epithelial cells was quantified at 24 and 48 hours by phase contrast microscopy.

There was substantial invasion of epithelial cells from the central spheroid into the surrounding matrix in the spheroids containing $26b^{k/d}$ fibroblasts, but much less invasion from those with $con^{k/d}$ fibroblasts (**Figure 5.10A**). Using phase contrast microscopy the invading epithelial cells appeared as “tongue” shaped structures extending out of the spheroids with a typical epithelial cobble stone appearance. Invasion of epithelial cells contained in the spheroid comprising $con^{k/d}$ fibroblasts was negligible. In addition, spindle shaped cells ($26b^{k/d}$ or $con^{k/d}$ fibroblasts) were seen to be radiating out of the spheroids from the periphery. The above experiment was repeated using $26b^{k/d}2$ or $con^{k/d}2$ fibroblasts, with similar results.

To confirm the individual cell type seen on the phase contrast microscopy, H&E staining was performed on these spheroids. For this, the spheroid were fixed in 4% formalin

overnight and embedded in paraffin. Paraffin sections were stained with H&E stain. This confirmed the predominance of epithelial cells in the “tongue” shaped structures (**Figure 5.10B).** The fibroblasts were seen enveloping the spheroids and extending out of the spheroids in a radial pattern. A minor population of fibroblasts was also present within the “tongue” shaped structures. The central area of the spheroids was necrotic. To confirm that the “tongue” shaped structures contained predominantly epithelial cells, immunohistochemistry staining using a broad-spectrum anti-epithelial cytokeratins antibody was also performed. The majority of the cells in these structures were positive for cytokeratin staining, whereas the spindle shaped cells were negative (**Figure 5.10C**) thus confirming the cell type.

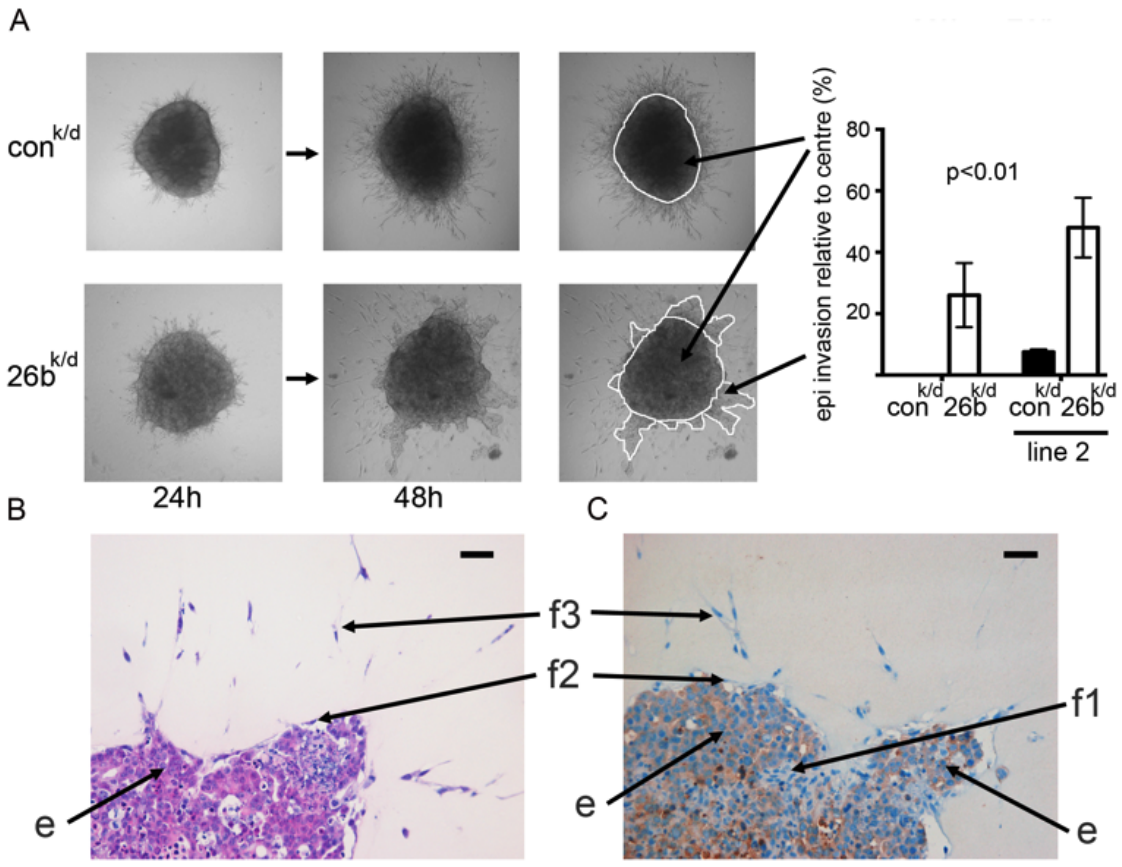

Figure 5.10. Invasion of breast cancer epithelial cells is increased by breast fibroblasts with reduced miR-26b in 3D spheroidal assays. MCF-7 cells and miR26$^{k/d}$ fibroblasts or con$^{k/d}$ fibroblasts were aggregated to form 3-dimensional spheroids. These spheroids were suspended in a collagen-I/Matrigel matrix for up to 48 hours. Invasion of MCF-7 cells extending away from the central spheroid was quantified at 48 hours (A). A similar assay was performed using 26$^{k/d}$2 or con$^{k/d}$2 fibroblasts (line 2). The spheroids were formalin fixed and paraffin embedded. Sections were made and H&E staining and broad spectrum cytokeratin staining was performed (B & C). e= epithelial cells, f1= fibroblasts present within the spheroids, f2=fibroblasts present on the outer surface of the spheroids, f3=fibroblasts extending outwards from the spheroids. Data points

represent the mean of at least three biological replicates and the results are representative of at least 2 independent experiments.

5.4 GROWTH OF MCF-7-LUC BUT NOT FIBROBLASTS WAS ESTABLISHED IN AN *IN-VIVO* XENOGRAFT MODEL

I wanted to examine if the findings observed *in-vitro* could be replicated *in-vivo*. Therefore we decided to use a co-culture xenograft model to examine the influence of miR-26b in fibroblasts on breast cancer epithelial cells **(Orimo et al., 2005)**. I initially performed a pilot study to establish the xenograft model.

We used 8 athymic Ncr-nu/nu nude female mice with two experimental groups containing 4 mice each. A single oestrogen pellet was inserted into each mouse one week before the cell inoculation. MCF-7-Luc and $26b^{k/d}$ or $con^{k/d}$ fibroblasts were pooled and inoculated into the right lower mammary gland. The ratio of epithelial cells to fibroblasts was 1:3 (Total cell count 4 x 10^6). Fluorescence images (GFP positive fibroblasts) and luciferase bioluminescence images (luciferase positive MCF-7-Luc cells) were obtained at various time points from 1 to 25 days.

On day 1, fluorescence was detected in 2 $26^{k/d}$-mice (**Figure 5.11A**) and 3 $con^{k/d}$-mice (**Figure 5.11B**). There was an initial increase in the fluorescence intensity (indicating growth of fibroblasts) in all 3 $con^{k/d}$-mice on Day 4. However, subsequently there was a gradual decrease in the fluorescence intensity. For the $26^{k/d}$-mice there was a gradual decrease in fluorescence intensity after Day 1.

Luciferase bioluminescence indicating the presence of MCF-7-Luc was detected in all 8 mice on Day 1 (**Figure 5.11C**). However two of the $26^{k/d}$-mice did not show any further increase in total luciferase flux (**Figure 5.11D**). Incidentally, these were the mice that did not show fluorescence intensity on Day 1. The other two showed a gradual increase in

the total luciferase flux indicating an increase in the number of MCF-7-Luc cells in the xenografts. Data from only these two mice were taken forward for further analysis. All four $con^{k/d}$-mice showed a gradual increase in the total luciferase flux. There was no difference in the growth rates of MCF-7-Luc that were grown with either $26^{k/d}$ or $con^{k/d}$ up to day 25 (**Figure 5.11E**).

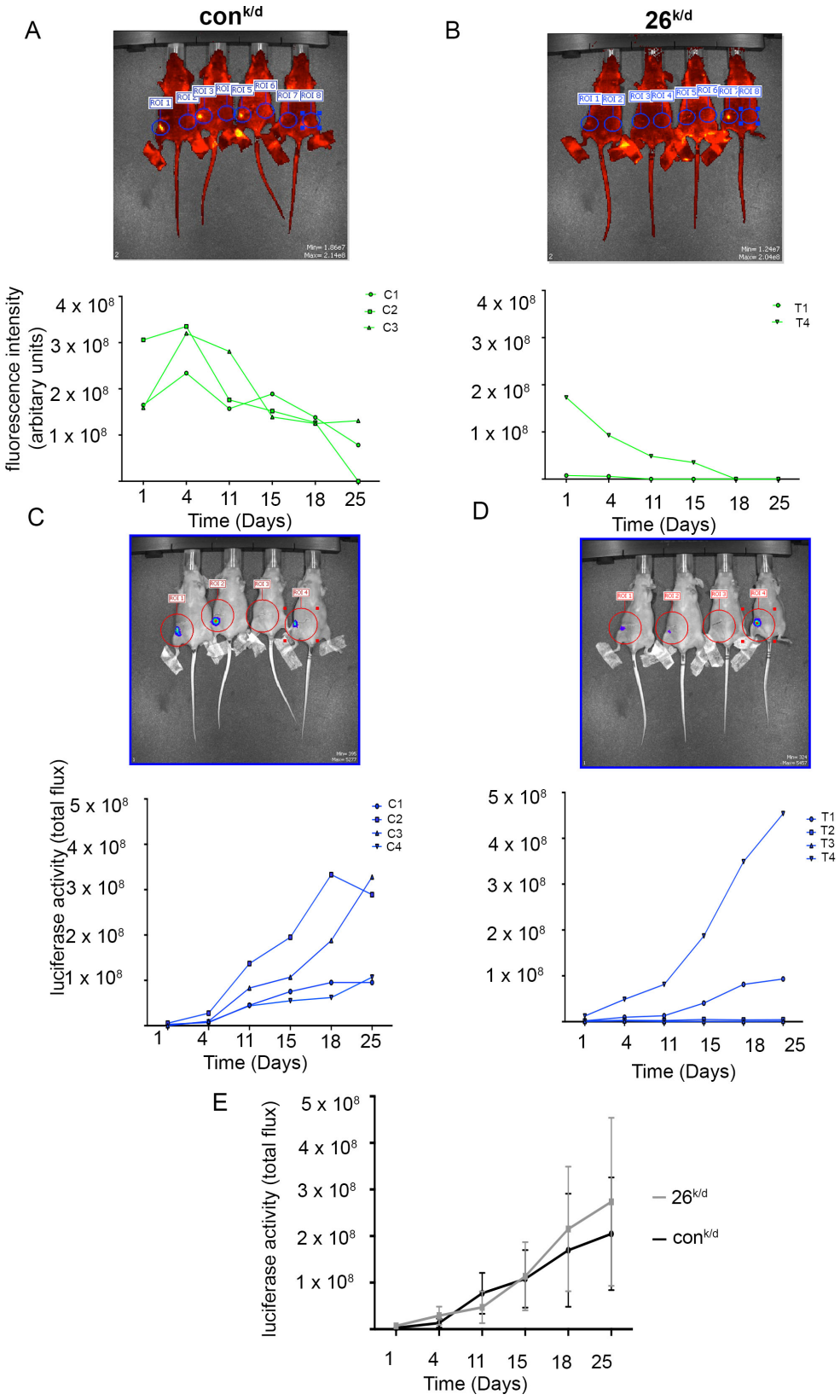
A
con k/d
B
26 k/d
fluorescence intensity (arbitary units)
4 x 10^8
3 x 10^8
2 x 10^8
1 x 10^8
C1
C2
C3
T1
T4
1 4 11 15 18 25
Time (Days)
C
D
luciferase activity (total flux)
5 x 10^8
4 x 10^8
3 x 10^8
2 x 10^8
1 x 10^8
C4
T2
T3
E
26 k/d
con k/d

Figure 5.11 Growth of breast cancer epithelial cells was established in a xenograft model, but immortalised breast fibroblasts were unable to grow. GFP fluorescence (A) and luciferase activity (C) in con$^{k/d}$-mice (containing luciferase positive MCF-7 cells [blue/green colour] and GFP positive con$^{k/d}$ fibroblasts [yellow colour]). GFP fluorescence (B) and luciferase activity (D) in 26$^{k/d}$-mice (containing luciferase positive MCF-7 cells and GFP positive 26b$^{k/d}$ fibroblasts). In both mice, the initial seeding ratio of epithelial cells to fibroblasts was 1:3. Fluorescence intensity (as a surrogate measure for the number of fibroblasts) and luciferase activity (as a surrogate measure for the number of MCF-7-Luc cells) was measured for individual mice on the days shown. Mean luciferase activity was calculated for the con$^{k/d}$-mice (black line) and 26$^{k/d}$-mice (grey line). Error bars represent the standard error of mean.

It was decided to sacrifice the mice on Day 25 because there was concern about the validity of the co-culture xenograft experiment since fibroblast growth was not established. In addition, although the decrease in fluorescence could indicate a decrease in the fibroblast population it was also possible that the decrease could be due to silencing of the lentiviral knockdown vector in the cell lines. To investigate this further, expression of human vimentin was examined as a marker to indicate the presence of human fibroblasts in the formalin fixed tumour xenograft samples.

Tumours were dissected out of the mice and fixed in 10% buffered formalin for 1 day, prior to paraffin embedding. I examined the vimentin expression in three tumours: one GFP positive at day 25, one GFP positive earlier but lost by 25, and one never positive. IHC with vimentin was performed on the paraffin sections from the FFPE tumour blocks.

Vimentin expression was present in a small percentage of cells only in the tumours from C2 (**Figure 5.12A**). In C4 there were a few scattered cells that were positive for vimentin (**Figure 5.12C**). Interestingly there was no staining for vimentin in C1 (**Figure**

5.12B). The epithelial cells in all tumours were negative for vimentin expression. Therefore the IHC expression of vimentin did not correlate with the GFP expression.

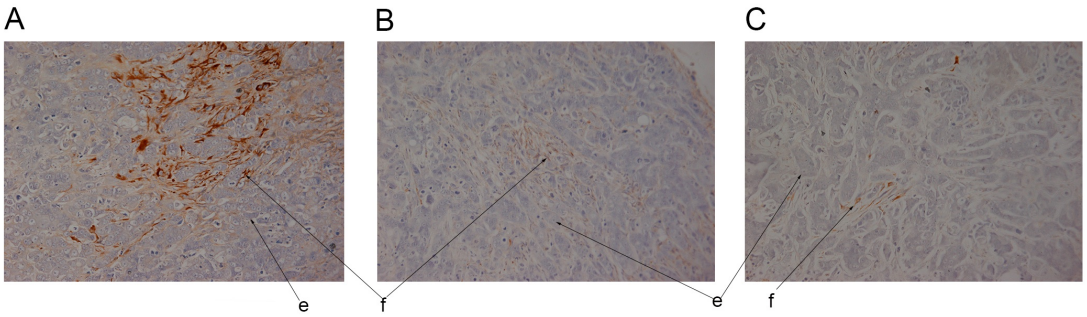

Figure 5.12. Vimentin expression in FFPE sections of the tumour grafts from C2 (A), C1 - positive control (B) and C4 - negative control (C). e=epithelial cells, f=fibroblasts. Brown staining of the cells indicates the presence of vimentin.

Co-culture xenografts have been used by a number of different investigators to study the influence of CAFs on cancer epithelial cells **(Cheng et al., 2005; Cheng et al., 2007; Noel et al., 1993; Orimo et al., 2005)**. In most of these studies investigators have used immunodeficient mice. The primary outcomes in all such experiments have been either the total tumour volume or the growth of cancer epithelial cells alone. Unfortunately, in most of these studies investigators did not quantity the growth of CAFs in the xenografts. Only one study specifically reported on the growth of CAFs alone in a xenograft model using Balb/c nu/nu mice. Although they showed an increase in the tumour growth when cancer epithelial cells were co-injected with CAFs, CAFs on their own failed to establish in the mice **(Olsen et al., 2010)**. This was similar to our results, where CAFs failed to establish whereas cancer epithelial cells in the same mice did form tumours. Perhaps it is not surprising that CAFs are unable to establish *in-vivo* as they are of course non-tumourigenic; however there are doubts as to the validity of this model in the absence of such growth.

My results show that cancer epithelial cells had established and were continuing to grow at day 25. However, this was a pilot experiment and there were insufficient numbers of mice in each experimental group to hope to gain convincing data from continued monitoring of this growth. A repeat experiment with larger numbers of mice is required to draw definite conclusions from the xenograft experiments.

5.5 DISCUSSION

The effects on growth and apoptosis of fibroblasts associated with overexpression of miR-26b are similar to the results that most investigators have obtained when they have overexpressed miR-26b in various other cells types (see Section 7.1.3). Although there is no doubt regarding the technical validity of these results, the biological relevance is uncertain. I have shown that the experimental degree of change in the levels of miR-26b far exceeds the physiological difference between breast NFs and CAFs. Therefore, it is a possibility that the changes seen when overexpressing miR-26b in fibroblasts might represent an “off-target” effect.

There was a significant difference in fibroblast growth after transient knockdown of miR-26b as compared to control knockdown at 48 hours. However this was lost at 72 hours. This might reflect the loss of knockdown efficiency of anti-miR-26b™ miRNA inhibitors at 72 hours. However, this was not formally tested.

In the co-culture transwell invasion assays there was increased invasion of MCF-7-Luc cells when they were grown along with $26b^{k/d}$ fibroblasts compared to when they were grown with $con^{k/d}$ fibroblasts. It is likely that this did not reach statistical significance due to the limited number of replicates.

Although, growth of CAFs could not be established in the xenograft model, a pertinent observation was the pattern of growth of MCF-7-Luc cells in the various mice. While there was an increase in the number of MCF-7-Luc cells in most mice, there was a striking difference in the rate of growth between individual mice. This probably reflects the intrinsic variability between individual mice. In addition, technical consideration can

also result in such differences. For example, variation in the amount of inoculum that has been retained in the mammary gland can also generate this variation. Although, the xenograft experiment was a pilot experiment, it highlighted the need to have large numbers mice in each experimental arm to perform a full experiment.

6 IDENTIFICATION OF FUNCTIONALLY RELEVANT MIR-26B TARGETS

6.1 ABSTRACT

To identify the potential direct targets of miR-26b, I initially used an in-silico analysis. However, this approach identified a large number of potential targets that were experimentally intractable. Therefore, I combined in-silico analysis with label-free quantification-mass spectrometry (LFQ-MS) to identify proteins that were differentially expressed between fibroblasts that were knocked down for miR-26b compared to controls.

Of 3369 proteins that were identified using LFQ-MS, 218 proteins were significantly up-regulated in breast fibroblasts that had reduced miR-26b ($p<0.05$). Using a false discovery rate set to <0.1, I identified three potential direct targets of miR-26b, namely COL12-A1, CPSF7 and TNKS1BP1. To examine the clinical relevance of these targets I interrogated publically available datasets. Increased stromal mRNA expression of two of these targets was significantly associated with increased recurrence of breast cancer on multivariate analysis: COL12A1 ($p=0.01$, HR=17, C.I. 1.9-159) and TNKS1BP1 ($p=0.034$, HR=9.1, C.I. 1.2-69). Of these three, only COL12A-1 was confirmed as a direct target of miR-26b using luciferase assays.

Pathway analysis of all significantly ($p<0.05$) differentially expressed proteins from the mass spectrometry dataset, revealed that proteins in the “Glycolysis” and “Cytoskeleton regulation by RHO GTPase” pathways were significantly enriched, indicating that deregulation of these pathways is likely to be downstream of miR-26b in breast fibroblasts.

In conclusion, mir-26b might potentially regulate the “Glycolysis” and “Cytoskeleton regulation by RHO GTPase” pathways and COL12-A1 is a likely direct target of miR-26b.

6.2 SCREENING USING IN-SILICO TOOLS IDENTIFIED MULTIPLE TARGETS

Having established that the presence of $26b^{k/d}$ fibroblasts results in a change in the behaviour of MCF-7-Luc cells, my next step was to dissect the mechanisms responsible for the change in behaviour of both fibroblasts and epithelial cells. It is likely that this could be due to one or more of the direct targets of miR-26b and/or their downstream effector molecules.

I initially decided to examine the direct targets of miR-26b to identify potential molecules that might be involved in these pathways. For this I performed an *in-silico* screen to identify potential targets.

A total of 3477 potential targets were identified using TargetScan. By filtering for only targets that are conserved across species, this number was reduced to 885. To examine if another algorithm identified the same targets, I generated another list of potential miR-26b targets using PicTar: 597 potential targets were identified. PicTar only identifies target that are conserved across species. When I compared the list of conserved targets, only 250 were common to both PicTar and TargetScan (Appendix, **Table A4**).

The limited overlap of the potential targets identified by the two algorithms confirmed that the different approaches do in fact identify different potential targets. So there might be limited utility and/or validity using this combinatorial approach since the true targets might be the ones that do not fall into the overlapping category. In addition this *in-silico* analysis resulted in large lists of potential targets that were not easily manageable experimentally. Furthermore, these algorithms do not take into consideration the cellular

context and therefore it is difficult to decide using *in-silico* analyses alone the true targets of miR-26b in breast CAFs.

6.3 LABEL FREE QUANTIFICATION-MASS SPECTROMETRY IDENTIFIED MULTIPLE PROTEINS THAT ARE DIFFERENTLY EXPRESSED

Since the initial approach of using *in-silico* methods alone raised problems identifying the key targets, another screening tool was required to complement the *in-silico* analysis. Therefore I used label free quantification-mass spectrometry (LFQ-MS) to identify the actual proteins that were differentially expressed between $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts. The differentially expressed proteins were then compared against the database of miR-26b targets generated by TargetScan to attempt to identify the true targets.

Protein lysates of $26b^{k/d}$ or $con^{k/d}$ fibroblasts were prepared in triplicate, further processed and analysed using LC-MS/MS. The complete data is available in the "Mass Spectrometry data folder" on the attached CD. I initially examined the correlation of protein expressions between the triplicates. A Pearson correlation coefficient was calculated that demonstrated a very good correlation between triplicate samples of both $26b^{k/d}$ and $con^{k/d}$ fibroblasts (**Figure 6.1, Table 6**).

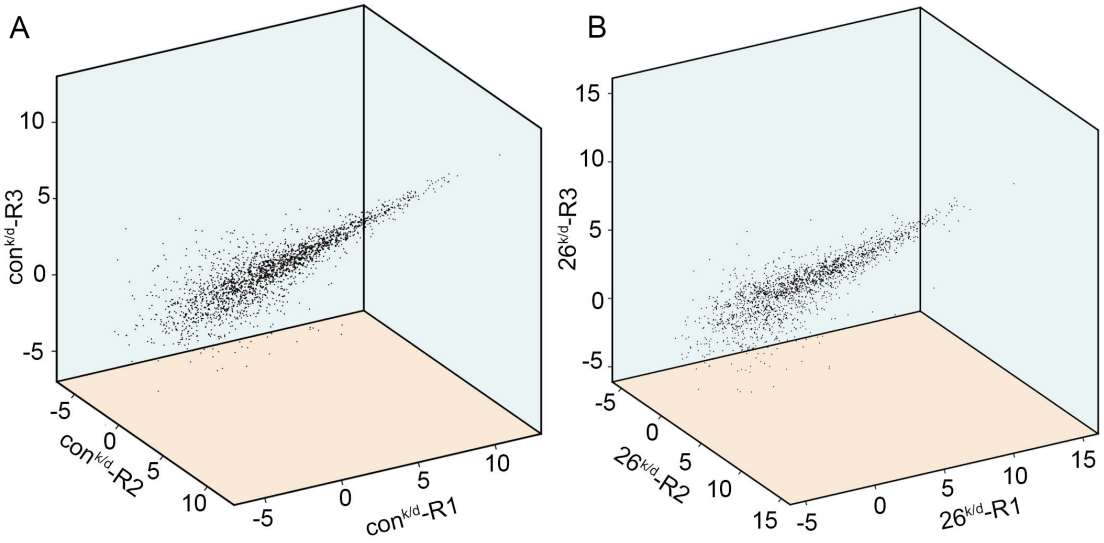

Figure 6.1. There is good correlation between the quantification of proteins using mass spectrometry in triplicate samples. Graphical representation of the quantity of protein (arbitrary units) in the biological triplicates (labelled R1, R2 and R3) of con$^{k/d}$ fibroblasts (A) and 26$^{k/d}$ fibroblasts (B).

Comparison	Correlation coefficient	p-value
$con^{k/d}$-1 v/s $con^{k/d}$-2	0.854	<0.0001
$con^{k/d}$-1 v/s $con^{k/d}$-3	0.862	<0.0001
$con^{k/d}$-2 v/s $con^{k/d}$-3	0.877	<0.0001
$26b^{k/d}$-1 v/s $26b^{k/d}$-2	0.793	<0.0001
$26b^{k/d}$-1 v/s $26b^{k/d}$-3	0.835	<0.0001
$26b^{k/d}$-2 v/s $26b^{k/d}$-3	0.849	<0.0001

Table 6.1. Correlation coefficients of pairwise comparison of the quantity of proteins assessed using mass spectrometry in biological triplicates of $con^{k/d}$ fibroblasts and $26^{k/d}$ fibroblasts (in each case individual samples are designated samples 1, 2 or 3). Pearson correlation coefficients were examined to determine correlations between samples.

A large number of proteins were differentially expressed ($p<0.05$, t-test) between $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts. This was performed by comparing the expression of triplicate values of the lysates using a t-test. 218 (60%) of the 360 proteins that were differentially expressed proteins were up-regulated. These up-regulated proteins would include the potential direct targets of miR-26b.

Next, I analysed the data using two cut-off values of False Discovery Rate (FDR): FDR<0.1 and FDR<0.2. The appropriate cut-off for the FDR is arbitrary to an extent and is based on how stringently one needs to control for false discoveries. Using a FDR<0.1, 11 proteins were identified as differentially expressed between $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts. Of these, 5 proteins were up-regulated. Three (60%) of the up-regulated proteins were predicted targets of miR-26b. These were COL12A1,

CPSF-7 and TNKS1BP1. None of the down-regulated proteins were predicted miR-26b targets (**Table 6.2**). Using a FDR<0.2, 103 proteins were identified to be differentially expressed between $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts. Of these, 53 proteins were up-regulated. Interestingly, only 12 (25%) of these were predicted miR-26b targets, demonstrating that the majority of changes may be downstream of the direct targets of miR-26b. Furthermore, 17 (32%) of the down-regulated proteins were predicted miR-26b targets (**Appendix, Table A5**), despite the fact that direct targeting of these by miR-26b would drive their up-regulation. Based on this analysis it is clear that using a FDR cut-off of <0.2 may have resulted in a large number of direct targets and indirect targets. Therefore, further analysis of individual molecules was restricted to the up-regulated proteins identified using a FDR<0.1.

	Regulation / protein fold change*	Predicted target
TMEM119	Down	No
NUDCD3	Down	No
TPM2	Down	No
FAM3C	Down	No
CLTB	Down	No
THOC5	Down	No
USP19	Up	No
APIP	Up	No
TNKS1BP1	Up / 3.32	Yes
CPSF7	Up / 4.35	Yes
COL12A1	Up / 7.20	Yes

Table 6.2. Proteins that were differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts using an FDR cut off value of <0.1. Fold changes were calculated using mean expression levels in triplicate con$^{k/d}$ samples and triplicate 26$^{k/d}$ samples. *Fold changes are shown only when the protein was detected in both cell lines – when it was not detected in either line only the direction of differential expression is indicated ('up' meaning more highly expressed in 26$^{k/d}$ cells).

6.4 COL12A-1 IS A DIRECT TARGET OF MIR-26B

I used Western blotting to confirm the up-regulation of the three potential direct miR-26b targets identified using LFQ-MS. In addition, it was possible that the 'direct' targets that have been identified using the above described combinatorial approach of mass spectrometry and *in-silico* analysis might not be entirely accurate, since the up-

regulated potential targets could be secondary to a downstream effect of another target. Therefore luciferase reporter assays for the three targets were also performed.

Equal masses of protein in RIPA lysates of $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts were analysed using Western blots probed using antibodies to COL12A1, TNKS1BP1 and CPSF-7 (see section 3.11 for antibodies details). α-SMA, which is a common marker for CAFs, is neither a predicted target of miR-26b nor was it differentially expressed by mass spectrometry; therefore this was used as the loading control. This was necessary because many commonly used loading controls, such as GAPDH and β-actin, were found to be differentially expressed in the LFQ-MS dataset. However, since these antibodies were available, Western blots were also performed for these proteins to validate the mass spectrometry data further. TNKS1BP1, GAPDH and β-actin were overexpressed in $26b^{k/d}$ fibroblasts compared to $con^{k/d}$ fibroblasts, whereas the expression of α-SMA did not differ. This pattern of protein expression is in keeping with the results obtained using mass spectrometry (**Figure 6.2**). Unfortunately, antibodies against COL12A1 (two different antibodies used) and CPSF-7 (one antibody only) did not allow detection of any proteins on Western blots.

Luciferase reporter plasmids containing the predicted binding sites for miR-26b from the 3'UTRs of COL12A1, TNKS1BP1 and CPSF-7 (see section 3.13) were co-transfected along with anti-miR-26b inhibitor molecules or control molecules into IF1 cells. Luciferase activities were measured. There was a significant increase in luciferase activity when COL12A1 plasmid was co-transfected with anti-miR-26b inhibitors compared to when it was co-transfected with control molecules (**Figure 6.3**). This suggested that COL12A1 is indeed a direct miR-26b target in these cells, since increased expression of miR-26b targeted transcript would be expected when miR-26b was knocked down. This difference was not seen for TNKS1BP1 or CPSF-7.

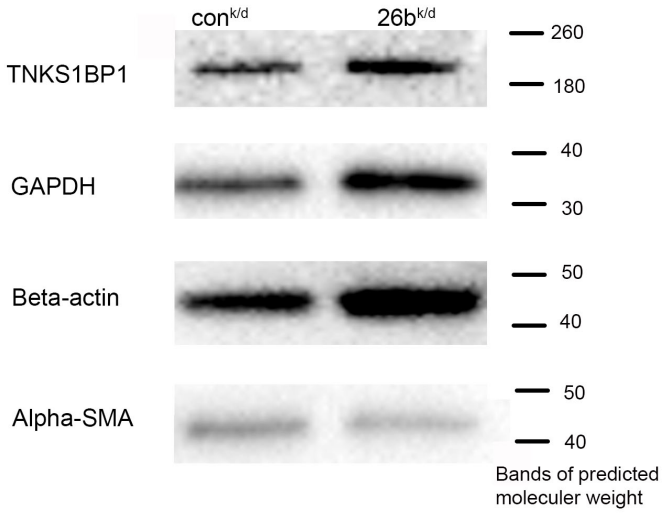

Figure 6.2. Western blots were used to confirm the up-regulation of proteins identified by mass spectrometry in fibroblasts knocked down for miR-26b. Equal amount of proteins from con$^{k/d}$ fibroblasts and 26b$^{k/d}$ fibroblasts were probed using antibodies against TNKS1BP1, GAPDH and β-actin. α-SMA was used as the loading control.

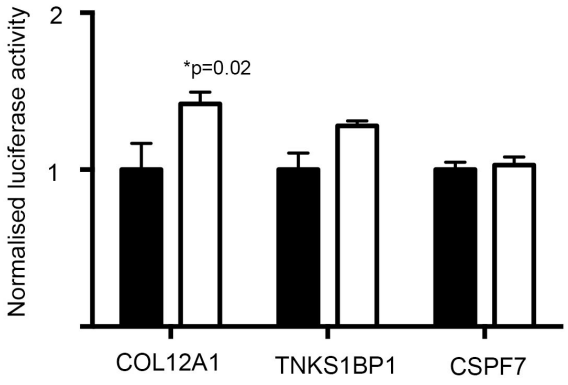

Figure 6.3. COL12A1 is a direct target of miR-26b. Immortalised breast fibroblasts were transfected with anti-miR-26b inhibitor molecules (open bar) or control molecules (black bar) and luciferase reporter plasmids encoding the predicted binding sites for miR-26b from the COL12A1, TNKS1BP1 or CPSF-7 genes in the 3'UTR for firefly luciferase. Luciferase activity (normalised to control renilla luciferase activity) was measured 6 hours after transfection. Data points represent means of biological triplicates.

The Western blot results indicate that mass spectrometry has accurately identified proteins that are truly differentially expressed. Secondly, the luciferase reporter assay indicates that COL12A1 is a direct target of miR-26b whereas TNKS1BP1 and CPSF-7 are apparently not. Although TNKS1BP1 and CPSF-7 are not true targets of miR-26b, they could represent true downstream changes as a consequence of knocking down miR-26b.

6.5 STROMAL EXPRESSION OF COL12A1, TNKS1BP1 AND CPSF-7 PREDICTED RECURRENCE OF BREAST CANCERS

On reviewing the literature, it was apparent that the presence or relevance of fibroblast/stromal expression of COL12A1, TNKS1BP1 or CPSF-7 in breast cancer had not been examined. Therefore the clinical relevance of the expression of these proteins is uncertain. Therefore to address this question, I decided to examine the mRNA expression of these genes in publicly available databases of gene expression. Although this type of analysis might not accurately reflect the relationship of protein expression to clinical outcomes such as tumour recurrence, it does provide an indication of the clinical relevance of the gene in question.

The expression levels for these genes in Finak et al's dataset were obtained using the Oncomine web platform. This dataset contained the mRNA expression profile of laser capture microdissected stroma of 53 primary breast tumours (Section 3.19). The key clinico-pathological parameters for these patients were obtained from the supplementary data file and are listed in (**Table 6.3**). Initially, I tested whether the genes showed differential expression between normal stroma and tumour stroma. There was increased expression of COL12A1 and TNKS1BP1 in the tumour stroma. The mean expression level was 19.3 and 1.3 respectively in tumour compared to 0.97 and 0.96 respectively in normal (Wilcoxon signed rank test, $p=0.035$, $p=0.035$). Interestingly, there was decreased expression of CPSF-7 in the tumour stroma compared to the normal, but this was not significant.

Parameter	Number (%)
Number in cohort	53
Age: Range, 32-83	
Mean, 54	
Median, 54	
Tumour size (mm): <20	21 (40)
≥20	32 (60)
Histological grade: 1	3 (5.7)
2	23 (43.4)
3	27 (50.9)
Number positive lymph nodes: Positive	25 (47.2)
Negative	25 (47.2)
data	3 (5.7)
ER+	43 (81)
PR +	27 (51)
Her-2 +	10 (19)
Systemic Treatment: Endocrine therapy	36 (68)
Adjuvant	37 (70)
Neo-adjuvant	Nil
Radiotherapy	42 (79)
Cases with survival data	53 (100)
Follow up (months): Median, 41	
Mean, 39	
Range, 0-75	

Table 6.3. Clinico-pathological parameters of breast cancer patient cohort used by Finak et al. to examine mRNA expression profiles of tumour stroma (Finak et al., 2008)

Next, expression levels of COL12A1, TNKS1BP1 and CPSF-7 in the tumour stroma were dichotomised and their relationships to various clinico-pathological parameters and clinical outcomes were examined. There were no significant association between COL12A1 or CPSF-7 expression and grade, lymph node positivity, tumour size, ER, PR or Her-2 status. Similarly, there were no significant association between TNKS1BP1 expression and grade, tumour size, ER, PR or Her-2 status. However, there was a significant association between higher TNKS1BP1 and a positive lymph node status (p=0.001, chi squared test). The median follow up time was 3.58 years and the clinical outcome was defined as local or distant recurrence. On univariate analysis higher expression of COL12A1 (log rank p=0.043), TNKS1BP1 (log rank p=0.002) and CPSF-7 (log rank p=0.007) was associated with reduced recurrence free survival (**Figure 6.4**). On multivariate analysis including the clinico-pathological parameters listed in Table 6.3 (tumour size, grade, lymph node status, ER and Her-2 status) the associations between expression of COL12A1 (p=0.01, HR=17, C.I. 1.9, 159) or TNKS1BP1 (p=0.034, HR=9.1, C.I. 1.2, 69) and outcome remained significant. However, differential expression of CPSF-7 and recurrence free survival was not significant in multivariate analysis.

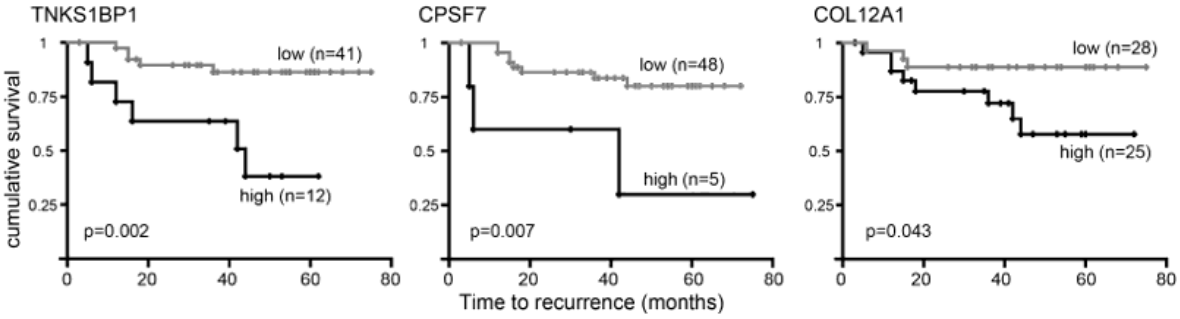

Figure 6.4. Kaplan-Meier analyses showing higher expression of TNKS1BP1, CPSF7 or COL12A1 to be associated with reduced recurrence free survival. mRNA expression data from 53 laser capture microdissected stromal breast cancer tissue was obtained using Oncomine (www.oncomine.org). The patient cohort was divided into two groups (low & high) based on the degree of mRNA expression of TNKS1BP1, CPSF7 or COL12A1. Log Rank tests were used to compare the survival distributions between the groups.

6.6 PATHWAY ANALYSIS IDENTIFIED THE GLYCOLYSIS PATHWAY AND THE CYTOSKELETON REGULATION BY RHO GTPASE PATHWAY AS DOWNSTREAM OF MIR-26B

I wanted to identify the pathways that might be influenced when miR-26b is down-regulated in fibroblasts. There are a number of different tools available online for pathway analysis. These tools examine a gene list for enrichment of a set of genes that belong to a particular pathway by comparing the test list to gene groupings within a database. Of these the tools available, ToppGene and ENDEAVOUR **(Aerts et al., 2006)** focus on using multiple data sources for their analyses. For my analysis I decided to use "ToppGene suite", http://toppgene.cchmc.org/, since ToppGene reportedly performs better than ENDEAVOUR for gene enrichment analysis **(Chen et al., 2007)**. A Bonferroni correction with a p-value cut-off of 0.05 was included in the ToppGene pathway analysis to correct for multiple comparisons.

Proteins that were differentially expressed (with a p-value<0.05) between $26b^{k/d}$ fibroblasts and $con^{k/d}$ fibroblasts as identified by mass spectrometry were analysed. This included 373 proteins (Appendix Table A2). Initially, I used all genes in the genome as my background set. I also repeated the analyses using a background set restricted to the set of proteins that were detected using mass spectrometry. For this, I could not use the "ToppGene suite", since it does not have the option to specify a customised reference list. Therefore, I used the "PANTHER" system (http://www.pantherdb.org), which is one of the sources that "ToppGene suite" uses to perform gene enrichment pathway analysis.

When I used whole genome as the background set, the differentially expressed proteins were significantly enriched for molecules that were involved in three pathways: 1) the glycolysis pathway; 2) the "superpathway" of glycolysis, pyruvate dehydrogenase, TCA, and glyoxylate bypass; and 3) cytoskeleton regulation by RHO GTPases. 5 of the 13 components in the glycolysis pathway were present and all of these were up-regulated (**Table 6.4**). 8 of the 34 components of the "superpathway" were present (**Table 6.5**) and

all 8 were up-regulated. 14 of the 72 components of the Cytoskeleton regulation by RHoGTPase pathway were present. Of these, 10 were up-regulated (**Table 6.6**).

When I performed a similar analysis using the restricted reference list, molecules belonging to the "cytoskeleton regulation by RHO GTPases" pathway were over represented, but this fell just short of statistical significance (p=0.0588). The molecules in the remaining two pathways were not significantly enriched.

Gene symbol	Gene name	Regulation
ALDOA	aldolase A, fructose-bisphosphate	Up
PGK1	phosphoglycerate kinase 1	Up
PGAM1	phosphoglycerate mutase 1 (brain)	Up
PGM1	phosphoglucomutase 1	Up
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	Up

Table 6.4. Components of glycolysis pathway identified by Gene Ontology analyses as enriched in the proteins differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts. Gene names and regulation are shown.

Gene symbol	Gene name	Regulation
ALDOA	aldolase A, fructose-bisphosphate	Up
PGK1	phosphoglycerate kinase 1	Up
PGAM1	phosphoglycerate mutase 1 (brain)	Up
ENO1	Enolase 1	Up
SDHA	succinate dehydrogenase complex, subunit A	Up
MDH1	malate dehydrogenase	Up
IDH1	isocitrate dehydrogenase	Up
DLST	dihydrolipoamide S-succinyltransferase	Up

Table 6.5. Components of "superpathway" of glycolysis, pyruvate dehydrogenase, TCA, and glyoxylate bypass identified by Gene Ontology analyses as enriched in the proteins differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts. Gene names and regulation are shown.

Gene symbol	Gene name	Regulation
ACTB	actin, beta	Up
ACTG1	actin, gamma 1	Up
ACTBL2	actin, beta-like 2	Up
ARHGAP1	Rho GTPase activating protein 1	Up
PAK1	p21 protein (Cdc42/Rac)-activated kinase 1	Down
MYH10	myosin, heavy chain 10, non-muscle	Up
ARPC1A	actin related protein 2/3 complex, subunit	Down
CFL2	cofilin 2 (muscle)	Up
CFL1	cofilin 1 (non-muscle)	Up
TUBB6	tubulin, beta 6 class V	Up
PFN1	profilin 1	Up
ARPC5	actin related protein 2/3 complex, subunit 5	Down
ROCK2	Rho-associated, coiled-coil containing	Down
STMN1	stathmin 1	Up

Table 6.6. Components of "cytoskeleton regulation by RHO GTPase" pathway identified by Gene Ontology analyses as enriched in the proteins differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts. Gene names and regulation are shown.

6.7 DISCUSSION

I have relied on bio-informatics tools for the identification of functionally relevant target. Most these tools use algorithms that are based on rules that have been elucidated by earlier studies on miRNA:mRNA interactions. Therefore, such tools have an intrinsic limitation. The most recent discoveries on miRNA:mRNA interactions and, of course, hitherto unknown modes of interaction are not incorporated into such algorithms. For example, recent studies have shown that upto 60% of miRNA:mRNA seed region binding is non-canonical. In addition, most of the seed interactions are accompanied by specific non-seed binding. Even more surprising was that up to 18% of the interactions involve the 3' UTR of miRNA **(Helwak et al., 2013)**. Therefore results from bio-informatics tools for in-silico analyses, might not accurately reflect all the direct targets.

Similarly, the proteomic analyses I have used are beset by limitations. In particular, a number of considerations must be taken into account when performing gene enrichment analysis using mass spectrometry data.

This analysis usually involves the integration of proteome data with annotation databases such as Gene ontology databases and pathway databases. There are several software programmes that allow the end user to perform these types of analyses. Software tools such as DAVID **(Huang da et al., 2007)**, Panther **(Mi et al., 2013)** and Bingo with Cytoscape **(Maere et al., 2005)** are some of the commonly used ones. Inherent differences between these tools, such as the differing databases, might alter the final results, although there is surprisingly little literature in which these possible differences are identified or discussed.

One key limitation of mass spectrometry is that the coverage of proteins that are identified is highly limited, particularly in comparison with transcriptomic techniques **(Cox et al., 2009)**. In addition, mass spectrometry tends to identify more easily abundant proteins compared to those that are less abundant. For example, of all the proteins that I identified using mass spectrometry, 77% of the proteins were annotated as being cytoplasmic. However, of all the proteins in the Swiss-port human protein

database, only 43% are annotated as being cytoplasmic **(Kobayashi et al., 2009)**. Clearly this indicates that the proteins identified using mass spectrometry was biased towards cytoplasmic proteins, which may be a reflection of higher expression levels of cytoplasmic proteins simply on the basis of the cytoplasm being a much greater proportion of the cell than, for example, the nucleus.

In my initial analysis, I performed gene enrichment analysis using genes representing the whole genome as the background set. However, this analysis suffers from some bias since the proteins included in the "test set" are already enriched for genes from the cytoplasmic compartment. Pathway analysis using this background set could potentially falsely implicate pathways whose components are predominantly within the cytoplasm.

A number of strategies have been used to partly overcome this problem. One approach is to only use the proteins that were detected using mass spectrometry as the background set **(Kobayashi et al., 2009)**. The second analysis of the mass spectrometry data using this "restricted background set" identified one of the pathways as being enriched, but this was not statistically significant, indicating that the results from the initial analysis may be a false positive. Another approach is to use a "test set" that only contains differentially expressed proteins with a very low false discovery rate (e.g. FDR<0.2 or FDR <0.3). A more complex approach that has been used in some studies is to initially divide the fold change distribution of the proteome into different quantiles. Subsequent gene ontology (GO) enrichment analysis and pathway analysis was performed separately for these quantiles **(Pan et al., 2009)**. For this approach the whole proteome was used as the background set. However, it is worth noting that these attempts to correct for bias at the analysis stage represent a very small proportion of the literature.

The inherent bias of mass spectrometry can also be dealt with at the sample preparation stage. For example in some studies, subcellular fractions of nuclei, mitochondria, membrane, and cytosol are initially prepared. Subsequently mass spectrometry is performed on the individual compartments to identify the individual

compartment's global proteome. These data can be either combined as one set or can be analysed individually to identify GO categories or pathways that are enriched **(Adachi et al., 2007)**. Another possibility is depletion of more abundant proteins during the sample preparation stage. This will increase the detection and coverage of less abundant proteins **(Fonslow et al., 2013)**. However it is likely that these processes will themselves introduce biases.

It is clear that there is not yet an agreed robust methodology for these types of pathway analyses, and therefore these types of analyses should be used as hypothesis generating tools that will require subsequent testing and validation.

7 DISCUSSION AND FUTURE WORK

7.1 MIR-26B

MiR-26b belongs to a family of microRNAs that consists of miR-26a, miR-26b, miR-1297 and miR-4465. These miRNAs belong to a single family due to the similarity in their seed regions. The miR-26b gene is located on the long arm of chromosome 2, within an intron of the C-terminal domain small phosphatases-2 (CTDSP2) gene **(Dill et al., 2012)**. MiR-26b is transcribed as part of the primary transcript from this protein-coding gene and, interestingly, miR-26b also regulates the translation of this gene product from the host gene.

MiR-26b has been shown to play a role in normal physiological processes, such as follicular atresia of the ovaries **(Lin et al., 2012)** and neuronal differentiation **(Dill et al., 2012)**. It also has varied roles in a number of disease processes such as polycythemia vera, pre-eclampsia, lichen planus and cancer **(Bruchova et al., 2008; Caputo et al., 2011; Danielsson et al., 2012; Hu et al., 2009)**. However, the overwhelming majority of the published literature on miR-26b examines its role in cancer and this will be discussed further.

7.1.1 MIR-26B EXPRESSION IN CANCER

MiR-26b is thought to have a tumour suppressor activity since it is down-regulated in most cancers. It is down-regulated in squamous cell carcinoma of tongue **(Wong et al., 2008)**, hepatocellular carcinoma **(Ji et al., 2009)**, breast cancer **(Li et al., 2013; Liu et al., 2011)**, colorectal cancer **(Ma et al., 2011b)**, parathyroid carcinoma **(Rahbari et al.,**

2011), glioma **(Wu et al., 2011)**, lymphoma **(Peveling-Oberhag et al., 2012)** and squamous cell carcinoma of lungs **(Solomides et al., 2012)**. However, in gastric cancer **(Inoue et al., 2012)** and pituitary tumours **(Palumbo et al., 2012)** miR-26b was up-regulated. Therefore, the function of miR-26b might not be restricted to a tumour suppressor activity and could vary depending on the cancer type. In most of these studies the deregulation of miR-26b was identified by expression profiling in either cell lines, or in whole tumours and normal tissues. Critically in relation to my study, use of whole tumour tissues would include the target cancer cells and the stroma. In addition, the use of whole normal tissue as a comparator is debatable since the normal tissue might not contain the tumour's cell of origin as one of its major components.

There are two studies in which investigators examined the expression of miR-26b in breast cancer. In both studies, the hypothesis that miR-26b could be important in breast cancer was based on the initial miRNA profiling studies by Iorio et al **(Iorio et al., 2005)**. In the first study, expression of miR-26b was examined in 27 breast cancer samples and compared to the matched normal tissue **(Liu et al., 2011)**. There was reduced expression of miR-26b in breast cancer compared to the matched normal tissue in most of the cases. In addition, miR-26b expression was lower in a number of breast cancer cell lines compared to some cultured non-cancer human cells - normal skin fibroblasts. However the relevance of such a comparison is doubtful. In the second study, investigators confirmed the down regulation of miR-26b in further clinical samples using the same approach as above (n=38) **(Li et al., 2013)**. Importantly, in both studies investigators did not use pure/enriched populations of epithelial cells when they examined expression of miR-26b in clinical samples. However, in the single case from which I microdissected the matched normal and tumour epithelial compartments and examined the miRNA expression profiles, miR-26b was down-regulated in the tumour epithelium compared to its matched normal epithelium (Section 4.1.1). Therefore it is possible that down-regulation of miR-26b occurs not only in the fibroblast compartment, as studied in this thesis, but also in the epithelial compartment, as inferred from the above studies.

With respect to expression in fibroblasts, Zhao L et al. specifically examined the expression profile of breast CAFs in culture compared to NFs in culture and showed that miR-26b was down-regulated in 6 cases of CAFs compared to their matched NFs **(Zhao et al., 2012a)**. This is in keeping with my initial findings, where I showed that miR-26b was commonly down-regulated in CAFs. Although Zhao L et al. performed various computational analyses to identify potential roles for miR-26b, they did not perform any functional experiments to confirm these, nor did they examine the influence of miR-26b down-regulation in CAFs on cancer epithelial cell behaviour.

7.1.2 FUNCTIONAL ROLES FOR MIR-26B

To examine further the functional roles of miR-26b, experimental studies have been performed in a number of different tumours cell lines where expression of miR-26b was manipulated. Importantly, all these experiments were performed in the cancer cells, and not any components of tumour stroma. These cancer types include glioma, colorectal cancer, pituitary tumour, nasopharyngeal carcinoma and breast cancer. The functional influence of miR-26b in glioma was studied by overexpressing miR-26b in glioma cell lines **(Wu et al., 2011)**. Overexpression of miR-26b resulted in decreased growth, migration and invasion of glioma cell lines. Similar experiments in colorectal cancer cell lines also showed analogous results **(Ma et al., 2011a)**. In addition, Ma et al. also used an *in-vivo* model to validate the *in-vitro* findings, where they generated tumour xenografts using human colorectal cancer cell lines and injected them with either miR-26b mimics or control mimics. Tumours treated with miR-26b mimics showed significantly lower tumour volumes at the end of the experiments. Overexpression of miR-26b in nasopharyngeal carcinoma cell lines showed reduced growth and colony formation **(Ji et al., 2010)**. Interestingly, in pituitary cancer cell lines, inhibition of miR-26b expression showed reduced invasion and colony formation **(Palumbo et al., 2012)**.

Although this finding in pituitary cancer is contradictory to the other studies it is compatible with the earlier assertion that miR-26b might have differing roles in different cell types. This is the only study where the authors performed knockdown experiments with miR-26b. The results from this study were contradictory to those I obtained in fibroblasts (**Figure 5.4 & 5.7**). However these were performed in pituitary epithelial cell lines, which will have a transcriptome that is different from fibroblasts, and therefore no doubt contain a different pool of mRNAs for miR-26b to act upon.

In breast cancer, functional studies were performed in the MCF-7 cell line, as I have used for my luminal breast cancer epithelial model line, and in the MDA-MB-231 cell line, which is a more invasive breast cancer cell line and represents the triple negative/basal type of breast cancer. Overexpression of miR-26b in cells caused a decrease in cell viability and increased apoptosis and overexpression of miR-26b in MDA-MB-231 resulted showed reduced proliferation **(Li et al., 2013; Liu et al., 2011)**. Most studies that overexpressed miR-26b and examined cell viability, proliferation or apoptosis showed results that were consistent with my initial findings in fibroblasts, when I overexpressed miR-26b (**Figure 5.1 & 5.2**).

7.1.3 TARGETS OF MIR-26B

To date 14 targets of miR-26b have been experimentally verified (**Table 7.1**). As is evident from the table, degrees of experimental verification associated with these claimed targets are highly variable.

It is interesting to note that in most studies there does not always appear to be a coherent reason as to why a particular target was chosen and followed up. Some studies state that they identified the targets by selecting commonly predicted ones from multiple in-silico prediction tools. A few studies applied a further filter of target

conservation across species to achieve greater “biological significance”. However, I did not find identification of the true targets of miR-26b in fibroblasts straightforward. In-silico target prediction tools predicted a large number of targets for miR-26b with relatively little overlap between different prediction tools. Examining only the targets that are conserved across different species has the inherent problem that such a strategy will miss potentially important targets that are unique to the species being studied. In addition, these targets are only a theoretical possibility and it is more than likely that the actual true targets of miR-26b will depend on the cell type, its molecular profile and the micro-environmental conditions that it sits in.

PTEN is the only miR-26b target that has been shown to be involved in migration/invasion of cells. This was seen in pituitary tumour cell lines. PTEN was confirmed to be a direct target of miR-26b using luciferase reporter assays (**Table 7.1**). Although the targeting was not confirmed at the protein level, siRNA knockdown of PTEN showed a decrease in invasion of pituitary tumour cells. This was similar to the result obtained when miR-26b was overexpressed in these cell lines, therefore the authors claimed that PTEN must be a mediator of miR-26b’s functional influence.

Target	Functional role of miR-26b		Target identification	Cell type	Length of 3'UTR in luciferase reporter (bp) (O & K), Control reporter	Western blots (O & K)	Target linked to miR-26b function
	Knockdown (K)	Overexpression (O)					
GATA4 **(Han et al., 2012)**	-l	hypertrophy of myocytes	TargetScan	myocytes	23 x 4 (O & K), No	Yes (O & K)	NA
PLCβ1 **(Han et al., 2012)**	-	hypertrophy of myocytes	TargetScan	myocytes	161 (O), No	Yes (O)	NA
PTEN **(Palumbo et al., 2012)**	decreases invasion, colony formation	-	miRBase, miRanda and TargetScan	pituitary tumour cell lines	"3' UTR". (O), No	No	Yes
ATF2 **(Arora et al., 2011)**	-	-	DIANA	lung cancer cell lines	"3' UTR", (O) No	Yes (O)	NA
ATM **(Lin et al., 2012)**	-	increases apoptosis	Bioinformatics tools (names not mentioned	porcine granulosa cells	27 (O), Yes	No	No
CTDSP2 **(Dill et al., 2012)**	impairs neurogenesis	-l	NA	zebra fish	"3' UTR", (O), Yes	Yes (O)	No
BDNF **(Caputo et al., 2011)**	-	-	NA	HeLa	"3' UTR", (O), Yes	No	NA
COX2	-	decreases proliferation and colony	Literature	carcinoma of nasopharyngeal epithelial cells	"3' UTR", (O), Yes	No	No
DNMT3b **(Sandhu et al., 2012)**	-	-	NA	breast cancer	No	No	NA

LEF1 **(Zhang et al., 2010)**	-	-	NA	pituitary tumour cell lines	“3’ UTR”, (O) Yes	Yes (O)	Nil
EphA2 **(Wu et al., 2011)**	-	inhibits proliferation, migration and invasion	miRanda, TargetScan, PicTar	glioma cell lines	“3’ UTR” (O), Yes	Yes (O)	No
SLC7A11 **(Liu et al., 2011)**	-	increases apoptosis	Bioinformatics tools (names not mentioned	breast cancer cell lines	3971 (O & K), Yes	Yes (O & K)	Yes
EZH2 **(Koh et al., 2011)**	-	increases apoptosis	NA	prostate cancer cell line	“3’ UTR”, (O), Yes	No	No
PTGS2 **(Li et al., 2013)**	-	inhibits proliferation	Bioinformatics tools (names not mentioned	breast cancer cell line	“3’ UTR”, (O), Yes	Yes (O)	Yes
CDK6 **(Zhu et al., 2012)**	-	cell cycle arrest	NA	liver cancer cell lines	“3’ UTR fragment”, (O), Yes	Yes (O)	Yes
Cyclin E1 **(Zhu et al., 2012)**	-	cell cycle arrest	NA	liver cancer cell lines	“3’ UTR fragment”, (O), Yes	Yes (O)	Yes

Table 7.1. miR-26b targets that have been verified *in-vitro*. Details of experimental validation are listed including the type of miR-26b manipulation and its functional readout, cell type in which the target verification was performed, type of verification that was performed (either luciferase reporter assay or Western blots) and whether manipulation of the proposed target phenocopied manipulation of expression miR-26b.

7.2 TARGETS AND PATHWAYS REGULATED BY MIR-26B IN BREAST FIBROBLASTS

By using mass spectrometry I have identified 3 potential targets of miR-26b and 2 distinct pathway that might be regulated downstream of miR-26b. I will discuss the potential roles of these molecules and pathways in the context of cancer in the following sections. However, it is to be kept in mind that these pathways as targets of miR-26b is subject to questions related to the ontology enrichment method used - as discussed in section 6.7.

7.2.1 COL-XII

Collagen XII is present in the extracellular matrix of different tissues such as skin, bone and tendon **(Berthod et al., 1997; Izu et al., 2011; Posthumus et al., 2010)**. It exists as a homotrimer of COL12A1 molecules. Collagen XII binds to collagen 1 and helps to stabilise the collagen I fibrils **(Gerecke et al., 1997)**. This interaction can modify the extracellular matrix; for example in fibrosis, interactions between collagen XII and collagen I have been shown to occur at an early stage and might be responsible for rigidity of the fibrous tissue **(Nishiyama et al., 1994; Tzortzaki et al., 2006)**.

Up-regulation of collagen XII occurs in various conditions such as corneal scaring, systemic sclerosis and idiopathic pulmonary fibrosis, where interestingly enough the underlying pathological process is fibrosis **(Massoudi et al., 2012; Moinzadeh et al., 2013; Nishiyama et al., 1994; Tzortzaki et al., 2006)**. However, there is very little literature on the role of Collagen XII in cancer. Below, I describe all the studies where Collagen XII has been show to play a role in tumours.

In chondromyxoid fibroma, a benign tumour of the bone, rearrangements of 6q13 locus are relatively frequent (up to 50%). Using a FISH (fluorescence in-situ hybridisation) based positional cloning strategy; the breakpoint in this locus was mapped to the COL12A1 gene, suggesting its involvement in the development of chondromyxoid fibroma **(Yasuda et al., 2009)**. However, functional studies were not performed and furthermore it is not known if collagen XII is over or under expressed in these tumours compared to normal tissue. Myxofibrosarcoma is a malignant soft tissue tumour, with abundant extracellular mucin, that is commonly seen in the extremities of elderly people. Collagen XII was shown to be present in these tumours but not in myxomas (benign soft tissue tumours that also have abundant extracellular mucin) using mass spectrometry **(Willems et al., 2009)**. This was subsequently confirmed by immunohistochemistry and qPCR. In colorectal cancer, increased expression of collagen XII was associated with the progression of normal colonic mucosa (NC) to colonic adenoma (AD), and onwards to colorectal adenocarcinoma (AC) **(Mikula et al., 2010)**. Using a combination of mass spectrometry and gene expression profiling, proteins/genes that were progressively up or down-regulated in the NC/AD/AC sequence were identified. After validation using qRT-PCR, collagen XII was the only protein/gene that was seen to be progressively up-regulated in the NC/AD/AC sequence. Increased expression of collagen XII in colorectal cancer was also seen in another study. In this work, the increased expression of collagen XII was seen mainly in CAFs that were present at the leading edge of invasive tumours **(Karagiannis et al., 2012)**. However, in all the above studies no functional experiments were performed to examine if collagen XII deregulation was pathogenic and drove tumour progression or was simply a reactive phenomenon with no functional consequences **(Willems et al., 2009; Yasuda et al., 2009)**.

Recessive Dystrophic Epidermolysis Bullosa (RDEB) is a skin condition associated with a mutation in another collagen gene, COL7A1. Patients with this condition are prone to aggressive squamous cell carcinoma of the skin. Messenger RNA expression profiles of “normal” skin fibroblasts (RDEB fibroblasts) from these patients were similar to the

profile of squamous cell carcinoma CAFs (SCC CAFs) **(Ng et al., 2012)**. The RDEB fibroblasts and SCC CAFs showed up-regulation of collagen XII compared to fibroblasts from non-REDB patients. Overexpression of COL7A1 in RDEB fibroblasts modified the extracellular matrix composition by decreasing the production of collagen XII. Furthermore, re-expression of collagen COL7A1 in these fibroblasts slowed cancer cell invasion and in-vivo tumour progression. When I knocked down miR-26b in immortalised breast fibroblasts, the $26b^{k/d}$ fibroblasts also showed increased expression collagen XII (Section 6.3) and an ability to enhance cancer cell invasion, demonstrating the same correlation between stromal collagen XII expression and epithelial invasion.

Although the evidence from the above experiments is indirect, there appears to be a link between increased expression of collagen XII by fibroblasts and cancer epithelial cell invasion. This is further supported by outcome data in breast cancer patients where higher expression of collagen XII in the tumour stroma of breast cancer is associated with decreased recurrence free survival (**Figure 6.4**). However, there is no other evidence to suggest that increased production of collagen XII in CAFs leads to an increase in the motility of the CAFs themselves, although it potentially increases the motility of epithelial cells.

There is evidence in the literature pointing to the potential mechanism by which collagen XII can increase invasion of cancer epithelial cells. This is further discussed in section 7.3.

7.2.2 CPSF7 & TNKS1BP1

There is only a very limited literature on CPSF7. It is one of the components of ubiquitous machinery that is involved in the endonucleotide specific cleavage and polyadenylation of 3' ends of mRNAs. There are two main protein complexes involved in

this process, CPSF (Cleavage and polyadenylation specificity factor) and CSF (cleavage stimulatory factor). In addition to this, two cleavage factors (CFs) CFIm and CFIIm are required. CPSF7 (or CFIm59) is one of the subunits of CFIm **(Kim et al., 2010)**.

Polyadenylation is a normal process in all cells. However the phenomenon of Alternative PolyAdenylation (APA) – a process, by which the length of transcript is altered at the 3' end by using alternate sites in the mRNA transcript for polyadenylation, is thought to play a role in cancer. Cancer-associated APA commonly results in shorter 3' UTR. Therefore widespread APA in cancer cells can results in increased proliferation due to shorter 3' UTR for proto-oncogenes that can escape the translational repression by miRNAs. CFIm regulates APA by controlling the sites at which polyadenylation occurs. Interestingly, the presence of CFIm results in longer transcripts **(Elkon et al., 2013)**, which may cause reduced growth. Although it is tempting to speculate that the reduced growth seen in fibroblasts knocked down with miR-26b could be a result of increased CFIm activity through increased levels of CPSF7, current evidence suggest that this subunit is functionally redundant **(Kim et al., 2010)**. However, this might not be the case in fibroblasts and in addition these factors may have other roles that are as yet unknown.

The literature on TNKS1BP1 is also limited. TNKS1BP1 binds to the Tankyrase 1 **(Seimiya and Smith, 2002)**. Tankyrase 1 is a member of the poly ADP-Ribose polymerases (PARPs), and is involved in the regulation of telomere length. In cancer cells, Tankyrase 1 is a potential therapeutic target, since inhibition of Tankyrase 1 can result in shortening of telomere length and cell death **(Seimiya et al., 2005)**. However the role of TNKS1BP1 in this context is unknown. Interestingly, TNKS1BP1 has a subcellular localisation not only in the nucleus but also in the cytoplasm. The cytoplasmic staining is co-localised to cortical actin that is in close contact with the plasma membrane **(Seimiya and Smith, 2002)**. The localisation is particularly prominent at the sites of cell-to-cell contact. Although the exact role of TNKS1BP1 is

unknown it is appealing to speculate that it has a role in migration of cells because of its association with actin.

Both these molecules are up-regulated in $26b^{k/d}$ fibroblasts and their stromal expression can predict recurrence. However, they do not appear to be direct targets of miR-26b and further study is required to decipher their roles in CAFs in breast cancer.

7.2.3 GLYCOLYSIS AND TCA CYCLE

Glycolysis is the primary metabolic pathway by which cells are able to utilise glucose (**Figure 7.1**). In addition to ATP glycolysis also generates pyruvate as the end product. Pyruvate in normal cells is further processed in the mitochondria by aerobic respiration to generate chemically-accessible energy for the cell in the form of ATP. The series of reactions that takes place in the mitochondria is called the tricarboxylic acid cycle (TCA) cycle (**Figure 7.2**). Occasionally in some cells, such as skeletal myocytes, glucose undergoes glycolysis followed by lactic acid fermentation in the cytoplasm. This is a less energy efficient process compared to the TCA cycle, since it does not result in further energy production. Cancer cells are thought to be different from normal cells in that most of the energy source comes from a high rate of glycolysis rather than the TCA cycle. This has been termed the “Warburg” effect. Instead of entering the TCA cycle, pyruvate, which is the end product of glycolysis, undergoes lactic acid fermentation. The Warburg effect may be a response to the hypoxic environment of tumours since glycolysis and lactic acid fermentation does not require oxygen whereas TCA cycle requires this **(Gasparre et al., 2013)**.

In recent years a new model has been proposed in breast cancer, termed the “Reverse Warburg effect” **(Sotgia et al., 2012)**. This states that the primary cell type that is responsible for the high rate of glycolysis is the CAFs. Pyruvate is transferred to the

adjacent cancer epithelial cells via the monocarboxylate transporters where it undergoes aerobic respiration **(Martinez-Outschoorn et al., 2010)**.

The results of the pathway analysis for the differentially regulated proteins in $26b^{k/d}$ fibroblasts are entirely compatible with this model. This analysis identified the glycolytic pathway to be enriched. And importantly all the components that were identified were up-regulated (**Figure 7.1**). This indicates that this pathway might be overactive, resulting in increased pyruvate production, which can then be used by the cancer epithelial cells. However, in the *in-vitro* system where the difference in protein expression was examined, the fibroblasts were present on their own, without the cancer epithelial cells. Therefore it is likely that the increased pyruvate being produced by the overactive glycolysis pathway is utilised by the fibroblasts themselves. Again there is evidence to suggest that this is the case, since the pathway analysis also identified a "superpathway" containing glycolysis pathway, pyruvate dehydrogenase and components of TCA cycle to be enriched. Here again, all the components that were identified were up-regulated.

There is evidence, at least in some cancer cells, that increased migration of cells can be a consequence of increased rate of glycolysis **(Liu et al., 2013; Sottnik et al., 2011)**. For example, addition of glucose to pancreatic cancer cell lines resulted in increased migration in hypoxic conditions. Similarly, pharmacological inhibition of glycolysis using 2-deoxy-D-Glucose (2DG) in osteosarcoma cell lines reduced migration and invasion. Therefore, the overactive glycolytic pathway might explain the increased rate of migration of $26b^{k/d}$ fibroblasts (**Figure 5.3 & 5.7)**. However, the increased motility of cancer epithelial cells is difficult to explain solely on the basis the reverse Warburg effect, since the increased motility was seen only in transwell migration assays and not in indirect co-culture assays or transwell invasion assays.

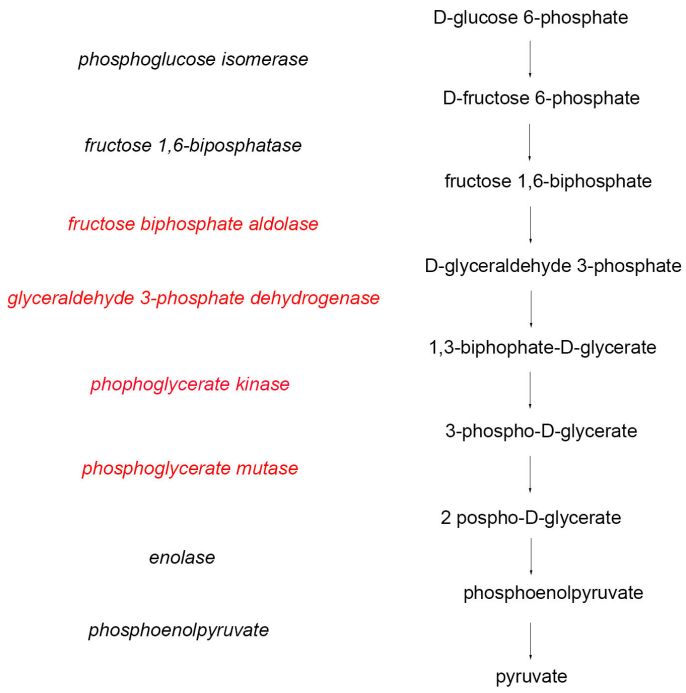

Figure 7.1. Flow chart showing the various intermediate compounds formed during the glycolysis pathway. The enzymes responsible for converting one compound to the next are listed left to the flowchart. Those enzymes that are highlighted in red were identified as up-regulated in 26b$^{k/d}$ as compared to con$^{k/d}$ fibroblasts in the pathway analysis. Note: one of the up-regulated components, phosphoglucomutase 1 that converts glucose1-phosphate into glucose-6-phosphate, is not represented in this flowchart, but was identified to be part of this in the pathway analysis.

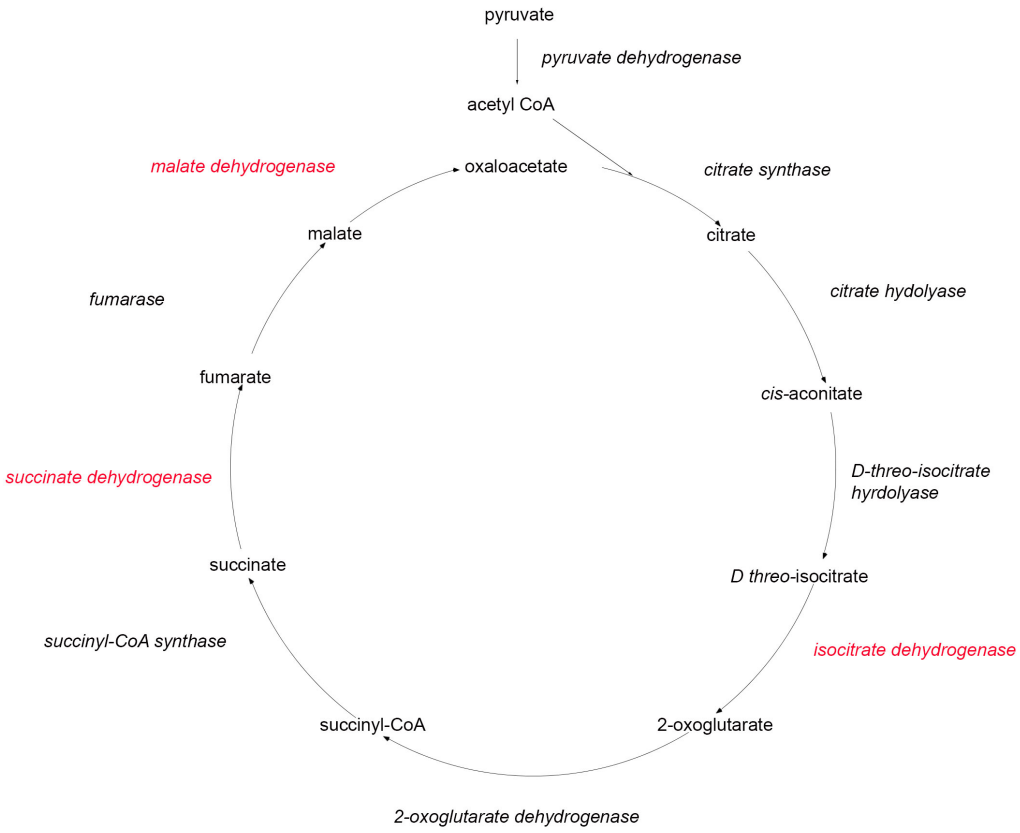

Figure 7.2. Flow chart showing the various intermediate compounds formed during the TCA cycle. The enzymes responsible for converting one compound to the next are listed on the outside of the circle. Those enzymes that are highlighted in red were identified as up-regulated in 26b$^{k/d}$ as compared to con$^{k/d}$ fibroblasts in the pathway analysis.

7.2.4 CYTOSKELETON REGULATION BY RHO GTPASE PATHWAY

The cytoskeleton of the cell consists of thin filaments, intermediate filaments and microtubules. Many different cellular functions are dependent on this cytoskeleton. The thin filaments are composed of actin and are particularly relevant to my findings since they have an important role in the motility of cells. There is a wide variety of actin binding proteins, which include a number of regulatory molecules that control actin mediated cell motility.

Of these regulatory molecules, members of the RHO GTPase family have been studied extensively. This family consists of 21 molecules that have been sub-classified into 6 major subfamilies **(Bustelo et al., 2007)**. These groups are the Rac, Cdc42, Rho, Rnd, RhoT and RhoBTB subfamilies. These molecules are all G proteins and act as molecular switches. They exist in an active form that is bound to GTP and an inactive form that is bound to GDP **(Figure 7.3**). The switching between the active and inactive states is regulated by a number of regulatory molecules that are grouped into three classes: guanine nucleotide exchange factor (GEFs); GTPase-activating proteins (GAPs); and guanine nucleotide dissociation inhibitors (GDIs). The active form of RHO-GTPases acts on a number of effector molecules that regulate the cell cytoskeleton. Molecules identified as differentially expressed in my pathway analysis included those belonging to various different groups: GAP - ARHGAP1; effector molecules - STMN1, ROCK2, ARPC5, PFN1, CFL1, CFL2, ARPC1A and PAK1; constituent of cytoskeleton - TUBB6, MYH10, ACTBL2, ACTG1 and ACTB.

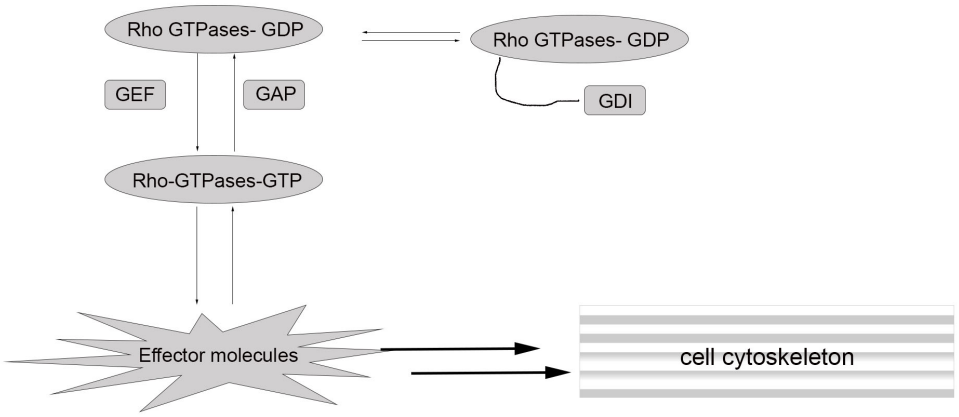

Figure 7.3. RHO GTPase pathway. GEFs (guanine nucleotide exchange factor), GAPs (GTPase-activating proteins), GDIs (guanine nucleotide dissociation inhibitors)

Members of the Rac, Rho and Cdc42 subfamilies have been studied in great detail in the context of cell motility **(Nobes and Hall, 1995)**. Rac activity results in the formation of “focal complexes” and polymerization of actin into lamellipodia at the leading edge of the cell. Cdc42 activity induces actin polymerization and assembly of filopodia. Activity of Rho results in the formation of “stress fibres”, formed by binding of actin by the ATP-depended motor protein myosin. Rac, Rho and Cdc42 are all involved in formation of focal adhesion complexes **(Nobes and Hall, 1995)**. Lamellipodia and filopodia are actin containing cytoplasmic extensions of the leading edge of cells. Focal adhesion complexes are clusters of integrin molecules that are linked to various components of the extracellular matrix and to intracellular components such as actin. The lamellipodia,

filopodia, stress fibres and focal adhesion complexes are all integral to cell motility. Although a number of these molecules have established roles in the motility of cells, they are also involved in a wide variety of cellular mechanism such as vesicular trafficking, cell cycle and gene expression.

There is evidence for involvement of different members of RHO GTPase pathway in the breast cancer invasion and metastasis **(Jiang et al., 2009; Keely et al., 1997)**. These studies focus on the roles of this pathway in epithelial cells, and its influence on epithelial cell motility. I could not identify any studies in which the role of RHO GTPase pathways in CAFs and their influence on breast cancer epithelial cell motility was examined. However, there is emerging evidence from other tumour types such cholangiocarcinoma and squamous cell carcinoma, that this pathway is important in the motility of CAFs and this in turn influences the motility of cancer epithelial cells. For example inhibition of small RHO GTPases decreased migration of CAFs in cholangiocarcinoma. However, the role of these fibroblasts in the invasion of epithelial cells was not examined. In squamous cell carcinoma inhibition of ROCK prevented the fibroblast mediated matrix remodelling that helped in the invasion of epithelial cells (Section 7.3)

The pathway analysis indicates that components of RHO GTPase that have a role in cytoskeleton regulation are deregulated. Given the key role of this pathway in cell motility it is possible that these changes could have resulted in the increased motility of $26b^{k/d}$ fibroblasts. Although it is tempting to examine the direction of change and speculate on the potential roles for each member that has been identified, this is likely to be an oversimplification of a very complex pathway and unlikely to yield much insight. Furthermore, members of the above pathways are not predicted targets of miR-26b. Nevertheless, it is likely that changes in this pathway are downstream effects of one or many miR-26b targets. A detailed study of all potential targets of miR-26b identified on the mass spectrometry screen will be required to understand this further.

The increased motility of $26b^{k/d}$ fibroblasts could contribute to the increased motility of cancer epithelial cells, but as discussed below (section 7.3) this is unlikely to be the sole reason for enhanced epithelial invasion. This is further explored in section 7.3.

7.3 MOTILITY OF CANCER EPITHELIAL CELLS

Migration of cells is a multistep process that involves multiple components. The commonly recognised model includes the following steps: (i) polymerisation of the actin cytoskeleton at the leading edge of the cell; (ii) interaction with the surrounding matrix through various integrins by formation of focal adhesion complexes; (iii) protease digestion of matrix; (iv) contraction of cytoskeletal structure of the cell resulting in the forward movement of the cell; and (v) resolution of these events at the trailing edge. This model is also relevant to cancer cell migration **(Friedl and Wolf, 2003)**.

Over the years various pathways have been identified in breast cancer epithelial cells that regulate the different steps in the above model and subsequently affect cellular motility **(Jiang et al., 2009)**. In addition to this plethora of changes in epithelial cells, various components of the surrounding microenvironment also have an influence on the motility of breast cancer epithelial cells **(Al-Ansari et al., 2013; Cheng et al., 2007; Goswami et al., 2005; Studebaker et al., 2008)**. The most commonly studied component of the microenvironment is the CAFs and their ability to influence the motility of cancer epithelial cells. In most of these studies the commonly described mechanism for CAFs to influence the epithelial cells is the secretion of soluble factors by the CAFs that promote the invasive capacity of cancer epithelial cells. For example, p16 deficient fibroblasts promote the migration and invasion of breast cancer epithelial cells through secretion of SDF-1 **(Al-Ansari et al., 2013)**. Similarly IL-6 secreted by fibroblasts promotes growth and invasion of breast cancer cells by up-regulation of Notch-3, Jagged-1, and carbonic anhydrase IX **(Studebaker et al., 2008)**. Here, I have shown that the presence of $26^{k/d}$ fibroblasts in a 3D spheroidal invasion assay model

contributes to the migratory capacity of the cancer epithelial cells. However, this is unlikely to be solely related to a soluble factor since there was no difference in migration of cancer epithelial cells in the indirect co-culture assays (**Figure 5.9**).

Another possible explanation is that a more motile cell might be dragging a less motile cell as it migrates into the surrounding tissue. This has been shown to be the case in melanoma where a group of highly motile melanoma cells in tumours, termed “guiding cells” generate migratory traction, and pull the trailing cells along **(Hegerfeldt et al., 2002)**. The more motile $26^{k/d}$ fibroblasts could be dragging the cancer epithelial cells along. Therefore, a larger proportion of cancer epithelial cells will be seen to migrate. This could explain the result seen in the direct transwell migration assays. However, it is unlikely to be the sole reason for the increased invasion of epithelial cells, since there was no significant increase in the invasion of cancer epithelial cells in 2D direct invasion assays (**Figure 5.9).**

The key difference between these migration and invasion assays is the presence of extracellular matrix components. CAFs have been shown to generate tracks in the ECM through which cancer epithelial cells can travel. For example in squamous cell carcinoma the leading cells in collective cell migration are always CAFs with the cancer epithelial cells following the tracks generated by CAFs **(Gaggioli et al., 2007)**. The results from the 3D spheroidal invasion assays where $26^{k/d}$ fibroblasts are present at the leading edges of invasive fronts (**Figure 5.10**) are consistent with this observation. Therefore, the more motile $26^{k/d}$ fibroblasts could generate more tracks through which epithelial cells can migrate more easily. However, this model alone fails to explain the difference in migration of epithelial cells between the 3D spheroidal invasion and direct transwell invasion assays (**Figure 5.9 & 10**).

This suggests that the mechanism involved in the CAF mediated migration of cancer epithelial cells might be more complex and/or the type of extracellular matrix might also be relevant. There is emerging evidence that matrix remodelling involves not only degradation of matrix by protease but also deposition of new collagen and formation of new cross-links **(Levental et al., 2009)**. Thickened matrix caused by increased collagen

cross-linking can promote cancer epithelial cell motility through enhanced integrin signalling. In the study by Gaggioli et al, the authors noted that the tracks laid out by CAFs were characterised by thick bundles of collagen in the periphery of the tracks. Therefore, the tracks formed by CAFs not only had a central area that was free of collagen but also more firm areas in the periphery that allowed the cancer epithelial cells to form multiple adhesion complexes that eventually resulted in integrin mediated epithelial cell motility. The specific components of the extracellular matrix are critical in this context since this cross-linking is depended on their chemistry. In the transwell invasion assay, Matrigel, which contains collagen IV, was used to model the ECM. Whereas in the 3D spheroidal invasion assay, collagen I was the predominant ECM component. Collagen XII, one of the up-regulated proteins in $26^{k/d}$ fibroblasts, can form cross-links with collagen I and this interaction can result in changes in the biomechanical properties of tissue, resulting in increased stiffness (Section 7.2.1). However, this crosslinking would occur only in the 3D spheroidal invasion assay model (containing collagen I) and not in the transwell invasion assay model (lacking collagen I). Hence one hypothesis is that that collagen XII, one of the direct targets of miR-26b that is overexpressed in $26b^{k/d}$ fibroblasts, forms crosslinks with collagen I which in turn promote motility of epithelial cells by increased integrin signalling.

7.4 SUMMARY

This study has provided further understanding into the role of the stroma, particularly the role of fibroblasts in the biology of breast cancer. I have shown that expression of miR-26b is different between normal fibroblasts and CAFs and miR-26b influences migration and invasion of CAFs. Fibroblasts that are knocked down for miR-26b increases the invasion of cancer epithelial cells. In addition, I have also identified potential downstream targets of miR-26b in CAFs that might explain this behaviour of cancer epithelial cells. These findings support my primary hypothesis that changes in the level of miRNAs regulate the behaviour of carcinoma-associated fibroblasts (CAFs) and these CAFs in turn regulate the behaviour of malignant epithelial cells.

However, many questions remain unanswered at present and the areas of focus for further research will be the mechanisms by which CAFs migrate/invade faster and the mechanism by which the CAFs promote the invasion of epithelial cells. It is attractive to consider that enhanced understanding of this process might lead to novel therapeutic strategies based on targeting the invasion promoting CAFs.

8 REFERENCES

Adachi, J., Kumar, C., Zhang, Y., and Mann, M. (2007). In-depth analysis of the adipocyte proteome by mass spectrometry and bioinformatics. Molecular & cellular proteomics : MCP *6*, 1257-1273.

Aerts, S., Lambrechts, D., Maity, S., Van Loo, P., Coessens, B., De Smet, F., Tranchevent, L. C., De Moor, B., Marynen, P., Hassan, B.*, et al.* (2006). Gene prioritization through genomic data fusion. In Nat Biotechnol, (United States), pp. 537-544.

Al-Ansari, M. M., Hendrayani, S. F., Shehata, A. I., and Aboussekhra, A. (2013). p16(INK4A) represses the paracrine tumor-promoting effects of breast stromal fibroblasts. Oncogene *32*, 2356-2364.

Allinen, M., Beroukhim, R., Cai, L., Brennan, C., Lahti-Domenici, J., Huang, H., Porter, D., Hu, M., Chin, L., Richardson, A.*, et al.* (2004). Molecular characterization of the tumor microenvironment in breast cancer. Cancer Cell *6*, 17-32.

Aprelikova, O., Yu, X., Palla, J., Wei, B. R., John, S., Yi, M., Stephens, R., Simpson, R. M., Risinger, J. I., Jazaeri, A., and Niederhuber, J. (2010). The role of miR-31 and its target gene SATB2 in cancer-associated fibroblasts. Cell cycle (Georgetown, Tex *9*, 4387-4398.

Arora, H., Qureshi, R., Park, A. K., and Park, W. Y. (2011). Coordinated regulation of ATF2 by miR-26b in gamma-irradiated lung cancer cells. PLoS One *6*, e23802.

ATCC (2014).

Baek, D., Villen, J., Shin, C., Camargo, F. D., Gygi, S. P., and Bartel, D. P. (2008). The impact of microRNAs on protein output. Nature *455*, 64-71.

Barber, R. D., Harmer, D. W., Coleman, R. A., and Clark, B. J. (2005). GAPDH as a housekeeping gene: analysis of GAPDH mRNA expression in a panel of 72 human tissues. Physiological genomics *21*, 389-395.

Barcellos-Hoff, M. H., and Ravani, S. A. (2000). Irradiated mammary gland stroma promotes the expression of tumorigenic potential by unirradiated epithelial cells. Cancer Res *60*, 1254-1260.

Bartel, D. P. (2009). MicroRNAs: target recognition and regulatory functions. Cell *136*, 215-233.

Bauer, M., Su, G., Casper, C., He, R., Rehrauer, W., and Friedl, A. (2010). Heterogeneity of gene expression in stromal fibroblasts of human breast carcinomas and normal breast. In Oncogene, (England), pp. 1732-1740.

Bergamaschi, A., Tagliabue, E., Sorlie, T., Naume, B., Triulzi, T., Orlandi, R., Russnes, H. G., Nesland, J. M., Tammi, R., Auvinen, P*., et al.* (2008). Extracellular matrix signature identifies breast cancer subgroups with different clinical outcome. J Pathol *214*, 357-367.

Berthod, F., Germain, L., Guignard, R., Lethias, C., Garrone, R., Damour, O., van der Rest, M., and Auger, F. A. (1997). Differential expression of collagens XII and XIV in human skin and in reconstructed skin. J Invest Dermatol *108*, 737-742.

Bertucci, F., and Birnbaum, D. (2008). Reasons for breast cancer heterogeneity. J Biol *7*, 6.

Biosystems, A. (2007). Application note TaqMan® MicroRNA Assays. Application note, 8.

Blows, F. M., Driver, K. E., Schmidt, M. K., Broeks, A., van Leeuwen, F. E., Wesseling, J., Cheang, M. C., Gelmon, K., Nielsen, T. O., Blomqvist, C*., et al.* (2010). Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype

and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med *7*, e1000279.

Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C. A., Causton, H. C.*, et al.* (2001). Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet *29*, 365-371.

Bruchova, H., Merkerova, M., and Prchal, J. T. (2008). Aberrant expression of microRNA in polycythemia vera. Haematologica *93*, 1009-1016.

Buchanan, C. F., Szot, C. S., Wilson, T. D., Akman, S., Metheny-Barlow, L. J., Robertson, J. L., Freeman, J. W., and Rylander, M. N. (2012). Cross-talk between endothelial and breast cancer cells regulates reciprocal expression of angiogenic factors in vitro. J Cell Biochem *113*, 1142-1151.

Bustelo, X. R., Sauzeau, V., and Berenjeno, I. M. (2007). GTP-binding proteins of the Rho/Rac family: regulation, effectors and functions in vivo. Bioessays *29*, 356-370.

Buyse, M., Loi, S., van't Veer, L., Viale, G., Delorenzi, M., Glas, A. M., d'Assignies, M. S., Bergh, J., Lidereau, R., Ellis, P.*, et al.* (2006). Validation and clinical utility of a 70-gene prognostic signature for women with node-negative breast cancer. J Natl Cancer Inst *98*, 1183-1192.

Camp, J. T., Elloumi, F., Roman-Perez, E., Rein, J., Stewart, D. A., Harrell, J. C., Perou, C. M., and Troester, M. A. (2011). Interactions with fibroblasts are distinct in Basal-like and luminal breast cancers. Mol Cancer Res *9*, 3-13.

CancerResearchUK http://info.cancerresearchuk.org/cancerstats

Caputo, V., Sinibaldi, L., Fiorentino, A., Parisi, C., Catalanotto, C., Pasini, A., Cogoni, C., and Pizzuti, A. (2011). Brain derived neurotrophic factor (BDNF) expression is

regulated by microRNAs miR-26a and miR-26b allele-specific binding. PLoS One *6*, e28656.

Chang, H. Y., Nuyten, D. S., Sneddon, J. B., Hastie, T., Tibshirani, R., Sorlie, T., Dai, H., He, Y. D., van't Veer, L. J., Bartelink, H.*, et al.* (2005). Robustness, scalability, and integration of a wound-response gene expression signature in predicting breast cancer survival. Proceedings of the National Academy of Sciences of the United States of America *102*, 3738-3743.

Chen, J., Xu, H., Aronow, B. J., and Jegga, A. G. (2007). Improved human disease candidate gene prioritization using mouse phenotype. In BMC Bioinformatics, (England), p. 392.

Chendrimada, T. P., Gregory, R. I., Kumaraswamy, E., Norman, J., Cooch, N., Nishikura, K., and Shiekhattar, R. (2005). TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature *436*, 740-744.

Cheng, N., Bhowmick, N. A., Chytil, A., Gorksa, A. E., Brown, K. A., Muraoka, R., Arteaga, C. L., Neilson, E. G., Hayward, S. W., and Moses, H. L. (2005). Loss of TGF-beta type II receptor in fibroblasts promotes mammary carcinoma growth and invasion through upregulation of TGF-alpha-, MSP- and HGF-mediated signaling networks. Oncogene *24*, 5053-5068.

Cheng, N., Chytil, A., Shyr, Y., Joly, A., and Moses, H. L. (2007). Enhanced hepatocyte growth factor signaling by type II transforming growth factor-beta receptor knockout fibroblasts promotes mammary tumorigenesis. Cancer Res *67*, 4869-4877.

Coombe, D. R., Nakhoul, A. M., Stevenson, S. M., Peroni, S. E., and Sanderson, C. J. (1998). Expressed luciferase viability assay (ELVA) for the measurement of cell growth and viability. Journal of immunological methods *215*, 145-150.

Coussens, L. M., and Werb, Z. (2002). Inflammation and cancer. Nature *420*, 860-867.

Cox, B., Kotlyar, M., Evangelou, A. I., Ignatchenko, V., Ignatchenko, A., Whiteley, K., Jurisica, I., Adamson, S. L., Rossant, J., and Kislinger, T. (2009). Comparative systems biology of human and mouse as a tool to guide the modeling of human placental pathology. Molecular systems biology *5*, 279.

Danielsson, K., Ebrahimi, M., Wahlin, Y. B., Nylander, K., and Boldrup, L. (2012). Increased levels of COX-2 in oral lichen planus supports an autoimmune cause of the disease. Journal of the European Academy of Dermatology and Venereology : JEADV *26*, 1415-1419.

Davis, E., Caiment, F., Tordoir, X., Cavaille, J., Ferguson-Smith, A., Cockett, N., Georges, M., and Charlier, C. (2005). RNAi-mediated allelic trans-interaction at the imprinted Rtl1/Peg11 locus. Curr Biol *15*, 743-749.

DeCosse, J. J., Gossens, C., Kuzma, J. F., and Unsworth, B. R. (1975). Embryonic inductive tissues that cause histologic differentiation of murine mammary carcinoma in vitro. J Natl Cancer Inst *54*, 913-922.

Diab, S. G., Clark, G. M., Osborne, C. K., Libby, A., Allred, D. C., and Elledge, R. M. (1999). Tumor characteristics and clinical outcome of tubular and mucinous breast carcinomas. J Clin Oncol *17*, 1442-1448.

Dill, H., Linder, B., Fehr, A., and Fischer, U. (2012). Intronic miR-26b controls neuronal differentiation by repressing its host transcript, ctdsp2. Genes Dev *26*, 25-30.

Dirat, B., Bochet, L., Dabek, M., Daviaud, D., Dauvillier, S., Majed, B., Wang, Y. Y., Meulle, A., Salles, B., Le Gonidec, S.*, et al.* (2011). Cancer-associated adipocytes exhibit an activated phenotype and contribute to breast cancer invasion. Cancer Res *71*, 2455-2465.

Doi, N., Zenno, S., Ueda, R., Ohki-Hamazaki, H., Ui-Tei, K., and Saigo, K. (2003). Short-interfering-RNA-mediated gene silencing in mammalian cells requires Dicer and eIF2C translation initiation factors. Curr Biol *13*, 41-46.

Duffy, M. J., Maguire, T. M., Hill, A., McDermott, E., and O'Higgins, N. (2000). Metalloproteinases: role in breast carcinogenesis, invasion and metastasis. Breast Cancer Res *2*, 252-257.

Elkon, R., Ugalde, A. P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet *14*, 496-506.

Ellis, I. O., Galea, M., Broughton, N., Locker, A., Blamey, R. W., and Elston, C. W. (1992). Pathological prognostic factors in breast cancer. II. Histological type. Relationship with survival in a large study with long-term follow-up. Histopathology *20*, 479-489.

Elston, C. W., and Ellis, I. O. (1991). Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology *19*, 403-410.

Fan, C., Oh, D. S., Wessels, L., Weigelt, B., Nuyten, D. S., Nobel, A. B., van't Veer, L. J., and Perou, C. M. (2006). Concordance among gene-expression-based predictors for breast cancer. N Engl J Med *355*, 560-569.

Fiegl, H., Millinger, S., Goebel, G., Muller-Holzner, E., Marth, C., Laird, P. W., and Widschwendter, M. (2006). Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. Cancer Res *66*, 29-33.

Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., Chen, H., Omeroglu, G., Meterissian, S., Omeroglu, A.*, et al.* (2008). Stromal gene expression predicts clinical outcome in breast cancer. Nat Med *14*, 518-527.

Fonslow, B. R., Stein, B. D., Webb, K. J., Xu, T., Choi, J., Park, S. K., and Yates, J. R., 3rd (2013). Digestion and depletion of abundant proteins improves proteomic coverage. Nature methods *10*, 54-56.

Forne, T., Labourier, E., Antoine, E., Rossi, F., Gallouzi, I., Cathala, G., Tazi, J., and Brunel, C. (1996). Structural features of U6 snRNA and dynamic interactions with other spliceosomal components leading to pre-mRNA splicing. Biochimie *78*, 436-442.

Friedl, P., and Wolf, K. (2003). Tumour-cell invasion and migration: diversity and escape mechanisms. Nat Rev Cancer *3*, 362-374.

Gaggioli, C., Hooper, S., Hidalgo-Carcedo, C., Grosse, R., Marshall, J. F., Harrington, K., and Sahai, E. (2007). Fibroblast-led collective invasion of carcinoma cells with differing roles for RhoGTPases in leading and following cells. Nat Cell Biol *9*, 1392-1400.

Gasparre, G., Porcelli, A. M., Lenaz, G., and Romeo, G. (2013). Relevance of mitochondrial genetics and metabolism in cancer development. Cold Spring Harb Perspect Biol *5*.

Gee, H. E., Buffa, F. M., Camps, C., Ramachandran, A., Leek, R., Taylor, M., Patil, M., Sheldon, H., Betts, G., Homer, J.*, et al.* (2011). The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. British journal of cancer *104*, 1168-1177.

Gerecke, D. R., Olson, P. F., Koch, M., Knoll, J. H., Taylor, R., Hudson, D. L., Champliaud, M. F., Olsen, B. R., and Burgeson, R. E. (1997). Complete primary structure of two splice variants of collagen XII, and assignment of alpha 1(XII) collagen (COL12A1), alpha 1(IX) collagen (COL9A1), and alpha 1(XIX) collagen (COL19A1) to human chromosome 6q12-q13. Genomics *41*, 236-242.

Goswami, S., Sahai, E., Wyckoff, J. B., Cammer, M., Cox, D., Pixley, F. J., Stanley, E. R., Segall, J. E., and Condeelis, J. S. (2005). Macrophages promote the invasion of breast carcinoma cells via a colony-stimulating factor-1/epidermal growth factor paracrine loop. Cancer Res *65*, 5278-5283.

Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., and Shiekhattar, R. (2004). The Microprocessor complex mediates the genesis of microRNAs. Nature *432*, 235-240.

Grimson, A., Farh, K. K., Johnston, W. K., Garrett-Engele, P., Lim, L. P., and Bartel, D. P. (2007). MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell *27*, 91-105.

Gusterson, B. (2009). Do 'basal-like' breast cancers really exist? Nat Rev Cancer *9*, 128-134.

Han, M., Yang, Z., Sayed, D., He, M., Gao, S., Lin, L., Yoon, S., and Abdellatif, M. (2012). GATA4 expression is primarily regulated via a miR-26b-dependent post-transcriptional mechanism during cardiac hypertrophy. Cardiovasc Res *93*, 645-654.

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

Hegerfeldt, Y., Tusch, M., Brocker, E. B., and Friedl, P. (2002). Collective cell movement in primary melanoma explants: plasticity of cell-cell interaction, beta1-integrin function, and migration strategies. Cancer Res *62*, 2125-2130.

Helwak, A., Kudla, G., Dudnakova, T., and Tollervey, D. (2013). Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell *153*, 654-665.

Holliday, D. L., Hughes, S., Shaw, J. A., Walker, R. A., and Jones, J. L. (2007). Intrinsic genetic characteristics determine tumor-modifying capacity of fibroblasts: matrix metalloproteinase-3 5A/5A genotype enhances breast cancer cell invasion. Breast Cancer Res *9*, R67.

Hu, M., Yao, J., Cai, L., Bachman, K. E., van den Brule, F., Velculescu, V., and Polyak, K. (2005). Distinct epigenetic changes in the stromal cells of breast cancers. Nature genetics *37*, 899-905.

Hu, Y., Li, P., Hao, S., Liu, L., Zhao, J., and Hou, Y. (2009). Differential expression of microRNAs in the placentae of Chinese patients with severe pre-eclampsia. Clinical chemistry and laboratory medicine : CCLM / FESCC *47*, 923-929.

Huang da, W., Sherman, B. T., Tan, Q., Kir, J., Liu, D., Bryant, D., Guo, Y., Stephens, R., Baseler, M. W., Lane, H. C., and Lempicki, R. A. (2007). DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. Nucleic Acids Res *35*, W169-175.

Hugh, J., Hanson, J., Cheang, M. C., Nielsen, T. O., Perou, C. M., Dumontet, C., Reed, J., Krajewska, M., Treilleux, I., Rupin, M.*, et al.* (2009). Breast cancer subtypes and response to docetaxel in node-positive breast cancer: use of an immunohistochemical definition in the BCIRG 001 trial. J Clin Oncol *27*, 1168-1176.

Inoue, T., Iinuma, H., Ogawa, E., Inaba, T., and Fukushima, R. (2012). Clinicopathological and prognostic significance of microRNA-107 and its relationship to DICER1 mRNA expression in gastric cancer. Oncol Rep *27*, 1759-1764.

Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M.*, et al.* (2005). MicroRNA gene expression deregulation in human breast cancer. Cancer Res *65*, 7065-7070.

Izu, Y., Sun, M., Zwolanek, D., Veit, G., Williams, V., Cha, B., Jepsen, K. J., Koch, M., and Birk, D. E. (2011). Type XII collagen regulates osteoblast polarity and communication during bone formation. J Cell Biol *193*, 1115-1130.

Jacobs, T. W., Byrne, C., Colditz, G., Connolly, J. L., and Schnitt, S. J. (1999). Radial scars in benign breast-biopsy specimens and the risk of breast cancer. N Engl J Med *340*, 430-436.

Jarocka-Cyrta, E., and Bankowski, E. (1991). [Molecular collagen content and polymorphism in the breast with benign dysplasia]. Pol Tyg Lek *46*, 981-983.

Jedeszko, C., Victor, B. C., Podgorski, I., and Sloane, B. F. (2009). Fibroblast hepatocyte growth factor promotes invasion of human mammary ductal carcinoma in situ. Cancer Res *69*, 9148-9155.

Ji, J., Shi, J., Budhu, A., Yu, Z., Forgues, M., Roessler, S., Ambs, S., Chen, Y., Meltzer, P. S., Croce, C. M.*, et al.* (2009). MicroRNA expression, survival, and response to interferon in liver cancer. N Engl J Med *361*, 1437-1447.

Ji, Y., He, Y., Liu, L., and Zhong, X. (2010). MiRNA-26b regulates the expression of cyclooxygenase-2 in desferrioxamine-treated CNE cells. FEBS letters *584*, 961-967.

Jiang, P., Enomoto, A., and Takahashi, M. (2009). Cell biology of the movement of breast cancer cells: intracellular signalling and the actin cytoskeleton. Cancer Lett *284*, 122-130.

Jones, J. L., Shaw, J. A., Pringle, J. H., and Walker, R. A. (2003). Primary breast myoepithelial cells exert an invasion-suppressor effect on breast cancer cells via paracrine down-regulation of MMP expression in fibroblasts and tumour cells. J Pathol *201*, 562-572.

Karagiannis, G. S., Petraki, C., Prassas, I., Saraon, P., Musrap, N., Dimitromanolakis, A., and Diamandis, E. P. (2012). Proteomic signatures of the desmoplastic invasion front reveal collagen type XII as a marker of myofibroblastic differentiation during colorectal cancer metastasis. Oncotarget *3*, 267-285.

Keely, P. J., Westwick, J. K., Whitehead, I. P., Der, C. J., and Parise, L. V. (1997). Cdc42 and Rac1 induce integrin-mediated cell motility and invasiveness through PI(3)K. Nature *390*, 632-636.

Kim, S., Yamamoto, J., Chen, Y., Aida, M., Wada, T., Handa, H., and Yamaguchi, Y. (2010). Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. Genes to cells : devoted to molecular & cellular mechanisms *15*, 1003-1013.

Kim, V. N. (2004). MicroRNA precursors in motion: exportin-5 mediates their nuclear export. Trends Cell Biol *14*, 156-159.

Kiss-Laszlo, Z., Henry, Y., Bachellerie, J. P., Caizergues-Ferrer, M., and Kiss, T. (1996). Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. Cell *85*, 1077-1088.

Kobayashi, D., Kumagai, J., Morikawa, T., Wilson-Morifuji, M., Wilson, A., Irie, A., and Araki, N. (2009). An integrated approach of differential mass spectrometry and gene ontology analysis identified novel proteins regulating neuronal differentiation and survival. Molecular & cellular proteomics : MCP *8*, 2350-2367.

Koh, C. M., Iwata, T., Zheng, Q., Bethel, C., Yegnasubramanian, S., and De Marzo, A. M. (2011). Myc enforces overexpression of EZH2 in early prostatic neoplasia via transcriptional and post-transcriptional mechanisms. Oncotarget *2*, 669-683.

Kojima, Y., Acar, A., Eaton, E. N., Mellody, K. T., Scheel, C., Ben-Porath, I., Onder, T. T., Wang, Z. C., Richardson, A. L., Weinberg, R. A., and Orimo, A. (2010). Autocrine TGF-beta and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts. Proceedings of the National Academy of Sciences of the United States of America *107*, 20009-20014.

Krek, A., Grun, D., Poy, M. N., Wolf, R., Rosenberg, L., Epstein, E. J., MacMenamin, P., da Piedade, I., Gunsalus, K. C., Stoffel, M., and Rajewsky, N. (2005). Combinatorial microRNA target predictions. Nat Genet *37*, 495-500.

Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet *11*, 597-610.

Kuperwasser, C., Chavarria, T., Wu, M., Magrane, G., Gray, J. W., Carey, L., Richardson, A., and Weinberg, R. A. (2004). Reconstruction of functionally normal and malignant human breast tissues in mice. Proceedings of the National Academy of Sciences of the United States of America *101*, 4966-4971.

Kurose, K., Gilley, K., Matsumoto, S., Watson, P. H., Zhou, X. P., and Eng, C. (2002). Frequent somatic mutations in PTEN and TP53 are mutually exclusive in the stroma of breast carcinomas. Nat Genet *32*, 355-357.

Lakhani, S. R., and Ellis, I. O. (2012). WHO Classification of Tumours of the Breast.

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

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

Lee, Y., Hur, I., Park, S. Y., Kim, Y. K., Suh, M. R., and Kim, V. N. (2006). The role of PACT in the RNA silencing pathway. Embo J *25*, 522-532.

Levental, K. R., Yu, H., Kass, L., Lakins, J. N., Egeblad, M., Erler, J. T., Fong, S. F., Csiszar, K., Giaccia, A., Weninger, W.*, et al.* (2009). Matrix crosslinking forces tumor progression by enhancing integrin signaling. Cell *139*, 891-906.

Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787-798.

Li, J., Kong, X., Zhang, J., Luo, Q., Li, X., and Fang, L. (2013). MiRNA-26b inhibits proliferation by targeting PTGS2 in breast cancer. Cancer cell international *13*, 7.

Liang, C. C., Park, A. Y., and Guan, J. L. (2007). In vitro scratch assay: a convenient and inexpensive method for analysis of cell migration in vitro. Nat Protoc *2*, 329-333.

Lin, F., Li, R., Pan, Z. X., Zhou, B., Yu de, B., Wang, X. G., Ma, X. S., Han, J., Shen, M., and Liu, H. L. (2012). miR-26b promotes granulosa cell apoptosis by targeting ATM during follicular atresia in porcine ovary. PLoS One *7*, e38640.

Liu, X. X., Li, X. J., Zhang, B., Liang, Y. J., Zhou, C. X., Cao, D. X., He, M., Chen, G. Q., He, J. R., and Zhao, Q. (2011). MicroRNA-26b is underexpressed in human breast cancer and induces cell apoptosis by targeting SLC7A11. FEBS letters *585*, 1363-1367.

Liu, Z., Jia, X., Duan, Y., Xiao, H., Sundqvist, K. G., Permert, J., and Wang, F. (2013). Excess glucose induces hypoxia-inducible factor-1alpha in pancreatic cancer cells and stimulates glucose metabolism and cell migration. Cancer Biol Ther *14*.

Longacre, T. A., Ennis, M., Quenneville, L. A., Bane, A. L., Bleiweiss, I. J., Carter, B. A., Catelano, E., Hendrickson, M. R., Hibshoosh, H., Layfield, L. J.*, et al.* (2006). Interobserver agreement and reproducibility in classification of invasive breast carcinoma: an NCI breast cancer family registry study. Mod Pathol *19*, 195-207.

Ma, L., Teruya-Feldstein, J., and Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature *449*, 682-688.

Ma, X. J., Wang, Z., Ryan, P. D., Isakoff, S. J., Barmettler, A., Fuller, A., Muir, B., Mohapatra, G., Salunga, R., Tuggle, J. T.*, et al.* (2004). A two-gene expression ratio predicts clinical outcome in breast cancer patients treated with tamoxifen. Cancer Cell *5*, 607-616.

Ma, Y. L., Zhang, P., Wang, F., Moyer, M. P., Yang, J. J., Liu, Z. H., Peng, J. Y., Chen, H. Q., Zhou, Y. K., Liu, W. J., and Qin, H. L. (2011a). Human embryonic stem cells and metastatic colorectal cancer cells shared the common endogenous human microRNA-26b. Journal of cellular and molecular medicine *15*, 1941-1954.

Ma, Y. L., Zhang, P., Wang, F., Moyer, M. P., Yang, J. J., Liu, Z. H., Peng, J. Y., Chen, H. Q., Zhou, Y. K., Liu, W. J., and Qin, H. L. (2011b). Human embryonic stem cells and metastatic colorectal cancer cells shared the common endogenous human microRNA-26b. J Cell Mol Med *15*, 1941-1954.

Maeda, T., Desouky, J., and Friedl, A. (2006). Syndecan-1 expression by stromal fibroblasts promotes breast carcinoma growth in vivo and stimulates tumor angiogenesis. Oncogene *25*, 1408-1412.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics (Oxford, England) *21*, 3448-3449.

Martinez-Outschoorn, U. E., Whitaker-Menezes, D., Pavlides, S., Chiavarina, B., Bonuccelli, G., Casey, T., Tsirigos, A., Migneco, G., Witkiewicz, A., Balliet, R.*, et al.* (2010). The autophagic tumor stroma model of cancer or "battery-operated tumor growth": A simple solution to the autophagy paradox. Cell cycle (Georgetown, Tex *9*, 4297-4306.

Massoudi, D., Malecaze, F., Soler, V., Butterworth, J., Erraud, A., Fournie, P., Koch, M., and Galiacy, S. D. (2012). NC1 long and NC3 short splice variants of type XII collagen are overexpressed during corneal scarring. Invest Ophthalmol Vis Sci *53*, 7246-7256.

Mestdagh, P., Feys, T., Bernard, N., Guenther, S., Chen, C., Speleman, F., and Vandesompele, J. (2008). High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. Nucleic Acids Res *36*, e143.

Mi, H., Muruganujan, A., and Thomas, P. D. (2013). PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. Nucleic Acids Res *41*, D377-386.

Mikula, M., Rubel, T., Karczmarski, J., Goryca, K., Dadlez, M., and Ostrowski, J. (2010). Integrating proteomic and transcriptomic high-throughput surveys for search of new biomarkers of colon tumors. Functional & integrative genomics.

Mitra, A. K., Zillhardt, M., Hua, Y., Tiwari, P., Murmann, A. E., Peter, M. E., and Lengyel, E. (2012). MicroRNAs reprogram normal fibroblasts into cancer-associated fibroblasts in ovarian cancer. Cancer discovery *2*, 1100-1108.

Moinfar, F. (2008). Is 'basal-like' carcinoma of the breast a distinct clinicopathological entity? A critical review with cautionary notes. Pathobiology *75*, 119-131.

Moinfar, F., Man, Y. G., Arnould, L., Bratthauer, G. L., Ratschek, M., and Tavassoli, F. A. (2000). Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. Cancer Res *60*, 2562-2566.

Moinzadeh, P., Agarwal, P., Bloch, W., Orteu, C., Hunzelmann, N., Eckes, B., and Krieg, T. (2013). Systemic sclerosis with multiple nodules: characterization of the extracellular matrix. Archives of dermatological research.

Musumeci, M., Coppola, V., Addario, A., Patrizii, M., Maugeri-Sacca, M., Memeo, L., Colarossi, C., Francescangeli, F., Biffoni, M., Collura, D.*, et al.* (2011). Control of tumor and microenvironment cross-talk by miR-15a and miR-16 in prostate cancer. Oncogene.

Network, C. G. A. (2012). Comprehensive molecular portraits of human breast tumours. In Nature, (England), pp. 61-70.

Ng, Y. Z., Pourreyron, C., Salas-Alanis, J. C., Dayal, J. H., Cepeda-Valdes, R., Yan, W., Wright, S., Chen, M., Fine, J. D., Hogg, F. J.*, et al.* (2012). Fibroblast-derived dermal matrix drives development of aggressive cutaneous squamous cell carcinoma in patients with recessive dystrophic epidermolysis bullosa. Cancer Res *72*, 3522-3534.

Nielsen, B. S., Jorgensen, S., Fog, J. U., Sokilde, R., Christensen, I. J., Hansen, U., Brunner, N., Baker, A., Moller, S., and Nielsen, H. J. (2011). High levels of microRNA-21 in the stroma of colorectal cancers predict short disease-free survival in stage II colon cancer patients. Clinical & experimental metastasis *28*, 27-38.

Nishiyama, T., McDonough, A. M., Bruns, R. R., and Burgeson, R. E. (1994). Type XII and XIV collagens mediate interactions between banded collagen fibers in vitro and may modulate extracellular matrix deformability. J Biol Chem *269*, 28193-28199.

Nobes, C. D., and Hall, A. (1995). Rho, rac, and cdc42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia. Cell *81*, 53-62.

Noel, A., De Pauw-Gillet, M. C., Purnell, G., Nusgens, B., Lapiere, C. M., and Foidart, J. M. (1993). Enhancement of tumorigenicity of human breast adenocarcinoma cells in nude mice by matrigel and fibroblasts. Br J Cancer *68*, 909-915.

Nowicki, M. O., Dmitrieva, N., Stein, A. M., Cutter, J. L., Godlewski, J., Saeki, Y., Nita, M., Berens, M. E., Sander, L. M., Newton, H. B.*, et al.* (2008). Lithium inhibits invasion of glioma cells; possible involvement of glycogen synthase kinase-3. Neuro-oncology *10*, 690-699.

Olsen, C. J., Moreira, J., Lukanidin, E. M., and Ambartsumian, N. S. (2010). Human mammary fibroblasts stimulate invasion of breast cancer cells in a three-dimensional culture and increase stroma development in mouse xenografts. In BMC Cancer, (England), p. 444.

Orimo, A., Gupta, P. B., Sgroi, D. C., Arenzana-Seisdedos, F., Delaunay, T., Naeem, R., Carey, V. J., Richardson, A. L., and Weinberg, R. A. (2005). Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. Cell *121*, 335-348.

Paik, S., Shak, S., Tang, G., Kim, C., Baker, J., Cronin, M., Baehner, F. L., Walker, M. G., Watson, D., Park, T.*, et al.* (2004). A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. N Engl J Med *351*, 2817-2826.

Palumbo, T., Faucz, F. R., Azevedo, M., Xekouki, P., Iliopoulos, D., and Stratakis, C. A. (2012). Functional screen analysis reveals miR-26b and miR-128 as central regulators of pituitary somatomammotrophic tumor growth through activation of the PTEN-AKT pathway. Oncogene.

Pan, C., Kumar, C., Bohl, S., Klingmueller, U., and Mann, M. (2009). Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions. Molecular & cellular proteomics : MCP *8*, 443-450.

Parker, J. S., Mullins, M., Cheang, M. C., Leung, S., Voduc, D., Vickery, T., Davies, S., Fauron, C., He, X., Hu, Z.*, et al.* (2009). Supervised risk predictor of breast cancer based on intrinsic subtypes. J Clin Oncol *27*, 1160-1167.

Perou, C. M., Sorlie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S., Rees, C. A., Pollack, J. R., Ross, D. T., Johnsen, H., Akslen, L. A.*, et al.* (2000). Molecular portraits of human breast tumours. Nature *406*, 747-752.

Peveling-Oberhag, J., Crisman, G., Schmidt, A., Doring, C., Lucioni, M., Arcaini, L., Rattotti, S., Hartmann, S., Piiper, A., Hofmann, W. P.*, et al.* (2012). Dysregulation of global microRNA expression in splenic marginal zone lymphoma and influence of chronic hepatitis C virus infection. Leukemia *26*, 1654-1662.

Place, A. E., Jin Huh, S., and Polyak, K. (2011). The microenvironment in breast cancer progression: biology and implications for treatment. In Breast Cancer Res, (England), p. 227.

Posthumus, M., September, A. V., O'Cuinneagain, D., van der Merwe, W., Schwellnus, M. P., and Collins, M. (2010). The association between the COL12A1 gene and anterior cruciate ligament ruptures. British journal of sports medicine *44*, 1160-1165.

Qi, L., Bart, J., Tan, L. P., Platteel, I., Sluis, T., Huitema, S., Harms, G., Fu, L., Hollema, H., and Berg, A. (2009). Expression of miR-21 and its targets (PTEN, PDCD4, TM1) in flat epithelial atypia of the breast in relation to ductal carcinoma in situ and invasive carcinoma. BMC Cancer *9*, 163.

Qiu, W., Hu, M., Sridhar, A., Opeskin, K., Fox, S., Shipitsin, M., Trivett, M., Thompson, E. R., Ramakrishna, M., Gorringe, K. L.*, et al.* (2008). No evidence of clonal somatic genetic alterations in cancer-associated fibroblasts from human breast and ovarian carcinomas. Nat Genet *40*, 650-655.

Rahbari, R., Holloway, A. K., He, M., Khanafshar, E., Clark, O. H., and Kebebew, E. (2011). Identification of differentially expressed microRNA in parathyroid tumors. Annals of surgical oncology *18*, 1158-1165.

Rask, L., Balslev, E., Jorgensen, S., Eriksen, J., Flyger, H., Moller, S., Hogdall, E., Litman, T., and Nielsen, B. S. (2011). High expression of miR-21 in tumor stroma correlates with increased cancer cell proliferation in human breast cancer. APMIS : acta pathologica, microbiologica, et immunologica Scandinavica *119*, 663-673.

Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J., Briggs, B. B., Barrette, T. R., Anstet, M. J., Kincead-Beal, C., Kulkarni, P.*, et al.* (2007). Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. Neoplasia *9*, 166-180.

Ronnov-Jessen, L., Petersen, O. W., and Bissell, M. J. (1996). Cellular changes involved in conversion of normal to malignant breast: importance of the stromal reaction. Physiol Rev *76*, 69-125.

Ronnov-Jessen, L., Petersen, O. W., Koteliansky, V. E., and Bissell, M. J. (1995). The origin of the myofibroblasts in breast cancer. Recapitulation of tumor environment in culture unravels diversity and implicates converted fibroblasts and recruited smooth muscle cells. J Clin Invest *95*, 859-873.

Sandhu, R., Rivenbark, A. G., and Coleman, W. B. (2012). Loss of post-transcriptional regulation of DNMT3b by microRNAs: a possible molecular mechanism for the hypermethylation defect observed in a subset of breast cancer cell lines. Int J Oncol *41*, 721-732.

Schetter, A. J., Leung, S. Y., Sohn, J. J., Zanetti, K. A., Bowman, E. D., Yanaihara, N., Yuen, S. T., Chan, T. L., Kwong, D. L., Au, G. K.*, et al.* (2008). MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. Jama *299*, 425-436.

Scott, G. K., Mattie, M. D., Berger, C. E., Benz, S. C., and Benz, C. C. (2006). Rapid alteration of microRNA levels by histone deacetylase inhibition. Cancer Res *66*, 1277-1281.

Seimiya, H., Muramatsu, Y., Ohishi, T., and Tsuruo, T. (2005). Tankyrase 1 as a target for telomere-directed molecular cancer therapeutics. Cancer Cell *7*, 25-37.

Seimiya, H., and Smith, S. (2002). The telomeric poly(ADP-ribose) polymerase, tankyrase 1, contains multiple binding sites for telomeric repeat binding factor 1 (TRF1) and a novel acceptor, 182-kDa tankyrase-binding protein (TAB182). J Biol Chem *277*, 14116-14126.

Sempere, L. F., Christensen, M., Silahtaroglu, A., Bak, M., Heath, C. V., Schwartz, G., Wells, W., Kauppinen, S., and Cole, C. N. (2007). Altered MicroRNA expression confined to specific epithelial cell subpopulations in breast cancer. Cancer Res *67*, 11612-11620.

Shieh, A. C., Rozansky, H. A., Hinz, B., and Swartz, M. A. (2011). Tumor cell invasion is promoted by interstitial flow-induced matrix priming by stromal fibroblasts. Cancer Res *71*, 790-800.

Si, M. L., Zhu, S., Wu, H., Lu, Z., Wu, F., and Mo, Y. Y. (2007). miR-21-mediated tumor growth. Oncogene *26*, 2799-2803.

Singletary, S. E., Allred, C., Ashley, P., Bassett, L. W., Berry, D., Bland, K. I., Borgen, P. I., Clark, G., Edge, S. B., Hayes, D. F.*, et al.* (2002). Revision of the American Joint Committee on Cancer staging system for breast cancer. J Clin Oncol *20*, 3628-3636.

Solomides, C. C., Evans, B. J., Navenot, J. M., Vadigepalli, R., Peiper, S. C., and Wang, Z. X. (2012). MicroRNA profiling in lung cancer reveals new molecular markers for diagnosis. Acta Cytol *56*, 645-654.

Sorlie, T., Perou, C. M., Tibshirani, R., Aas, T., Geisler, S., Johnsen, H., Hastie, T., Eisen, M. B., van de Rijn, M., Jeffrey, S. S.*, et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proceedings of the National Academy of Sciences of the United States of America *98*, 10869-10874.

Sotgia, F., Martinez-Outschoorn, U. E., Howell, A., Pestell, R. G., Pavlides, S., and Lisanti, M. P. (2012). Caveolin-1 and cancer metabolism in the tumor microenvironment: markers, models, and mechanisms. Annual review of pathology *7*, 423-467.

Sotiriou, C., Wirapati, P., Loi, S., Harris, A., Fox, S., Smeds, J., Nordgren, H., Farmer, P., Praz, V., Haibe-Kains, B.*, et al.* (2006). Gene expression profiling in breast cancer: understanding the molecular basis of histologic grade to improve prognosis. J Natl Cancer Inst *98*, 262-272.

Sottnik, J. L., Lori, J. C., Rose, B. J., and Thamm, D. H. (2011). Glycolysis inhibition by 2-deoxy-D-glucose reverts the metastatic phenotype in vitro and in vivo. Clin Exp Metastasis *28*, 865-875.

Stingl, J., and Caldas, C. (2007). Molecular heterogeneity of breast carcinomas and the cancer stem cell hypothesis. Nature reviews Cancer *7*, 791-799.

Studebaker, A. W., Storci, G., Werbeck, J. L., Sansone, P., Sasser, A. K., Tavolari, S., Huang, T., Chan, M. W., Marini, F. C., Rosol, T. J.*, et al.* (2008). Fibroblasts isolated

from common sites of breast cancer metastasis enhance cancer cell growth rates and invasiveness in an interleukin-6-dependent manner. Cancer Res *68*, 9087-9095.

Subik, K., Lee, J. F., Baxter, L., Strzepek, T., Costello, D., Crowley, P., Xing, L., Hung, M. C., Bonfiglio, T., Hicks, D. G., and Tang, P. (2010). The Expression Patterns of ER, PR, HER2, CK5/6, EGFR, Ki-67 and AR by Immunohistochemical Analysis in Breast Cancer Cell Lines. Breast cancer : basic and clinical research *4*, 35-41.

Sung, C. O., Lee, K. W., Han, S., and Kim, S. H. (2011). Twist1 is up-regulated in gastric cancer-associated fibroblasts with poor clinical outcomes. Am J Pathol *179*, 1827-1838.

Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., Gerald, W. L., and Massague, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. Nature *451*, 147-152.

Trimboli, A. J., Cantemir-Stone, C. Z., Li, F., Wallace, J. A., Merchant, A., Creasap, N., Thompson, J. C., Caserta, E., Wang, H., Chong, J. L.*, et al.* (2009). Pten in stromal fibroblasts suppresses mammary epithelial tumours. Nature *461*, 1084-1091.

Tzortzaki, E. G., Koutsopoulos, A. V., Dambaki, K. I., Lambiri, I., Plataki, M., Gordon, M. K., Gerecke, D. R., and Siafakas, N. M. (2006). Active remodeling in idiopathic interstitial pneumonias: evaluation of collagen types XII and XIV. J Histochem Cytochem *54*, 693-700.

van 't Veer, L. J., Dai, H., van de Vijver, M. J., He, Y. D., Hart, A. A., Mao, M., Peterse, H. L., van der Kooy, K., Marton, M. J., Witteveen, A. T.*, et al.* (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530-536.

van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A., Voskuil, D. W., Schreiber, G. J., Peterse, J. L., Roberts, C., Marton, M. J.*, et al.* (2002). A gene-expression signature as a predictor of survival in breast cancer. N Engl J Med *347*, 1999-2009.

Verghese, E. T., Hanby, A. M., Speirs, V., and Hughes, T. A. (2008). Small is beautiful: microRNAs and breast cancer-where are we now? The Journal of pathology *215*, 214-221.

Verghese, E. T., Shenoy, H., Cookson, V. J., Green, C. A., Howarth, J., Partanen, R. H., Pollock, S., Waterworth, A., Speirs, V., Hughes, T. A., and Hanby, A. M. (2011). Epithelial-mesenchymal interactions in breast cancer: evidence for a role of nuclear localized beta-catenin in carcinoma-associated fibroblasts. Histopathology *59*, 609-618.

Wang, Y., Klijn, J. G., Zhang, Y., Sieuwerts, A. M., Look, M. P., Yang, F., Talantov, D., Timmermans, M., Meijer-van Gelder, M. E., Yu, J.*, et al.* (2005). Gene-expression profiles to predict distant metastasis of lymph-node-negative primary breast cancer. Lancet *365*, 671-679.

Whitley, E., and Ball, J. (2002). Statistics review 4: sample size calculations. Crit Care *6*, 335-341.

Wienholds, E., Kloosterman, W. P., Miska, E., Alvarez-Saavedra, E., Berezikov, E., de Bruijn, E., Horvitz, H. R., Kauppinen, S., and Plasterk, R. H. (2005). MicroRNA expression in zebrafish embryonic development. Science *309*, 310-311.

Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855-862.

Willems, S. M., Mohseny, A. B., Balog, C., Sewrajsing, R., Briaire-de Bruijn, I. H., Knijnenburg, J., Cleton-Jansen, A. M., Sciot, R., Fletcher, C. D., Deelder, A. M.*, et al.* (2009). Cellular/intramuscular myxoma and grade I myxofibrosarcoma are characterized by distinct genetic alterations and specific composition of their extracellular matrix. J Cell Mol Med *13*, 1291-1301.

Wong, T. S., Liu, X. B., Wong, B. Y., Ng, R. W., Yuen, A. P., and Wei, W. I. (2008). Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. Clin Cancer Res *14*, 2588-2592.

Wu, N., Zhao, X., Liu, M., Liu, H., Yao, W., Zhang, Y., Cao, S., and Lin, X. (2011). Role of microRNA-26b in glioma development and its mediated regulation on EphA2. PLoS One *6*, e16264.

Wyckoff, J., Wang, W., Lin, E. Y., Wang, Y., Pixley, F., Stanley, E. R., Graf, T., Pollard, J. W., Segall, J., and Condeelis, J. (2004). A paracrine loop between tumor cells and macrophages is required for tumor cell migration in mammary tumors. Cancer Res *64*, 7022-7029.

Wyckoff, J. B., Wang, Y., Lin, E. Y., Li, J. F., Goswami, S., Stanley, E. R., Segall, J. E., Pollard, J. W., and Condeelis, J. (2007). Direct visualization of macrophage-assisted tumor cell intravasation in mammary tumors. Cancer Res *67*, 2649-2656.

Yan, L. X., Wu, Q. N., Zhang, Y., Li, Y. Y., Liao, D. Z., Hou, J. H., Fu, J., Zeng, M. S., Yun, J. P., Wu, Q. L.*, et al.* (2011). Knockdown of miR-21 in human breast cancer cell lines inhibits proliferation, in vitro migration and in vivo tumor growth. Breast cancer research : BCR *13*, R2.

Yasuda, T., Nishio, J., Sumegi, J., Kapels, K. M., Althof, P. A., Sawyer, J. R., Reith, J. D., and Bridge, J. A. (2009). Aberrations of 6q13 mapped to the COL12A1 locus in chondromyxoid fibroma. Mod Pathol *22*, 1499-1506.

Zhang, Z., Florez, S., Gutierrez-Hartmann, A., Martin, J. F., and Amendt, B. A. (2010). MicroRNAs regulate pituitary development, and microRNA 26b specifically targets lymphoid enhancer factor 1 (Lef-1), which modulates pituitary transcription factor 1 (Pit-1) expression. J Biol Chem *285*, 34718-34728.

Zhao, L., Sun, Y., Hou, Y., Peng, Q., Wang, L., Luo, H., Tang, X., Zeng, Z., and Liu, M. (2012a). MiRNA expression analysis of cancer-associated fibroblasts and normal fibroblasts in breast cancer. Int J Biochem Cell Biol *44*, 2051-2059.

Zhao, L., Sun, Y., Hou, Y., Peng, Q., Wang, L., Luo, H., Tang, X., Zeng, Z., and Liu, M. (2012b). MiRNA expression analysis of cancer-associated fibroblasts and normal fibroblasts in breast cancer. Int J Biochem Cell Biol *44*, 2051-2059.

Zhu, Y., Lu, Y., Zhang, Q., Liu, J. J., Li, T. J., Yang, J. R., Zeng, C., and Zhuang, S. M. (2012). MicroRNA-26a/b and their host genes cooperate to inhibit the G1/S transition by activating the pRb protein. Nucleic Acids Res *40*, 4615-4625.

9 APPENDIX

9.1 ETHICS

Leeds (East) Research Ethics Committee
Room 5.2, Clinical Sciences Building
St James's University Hospital
Beckett Street
Leeds
LS9 7TF

Telephone: 0113 2065652
Facsimile: 0113 2066772

17 November 2006

Professor Andrew Hanby
Professor of Breast Pathology
Department of Histopathology
St James's University Hospital
Beckett Street
Leeds
LS9 7TF

Dear Professor Hanby

Full title of study: **Studies of the biological significance of breast cancer subtype, using molecular and cytogenetic profiling and in vivo models.**

REC reference number: **06/Q1206/180**

The Research Ethics Committee reviewed the above application at the meeting held on 7 November 2006.

Ethical opinion

The members of the Committee present gave a favourable ethical opinion of the above research on the basis described in the application form, protocol and supporting documentation.

Ethical review of research sites

,The Committee agreed that all sites in this study should be exempt from site-specific assessment (SSA). There is no need to complete Part C of the application form or to inform Local Research Ethics Committees (LRECs) about the research. The favourable opinion for the study applies to all sites involved in the research.

Conditions of approval

The favourable opinion is given provided that you comply with the conditions set out in the attached document. You are advised to study the conditions carefully.

Approved documents

The documents reviewed and approved at the meeting were:

Document	*Version*	*Date*
Application		03 October 2006
Investigator CV		

An advisory committee to West Yorkshire Strategic Health Authority

Protocol	1	09 October 2006
Letter from Sponsor		26 October 2006

Research governance approval

You should arrange for the R&D Department at all relevant NHS care organisations to be notified that the research will be taking place, and provide a copy of the REC application, the protocol and this letter.

All researchers and research collaborators who will be participating in the research at a NHS site must obtain final research governance approval before commencing any research procedures. Where a substantive contract is not held with the care organisation, it may be necessary for an honorary contract to be issued before approval for the research can be given.

Membership of the Committee

The members of the Ethics Committee who were present at the meeting are listed on the attached sheet.

Statement of compliance

The Committee is constituted in accordance with the Governance Arrangements for Research Ethics Committees (July 2001) and complies fully with the Standard Operating Procedures for Research Ethics Committees in the UK.

06/Q1206/180 **Please quote this number on all correspondence**

With the Committee's best wishes for the success of this project

Yours sincerely

Ann Prothero

p.p. **Dr John Holmes**
Chair

Email: ann.prothero@leedsth.nhs.uk

Enclosures: List of names and professions of members who were present at the meeting and those who submitted written comments

Standard approval conditions

Copy to: Clare Skinner
Research Office
Room 7.11, Level 7
Worsley Building
University of Leeds

R&D Department, Leeds Teaching Hospitals NHS Trust

9.2 REAGENTS AND BUFFERS

Collagen Medium (1X):

Components	**Total Volume 50 ml**
alpha MEM powder (5X)*	2.5 g
1M Hepes pH7.5	5 ml (0.1 M final)
NaHCO3	1g (2% final)
H_20	Up to 50 ml

alpha MEM powder(5X) was kindly provided by Steven Hooper, UCL

The above components were mixed and then sterilised by passing it through a 0.2µm filter.

Collagen-1/Matrigel Mix:

The following components were maintained on ice and mixed in the order shown.

Components	**Total Volume1000 µl**
Matrigel (BD Cat no. 356231)	200 µl
Collagen-1 (BD Cat no. 354236)	400 µl
5X collagen medium	80 µl

FCS	100 µl
DMEM	220 µl

Lysogeny broth (LB):

5 g of Bacto-Tryptone tryptone, 2.5 g of yeast extract, and 5 g of NaCl was added to 500 ml deionized water and autoclaved. Ampicillin was added at final concentration of 50 µg/ml.

Phosphate Buffered Saline (PBS):

PBS was prepared by mixing 60 ml of 2.5 M NaCl with 20 ml 1M Tris-HCl (pH-7.4). It was then made up to 1000 ml using deionised water.

PBS- Tween (PBST):

PBST was prepared by adding 1.2 ml 10% v/v Tween-20 (Fisher Scientific Cat no. BP337-100) to 1000 ml of PBS.

5% Skimmed milk:

5 mg of skimmed milk was mixed with 100 ml of PBST.

RIPA Buffer:

A working solution of RIPA buffer was prepared by adding 1 tablet of Complete protease inhibitor cocktail tablet (Roche Cat no. 11836153001) to 50 ml of RIPA buffer (Sigma Cat no. R0278).

1X Running Buffer:

A working solution of 1X running buffer was made by mixing 25 ml of NuPAGE® MOPS SDS Running Buffer (Life technologies Cat No. NP0001) with 475 ml of deionised water.

1X Transfer Buffer:

A working solution of 1X transfer buffer was made by mixing 25 ml of NuPAGE® Transfer Buffer (Life technologies Cat no. NP 0006-1) with 50 ml of methanol and 425 ml of deionised water.

DNA Gels:

To prepare the gel, 0.8 to 2.5 g agarose was added to 100 ml of 1 x Tris Borate EDTA buffer (Severn Biotech Ltd Cat no. 20-6000-50). This was warmed to dissolve the agarose. It was then cooled to 50^0C and 6 µl of ethidium bromide (10 mg/ml solution; Promega Cat no. H5041) was added. The gel was poured into the rig with a comb in place. Once cooled, the comb was removed to reveal the wells. 10 µl of Bioline hyperladder-1 (Cat no. BIO-33053) or 12 µl of the sample was loaded into each well. The tank was filled with 1X TBE. It was then run for 0.5 to 1 hour at 50-100V. The gel was visualized using a UV transilluminator (BioRad)

Proteinase-K Buffer:

This was prepared by adding 5 ml of 1 M Tris-HCl (pH 7.4), 2 ml of 0.5 M EDTA and 0.2 ml of 5 M NaCl to 900 ml nuclease free water. This was made up to 1000 ml and autoclaved.

Proteinase-K Reagent:

This was prepared by diluting the proteinase K stock solution (20mg/ml) using the proteinase K buffer. For example, for a proteinase-K reagent containing proteinase K at a concentration of 20 µg/ml, 10 µl of proteinase K stock solution was added to 10 ml of proteinase K buffer.

SSC solutions:

A 20x Saline Sodium Citrate (SSC) solution (Life technologies Cat no. 15557-036) containing 3.0 M NaCl and 0.3 M sodium citrate, at pH 7.0 was used as the stock solution. The lower concentrations of SSC were made by diluting 20x SSC with the appropriate amount of nuclease free water.

KTBT (AP stop solution):

7.9g Tris-HCl, 8.7g NaCl and 0.75g KCl were added to 900 ml of nuclease free water. This was made up to 1000 ml and autoclaved.

DIG Blocking solution:

1x blocking solution was prepared by diluting the 10X blocking solution with 1x Maleic acid buffer (Roche Cat no.11585762 001).

Anti-DIG reagent:

Sheep-anti-DIG-AP antibody (Roche Cat no. 11093274910) was diluted (1:800) in DIG blocking solution containing 2% sheep serum (Jackson Immunoresearch).

AP substrate:

Immediately prior to use, a single tablet of NBT (nitro-blue tetrazolium chloride) / BCIP (5-bromo-4-chloro-3'-indolyphosphate p-toluidine salt) (Roche Cat no. 11697471001) was dissolved in Milli-Q water. To this Levamisole (Fluka, Cat no. 31742) was added for a final concentration of 0.2 mM.

9.3 TABLES

	Down- regulated	Up-regulated
Fold Change > 10	hsa-let-7f	Nil
	hsa-let-7a	
	hsa-miR-26b	
	hsa-miR-199b-3p	
	hsa-miR-365	
	hsa-let-7e	
	hsa-let-7d	
	hsa-let-7g	
	hsa-miR-10b	
	hsa-miR-15b	
	hsa-miR-451	
	hsa-miR-107	
	hsa-miR-195	
	hsa-miR-10a	
	hsa-miR-30b	
	hsa-miR-16	
	hsa-miR-23a	
	hsa-miR-126	
Fold change > 5	hsa-miR-148a	hsa-miR-933
	hsa-miR-23b	hsa-miR-767-3p
	hsa-let-7c	hsa-let-7b*
	hsa-miR-100	hsa-miR-563
	hsa-miR-199b-5p	hsa-miR-551b
	hsa-miR-150	
	hsa-miR-20a	
	hsa-miR-98	

	hsa-miR-125b	
	hsa-miR-21	
	hsa-miR-34a	
	hsa-miR-151-5p	
	hsa-miR-29b	
	hsa-miR-19b	
	hsa-miR-768-5p	
	hsa-miR-223	
	hsa-miR-27a	
	hsa-miR-15a	
	hsa-miR-196a	
	hsa-miR-214	
	hsa-miR-142-3p	
	hsa-miR-29a	
	hsa-miR-30c	
	hsa-miR-27b	
	hsa-miR-331-3p	
	hsa-miR-103	
	hsa-miR-17	
	hsa-miR-125a-5p	
	hsa-miR-200b	
Fold change < 5	hsa-miR-24	hsa-miR-136
	hsa-miR-146b-5p	hsa-miR-144
	hsa-miR-29c	hsa-miR-483-3p
	hsa-miR-25	hsa-miR-516a-5p
	hsa-miR-455-3p	hsa-miR-1229
	hsa-miR-199a-5p	hsa-let-7f-1*
	hsa-miR-130a	hsa-miR-580
	hsa-miR-886-3p	hsa-miR-1237

hsa-miR-99a	hsa-miR-636
hsa-miR-146a	hsa-miR-149*
hsa-let-7b	hsa-miR-188-5p
hsa-miR-342-5p	hsa-miR-296-5p
hsa-miR-128	hsa-miR-425*
hsa-miR-425	hsa-miR-602
hsa-miR-29c*	hsa-miR-125a-3p
hsa-miR-28-5p	hsa-miR-518e*
hsa-miR-34b*	hsa-miR-557
hsa-miR-30e*	hsa-miR-1225-3p
hsa-miR-361-3p	hsa-miR-204
hsa-miR-768-3p	hsa-miR-1228
hsa-miR-424	hsa-miR-572
hsa-miR-132	hsa-miR-129*
hsa-miR-106b	hsa-miR-671-5p
hsa-miR-93	hsa-miR-1238
hsa-let-7i	hsa-miR-135a*
hsa-miR-26a	hsa-miR-877
hsa-miR-374b	hsa-miR-191*
hsa-miR-20b	hsa-miR-550
hsa-miR-101	hsa-miR-187*
hsa-miR-149	hsa-miR-769-3p
hsa-miR-193a-3p	hsa-miR-18b*
hsa-miR-96	hsa-miR-638
hsa-miR-342-3p	hsa-miR-760
hsa-miR-185	hsa-miR-373*
hsa-miR-513a-5p	hsa-miR-650
hsa-miR-221	hsa-miR-371-5p
hsa-miR-200c	hsa-miR-625*

hsa-miR-193b	hsa-miR-298
hsa-miR-652	hsa-miR-877*
hsa-miR-374a	hsa-miR-30c-1*
hsa-miR-30a*	hsa-miR-518a-3p
hsa-miR-625	hsa-miR-601
hsa-miR-143	hsa-miR-328
hsa-miR-155	hsa-miR-623
hsa-miR-22*	hsa-miR-202
hsa-miR-127-3p	hsa-miR-10b*
hsa-miR-324-5p	hsa-miR-1226*
hsa-miR-376c	hsa-miR-564
hsa-miR-145	hsa-miR-21*
hsa-miR-940	hsa-miR-498
hsa-miR-99b	hsa-miR-1234
hsa-miR-30e	hsa-miR-1225-5p
hsa-miR-497	hsa-miR-139-3p
hsa-miR-151-3p	hsa-miR-887
hsa-miR-140-3p	hsa-miR-939
hsa-miR-200b*	hsa-miR-33b*
hsa-miR-574-3p	hsa-miR-936
hsa-miR-92a	hsa-miR-634
hsa-miR-186	hsa-miR-483-5p
hsa-miR-375	hsa-miR-575
hsa-miR-22	hsa-miR-299-5p
hsa-miR-376a	hsa-miR-1224-5p
hsa-miR-205	hsa-miR-378
hsa-miR-210	hsa-miR-378*
hsa-miR-494	hsa-miR-512-3p
hsa-miR-484	hsa-miR-150*

hsa-miR-324-3p	hsa-miR-134
hsa-miR-197	hsa-miR-320
hsa-miR-424*	hsa-miR-518c*
hsa-miR-532-3p	hsa-miR-218
hsa-miR-30a	hsa-miR-622
hsa-miR-141	hsa-miR-29b-1*
hsa-miR-181d	hsa-miR-513c
hsa-miR-892b	hsa-miR-874
hsa-miR-193a-5p	hsa-miR-766
hsa-miR-513b	hsa-miR-31
hsa-miR-200a	hsa-miR-193b*
hsa-miR-148b	hsa-miR-30d
hsa-miR-429	hsa-miR-370
hsa-miR-152	hsa-miR-139-5p
hsa-miR-532-5p	hsa-miR-765
hsa-miR-140-5p	hsa-miR-423-3p
hsa-miR-744	hsa-miR-125b-2*
hsa-miR-30c-2*	hsa-miR-663
hsa-miR-183	hsa-miR-630
hsa-miR-196b	hsa-miR-222
hsa-miR-362-5p	hsa-miR-423-5p
hsa-miR-665	hsa-miR-501-5p
hsa-miR-19a	hsa-miR-345
hsa-miR-409-3p	hsa-miR-23a*
hsa-miR-875-5p	hsa-miR-181a
hsa-miR-130b	hsa-miR-486-5p
hsa-miR-492	hsa-miR-145*
hsa-miR-505	hsa-miR-801
hsa-miR-23b*	hsa-miR-181b

hsa-miR-629*	hsa-miR-379
hsa-miR-181c	hsa-miR-923
hsa-miR-382	hsa-miR-198
hsa-miR-30b*	hsa-miR-662
hsa-miR-769-5p	hsa-miR-574-5p
hsa-miR-487b	hsa-miR-361-5p
hsa-miR-941	hsa-miR-500
hsa-miR-1236	hsa-miR-654-5p
hsa-miR-770-5p	hsa-miR-224
hsa-miR-106b*	
hsa-miR-183*	

Table A1. List of miRNAs deregulated in CAFs when compared to normal fibroblasts in FFPE samples.

	Down-regulated	Up-regulated
Fold Change > 10	hsa-miR-19a	Nil
	hsa-miR-575	
	hsa-miR-205	
	hsa-miR-663	
	hsa-miR-26b	
	hsa-miR-513a-5p	
	hsa-miR-630	
	hsa-miR-483-5p	
	hsa-miR-20b	
	hsa-miR-376a	
	hsa-miR-30b	
	hsa-miR-423-5p	
	hsa-miR-134	
	hsa-miR-185	
	hsa-miR-150*	
	hsa-miR-29b-1*	
	hsa-miR-768-5p	
	hsa-miR-765	
	hsa-miR-572	
	hsa-miR-200b	
	hsa-miR-200c	
	hsa-miR-376c	
	hsa-miR-203	
	hsa-miR-374b	
	hsa-miR-188-3p	

hsa-miR-18a
hsa-miR-484
hsa-miR-7
hsa-miR-191*
hsa-miR-940
hsa-miR-31*
hsa-miR-425
hsa-miR-370
hsa-miR-671-5p
hsa-miR-20a
hsa-miR-487b
hsa-miR-138
hsa-miR-625
hsa-miR-361-3p
hsa-miR-22*
hsa-miR-98
hsa-miR-132
hsa-miR-30d
hsa-miR-151-5p
hsa-miR-101
hsa-miR-495
hsa-miR-23b
hsa-miR-30a
hsa-miR-361-5p
hsa-miR-886-3p
hsa-miR-376b
hsa-miR-137
hsa-miR-107
hsa-let-7g

	hsa-miR-15b	
	hsa-miR-130a	
	hsa-miR-135a*	
	hsa-miR-188-5p	
	hsa-let-7f	
Fold change > 5	hsa-miR-29c	hsa-miR-296-3p
	hsa-miR-19b	hsa-miR-889
	hsa-miR-939	hsa-miR-452*
	hsa-miR-379	hsa-miR-609
	hsa-miR-801	hsa-miR-654-5p
	hsa-miR-503	hsa-miR-106b*
	hsa-miR-16	
	hsa-miR-1224-5p	
	hsa-miR-106b	
	hsa-let-7d	
	hsa-miR-193b*	
	hsa-miR-424	
	hsa-miR-768-3p	
	hsa-miR-155	
	hsa-miR-1225-5p	
	hsa-miR-10a	
	hsa-miR-193b	
	hsa-miR-874	
	hsa-miR-181a	
	hsa-miR-17	
	hsa-miR-1225-3p	
	hsa-miR-1228	
	hsa-miR-574-3p	

hsa-miR-210
hsa-miR-214
hsa-let-7e
hsa-miR-28-5p
hsa-miR-92a
hsa-let-7a
hsa-miR-25
hsa-miR-93
hsa-miR-365
hsa-miR-103
hsa-miR-21
hsa-miR-10b
hsa-miR-923
hsa-miR-494
hsa-miR-99b
hsa-miR-125a-5p
hsa-miR-31
hsa-miR-638
hsa-miR-196a
hsa-miR-29b
hsa-let-7i
hsa-miR-27a
hsa-miR-24
hsa-miR-29a
hsa-miR-27b
hsa-miR-320
hsa-miR-199b-3p
hsa-miR-34a
hsa-miR-23a

	hsa-miR-21*	
	hsa-miR-299-5p	
	hsa-let-7b	
	hsa-miR-193a-3p	
	hsa-miR-377	
Fold change < 5	hsa-miR-221*	hsa-miR-541
	hsa-miR-152	hsa-miR-769-3p
	hsa-miR-409-3p	hsa-miR-639
	hsa-miR-218	hsa-miR-578
	hsa-miR-1234	hsa-miR-191
	hsa-miR-513b	hsa-miR-660
	hsa-miR-574-5p	hsa-miR-520h
	hsa-miR-26a	hsa-let-7d*
	hsa-miR-99a	hsa-miR-758
	hsa-let-7c	hsa-miR-130b*
	hsa-miR-125b	hsa-miR-615-3p
	hsa-miR-151-3p	hsa-miR-20b*
	hsa-miR-141	hsa-miR-593
	hsa-miR-30e	hsa-miR-30b*
	hsa-miR-331-3p	hsa-let-7b*
	hsa-miR-22	hsa-miR-621
	hsa-miR-100	hsa-miR-554
	hsa-miR-202	hsa-miR-1236
	hsa-miR-766	hsa-miR-548b-5p
	hsa-miR-181b	hsa-miR-326
	hsa-miR-371-5p	hsa-miR-606
	hsa-miR-622	hsa-miR-539
	hsa-miR-15a	hsa-miR-101*

hsa-miR-145	hsa-miR-647
hsa-miR-127-3p	hsa-miR-15a*
hsa-miR-324-5p	hsa-miR-30d*
hsa-miR-34b*	hsa-miR-610
hsa-miR-222	hsa-miR-662
hsa-miR-654-3p	hsa-miR-641
hsa-miR-221	hsa-miR-644
hsa-miR-130b	hsa-miR-542-3p
hsa-miR-128	hsa-miR-548c-5p
hsa-miR-299-3p	hsa-miR-525-3p
hsa-miR-197	hsa-miR-212
hsa-miR-381	hsa-miR-183*
hsa-miR-199a-5p	hsa-miR-26a-2*
hsa-miR-193a-5p	hsa-miR-557
hsa-miR-1238	hsa-miR-125a-3p
hsa-miR-224	hsa-miR-182*
hsa-miR-324-3p	hsa-miR-302b*
hsa-miR-140-3p	hsa-miR-922
hsa-miR-30c	hsa-miR-490-5p
hsa-miR-140-5p	hsa-miR-30c-1*
hsa-miR-382	hsa-miR-450b-3p
hsa-miR-342-3p	hsa-miR-422a
hsa-miR-590-5p	hsa-miR-548a-3p
hsa-miR-455-3p	hsa-miR-623
hsa-miR-493*	hsa-miR-562
hsa-miR-196b	hsa-miR-1229
hsa-miR-877*	hsa-miR-625*
	hsa-miR-92b*
	hsa-miR-770-5p

hsa-miR-576-3p
hsa-miR-448

Table A2. List of miRNAs deregulated in breast fibroblasts that were grown with MCF-7 cells compared to breast fibroblasts that were grown with HB2 cells.

Down-regulated	Up-regulated
hsa-let-7f	hsa-miR-125a-3p
hsa-let-7a	hsa-miR-623
hsa-miR-26b	hsa-miR-30c-1*
hsa-miR-199b-3p	hsa-miR-557
hsa-miR-365	hsa-miR-662
hsa-let-7e	hsa-let-7b*
hsa-let-7d	hsa-miR-1229
hsa-let-7g	hsa-miR-625*
hsa-miR-10b	hsa-miR-654-5p
hsa-miR-15b	hsa-miR-769-3p
hsa-miR-107	
hsa-miR-10a	
hsa-miR-30b	
hsa-miR-16	
hsa-miR-23a	
hsa-miR-23b	
hsa-let-7c	
hsa-miR-100	
hsa-miR-20a	
hsa-miR-98	
hsa-miR-125b	
hsa-miR-21	
hsa-miR-34a	
hsa-miR-151-5p	
hsa-miR-29b	
hsa-miR-19b	

hsa-miR-768-5p
hsa-miR-27a
hsa-miR-15a
hsa-miR-196a
hsa-miR-214
hsa-miR-29a
hsa-miR-30c
hsa-miR-27b
hsa-miR-331-3p
hsa-miR-103
hsa-miR-17
hsa-miR-125a-5p
hsa-miR-200b
hsa-miR-24
hsa-miR-29c
hsa-miR-25
hsa-miR-455-3p
hsa-miR-199a-5p
hsa-miR-130a
hsa-miR-886-3p
hsa-miR-99a
hsa-let-7b
hsa-miR-128
hsa-miR-425
hsa-miR-28-5p
hsa-miR-34b*
hsa-miR-361-3p
hsa-miR-768-3p
hsa-miR-424

hsa-miR-132
hsa-miR-106b
hsa-miR-93
hsa-let-7i
hsa-miR-26a
hsa-miR-374b
hsa-miR-20b
hsa-miR-101
hsa-miR-193a-3p
hsa-miR-152
hsa-miR-140-5p
hsa-miR-342-3p
hsa-miR-185
hsa-miR-513a-5p
hsa-miR-221
hsa-miR-200c
hsa-miR-193b
hsa-miR-625
hsa-miR-155
hsa-miR-22*
hsa-miR-127-3p
hsa-miR-324-5p
hsa-miR-376c
hsa-miR-196b
hsa-miR-145
hsa-miR-940
hsa-miR-99b
hsa-miR-30e
hsa-miR-151-3p

hsa-miR-140-3p
hsa-miR-19a
hsa-miR-574-3p
hsa-miR-92a
hsa-miR-409-3p
hsa-miR-22
hsa-miR-376a
hsa-miR-205
hsa-miR-210
hsa-miR-494
hsa-miR-484
hsa-miR-324-3p
hsa-miR-197
hsa-miR-130b
hsa-miR-30a
hsa-miR-141
hsa-miR-193a-5p
hsa-miR-513b
hsa-miR-382
hsa-miR-487b

Table A3. List of miRNAs that were up-regulated or down-regulated consistently in both the FFPE sample and cell co-culture model.

Gene symbols								
PLOD2	ACSL3	TMEM2	PDHX	RANBP9	CCNJ	INHBB	CHORDC1	PGR
NAB1	COPS2	HPGD	KIAA1539	GNPNAT1	SMAD1	FLJ14213	PPP1R15B	DMXL1
NHS	CHSY1	STRBP	SLC6A6	KCNK1	EAF1	BICD2	CAMSAP1	ZFHX4
ARHGAP26	STAC2	TP53INP2	MAP1A	PLCB1	CHFR	ATP1A2	ASPN	SPOCK2
ARMCX2	KCNJ2	HIAT1	PELI2	SALL1	SNN	BLOC1S2	SLC24A4	ADAM12
CDK6	C10orf137	PIK3R3	KIAA0528	FLJ20160	ERO1LB	HOXA5	ACVR1C	B4GALT1
OSBPL11	BTBD7	ITPR1	COL22A1	COL1A2	RCN1	PGRMC2	WNK3	CCND2
ALS2CR2	CACNA1C	MTX2	COL19A1	GTF3C2	SENP5	BCL11A	TRPS1	DNMT3B
YPEL1	GGA3	ACBD5	ARHGAP21	USP3	CDH4	SSH2	CDC2L5	ZNF238
HMGA2	POM121	TTC13	PAK2	SLC12A2	FA2H	SUV420H2	FAM45A	NUP153
TP53INP1	ZNF462	ATPAF1	ZNF410	SLC25A20	SRP19	PDGFRA	KCNH7	PIM1
ARPP-19	ITGA5	SSX2IP	RANBP10	PRELP	TLK2	FBXL19	UBE2E2	RBM24
HIPK2	PRKAG2	DAPK1	UTX	ABHD2	MAP3K2	ZDHHC18	PHF6	KLF4
ULK2	ADM	PHLDB2	COPS7B	SACS	NAP1L5	ALDH5A1	EPHA7	USP25
ULK1	RCN2	FBXO28	TRIB2	BHLHB2	USP15	STK39	OSBPL2	SGEF
DYRK1A	ADAM19	TFAP2C	ESR1	SMAD4	LEF1	SERP1	RAP2C	PCDH18

NLK	DCDC2	AKAP7	MAP2	PTP4A1	SRGAP1	RLF	CTDSP2	ABR
MAP1B	DEPDC1	NFE2L3	EHD1	PAN3	DLG5	EIF5	FAM49B	HAO1
ARPP-21	CEBPG	MTPN	WNT5A	PAPD4	ARID3A	SOX17	KPNA2	HIPK1
C9orf40	BHLHB5	COX5A	DEPDC1B	CPEB2	NEK6	ALS2	MXI1	DMRT3
EPC2	EZH2	CPSF2	CD200	HAS3	RPS6KA6	DUSP5	UBE4B	PDCD10
PSD3	FAM44B	TNKS2	CDH11	CPEB3	ATP11C	ZDHHC6	DCBLD2	CTGF
HOXD13	COL5A1	JAG1	CDK2AP1	PTER	EPHA2	MMP14	SLC38A2	XKRX
UBE2G1	RTN1	DNAJA2	TRPC3	CREBBP	PRKCD	GRSF1	ABCC4	COL10A1
KCNQ4	FBXO11	FAM46C	NAGPA	TOB1	CELSR1	KIAA1468	SYT10	PITPNC1
ARHGEF12	KIAA0152	ZNF217	PCDH9	KIAA1598	ENC1	UBE2D1	RGS4	ABCA1
MARK1	CUGBP2	PGM2L1	MRAS	SLC4A4	LRRTM4	SEMA6D	G3BP2	KIAA1128
RNF138	CPEB4	MAB21L1	PTEN	E2F7	LSM11	MUM1L1		

Table A4. List of predicted miR-26b targets identified in common by both PicTar and TargetScan

Up-regulated proteins (Gene symbols)							
TNKS1BP1	ALDH18A1	CACYBP	STAT2	SMAD2	MTHFS	ACACA	SSRP1
APIP	PSMB4	CHI3L1	SYNE1	SMAD9	PA2G4	UGGT1	MYH10
COL12A1	FAM96B	HSPA5	SFXN3	PSMA1	ACOT7	PGAM5	SH3GL1
CPSF7	MSN	GNPDA1	RPS27A	COL3A1	VCP	HNRNPK	HLA-C
USP19	GLUD1	RSU1	UBC	PGM1	PGAM1	CNDP2	HIST1H1E
COL6A3	GLUD2	COL6A2	UBB	HADHB	ASPH	HSPA8	PDHA1
ILF2	VCL	CCT8	UBA52	LSM5	HNRNPA3	ANXA2	LOC79064
COL6A1	MAP1B	HSPB1	VPS26A	HSPA4	PRDX4	ANXA2P2	HSPG2
DAB2	FN1	YAP1	SERBP1	DEK	RPS20	BCLAF1	IMPDH2
MIA3	NANS	NACA	FIS1	GSTO1	GSTP1	AKAP2	ARHGAP1
ACSL4	ACAT1	HEATR2	NCKAP1	PPAP2B	IKBKAP	KTN1	ACTR1A
LDHA	EIF3C	SF3B14	TOP1	COPB2	OTUB1	RAB5C	IPO5
UAP1L1	EIF3CL	PDCD10	PPT1	PSMD13	MTHFD1	TMSB4X	ACADM
IFITM2	SSB	SOD2	CST6	ACTBL2	MAP4	ALDH6A1	NT5E
IFITM3	NME1	KHDRBS1	CSTB	GSN	NCL	CKAP5	SH2D4A
IFITM1	APOOL	FAM120A	EEF1B2	SNX1	CAST	ESD	PFN1
SEC24C	GNL3L	STRAP	CLIC4	LMNB1	LMNA	PGK1	TCEA1

SCAMP1	PDCD4	ENO1	NPM3	SET	TMPO	ACOT2	RPS15
HSPD1	KHSRP	AK2	EPHB4	MRPS9	GAPDH	ACOT1	PDIA4
CTSD	RAB2A	ALDOA	PPIF	LARS	CUL4B	RUVBL1	ANXA11
CHMP4A	LGALS1	DLST	MVP	IFI16	ATIC	CAV1	ACTB
GANAB	PDCD6IP	MTDH	RPS6KA2	SCAMP3	CYB5R3	PTMS	ACT
CFL2	CCDC6	CTTN	A2M	STMN1	DARS	IDH1	ACTG1
LARS2	TST1	TKT	RBM25	LRRFIP1	CRK	GLYR1	PDHB
ACTL6A	TUBB6	CFL1	PRKDC	SDHA	PLAA	RPS2	MAPK1IP1L
STOML2	GLMN	CTSB	ANXA1	COPA	CD3EAP	CAMK2D	LARP1
PRKCSH	KIAA1407	XRCC5	IGF2BP2	RPL22	EEF1D	ATP6V1G1	
LONP1	AK3	HMGB2	AHNAK2	MDH1	KIF5B	ATP6V1G2	
HSPA9	FAHD2A	EHD2	HNRNPC	HNRNPCL1	RAD23B	LRRC47	

Down-regulated proteins (Gene symbols)							
TMEM119	SLC7A5	CHMP6	CTGF	PAK1	URM1	PNKP	TRAM1
NUDCD3	NSDHL	TRMT1	CSPG4	PKN1	HDAC2	LRRFIP1	PSMD8
FAM3C	H2AFZ	BICD2	STRBP	TBC1D2	AARS2	C16orf62	KIF23
CLTB	H2AFV	NIPSNAP3A	PRG2	SUPT6H	TRIOBP	DCTPP1	ARPC5

THOC5	RAB18	TUBA4A	PXDN	ZCCHC9	SLC3A2	PHF23	IST1
DDX20	FAM134C	RNMTL1	DNAJB11	PPP6C	MCM5	HMOX2	PLA2G4A
CASP6	SUMO3	NKRF	GALNT10	TOE1	TRIM16	CDH6	LAMB1
IDH3B	ZNF598	BLMH	CASP3	KRT10	JAGN1	HELLS	NONO
STIM1	COL5A1	ZFPL1	UAP1	EIF2B4	BAG3	FNDC3B	ATXN2L
CHCHD2	MORF4L2	ROCK2	YIF1B	POLD2	RRP1	SSR3	KRT1
MLF2	GATAD2A	SC4MOL	DHX38	IKBIP	MAP2K1	SMTN	TPX2
GBP1	CDK2	SLAIN2	ARPC1A	ETHE1	S100A4	PACSIN2	WDR3
MAD2L1	PTPN12	SLC9A3R1	PREB	CBS	ANXA4	NEXN	HPP1
RBM7	SUN2	VPS11	SEPT8	CASC4	GOLGA4	OSBPL3	TRIP12
HIST3H2BB	CRELD2	MEN1	FAM160B1	PSMD14	CIAPIN1	EGFR	NAA15
GRD1	SEC62	SCG2	PNPO	SMARCA2	DDRGK1	TRAP1	SPTBN1
GSR	OSTF1	XAB2	CARF	PLRG1	LOX	KIAA0368	MTHFD1L
TPM4	FAM107B	NAE1	CDKN2AIP	SEC31A	CORO1C	RANBP9	

Table A5. Protein identified by mass spectrometry to be differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts.

Up-regulated proteins (Gene symbols)							
Predicted miR-26b targets							
TNKS1BP1	COL12A1	CPSF7	DAB2	ACSL4	SCAMP1	MAP1B	FN1
APOOL	PDCD6IP	CCDC6	TST1	TUBB6			
Other downstream targets							
APIP	USP19	COL6A3	ILF2	COL6A1	MIA3	LDHA	UAP1L1
IFITM2	IFITM3	IFITM1	SEC24C	HSPD1	CTSD	CHMP4A	GANAB
CFL2	LARS2	ACTL6A	STOML2	PRKCSH	LONP1	ALDH18A1	PSMB4
FAM96B	MSN	GLUD1	GLUD2	VCL	NANS	ACAT1	EIF3C
EIF3CL	SSB	NME1	GNL3L	PDCD4	KHSRP	RAB2A	LGALS1

Down-regulated proteins (Gene symbols)							
Predicted miR-26b targets							
DDX20	CASP6	GBP1	MAD2L1	GRD1	GSR	TPM4	TPX2
H2AFZ	H2AFV	RAB18	FAM134C	COL5A1	WDR3	BICD2	NIPSNAP3A
NKRF	MSMO1	SLAIN2					
Other downstream targets							
TMEM119	NUDCD3	FAM3C	CLTB	THOC5	IDH3B	STIM1	CHCHD2

MLF2	RBM7	HIST3H2BB	ARPC5	SLC7A5	NSDHL	SUMO3	ZNF598
MORF4L2	GATAD2A	CDK2	PTPN12	SUN2	CRELD2	SEC62	OSTF1
FAM107B	IST1	CHMP6	TRMT1	TUBA4A	RNMTL1	BLMH	ZFPL1
ROCK2							

Tables A6. Proteins that were differentially expressed between con$^{k/d}$ and 26$^{k/d}$ fibroblasts using an FDR cut off value of <0.2

COL-12A: Transcript ID: ENST00000322507

Site type: 7mer-1A
Seed region: 173 to 179
3' UTR (2232 bp):
5'acgtgcttattaggaaatgagtctgtatggaaatctcaccacagataatggttaacgaaccgggtcgacatcacaaaggagggtgga
gactctttttact**aacttgaa**tgagacaaaagcagtggtgtcagtttataatcctgatgcatttcagtaataatgtagaaaaacattattttaaa
aaagttccaacacacagccatgagg

TNKS1BP1 : Transcript ID: ENST00000358252
Site type: 7mer-1A
Seed region: 410 to 416
3' UTR (440 bp):
5'gttcttcccacctggcagtctcaggcagtgcccattcctgtggggtccctgggtgaggagacggctggagccccaccatgccccagg
ctgcagcctctgtcccctccacctctgaggagcgtctggggaggcacatttatgcactttgtatcaccctccgaactccccccacaccttcc
cttccctggatttcatcactagtggttgaaggttttgtcccttcctctcctccttcccctctccctctctgcttcctcctccagcctcccttgggttttctttt
gataccaatttatagcatttttataaaagcctttgatttttataatgggtgggactgtatccctgcctcaccccaggtctccgtctgccccgcca
ggtaccccacagagaccaatgacattttgc**cacttgaa**acaataaataaagtttttgggaa

CPSF7 : Transcript ID: ENST00000340437
Site type: 7mer-1A
Seed region: 1202 to 1206
3' UTR (2056 bp):
5'tttacttattgtgggataacctggcactagtaggcaggtaaagtcacaaatttggtgtcttttcacctttgactgttgacttaatagctcctctc
actctgcctggagatacttcctgcctcagatgaggagccagaagaaacagagccc**gacttgaa**tgaactcagctcagagttctaagg
accagcattctgggggccattttctctacaggcaaatggaattgcttttccataacatccaaattgtaatgtggttgctgctgaaggaggagg
cagcagcgaggtcctgcggtaccca

Table A7. MiR-26b binding site type, position within 3'UTR, and sequence of 3' UTR that was cloned into the luciferase reporter plasmid for COL12A1, CPSF7 and TNKS1BP1. The highlighted nucleotide sequences are the seed regions.

Gene symbol	Forward primer	Reverse primer
COL12A1	5’ gcgcatgagctctgagtctgtatggaaatctacc	5’ cgcgattctagacctcatggctgtgtgttgg
TNKS1BP1	5’ gcgcatgagctcggttttgtcccttcctctcc	5’ cgcgattctagactttatttattgtttcaagtggc
CPSF7	5’ gcatatgagctcatagctcctctcactctgcc	5’ cgctattctagactcctccttcagcagcaacc

Table A8. Sequences of forward and reverse primers used for cloning the 3’ UTRs of COL12A1, TNKS1BP1 and CPSF7