The secretome of myofibroblasts and its significance in gastric cancer

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Christopher Holmberg

September 2009

This thesis is the result of my own work. The material contained in the thesis has not been presented, nor is currently being presented, either wholly or in part for any other degree or other qualification. The research was performed in the Department of Physiology, School of Biomedical Sciences, University of Liverpool. All other parties involved in the research presented here, and the nature of their contribution, are listed in the 'Acknowledgements' section of this thesis.

Acknowledgements

I would first like to say a big thank you to my supervisor Prof. Andrea Varro for her continuous support, guidance, enthusiasm and encouragement. I have learned an incredible amount in what seems like a very short time, and I am certain that these experiences will serve me well. Thank you for giving me the opportunity to develop my passion for science and research.

I would also like to thank Prof. Graham Dockray for his insight and wisdom throughout the course of my research and for teaching me the art of scientific writing. Thank you to Dr. Peter Hegyi and colleagues at the Univeristy of Szeged for help in culturing the myofibroblasts used in this study. Thanks also to Dr. Cédric Duval for teaching me many of the techniques that I used to get me started with this work. Thank you to Dr. Islay Steele for help and long discussions on gene expression arrays and bioinformatics. Thanks to Dr. Roz Jenkins and Dr. Vicki Elliott for all of their help with mass spectrometry. Thank you to Dr Bahram Ibrahimi and Dr. Lucille Rainbow for their help in perofrming the Affymetrix GeneChip arrays. Special thanks to Prof. Kris Gevaert, Dr. Bart Ghesquière, Veerle Vandenbussche, Francis Impens and colleagues from the University of Ghent for their collaboration with the COFRADIC experiments. Thanks you to all members of Green Block, past and present, for your friendship and support during my PhD.

A huge thank you to the Wellcome Trust for funding this work.

Thank you to my Mum, Dad, Katie and Nan for always cheering me on. Last, but definitely not least, thank you to my wife Suzanne for her constant love, support and encouragement, and for putting up with me as I was writing this thesis!

Contents

List of Figures

List of Tables

Abbreviations

The abbreviations used in this thesis are listed below.

2D-PAGE	Two-dimensional polyacrylamide gel electrophoresis
A	Absent
A2M	Alpha-two-macroglobulin
ACN	Acetonitrile
ANM	Adjacent non-cancer tissue-derived myofibroblast
ANOVA	Analysis of variance
APP	Amyloid precursor protein
Arg	Arginine
α-SMA	Alpha smooth muscle actin
BMDC	Bone marrow derived cell
BMP4	Bone morphogenic protein 4
CAM	Cancer-associated myofibroblast
CEX	Cation exchange chromatography
CHCA	Alpha-cyano-4-hydroxycinnamic acid
CM	Conditioned media
COFRADIC	Combined fractional diagonal chromatography
D3-NHS	Trideutero-N-hydroxysuccinimide
DAPI	4',6-diamidino-2-phenylindole
DMEM	Dubecco's modified Eagle's medium
DTT	Dithiothreitol
EACA	Epsilon-aminocaproic acid
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid

EGF	Epidermal growth factor
EGFR	Epidermal growth factor receptor
EMI	EMILIN-like domain
EMT	Epithelial-mesenchymal transition
FAS1	Fascilin 1-like domain
FBS	Fetal bovine serum
FGF	Fibroblast growth factor
FITC	Fluorescein isothiocyanate
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GI	Gastrointestinal
GIST	Gastrointestinal stromal tumour
HGF/SF	Hepatocyte growth factor/scatter factor
Hh	Hedgehog
HPLC	High performance liquid chromatography
HPRD	Human protein resource database
HRP	Horseradish peroxidase
IAA	Iodoacetamide
ICAT	Isotope-coded affinity tags
IDA	Information-dependent acquisition
IEF	Isoelectric focussing
IFN	Interferon
IGF	Insulin-like growth factor
IGFBP	Insulin-like growth factor binding protein
IgG	Immunoglobulin G
IL	Interleukin

IMAC	Immobilised metal ion chromatography
IPF	Idiopathic Pulmonary Fibrosis
IPG	Immobilised pH gradient
iTRAQ	Isobaric tagging for relative and absolute quantitation
KGF	Keratinocyte growth factor
KO	Knock-out
LC	Liquid chromatography
LOH	Loss of heterozygosity
LRP-1	Pro low-density lipoprotein receptor-related protein 1
Lys	Lysine
M	Marginal
MALDI	Matrix assisted laser desorption ionization
MALT	Mucosa-associated lymphoid tissues
MAS5	Microarray suite 5.0
MMP	Matrix metalloproteinase
MMTS	Methyl methanethiosulfonate
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
NGF	Nerve growth factor
Oligo	Oligonucleotide probe
P	Present
PAGE	Polyacrylamide gel electrophoresis
PAI-1	Plasminogen activator inhibitor-1
PAI-2	Plasminogen activator inhibitor-2
PBS	Phosphate buffered saline

PDGF	Platelet-derived growth factor
PFA	Paraformaldehyde
RGD-CAP	RGD-containing collagen-associated protein
rh	Recombinant human
RIPA	Radio immuno precipitation assay
ROS	Reactive oxygen species
RP-LC	Reverse phase liquid chromatography
SEM	Standard error of the mean
SCX	Strong cation exchange
SDF-1	Stromal derived factor-1
SDS	Sodium dodecyl sulfate
SFM	Serum free media
SILAC	Stable isotope labelling by amino acids in cell culture
SLRP	Small leucine-rich proteoglycan
SNP	Single nucleotide polymorphism
SS	Signal sequence
TBS	Tris-buffered saline
TBST	Tris-buffered saline with Tween
TCA	Trichloroacetic acid
TCEP	Tris-(2-carboxyethyl)phosphine
TEAB	Triethylammonium bicarbonate
TFA	Trifluoroacetic acid
Tgfbr2	Transforming growth factor beta receptor 2
TGFα	Transforming growth factor alpha
TGFβ	Transforming growth factor beta

TIMP	Tissue inhibitor of matrix metalloproteinases
TNBS	2,4,6-trinitrobenzenesulphonic acid
TNFα	Tumour necrosis factor alpha
TOF	Time of flight
TRITC	Tetramethyl rhodamine iso-thiocyanate
uPA	Urokinase plasminogen activator
uPAR	Urokinase plasminogen activator receptor
VEGF	Vascular endothelial growth factor

Publications and Presentations

Oral Presentations

5th International Conference of Postgraduate Medical Students, Hradec Kravlove Medical School, Charles University, Czech Republic - 26-29 November 2008. "The secretome of myofibroblasts and significance in gastric cancer"

Bristol-Liverpool Postgraduate Symposium, Liverpool - September 2008. "The secretome of myofibroblasts and significance in gastric cancer"

Poster Presentations

Keystone Symposium – Extrinsic control of tumour genesis and progression, Vancouver, BC - 15-20 March 2009.

Christopher Holmberg, Cedric Duval, Peter Hegyi, Gyorgy Lazar, Roz Jenkins, Islay Steele, Graham Dockray, Andrea Varro "TGFbigh3 is a target of urokinase plasminogen activator in gastric myofibroblasts: implications for tumor and stromal cell migration".

Abstract

The microenvironment is important in regulating the behaviour of cancer cells. Myofibroblasts are a key stromal cell type with important roles in defining microenvironments in both health and disease by secreting proteins including growth factors, proteases and extracellular matrix (ECM) components. In the case of gastric cancer, the myofibroblast population is significantly expanded compared with the adjacent non-cancer tissue, and with healthy gastric tissue. It has been shown in other cancers that cancer-associated myofibroblasts (CAMs) can promote tumour development and progression via secreted proteins such as hepatocyte growth factor (HGF) and stromal cell-derived factor 1 (SDF-1). In the stomach it has been shown that changes in the signalling between myofibroblasts and the adjacent epithelial cells are important in mediating the effects of the gastric pathogen *Helicobacter pylori.* Proteomic studies showed that increased secretion of matrix metalloproteinase 7 (MMP7) by the epithelial cells during *H. pylori* infection causes cleavage of insulin-like growth factor binding protein-5 (IGFBP-5) in the myofibroblast secretome, resulting in increased bioavailability of IGF-II. This thesis is the first comprehensive study of gastric myofibroblast secretomes. The aim of this thesis was to define the differences between gastric CAMs and their adjacent non-cancer tissue myofibroblast (ANM) counterparts.

This study made use of the unique availability in the group of cultures of myofibroblasts derived from human gastric carcinoma samples and adjacent non-cancer tissue. *In vitro* assays revealed that CAMs were more migratory than ANMs, both in response to IGF-II and when unstimulated. Comparison of CAM and ANM transcriptomes by gene expression array revealed significant differences, particularly in genes related to transforming growth factor beta (TGFβ) signalling. Myofibroblast

conditioned media was used to stimulate AGS cells, a gastric cancer cell line. This revealed that CAM conditioned media induced greater AGS migration and invasion than ANM conditioned media. Gene expression array analysis of the stimulated AGS cells showed that these phenotypic differences were accompanied by significant changes in gene expression. Again, many of the differences were in genes involved in TGFβ signalling. These experiments suggested that the most important differences between CAMs and ANMs were in their secretomes.

Quantitative proteomic analysis of the myofibroblast secretomes and proteomes using the isobaric tagging for relative and absolute quantitation (iTRAQ) system revealed that a large proportion of the secreted proteins were differentially abundant in CAM media compared with ANM media. Most of these proteins, including ECM components and protease inhibitors, were less abundant in the CAM media. Only a small number of proteins, including proteases, were more abundant in CAM media than in ANM media. These results suggested a higher level of proteolytic activity in the CAM secretome. Analysis of the secretome of one CAM and ANM pair by combined fractional diagonal chromatography (COFRADIC) revealed greater proteolytic processing of several proteins in CAM media compared with ANM media. These included proteins with roles in TGFβ signalling, ECM components and proteases.

One protein which was often identified in proteomic experiments as being less abundant in CAM media than ANM media was TGFβ ig-h3. This is a TGFβ-induced extracellular protein with a known role in cellular adhesion. It was shown by Western blotting that in several cases βig-h3 was either less abundant or more degraded in CAM media than ANM media. With proteomic and Western blot data taken together

there was evidence for either reduced abundance or increased degradation of βig-h3 in 9 out of 11 CAMs compared with their matching ANMs. *In vitro* assays revealed that βig-h3 inhibited AGS cell migration in response to myofibroblast conditioned media. It was also shown that βig-h3 could be cleaved by plasmin, and that inhibiting urokinase plasminogen activator (uPA), which can convert plasminogen into active plasmin, also inhibited AGS cell migration in response to myofibroblast conditioned media.

Together these data show that there are significant differences between gastric CAMs and ANMs. Many of these differences involve genes and proteins with known roles in TGFβ signalling, and the key differences are in the secretome. These differences, including a higher level of proteolytic activity in CAM media, result in a more migratory response in AGS cells. This suggests that CAMs could be important in defining a microenvironment in cancer which might promote a more aggressive cancer phenotype.

Chapter 1

Introduction

1.1 Microenvironments - An historical perspective

1.1.1 Homeostasis and hormones

The concept that physiological and patho-physiological processes are regulated by an internal environment was introduced in 1865 by Claude Bernard who wrote "The constancy of the internal environment is the condition for a free and independent life" (Noble, 2008; Toledo-Pereyra, 2009). This has become the underlying principle of modern concepts of homeostasis. Bernard had studied the control of blood glucose, but it is only in the last century that the factors which regulate the internal environment have been elucidated. In 1902, Bayliss and Starling identified the first endocrine system when they showed that a blood-borne factor, secretin, stimulates pancreatic secretions in response to acid in the duodenum (Bayliss and Starling, 1902). Starling then gave the name 'hormone' to all similar substances that act via delivery in the blood stream to regulate the internal environment. Much later, the idea emerged that there might be analogous mechanisms at the level of cell-cell interactions within a tissue, and the term paracrine was introduced to describe this. In some cases the distinction between endocrine and paracrine can become blurred, since certain hormones can act both on distant and local cells.

1.1.2 Neurotransmitters

Another major advancement in the understanding of signalling molecules came with the discovery of acetylcholine (Dale, 1914) and the demonstration that it acts as a biochemical messenger across synapses (Loewi, 1921). The work of Loewi and Dale addressed a long standing question concerning the nature of synaptic transmission, revealing it to be of a biochemical nature rather than electrical, and winning them the 1936 Nobel Prize in Physiology or Medicine. Subsequent research has revealed

neurotransmitters of several types, including biogenic amines, amino acids, neuropeptides, nitric oxide and carbon monoxide.

1.1.3 Discovery of growth factors

Growth factors are now recognised to act as paracrine signalling molecules. The first growth factor to be identified was neural growth factor (NGF), as the result of work done by Levi-Montalcini and colleagues in the early 1950's (Levi-Montalcini, 1987; Levi-Montalcini and Hamburger, 1951; Levi-Montalcini et al., 1954). This was shortly followed by the discovery of epidermal growth factor (EGF) by Cohen and co-workers, who worked to isolate growth factors from the salivary glands (Carpenter and Cohen, 1979; Cohen, 1959; Cohen, 1962). They originally identified EGF as a peptide which accelerated growth and development of mouse embryos, although further work revealed a myriad of functions for this growth factor (Jorissen et al., 2003). The importance of these discoveries was such that Cohen and Montalcini were awarded the 1986 Nobel Prize in Physiology or Medicine. Their work triggered a flurry of research into the functions of growth factors, which included the identification of the EGF receptor in the 1970's (Das et al., 1977; Pruss and Herschman, 1977). Later, the concept of autocrine signalling, i.e. factors which act on the same cell type as secretes them, was put forward in order to explain the observation that malignant cells have a lower requirement for exogenous growth factors than normal cells (Sporn and Todaro, 1980). These studies also led to the discovery of TGFα and TGFβ, so called because of their apparent 'transforming' effect on normal fibroblasts (Anzano et al., 1982; Roberts et al., 1981). The complex action of TGFβ was highlighted when Moses and colleagues demonstrated that TGFβ could actually inhibit cell growth as well a promote it (Tucker et al., 1984). It

was soon recognised that these growth factors are not unique to malignant cells, but are involved in processes including embryogenesis (Flanders et al., 1991; Thompson et al., 1989) and inflammation (Wahl, 1994). Figure 1.1 highlights the distinctions between endocrine, paracrine and autocrine signalling.

1.1.4 Stem cell niches

Collectively the work on hormones, neurotransmitters and growth factors led to the concept of microenvironments that define cellular function by mechanisms influencing the immediate surroundings of a cell. This has proved to be crucial for understanding the behaviour of many cell types, particularly stem cells. Stem cells are undifferentiated cells defined by their ability to give rise to differentiated cell lineages and self-renew (Becker et al., 1963; Smith, 2001). While embryonic stem cells are capable of giving rise to every cell in an organism, there are various tissue stem cells which exist in the adult with a more restricted capacity to differentiate,

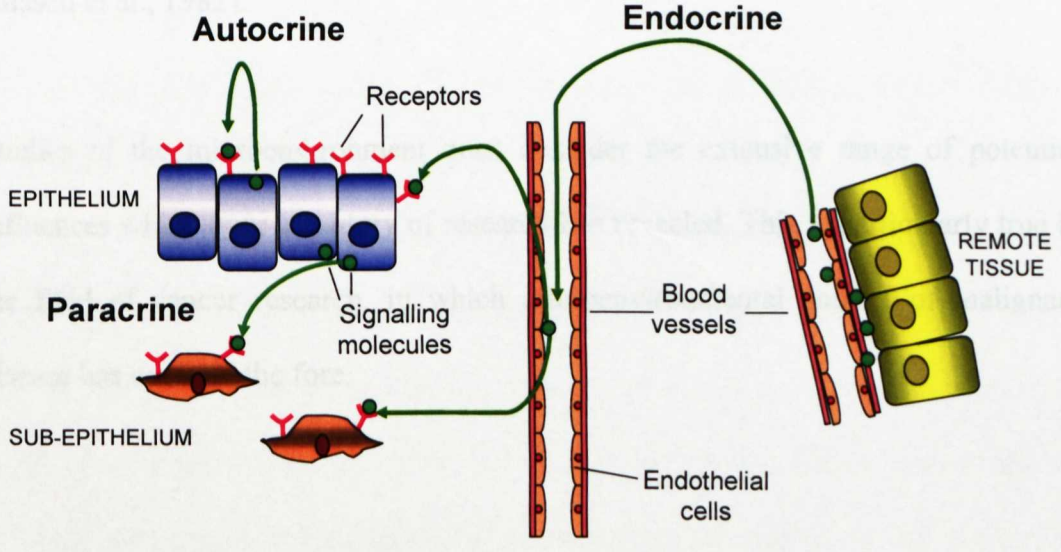


Figure 1.1

Endocrine, paracrine and autocrine signalling
Schematic shows three mechanisms of inter-cellular signalling: endocrine signalling from a remote tissue, paracrine signalling from an adjacent cell, and autocrine signalling from the same cell.

which give rise to differentiated cells within their host organ or tissue. An important development in the understanding of microenvironments came with the proposal of the 'stem cell niche' hypothesis (Schofield, 1978). These specialised microenvironments harbour the tissue stem cells and provide signalling cues which regulate their proliferation and differentiation, thus helping to maintain homeostasis of the tissue (Watt and Hogan, 2000). This paracrine regulation of the niche can come from populations of mesenchymal, endothelial, epithelial and immune cells which exist in close proximity to the stem cell(s) (Fuchs et al., 2004). Research into the tight regulation of the stem cell niche revealed a host of ligands which are produced by the niche cells, including hedgehog, wnt, bone morphogenic proteins and Notch (Artavanis-Tsakonas et al., 1999; Brittan and Wright, 2002; Reya et al., 2003; Taipale and Beachy, 2001). The extracellular matrix (ECM) has also been shown to play an important role in regulating the microenvironment. In 1982, Bissell and colleagues proposed a 'dynamic reciprocity' between the ECM and cell fate (Bissell et al., 1982).

Studies of the microenvironment must consider the extensive range of potential influences which over a century of research has revealed. This is particularly true in the field of cancer research, in which microenvironmental control of malignant disease has come to the fore.

1.2 Main components of microenvironments

Microenvironments and their specific composition uniquely define the specificity of function in organs or regions of organs. It is possible, however, to define the common molecular and cellular constituents of physiological microenvironments.

1.2.1 Fibroblasts

The term 'fibroblast' is often applied in a broad way to describe a ubiquitous class of mesenchymal cells which express vimentin (Franke et al., 1978) and produce extracellular matrix proteins, and in particular collagen. They are important in maintaining the structure of tissues. Different populations of fibroblasts can vary greatly, with some populations of specialised fibroblasts being important regulators of the microenvironment in various stem cell niches (Moore and Lemischka, 2006). They may also give rise to myofibroblasts (see later).

1.2.2 Blood vessels and angiogenesis

Mammalian cells require proximity to blood vessels in order to survive. A distance of greater than approximately 100-200μm becomes prohibitive since this is the diffusion limit of molecular oxygen in the interstitial fluid (Carmeliet and Jain, 2000). Blood vessel formation therefore needs to be a dynamic process as tissues change, for example in wound healing (Folkman and Shing, 1992). There is strict microenvironmental control over the formation of new blood vessels which balances pro- and anti-angiogenic signals. This is sometimes referred to as the 'angiogenic switch', since angiogenesis can be switched 'on' and 'off' depending on the balance of these cues (Folkman, 1995). Metabolic and mechanical stresses, inflammation, immune reactions and discrete signalling molecules can all influence angiogenesis.

Some of the best characterised angiogenic factors are fibroblast growth factors (FGFs), vascular endothelial growth factor (VEGF), platelet-derived growth factor (PDGF), TGFα and TGFβ, tumour necrosis factor alpha (TNFα), hepatocyte growth factor/scatter factor (HGF/SF) and angiogenein (Carmeliet and Jain, 2000; Folkman and Klagsbrun, 1987). Proteases such as matrix metalloproteinases (MMPs, see later) and the plasminogen activator family are also important in allowing newly forming blood vessels to penetrate the extracellular matrix (Andreasen et al., 2000; Egeblad and Werb, 2002).

1.2.3 Inflammatory cells

Circulating leukocytes are normally recruited to specific tissues during wound healing. Leukocytes include circulating neutrophils, monocytes, eosinophils, lymphocytes and basophils, as well as cells which reside permanently in tissues, such as mast cells. In the case of wound healing, neutrophils, and sometimes eosoniphils, are the first leukocytes to be recruited to an acute injury by circulating factors such as complement factor 5a, kallikrein and TGFβ (Singer and Clark, 1999). Neutrophils produce TNFα (Feiken et al., 1995), IL-1α and IL-1β (Hubner et al., 1996), initiating expression of MMPs and keratinocyte growth factor (KGF) by fibroblasts (Chedid et al., 1994). Monocytes, which are precursors of macrophages, are often the next cells to be recruited to the injury. Macrophages are the main source of growth factors and cytokines in the healing wound microenvironment, orchestrating the involvement of endothelial, epithelial and mesenchymal cells.

Inflammation in wound healing is usually self-limiting, but in certain diseases inflammation can become chronic, which can lead to pre-neoplastic changes in the tissue (Coussens and Werb, 2002). The association between inflammation and

cancer has been recognised since 1863, when Virchow postulated that cancer arises from sites of inflammation (Balkwill and Mantovani, 2001).

1.2.4 Growth factors and peptide hormones

There is an impressive array of polypeptide signalling molecules which, amongst others, include hormones, growth factors and cytokines. These molecules are effectors of biological signalling, binding to cell surface receptors. Hormones are more commonly associated with the endocrine system, whilst growth factors and cytokines are paracrine and autocrine signalling molecules. Growth factors are a diverse group of ligands, all of which induce proliferation and/or differentiation responses.

1.2.5 Cytokines

Cytokines are a large group of signalling polypeptides involved in a range of cellular communication events, but predominantly in immune responses. Cytokines may be produced by almost any cell in the body, and as such are important paracrine and autocrine signalling molecules. Examples of cytokines include interleukins (ILs), interferons (IFNs) and tumour necrosis factor (TNF).

Chemokines are a family of chemotactic cytokines, with approximately 41 members in humans (Luster, 1998). Most chemokines contain four cysteine residues, which form two disulphide bridges. Chemokines can be sub-classified into CXC, CC, C and CX3C chemokines, as defined by the number of residues which separate the cysteine residues. Their receptors are G-protein coupled, seven-transmembrane proteins. Chemokines were initially identified as molecules which regulate the

recruitment of leukocytes (Rossi and Zlotnik, 2000), although they now have recognised roles in numerous process, including angiogenesis (Folkman, 1995).

1.2.6 Lipids and small signalling molecules

Lipid hormones are another important class of signalling molecules. Prostaglandins, for instance, are a diverse group of fatty acids derived from arachidonic acid. They have diverse functions in the immune systems, regulating gastric secretions, and the reproductive system, amongst others. Other important features of microenvironments include water, ions, pH and small signalling molecules such as nitric oxide. All of these can have significant effects on cellular function, and all are tightly regulated in order to maintain tissue homeostasis.

1.2.7 Extracellular matrix proteins

The extracellular matrix (ECM) provides structural integrity to tissues, acting as a scaffold for cell interactions and movement, as well as maintaining tissue homeostasis by regulating bioavailability of ions, water and signalling molecules. The ECM is largely proteinaceous, comprising several families of fibrous proteins, such as collagens, and proteoglycans, such as heparin sulphate. This section describes some of the major ECM proteins, though this list is by no means exhaustive.

Collagen type I is the most abundant protein in the human body, being a major component of bones, tendons, skin and extracellular matrix (Kadler et al., 2007; Prockop and Kivirikko, 1995). There are 29 members of the collagen superfamily, and some of the other collagens are major constituents of cartilage, granulation tissue of healing wounds, reticular fibers of connective tissue and the basal lamina, a

layer of specialised ECM on which epithelial cells sit. Collagens form microfibrils consisting of supramolecular helices of five collagen molecules. These microfibrils are the subunit of collagen fibrils.

Fibronectin is a large glycoprotein with roles in a range of biological functions including adhesion, migration, growth, differentiation (Pankov and Yamada, 2002). There are several splice variants of fibronectin, all encoded by a single gene. The two main forms are soluble fibronectin, which is an abundant plasma protein which is deposited at blood clots, and cellular fibronectin which condenses in the ECM (Hynes and Yamada, 1982). Fibronectin contains binding sites for intergrins, collagens and heparin.

Vitronectin is an abundant glycoprotein in serum and ECM. It helps to regulate cellular adhesion and migration by binding integrin via an RGD domain. Vitronectin also binds and stabilises PAI-1 and binds UPAR, thus regulating both adhesion and proteolytic activity (Stefansson and Lawrence, 1996).

Proteoglycans are a large and diverse family of glycoproteins, many of which are associated with ECM in various tissues (Iozzo, 1998; Iozzo and Murdoch, 1996). One of the most abundant proteoglycans is heparin sulphate, which is commonly found in the ECM where it binds and regulates bioavailability of numerous ligands, cations and water. Another important group of proteoglycans are the small leucine-rich repeat proteoglycans (SLRPs), which include decorin, lumican, biglycan and fibromodulin. The SLRPs bind to collagen fibres and are believed to regulate factors such as TGFβ and BMP4 (Iozzo, 1999).

Numerous other proteins can participate in the ECM, largely via interactions with the fibrous ECM proteins such as collagen. Beta ig-h3 is one such protein, which binds collagens using four Fascilin-1-like domains, as well as interacting with cell surface integrins via an RGD domain (Thapa et al., 2007). Proteins such as beta ig-h3 are important in mediating the interactions between cells and the ECM, and as such can have important roles in adhesion and migration.

1.2.8 Proteases

Proteases are a large family of enzymes consisting of 570 members in humans, and accounting for 2% of all human genes (Rawlings et al., 2006). The various members play essential parts in a diverse range of physiological processes, including embryogenesis (Werb and Chin, 1997), wound healing (Lund et al., 1999) and angiogenesis (Pepper, 2001). Proteases represent 5-10% of all drug targets (Overall and Kleifeld, 2006; Turk, 2006), and so are of great interest to pharmaceutical research as well.

All proteins will undergo proteolysis at some stage in their life, whether it be degradation, or proteolytic processing resulting in a specific change in protein function (Lopez-Otin and Overall, 2002). Proteolysis is one of the most ubiquitous post-translational modifications of proteins, and as such proteases might be considered as a class of signalling molecules in their own right (Overall, 2004). The repertoire of targets of a given protease is known as its substrate degradome (Lopez-Otin and Overall, 2002).

Targeted proteolytic processing is an essential feature of numerous processes, including: the blood clotting cascade (Davie and Ratnoff, 1964; Macfarlane, 1964);

activation of many proteases, such as trypsinogen (Davie and Neurath, 1955); ectodomain shedding, important in regulating factors including TGFβ and TNFα (Black et al., 1997; Moss et al., 1997; Peschon et al., 1998); chemokine processing during an inflammatory response (McQuibban et al., 2000; Parks et al., 2004); the release of cryptic epitopes. Cleavage of laminin by MMPs, for example, can reveal sites which enhance migration (Hintermann and Quaranta, 2004); generation of neoproteins, such as angioststin from plasminogen (Oreilly et al., 1996), or endoststin from collagen XVIII (Bergers et al., 1999)

Proteases are grouped into six families, based on the component of the active site responsible for nucleophilic attack of peptide bonds. The families are metallo-, serine-, cysteine-, aspartic-, threonine- and glutamic acid proteases.

1.2.8.1 Matrix metalloproteinases

The matrix metalloproteinases (MMPs) are zinc-dependent proteases capable of cleaving most components of the ECM, amongst other substrates. They were historically grouped into the collagenases, gelatinases, stromelysins and matrilysins on the basis of their specificity for various ECM proteins, although it is now recognised that there is a large overlap between their substrate specificities. The MMPs were originally thought to degrade the ECM and little else. As such they have been shown to promote cell migration and invasion, with important roles in cancer development and metastasis (Darmiento et al., 1995; Rudolph-Owen et al., 1998; Sternlicht et al., 1999; Sympson et al., 1995; Thomasset et al., 1998). Lack of MMP7, for example, was shown to inhibit intestinal tumorigenesis (Wilson et al., 1997). These findings generated great interest in the MMPs as targets for anti-cancer therapies. Unfortunately, these drugs were largely unsuccessful in clinical trials

(Overall and Lopez-Otin, 2002). Subsequent research into the functions of MMPs has shown them to regulate numerous signalling events (Egeblad and Werb, 2002; Overall, 2004), meaning that broad-spectrum inhibition of MMPs affects multiple processes, not just cancer cell invasiveness.

1.2.8.2 The urokinase plasminogen activator system

The urokinase plasminogen activator (uPA) consists of a system of tightly regulated serine proteases and inhibitors with roles in blood coagulation, wound healing and cancer, among others (Irigoyen et al., 1999). Cleavage of plasminogen by uPA produces active plamsin, which can occur either following the binding of the high molecular weight form of uPA to its receptor uPAR, or via the soluble low molecular weight form of uPA (Alfano et al., 2005; Kasai et al., 1985a; Kasai et al., 1985b), though the receptor bound form has a much higher activity (Ellis et al., 1989; Quax et al., 1991). There is a positive-feedback mechanism by which plasmin cleaves pro-uPA, although uPA activation can also occur via cathepsin B and G (Kobayashi et al., 1991), plasma kallikrein (Ichinose et al., 1986), mast cell tryptase (Stack and Johnson, 1994), nerve growth factor gamma (Wolf et al., 1993) and epithelial serine membrane protease matriptase (Lee et al., 2000). The main physiological inhibitors of uPA are PAI-1 (Serpin E1) (Carmeliet et al., 1993a; Carmeliet et al., 1993b), PAI-2 (Serpin B2) and glia-derived nexin (Serpin E2), whilst the main inhibitor of plasmin is alpha 2-antiplasmin (Serpin F2). As well as causing receptor mediated internalisation and degradation of uPA (Cubellis et al., 1990; Olson et al., 1992), PAI-1 is also involved in cell adhesion through interactions with vitronectin (Irigoyen et al., 1999).

Plasmin degrades fibrin during fibrinolysis of blood clots, as well as cleaving laminin basement membrane components fibrinogen and fibronectin (Goldfarb et al., 1986; Liotta et al., 1981a; Liotta et al., 1981b; Mochan and Keler, 1984). The plasminogen activator system can, however, influence a wide range of cellular processes and disease states, particularly cancer. Increased plasminogen activator activity is associated with cancer cell migration and invasion (Blasi, 1993; Pyke et al., 1991; Schmitt et al., 1991) and metastasis (Andreasen et al., 2000). The broad spectrum of functions for the uPA system is in part explained by a complex interplay with MMPs and growth factor signalling. The uPA system can activate TGFβ (Yee et al., 1993), whilst TGFβ can in turn induce PAI1 expression (Lund et al., 1987). Hepatocyte growth factor (HGF) can be activated by uPA cleavage (Mars et al., 1993; Naldini et al., 1992), and IGF-II bioavailability can be increased by cleavage IGFBP4 (RemacleBonnet et al., 1997). Matrix metalloproteinases, such as MMP2 and MMP9, are activated by plasmin (Baramova et al., 1997; Keskioja et al., 1992), and uPA (activates MMP2) (Kazes et al., 1998), while MMP2 can activate uPA in response to VEGF (Prager et al., 2004).

Together the various aspects of a microenvironment tightly regulate the cells which inhabit it and allow them to communicate appropriately, allowing tissues to maintain their homeostasis and to respond appropriately to injury and infection. As in all aspects of physiology, however, things can go wrong. Some examples have already been given of components of the microenvironment contributing to malignant disease. Where microenvironments are complex in normal physiological states, they are equally complex in cancer, and their involvement in these diseases provides researchers with both challenges and opportunities.

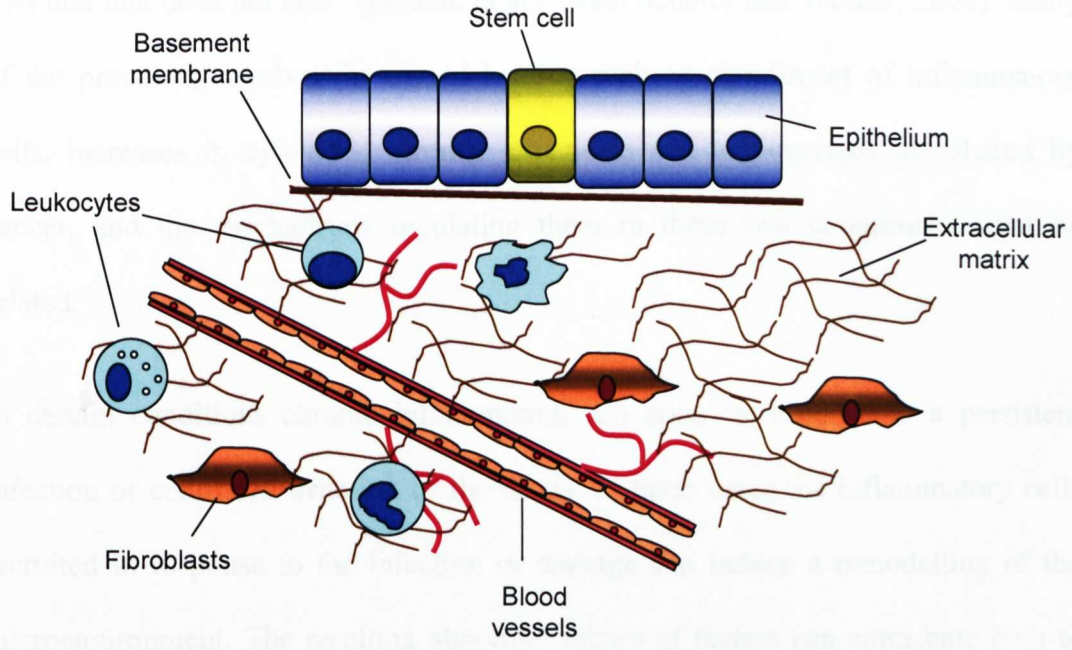


Figure 1.2

Key components of the microenvironment
Schematic shows some of the key components which define the cellular microenvironment, including epithelial cells, mesenchymal cells such as fibroblasts, infiltrating leukocytes of various types, vasculature and extracellular matrix.

1.3 Cancer microenvironments

1.3.1 Cancer-associated stroma

For many years the focus in cancer research was on those cells which accumulate mutations and other genetic lesions, causing a deregulation of cellular proliferation, cell death, genomic integrity and cell motility, ultimately leading to formation of tumours. Decades of research have identified many of the genetic lesions (Lengauer et al., 1998), and have provided a detailed understanding of the molecular processes involved in cancer development and progression (Vogelstein and Kinzler, 2004). Despite this progress there is still much about cancer that we do not yet fully understand. In recent years the emphasis in cancer research has shifted towards studying not only the cancer cells themselves, but the microenvironment which they inhabit. Much of this new thinking centres on an old hypothesis of cancer as a

"Wound that does not heal" (Dvorak et al., 1986; Schafer and Werner, 2008). Many of the processes involved in wound healing such as recruitment of inflammatory cells, increases in cytokines, stromal expansion and angiogenesis are shared by cancer, and the mechanisms regulating them in these two conditions might be related.

In certain conditions chronic inflammation can arise in response to a persistent infection or continued irritation of the tissue. In these cases the inflammatory cells recruited in response to the infection or damage can induce a remodelling of the microenvironment. The resulting abnormal milieu of factors can contribute both to pro-inflammatory and proliferative signals, which can ultimately initiate neoplastic progression (Coussens and Werb, 2002). Examples of cancers arising through this type of mechanism include gastric cancer (see later) due to *Helicobacter pylori* (*H. pylori*) infection, oesophageal adenocarcinoma due to gastroesophageal reflux (Lagergren et al., 1999) and lung cancer due to smoking, to name a few.

The stroma, together with the cancer cells themselves, defines the microenvironment of a cancer. The stroma comprises of extracellular matrix, fibroblasts/ myofibroblasts, vasculature/endothelial cells and immune cells. The role of stroma in organogenesis has long been recognised by developmental biologists (Cunha et al., 1983; Sakakura et al., 1976). The development and progression of many epithelial cancers is typically accompanied by stromal expansion (RonnovJessen et al., 1996; Tlsty and Hein, 2001). In addition to an expansion of the stroma the normal architecture of the tissue can be lost, suggesting aberrant cancer-stromal cell interactions. It is important, therefore, to know how the stroma associated with tumours differs from that found under normal physiological conditions.

1.3.2 Stroma as a therapeutic target

Progress in elucidating the roles of stroma in cancer development and progression have made it a rich source of potential targets for anti-cancer therapy (Albini and Sporn, 2007). One example is Avastin which targets vascular endothelial growth factor (VEGF) in order to reduce angiogenesis and thus restrict the blood flow to the tumour (Sandler et al., 2003). Other promising targets include members of the hedgehog (Hh) signalling pathway, which has been described as important in a number of cancers (Chari and McDonnell, 2007). More recently it was shown that Hh signalling is necessary for the tumour-promoting action of the stroma, and inhibition of genetic deletion inhibits growth in a mouse xenograft model (Yauch et al., 2008). Another class of proteins which have been identified as a target for anti-cancer therapies are the matrix metalloproteinases (MMPs) (see section 1.9.1). Inhibitors of the MMPs have been in development for over 20 years (Overall and Kleifeld, 2006), and were a promising prospect until phase III trails failed to show any improvement in patient survival. It is now clear that MMPs are not simply the invasion promoting proteases they were once thought to be, and that different members of the MMP family can have diverse and opposing roles in cancer. The consequences of the changes in tumour-stroma interactions which result from broad-spectrum MMP inhibition were not anticipated, and it is only now, armed with a better understanding of the stroma in cancer that the prospect of inhibiting MMPs in cancer is being revisited as a potential therapeutic strategy (Overall and Lopez-Otin, 2002). The stroma-targeted therapies developed to date, both the successes and failures, highlight the importance of this field of research.

1.3.3 Cancer stroma differs from normal stroma

In the last decade or so, cancer researchers have begun to define the changes that occur in the various stromal components during cancer initiation and progression. Angiogenesis, for instance, has long been known to be essential for tumour survival, and mechanisms by which cancer stroma promotes the growth of new vasculature, such as upregulation of vascular endothelial growth factor (VEGF), have now been identified (Carmeliet and Jain, 2000). The recruitment of immune cells to a tumour by cancer-derived chemokines can also help to create a microenvironment which allows cancer cells to develop (Coussens and Werb, 2002). For example, an increase in the levels of reactive oxygen species (ROS) generated by infiltrating leukocytes can not only cause DNA damage, thereby increasing the risk of mutations, but also provide selective pressure for cancer cells which are more stress-resistant (Blagosklonny, 2002; Marusyk and DeGregori, 2008). Similarly, changes in the interactions between cancer cells and ECM have been shown to have wide-ranging effects on the cancer cells, including their proliferation, survival and invasive potential (Ronnovjessen et al., 1995; Weaver et al., 1997). Some of these changes are orchestrated by the cancer cells themselves, which are involved in the production of the basement membrane, a specialised form of sub-epithelial ECM containing collagen IV, laminin, heparin-sulphate proteoglycans (Kalluri, 2003). Furthermore, changes in the composition of the ECM can influence paracrine signalling in the tumour, since ECM can act as a pool for growth factors (Schmidt and Kao, 2007). Fibroblasts, and in particular myofibroblasts, have, however, received the most attention of all of the stromal components in recent years for their pivotal role in defining the cancer microenvironment.

1.3.4 Myofibroblasts

The term myofibroblast describes a sub-population of fibroblasts which exhibit smooth muscle-like characteristics (Gabbiani, 1994). Small populations of myofibroblasts exist in normal tissues, where they secrete a range of growth factors, proteases and extracellular matrix (ECM) components, and as such are considered to be a key mesenchymal cell type important in maintaining the cellular microenvironment (Powell et al., 1999a; Powell et al., 1999b). Myofibroblasts may derive from a number of sources, including: cancer cells by epithelial-to-mesenchymal transition; transdifferentiation from fibroblasts; acute differentiation from fibroblasts e.g. by TGFβ stimulation; bone marrow-derived mesenchymal stem cells (Pittenger et al., 1999). Epithelial-to-mesenchymal transition (EMT) is a phenomenon in which epithelial cells acquire characteristics more usually associated with mesenchymal cells (Yang and Weinberg, 2008), and which can be induced by TGFβ signalling (Bierie and Moses, 2006; Zavadil and Bottinger, 2005).

As well as their role in maintaining the homeostasis of the cellular microenvironment, myofibroblasts play an important role in wound healing in a number of organs (Powell et al., 1999a). In this physiological process myofibroblasts are one of a number of cell types recruited to the healing wound by inflammatory signals where their contractile behaviour helps to close the wound. Through the production of various paracrine factors they also promote the migration of other cell types to the wound site as well as regulating the proliferating cells. It is these very characteristics of myofibroblasts which enable them to contribute to the development of neoplastic disease. In the setting of chronic inflammation the expansion of the myofibroblast population leads to an altered microenvironment

which promotes the migration, invasion and proliferation of myofibroblasts and other cell types, including cancer cells (Orimo et al., 2005).

1.3.5 Cancer-associated myofibroblasts

Myofibroblasts have been identified as an important cell type in many cancers by numerous studies (Bhowmick et al., 2004; Tlsty and Hein, 2001). These cells have been referred to using several terms, including carcinoma-associated fibroblasts, activated fibroblasts and myofibroblasts. Here they will be henceforth referred to as cancer-associated myofibroblasts (CAMs). These cells are well known to play an important role in several solid tumours, including prostate (Cunha et al., 2002), breast (Kuperwasser et al., 2004) and colorectal cancers (De Wever et al., 2004). In all of these cases they were characterised by increased motility and expression of α-smooth muscle actin (Vandenhooff, 1988). It is generally recognised that CAMs are more migratory and proliferative than fibroblasts from non-tumorigenic sites (Schor et al., 1988; Schor et al., 1985).

1.3.6 CAMs promote cancer development and progression

It has been shown that myofibroblasts isolated from various tumours maintain their altered phenotype in isolation from cancer cells (Olumi et al., 1999). These changes must therefore be stable alterations induced by some aspect of the disease, or these cancer associated myofibrobasts may be recruited from a different pool of cells such as circulating bone marrow derived cells (BMDCs). It is now recognised that changes arising in cancer-associated myofibroblasts (CAMs) can contribute to the progression of cancer by altering the microenvironment (Bhowmick et al., 2004; Tlsty and Hein, 2001). An example of this is breast carcinoma, where CAMs

promote tumour growth more than normal fibroblasts due to an increased secretion of stromal cell-derived factor 1 (SDF-1) (Orimo et al., 2005). Xenograft models of tumours have provided compelling evidence for the role of myofibroblasts in promoting tumour development (Camps et al., 1990; Gleave et al., 1991). Co-injection of CAMs with cancerous cells was shown to increase the rate of tumour development and growth. Sub-lethally irradiated fibroblasts have also been shown to be more potent inducers of tumour xenografts than normal human fibroblasts, in breast (Barcellos-Hoff, 1998), prostate (Cunha et al., 2002) and pancreas (Ohuchida et al., 2004). These irradiated fibroblasts resemble senescent aging fibroblasts (Krtolica et al., 2001; Parrinello et al., 2005), as well as CAMs. The risk of developing cancer is well known to increase with age, and the altered phenotype of aging fibroblasts might offer an additional mechanism to explain this. Normal epithelial cells, however, do not respond to CAMs in the same way as cancer cells, indicating that both the way in which CAMs define the microenvironment and the way cancer cells respond to the microenvironment are important (Cunha et al., 2002). Furthermore, it has been observed that whilst co-injection of human fibroblasts with cancer cells promotes the initial tumour xenograft, the human fibroblasts don't persist very long in the host (Hu et al., 2008; Yauch et al., 2008). While this might suggest that they are not necessary for subsequent tumour growth and survival, it is possible that the altered microenvironment which they helped to define results in the conversion of host fibroblasts to an activated myofibroblast phenotype.

A critical development in defining the roles of CAMs came when Bhowmick et al. (2004) demonstrated that targeting mutations of the *Tgfbr2* gene to fibroblasts in a murine model system caused spontaneous tumour formation in several organs,

proving that changes in a mesenchymal cell type can precede neoplastic changes and even induce them. Further work on this model system suggested that this might be the result of increased secretion of hepatocyte growth factor (HGF) from the mutated fibroblasts, acting through its receptor (encoded by the MET proto-oncogene) which is expressed by epithelial cells (Kuperwasser et al., 2004). An implication of this model is that changes in the molecular cross-talk between mesenchymal and epithelial cells may result in genetic or epigenetic changes within either cell type, and these changes can further contribute to the development and progression of cancer.

1.4 Gastric cancer

Gastric adenocarcinoma is the second leading cause of cancer-related death worldwide (Peek and Blaser, 2002), after lung cancer. It is well recognised that gastric adenocarcinoma can arise on the background of chronic inflammation of the stomach. The primary aetiological agent responsible for this chronic inflammation in humans in *Helicobacter pylori* infection. This bacterium was first identified as a causative agent in the development of gastritis and duodenal ulcer (Marshall and Warren, 1984), a discovery for which Marshall and Warren were awarded the Nobel Prize for Physiology or Medicine in 2005. Subsequently it became clear that there was also a relationship between *H.pylori* infection and gastric cancer arising from chronic gastritis (Forman et al., 1991; Nomura et al., 1991; Parsonnet et al., 1991). In the majority of instances, *H. pylori* infections remain asymptomatic. In some individuals, however, long term infection can lead to chronic inflammation of the gastric mucosa, which in turn significantly increases the risk of developing gastric cancer (Forman et al., 1991; Fox and Wang, 2007; Karnes et al., 1991; Parsonnet et

al., 1991). Approximately 95% of all gastric tumours are adenocarcinomas arising from the glandular epithelium. Most other gastric cancers are accounted for by gastrointestinal stromal tumours (GISTs) and tumours of mucosa-associated lymphoid tissues (MALT). The occurrence of MALT tumours is also related to *H.pylori* infection, since they arise from B cells involved in the immune response to the infection (Isaacson, 1994). The GISTs can occur anywhere in the GI tract, and consist of KIT (CD117) positive mesenchymal cells. Mutation in the c-kit gene are often found in GISTs, since their occurrence relies on activation of c-kit (Miettinen and Lasota, 2001).

A widely adopted histological classification system for gastric adenocarcinomas defines the two mains types; intestinal and diffuse (Lauren, 1965). Intestinal-type gastric carcinomas are more common in males and in old age (Sipponen and Marshall, 2000), and develop via a series of distinct histological steps (Correa, 2004; Fox and Wang, 2001; Peek and Blaser, 2002). These begin with chronic superficial gastritis, leading to atrophic gastritis. This is characterised by the loss of parietal cells or gastric glands, and also involves infiltration of the tissue with lymphocytes and macrophages (Dixon et al., 1996), and the development of fibrous stromal tissue. This is followed by intestinal metaplasia, in which the gastric glands are replaced with an intestinal-like epithelium containing mucin-filled goblet cells (Matsukura et al., 1980). Finally, the disease progresses to dysplasia and adenocarcinoma. Diffuse gastric carcinomas do not form gland-like structure as is the case in intestinal-type, but rather involve individual neoplastic cells infiltrating throughout the gastric epithelium (Houghton et al., 2001). Unlike intestinal-type adenocarcinomas, diffuse gastric carcinoma affects men and women equally, and is more likely to occur at a younger age.

Remodelling of the microenvironment is a critical step in the initiation and progression of gastric adenocarcinoma (Figure 1.3). In the stomach a small population of sub-epithelial myofibroblasts exist under normal physiological conditions. They are found in association with the isthmus region of the gastric gland, which is the presumed location of the gastric epithelial stem cell (McDonald et al., 2008). The close juxtaposition of myofibroblasts to the epithelium (Leedham et al., 2006) means that they are well placed for a role defining the microenvironment of the epithelial cells. They secrete HGF, TGFbeta and KGF (Modlin et al., 2003), the receptors for which are expressed on gastric epithelial cells (Powell et al., 1999b). Despite the gastric adenocarcinoma being epithelial in origin, it is estimated that 60-90% of the mass of gastrointestinal tumours consists of associated stromal tissue (Powell et al., 2005). This is not to be confused with the GISTs, which are wholly stromal in origin. The myofibroblast population has been shown to be expanded in *H.pylori* infection (McCaig et al., 2006) and gastric cancer (Jiang et al., 2008; Nakayama et al., 2000). This has been shown to result in part from a disregulation of epithelial-mesenchymal signalling. In *H.pylori* infection and cancer, there is an increase in the secretion of MMP7 by epithelial cells (McCaig et al., 2006). It was shown that MMP7 stimulates the proliferation and migration of gastric myofibroblasts (Hemers et al., 2005) via stimulation of the mitogen-activated protein kinase and phosphatidylinositol 3-kinase signaling pathways. Using a proteomics approach, insulin-like growth factor binding protein 5 (IGFBP5) was shown to be a target for MMP7, the cleavage of which released bioactive IGF-II, which could then act as an autocrine growth factor for gastric myofibroblasts, and a paracrine factor for the epithelium (Hemers et al., 2005; McCaig et al., 2006). These findings also highlighted the power of proteomic approaches in elucidating

signalling mechanisms. Gastric cancer-associatiated myofibroblasts (CAMs) were further characterised by comparing DNA methylation with myofibroblasts derived from adjacent, macroscopically normal tissue, revealing global hypomethylation of genomic DNA in the CAMs (Jiang et al., 2008). We do not yet know, however, whether the influence of CAMs and adjacent non-cancer tissue myofibroblasts (ANMs) on the microenvironment differs, or whether there are differences between the two cell populations which are functionally significant in gastric cancer.

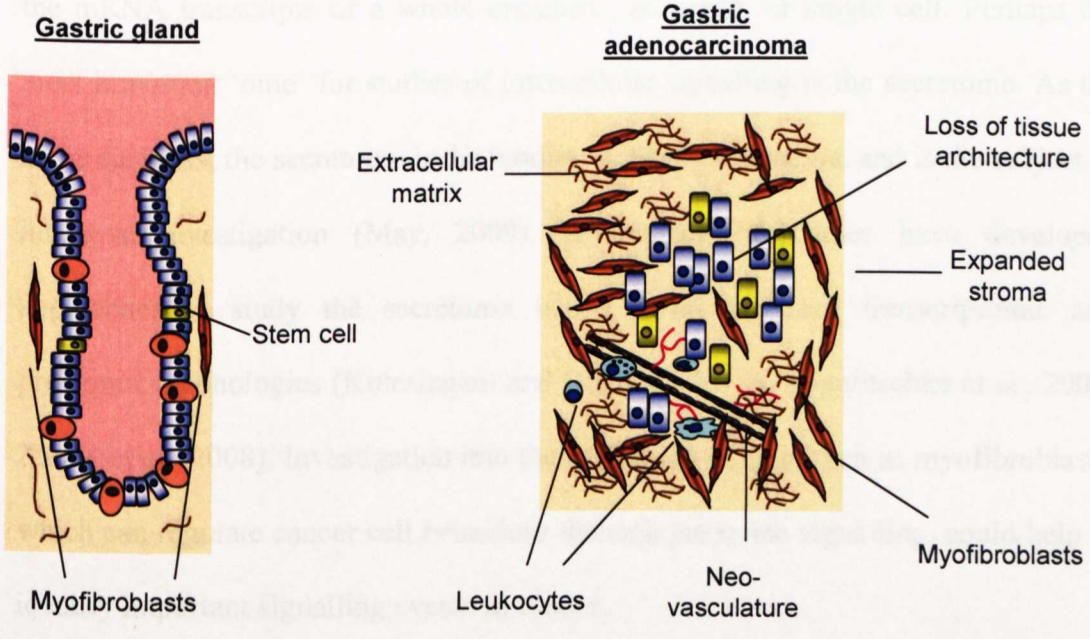


Figure 1.3

Microenvironment of gastric adenocarcinoma
Schematic illustrates the key components of the microenvironment in a healthy gastric gland and in a gastric adenocarcinoma. During the development of gastric adenocarcinoma there is significant remodelling of the microenvironment, including expansion of the myofibroblast population, loss of gland architecture and infiltration of inflammatory cells.

1.5 The '-omics' revolution

Modern technical advances have allowed a transformation in the way researchers tackle biological questions. The sequencing of the human genome (Lander et al., 2001; Venter et al., 2001), together with technical advances, means that it is now possible to study every human gene in a single experiment. The suffix 'ome' is used in the case of the genome to describe the complete set of chromosomes and DNA sequence of an organism. Over the past decade numerous new 'omes' have been described, such as the transcriptome, metabolome and secretome. Depending on the context in which these terms are used they can describe the totality of, for example, the mRNA transcripts of a whole organism, an organ, or single cell. Perhaps the most important 'ome' for studies of intercellular signalling is the secretome. As the name suggests, the secretome is the totality of secreted proteins, and is the subject of intensive investigation (May, 2009). A number of studies have developed approaches to study the secretome which have combined transcriptomic and proteomic technologies (Kulasingam and Diamandis, 2007; Paulitschke et al., 2009; Zhong et al., 2008). Investigation into the secretome of cells such as myofibroblasts, which can regulate cancer cell behaviour through paracrine signalling, could help to identify important signalling events in cancer.

1.6 Genomics and transcriptomics

1.6.1 Definition and use

The genome is defined as the full set of chromosomes of a cell, the DNA sequence of each chromosome, as well as the mitochondrial DNA. The transcriptome is the full set of mRNA transcripts produced by a cell, or a whole organism. Recent technical advances have allowed researchers to compare genomes and

transcriptomes to identify the molecular basis of diseases such as cancer. Global genomic analyses have successfully identified the most frequently occurring mutations in cancers of the breast, colon (Wood et al., 2007), pancreas (Jones et al., 2008) and brain (Parsons et al., 2008). These studies also employed a systems biology approach in order to gain more information from the genomic experiments by identifying the key signalling pathways affected by the mutations identified. Transcriptomic studies can potentially yield even more information, as they can measure the relative abundance of every known gene or exon transcript. Of particular interest in this thesis is the use of transcriptomics to investigate stromal gene expression in cancer. It was recently shown that stromal gene expression of breast cancer can predict clinical outcome (Finak et al., 2008), highlighting both the importance of the stroma in defining cancer, and the power of transcriptomic approaches. It is not known, however, how the transcriptomes of myofibroblast populations associated with a tumour and the adjacent non-cancer tissue compare. Analysis of the transcriptomes of these cells could therefore provide important insight into the roles of this important stromal cell type.

1.6.2 Gene expression arrays

Gene expression arrays allow researchers to measure the relative abundance of mRNAs across entire transcriptomes (Schulze and Downward, 2001). They consist of thousands of oligonucleotide probes immobilised on a solid surface, such as glass, plastic or silicon. The samples to be analysed are labelled, often using fluorescent tags, and allowed to hybridise to the array. The intensity of fluorescence is used as a measurement of mRNA abundance. This technology has given rise to an entire field of research dedicated to applying gene expression arrays to biological questions, and

developing the statistical processes, software and equipment necessary to produce valid data (Breitling, 2006). Cancer research in particular has benefited from the development of gene expression arrays, which have allowed researchers to define new disease classes (Alizadeh et al., 2000; Golub et al., 1999; van't Veer et al., 2002).

With a large selection of gene arrays now available, covering entire transcriptomes or focussing on specific subsets of genes, gene array technology offers great flexibility when designing experiments. For example, by combining gene expression profiling with signalling pathway analysis software a population of putative breast cancer stem cells were identified with increased TGFβ pathway activity. (Shipitsin et al., 2007). Others have used gene expression profiling together with whole genome mutational analyses to identify key mutations in cancers and how they affect gene expression and important functional networks (Jones et al., 2008; Parsons et al., 2008). Integrating gene array data in this way with other approaches helps to overcome some of the difficulties associated with interpreting gene expression data (Breitling, 2006). The availability of gene array datasets in repositories such as Oncomine has also been important in refining the interpretations of these data by allowing cross-comparisons and meta-analyses of related studies (Rhodes et al., 2007).

A number of studies have also used gene expression profiling to generate correlations with clinical outcomes, and in doing so identify subsets of genes which can be used as prognostic predictors (Sorlie et al., 2001; van't Veer et al., 2002; van de Vijver et al., 2002). It has now been shown in breast carcinoma that not only cancer cell transcriptomes, but also stromal gene expression can be correlated with

clinical outcome (Finak et al., 2008). While this is an interesting finding, it is not clear whether it is due to changes in a single cell type, or a change in the composition of the stroma. It is also not known whether this is true in other cancers, such as gastric adenocarcinoma. Nevertheless, studies of this sort could lead to gene expression profiling becoming an important clinical tool, with the potential for individualised medicine (Sotiriou and Piccart, 2007).

1.7 Proteomics

1.7.1 Definition and use

Proteomics is the study of proteomes, in a manner analogous to the study of genomes in the field of genomics. This distinguishes it from the protein chemistry of the past century which often focussed on the properties of individual proteins. A proteome is a complete set of proteins, including all modifications and structural variants, produced by a particular organism, system or cell. As such a proteome is not a static entity, but one which can vary greatly between different situations (Aebersold and Mann, 2003). One of the great advantages of proteomics over genomics is the capacity to study the proteome of specific subcellular compartments, thus giving information not only about which proteins are expressed and their abundance, but also their spatial distribution within a cell (Taylor et al., 2003). Proteomic analysis of the secretome, for example, is potentially an important tool for biomarker and drug target discovery (Gronborg et al., 2004), and for defining intercellular signalling events (May, 2009; Zhong et al., 2008).

1.7.2 Mass spectrometry

The primary tool of proteomics research is the mass spectrometer (MS) (Mann et al., 2001). Mass spectrometers comprise three principal components; an ion source, a mass analyser which sorts ions by electrical and/or magnetic fields, and a detector (see Appendix I). There are many variations on each of these three components, and there are a large number of possible combinations which can comprise a MS system. In proteomics experiments, MS is used to measure the molecular mass and charge of peptide fragments, and can also provide sequencing of peptides. Proteins are often cleaved with an endopeptidase such as trypsin prior to MS analysis in order to generate peptides of a suitable size. Trypsin cleaves peptide bonds on the C-terminal side of polar amino acid residues arginine and lysine, giving the advantage that every peptide is polar, making ionization easier. Various databases and bioinformatics tools have been developed to allow proteins to be identified and characterised based on MS spectra (Aebersold and Mann, 2003; Gasteiger et al., 2003; Mann and Jensen, 2003). Advances in proteomics are now such that a single experiment was able to identify 40,582 unique peptides, corresponding to 7,093 proteins (Wisniewski et al., 2009). Whilst development in MS technology have been important in achieving this, perhaps the most significant factor in improving coverage in proteomics experiments is sample processing, particularly separation of proteins or peptides.

1.7.3 Liquid chromatography

High performance liquid chromatography (HPLC) separates compounds as they pass through a column packed with material referred to as the stationary phase. The mobile phase which is pumped through the system contains the sample to be

separated, which can either be analysed directly as it elutes from the system, or collected in a series of discrete fractions. The retention time of molecules will vary depending on their interaction with the stationary phase, as determined by some aspect of their chemistry. For example, reverse phase chromatography separates a moderately polar mobile phase based on adsorption to a non-polar stationary phase (Horvath et al., 1976). Ion exchange chromatography utilises attractive forces between charged sites on the stationary phase and ions in the mobile phase (Small et al., 1975). Several other systems exist using similar principles to these, with each separating the compounds in terms of a specific feature. In proteomics the separation of protein and peptide samples can greatly improve proteome coverage and the accuracy of protein identifications by fractionating complex samples (Aebersold and Mann, 2003). In this way each fraction generated by HPLC can be run as a separate MS sample, which can reduce the masking of peptides by very abundant peptides, and reduce the amount of spectral overlap.

1.7.4 Quantitative proteomics

The term quantitative proteomics describes a range of methods which allow the determination of either absolute quantities of proteins with samples, or relative abundances of proteins when different samples are compared.

Much of the early work on proteomics utilised electrophoretic separation of proteins from complex samples, an approach which is still widely used (Gorg et al., 2004). These experiments commonly use polyacrylamide gels in order to separate proteins based on their apparent molecular weight, though this separation is also influenced by protein conformation. As such, sodium dodecyl sulphate (SDS) or other detergents can be included in the gels to denature the proteins. Additional protein

separation can be performed prior to polyacrylamide gel electrophoresis (PAGE) by using isoelectric focussing (IEF) to separate proteins based on their isoelectric point by applying electrophoresis along an immobilised pH gradient (Bjellqvist et al., 1982). This method is known as two-dimensional PAGE (2D-PAGE) (Ofarrell, 1975).

More recent advances in quantitative proteomics have utilised stable mass isotopes to tag proteins from different samples, allowing them to be distinguishable by MS (Gevaert et al., 2008). These methods do not require protein separation by gel electrophoresis, and are therefore referred to as 'gel-free' quantitative proteomics methods. When used with complex samples, such methods are usually combined with LC fractionation of the sample prior to MS. One of the first of these methods exploited the incorporation of two atoms of oxygen during hydrolysis of peptide bonds by trypsin (Mirgorodskaya et al., 2000). By performing the trypsinisation of one of a pair of samples in water containing the ^{18}O isotope of oxygen a mass difference of 4Da is introduced into each tryptic peptide (Staes et al., 2004). Since then more advanced techniques such as isotope coded affinity tags (ICAT), stable isotopic labelling by amino acids in cell culture (SILAC) and isobaric tags for relative and absolute quantitation (iTRAQ) have emerged and quantitative proteomics has become a widely used technique.

Stable isotopic labelling by amino acids in cell culture (SILAC) (Ong et al., 2002) uses isotopically labelled amino acids such as Lys and Arg which are incorporated into proteins during cell culture. These labelled amino acids contain one or more stable isotopes of carbon, nitrogen, hydrogen or oxygen, with the various possible combinations enabling the generation of a range of mass differences between

peptides from different samples. Whilst it is possible to use a number of different SILAC amino acids, Lys and Arg are a good choice for most experiments in which the sample is to be trypsinised, since this results in each peptide bearing a single isotopic amino acid at its C-terminal end. When the peptides are analysed by MS, this isotopic difference is apparent during the first MS stage at which the peptide masses are determined. Relative abundances can then be calculated as a ratio of the area under the two sets of isotopic peaks.

Isotope coded affinity tags (ICAT) (Gygi et al., 1999) uses chemical probes containing a reactive group for targeting a specific amino acid side chain, such as cysteine residues by incorporating iodoacetamide in the probe. To allow for relative quantitation of samples, the probes also incorporate stable mass isotopes, and finally a group such as biotin to allow for affinity capture of labelled peptides. Unlike SILAC, the proteins are tagged after the sample has been collected.

Isobaric tags for relative and absolute quantitation (iTRAQ) uses an approach similar to that of ICAT, labelling the peptides after they have been extracted and trypsinised (Wiese, 2007). The tags used in iTRAQ are termed isobaric, because they in fact all have the same mass. This means that when the peptides are analysed by MS, equivalent peptides from each experiment appear as a single peak in the first round of MS. It is only when a peptide ion is selected and fragmented for fragmentation and MS/MS sequencing that relative quantitation is possible. Each different iTRAQ label is designed to fragment at a different point during collision, resulting in reporter ions of different molecular weights. It is these reporter ions which are used to calculate the relative abundance of the peptide in different samples. The iTRAQ method has the advantage of potentially higher sensitivity,

since the peptides from individual samples are effectively summed during the first round of MS, increasing the chance of a low abundance peptide passing the sensitivity threshold. Users of the iTRAQ method must, however, ensure that their mass spectrometer is capable of MS/MS of a high and reproducible quality to ensure accuracy of the quantitation.

1.7.5 Advanced proteomics - COFRADIC

Combined fractional diagonal chromatography (COFRADIC) is a method for enriching a sample for a desired subset of peptides prior to MS analysis in a proteomic experiment (Vandekerckhove and Gevaert, 2005). The method utilises two identical reverse phase liquid chromatography (RP-LC) steps, separated by a peptide modification reaction. Depending upon the application, this step is tailored to alter the column retention time of either the peptides of interest or the peptides which are to be removed from the sample. In either situation the result is the separation of the desired peptides into elution fractions which are distinct from the majority of other peptides in the sample. This approach has now been successfully applied to enrich for a range of peptides and modifications, including methionine-containing peptides (Gevaert et al., 2002), cysteine-containing peptides (Gevaert et al., 2004), phosphorylated peptides (Gevaert et al., 2005), N-glycosylation (Ghesquiere et al., 2006), sialyation (Ghesquiere et al., 2007) and N-terminal peptides (Gevaert et al., 2003; Impens et al., 2008; Van Damme et al., 2005). All of these applications have opened up new possibilities for addressing specific biological questions using proteomics. Of particular interest here, however, is the enrichment for N-terminal peptides. This provides the maximum possible simplification of a protein sample, since each protein is normally represented by a

single peptide. More importantly, it can be used to identify sites of proteolytic cleavage, since each cleavage will generate a new N-terminal peptide. These peptides will be enriched for together with the other N-terminal peptides, and can be distinguished based on their position in the peptide sequence of the protein (see Appendix II). The role of proteases in cancer is well recognised, and the study of proteases and their substrates on a global scale has been highlighted as a promising new approach in cancer research (Doucet et al., 2008).

COFRADIC protocols have now been successfully employed to address a variety of biological questions, including identification of the targets of caspases which mediate Fas-induced cell death (Van Damme et al., 2005), and definition of the platelet proteome (Martens et al., 2005). Their potential is a good example of how advances in proteomic techniques have been an important force for driving innovation in biomedical research in recent years.

1.8 Systems biology, bioinformatics and pathway analysis

Systems biology aims to address important biological questions using an integrative approach which combines mathematical models, bioinformatics and data mining tools. This approach is intimately related with the fields of genomics and proteomics, which have provided a wealth of information in recent years. By pooling such data, a systems biologist can define signalling pathways and links between pathways. It is increasingly clear that it is pathways, not individual genes, which determine the course of oncogenesis (Vogelstein and Kinzler, 2004). As such, there has been a recent flurry of activity in developing software tools which are capable of pathway-centric analysis of '-omics' datasets, with MetaCore (GeneGo Inc.) being a

notable example (http://www.genego.com/metacore.php). MetaCore is unique in that every interaction within its database has been manually curated from peer-reviewed publications. Whilst MetaCore does not offer as much detailed control over the analyses performed as other software packages, such as Ingenuity, the interface is well suited to biologists who are not necessarily bioinformatics specialists. The validity of a pathway-centric approach is demonstrated in a number of recent high-profile publications which have made tremendous progress in defining the key signalling pathways responsible for the development of breast (Shipitsin et al., 2007), colorectal (Wood et al., 2007) and pancreatic cancers (Jones et al., 2008).

Systems biology and the '-omics' are developing rapidly, and it is essential for researchers to stay abreast of these evolving fields in order to maximise the information that they can offer. This will allow them to continue to be applied to new biological questions, such as elucidating the role of myofibroblasts in gastric cancer.

1.9 Aims and objectives

1.9.1 To characterise gastric cancer associated myofibroblasts

Previous studies have highlighted the importance of myofibroblasts in gastric cancer (Hemers et al., 2005; McCaig et al., 2006) and revealed differences in genomic DNA methylation between CAMs and ANMs (Jiang et al., 2008). It is not known, however, whether there are any functional differences between CAMs and ANMs, and so the initial aims of this study were to:

1. Confirm the phenotype of primary human myofibroblasts derived from gastric carcinoma and adjacent non-cancer tissue.

2. Compare the migratory ability of cancer-associated myofibroblasts (CAMs) and the adjacent non-cancer tissue myofibroblast (ANM) counterparts.
3. Compare the transcriptomes of CAMs and ANMs using gene expression arrays and pathway analysis bioinformatics tools.

1.9.2 To determine the effects of myofibroblast secreted factors on gastric cancer cells

Cancer-associated myofibroblast secreted factors have been shown to play important roles in prostate (Cunha et al., 2002), breast (Kuperwasser et al., 2004) and colorectal cancers (De Wever et al., 2004), and to differ from those secreted by normal fibroblasts (Orimo et al., 2005; Tlsty and Hein, 2001). It is not known, however, whether this is the case in gastric cancer, and so the aims of this section were to:

1. Compare the effects of CAM and ANM media on gastric cancer cell phenotypes.
2. Compare the effects of CAM and ANM media on gastric cancer cell transcriptomes using gene expression arrays and pathway analysis bioinformatics tools.

1.9.3 To compare the secretomes of CAMs and ANMs

While CAMs have been shown to be an important cell type in several cancers due to changes in secreted factors (Bhowmick et al., 2004; Olumi et al., 1999; Orimo et al., 2005), a comparison of the secretomes of CAMs and ANMs has not been reported. The aims of this section were therefore to:

1. Define the differences between CAM and ANM secretomes using quantitative proteomics approaches.
2. Identify signalling pathways likely to be affected by any differences identified in the secretomes using pathway analysis bioinformatics tools.
3. Investigate proteolytic processing in CAM and ANM secretomes using N-terminal COFRADIC.

1.9.4 Validation of hits

The identification of differentially abundant proteins by proteomics techniques does not necessarily mean that the proteins are functionally important in this context. The aims of this section were therefore to:

1. Use Western blotting to validate proteomics data for protein of interest.
2. Determine the functional significance of the protein of interest in defining CAM and ANM secretomes, and their effect on gastric cancer cells.

Chapter 2

Materials and Methods

2.1 Materials

All chemicals purchased from Sigma unless otherwise stated.

AGS gastric cancer cell line obtained from the American Type Culture Collection (VA, USA)

Amicon Ultra-15 3kDa centrifugal filter devices (Millipore, Watford, UK)

Antibiotic-antimycotic solution (Sigma, Poole, Dorset, UK)

BD Control Cell Culture Inserts and BD BioCoat Matrigel Invasion Chambers (BD Biosciences, Massachusetts, USA)

Bovine serum albumin (BSA; Jackson Immuno Research Laboratories, Suffolk, UK)

Chamber slides (BD Flacon, Belgium)

Coomassie brilliant blue G-250 (Merck, Hertfordshire, UK)

DALT Buffer Kit (GE Healthcare Bio-Sciences AB, Uppsala, Sweden)

DALT Gel 12.5 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden)

DC microplate protein assay (BioRad, Hertfordshire, UK)

DiffQuick (Dade Behring Inc. Newark, DE, USA)

Donkey serum (Jackson Immuno Research Laboratories, Suffolk, UK)

Dulbecco's Modified Eagle's Medium (DMEM, Sigma, Poole, Dorset, UK)

Fetal bovine serum (FBS; Perbio, Cheshire, UK)

GeneChip©Human Genome U133 Plus 2.0 arrays (Affymetrix, Santa Clara, CA, USA)

ICAT® cartridge cation exchange (Applied Biosystems, Foster City, CA, USA)

Immobiline™ DryStrips pH 3-10NL, 18cm (GE Healthcare Bio-Sciences AB, Uppsala, Sweden)

IPG buffer pH 3-10 (GE Healthcare Bio-Sciences AB, Uppsala, Sweden)

IPG cover fluid (GE Healthcare Bio-Sciences AB, Uppsala, Sweden)

iTRAQ® 4-plex reagents kit (Applied Biosystems, Foster City, CA, USA)

Monoclonal antibody to Glyceraldehyde-2-PDH (GAPDH; Biodesign International, Saco, ME, USA)

Monoclonal antibody against human uPA B-chain (American Diagnostica, Inc., CT, USA)

Non-essential amino acid solution (Sigma, Poole, Dorset, UK)

PageRuler™ Plus Prestained Protein Ladder (Fermentas, York, UK)

Penicillin-streptomycin solution (Sigma, Poole, Dorset, UK)

Plasmin purified from human plasma and certified lysine- and epsilon-aminocaproic acid (EACA)-free (Calbiochem, Beeston, UK)

Phosphatase Inhibitor Cocktail set II, EDTA-Free (Calbiochem, USA)

Protease Inhibitor Cocktail Set III, EDTA-Free (Calbiochem, USA)

Recombinant human insulin-like growth factor 2 (IGF-II) (Calbiochem, Beeston, UK).

Recombinant human βig-h3 protein (R&D Systems Inc., Oxfordshire, UK)

RNeasy kit (Qiagen, West Sussex, UK)

Sequencing-grade modified trypsin (Promega, WI, USA)

SILAC™ Protein Identification and Quantitation Kit (Invitrogen, Paisley, Renfrew, UK)

Synthetic uPA inhibitor uPA-STOP™ (American Diagnostica, CT, USA)

[^{3}H]-thymidine (Amersham, Little Chalfont, Bucks, UK)

TRITC-conjugated Phalloidin (Sigma, Poole, Dorset, UK)

Trypsin solution (Sigma, Poole, Dorset, UK)

Ultrapure α-cyano-4-hydroxycinnamic acid (CHCA; Proteabio, WV, USA)

Urokinase plasminogen activator (uPA) prepared from human urine (Millipore, Watford, UK)

Vectashield containing 4',6-diamidino-2-phenylindole DAPI (Vector laboratories, Peterborough, UK)

Water (H_2O) used in all experiments was Millipore double filtered water.

2.2 Derivation of human primary myofibroblasts

Human primary myofibroblasts were derived from resected gastric cancer and adjacent normal gastric tissue obtained during surgery to remove tumours at Szeged Hospital, Hungary. Myofibroblasts were prepared in collaboration with Dr Peter Hegyi, First Department of Medicine, University of Szeged, Hungary, as published previously (McCaig et al., 2006). Myofibroblasts were transported to Liverpool in a liquid nitrogen tank and kept in liquid nitrogen until use. For each of the 12 sets of patient samples, cancer associated myofibroblasts (CAMs) and adjacent non-cancer tissue-derived myofibroblasts (ANMs) were cultured. The details of the gastric cancers are shown in table 2.1, including the histological classification of the gastric cancer using the Lauren system (Lauren, 1965), and the dynamic scores which indicate the severity of the disease (higher numbers indicating poorer prognosis). This work was approved by the Ethics Committees of the University of Szeged, Hungary.

Patient Number	Patient	Gastric Cancer	Lauren Classification	Dynamic score	Gender	Year born	Myofibroblasts Generated
1	Sz42	Distal	Medullar (non-Lauren)	5	M	1934	CAM; ANM
2	Sz45	Distal	Intestinal	11	M	1923	CAM; ANM
3	Sz190	Distal	Mixed	11	F	1941	CAM; ANM
4	Sz192	Distal	Diffuse	12	F	1957	CAM; ANM
5	Sz194	Distal	Intestinal	4	M	1931	CAM; ANM
6	Sz195	Distal	Diffuse	9	F	1922	CAM; ANM
7	Sz198	Distal	Intestinal	7	M	1930	CAM; ANM
8	Sz268	Distal	Intestinal	11	M	1930	CAM; ANM
9	Sz271	Distal	Mixed	12	M	1935	CAM; ANM
10	Sz294	Distal	Intestinal	7	F	1923	CAM; ANM
11	Sz305	Distal	Diffuse	13	F	1948	CAM; ANM
12	Sz308	Distal	Mixed	13	M	1956	CAM; ANM

Key:
CAM – Cancer
ANM – Adjacent non-cancer tissue

Table 2.1

Details of gastric cancers used to generate primary human myofibroblast cultures
Cancer and adjacent normal gastric tissue was resected from 12 patients and used to derive primary cultures of gastric myofibroblasts. The table indicates gastric cancer classification, dynamic scores and the myofibroblast cultures generated.

2.3 Tissue culture

2.3.1 Human primary myofibroblast culture

Myofibroblasts were maintained in DMEM supplemented with 10% v/v FBS, 2% v/v antibiotic-antimycotic, 1% v/v penicillin-streptomycin solution and 1% v/v non-essential amino acids, unless otherwise stated. Myofibroblasts were cultured at 37°C in an atmosphere of 5% v/v CO_2 with media changes every 48-60 hours and passaged at confluence by incubation with 0.25% w/v trypsin solution for 10 minutes. Myofibroblasts were used after a minimum of 4 passages and up to a maximum of 12 passages. For experiments using myofibroblast conditioned media (CM), $1.1x10^6$ myofibroblasts were plated in 10cm diameter dishes to give 80-90% confluency. The myofibroblasts were then cultured for 24 hours before being washed three times with phosphate buffered saline (PBS) and cultured for a further 24 hours in media minus FBS. This serum-free media was then collected and either used immediately or stored at -80°C.

2.3.2 Human gastric carcinoma cell line culture

AGS gastric carcinoma cell line was cultured in DMEM supplemented with 10% v/v FBS and 1% v/v penicillin-streptomycin, as published earlier (Varro et al., 2002).

2.4 Immunocytochemistry

Cells were cultured in chamber slides and fixed using 4% w/v paraformalydehyde (PFA) for 30 minutes at room temperature. Fixed cells were washed three times with PBS and permeabilised by incubating with PBS containing 0.1% v/v triton X-100 for 30 minutes at room temperature. The cells were then washed a further three times with PBS and incubated with 5% w/v BSA for 30 minutes at room temperature. Cells were washed twice with PBS and incubated with 10% v/v donkey serum for one hour at room temperature to prevent the non-specific binding of the secondary antibodies. The serum was removed and cells washed twice in PBS before primary antibody was added, diluted in PBS, and incubated overnight at 4°C in a humidified atmosphere. Details of the primary antibodies used are shown in table 2.2. The cells were then washed sequentially with 0.14M NaCl, 0.5M NaCl and 0.14M NaCl. The appropriate secondary antibody, as indicated in table 2.2, was diluted in 10mM Hepes and added to the cells for one hour at room temperature in the dark. The secondary antibodies were then removed, the cells washed three times with PBS (10 minutes), and mounted with Vectashield containing DAPI. For visualisation of F-actin, the primary antibody was substituted with 50μg/mL TRITC-conjugated phalloidin in tris-buffered saline (TBS), and no secondary antibody was used. When using phalloidin, PBS was substituted with TBS at all stages. Fixed cells were incubated with phalloidin for 40 minutes at room temperature before washing with TBS. Slides were viewed using a Zeiss Axioplan-2 microscope (Zeiss Vision, Welwyn Garden City, UK). Images were captured using a JVC-3 charge-coupled device camera at 40X magnification (air objective) with KS300 software (Imaging Associates, Bicester, Oxfordshire, UK).

Primary Antibody	Dilution	Supplier	**Secondary antibody**	Dilution	Supplier
Anti-α smooth muscle actin (Mouse)	1:100	Fitzgerald, MA, USA	**FITC-conjugated donkey-anti-mouse**	1:400	Jackson, PA, USA
Anti-vimentin (Guinea pig)	1:400	Fitzgerald, MA, USA	**Texas Red-conjugated donkey-anti-guinea pig**	1:400	Jackson, PA, USA
Anti-desmin (Rabbit)	1:100	Fitzgerald, MA, USA	**FITC-conjugated donkey anti-rabbit**	1:400	Jackson, PA, USA
Anti-pan cytokeratin (Rabbit)	1:100	Dako, Denmark	**FITC-conjugated donkey-anti-rabbit**	1:400	Jackson, PA, USA

Table 2.2

Antibodies and working dilutions for immunocytochemistry
Table shows the primary and corresponding secondary antibodies used for immunocytochemistry of primary gastric myofibroblasts and AGS gastric carcinoma cell line, together with the working titres and suppliers.

2.5 Cellular functional assays

2.5.1 Cell migration and invasion assays

Cell migration and invasion was assayed using BD Control Cell Culture Inserts and BD BioCoat Matrigel Invasion Chambers (BD Biosciences, Massachusetts, USA) respectively following manufacturer's instructions. Both assays consisted of a Boyden chamber-type insert with a porous membrane, onto which cells were plated, which sat in multi-well plate. The two assay types differed in that the invasion assays had a coating of Matrigel™ on the porous membrane, such that a proteolytic or invasive phenotype would be required for cells to pass through the membrane. The cell types, seeding densities and media used in the experiments are detailed in the relevant results chapters. Plates were incubated for 18 hours before migrating or invading cells were stained using DiffQuick (Dade Behring Inc. Newark, DE). The

cells on the lower side of the membrane were counted at 10x magnification. Five fields of view were counted for each membrane, and the mean value used.

2.5.2 [^{3}H]-thymidine incorporation proliferation assay

Synthesis of DNA was measured as an indicator of cell proliferation by incorporation of [^{3}H]-thymidine as reported previously (Varro et al., 2002). AGS cells were cultured in six-well plates at a seeding density of $5x10^4$/well in 2mL of DMEM supplemented with 10% v/v FBS and 1% v/v penicillin/streptomycin solution. The cells were grown for 24 hours before being synchronised in G0/G1 by incubation with serum-free DMEM for 48 hours. The media was then changed for myofibroblast media and incubated for 18 hours, with 2μCi [^{3}H]-thymidine included for the last 2 hours. Cells were then washed three times with cold PBS, then incubated with 2mL 5% w/v trichloroacetic acid (TCA) for 20 minutes at 4°C. Cells were then washed twice with 2mL cold ethanol before DNA was solubilised by incubation in 1mL of 0.1M NaOH for 60 minutes at 60°C. An aliquot of solubilised DNA was then taken for scintillation counting, with each experiment performed in triplicate.

2.6 Gene expression arrays

Myofibroblasts and AGS cells were cultured as described in sections 2.3.1. and 2.3.2 respectively, and RNA extracted using the RNeasy kit (Qiagen), performed by Dr. Islay Steele as per the manufacturers instructions. Samples were analysed using GeneChip©Human Genome U133 Plus 2.0 arrays (Affymetrix, Santa Clara, CA, USA), at the Liverpool Genome Facility by Dr. Lucille Rainbow, as per manufacturers instructions:

http://www.affymetrix.com/products_services/arrays/specific/hgu133plus.affx#1_2. The U133 Plus 2.0 arrays measure the relative expression levels of over 47,000 transcripts. A GeneChip® Scanner 3000 (Affymetrix) was used to image the arrays. Quality control was performed using Microarray Suite 5.0 QC Metrics (Affymetrix) to ensure that each array experiment was performed correctly and that the data from different experiments could therefore be compared. Gene array data was analysed using GeneSpring GX v.10 (Agilent). Experiments were normalised using the MAS5 algorithm and filtered using the present (P), marginal (M) and absent (A) calls generated by MAS5 to indicate the likelihood that a given gene is expressed, such that in each comparison performed the gene must be flagged as P in at least 1 of the experiments. Statistical analysis was performed using paired t-tests and ANOVAs within GeneSpring as indicated

2.7 Proteomics

2.7.1 Protein sample preparation

Myofibroblast media, prepared as described in section 2.3.1, was concentrated to approximately 0.5mL using Amicon Ultra-15 3kDa centrifugal filter devices (Millipore). Immediately after media collection, cells were washed 3 times with cold PBS and scraped in 100µL of RIPA cell lysis buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 1% w/v NP-40, 1% w/v sodium deoxycholate and 0.1% w/v sodium dodecyl sulfate), 1% v/v Protease Inhibitor Cocktail Set III, EDTA-Free (Calbiochem, USA) and 1% v/v Phosphatase Inhibitor Cocktail set II, EDTA-Free (Calbiochem, USA). The cell suspension was then sonicated in a water bath for 5 minutes to promote cell lysis and incubated on ice for 30 minutes. The suspension was then centrifuged for 3 minutes at 10,000 x G, 4°C, and the supernatant collected.

Media and cellular protein samples were assayed for protein concentration using the BioRad DC microplate protein assay (BioRad). The assay was calibrated using BSA standards prepared in either DMEM or RIPA buffer for media or cell samples respectively.

2.7.2 Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)

Concentrated myofibroblast media samples were precipitated by addition of an equal volume of cold 20% w/v trichloracetic acid (TCA) and incubated on ice for 2 hours. Precipitated protein samples were resuspended in cell lysis buffer. Protein samples were added to 9M urea containing 2% w/v CHAPS, 0.01% w/v bromophenol blue with freshly added 2% v/v immobilised pH gradient (IPG) buffer and 0.028% w/v dithiothreitol (DTT) and added to 180mm pH 3-10 nonlinear IPG strips which were then overlaid with IPG Cover Fluid and incubated for 18 hours at room temperature to rehydrate. Isoelectric focussing (IEF) was performed using Immobiline™ DryStrips pH 3-10NL, 18cm (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) following manufacturer's instructions. Following the IEF run, strips were stored at -80°C. Prior to running samples on second dimension polyacrylamide gels, IPG strips were equilibrated in solutions of 50mM Tris, 6M urea, 30% v/v glycerol, 2% w/v SDS, 0.01% bromophenol blue. The equilibration solution 1 also contained freshly added 1% w/v DTT and solution 2 contained 2.5% w/v iodoacetamide (IAA). Strips were incubated for 15 minutes in solution 1 followed by 15 minutes in solution 2; second dimension was then run using the Ettan DALT Gel 12.5 and DALT Buffer Kit (GE Healthcare Bio-Sciences AB, Uppsala, Sweden) at 2.5 W per gel for 30 minutes then at 20W per gel until the dye front reached the bottom of the gel. Following electrophoresis, proteins were fixed in a solution of 7% v/v glacial acetic

acid, 40% v/v methanol for 1 hour. Gels were then stained in a solution of 4 parts Coomassie stain (0.1% w/v Coomassie brilliant blue G-250 (Merck) in 2% w/v phosphoric acid, 10% w/v ammonium sulphate) and 1 part methanol with agitation for 1-2 hours. Gels were then destained with 10% v/v acetic acid, 25% v/v methanol for 60 seconds, rinsed with 25% v/v methanol and destained further in 25% v/v methanol for 24 hours. Protein spots from 2D gels that were selected for identification were excised from the gel into 25% v/v methanol.

2.7.3 MALDI-TOF identification of proteins

The 25% v/v methanol was removed from the excised gel fragments and 100µL 50mM NH_4HCO_3, 50% v/v acetonitrile (ACN) was added to each sample. Gel fragments were incubated for 15 minutes at room temperature with occasional vortexing to remove Coomassie stain. The supernatants were then discarded and the fragments dried by vacuum centrifugation. Sequencing-grade modified trypsin was added (2.5µg) and incubated at 37°C for 18 hours. To each sample 30µL of 60% v/v ACN, 1% w/v trifluoroacetic acid (TFA) were then added and samples sonicated in a water bath for 5 minutes. The supernatants were collected, and a further 30µL of 60% v/v ACN, 1% w/v TFA were added, sonicated for 5 minutes and the supernatants pooled with the previously collected ones. The supernatants were then dried in a vacuum with centrifugation. Samples were resuspended in a solution of ultrapure CHCA in 50% v/v ACN, 0.1% w/v TFA. Samples were spotted onto a MALDI loading plate and allowed to dry before analysis using a Voyager-DE PRO (Applied Biosystems) instrument. Database searching of peptide mass fingerprints was performed using the online version of Mascot (Matrix Science Inc, MA, USA).

2.7.4 Stable isotope labelling by amino acids in cell culture (SILAC)

SILAC labelling was performed using the SILAC™ Protein Identification and Quantitation Kit (Invitrogen, Paisley, Renfrew, UK). Lysine and arginine containing stable mass isotopes of their C and N atoms were used to label proteins from different samples in tissue culture. The lysine isotopes used were $^{12}C_6$, $^{14}N_2$ l-lysine (light isotope) and $^{13}C_6$, $^{14}N_2$ l-lysine (heavy isotope), which have a total mass difference of 6 Da. The arginine isotopes used were $^{12}C_6$, $^{14}N_4$ l-lysine (light isotope) and $^{13}C_6$, $^{15}N_4$ l-lysine (heavy isotope), which have a total mass difference of 10 Da. Myofibroblasts were labelled incubation with DMEM containing SILAC amino acids, as indicated in section 5.2.2, supplemented with 10% v/v dialysed FBS, 2% v/v antibiotic/antimycotic and 1% penicillin/streptomycin for at least 6 population doublings, which for primary gastric myofibroblasts requires approximately 28 days. Protein samples were prepared as described in section 2.7.1. Concentrated media samples were combined, precipitated by addition of 20% w/v TCA, resuspended in 50mM ammonium bicarbonate, and incubated with 5mg/mL trypsin for 18 hours at 37°C. The samples were then analysed by LC-MS/MS. The cell extract samples were electrophoresed using 12% polyacrylamide gels, which were stained using Coomassie stain as in section 2.7.2. Visible bands were excised from the gel, and the proteins were digested in gel as described in section 2.7.3. Once the supernatants containing the digested proteins were collected and dried, they were resuspended in 50mM ammonium bicarbonate and analysed by LC-MS/MS.

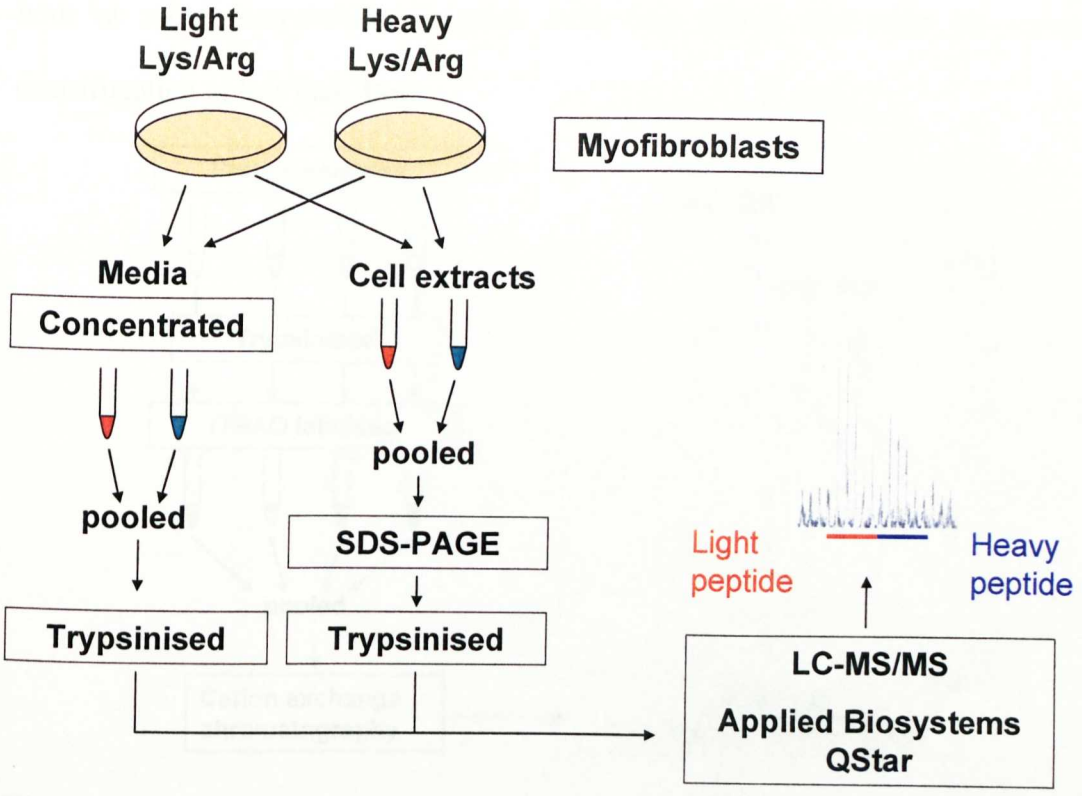


Figure 2.1

Stable isotope labelling by amino acids in cell culture (SILAC) schematic Workflow for the processing of myofibroblast media and cell extract samples following SILAC labelling.

2.7.5 Isobaric tagging for relative and absolute quantitation (iTRAQ)

Myofibroblast concentrated media and cell extract samples were prepared as described above, and 80µg of each protein sample was precipitated in cold acetone for 2 hours at -20°C, then centrifuged at 14,000rpm at 4°C for 10 minutes. The protein pellets were resuspended in 20µL of 0.5M triethylammonium bicarbonate (TEAB, pH8.5), reduced with 2µL 50mM tris-(2-carboxyethyl)phosphine (TCEP) at 60°C for one hour and alkylated with 1µL 200mM methyl methanethiosulfonate (MMTS) for 10 minutes at room temperature. Samples were then digested using 8µg sequencing-grade modified trypsin overnight at 37°C. The iTRAQ tag reagents were prepared by adding 70µL ethanol, then added to the samples and incubated for one

hour at room temperature. Samples were then mixed and dried by vacuum centrifugation to less than 30µL.

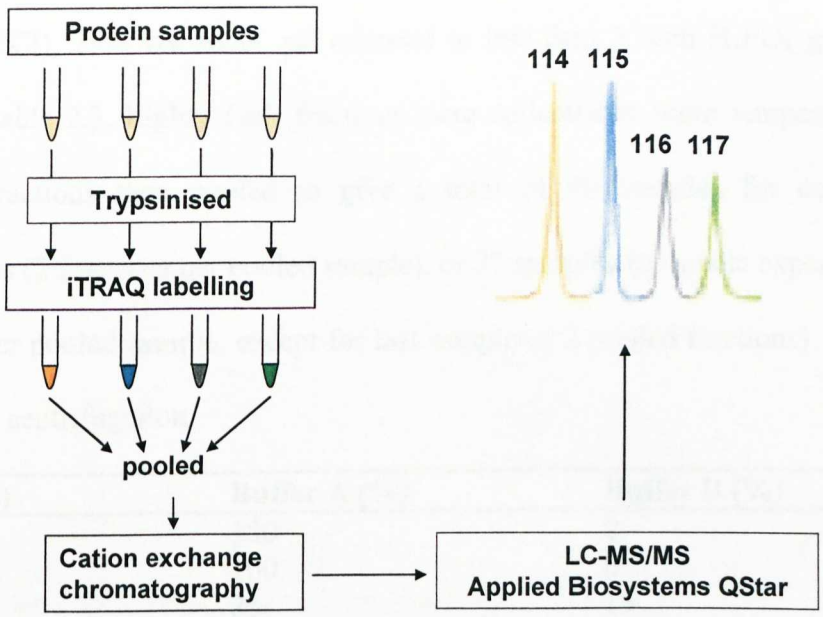


Figure 2.2

Isobaric tagging for relative and absolute quantitation (iTRAQ) schematic
Schematic describing the workflow for iTRAQ labelling of protein samples.

2.7.6 Cation-exchange cartridge

Samples labelled by iTRAQ require strong cation exchange chromatography prior to mass spectrometry analysis in order to remove various substances using in the labelling which interfere with subsequent steps. Strong cation exchange (SCX) cartridge processing was performed using ICAT® cartridge cation exchange (Applied Biosystems) as per the manufacturer's instructions.

2.7.7 Cation-exchange chromatography

Protein samples labelled by iTRAQ were diluted to 2mL with 25% v/v ACN and pH adjusted to less than 3 with H_3PO_4. The sample was then loaded onto a PolyLC PolySULFOETHYL A (4.6 x 200mm i.d.) cation exchange (CEX) column using an Agilent 1100 HPLC system at a flow rate of 1 mL/min. Buffer A comprised 10mM

potassium phosphate (KH_2PO_4), 25% v/v ACN, pH adjusted to less than 3 with H_3PO_4, buffer B comprised 10mM potassium phosphate (KH_2PO_4), 1M potassium chloride (KCl), 25% v/v ACN, pH adjusted to less than 3 with H_3PO_4, gradient as shown in table 2.3. Eighty 1mL fractions were collected at room temperature, and adjacent fractions then pooled to give a total of 40 samples for cell extract experiments (2 fractions per pooled sample), or 27 samples for media experiments (3 fractions per pooled sample, except for last sample of 2 pooled fractions) and dried by vacuum centrifugation.

Time (min)	Buffer A (%)	Buffer B (%)
0:00	100	0
15:00	100	0
60:00	85	15
75:00	50	50
90:00	50	50
90:05	100	0
95:00	100	0

Table 2.3

Cation exchange HPLC flow gradient
The flow gradient of buffer A (10mM KH_2PO_4, 25% v/v ACN, pH less than 3) and buffer B (1M KCl, 10mM KH_2PO_4, 25% v/v ACN, pH less than 3) during strong cation exchange chromatography of iTRAQ labelled protein samples.

2.7.8 LC-MS/MS

Samples were resuspended in 180µL 5% v/v ACN, 0.05% w/v TFA and loaded onto a PepMap C_{18} (75µm x 180 mm i.d.) reverse phase column with C_{18} Trap (300µm x 5 mm i.d.) using an LC Packings Ultimate nano-LC system run in-line with an Applied Biosystems QStar Pulsar mass spectrometer via a nano-electrospray source head and 10µm inner diameter PicoTip (New Objective, Massachusetts, USA). Following loading of samples the trap column was washed for 30 minutes with 5% v/v ACN, 0.05% w/v TFA to desalt. The system was run at a flow rate of 350nL/min using gradient as shown in table 2.4 where buffer A was 5% v/v ACN, 0.05% w/v

TFA, and buffer B was 95% v/v ACN, 0.05% w/v TFA. Spectra from both MS and MS/MS were acquired using an information-dependant acquisition (IDA) in positive ion mode.

Time (min)	Buffer A (%)	Buffer B (%)
0:00	100	0
30:00	100	0
100:00	40	60
110:00	20	80
120:00	1	99
121:00	100	0
140:00	100	0

Table 2.4

Reverse phase LC-MS/MS HPLC flow gradient
The flow gradient of buffer A (5% v/v ACN, 0.05% w/v TFA) and buffer B (95% v/v ACN, 0.05% w/v TFA) during reverse phase HPLC of protein samples in-line with a QStar Pulsar mass spectrometer.

2.7.9 Proteomic data handling

Protein identification and quantification were performed using the ProteinPilot™ v2.0.1 software (Applied Biosystems; MDS-Sciex). The Paragon algorithm was selected as the default search program, with the digestion agent set as trypsin and cysteine modification as either iodoacetic acid or methyl methanethiosulfonate for SILAC and iTRAQ experiments respectively. Spectra were searched against the June 2008 Uniprot database. Only proteins with an identification confidence of at least 95% were reported. Differentially abundant proteins were defined as having a fold change of 1.3 or more.

2.8 Combined fractional diagonal chromatography (COFRADIC)

Combined fractional diagonal chromatography (COFRADIC) (Gevaert et al., 2002) is a technique which uses two rounds of HPLC, separated by a peptide modification

step, to enrich for chosen classes of peptides such as, in this case, methionine-containing peptides or N-terminal peptides. Methionyl COFRADIC was performed together with Dr. Bart Ghésquiere (University of Ghent), N-terminal COFRADIC was performed together with Ms Veerle Vandenbussche (University of Ghent). Mass spectrometry of COFRADIC samples was performed using a LTQ Orbitrap LC-MS/MS system (Thermo) and database searching performed using an in-house version of Mascot (Matrix Science) by Mr Pieter-Jan De Bock and Ms Evy Timmerman (University of Ghent). Myofibroblasts were labeled using SILAC as described previously, and media collected and concentrated as before. Protein samples were precipitated with 20% w/v cold trichloroacetic acid (TCA), which were incubated on ice for 2 hours and centrifuged at 14,000 rpm at 4°C for 10 minutes. The precipitated protein was resuspended in 50 mM HEPES, pH 7.4, 100 mM NaCl, 0.8% w/v CHAPS, 1% v/v Protease Inhibitor Cocktail Set III, EDTA-Free (Calbiochem, USA).

2.8.1 Methionyl COFRADIC

Guanidinium HCl was added to each sample to a final concentration of 0.8mM, and 200µL 1M Tris pH 8.7 added to buffer the samples. Iodoacetamide (IAA) was added to a final concentration of 16mM, and tris(2-carboxyethyl)phosphine (TCEP) was added to a final concentration of 8mM. Samples were vortexed to mix and incubated at 37°C for 30 minutes. Samples were then desalted using NAP-5 columns (GE Healthcare) in a solution of 50mM triethylammonium buffer (Fluka), and eluted in 500µL of 50mM triethylammonium buffer. Equal amounts of the samples were then pooled and 5µg sequencing-grade modified trypsin (Promega, Madison, WI, USA) added and samples were incubated at 37°C for 18 hours. The digested samples were

acidified by addition of 50μL 10% w/v acetic acid. The samples were then dried to 100μL by vacuum centrifugation. The pH of the samples was measured as it was important to ensure that the sample pH is the same during both HPLC runs so that the column retention times will not be influenced by pH. The samples were run using an Agilent 1100 HPLC system with a Zorbax 300SB-C_{18} column (Agilent) at a flow rate of 80μL/min, and 60 1 minute fractions were collected between minutes 20 and 80 of the run. Solvent A was 10mM ammonium acetate, 2% v/v ACN, pH5.5, and solvent B was 10mM ammonium acetate, 70% v/v ACN, pH 5.5. Following injection of the samples onto the column, a 10 minute isocratic run with 100% solvent A was applied at a constant flow rate of 80μL/min, followed by a linear binary gradient from 100% solvent A to 100% solvent B over 100 minutes, a 10 minute isocratic wash with 100% solvent B and finally a 5 minute linear gradient back to 100% solvent A. Collected fractions were pooled as shown in table 2.5 to give 15 samples, which were then dried by vacuum centrifugation. Each sample was resuspended in 90μL 0.5% w/v acetic acid, 5μL of which was analysed by LC-MS/MS. The samples were then run on the reverse-phase HPLC under the same conditions as the primary run, with the exception of a pre-incubation step with 3% v/v hydrogen peroxide at 30°C for 30 minutes in order to oxidise methionine residues. Blank runs were performed after every 5 samples in order to clear the column. Secondary fractions were collected according to the strategy shown in table 2.6, with 24 secondary fractions collected during each of the 15 runs. These secondary fractions were pooled such that, within each of the 15 sets of 24 secondary fractions produced, every sixth fraction was pooled to give 6 pooled samples per set of 24 secondary fractions, and dried by vacuum centrifugation. The samples were then resuspended in 20μL 2% v/v ACN and analysed by LC-MS/MS.

Pooled sample	Primary interval (min)			Primary interval (min)			Primary interval (min)			Primary interval (min)		
1	20	-	21	35	-	36	50	-	51	65	-	66
2	21	-	22	36	-	37	51	-	52	66	-	67
3	22	-	23	37	-	38	52	-	53	67	-	68
4	23	-	24	38	-	39	53	-	54	68	-	69
5	24	-	25	39	-	40	54	-	55	69	-	70
6	25	-	26	40	-	41	55	-	56	70	-	71
7	26	-	27	41	-	42	56	-	57	71	-	72
8	27	-	28	42	-	43	57	-	58	72	-	73
9	28	-	29	43	-	44	58	-	59	73	-	74
10	29	-	30	44	-	45	59	-	60	74	-	75
11	30	-	31	45	-	46	60	-	61	75	-	76
12	31	-	32	46	-	47	61	-	62	76	-	77
13	32	-	33	47	-	48	62	-	63	77	-	78
14	33	-	34	48	-	49	63	-	64	78	-	79
15	34	-	35	49	-	50	64	-	65	79	-	80

Table 2.5

Fraction collection and pooling strategy for primary Methionyl-COFRADIC HPLC run
Fractions were collected at one minute intervals between 20 and 80 minutes, as shown in the columns of the table. These 60 fractions were pooled into 15 samples, such that the fractions shown in each of the 15 rows of the table were pooled together. In this way the number of samples is reduced whilst maintaining the peptide separation generated by reverse-phase HPLC, since the pooled fractions in the 15 samples each elute 15 minutes apart from each other.

Sample	Fraction interval (minutes)							
	Primary	Secondary	Primary	Secondary	Primary	Secondary	Primary	Secondary
1	20-21	8-18	35-36	23-33	50-51	38-48	65-66	53-63
2	21-22	9-19	36-37	24-34	51-52	39-49	66-67	54-64
3	22-23	10-20	37-38	25-35	52-53	40-50	67-68	55-65
4	23-24	11-21	38-39	26-36	53-54	41-51	68-69	56-66
5	24-25	12-22	39-40	27-37	54-55	42-52	69-70	57-67
6	25-26	13-23	40-41	28-38	55-56	43-53	70-71	58-68
7	26-27	14-24	41-42	29-39	56-57	44-54	71-72	59-69
8	27-28	15-25	42-43	30-40	57-58	45-55	72-73	60-70
9	28-29	16-26	43-44	31-41	58-59	46-56	73-74	61-71
10	29-30	17-27	44-45	32-42	59-60	47-57	74-75	62-72
11	30-31	18-28	45-46	33-43	60-61	48-58	75-76	63-73
12	31-32	19-29	46-47	34-44	61-62	49-59	76-77	64-74
13	32-33	20-30	47-48	35-45	62-63	50-60	77-78	65-75
14	33-34	21-31	48-49	36-46	63-64	51-61	78-79	66-76
15	34-35	22-32	49-50	37-47	64-65	52-62	79-80	67-77

Table 2.6

Fraction collection strategy for secondary Methionyl-COFRADIC HPLC run
The 15 samples produced by the pooling of primary fractions were each run as separate reverse-phase secondary HPLC runs. The rows in the table show the primary fractions were used to make up each of these 15 samples. To the right of each of the primary fractions are the intervals at which corresponding secondary fractions were collected following methionine oxidation. Within these 10 minute secondary intervals, 6 fractions were collected such that the first 2 fractions were each collected over a 3 minute interval, and the remaining fractions were collected over a 1 minute interval. Thus each of the 15 samples produced a total of 24 secondary fractions.

2.8.2 N-terminal COFRADIC

Equal amounts of protein were combined and guanidinium hydrochloride added to a final concentration of 4M. Iodoacetamide (IAA) was added to a final concentration of 10mM, and tris(2-carboxyethyl)phosphine (TCEP) was added to a final concentration of 5mM. Samples were vortexed to mix and incubated at 37°C for 30 minutes. The samples were then desalted using a NAP-5 column (GE Healthcare), which was first equilibrated with three washes of 1.4M guanidinium HCl in 50mM sodium phosphate (pH8) buffer. The samples were eluted in 850µL of the same buffer. Trideutero-N-hydroxysuccinimide (D3-NHS) was added to the samples in a ratio of 150g D3-NHS to 1g protein, and the samples incubated for 120 minutes at 30°C. Hydroxylamine was added in a ratio of 4mol hydroxylamine to 1mol D3-NHS, in order to revert the partial acetylation of hydroxyl groups, and incubated for 15 minutes at 30°C. Glycine was then added in a ratio of 2mol glycine to 1mol D3-NHS in order to quench the D3-NHS reaction. The samples were then desalted again, this time using a NAP-10 column (GE Healthcare) and eluted in 1.3mL of 10mM NH_4HCO_3. The samples were incubated at 95°C for 5 minutes, then transferred to ice for 5 minutes, before adding sequencing-grade modified trypsin (Promega, Madison, WI, USA) in a ratio of 1g trypsin to 100g protein and incubated for 18 hours at 37°C. The samples were then dried by vacuum centrifugation. The samples were redissolved in 300 µl solvent A for strong cation exchange (SCX) (50 mM sodium phosphate, 50% v/v ACN, pH3) and further acidified to pH3 with careful addition of 50% v/v ACN containing 0.1%, 0.5% or 1% w/v TFA as required, such that the final TFA concentration is less than 0.08% w/v. Strong cation exchange (SCX) was performed on an Accu Bond cartridge (Agilent). The cartridge was rinsed with 5 ml of ddH_2O and then pre-equilibrated with 10 mL of solvent A.

Samples (1mL) were loaded onto the column and washed with solvent A (5mL) and NaCl in solvent A (5mM, 2mL, total 8 ml). Eluates were brought to pH 7 by addition of 5M NaOH and dried by vacuum centrifugation. The samples were resuspended in 300µL of 50mM sodium phosphate (pH 3) and the pH brought to pH 3 by addition of TFA in 50% v/v ACN as before, again ensuring to keep the final concentration of TFA below 0.08% w/v. The samples were made up to a final volume of 1mL with solvent A. The samples were again run through an Accu Bond (Agilent) SCX cartridge, which was first equilibrated with 5mL ddH_2O and 10mL solvent A. The samples were eluted in 5mL solvent A, followed by 2mL 5mM NaCl in solvent A. The samples were then adjusted to a pH between 6-10 using NaOH in order to facilitate enzyme activity in subsequent steps, then dried by vacuum centrifugation. A pyro-glutamate removal step was then performed utilising the pGAPAse from the TagZyme® kit (Qiagen) which was purified by Ni^{2+}-immobilised metal ion chromatography (IMAC). This resulted in the pGAPase being dissolved in 25µl 20mM ammonium bicarbonate, 500mM immidazole, and 25µL of this solution was dissolved in 1.25µl 800mM NaCl, 1µl 50mM EDTA and 11µl 50mM fresh cysteamine. The pGAPase was activated by incubation for 10 minutes at 37°C (a total volume of 38.25 µl). The dried samples were redissolved in 27µl of 100mM sodium phosphate and the activated pGAPase was added as well as 25µl Q-cyclase (Qiagen). The whole mixture (total volume 125µl) was incubated for 60 minutes at 37°C. The samples were acidified with 0.5% acetic acid to pH5, then centrifuged at 16,000g to remove insoluble peptides. The samples were then run on an Agilent 1100 HPLC system with a Zorbax 300SB-C_{18} column (Agilent) under the same conditions as used in the Methionyl-COFRADIC method, with the methionine oxidation step included in this case in the primary run. In this case only 15 primary

fractions were collected at equal intervals between 20 and 100 minutes. The fractions were dried by vacuum centrifugation and resuspended in 20μL 2% w/v pyridine in H_2O in order to help to remove ammonium from the sample. The samples were again dried and then resuspended in 50μL of 50mM sodium borate, pH 9.5, and 15nmol of 2,4,6-trinitrobenzenesulphonic acid (TNBS) added to each. The TNBS was used to produce highly hydrophobic conjugates to free α-amine groups, which will be present at the N-terminus of all peptides except for those which have been blocked by acetlyation, i.e. the N-terminal peptides. The samples were incubated at 37°C for 30 minutes, and the addition of TNBS and incubation repeated a further three times. The samples were acidified by addition of 3μL 100% w/v acetic acid, then centrifuged for 10 minutes at 13,000 rpm to remove insoluble material. A second reverse-phase HPLC run was then performed under the exact same conditions as the primary run, with each of the 15 primary fractions run separately and each producing 15 secondary fractions. Fractions were dried and resuspended in 20μL 2% v/v ACN. Those fractions which could be seen to contain peptides which had not shifted following the TNBS reaction were analysed by LC-MS/MS, since these fractions were enriched for N-terminal peptides.

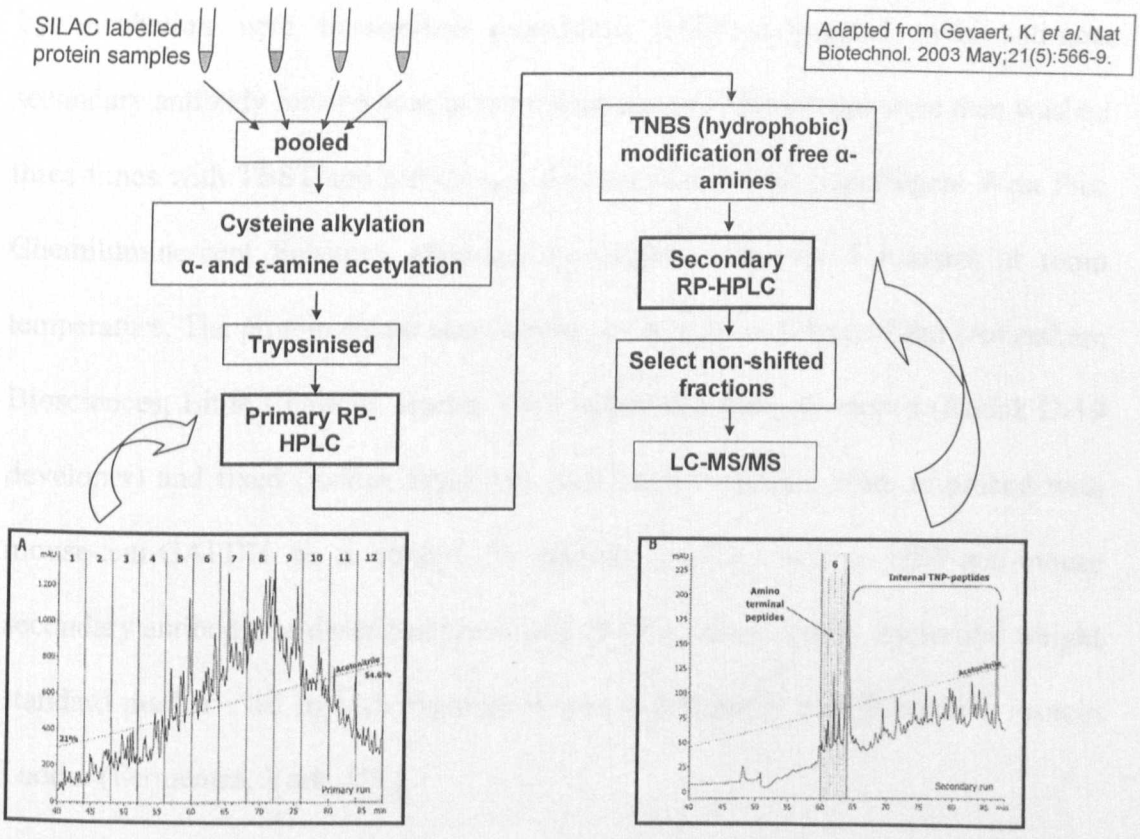


Figure 2.3

N-terminal COFRADIC schematic
The schematic depicts the strategy for enriching for N-terminal peptides in a SILAC labeled proteomic experiment using N-terminal COFRADIC.

2.9 Western blotting

Protein samples, prepared as described in section 2.7.1, were electrophoresed and resolved on 10% w/v SDS-polyacrylamide gels and electrotransferred onto nitrocellulose membranes (Amersham Pharmacia Biotech, Buckinghamshire, UK). Following transfer of proteins, membranes were washed three times with tris buffered saline (TBS) containing 0.1% v/v Tween (TBST), then incubated in blocking buffer (TBST containing 5% w/v Marvel) for one hour at room temperature. Membranes were probed with goat-anti-βig-h3 primary antibody (R&D systems Inc., Oxfordshire, UK) diluted in blocking buffer for 18 hours at 4°C with shaking. Membranes were then washed a further three times with TBST, followed

by incubation with horseradish peroxidase (HRP)-conjugated rabbit-anti-goat secondary antibody for one hour at room temperature. Membranes were then washed three times with TBST and membranes then incubated with SuperSignal West Pico Chemiluminescent Substrate (Pierce; Cramlington, UK) for 5 minutes at room temperature. The protein signal was detected by exposure of HyperFilm (Amersham Biosciences; Little Chalfont, Bucks, UK), which was then developed (Kodak D-19 developer) and fixed (Kodak rapid fix). Cell extract samples were re-probed with mouse-anti-GAPDH as a control for protein loading, with a HRP-anti-mouse secondary antibody, as described previously (Hemers et al., 2005). Molecular weight standard marker used in each experiment was PageRuler™ Plus Prestained Protein Ladder (Fermentas, York, UK).

2.10 Statistics

Results are expressed as mean ± standard error of the mean (SEM), unless otherwise stated. Student's *t*-test was used to determine statistical significance of results and considered significant (*) at $p<0.05$, unless otherwise stated. *N* = number of independent samples. Where appropriate, correlation of data were measured using the coefficient of determination (R^2).

2.11 Software

GeneSpring GX (Agilent)

Mascot (Matrix Science)

MetaCore (GeneGo Inc.). Analyses performed May 2009.

Protein Pilot 2.0.1 (Applied Biosystems)

SigmaPlot 11.0 (Systat software Inc.)

Chapter 3

The functional characterisation and transcriptome analysis of gastric cancer-associated myofibroblasts

3.1 Introduction

Myofibroblasts can be an important cell type in cancer (Bhowmick et al., 2004), and it has recently been identified that gastric myofibroblast populations are expanded by *H. pylori* infection (McCaig et al., 2006) and in gastric adenocarcinoma (Jiang et al., 2008). While previous work on cancer associated myofibroblast (CAMs) has focussed largely on cancers of the breast (Orimo et al., 2005) and colon (De Wever et al., 2004), very little is known about myofibroblasts from gastric cancer. This study makes use of the unique availability in the group of cultures of myofibroblasts derived from human gastric carcinoma samples. Collaborative work with these cells revealed global hypomethylation of the DNA of CAMs compared with ANMs (Jiang et al., 2008). These findings firmly established that differences exist between gastric CAMs and ANMs. It is now important to investigate the functional differences between these two populations of myofibroblasts and identify the genes and proteins involved. This will allow us to better understand how myofibroblasts might be involved in redefining the microenvironment in gastric cancer and the mechanisms involved.

Gene expression arrays are a useful tool for comparing the transcriptomes of different populations of cells in order to identify genes which are differentially expressed. Gene expression profiling was shown to be able to distinguish between diffuse and intestinal type gastric adenocarcinomas, as well as tissue from intestinal metaplasia and chronic gastritis (Boussioutas et al., 2003). Their study used resected tissue from which RNA was isolated, and so it is likely that this included both the cancer or epithelial cells and stromal cells such as myofibroblasts. Others have used gene expression arrays to investigate the roles of myofibroblasts in various diseases. For example, a comparison of primary human myofibroblasts derived from patients

with Idiopathic Pulmonary Fibrosis (IPF) and from control tissue using Affy U133 plus 2.0 gene expression arrays identified TGFbeta-driven EMT as a source of the IPF myofibroblasts (Larsson et al., 2008). Similarly, a comparison of hepatic stellate cells and myofibroblast-like cells from livers with and without cirrohsis identified different differences in NF-κB signalling and expression of interleukins. (Estep et al., 2009)

While stromal gene expression has been shown to correlate with clinical outcome in breast cancer (Finak et al., 2008), surprisingly little is known about changes in the gene expression of myofibroblasts themselves. Since the role of myofibroblasts in gastric cancer is poorly defined, a comparison of cancer-associated and adjacent normal myofibroblasts could provide useful insights.

3.1.1 Aims

The aims of the work described in this chapter were to functionally characterise gastric cancer-associated myofibroblasts (CAMs) and adjacent non-cancer tissue-derived myofibroblasts (ANMs), and define changes in their transcriptomes. Specifically, the objectives were to;

1. Establish the myofibroblast identity and purity of the tissue-derived cultures.
2. Compare the motility of CAMs and ANMs *in vitro*.
3. Use gene expression arrays to identify differences in the transcriptomes of CAMs and ANMs

3.2 Materials and Methods

3.2.1 Primary human myofibroblasts

Primary myofibroblasts were derived from resected gastric carcinoma and adjacent normal tissue from 12 patients as described in section 2.2.

3.2.2 Immunocytochemistry

Myofibroblasts were grown in 4-well chamber slides at a seeding density of $2x10^4$cells/well for 24 hours prior to fixation and staining, as described in section 2.3. For each marker, 1000 cells per condition were manually inspected and counted.

3.2.3 Migration assays

Migration assays were performed using BD Control Cell Culture Inserts (BD Biosciences, Massachusetts, USA) as described in section 2.4.1. Myofibroblasts were split by incubating with 0.25% trypsin solution (Sigma) for 10 minutes, pelleted by centrifugation at 800G, washed with serum-free DMEM (SFM) and resuspeded in SFM. Myofibroblasts were seeded at a density of $1x10^5$ per well, with each condition duplicated. The lower well of the assay was filled with either SFM, or SFM containing 100ng/mL recombinant human (rh) IGF-II (R&D Systems, MN, USA). Migration data are presented as mean number of cells per field +/- standard error of the mean (S.E.M.), and significance determined by two-tailed paired t-test $p<0.05$.

3.2.4 Gene expression arrays

Gene expression arrays were performed using GeneChip©Human Genome U133 Plus 2.0 arrays (Affymetrix, Santa Clara, CA, USA) as described in section 2.5.

Myofibroblasts were cultured as described in section 3.2.1 until confluent, at which point RNA was extracted for transcriptome analysis as detailed in section 2.5. Data was analysed using GeneSpring GX (Agilent). Gene expression data from individual pairs of myofibroblasts were normalised and baseline corrected to the median of all samples using the MAS5 algorithm. This also flagged each data point generated by the array as present (P), marginal (M) or absent (A) to identify data from transcripts for which the likelihood of their being genuinely expressed is consider reliable, marginal and unreliable respectively. Gene expression data were analysed in pairs of experiments using the matched CAM and ANM experiments, and the data were filtered such that, for each individual pair of myofibroblasts, expression data were removed unless flagged as P in at least 1 of the 2 gene arrays from the pair.

3.2.5 Analysis of gene array data

MetaCore (GeneGo) was used to analyse gene expression data in the context of signalling pathways. MetaCore utilises databases of manually curated protein interactions and signalling pathways. MetaCore was used here to compile data and to identify signalling pathways which were significantly represented by myofibroblast gene expression data.

3.3 Results

3.3.1 Gastric myofibroblasts express α-SMA and vimentin, but not cytokeratin or desmin

Immunocytochemistry analysis of CAMs and ANMs revealed strong expression of both α-SMA and vimentin in all cells (Figure 3.1 A - H), with cytoskeletal fibres of both clearly visible. Both CAMs and ANMs had no detectable cytokeratin or desmin expression (Figure 3.1 I and K respectively),

consistent with a myofibroblast phenotype. The majority of myofibroblasts had a spindle-like morphology, with no obvious differences in size or morphology between the CAMs and ANMs. In each of the myofibroblast cultures over 99% of cells exhibited this expression profile and morphology, indicating that these are essentially pure myofibroblast cultures with no contamination by epithelial, endothelial or other cell types. There were no differences in this respect between the 12 different sets of CAMs and ANMs.

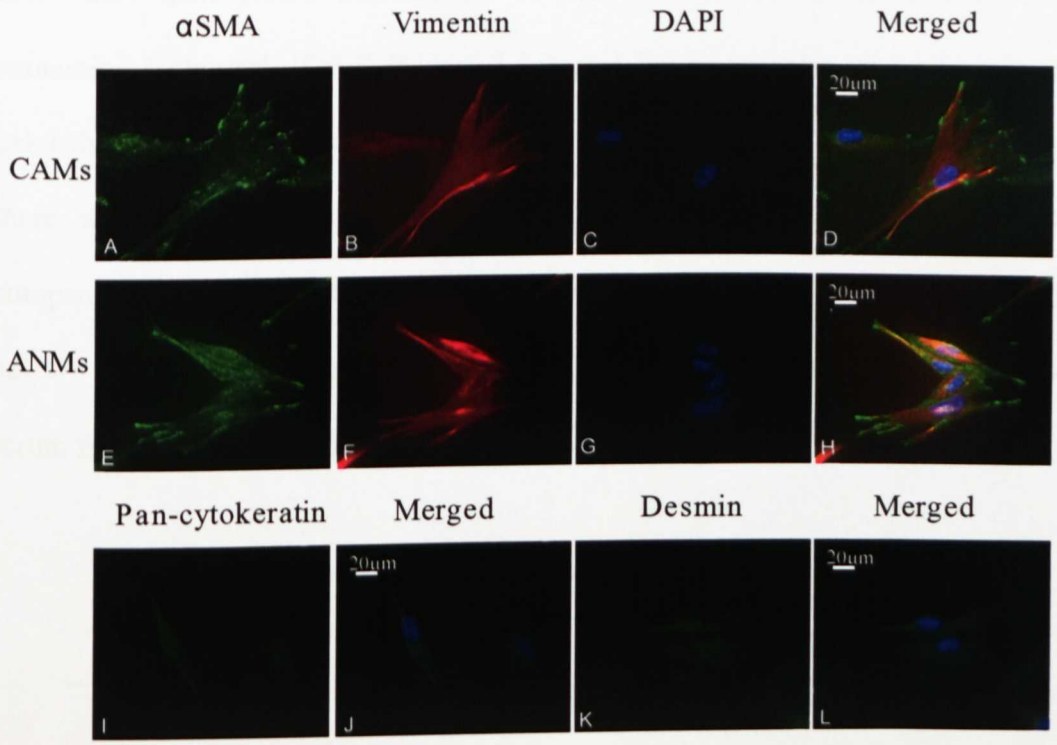


Figure 3.1

Cultured gastric myofibroblasts express α-SMA and vimentin, but not pan-cytokeratin or desmin.
Myofibroblast cultures derived from both gastric carcinoma (CAMs) (A-D) and adjacent non-cancer tissue (ANMs) (E-H) expressed both α-SMA (green) and vimentin (red). The myofibroblasts did not express pan-cytokeratin (green, I-J) or desmin (green K-L). Nuclei shown in blue (C, G, merged in D, H, J, L). Magnification x40.

3.3.2 Cancer derived myofibroblasts show greater basal and IGF-II induced migration than their adjacent normal counterparts

Gastric myofibroblasts have previously been shown to be responsive to IGF-II (Hemers et al., 2005), which was therefore used here to induce migratory responses. Migration of CAMs and ANMs under serum free conditions, both in the presence and absence of 100ng/mL rhIGF-II, was compared using modified Boyden-chamber migration assays. In 9 out of the 12 pairs of myofibroblasts the CAMs migrated more than their ANM counterparts in both serum-free media, and in media containing 100ng/mL IGF-II (Figure 3.2 A and B respectively). In the remaining 3 pairs the CAMs were less migratory under both conditions. When taken together there was a statistically significant ($p<0.05$) increase in migration of CAMs compared with ANMs, both in serum free media and with IGF-II stimulation (Figure 3.2 C). The increase in migration stimulated by IGF-II treatment compared with serum free in both the CAMs and ANMs was also significant.

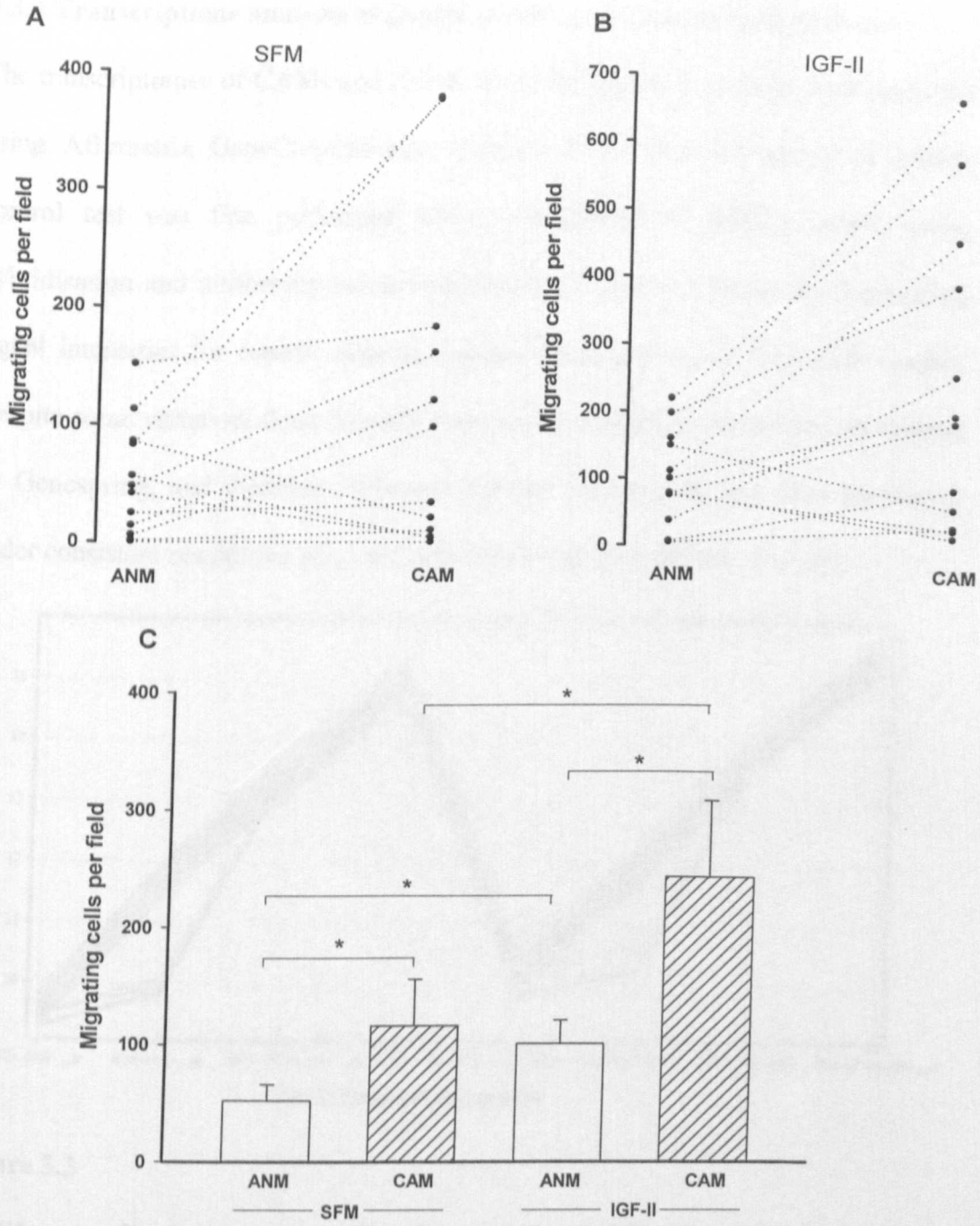


Figure 3.2

Basal and IGF-II-stimulated migration is increased in cancer derived myofibroblasts compared with their adjacent normal counterparts.
A, in 9 out of 12 pairs of gastric myofibroblasts the cancer-derived cells (CAMs) migrated more than adjacent non-cancer tissue-derived cells (ANMs) from the same patient in a modified Boyden chamber assay under both serum-free conditions and B, in the presence of 100ng/mL IGF-II. C, mean values from each condition reveal a significant increase in the migration of cancer derived myofibroblasts compared with adjacent normal derived myofibroblasts under both conditions. IGF-II stimulated cells are significantly more migratory than unstimulated myofibroblasts, with 2-fold increase in the mean migration observed with both cancer and adjacent normal derived myofibroblasts. N=12, $p<0.05$, t-test.

3.3.3 Transcriptome analysis of gastric cancer associated myofibroblasts

The transcriptomes of CAMs and ANMs from all 12 patient samples were analysed using Affymetrix GeneChip©Human Genome U133 Plus 2.0 arrays. A quality control test was first performed within GeneSpring to ensure correct probe hybridisation and uniformity across experiments. Figure 3.3 shows the normalised signal intensities for control oligonucleotides which are spiked into each sample. Despite some variation, these controls were within acceptable boundaries as defined by Genespring, and therefore indicated that the experiments had been performed under consistent conditions and were therefore suitable for further analysis.

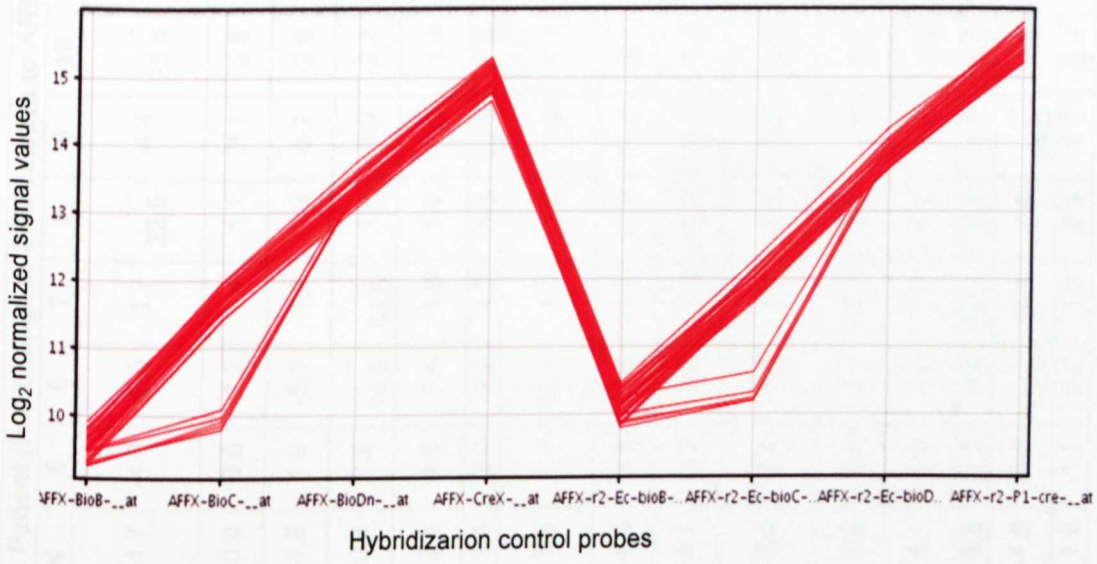


Figure 3.3

MAS5 normalized expression values for AFFX hybridisation controls for Affy U133 Plus 2 expression arrays of myofibroblasts.
Each experiment, represented by a single red line, had similar expression values for hybridisation controls, indicated good quality uniform hybridisation across experiments.

The 12 pairs of CAM and ANM experiments were analysed by paired t-test using GeneSping GX (Agilent). The genes which were significantly different ($p<0.01$) between the two groups, and with a mean fold change of greater than 2, are shown in Table 3.1.

Protein name	Patient (fold change in CAM relative to ANM)													
	1	2	3	4	5	6	7	8	9	10	11	12	Mean	SD
1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase gamma-2	2.7	-12.0	-11.7	-1.7	-4.8	-2.6	1.7	-22.5	-4.4	-25.4	-2.4	2.2	**-6.7**	8.9
Acetylcholine receptor subunit alpha	3.8	5.7	1.7	10.0	8.0	-1.2	1.4	1.1	7.7	5.8	3.5	3.2	**4.2**	3.2
AIG2-like domain-containing protein 1	-7.3	-5.8	-2.3	-1.9	1.9	-5.1	1.3	1.9	-5.2	-2.9	-5.8	-6.9	**-3.2**	3.2
Alcohol dehydrogenase 1C	-1.2	-3.8	-35.8	-10.3	2.2	-14.5	-44.6	1.4	-5.0	-2.2	-2.5	-1.2	**-9.8**	14.4
Aldehyde dehydrogenase family 8 member A1	1.2	1.3	8.9	2.5	4.8	1.4	1.5	2.6	-1.0	1.9	-1.1	1.5	**2.1**	2.5
Amphiregulin	16.7	3.5	9.8	1.4	2.0	-1.2	1.7	1.3	15.5	-1.2	1.7	1.8	**4.4**	5.9
ANKRD26-like family C member ENSP00000349402	1.7	1.6	1.1	3.6	2.7	10.1	1.1	1.6	1.5	2.7	1.1	-1.2	**2.3**	2.6
Ankyrin repeat and SOCS box protein 5	1.6	-2.4	-1.6	-4.5	-3.3	-3.5	-1.6	-1.7	-4.1	-3.7	-2.0	-5.1	**-2.7**	1.7
Atrial natriuretic peptide receptor A	-2.2	-5.8	-18.0	-14.7	-4.3	-5.6	-7.4	-2.8	-5.6	-9.4	-22.7	1.6	**-8.1**	6.7
Bromodomain and WD repeat-containing protein 1	1.0	-4.4	-1.8	-3.0	-3.4	-1.8	1.4	-1.1	-4.2	-4.3	-3.9	-1.2	**-2.2**	1.9
cAMP-dependent protein kinase inhibitor beta	2.5	8.8	-11.7	8.2	6.5	2.7	13.7	7.1	3.4	2.7	1.7	4.5	**4.2**	5.8
cAMP-specific 3',5'-cyclic phosphodiesterase 7B	2.7	2.9	4.8	5.1	1.9	-1.6	2.7	1.6	1.3	-1.2	1.2	2.5	**2.0**	1.9
Cannabinoid receptor 1	2.2	2.4	5.2	8.5	-1.2	3.4	22.3	9.7	-1.7	15.5	-1.3	1.1	**5.5**	7.1
CASP8-associated protein 2	-10.4	-30.7	1.3	-4.8	-5.8	-11.1	-9.1	7.8	-14.0	1.0	-1.6	-28.5	**-8.8**	11.0
Cell adhesion molecule 4	2.5	2.1	3.7	1.9	-1.1	2.0	1.3	2.9	1.3	5.5	3.0	1.2	**2.2**	1.5

Cell cycle control protein 50B	-2.1	-2.4	-2.6	-17.0	-1.4	-2.1	-2.1	-7.3	-2.3	-1.6	-1.1	-1.0	**-3.6**	4.3
Ceruloplasmin	4.3	2.8	-1.1	12.0	15.3	6.9	9.9	6.1	38.2	3.4	2.4	-1.8	**8.2**	10.3
Condensin complex subunit 2	-2.3	1.1	-1.4	-2.9	1.7	-3.2	-1.1	-3.7	-11.4	-1.2	-9.1	-1.9	**-3.0**	3.6
Dymeclin	-1.9	-1.2	-2.5	-4.0	-1.9	-4.7	-2.7	-47.5	-1.3	1.5	-4.0	-3.3	**-6.1**	12.6
EMILIN-2	-3.3	1.3	-3.3	-48.0	-4.1	-3.5	-5.0	-6.0	-1.8	-1.4	-3.0	1.1	**-6.4**	12.7
Epididymal secretory protein E3-alpha	1.9	2.2	11.4	1.3	11.4	5.5	7.2	8.8	2.2	9.7	6.3	4.0	**6.0**	3.6
FGGY carbohydrate kinase domain-containing protein	-1.9	1.6	-22.3	-5.4	1.2	-1.6	-7.7	-1.6	-7.7	-9.4	-2.3	-2.2	**-5.0**	6.2
FYVE, RhoGEF and PH domain-containing protein 4	-1.5	-1.9	-3.7	-1.4	-1.8	-1.3	-2.1	-3.6	-1.5	-1.6	-2.3	-1.7	**-2.0**	0.8
G protein-activated inward rectifier potassium channel 2	61.2	46.7	277.0	1.7	-1.6	2.5	8.2	3.7	58.4	1.3	-2.8	4.8	**38.4**	75.5
Glycoprotein Xg	-3.7	-8.1	-40.0	-19.2	-6.2	-1.5	-5.0	-2.9	-5.9	-6.4	-6.9	-8.8	**-9.5**	10.1
Headcase protein homolog	-4.5	-16.1	-1.9	-1.9	1.3	-2.8	-3.2	-1.1	1.1	-3.2	-1.9	-6.8	**-3.4**	4.4
Heparan sulfate 2-O-sulfotransferase 1	-1.3	-3.2	-2.8	-4.7	-1.2	-8.2	-2.2	-2.3	1.6	-1.4	-1.1	1.0	**-2.2**	2.4
Homeobox protein Hox-C8	-9.9	-8.6	-13.2	-2.5	-3.2	1.1	-3.5	-1.5	-1.4	-1.4	-1.6	-1.1	**-3.9**	4.1
Homogentisate 1,2-dioxygenase	1.8	-2.1	-15.2	-11.8	-1.2	-2.6	-2.7	-3.3	-2.2	1.1	-3.2	-5.3	**-3.9**	4.7
Interleukin-27 receptor subunit alpha	10.2	3.9	9.0	30.9	4.3	1.8	2.1	19.4	3.0	11.8	1.2	-4.1	**7.8**	9.1
Kremen protein 1	3.5	3.4	1.8	1.5	2.1	-1.7	2.0	2.4	2.4	1.8	1.6	2.8	**2.0**	1.3
Leucine-rich repeat and death domain-containing protein LOC401387	2.1	-3.7	-2.0	-11.0	1.4	-14.2	-8.4	-1.4	-1.1	-18.9	-1.4	-28.0	**-7.2**	8.8

LOC100130740	8.6	26.7	4.0	-1.6	2.3	1.7	1.2	8.1	3.1	9.7	1.4	8.4	**6.1**	7.1
LOC100132552	1.8	-2.1	-15.2	-11.8	-1.2	-2.6	-2.7	-3.3	-2.2	1.1	-3.2	-5.3	**-3.9**	4.7
LOC286434	1.8	7.9	2.3	2.7	1.2	3.1	12.4	3.0	2.5	1.5	1.2	42.1	**6.8**	11.1
LOC440863	4.4	2.0	13.1	2.3	-1.1	11.6	1.5	2.3	-1.3	1.4	-1.1	3.8	**3.2**	4.4
LOC641364	-4.7	-12.6	-1.9	-2.1	-2.9	-4.2	-15.0	-2.9	-3.8	3.0	-2.1	-3.0	**-4.4**	4.6
LOC652565	1.2	-5.2	-2.6	1.8	-4.3	-2.3	-8.7	-2.7	-2.6	-2.0	1.7	-8.7	**-2.9**	3.4
LOC727722	1.8	-2.1	-15.2	-11.8	-1.2	-2.6	-2.7	-3.3	-2.2	1.1	-3.2	-5.3	**-3.9**	4.7
LOC730139	-1.1	-1.9	-1.9	-2.9	-1.7	-5.8	-1.1	-1.7	-1.9	-1.7	-1.9	-1.8	**-2.1**	1.2
LOC730259	2.1	1.3	3.8	-1.1	-1.1	2.6	6.3	1.8	1.8	1.6	4.3	9.4	**2.8**	2.8
LOC84856	-1.7	2.0	1.0	3.2	1.2	10.3	4.0	10.3	13.0	1.4	22.9	5.1	**6.1**	6.6
Matrix-remodeling-associated protein 5	6.0	15.6	-1.0	2.4	5.0	1.3	1.9	1.4	5.7	1.3	-1.7	6.2	**3.7**	4.4
Membrane protein FAM174B	-1.4	-1.2	-14.8	-1.8	-1.3	-2.4	-1.1	-1.8	-3.3	-1.8	-3.2	-1.7	**-3.0**	3.6
Neighbor of punc e11	3.5	5.2	5.3	5.1	2.1	-1.9	3.5	1.3	4.8	1.8	1.4	2.7	**2.9**	2.0
Neurofascin	-2.3	-6.6	-28.6	-4.1	-2.9	-1.2	-1.6	-3.4	-1.5	-1.5	1.2	-1.8	**-4.5**	7.5
Neuropilin-2	21.3	2.1	3.5	7.2	11.6	-1.2	6.6	9.1	2.1	1.8	4.6	2.4	**5.9**	5.8
Normal mucosa of esophagus-specific gene 1 protein	1.7	5.0	8.6	-1.3	1.6	2.6	2.5	1.2	2.1	2.1	3.3	1.4	**2.6**	2.3
NUDT16P	-2.1	-2.1	1.2	-5.4	-1.8	-1.2	-1.7	-4.1	-1.2	-2.7	-5.7	-1.5	**-2.3**	1.8
Olfactory receptor 2F2	3.0	16.7	3.5	1.3	4.9	3.0	1.8	2.2	4.0	-1.7	1.6	1.3	**3.5**	4.3
Outer dense fiber protein 4	-2.3	-2.4	-6.2	-3.9	-1.1	-1.3	-1.3	-1.1	-1.8	-1.3	-4.8	-1.5	**-2.4**	1.6
Paired amphipathic helix protein Sin3a	1.6	-4.1	-2.3	-1.3	-1.1	-11.7	-2.3	-1.0	-2.0	-3.7	-2.1	-5.8	**-3.0**	3.2
Patatin-like phospholipase domain-containing protein 5	-1.0	5.0	-1.0	5.8	3.1	1.8	13.7	1.9	2.1	1.6	2.1	7.9	**3.6**	3.9
Phospholipase A1 member A	-7.9	-8.2	-31.8	-290.8	-4.6	-1.4	-17.0	-8.3	-2.8	-3.3	1.4	1.5	**-31.1**	78.8

Platelet glycoprotein 4	-7.8	-1.1	-35.7	-5.2	-2.6	-1.4	-3.0	1.2	-1.9	-1.2	-87.1	-21.9	**-14.0**	24.4
Poly(rC)-binding protein 3	3.4	14.2	-1.5	6.8	-1.3	2.0	2.4	4.6	4.3	1.4	2.4	9.6	**4.0**	4.2
PPP1R2P4	-2.4	-1.6	-9.4	-6.2	-1.4	-2.9	-4.1	-2.9	-1.8	-2.9	1.0	1.4	**-2.8**	2.8
Probable G-protein coupled receptor 144	-1.5	-4.2	1.6	-5.5	-4.1	-1.0	-3.3	1.5	-3.5	-3.8	-5.0	-1.9	**-2.6**	2.2
Pro-epidermal growth factor	-10.7	-1.4	-4.0	1.3	-2.9	-2.0	-2.3	-4.6	-6.2	-2.7	1.6	1.6	**-2.7**	3.4
Progestin and adipoQ receptor family member 9	3.9	3.6	-1.1	1.3	2.1	-2.1	4.5	6.8	2.7	1.4	-1.0	2.2	**2.0**	2.4
Protein FAM105A	1.2	4.3	8.7	26.4	6.3	-2.0	2.0	1.5	1.9	5.8	7.7	3.2	**5.6**	6.9
Protein FAM149A	-2.6	-2.0	-4.5	-4.2	-3.5	-1.3	-3.6	-1.6	-2.5	-1.2	-1.4	-1.4	**-2.5**	1.1
Protein kinase C eta type	-2.3	-13.9	-3.4	-8.5	1.3	-1.4	-1.5	-7.9	-2.5	-6.5	-12.7	-1.3	**-5.1**	4.6
Protein Wnt-5a	3.1	2.5	3.8	3.3	1.5	-1.6	2.1	2.7	4.5	-1.1	2.2	2.4	**2.1**	1.7
Putative DNA repair and recombination protein RAD26-like	1.4	-7.4	-4.6	1.4	-16.1	-3.9	-3.4	-1.8	1.1	-3.0	-4.8	-1.4	**-3.5**	4.6
Putative T-cell surface glycoprotein CD8 beta-2 chain	1.2	-5.2	-2.6	1.8	-4.3	-2.3	-8.7	-2.7	-2.6	-2.0	1.7	-8.7	**-2.9**	3.4
Putative uncharacterized protein FLJ43080	-1.3	-14.4	-3.4	-13.3	-2.1	-53.1	-7.4	-1.9	-2.0	-1.5	1.4	1.3	**-8.1**	14.4
Putative uncharacterized protein LOC404266	5.0	10.3	7.0	-1.5	1.7	-1.1	1.0	4.6	1.9	3.6	2.9	3.9	**3.3**	3.2
Receptor-type tyrosine-protein phosphatase epsilon	4.4	7.9	18.7	3.9	3.6	1.9	5.0	4.4	1.5	2.6	-2.1	2.3	**4.5**	4.8
Receptor-type tyrosine-protein phosphatase O	-2.0	-2.8	-9.7	-1.7	-1.5	-4.4	1.2	-2.8	-1.3	-4.8	-5.0	1.7	**-2.7**	2.9
Retinol-binding protein 1	-1.7	-1.6	2.7	4.1	2.5	4.7	4.4	6.1	-1.3	1.6	5.4	2.2	**2.4**	2.6

Ribonucleoside-diphosphate reductase large subunit	2.1	-12.5	-2.7	-2.2	-2.0	-14.5	-7.8	-3.0	-2.1	-2.3	1.6	-26.9	**-6.0**	7.9
RUN domain-containing protein 2B	-8.2	1.3	-8.8	-7.6	-1.5	-3.9	-4.0	-1.1	-1.0	1.1	-3.2	-1.3	**-3.2**	3.3
Scm-like with four MBT domains protein 2	4.6	22.9	-1.0	11.2	2.6	2.6	2.9	1.3	-1.5	6.2	2.1	-1.3	**4.4**	6.5
Seizure 6-like protein	-2.4	-1.3	-3.5	-4.3	-4.5	-1.5	-1.8	-2.1	1.5	-3.4	-2.6	-10.4	**-3.0**	2.7
Serine/threonine-protein kinase DCLK1	-1.3	-1.5	-7.9	-8.3	-1.8	-2.2	-2.4	-10.6	-22.8	-1.6	-12.3	1.4	**-5.9**	6.5
Sialic acid-binding Ig-like lectin 15	1.6	-2.8	-4.4	-2.9	-2.1	-2.4	-4.8	-2.1	-3.8	1.5	-3.3	1.1	**-2.0**	2.2
SNHG10	-1.6	4.9	10.3	2.9	19.1	4.6	1.4	1.3	1.4	2.3	2.4	10.4	**4.9**	5.5
Spermatogenesis- and oogenesis-specific basic helix-loop-helix-containing protein 2	-1.9	-1.3	-1.7	-3.5	-2.6	-1.7	-2.2	-8.6	-1.3	2.2	-5.9	-1.5	**-2.5**	2.5
SRC-like-adapter	1.5	4.4	40.1	9.9	-1.3	6.3	-1.4	11.2	1.8	1.7	3.4	1.4	**6.6**	10.8
Synaptopodin-2	-2.8	-5.1	-4.3	-11.6	-1.1	-2.4	-1.5	-4.1	-4.3	-2.9	1.7	-2.5	**-3.4**	3.0
T-cell surface glycoprotein CD8 beta chain	1.2	-5.2	-2.6	1.8	-4.3	-2.3	-8.7	-2.7	-2.6	-2.0	1.7	-8.7	**-2.9**	3.4
Tenascin-X	-2.2	-13.1	-14.8	-34.9	-2.1	-1.9	-22.6	-2.3	-4.3	-4.6	1.5	-15.4	**-9.7**	10.3
Transcription factor AP-2 alpha	2.8	10.9	14.1	1.3	4.9	2.0	4.8	3.9	1.4	3.3	1.9	-1.7	**4.1**	4.2
Transcription factor COE2	-1.3	-3.0	-11.5	-24.5	-2.0	-5.9	-9.7	-1.7	-5.8	-10.4	-2.1	1.3	**-6.4**	6.7
Transforming growth factor beta-2	2.6	3.0	11.7	20.7	1.3	2.0	3.9	3.1	3.9	1.4	3.1	2.8	**5.0**	5.4
Transmembrane protein 86B	-1.1	-7.7	-4.9	-1.5	-4.0	-2.9	1.2	-1.0	-1.5	-2.6	1.0	-1.4	**-2.2**	2.4

Tumor necrosis factor ligand superfamily member 18	2.5	2.5	-1.6	4.0	9.2	3.6	2.1	3.1	1.3	4.5	1.2	1.9	**2.8**	2.4
Tyrosine-protein phosphatase non-receptor type 22	-1.1	8.4	6.4	-2.0	1.2	14.8	17.8	1.3	20.6	1.6	6.0	1.8	**6.4**	7.3
Uncharacterized protein C13orf18	-2.4	-1.6	-9.4	-6.2	-1.4	-2.9	-4.1	-2.9	-1.8	-2.9	1.0	1.4	**-2.8**	2.8
Uncharacterized protein C18orf16	-5.4	-1.1	-10.1	-1.4	-16.0	-3.6	-3.5	-2.4	-1.1	-1.1	-9.2	-4.0	**-4.9**	4.4
Uncharacterized protein C18orf2	22.4	9.1	7.9	2.2	1.9	2.5	4.9	5.8	-1.4	3.3	1.3	-3.1	**4.7**	6.3
Uncharacterized protein C6orf168	14.1	6.2	1.8	2.0	1.1	6.7	7.8	3.0	1.9	4.6	1.0	3.1	**4.4**	3.6
Uncharacterized protein C8orf77	-12.2	-2.1	-23.0	-4.1	1.6	-12.7	2.2	-2.2	-1.2	-21.8	-3.6	-7.9	**-7.3**	8.1
Uncharacterized protein C9orf109	1.6	20.6	3.2	4.4	8.0	2.6	4.8	-3.0	4.1	2.2	-1.5	2.0	**4.1**	5.7
WNT1-inducible-signaling pathway protein 1	-1.1	10.3	3.1	1.6	1.9	3.2	2.0	2.3	1.4	8.1	1.1	1.8	**3.0**	3.0
Zinc finger protein 718	1.1	-6.8	-3.4	-9.4	-3.5	-8.8	-2.8	1.3	-1.3	-3.4	-9.0	-1.2	**-3.9**	3.6

Table 3.1

Genes differentially expressed by CAMs compared to ANMs by paired t-test
The gene expression profiles of 12 pairs of CAMs and ANMs were compared using a paired t-test ($p<0.05$). Fold changes are expressed as gene expression in CAM relative to ANM for each pair.

3.3.4 Functional network analysis of gene expression data

The gene expression data was also analysed to identify functional networks which are affected by the differences between CAMs and ANMs. In order to do this, fold changes in expression intensities were generated for each oligonucleotide probe (oligo) between each CAM and ANM pair. These data were analysed using MetaCore (GeneGo), and filtered to remove all genes with an associated fold change of less than 2-fold. The remaining genes were analysed using MetaCore to determine the most enriched pathway maps from within the MetaCore database (Figure 3.4). The algorithm employed to do this assigned a p-value to each of the maps within the MetaCore database based on the number of proteins within the map which have associated transcriptome data within the dataset, together with how common each of these proteins are in the map database.

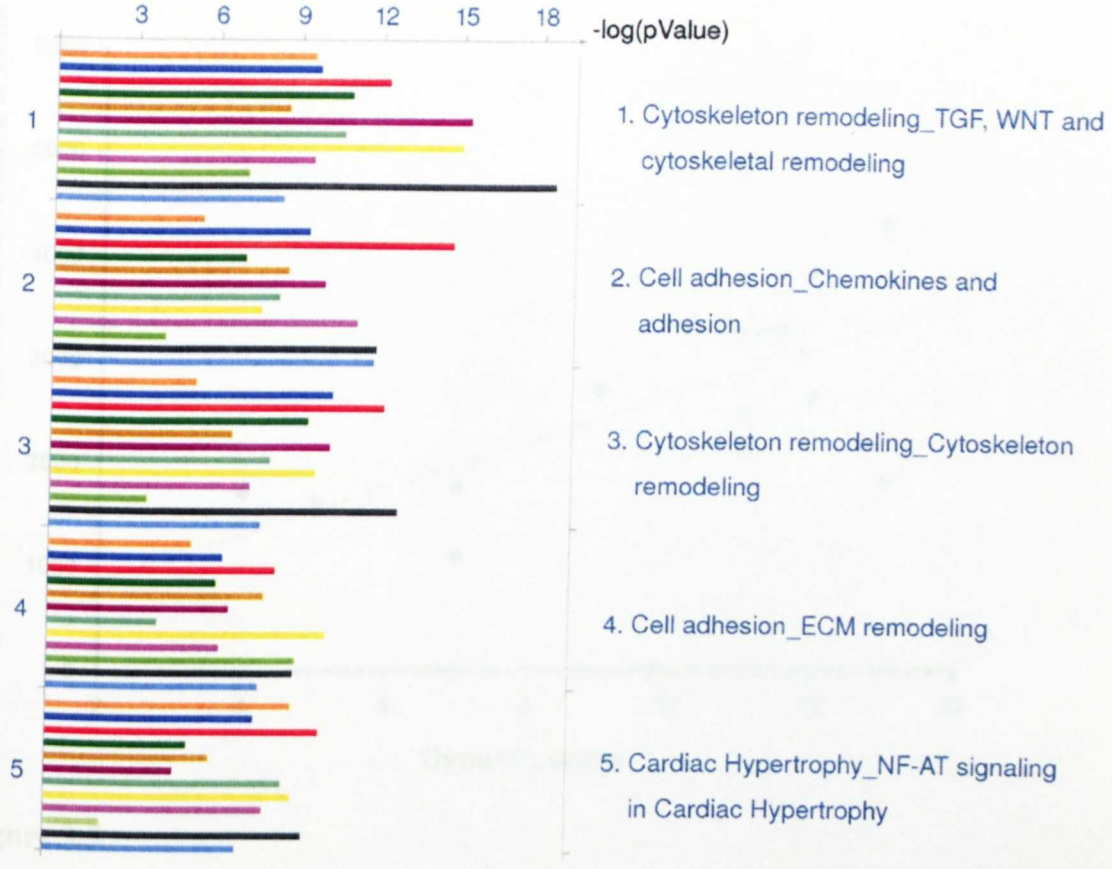


Figure 3.4

GeneGo maps most significantly enriched by differentially expressed genes from individually paired myofibroblast gene expression arrays
Individually paired gene arrays were used to calculate fold changes in gene expression between CAM and ANM experiments, and filtered for genes with fold changes greater than 1.5. The most significantly enriched GeneGo maps for these gene lists are shown, and the coloured bar represents the –log(p-value) of enrichment for a pair of samples from a single individual.

3.3.5 Correlation of gene expression data and clinical prognosis

The dynamic scores from each patient were used as measures of the severity of disease. These scores were compared with the gene expression data from the corresponding CAM and ANM pairs. Figure 3.5 shows a positive correlation between dynamic score and the number of genes whose expression differs more than 2-fold in the CAM relative to the corresponding ANM.

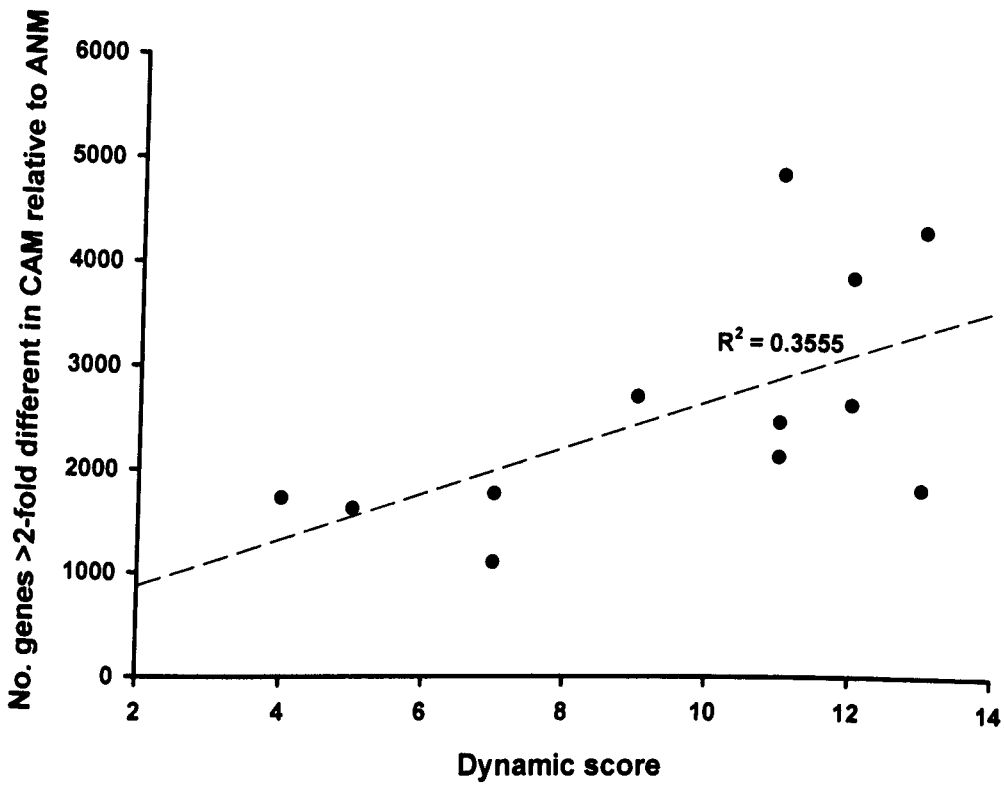


Figure 3.5

Correlation between changes in myofibroblast gene expression and dynamic scores.
The dynamic score assigned to each patient as a prognostic estimate (higher scores = worse prognosis) are plotted against the number of genes from the corresponding myofibroblast pair which show a fold change greater than 2 in CAM relative to ANM.

3.4 Discussion

This study is the first to compare the functional differences between gastric CAMs and ANMs, and the associated differences in gene expression. The cells used in this study were first confirmed as being myofibroblasts. Expression of both vimentin and α-smooth muscle actin, and the lack of expression of desmin and cytokeratin is defining of a myofibroblast. These data also indicate that the CAMs and ANMs are the same cell type, since there were no differences in either morphology or expression of markers.

It was apparent, however, that CAMs and ANMs have clear functional distinctions. It has been shown in our group that gastric CAMs proliferate faster than ANMs, with an accelerated progression through the G1 phase of the cell cycle (Charlotte Woodcock, unpublished data). The data presented here showed CAMs to also be more migratory than their ANM counterparts. This was consistent with previous observations of cancer associated fibroblasts from other tumours (Schor et al., 1988; Schor et al., 1985). It was previously shown that gastric myofibroblasts are responsive to IGF-II (Hemers et al., 2005), and this holds true for gastric CAMs and ANMs. Possible explanations for CAMs being more migratory include differences in the cytoskeletal migratory machinery, in secreted autocrine factors or the response to these factors, in ECM production or cellular adhesion to the ECM and plastics.

The comparison of gene expression profiles of CAMs and ANMs shed some light on the precise molecular differences between these two cell populations. The identification of 97 differentially expressed genes by paired t-test indicates a significant difference in gene expression between CAMs and their ANM counterparts. It is difficult, however, to identify the most important genes from this list. The genes with the greatest mean fold changes include those for G protein-activated inward rectifier potassium channel 2 and phospholipase A1 member A. In both cases though, some of the myofibroblast pairs have only a small change in relative expression of these genes. It seems then that the patterns of gene expression are different in each myofibroblast pair, and that the genes which are important in the functional differences between CAMs and ANMs might not be the same in each patient. It is worth noting here that care must be taken when using fold changes to compare gene expression profiles. Although care is taken during the normalisation and pre-processing of data in order to remove background signal noise, some of the

measured values might not be representative of the true level of gene expression. In these cases random fluctuations in signal noise could result in false estimations of fold changes. This is one of the reasons why statistical tests such as paired t-tests and ANOVAs are essential for an informative interpretation of gene expression array data (Allison et al., 2006; Breitling, 2006). These analyses, however, rely on finding consistency in the expression pattern of genes across many samples, which in this study each represent different patients. Since it is possible that the functional differences between CAM and ANMs might be the result of changes in the relative expression of a different set of genes in each patient, analyses at the level of genes might not be particularly informative. For this reason it is important to also analyse gene array data at the level of functional networks (Jones et al., 2008; Shipitsin et al., 2007; Wood et al., 2007).

MetaCore's database of signalling pathway maps was to identify the functional networks which were the most significantly enriched by the genes which were differentially expressed in CAMs relative to ANMs. This analysis revealed that the same small group of networks were the most significant for each patient. These pathways represent important molecular processes involved in cell adhesion, TGF and Wnt-regulated cytoskeletal remodelling, chemokines and ECM remodelling. These are all processes which have the potential to influence the proliferation and migration of cells. The identification of a TGFβ network is particularly interesting, since changes in the TGFβ signalling are strongly associated with cancer stroma (Bierie and Moses, 2006). Importantly, it was shown that mutating the gene for TGFβ receptor II in fibroblasts causes spontaneous tumour development in a murine model, including tumours in the forestomach (Bhowmick et al., 2004).

Whilst the functional networks identified were similar in each case, the probability of enrichment varied between patients. This is likely to be a consequence of variation in the number of genes which passed the 2-fold threshold assigned to the analysis. Since stromal gene expression has been shown to correlate to clinical outcome in breast cancer (Finak et al., 2008), the gene expression of these gastric myofibroblasts was compared with the clinical prognosis of the patients from which they were derived. This comparison revealed a positive correlation between the severity of the disease, as indicated by the dynamic scores, and the number of gene transcripts in the corresponding myofibroblast pair with a relative abundance greater than two-fold in CAM relative to ANM (Figure 3.5). It shows that the patients with a more severe disease also have a greater difference in the gene expression profile of CAMs relative to ANMs. Importantly, there was no correlation between changes in gene expression and any other known patient variable, including gender and age (not shown). It will be important to make similar comparisons with follow-up data from each patient as it becomes available, since these might give a more accurate indicator of disease severity.

The analyses presented here have provided some useful insights into the differences between CAMs and ANMs. It appears that genes encoding extracellular proteins account for many of the differences in gene expression (Figure 3.4). The extracellular environment is complex, with protein levels determined not only by gene expression, but translation of proteins, post-translation modifications, secretory mechanisms and rates of degradation. It will therefore be important to investigate the myofibroblast secreted proteins, known as the secretome, directly.

3.5 Conclusions

1. Both CAMs and ANMs have a myofibroblast phenotype.
2. Cancer derived myofibroblasts are more migratory than adjacent normal derived myofibroblasts.
3. There are large differences in gene expression between CAMs and their ANM counterparts.
4. Many of the differentially expressed genes encode extracellular proteins involved in processes such as TGFβ signalling.

Chapter 4

The influence of gastric cancer-associated myofibroblasts on cancer cell function and transcriptome

4.1 Introduction

The influence of the stroma on cancer cells is mediated largely via secreted factors (Bierie and Moses, 2006; Kuperwasser et al., 2004; Yauch et al., 2008). The influence of stroma on gastric carcinoma, however, is poorly understood. The cross-talk between myofibroblasts and epithelial cells in the stomach has been shown to be important in mediating some of the effects of *Helicobacter pylori* infection (Hemers et al., 2005; McCaig et al., 2006), which can induce pre-neoplastic remodelling of the stomach. It is important, therefore, to study the effects of gastric myofibroblasts on gastric carcinoma cells. The use of *in vitro* model systems of cancer is common since these systems and cell lines allow precise manipulation and analysis which are not always possible *in vivo*. A popular model system for gastric adenocarcinoma is the AGS cell line (Barranco et al., 1983). This is a well characterised and popular cell line derived from a single biopsy of an untreated human gastric adenocarcinoma (Barranco et al., 1983). When cultured they form dense monolayers of polygonal cells which are responsive to IGF-I (Guo et al., 1993) and gastrin (Ishizuka et al., 1991), both of which are important in gastric signalling events.

It was established in chapter 3 that there are significant differences between gastric CAMs and ANMs, both functionally and in terms of gene expression. Many of the differentially expressed genes were identified as being involved in extracellular signalling. It therefore becomes important to ask whether CAMs and ANMs differentially influence cancer cell behaviour. This chapter will investigate the effects of CAM and ANM media on AGS cells, using both functional assays of migration, invasion and proliferation, and comparing the effects on AGS cell gene expression.

4.1.1 Aims

The aim of this chapter is to determine the effect of myofibroblast secreted factors on gastric cancer cells. Specifically, the objectives were to;

1. Compare the functional effects of CAM and ANM conditioned media on a gastric cancer cell line.
2. Investigate the effects of CAM and ANM conditioned media on the gene expression profiles of a gastric cancer cell line.

4.2 Materials and Methods

4.2.1 Primary human myofibroblasts

Primary myofibroblasts were derived from resected gastric carcinoma and adjacent normal tissue from 12 patients as described in section 2.1.2. Myofibroblast conditioned media was prepared as described in section 2.2.1.

4.2.2 Immunocytochemistry

For immunocytochemistry staining experiments, AGS cells were cultured in 4-well chamber slides, at a seeding density of $4x10^4$cells/well. The AGS cells were grown in DMEM supplemented with 10% FBS and 1% penicillin/streptomycin for 24 hours prior to washing with PBS and treatment with either serum-free DMEM or myofibroblast conditioned media from pairs of CAMs and ANMs. Incubation with antibodies and visualisation of cells was performed as described in section 2.3.

4.2.3 Migration and invasion assays

Migration and invasion were assayed as described in section 2.5.1. AGS cells were split for migration and invasion assays by incubating with Versene EDTA solution

(Sigma) for 10 minutes and resuspended in serum-free DMEM (SFM). AGS cells were seeded at $2x10^4$ cells per well for migration assays, and $5x10^4$ cells per well for invasion assays, each condition performed in duplicate. The lower well of the assay was filled with either SFM or myofibroblast CM. Migration was measured after 18 hours incubation, and invasion after 24 hours.

4.2.4 [^{3}H]-thymidine incorporation proliferation assay

Proliferation assays were performed as described previously (Varro et al., 2002). Briefly, AGS cells were plated in 6-well plates at a density of $5x10^4$ cells/well. Following plating and synchronisation of cells in G0/G1 by serum starvation for 24 hours, AGS cells were treated with 2mL myofibroblast CM. Each experiment was performed in triplicate, and pairs of CAM and ANM CM were assayed on the same plates, such that one 6-well plate could assay one pair of myofibroblast CMs.

4.2.5 Statistical analysis of AGS functional assays

Where shown, data are presented as mean values +/- S.E.M., and significance determined by two-tailed paired t-test $p<0.05$.

4.2.6 Gene expression arrays

AGS cells were cultured at $1.5x10^6$ cells per $25cm^2$ tissue culture flask for 24 hours before being treated with either serum-free DMEM or myofibroblast CM for 24 hours. The CM used were from five randomly selected pairs of CAMs and ANMs from patients 3, 4, 9, 10 and 11. RNA extraction and array hybridisation were performed as described in Chapter 2. Raw image data from the gene chips was processed and analysed using GeneSpring GX (Agilent). Experiments were

normalised using the MAS5 algorithm with a baseline correction to the median of all samples, and expressed genes were identified by the flagging system consisting of 'Present' (P), 'Marginal' (M) and 'Absent' (A) calls generated by MAS5. In the case of comparisons made between data from individual pairs of CAM and ANM CM treatment experiments, data was filtered using the MAS5 flags such that in each comparison every transcript data point was flagged as P or M in at least one of the experiments. In the case of comparisons made between groups of experiments by either ANOVA or t-test the data were filtered such that every transcript data point was flagged as P or M in all of the experiments involved in the comparison. Further analysis of gene expression data was performed using MetaCore (GeneGo).

4.3 Results

4.3.1 Migration, invasion and proliferation of a gastric cell line is increased in response to cancer derived myofibroblast media

The influence of soluble factors derived from myofibroblasts on cancer cells was investigating by using myofibroblast conditioned media (CM) samples from 11 myofibroblast pairs to stimulate AGS gastric carcinoma cells in assays of migration, invasion and proliferation. In 11 out of the 11 pairs of myofibroblast conditioned media the CAM CM induced more migration and invasion of AGS cells than CM from their ANM counterparts (Figure 3.1 A and B respectively). As with the migration of the myofibroblasts themselves, there was variation in the absolute number of migrating cells in response to different CM samples. Even so, mean values of migration and invasion for all experiments (Figure 4.1 C and D respectively) revealed statistically significant ($p<0.05$) increases in response to CAM CM.

[^{3}H]-thymidine incorporation was used to measure the proliferation of AGS cells treated directly with myofibroblast CM. In every pair of myofibroblast CM tested AGS cell proliferation was greater in response to CAM CM than ANM CM (Figure 4.2 A). Taken together this increase was statistically significant ($p<0.05$) (Figure 4.2 B).

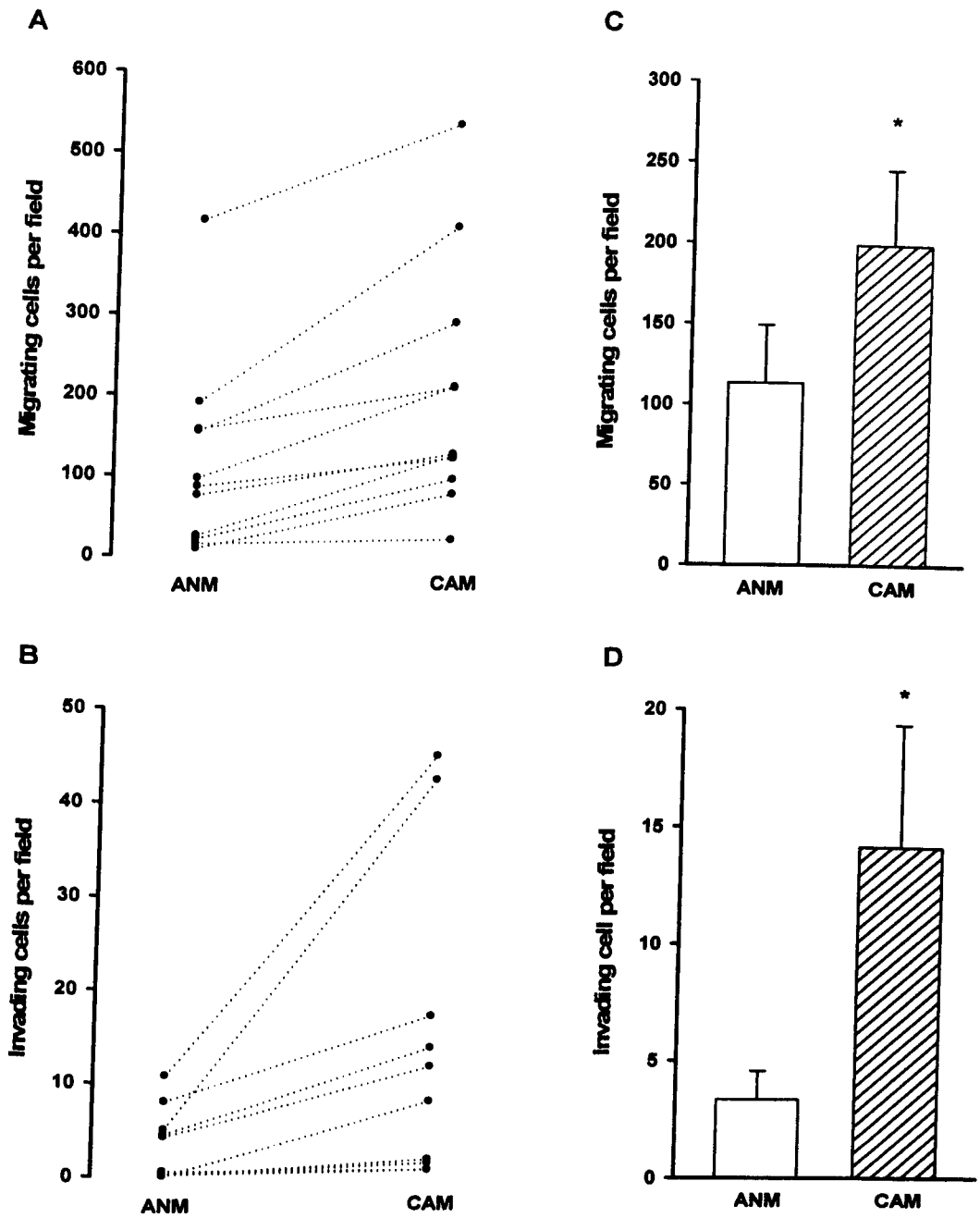


Figure 4.1

Cancer derived myofibroblasts induce a greater migration and invasion response in a gastric cancer cell line than their adjacent normal counterparts. Myofibroblast conditioned media induced migration and invasion of AGS cells in modified Boyden chamber assays. A, media from cancer-derived myofibroblasts (CAMs) induced both a greater migration (n=11) and B, invasion (n=9) response than media from adjacent non-cancer tissue-derived myofibroblasts (ANMs) from the same patient. C, mean values of migration and D, invasion across all experiments reveal a significant increase in migratory response ($p<0.05$, paired t-tests).

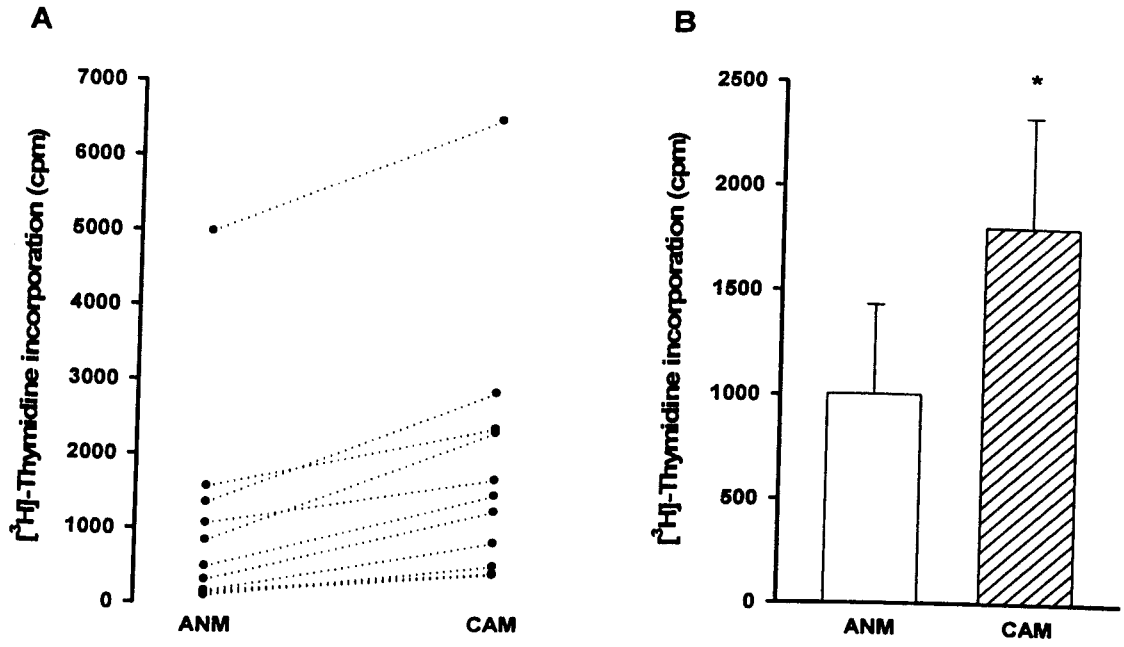


Figure 4.2

Cancer derived myofibroblasts induce a greater proliferation response in a gastric cancer cell line than their adjacent normal counterparts.
Myofibroblast conditioned media induced proliferation of AGS cells in [^{3}H]-thymidine incorporation assays. A, media from cancer-derived myofibroblasts (CAMs) induced a greater proliferation response than media from adjacent non-cancer tissue-derived myofibroblasts (ANMs) from the same patient. B, mean values of proliferation across all experiments reveal a significant increase in proliferative response. N=11, $p<0.05$, paired t-test.

4.3.2 Myofibroblast conditioned media induces EMT-like changes in a gastric cancer cell line

AGS cells were cultured with myofibroblast CM in order to further characterise the phenotypic changes induced. Immunocytochemistry was used to visualise F-actin filaments in order to reveal changes in cell morphology and cytoskeletal architecture. α-SMA expression was also examined since the changes in migration and invasion of AGS cells observed thus far suggested the acquisition of a mesenchymal-like phenotype. AGS cells cultured for 16 hours under control conditions form tight colonies of rounded cells (Figure 4.3 A and C). Under these

phenotype. AGS cells cultured for 16 hours under control conditions form tight colonies of rounded cells (Figure 4.3 A and C). Under these conditions AGS cells have little or no expression of α-SMA (Figure 4.3 B). When AGS cells were grown in the presence of myofibroblast CM for 16H there were dramatic changes in their appearance. Treated AGS cells were spindle-shaped, some with long projections, and exhibiting scattering. Staining for α-SMA revealed that myofibroblast CM had induced strong expression in this relatively short time. The acquisition of such properties suggest epithelial to mesenchymal transition (EMT) in the treated AGS cells. Treatments were performed using both CAM and ANM CM, but there was no observable difference between the two.

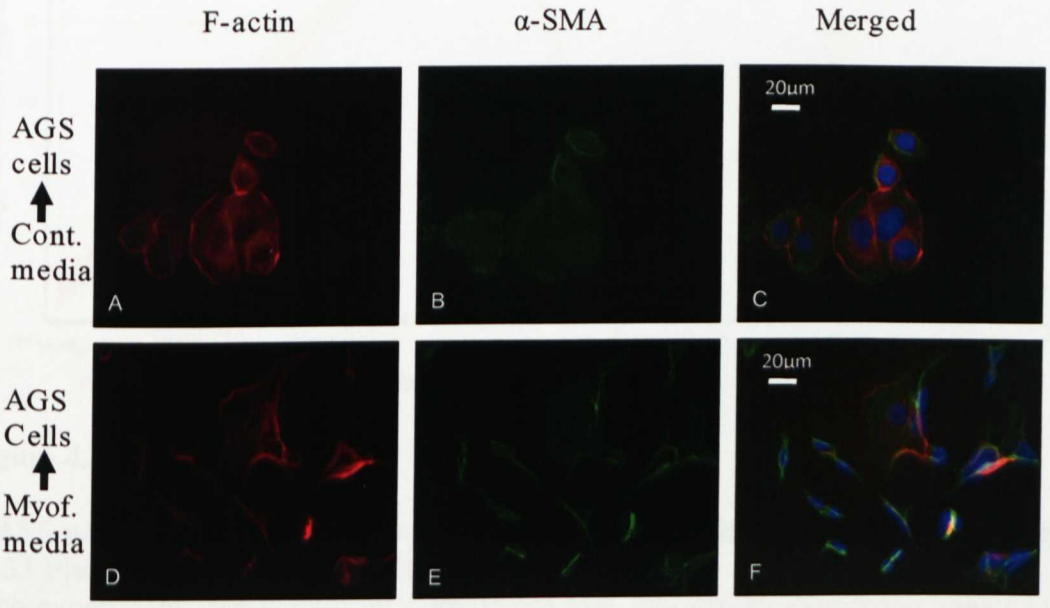


Figure 4.3

Myofibroblast conditioned media induces EMT-like changes in AGS cells within 16 hours.
A-C, AGS cells cultured under control conditions grow in tight colonies with a rounded morphology as revealed by the distribution of F-actin (red), and have little or no expression of α-SMA (green). D-F, AGS cells cultured in the presence of myofibroblast conditioned medium for 16 hours prior to fixation

4.3.3 Gene expression analysis of AGS cells treated with myofibroblast CM

A quality control test was first performed within GeneSpring to ensure correct probe hybridisation and uniformity across experiments. Figure 4.4 shows the normalised signal intensities for control oligonucleotides which are spiked into each sample. The good agreement observed between the experiments for each of the controls indicated that the experiments had been performed under consistent conditions and were therefore suitable for further analysis.

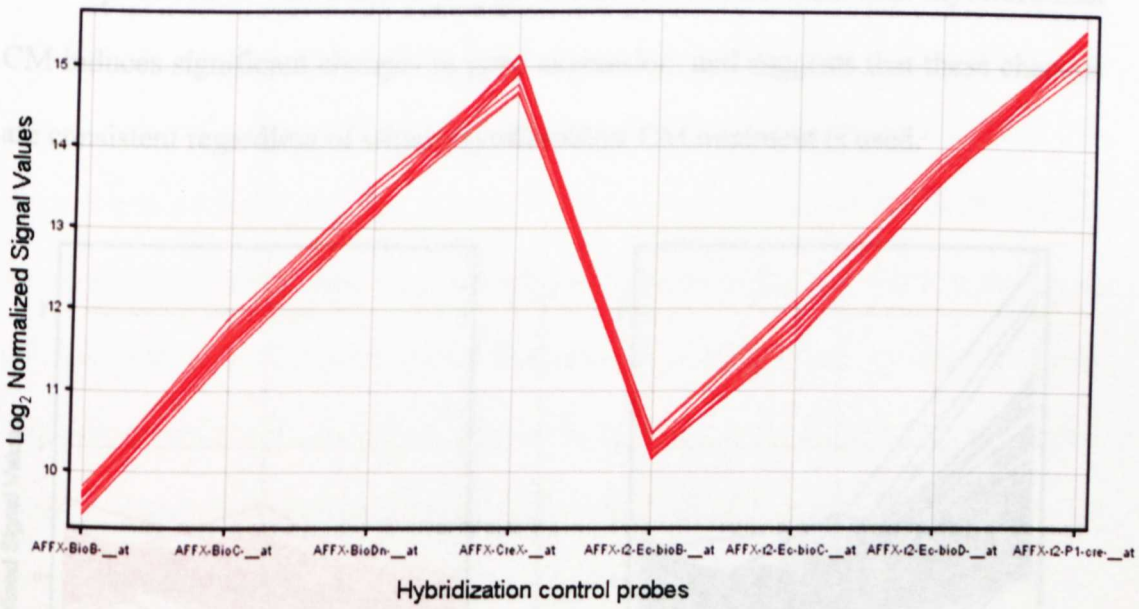


Figure 4.4

MAS5 normalized expression values for AFFX hybridisation controls for Affy U133 Plus 2 expression arrays of AGS cells.
Each experiment, represented by a single red line, had similar expression values for hybridisation controls, indicated good quality uniform hybridisation across experiments.

4.3.4 Comparison of AGS cell gene expression profiles by ANOVA

The experiments were grouped into three classes; SFM treated (Control), CAM CM treated (Cancer) and ANM CM treated (Adjacent). The data were analysed using a three way ANOVA test to generate a list of 887 oligonucleotide probes (oligos)

which showed differential expression of their associated transcripts across the three groups ($p<0.05$). Clustering of these 887 oligos by a K-means algorithm into 2 clusters revealed that the major differences in expression amongst these oligos were between the Control group and the two treatment groups; Adjacent Normal and Cancer (Figure 4.5). By choosing to split the oligos into two clusters we can see that a roughly even proportion of the signal values are increased and decreased in the control treated samples relative to both of the myofibroblast CM treatment groups. The implication of this analysis is that treatment of AGS cells with myofibroblast CM induces significant changes in gene expression, and suggests that these changes are consistent regardless of which myofibroblast CM treatment is used.

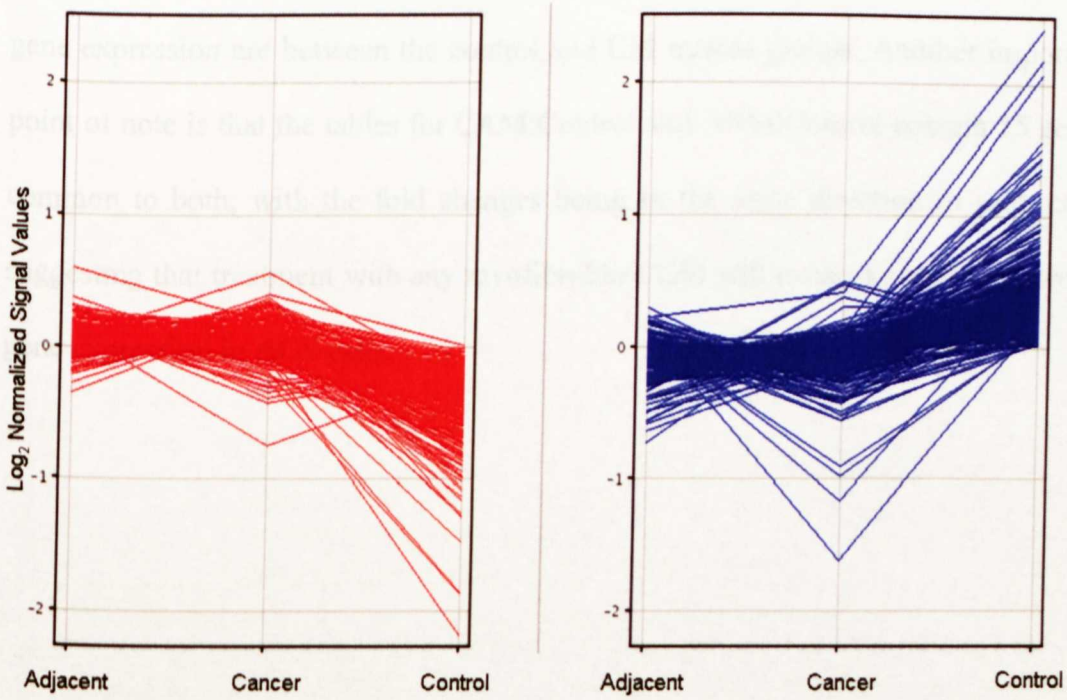


Figure 4.5

ANOVA results clustered by K-means algorithm
The differentially expressed oligos identified by ANOVA were clustered using the K-means algorithm to reveal that the majority of the differential expression is due to differences between the control and treatment groups. The figure shows the two clusters generated which show that oligos with low expression or high expression in the control group (red and blue respectively) show moderate expression in both the Adjacent and Cancer treatment groups.

For analysis at the gene level the oligo data were exported from GeneSpring and uploaded to MetaCore. MetaCore has the advantage of disregarding data from probes which Affymetrix have found to be unreliable and correcting errors in gene annotation. Table 4.1 shows the mean fold changes when comparing either CM treatment groups with the control group (Table 4.1 A and B respectively), and when comparing the CAM with the ANM CM treatments (Table 4.1 C). Only the ten most up-regulated and down-regulated for each group are shown, but this gives a useful insight into the changes in gene expression which have occurred. Most notable is the fact that the greatest increases and decreases between CAM and ANM are roughly 2-fold, whereas the greatest changes between these treatment groups and the control group are more than twice that, supporting the finding that the greatest differences in gene expression are between the control and CM treated groups. Another important point of note is that the tables for CAM:Control and ANM:Control contain 15 genes common to both, with the fold changes being in the same direction in each case, suggesting that treatment with any myofibroblast CM will evoke a similar pattern of gene expression in AGS cells.

A

Gene Symbol	Protein name	CAM: Control
RSAD2	Radical S-adenosyl methionine domain-containing protein 2	-4.16
MX2	Interferon-induced GTP-binding protein Mx2	-4.02
CREM	cAMP-responsive element modulator	-3.33
OAS1	2'-5'-oligoadenylate synthetase 1	-3.13
OAS3	2'-5'-oligoadenylate synthetase 3	-2.61
LAMP3	Lysosome-associated membrane glycoprotein 3	-2.61
BST2	Bone marrow stromal antigen 2	-2.54
USP18	Ubl carboxyl-terminal hydrolase 18	-2.50
IFIT5	Interferon-induced protein with tetratricopeptide repeats 5	-2.46
UNK	RING finger protein unkempt homolog	-2.40
MT1H	Metallothionein-1H	2.14
LOC100131177		2.16
SOCS3	Suppressor of cytokine signaling 3	2.55
MT1X	Metallothionein-1X	2.55
HIST1H2AD	Histone H2A type 1-D	2.60
REPS2	RalBP1-associated Eps domain-containing protein 2	2.66
HUS1B	Checkpoint protein HUS1B	3.06
CLN6	Ceroid-lipofuscinosis neuronal protein 6	3.13
MT1P2	MT-1H-like protein	3.17
IGFBP5	Insulin-like growth factor-binding protein 5	4.27

B

Gene Symbol	Protein name	ANM: Control
RSAD2	Radical S-adenosyl methionine domain-containing protein 2	-5.99
MX2	Interferon-induced GTP-binding protein Mx2	-4.60
OAS1	2'-5'-oligoadenylate synthetase 1	-3.54
OAS3	2'-5'-oligoadenylate synthetase 3	-3.13
BTN3A3	Butyrophilin subfamily 3 member A3	-3.09
PML	Probable transcription factor PML	-3.01
IFIT5	Interferon-induced protein with tetratricopeptide repeats 5	-2.92
BST2	Bone marrow stromal antigen 2	-2.87
USP18	Ubl carboxyl-terminal hydrolase 18	-2.74
LAMP3	Lysosome-associated membrane glycoprotein 3	-2.59
LOC647131		2.12
CLN6	Ceroid-lipofuscinosis neuronal protein 6	2.36
C3orf57	Protein ADMP	2.63
LOC100129285		2.63
MT1X	Metallothionein-1X	2.74
SOCS3	Suppressor of cytokine signaling 3	3.37
HIST1H2AD	Histone H2A type 1-D	3.61
MT1P2	MT-1H-like protein	3.85
HUS1B	Checkpoint protein HUS1B	4.50
IGFBP5	Insulin-like growth factor-binding protein 5	5.04

C

Gene Symbol	Protein name	CAM: ANM
CREM	cAMP-responsive element modulator	-2.36
LOC145783		-2.09
CYP1A2	Cytochrome P450 1A2	-1.98
CRIPAK	Cysteine-rich PAK1 inhibitor	-1.64
UNK	RING finger protein unkempt homolog	-1.63
RAB18	Ras-related protein Rab-18	-1.59
KIAA0415	Uncharacterized protein KIAA0415	-1.54
LOC100131354		-1.52
G3BP1	Ras GTPase-activating protein-binding protein 1	-1.47
HUS1B	Checkpoint protein HUS1B	-1.47
HFE	Hereditary hemochromatosis protein	1.46
ZBTB46	Zinc finger and BTB domain-containing protein 46	1.47
LOC100130740		1.53
LOC149684		1.61
PML	Probable transcription factor PML	1.62
STAU2	Double-stranded RNA-binding protein Staufen homolog 2	1.67
STXBP3	Syntaxin-binding protein 3	1.76
SLC40A1	Solute carrier family 40 member 1	1.77
EPOR	Erythropoietin receptor	1.84
FOXO4	Forkhead box protein O4	2.05

Table 4.1

Ten most highly up- and down- regulated genes between control, ANM and CAM CM treated AGS cells determined by ANOVA
Tables show the ten greatest increases and decrease in gene expression when comparing A, CAM treatment with control, B, ANM treatment with control, and C, CAM treatment with ANM treatment, using the gene list generated by three-way ANOVA analysis of all groups.

4.3.5 Comparison of CAM and ANM conditioned media treated AGS cells by paired t-test

Given the strong effect of any CM treatment on AGS gene expression, the CAM and ANM treatments were compared without the control group. A paired t-test was performed using gene expression data from AGS cells which had been treated by 5 pairs of myofibroblast CM. This analysis generated a list of 1160 oligos which showed differential expression of their associated transcripts ($p<0.05$). The normalized signal values for each experiment were exported and used to calculate a

fold change for each pair of experiments. These data were then uploaded to MetaCore for further analysis. Table 4.2 shows the ten greatest mean fold changes in gene expression between CAM and ANM CM treated AGS cells. When compared with the results from Table 4.1 C, it can be seen that the associated changes in gene expression are of a greater magnitude in the results from paired t-test than the ANOVA. Furthermore, there was no overlap of the genes identified by the paired t-test analysis and by ANOVA analysis when CAM and ANM CM treatments were compared.

MetaCore was used to identify the interaction maps from its database which were the most significantly enriched for these data, as shown in Table 4.3. MetaCore does this by assigning a p-value to each of the maps within its database based on the number of proteins within the map which have associated transcriptome data within the dataset, together with how common each of these proteins are in the map database. The most significant map 'Cytoskeleton remodeling_TGFβ, WNT and cytoskeletal remodeling' is shown in Figure 4.6.

Gene Symbol	Protein name	Fold change
TAGAP	T-cell activation Rho GTPase-activating protein	-6.65
C10orf46	Uncharacterized protein C10orf46	-6.62
RPS6KA5	Ribosomal protein S6 kinase alpha-5	-5.80
LRRC63	Leucine-rich repeat-containing protein 63	-4.98
SLC13A1	Solute carrier family 13 member 1	-4.11
IDS	Iduronate 2-sulfatase	-4.10
D4S234E	Neuron-specific protein family member 1	-4.01
NID2	Nidogen-2	-3.89
WDR19	WD repeat-containing protein 19	-3.83
ESM1	Endothelial cell-specific molecule 1	-3.70
CYLC1	Cylicin-1	4.11
C6orf182	Uncharacterized protein C6orf182	4.19
LOC93463		4.30
PLD1	Phospholipase D1	4.39
RXFP2	Relaxin receptor 2	4.98
KRTAP19-1	Keratin-associated protein 19-1	5.09
GIMAP1	GTPase IMAP family member 1	5.53
LIMS1	LIM and senescent cell antigen-like-containing domain protein 1	5.94
SLC10A7	Sodium/bile acid cotransporter 7	6.12
MCF2L2	Probable guanine nucleotide exchange factor MCF2L2	7.96

Table 4.2

Ten most highly up- and down- regulated genes when comparing CAM and ANM CM treated AGS cells
The genes showing the greatest fold changes when comparing CAM with ANM CM treatments using a paired t-test.

GeneGo Maps	pValue
Cytoskeleton remodeling_TGFβ, WNT and cytoskeletal remodeling	3.29E-06
Normal and pathological TGF-beta-mediated regulation of cell proliferation	1.38E-04
Cell adhesion_Alpha-4 integrins in cell migration and adhesion	1.64E-04
Transcription_Sin3 and NuRD in transcription regulation	3.09E-04
Development_Role of Activin A in cell differentiation and proliferation	4.13E-04
Apoptosis and survival_BAD phosphorylation	4.74E-04
Transcription_CREB pathway	5.41E-04
Immune response_IL-6 signaling pathway	6.56E-04
Proteolysis_Putative SUMO-1 pathway	6.56E-04
DNA damage_Role of SUMO in p53 regulation	6.81E-04

Table 4.3

Ten most highly enriched GeneGo maps by paired t-test results in MetaCore
The most significantly represented maps of signalling networks produced by MetaCore analysis of the paired t-test results. A low p-value indicates a low probability that the corresponding map is enriched in the data by chance.

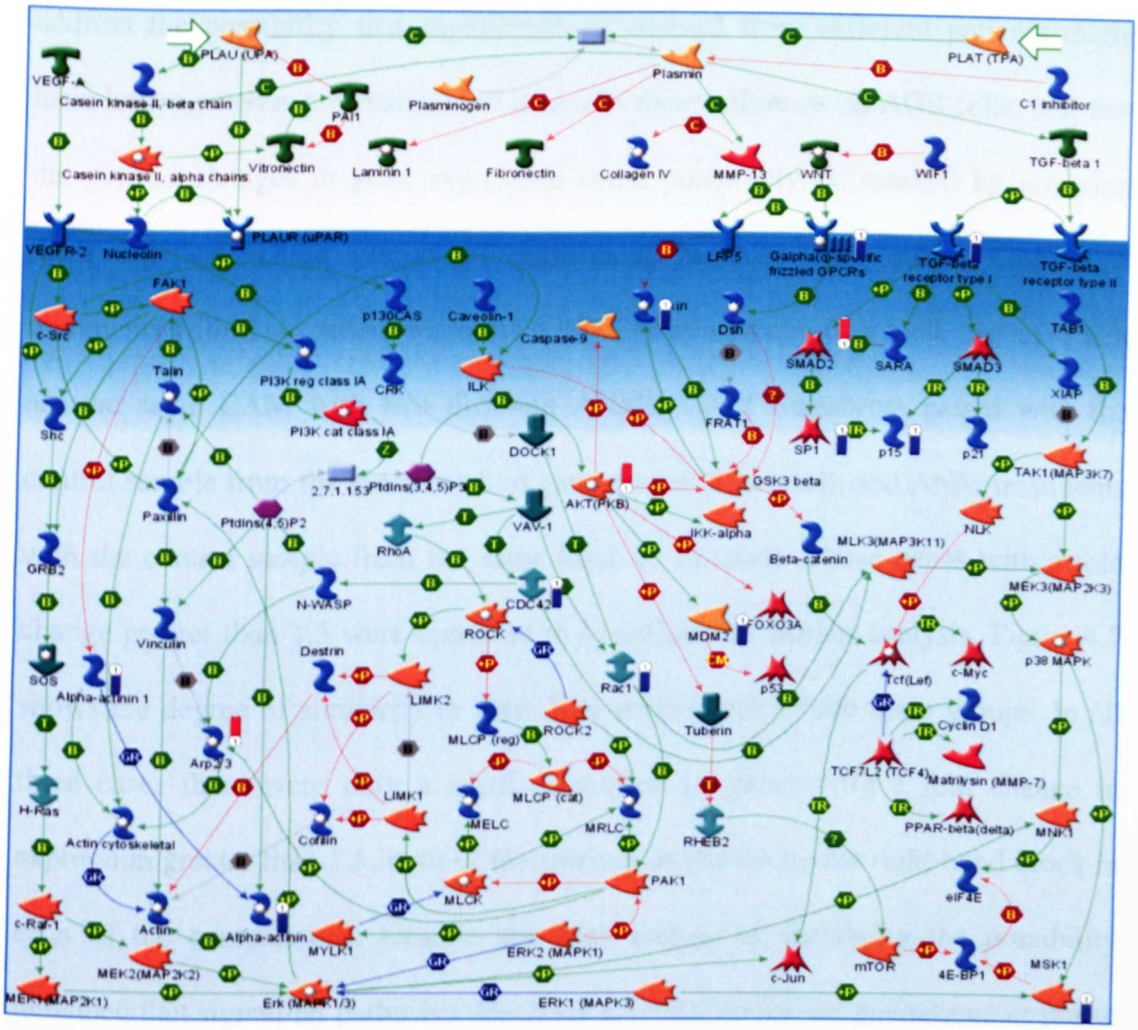


Figure 4.6

GeneGo map 'Cytoskeleton remodeling_TGFβ, WNT and cytoskeletal remodeling'
The most highly enriched GeneGo map in the analysis of paired t-test data. Proteins and other Network Objects are represented by various symbols and named, arrows and hexagons indicate published interactions. T-test data is shown by red and blue bars next to the relevant Network Object which represent increases and decreases in CAM relative to ANM respectively. (Key – See supplementary figure 1, page 192)

4.3.6 Comparison of CAM and ANM conditioned media treated AGS cells by functional networks

To ensure a thorough interrogation of the gene expression data, the experiments were separated into pairs to calculate fold changes and analysed separately without a prior statistical test. The reasoning behind this analysis is that it is important to

address the possibility that myofibroblasts derived from different patients might have heterogeneous mechanisms of action in their influence on AGS cells, and that the evoked changes in gene expression could potentially be masked by grouping them together. Three groups of experiment pairings were made; CAM CM treatments with their patient-matched ANM CM treatments (6 in total, two of which are the same CAM with two different ANM), CAM treatments paired with the control sample from the same batch of gene arrays (5 in total), and ANM treatments with the control sample from the same batch (6 in total). Those genes with a fold change greater than 1.5 were uploaded to MetaCore for further analysis. Figure 4.7 shows the degree of similarity in these lists within each of the three groups. In all three cases there were only a small proportion of genes with a fold change in expression greater than 1.5 in all of the pairings as shown by the right-hand block in each of the comparisons. Despite this high degree of variability the possibility remained that signalling pathways which are responsible for the phenotypic response were the same in each case. To test this hypothesis each group was analysed to determine the most significantly enriched maps as before. Figure 4.8 shows the three most significantly enriched maps for each of the three groups. It was clear from this that the same pathways were involved in each of the lists within the groups, despite the large differences at the individual gene level. There was also a large degree of similarity between the three groups, with the top two most significant maps being the same for each group, though not always in the same order. Of these maps only 'TGF, WNT and cytoskeletal remodelling' had already been identified in the previous analyses.

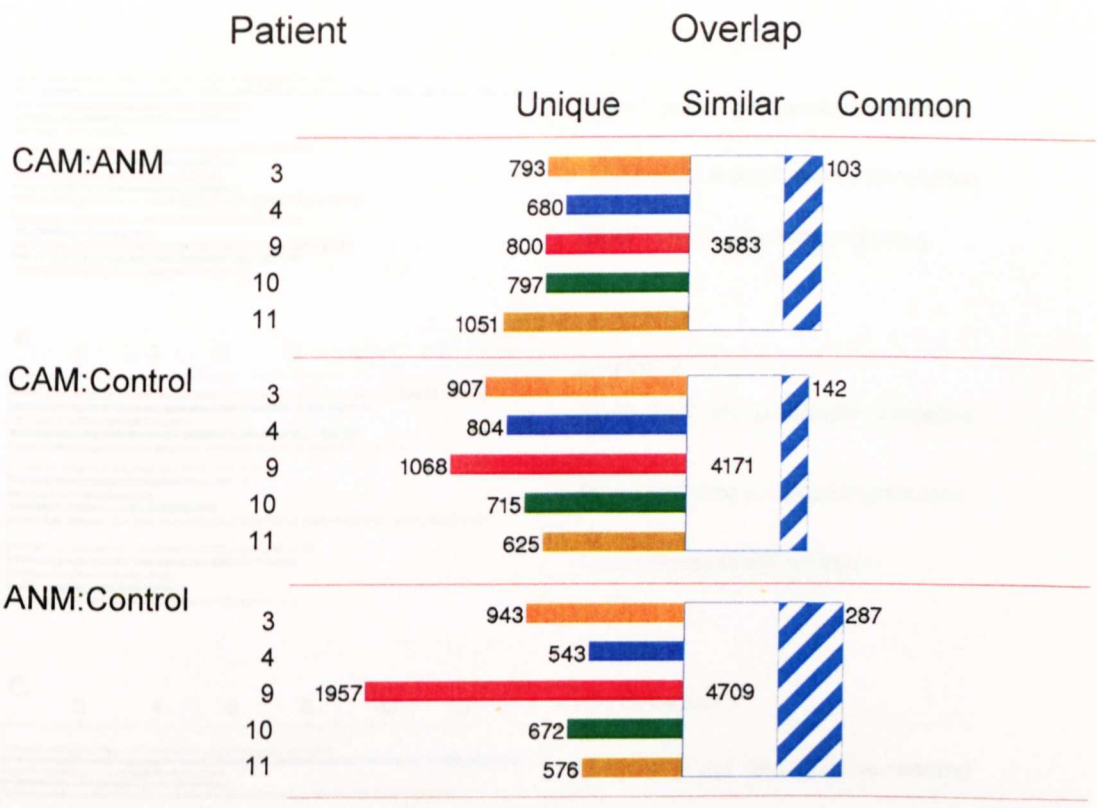


Figure 4.7

Overlap of genes with >1.5 fold changes in individual paired analyses of AGS gene array data

Comparison of genes with >1.5 fold change in paired analyses of CAM vs. matched ANM CM treatment (CAM:ANM), CAM CM treatment vs. control treatment from respective batch (CAM:Cont) and ANM CM treatment vs. control treatment from respective batch (ANM:Cont). Coloured bars on the left indicate genes which changed >1.5 fold in a single experiment only (unique), the central blocks indicate genes with >1.5 fold change in some but not all experiments (similar), and the blue and white striped blocks are those genes which are changed >1.5 fold in all experiments within the group.

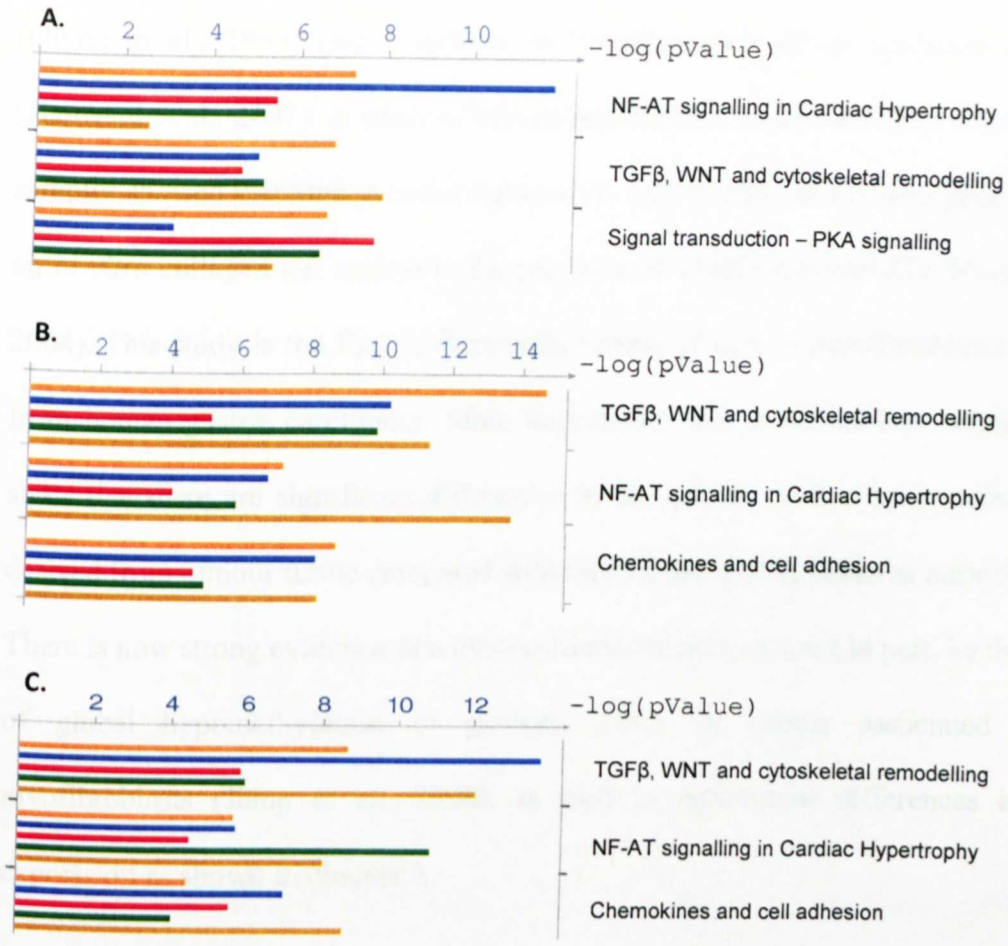


Figure 4.8

GeneGo maps most significantly enriched by differentially expressed genes from individually paired AGS gene expression arrays
Individually paired gene arrays were used to calculate fold changes between experiments and filtered for genes with >1.5 fold changes. The most significantly enriched GeneGo maps for pairs in the groups CAM:ANM, CAM:Control and ANM:Control are shown (A, B and C respectively). Coloured bars represent the –log(p-value) of enrichment for individual experiments, such that the longer the bar, the more significant the enrichment.

4.4 Discussion

The findings presented here highlight the potential for myofibroblasts to influence cancer cell behaviour and the potential to regulate gene expression. The myofibroblasts used in this study show similarities to myofibroblasts and CAMs from other studies. Work on human fibroblasts from normal and tumour prostate has shown that CAMs can increase the proliferation of prostate epithelial cancer cells

(Olumi et al., 1999) and contribute to transformation of an epithelial cell line (Hayward et al., 2001). A study of human myofibroblasts derived from colon cancer samples showed that human colon cancer cells only exhibit an invasive phenotype in an *in vitro* collagen gel system in the presence of myofibroblasts (De Wever et al., 2004). This study is the first to demonstrate these effects in myofibroblasts derived from human gastric carcinoma. More importantly the myofibroblasts studied here show that there are significant differences in the effects evoked by myofibroblasts derived from tumour tissue compared with those from nearby adjacent normal tissue. There is now strong evidence that these differences may, at least in part, be the result of global hypomethylation of genomic DNA in cancer associated gastric myofibroblasts (Jiang et al., 2008), as well as substantial differences in gene expression as shown in chapter 3.

The results of the AVONA analysis of gene array data offer some interesting insights, particularly into the general changes in gene expression in AGS cells caused by myofibroblast CM treatment, since both CAM and ANM CM evoked similar responses. The largest increase was in IGFBP5, whose main role is the binding of insulin-like growth factors to regulate their bioavailability. Interestingly, IGFBP5 has previously been shown to play an important role in preneoplastic changes in the stomach caused by *H. pylori* infection (Hemers et al., 2005; McCaig et al., 2006). In these studies it was shown that *H. pylori* infection causes an elevation in MMP-7 secretion from gastric epithelial cells, which in turn increases the release of both IGF-II and IGFBP5 in gastric myofibroblasts. The MMP-7 also cleaves IGFBP5, thereby releasing bioactive IGF-II which contributes to proliferation of both epithelial cells and myofibroblasts. It is therefore interesting that in this study it is factors secreted by myofibroblasts that are increasing

transcription of IGFBP5 in gastric cancer cells. This observation could possibly be explained by greater amounts of bioactive IGF in the CAM CM, since IGF-I has been shown to increase IGFBP-5 expression (Xin et al., 2004). Other notable increases in gene expression identified by ANOVA analysis include a number of metallothionein genes. These function as metal ion chelators, as well as having roles in cellular responses to oxidative stress and regulation of transcription. Furthermore, they have been identified as being upregulated in a number of cancers where they can have antiapoptotic, proliferative and angiogenic effects (Pedersen et al., 2009).

The gene expression profiles of CM treated AGS cells were also analysed separately, with the SFM treated AGS cells not included in the analyses. This allowed the CAM and ANM CM treated samples to be paired appropriately to reduce the effect of inter-patient variation. It also allowed for closer inspection of the differences between CAM and ANM CM treatments, which appeared to be less extreme than the differences between SFM treated AGS cells and any CM treatment. The differences in gene expression induced by CAM compared with ANM CM as determined by paired t-test suggest a role for TGF signalling (Table 4.3). There is certainly a strong precedent for TGFβ signalling in modulating the effects of stromal cells on cancer cells, although it seems that its role is both extremely complex and context dependent (Bierie and Moses, 2006; Yang and Moses, 2008). Furthermore, TGFβ has been shown to play a crucial role in re gulating the phenomenon of epithelial-to-mesenchymal transition (EMT) (Xu et al., 2009), and so could be involved in mediating the acquisition of the mesenchymal phenotype observed in the CM treated AGS cells. The wide ranging actions of TGFβ mean that it could be important in regulating all of the aspects of phenotypic change observed in this model system. A similar conclusion can be drawn from the results of the

individually paired experiments analysis (Figure 4.8), which strongly suggest a role for TGFβ and wnt signalling, leading to cytoskeletal remodelling. It has previously been shown that TGFβ signalling is impaired in some gastric cancers cell lines, including AGS cells (Ijichi et al., 2001). Similarly, it was shown that ELF, a β-spectrin involved in mediating transcriptional response to TGFβ signalling, is inactivated in some gastric cancers (Katuri et al., 2005). Despite this, levels of TGFβ has been shown to be increased in various cancers (Akhurst and Derynck, 2001; Bierie and Moses, 2006). In many cases it seems that disruption of the TGFβ response mechanisms of cancer cells impair its growth suppressor activities, whilst maintaining the resulting increases in EMT and invasive behaviour.

Compared with conventional statistical tests, these pathway-centric analyses allow for the possibility that in different experiments the same signalling pathway or pathways could be affected, but at different points and in different ways. The use of MetaCore to analyse the AGS gene expression data allowed a more detailed understanding of the mechanisms by which myofibroblast CM influences AGS cells. Care must be taken, however, in interpreting these analyses. The GeneGo maps, for instance, are by no means definitive in their assignment of function to a group of genes. The database of maps within MetaCore, though detailed, does not include every possible combination of cellular signalling pathway, but a large collection of well studied and annotated maps.

The data presented here indicate that secreted factors from myofibroblasts influence a number of attributes of cancer cells *in vitro* which could translate to aggressive behaviour *in vivo*, including migration, invasion and proliferation. Gene array analysis indicates that the induced changes in gene expression are extremely

complex, but suggests the involvement of TGFβ, growth factor and inflammatory mediator signalling in modulating these effects. Of these, TGFβ signalling could be particularly important due to its roles in both gastric cancer cells and in epithelial-stromal signalling. In order to dissect the signalling events occurring it will be necessary to explore the secretome of these myofibroblasts in detail.

4.5 Conclusions

1. AGS cells migrate, invade and proliferate to a greater extent in response to conditioned media from cancer derived myofibroblasts compared with conditioned media from adjacent normal derived myofibroblasts.
2. Myofibroblast conditioned media induce EMT-like changes in AGS cells.
3. Myofibroblast conditioned media induce a range of changes in gene expression in AGS cells, largely in genes involved in, or downstream of, TGFβ signalling, inflammation and adhesion.

Chapter 5

Proteomic analysis of myofibroblast secretomes and proteomes

5.1 Introduction

The secretome is the totality of proteins which are secreted by a given cell, and as such contains a wealth of information on intercellular communication. Recent efforts have employed various proteomic techniques in order to define the secretomes of cultured cells (Faca et al., 2008; Kulasingam and Diamandis, 2007; Lawlor et al., 2009). Proteomic characterisation of secretomes is emerging as an important tool in the study of diseases such as cancer, and has been used to gain insight into a number of cellular processes. These include the identification of secreted proteins which are regulated by p53 expression in a human glioma cell line (Khwaja et al., 2006) and by Smad4 in a colon cell line (Volmer et al., 2004). A recent study also analysed the effects of various stromal cell types on the secretome of a lung adenocarcinoma cell line in co-culture (Zhong et al., 2008).

There is very little known about the proteomic differences between CAMs and ANMs. It was shown in primary human myofibroblasts isolated from breast tumours that CAMs and ANMs have distinctive cellular proteomes which can be distinguished by image analysis of 2D-PAGE gels (Hawsawi et al., 2008), although the relevant proteins were not identified. It was also shown that p53/p21 DNA damage signalling is defective in CAMs but not ANMs. A recent study used primary cell cultures to compare the secretomes and proteomes of normal human skin fibroblasts, mouse melanoma xenograft associated fibroblasts and bone marrow fibroblasts from myeloma patients (Paulitschke et al., 2009). Whilst this study did not include quantitative proteomic measurements, or comparisons of patient-matched cell cultures, it was the first attempt to define the secretome of a cancer-associated stromal cell population.

Since myofibroblasts are a paracrine cell type with important roles in cancer development and progression (Bhowmick et al., 2004), it is important to define their secretome. This is the first time that myofibroblast secretomes have been comprehensively investigated, and the first time that the secretomes of cancer-associated and adjacent normal stromal cells have been quantitatively compared.

5.1.1 Aims

This chapter aims to define differences in protein abundances between CAMs and ANMs, with particular focus on the secretome. Specifically the objectives are to;

1. Evaluate methods for identification and quantitation of proteins and select the most appropriate proteomic approach.
2. Analyse the secretomes and cellular proteomes of CAMs and their respective ANMs.
3. Identify the proteins which are differentially abundant in CAMs and which may account for the functional differences between CAMs and ANMs.

5.2 Materials and Methods

5.2.1 Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE)

Gels were run as described in section 2.7.2, with 500μg concentrated media or cell extract. Excised protein spots were identified using a Voyager-DE PRO (Applied Biosystems) MALDI-TOF MS instrument and the online version of Mascot. A protein identification score of 67 or greater, which equates to $P<0.05$, was considered to be a positive identification (Protein score = $Log_{10}(P)$, where P = probability that observed match is a random effect).

5.2.2 SILAC labelling

Labelling by SILAC was performed as described in section 2.5.4. Myofibroblasts derived from patients 1 and 2 were labelled, the CAMs grown in the presence of heavy $^{13}C_6$ isotopic L-lysine and $^{13}C_6$, $^{15}N_4$ L-arginine, and the ANMs grown in the presence of light $^{12}C_6$ isoptopic L-lysine and $^{12}C_6$, $^{14}N_4$ L-arginine.

5.2.3 iTRAQ labelling

iTRAQ labelling was performed as described in section 2.5.5. Following labelling samples were pooled and cation exchange (CEX) purification of the sample was performed as described in Chapter 2. This was done using either a CEX mini-cartridge, which generated a single fraction, or by CEX HPLC, which was used to generate 27 fractions in the case of CM samples and 40 fractions in the case of cell extract samples. Myofibroblast proteomes and secretomes from all patients were analysed, with the exception of patient 5 due to insufficient cells.

5.2.4 Liquid chromatography-tandem mass spectrometry (LC-MS/MS)

All SILAC and iTRAQ samples were analysed using an Applied Biosystems QStar Pulsar mass spectrometer as described in section 2.5.8.

5.2.5 Proteomic data analysis

For SILAC and iTRAQ experiments, protein identification and quantitation calculations were performed using ProteinPilot 2.0 (Applied Biosystems). A 95% confidence threshold was set for positive protein identification. For COFRADIC experiments, protein identification was performed using an in-house verion of Mascot (Matrix Science). For methionyl-COFRADIC, only proteins which were identified by 2 or more unique peptides were considered to be a positive identification. Protein localisations were manually curated using the Uniprot database, or the Human Protein Resource Database (HPRD). Proteins which were not curated as secreted or membrane associated proteins were excluded from secretome data analyses. Proteins whose abundance in CAM was more than 1.3 fold different compared with their corresponding ANM were considered to be differentially abundant. This threshold was selected based on a previous experiment in our group which showed that when the same proteome sample is independently prepared and iTRAQ labelled four times, over 95% of the proteins identified have an associated fold change of less than 1.3 when any two of the sample preparations are compared (data unpublished). Pathway analysis was performed using MetaCore (GeneGo).

5.2.6 COFRADIC

Methionyl-COFRADIC and N-terminal COFRADIC were performed as described in sections 2.71 and 2.72 respectively. Myofibroblasts were first labelled using SILAC as described in section 2.5.4. Heavy isotopes of lysine and arginine were used to label the CAM samples, and light isotopes of lysine and arginine to label the ANM samples.

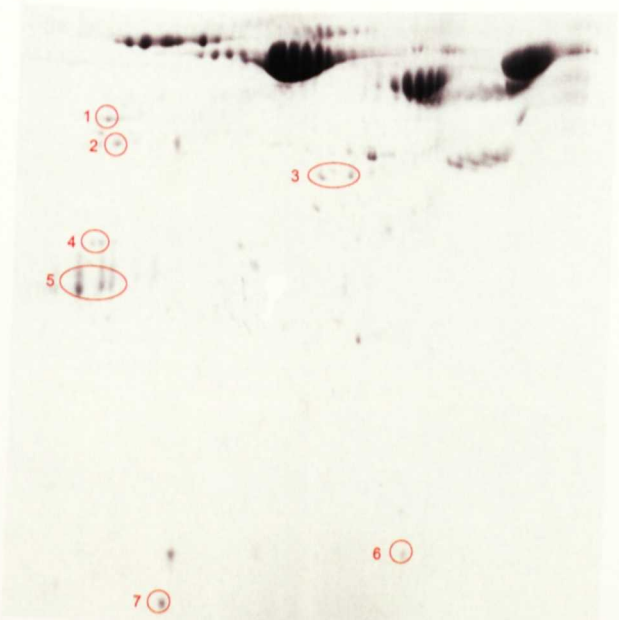
5.3 Results

5.3.1 2D-PAGE analysis of myofibroblast secretomes and proteomes

Media from CAM and ANM derived from patient 2 were run on two separate 2D-PAGE gels simultaneously. Excision of protein spots followed by MALDI-TOF identification of proteins successfully identified 7 secreted proteins (Figure 4.1). Of these, there were obvious differences in abundance of two; PAI-1 and TIMP1 being less abundant in the CAM secretome than the ANM. Spondin-2, IGFBP4 and TIMP1 appear as two or more spots, separated in the horizontal and with similar migration in the vertical axis. This indicates isoforms of the proteins which have the same apparent molecular weight, but different isoelectric points. This may be accounted for by the presence of post-translational modifications, such as sulphation or glycosylation.

A.

B.



Spot	Protein
1	MMP1
2	PAI-1
3	Spondin-2
4	IGFBP4
5	TIMP1
6	Galectin-1
7	Beta-2-microglobulin

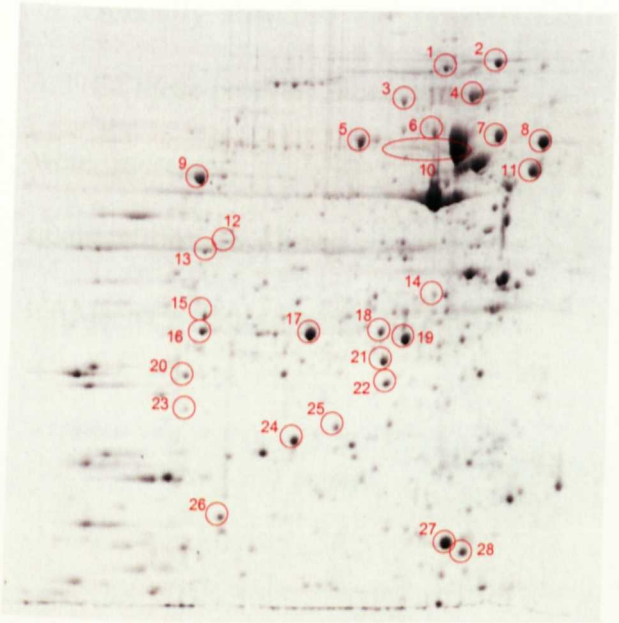
Figure 5.1

2D-PAGE analysis of myofibroblast secretomes
Two-dimensional gels of CAM and ANM CM (A. and B. respectively). Isoelectric focussing was used to separate proteins on the basis of their isoelectric points in the horizontal axis, and denaturing gel electrophoresis separated proteins in the vertical axis based on their size. Identified protein spots are highlighted on gel B., with protein identification shown in the adjacent table.

Analysis by 2D-PAGE of the cellular proteomes of one pair of CAM and ANM successfully identified 28 proteins (Figure 5.2). Of these, 8 were less abundant in CAM; calumenin, elfin, aldose reductase, phosphoglycerate mutase, hspb1, hspb6 and 40S ribosomal protein S12. Hspb1 appeared as two distinct spots (17 and 18), both of the same apparent molecular weight but with different isoelectric points, possibly indicating a post-translational modification. A similar effect was observed for vimentin, with a series of spots appearing in the CAM sample, and just a single spot in the ANM sample (10), again consistent with post-translational modification in CAM but not ANM. Only [Mn]-containing superoxide dismutase was identified as being more abundant in CAM than ANM.

A.

B.



Spot	Protein
1	Transitional endoplasmic reticulum ATPase
2	Endoplasmin
3	Stress-70 protein
4	78kDa glucose-regulated protein
5	Protein disulphide-isomerase A3
6	60kDa HSP
7	Protein disulphide-isomerase A1
8	Calreticulin
9	Calponin-2
10	Vimentin
11	Calumenin
12	Elfin
13	Aldose reductase
14	Cathepsin B
15	Phosphoglycerate mutase 1
16	Triosephosphate isomerase
17	HspB1
18	HspB1
19	Ubiquitin carboxy-terminal hydrolase isozyme L1
20	[Mn]-containing superoxide dismutase
21	Glutathione S-transferase P
22	Peroxiredoxin-2
23	Alpha crystallin B chain
24	HspB6
25	Nucleoside diphosphate kinase A
26	40S ribosomal protein S12
27	Galectin-1
28	Thioredoxin

Figure 5.2

2D-PAGE analysis of myofibroblast cellular proteomes

2D gels of CAM and ANM cell extracts (A and B respectively). Isoelectric focussing was used to separate proteins on the basis of their isoelectric points in the horizontal axis, and denaturing gel electrophoresis separated proteins in the vertical axis based on their size. Identified protein spots are highlighted on gel B, with protein identification shown in the adjacent table.

5.3.2 Quantitative proteomic analysis of the secretome by SILAC

In order to increase coverage and improve quantition, CAM and ANM CM from patients 1 and 2 were analysed using SILAC labelling and LC-MS/MS peptide identification. These samples were chosen for SILAC labelling as they were the first cell preparations to be generated. This approach increased secretome coverage from 7 proteins to 33 and 27 proteins in the two SILAC experiments (Table 5.1). The SILAC analysis also provided measures of relative quantitation which are shown here as fold changes in CAM relative to the respective ANM, with a negative value representing a decrease. Over half of the proteins identified in both patients were differentially abundant in CAM relative to ANM, with a fold change greater than 1.3. Of these proteins, 7 were decreased in both CAMs relative to their ANM, and 3 were increased in both CAMs relative to their ANM. All identifications and quantitations are shown in Table 5.2, which reveals decreases in collagens in the CAMs, as well as changes in proteins including IGFBPs and MMPs.

Patient	Secreted proteins	Increased in cancer	Decreased in cancer
1	33	7	13
2	27	6	13

Table 5.1

Summary of myofibroblast secretome coverage by SILAC
The total number of proteins identified and quantified in each experiment are shown, together with the number of proteins which are increased and decreased in CAM relative to ANM by more than 1.3 fold.

Protein name	Myofibroblast pair	
	1	2
Alpha-2-macroglobulin	-3.40	-1.98
Alpha-enolase	1.17	1.65
Annexin A2	1.33	1.40
Annexin A5	-1.27	
Basic proline-rich peptide IB-1	1.49	2.49
βig-h3	-1.31	-2.43
Cartilage oligomeric matrix protein	-3.48	
Cathepsin B		1.63
Clusterin		1.14
Collagen alpha-1(I) chain	-1.26	-1.63
Collagen alpha-1(III) chain	-4.61	-2.05
Collagen alpha-1(VI) chain		-1.15
Collagen alpha-2(I) chain	-1.26	-1.65
Complement C1r subcomponent		-1.36
Complement C3		-3.23
Decorin	-2.14	-2.43
Galectin-3-binding protein	1.02	-1.11
Gelsolin	1.00	
Glia-derived nexin	-2.11	-3.45
Heat shock cognate 71 kDa protein	1.50	1.27
Heat-shock protein beta-1	-1.12	1.06
IGFBP4	-1.01	-1.86
IGFBP5	3.13	-1.34
IGFBP6	1.09	
IGFBP7	1.38	-1.07
Inter-alpha-trypsin inhibitor heavy chain H2	-4.09	-2.67
Lumican	-1.57	-1.01
Microfibril-associated glycoprotein 4	-1.16	
MMP1	1.37	1.05
MMP10	-2.24	
MMP2	-1.21	1.13
MMP3	1.31	2.45
Moesin	1.24	1.31
Ryanodine receptor 2	-3.72	
SPARC		1.01
Spondin-2	1.01	
Sulfhydryl oxidase 1	-1.39	-1.44
TIMP1	-1.53	
Vitronectin	-2.21	

Table 5.2

Analysis of myofibroblast secretomes by SILAC
All of the identified and quantified secreted proteins from two SILAC experiments using myofibroblast CM. The change in relative protein abundances for each patient-derived pair of myofibroblasts is shown as fold changes in CAM relative to ANM. Blank boxes indicate that the protein was not identified in the corresponding experiment.

5.3.3 Improved secretome coverage using iTRAQ and cation exchange fractionation

In an attempt to further improve proteomic coverage a second isotopic tagging system, iTRAQ, was used to study both the secretomes and cellular proteomes of the myofibroblasts. The iTRAQ method requires a cation-exchange (CEX) step, which can be performed using either a CEX cartridge, or a CEX liquid chromatography (LC) column. Figure 5.3 shows the secretome coverage comparing CAM and ANM CMs from patient 2 using SILAC, iTRAQ with a CEX cartridge, and iTRAQ with a CEX column fractionation. iTRAQ with CEX fractionation gave a substantial increase in coverage compared to SILAC and iTRAQ with CEX cartridge, which gave similar coverages. Importantly only 1 of 27 proteins identified by SILAC were not identified by iTRAQ with column CEX (~4%), and 3 of 34 identified by iTRAQ/cartridge were not identified by iTRAQ/column (9%). Labelling using iTRAQ, followed by CEX fractionation was therefore selected as the method to analyse all of the remaining samples.

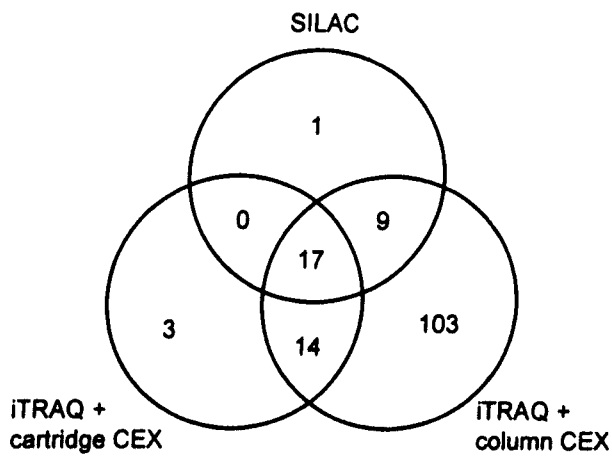


Figure 5.3

Venn diagram showing myofibroblast secretome coverage by three methods
The total number of secreted proteins identified and quantified by SILAC, iTRAQ with CEX cartridge sample cleanup and iTRAQ with CEX column fractionation in experiments comparing CAM and AN CM from a single pair of myofibroblasts.

The secretomes of 11 pairs of myofibroblasts were then analysed by iTRAQ, as summarised in Table 5.3. Of these secreted and membrane proteins, between 50-89% (mean 59%) of the proteins quantified differed by more than 1.3 fold. The results revealed that there is considerable heterogeneity between the pairs of myofibroblasts, with no single factor identified as being consistently up- or down-regulated in all of the pairs studied. Despite this heterogeneity, some proteins were differentially abundant more often than others. Table 5.4 shows the proteins which are increased (A) and decreased (B) in CAM by more than 1.3 fold in at least four pairs of samples. The variation between patients is exemplified by MMPs 1 and 3, which appear in both tables. These tables also reveal a greater number of decreases than increases in CAM relative to ANM. The proteins decreased in cancer include a number of protease inhibitors such as PAI-1 and glia-derived nexin, which are the main inhibitors of urokinase plasminogen activator (uPA), and TIMPs 1 and 2, which are the main inhibitors of MMPs.

Patient	Secreted proteins	Increased in cancer	Decreased in cancer
1	51	8	18
2	143	19	56
3	73	15	50
4	59	8	23
6	49	12	14
7	76	27	13
8	42	7	23
9	66	15	22
10	68	16	18
11	67	26	19
12	77	20	20

Table 5.3

Summary of myofibroblast secretome coverage by iTRAQ
The total number of secreted proteins identified and quantified in each experiment are shown, together with the number of proteins which are increased and decreased in CAM relative to ANM by more than 1.3 fold.

Secreted protein	Patient (fold change CAM vs ANM)												Total	
	1	2	3	4	6	7	8	9	10	11	12	Inc	Dec
Alpha-enolase	1.2	1.2	2.4	1.3		-1	2.9	1.4	1	-1	1.2	4	0
βig-h3	-1	-1	-2.6	1.2	-1	1.2	-2	1.4	1.1	1.3	-1	2	4
Chitinase-3-like protein 1		2.4	-2.6		9.2	1.7			2.7		4.7	5	1
Clusterin		-2	-3.6	-3	1.5		-2	-2		1.1	-1	1	5
Collagen alpha-1(I) chain	-2	-1	-8.8	1.4	-3	2.2	-1	-3	-1	-4	-1	2	6
Collagen alpha-1(III) chain	-4	-2	-11	1.1	-2	1.9		-2	-1	-4	-1	1	6
Collagen alpha-1(VI) chain	-1	-1	-3.5	-2	-2	1.5		-1	-1	1.3	-1	1	5
Collagen alpha-2(I) chain	-2	-1	-5.8	-1	-2	2	-2	-3	-1	-2	1.3	2	7
Complement C1r	-1	-2	-1.4	-2	-1	1.9	1.1	1	1.1	1.6	-1	2	4
Complement C1s	-1	-2	1.2	-2	-1	1.1	-2	1.1	1.1	1.7	-2	1	4
Cystatin-C		-1	-2.3	-2	1.2	1.5		-2	-2	-2	1.4	2	6
Decorin	-1	-2	-7.8	1	-2	-2	-3	-2	-2	2.1	1.6	2	8
Extracellular matrix protein 1	-1	-1	1	1.1			-1	1.6	-2	-1	1.2	1	4
Fibrillin-1	-2	-2	-2.1	-2		2.2		-2	1	-2	1.9	2	6
Fibronectin	-1	-1	-3.5	-2	-1	1.1	-2	-2	-1	-2	1.2	0	7
Fibulin-1 precursor.		-2	-3.1	-3	1.3	1.1		-2	-1	1.5	-1	2	4
Follistatin-related protein 1	-1	-2	-3	-1	-2	1.7	-1	-1	-1	-2	1.1	1	6
Galectin-1	1.1	1.1	3.8	1.5	1.5	-1	1.4	1.5	1	1.2	1.5	6	0

Glia-derived nexin	-3	-4	-13	-5	1.1	1.7		-5	-1	-2	3.6	2	7
IGFBP4	1.4	-2	-1.7	-2	-1		-3	1.3	-2	1.9	1.2	2	5
IGFBP5	1.8	-1	-4.8	1.3	-1	1.4	-2	-3	-1	3.8	-3	3	4
IGFBP7	1.1	-2	-1.9	-3	-1	1.3		1	1.6	-2	1.1	1	4
Lumican	-2	-1	-3.5	-1	-2	1.5		-2	-2	-2	1.3	1	7
MMP1	-2	-3	-1.2	-3	4	-3	-1	3.2	1.4	3	-2	4	6
MMP2	1.2	-2	-1.2	-3	1.1	1.2	-2	1.1	1	2.5	-1	1	4
MMP3	-1	-3	1.4	-3	5	-2	1.7	6.1	2.2	3.4	-2	6	4
Nucleobindin-1		-1	-2	-1	1.1			-1	-1	1.4	-1	1	4
PAI-1		-1	-3.9		-3	1			2.3	-4	-5	1	4
SPARC	-3	1.1	-4.9	-1	-1	2.1	1.5	-2	-1	-7	1.1	2	4
Spondin-2	1	2.1		1.2	1.4	-1	-1	-1	-2	2.3	1.8	4	1
Stanniocalcin-1	2.1	-2				-2	-1	1.7	1.4	2.3	1.6	5	2
Sulfhydryl oxidase 1	-1	-2	-6.4	-1	-2	1.5	-2	-2	-1	1.6	1.2	2	6
TIMP1	-1	-2	-9.6	-2	1.2	1.5	-2	-2	-1	-1	1.1	1	6
TIMP2	1	-2	-1.2	-2	-1	1.6	-2	-2	1	1.3	-1	1	4
Versican core protein		1.1	-3.5			2.1		-2	-2	-2		1	4

Table 5.4

Analysis of myofibroblast secretomes by iTRAQ
Secreted proteins which showed an increase or decrease of more than 1.3 fold in the secretomes of CAMs relative to their respective ANMs in at least 4 out of 11 iTRAQ-labelled pairs of samples. The change in relative protein abundances for each patient-derived pair of myofibroblasts is shown as fold changes in CAM relative to ANM.

5.3.4 Myofibroblast cellular proteome analysis by iTRAQ

The cellular proteomes of the myofibroblast pairs were studied using the same iTRAQ approach as for their secretomes. These experiments achieved a high coverage of the cellular proteomes which is summarised in Table 5.5. This table also shows that the number of proteins which were differentially abundant were proportionally less than observed in the secretomes; overall 7-40% (mean 21%) of the cellular proteins quantified differed by more than 1.3 fold, compared with 59% of the secreted proteins. The cellular proteins which were more than 1.3 fold different in four or more pairs of myofibroblasts are shown in Table 5.6. Of these, only two are proteins destined for secretion. These proteins; glia-derived nexin and collagen alpha-1, had already been identified as differentially represented in the secretome. There was, however, a positive correlation between the number of proteins in the secretome with a fold change greater than 1.3 and the number of proteins in the cellular proteome with a fold change greater than 1.3 for each myofibroblast pair (Figure 5.4).

Patient	Proteins quantified	Increased in cancer	Decreased in cancer
1	747	25	28
2	1260	110	114
3	669	133	120
4	902	59	73
6	770	79	52
7	753	40	47
8	518	95	113
9	547	75	85
10	791	51	33
11	869	121	117
12	919	60	83

Table 5.5

Summary of myofibroblast cellular proteome coverage by iTRAQ
The total number of proteins identified and quantified in each experiment are shown, together with the number of proteins which are increased and decreased in CAM relative to ANM by more than 1.3 fold.

Protein name	Patient (Fold change CAM vs ANM)											Total	
	1	2	3	4	6	7	8	9	10	11	12	Inc	Dec
6-phosphogluconate dehydrogenase	1.2	-1.2	1.4	-1.4	1.4	-1.2	-1.2	1.3	1	1.9	-1.4	4	2
Alcohol dehydrogenase 1B	-1.5	-1.9			-2.3	-1.2		-2.4	1	1.5	1.7	2	4
Aldose reductase	1	1	1.4	-3.3	3	-1.3	-1.8	2.3	1.4	2.9	-1.7	5	4
Alpha-crystallin B chain	-1.2	1.3	-2	-1.8	1.2	-1.9		-1.8	-1	-1.3	1.8	2	5
Beta-hexosaminidase subunit beta	-1.2	1.1	1.6	1.8	1.3	1.4	-2.5	-1.2	-1.1	1.1	1.5	5	1
Caldesmon	-1.1	1.2	-1.4	1.1	1.1	-1	-2.6	-1.5	-1.1	-1.7	1	0	4
CD166 antigen	-1.1	1.3	1.1	1.4	1.6	-1.1			1.2		1.4	4	0
Collagen alpha-1(I) chain	-1.3	2.7	-2.3	1.3	-2.1	1.1		-1	-1.3	-3.6	1.3	2	4
Collagen-binding protein 2 precursor	-1.1	1.4	-2.8	1.1	-1.5	1.2	-1.3	-1.4	1	-1.7	1.1	1	4
Cysteine and glycine-rich protein 1	-1.4	1.5	-1.2	1.8	-1.1	1.1	-3	-1.4	1.2	-1.4	-1.1	2	4
Cytoplasmic aconitate hydratase	1.3	-1.6	1.5		1.4	-1	-1.6		1.4	2	-1.3	5	2
D-3-phosphoglycerate dehydrogenase	-1.5	1.4	-6.5	1	-1.8	-1.4			-1.2	-2.4	1.7	2	5
Epoxide hydrolase 1	1.2	-1.2	2.3	-1.5	1.5	-1.3	-2.1	1.5	1	1.8	-1.5	4	3
Erythrocyte band 7 integral membrane	-1.1	-1.8	3.6	-2.5	1.3	-2.1	-1.8	1.7	1.2	3.2	-1.6	4	5
FKBP1A	1.3	-1.4			1.5	-1.5		1.3	2	1.3	-1.2	5	2
Four and a half LIM domains protein 1	-1.6	-1.5	1.6	-1.4	1.3	-1.8		1.1	1.5	-1.3	-1.4	2	5
Glia-derived nexin	-1.5	-1.2	-2.4	-2.4	2.1	1.4		-1.6	1.1	-2.5	2.4	3	5
Heat shock protein beta-1	-1.1	-1	-1.9	-1.4	1.4	-1.2	-1.4	-1.6	-1	1.1	-1	1	4
Heat shock protein beta-6	-1.5	1.1	-2.6	1.5	-1.8	1.2	-1.6	-1.9	-1.9	-2.1	1.8	2	7
HLA A-3 alpha chain		1.6	1.6	-1.1	1.1			2.2	1.4		-1.3	4	1
Long-chain-fatty-acid--CoA ligase 4	1	-1.5		-1.4	1.2	-1.5	-1.6		-1	2.4	1.1	1	4
Myosin light polypeptide 6	-1.3	1.2	-1.5	1.1	1	1	-1.5	-2	1	-2.7	-1.2	0	5
Myosin regulatory light polypeptide 9	-1.5	1.5	-1.8	-1.1	1	1.2		-2.7	1	-2.9	-1.4	1	5
Myosin-10	-1.5	1.2	-1.8	-1	-1	1.1		-1.8	-1.1	-2.7	-1.2	0	4
Myristoylated alanine-rich C-kinase substrate	-1.1	1.4	-1.4	1.4	-2	1.4	-1.1	-1.3	-1.3	-1.9	1.3	3	5
NAD(P)H dehydrogenase [quinone] 1	1.3	-1.3		-1.8	1.3	-1.3	-1.2	1.4	1.1	1.8	-1.5	3	4
NAMPT	1.1	-1.2	-1	-2.1	1.7	-1.3	1.1	1.7	1.4	1.8	-1.5	4	2
Niemann-Pick C1 protein	-1	-1	1.5	-1.1	1.5	-1.4		1.6		1.6		4	1

Phospholipase D3		1	2.4		1.7			1.3	1.9			4	0
Prostaglandin E synthase	-1	-1.8	-1.1	-2.7	2.1	-1.5	-2.1	1.3	1.1	1.3	1.1	2	4
Retinal dehydrogenase 1	-1.1	-1.6	1.7	-4.2	3.6	-1.8	-1.6	-1.3	1.6	2.5	-1.5	4	5
Seprase	-1.1	1.2	-3.1	1.4	1.3	1.3	-1.8	-1.3	1.1	-1.4	1	2	4
SH3 domain-binding glutamic acid-rich-like 3	-1	-1.3	-2.1	-1.7	1.1	1		-1.7	-1.1	-1.3	1.5	1	5
STOM	-1.1	-1.8	3.6	-2.5	1.3	-2.1	-1.8	1.7	1.2	3.2	-1.6	4	5
Superoxide dismutase [Mn], mitochondrial	1.1	-1.3	-1	-3.9	1.5	-1.2	-1.6	3	1.8	1.7	-2.2	4	4
Syntenin-1	1.2	-1.1		-1.5	-1.3	-1.5		1.6	1.2	1.6	-1.3	2	4
TGM2	1.3	1.1	3.3			1.4		1.5		-1.2	-1.4	4	1
Thioredoxin	1.3	-1.5	1.2	-1.5	1.4	-1.4	-1.9	1.5	1.1	1.5	-1.2	3	4
Transgelin	-1	2	-3.3	1.6	1	1.3	-1.7	-2.1	1	-2.1	-2	2	5
Tropomyosin alpha-1 chain	-1.3	1.7	-4	1.1	-1.4	-1.1		-2.2	-1.3	-2.7	-1.3	1	5
UDP-glucose 6-dehydrogenase	-1	1	-2.9	-1.3	1	1	-1.8	-1.6	-1.1	-1.1	1.3	0	4
Vimentin	1.6	-3.1	1.3	2	-1.3	1.3	-1.1	-1.2	-1.5	1.3	-1.9	5	3

Table 5.6

Analysis of myofibroblast cellular proteomes by iTRAQ

Myofibroblast cellular proteins which showed an increase or decrease of more than 1.3 fold in the cellular proteomes of CAMs relative to their respective ANMs in at least 4 out of 11 iTRAQ experiments. The change in relative protein abundances for each patient-derived pair of myofibroblasts is shown as fold changes in CAM relative to ANM.

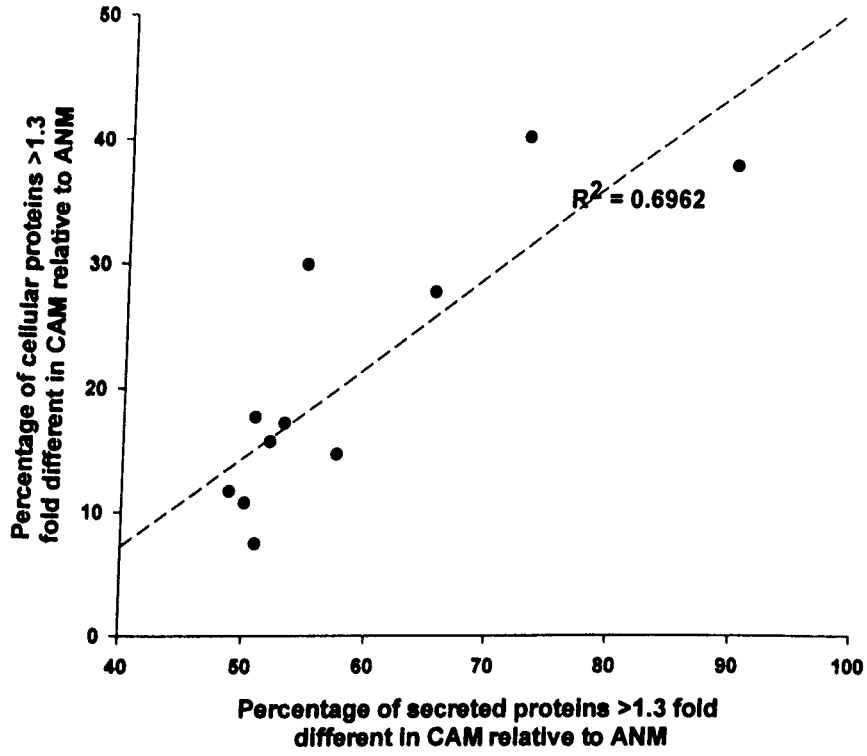


Figure 5.4

Correlation of differentially abundant proteins in secretomes and proteomes
The number of proteins in myofibroblast secretomes whose relative abundance in CAM was more than 1.3 fold different to their corresponding ANM correlate positively with the number of cellular proteins from the same myofibroblast pair with a fold change greater than 1.3. Coefficient of determination (R^2) = 0.6962.

5.3.5 Correlation of proteomes and secretomes with dynamic scores

Since a positive correlation was observed between patient's dynamic scores and the difference in gene expression between their CAMs and ANMs in Chapter 3, similar comparisons were made using the secretome and proteome data. Positive correlations were observed between dynamic scores and the number of proteins in the secretomes with an associated fold change greater than 1.3 (Figure 5.5), and between dynamic scores and the number of proteins in the cellular proteomes with an associated fold change greater than 1.3 (Figure 5.6). This correlation was stronger when using the cellular proteome data than then when using the secretome data, with an R^2 value of 0.3276 compared with 0.1354 for the secretome data.

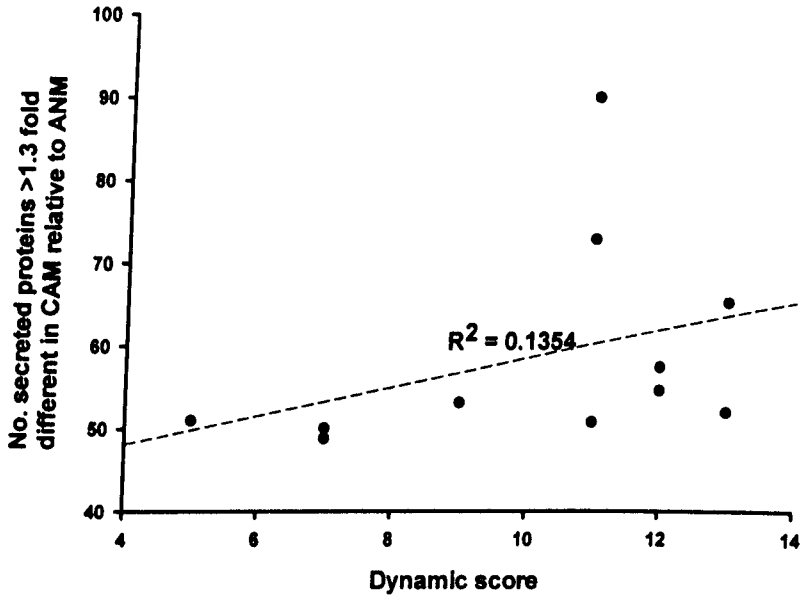


Figure 5.5

Correlation of differentially abundant proteins in secretomes and dynamic scores

The number of proteins in myofibroblast secretomes whose relative abundance in CAM was more than 1.3 fold different to their corresponding ANM correlate positively with the dynamic score of the patient. Coefficient of determination (R^2) = 0.1354.

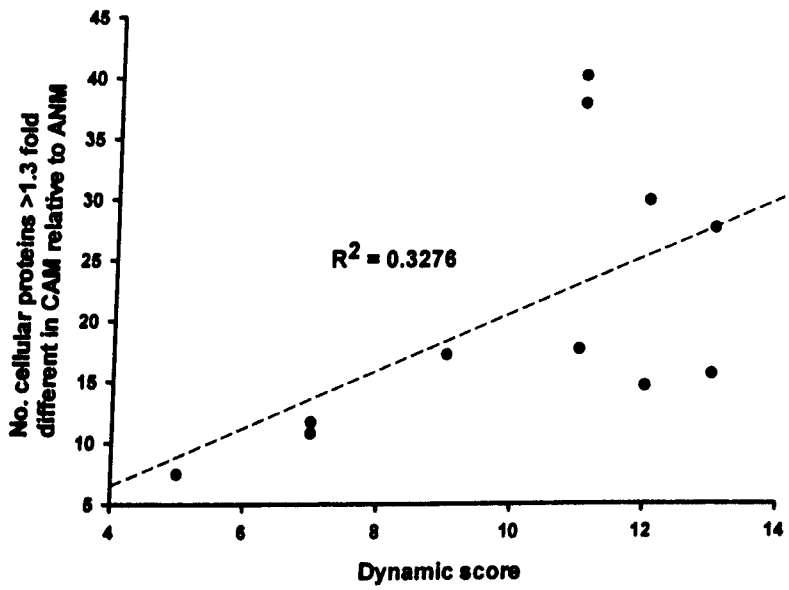


Figure 5.6

Correlation of differentially abundant proteins in cellular proteomes and dynamic scores

The number of proteins in myofibroblast cellular proteomes whose relative abundance in CAM was more than 1.3 fold different to their corresponding ANM correlate positively with the dynamic score of the patient. (R^2) = 0.3276.

5.3.6 Pathway anaylsis of iTRAQ secretome data by MetaCore

In order to further analyse the data generated by iTRAQ and to gain better insight into the functional networks involved, the results of the iTRAQ secretome analysis were analysed using MetaCore. A threshold of 1.3 was first applied so that only the proteins which were differentially abundant in CAM were used. An interactome analysis was next performed using the 'Interactions by protein function' tool. This uses MetaCore's database of curated protein-protein interactions to identify proteins which have a high number of interacting proteins within the dataset. The analysis focussed on receptors, since this is the class of proteins most likely to mediate the effects of myofibroblast secreted proteins on other cells. The receptors with the highest number of interacting proteins in the dataset are shown in Table 5.7.

In order to see which of the proteins in the dataset were reported to interact with these receptors, an interaction network was built using MetaCore. The four receptors with the highest numbers of interactions were used. In order to produce the most reliable network possible, only those interactions which have been confirmed in the literature were used. The resulting network using these four receptors and the secretome dataset of proteins with a fold change greater than 1.3 is shown in Figure 5.7. Several of the proteins from the iTRAQ dataset used in the network had interactions with more than one of the receptors, and with other proteins from the dataset. Table 5.8 summarises the changes in CAM CM of the secreted proteins interacting with these receptors in an activating or inhibitory fashion. These figures suggest a trend for increase in A2M receptor activating proteins, and a decrease in EGF receptor inhibitory proteins.

Object name	r	n	R	N	mean	z-score	p-value
Alpha-2 macroglobulin receptor	19	169	100	19836	0.85	19.80	1.24E-20
Amyloid precursor	19	169	196	19836	1.67	13.54	5.57E-15
Integrin beta-1	17	169	161	19836	1.37	13.46	3.88E-14
EGF receptor	15	169	361	19836	3.08	6.89	4.98E-07
LDL-related protein2 (Megalin)	11	169	64	19836	0.55	14.24	6.23E-12
CD44	8	169	101	19836	0.86	7.75	2.42E-06
Integrin alpha-2	7	169	32	19836	0.27	12.95	8.09E-09
Galectin-3-binding protein	6	169	27	19836	0.23	12.09	8.93E-08
Interleukin-8 receptor beta	6	169	29	19836	0.25	11.63	1.41E-07
uPA receptor	6	169	81	19836	0.69	6.43	6.7E-05
Myelin-associated glycoprotein	5	169	30	19836	0.26	9.43	5.07E-06
Integrin beta-3	5	169	82	19836	0.70	5.18	0.00068
Cubilin	4	169	13	19836	0.11	11.74	3.42E-06
VLDL receptor	4	169	16	19836	0.14	10.51	8.54E-06
Interleukin-8 receptor alpha	4	169	22	19836	0.19	8.85	3.3E-05
Integrin alpha-3	4	169	26	19836	0.22	8.07	6.57E-05
Integrin alpha-4	4	169	28	19836	0.24	7.74	8.87E-05
CD36	4	169	31	19836	0.26	7.31	0.000134
FK506 binding protein 12	4	169	32	19836	0.27	7.17	0.000152
E-selectin	4	169	34	19836	0.29	6.93	0.000193
Integrin beta-5	4	169	38	19836	0.32	6.49	0.000299
L-selectin	4	169	38	19836	0.32	6.49	0.000299
Integrin alpha-5	4	169	52	19836	0.44	5.37	0.001001
Dystroglycan	4	169	57	19836	0.49	5.07	0.001413
ErbB4	4	169	60	19836	0.51	4.91	0.00171
TGF-beta receptor type II	4	169	80	19836	0.68	4.04	0.004865
Integrin beta-2	4	169	82	19836	0.70	3.97	0.00531

Key	
Object name	network object name in MetaCore
r	number of network objects in the dataset which interact with the chosen object
n	Total number of network objects in the dataset
R	number of network objects in the complete database which interact with the chosen object
N	total number of gene-based objects in the complete database
mean	mean value for hypergeometric distribution (n*R/N)
z-score	z-score ((r-mean)/sqrt(variance))
p-value	probability to have the given value of r or higher

Table 5.7

MetaCore 'Interactions by protein function' analysis
The table shows the receptors identified by MetaCore as having the highest number of interacting proteins within the iTRAQ secretome dataset.

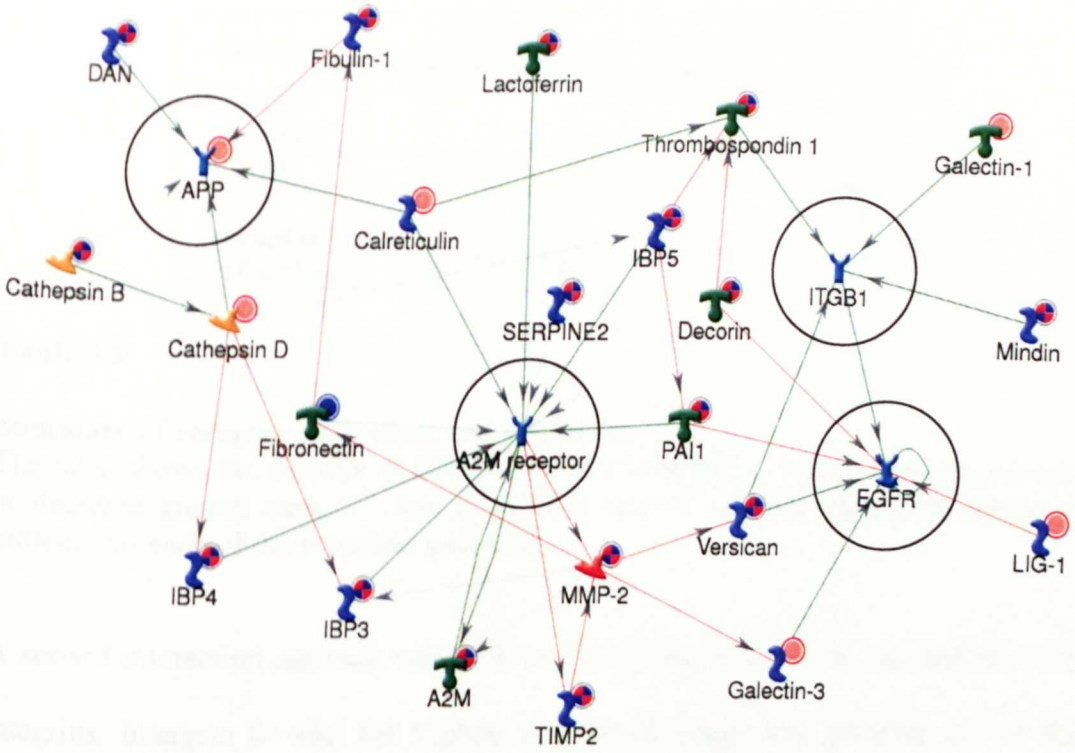


Figure 5.7

Interaction network of iTRAQ secretome data with A2M receptor, EGF receptor, integrin β1 and APP.
Proteins are shown as symbols representing their particular protein class (see MetaCore network key in supplement). The four receptors used to build the network are circled. Interactions between proteins are shown as arrows, with green representing an activating interaction, and red representing an inhibitory interaction. The arrow heads indicate the direction of the interaction. The blue and red circles in the upper right of the proteins represent fold change data from the iTRAQ results. Red and blue circles represent increases and decreases respectively, and a circle with both red and blue indicates that the protein has both increases and decreases in different experiments.

Receptor	Activating		Inhibitory	
	Increase	Decrease	Increase	Decrease
A2M receptor	9	6	-	-
APP	3	2	2	4
EGF receptor	4	4	4	10
ITGB1	10	10	-	-

Table 5.8

Summary of receptor activation and inhibition
The table shows the number of CAMs, out of a total of 11, which show an increase or decrease greater than 1.3 fold in at least one of the activating and inhibitory proteins for each of the receptors analysed.

A second interaction network was built in MetaCore to focus on interactions with integrins. Intergrin β1 had the highest number of interacting proteins out of the integrins, and was used in the previous MetaCore network. There were, however, a further seven integrins identified by the 'Interactions by protein class' tool, giving a total of eight integrins out of the 27 receptors in this analysis. All eight integrins were used to build a network in MetaCore showing the interactions with proteins from the dataset (Figure 5.8). This revealed a complex interaction web of proteins from the dataset which have curated direct and indirect interactions with integrins.

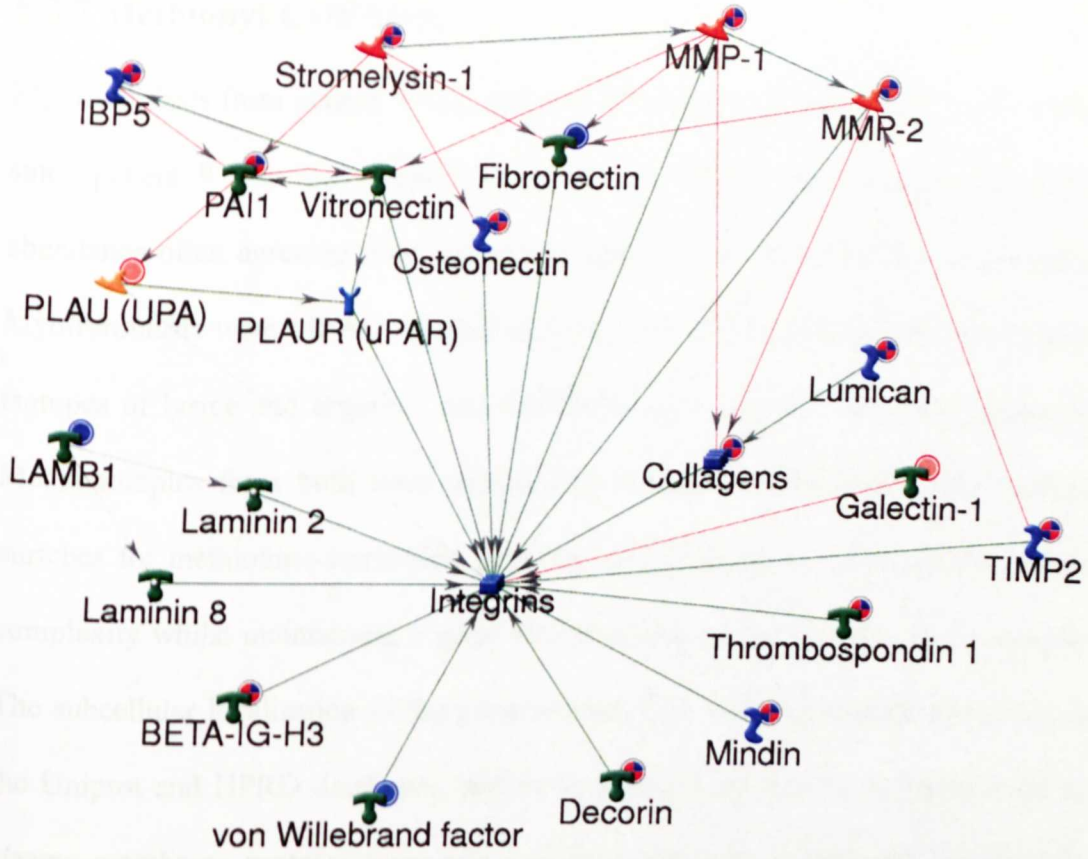


Figure 5.8

Interaction network of iTRAQ secretome data with integrins.
Proteins are shown as symbols representing their particular protein class (see MetaCore network key in supplement). The four receptors used to build the network are circled. Interactions between proteins are shown as arrows, with green representing an activating interaction, and red representing an inhibitory interaction. The arrow heads indicate the direction of the interaction. The blue and red circles in the upper right of the proteins represent fold change data from the iTRAQ results. Red and blue circles represent increases and decreases respectively, and a circle with both red and blue indicates that the protein has both increases and decreases in different experiments. For ease of presentation the integrins have been grouped together as a single object, as have collagens.

5.3.7 Methionyl-COFADIC

Myofibroblasts from patient 9 were selected for this initial study using COFRADIC since patient 9 was fairly representative of the cohort, with changes in protein abundance often agreeing with the general trend of all of the myofibroblast pairs. Myofibroblasts were SILAC labelled such that the CAM were labelled with heavy isotopes of lysine and arginine, and the ANM were labelled with light isoptopes. Media samples from both were analysed by methionyl-COFADIC. This method enriches for methionine-containing peptides, which effectively reduces the sample complexity whilst maintaining a good representation of the proteins in the sample. The subcellular localisation of the proteins identified was determined by searching the Uniprot and HPRD databases, and those proteins not known to be secreted or plasma membrane proteins were removed from subsequent analysis, leaving 175 proteins. An estimation of relative protein abundance was made by calculating the mean ratios of each peptide identified for each protein (Table 5.9). Of these, 27 were over 1.3 fold more abundant in CAM than ANM, and 103 were over 1.3 fold less abundant. This compares with the 66 secreted and membrane proteins identified by the iTRAQ for patient 9, which showed 15 proteins which were more abundant in CAM than ANM, and 22 which were less abundant. Secretome coverage for patient 9 was therefore 265% greater by Met-COFRADIC than by iTRAQ. Many of the biggest increases in abundance identified by Met-COFRADIC are proteases, including MMPs 1, 3 and 10. Of these, MMP1 and MMP3 were also identified by iTRAQ, and in both cases an increase in CAM relative to ANM was shown, which is in agreement with the Met-COFRADIC data. Many of the proteins which were less abundant in CAM than ANM media were ECM proteins, including collagens, laminins and fibronectin.

Protein	CAM:ANM
45 kDa calcium-binding protein	-1.8
5'-nucleotidase	-1.0
Adenosine deaminase	-17.3
Adenylyl cyclase-associated protein 1	1.0
Adenylyl cyclase-associated protein 2	-1.0
Adipocyte enhancer-binding protein 1	-5.8
ADM	1.4
Afadin	1.5
Alpha-2-macroglobulin	1.2
Alpha-type platelet-derived growth factor receptor	-1.0
Aminopeptidase N	1.2
Amyloid beta A4 protein	1.1
Amyloid-like protein 2	1.7
Angiopoietin-related protein 2	1.3
Annexin A1	-1.4
Annexin A11	-1.2
Annexin A4	1.0
Annexin A5	-1.2
Annexin A6	1.4
Annexin A7	-1.8
Apolipoprotein E	-2.6
Attractin	-1.6
Basement membrane-specific heparan sulfate proteoglycan core protein	-3.3
Basigin	-1.1
Beta-2-microglobulin	1.1
Biotinidase	-1.1
Brain acid soluble protein 1	-1.3
Cadherin-2	-3.1
Calpain small subunit 1	1.2
Calpain-1 catalytic subunit	1.0
Calsyntenin-1	-1.3
Calumenin	-6.0
Cartilage oligomeric matrix protein	-59.1
Catenin alpha-1	-1.4
Catenin beta-1	-1.2
Catenin delta-1	-1.4
Cathepsin B	-1.4
Cathepsin D	-1.9
Cathepsin F	-1.5
Cathepsin K	-1.1
Cathepsin L1	-2.2
Cathepsin Z	-1.2
CD109 antigen	1.4
CD166 antigen	1.8
CD44 antigen	1.4
CD63 antigen	1.0
Chloride intracellular channel protein 4	-3.2
Clathrin heavy chain 1	-1.7
Clathrin heavy chain 2	-1.3
Clusterin	-3.4
Coiled-coil domain-containing protein 80	-4.6

Collagen alpha-1(I) chain	-12.1
Collagen alpha-1(III) chain	-3.0
Collagen alpha-1(VI) chain	-1.7
Collagen alpha-1(VIII) chain	1.0
Collagen alpha-1(XII) chain	-1.3
Collagen alpha-1(XIV) chain	-2.2
Collagen alpha-2(I) chain	-4.3
Collagen alpha-2(IV) chain	-1.9
Collagen alpha-2(V) chain	-2.1
Collagen alpha-2(VI) chain	-1.7
Collagen alpha-3(VI) chain	-2.1
Complement C1q tumor necrosis factor-related protein 5	-2.6
Complement C1r subcomponent	-1.5
Complement C1s subcomponent	-1.4
Complement C3	-2.8
C-type lectin domain family 11 member A	-1.3
Cystatin-C	-3.3
Decorin	-1.5
Dickkopf-related protein 3	1.4
Dipeptidyl peptidase 4	-1.1
EGF-containing fibulin-like extracellular matrix protein 1	2.7
EGF-containing fibulin-like extracellular matrix protein 2	-1.3
EGF-like repeat and discoidin I-like domain-containing protein 3	1.1
EH domain-binding protein 1	-2.7
EH domain-containing protein 1	-2.0
Epididymis-specific alpha-mannosidase	-1.7
Extracellular matrix protein 1	-1.4
Fibrillin-1	-2.9
Fibrillin-2	3.0
Fibronectin	-8.6
Fibulin-1	-8.9
Fibulin-2	-8.2
Fibulin-5	1.2
Follistatin-related protein 1	-2.3
Galectin-1	-1.0
Galectin-3-binding protein	-2.6
Gelsolin	-2.3
Glia-derived nexin	-4.8
Glutathione peroxidase 3	8.6
Glypican-1	1.0
Hepatocyte growth factor	1.0
Insulin-like growth factor-binding protein 4	1.0
Insulin-like growth factor-binding protein 5	-3.0
Insulin-like growth factor-binding protein 6	-10.7
Insulin-like growth factor-binding protein 7	-1.6
Integrin alpha-V	2.8
Integrin beta-1	-1.1
Interleukin-6 receptor subunit beta	-1.7
Keratinocyte growth factor	1.7
Lactadherin	-4.2
Laminin subunit alpha-2	-3.3
Laminin subunit alpha-3	2.1

Laminin subunit alpha-4	-1.7
Laminin subunit alpha-5	1.6
Laminin subunit beta-1	-1.7
Laminin subunit gamma-1	-1.8
Latent-transforming growth factor beta-binding protein 2	-9.4
Leucyl-cystinyl aminopeptidase	-1.2
Lin-7 homolog C	-1.2
Liprin-beta-1	1.2
Lumican	-8.5
Lysyl oxidase homolog 1	-2.6
Lysyl oxidase homolog 2	-2.1
Macrophage colony-stimulating factor 1	1.6
Matrix metalloproteinase-14	-1.1
Matrix-remodeling-associated protein 7	1.2
Metalloproteinase inhibitor	-3.1
MMP1	7.8
MMP10	18.1
MMP2	1.0
MMP3	36.9
Multiple epidermal growth factor-like domains 8	-3.2
Myoferlin	1.3
Neogenin	1.8
Neural cell adhesion molecule L1-like protein	13.9
Neurogenic locus notch homolog protein 2	-1.4
Nidogen-1	1.1
Nidogen-2	-5.1
Nucleobindin-2	-2.4
Olfactomedin-like protein 3	-3.7
Osteopetrosis-associated transmembrane protein 1	1.1
Pappalysin-1	-2.2
Pentraxin-related protein PTX3	-11.2
Periostin	-5.3
Phosphatidylinositol-binding clathrin assembly protein	1.1
Phospholipid transfer protein	-2.9
Plasma glutamate carboxypeptidase	-1.6
Plasma protease C1 inhibitor	-21.8
Plasminogen activator inhibitor 1	-5.6
Poliovirus receptor	-1.5
Prolow-density lipoprotein receptor-related protein 1	1.2
Protein disulfide-isomerase	-1.7
Protein FAM3C	-3.5
Protein-lysine 6-oxidase	-4.7
Protocadherin Fat 1	1.1
Protocadherin gamma-A12	-1.5
Putative annexin A2-like protein	-1.5
Receptor-type tyrosine-protein phosphatase kappa	-1.3
Receptor-type tyrosine-protein phosphatase S	1.6
Retinoid-inducible serine carboxypeptidase	-1.8
Semaphorin-3C	-4.9
Semaphorin-7A	-13.6
Seprase	-2.0
Sodium/potassium-transporting ATPase subunit alpha-1	1.2

SPARC	-9.2
Spondin-2	1.1
Stanniocalcin-1	4.2
Sulfhydryl oxidase 1	-2.4
Synaptosomal-associated protein 23	-1.2
Talin-1	-1.4
Talin-2	1.4
Tenascin	-5.2
Testican-1	-2.7
Testin	-7.2
Thrombospondin-1	-1.9
Thrombospondin-2	1.8
Transforming growth factor beta-1	-1.5
Transforming growth factor-beta-induced protein ig-h3	-1.1
Tumor necrosis factor receptor superfamily member 11B	1.2
UPF0027 protein C22orf28	-1.3
UPF0556 protein C19orf10	-1.6
Versican core protein	-5.1
Voltage-dependent calcium channel subunit alpha-2/delta-1	-1.1
Xylosyltransferase 1	-16.2

Table 5.9

Methionyl COFRADIC analysis of myofibroblast secretomes
Comparison of CAM and ANM secretomes from patient 9 myofibroblasts by methionyl-COFRADIC. Changes in relative abundance of proteins was estimated by the mean of the ratios for all peptides identified for an individual protein.

5.3.8 N-terminal COFRADIC

The SILAC labelled media samples from patient 9 CAM and ANM were also enriched for N-terminal peptides using N-terminal COFRADIC. N-terminal peptides were also tri-deutero acetylated so that they were easily distinguished from other peptides by mass spectrometry. In this way it was possible to identify the N-termini of proteins as well as sites where the proteins had been proteolytically processed, which would generate neo-N-termini. The results were filtered such that only peptides which were tri-deuteroacetylated and had a valid Mascot identification were used in subsequent analyses. These results were then filtered further to remove peptides which began at residues 1 or 2 of the protein (i.e. at the initiator methionine

residue, or immediately after it), since these peptides did not represent proteolytic cleavage sites (other than possible removal of the initiator methionine). For each of the remaining peptides the ratio of relative abundance between the CAM and ANM samples were manually validated by inspecting the spectra and calculating the area under the peaks of the heavy and light isotopes of each. These ratios were compared with the mean ratio for the protein calculated from the methionyl-COFRADIC data. This was important since the N-terminal COFRADIC data identified proteolytic cleavage sites within the proteins, but it was otherwise difficult to know whether differences in the abundance of these peptides between the CAM and ANM samples was due to a change in the amount of proteolytic processing, or a change in the overall abundance of the protein. Using the Uniprot database the protein domains which contained the cleavage sites were also identified. The results were then separated into peptides which represented an increase or decrease in proteolytic processing with an associated ratio which was more than 1.5 fold different to the mean methionyl-COFRADIC ratio for that protein (Table 5.10). In some cases it was impossible to calculate the relative difference as only the heavy isotope of the peptide was identified (singletons), indicating that the peptide was only present in the CAM sample. Peptides with a fold change of less than 1.5 fold compared with the mean ratio for the protein were also grouped together (Table 5.11), as were peptides which represented the removal of signal sequences (SS) (Table 5.12).

Description	Protein CAM:ANM (Met-COFRADIC)	Cleaved peptide Can:Adj	Peptide vs. Protein	Domain	start	end
Collagen alpha-2(I) chain	0.23	0.71	3.04	Main chain	959	978
Collagen alpha-1(III) chain	0.33	0.66	2.00	Main chain	688	695
		5.36	16.34	Main chain	688	695
		6.04	18.42	Main chain	688	695
Collagen alpha-2(VI) chain	0.59	Singleton		Main chain	772	784
Insulin-like growth factor binding protein 5	0.34	2.73	8.09	PTM domain	162	176
		0.28	-1.20	PTM domain	164	176
		1.14	3.39	PTM domain	164	176
		0.42	1.24	PTM domain	165	176
		0.39	1.17	PTM domain	165	176
		0.45	1.33	PTM domain	165	176
		0.45	1.32	PTM domain	165	176
Lysyl oxidase homolog 2	0.48	Singleton		SRCR1	52	59
		Singleton		SRCR1	52	59
Latent-transforming growth factor beta-binding protein, isoform 1S	0.45	13.34	29.33	Main chain	449	461
Macrophage colony-stimulating factor 1	1.57	1.84	1.17	SS removed	33	53
		1.76	1.12	SS removed	33	53
		1.64	1.04	SS removed	33	53
		1.71	1.08	SS removed	33	53
		0.68	-2.32	Main chain	42	53
Major prion protein	No data	3.28		Main chain	121	136
		3.66		Main chain	121	136
		1.80		Main chain	121	136
		1.66		Main chain	121	136
		Singleton		Main chain	138	148
		Singleton		Main chain	138	148
		7.07		Main chain	138	148
		7.78		Main chain	138	148
Matrix metalloproteinase 2	1.04	8.75	8.41	Propeptide region	80	98
		13.59	13.07	Propeptide region	80	98
Matrix metalloproteinase 3	36.91	100.00	2.71	Propeptide region	92	101
Protein-lysine 6-oxidase	0.21	0.32	1.52	Lysyl-oxidase like	297	313
		0.32	1.53	Lysyl-oxidase like	297	313
Pro-neuregulin-1, membrane-bound isoform	No data	Singleton		Main chain	229	242
Urokinase-type plasminogen activator	12.66	7.91	-1.60	Kringle domain	109	123
		8.33	-1.52	Kringle domain	109	123

Table 5.10

Changes in secreted protein degradation in CAM relative to ANM identified by N-terminal COFRADIC

Table shows peptides which were identified as novel N-termini generated by proteolytic processing by N-terminal COFRADIC. Mean protein ratios calculated by Met-COFRADIC. Peptides with a ratio more than 1.5 fold different to their protein ratio shown in red (more degradation in CAM) and blue (less degradation in CAM).

Description	Protein CAM:ANM (Met-COFRADIC)	Cleaved peptide CAM:ANM	Peptide vs. Protein	Domain	start	end
Cathepsin B	0.72	0.88	1.22	Heavy chain	165	182
Collagen alpha-1(VI) chain	0.38	0.54	1.40	Main chain	188	199
		0.53	1.39	Main chain	188	199
Complement factor H	0.21	0.20	-1.03	Sushi 3 domain	167	175
Decorin	0.67	0.77	1.16	SS removed	17	22
		0.75	1.12	LRR-T 4	200	206
		0.49	-1.35	LRR-T 5	256	272
		0.53	-1.25	LRR-T 5	256	272
		0.80	1.20	LRR-T 7	331	345
		0.76	1.14	LRR-T 8	351	359
		0.72	1.08	LRR-T 8	351	359
Extracellular matrix protein 1	0.72	0.79	1.10	SS removed	20	31
		0.82	1.14	SS removed	20	31
		0.62	-1.15	Main chain	172	185
Fibulin-1	0.11	0.12	1.09	SS removed	30	42
		0.12	1.03	SS removed	30	42
		0.12	1.06	SS removed	30	42
		0.10	-1.14	EGF-like 4	344	353
Fibulin-5	1.22	1.09	-1.12	EGF-like 5	272	284
		1.22	-1.00	EGF-like 5	272	284
Insulin-like growth factor binding protein 4	1.04	1.28	1.23	SS removed	22	37
		1.08	1.03	SS removed	22	37
		1.25	1.20	Main chain	150	163
		1.24	1.19	Main chain	150	163
		1.31	1.26	C-term	192	206
		1.10	1.06	C-term	192	206
		1.22	1.17	C-term	192	206
Laminin subunit gamma-1	0.55	0.53	-1.05	EGF like 2	348	359
Lumican	0.12	0.14	1.20	LRR-T 7	304	310
		0.14	1.21	LRR-T 7	304	310
Matrix metalloproteinase 14	0.95	0.87	-1.09	Main chain	314	330
Spondin-2	1.10	0.64	-1.71	SS removed	27	38
		0.82	-1.34	Near SS	29	38
		1.07	-1.02	Spondin domain	217	229
Stanniocalcin-2	No data	1.47		Main chain	138	154
Thrombospondin-2	1.85	1.47	-1.26	Main chain	307	315
		1.44	-1.29	Main chain	307	315
Tissue inhibitor of metalloproteinases 1	0.33	0.41	1.24	NTR region	123	136
		0.37	1.12	NTR region	130	136

Table 5.11

Proteolytic processing of myofibroblast secreted proteins identified by N-terminal COFRADIC

Table shows peptides which were identified as novel N-termini generated by proteolytic processing by N-terminal COFRADIC, but whose abundance does not differ from the overall protein abundance by more than 1.5 fold.

Description	Protein CAM:ANM (Met-COFRADIC)	Cleaved peptide CAM:ANM	Peptide vs. Protein	Domain	start	end
Beta ig-h3	0.94	0.73	-1.29	SS removed	24	38
		0.75	-1.26	SS removed	24	38
C-X-C motif chemokine 5	No data	Singleton		SS removed	37	45
		Singleton		SS removed	37	45
Dentin matrix protein 4	No data	1.25		Within SS	11	18
Frizzled-6	No data	0.80		SS removed	19	31
		0.89		SS removed	19	31
Growth/differentiation factor 15	No data	0.95		SS removed	30	37
		1.15		SS removed	30	37
Insulin-like growth factor II	No data	0.22		SS removed	25	48
Periostin	0.19	0.15	-1.28	SS removed	22	33
Prostaglandin-H2 D-isomerase	0.93	0.84	-1.11	SS removed	23	42
Protein disulfide-isomerase	0.59	0.60	1.01	SS removed	18	30
		0.61	1.03	SS removed	18	30
Protein disulfide-isomerase A3	No data	0.57		SS removed	25	38
		6.14		SS removed	25	38
		0.60		SS removed	25	38
Protein-Arginine rich, mutated in early stage tumours	No data	0.76		SS removed	22	36
		0.78		SS removed	22	36
Sulfhydryl oxidase 1	0.42	0.44	1.05	Near SS	33	51
Tissue inhibitor of metalloproteinases 2	No data	0.64		SS removed	27	46
		0.73		SS removed	27	46
		0.65		SS removed	27	46
		0.60		SS removed	27	46
		0.61		SS removed	27	46
		0.58		SS removed	27	46
		0.62		SS removed	27	46
		0.62		SS removed	27	46
		0.65		SS removed	27	46
		0.60		SS removed	27	46
		0.60		SS removed	27	46
		0.59		SS removed	28	46
		0.61		SS removed	28	46
		0.73		SS removed	30	46
		0.71		SS removed	30	46
		0.71		SS removed	30	46
Twisted gastrulation protein homolog 1	No data	2.68		SS removed	26	47
UPF0556 protein C19orf10	0.63	0.65	1.03	SS removed	32	43
		0.66	1.05	SS removed	32	43
Versican core protein	0.19	0.21	1.07	SS removed	21	33

Table 5.12

Proteolytic processing of signal sequences of myofibroblast secreted proteins identified by N-terminal COFRADIC

Table shows peptides which were identified as novel N-termini generated by proteolytic processing by N-terminal COFRADIC. All peptides in the table identify cleavage sites at the protein signal sequence (SS).

5.4 Discussion

5.4.1 Comparison of proteomic techniques

Selecting proteomics techniques capable of quantitatively comparing proteomes is a challenge, and despite numerous published comparisons of methods including those used here there is by no means a definitive choice (Wu et al., 2006). The use of 2D-PAGE has previously proved very valuable (Gorg et al., 2004; Hemers et al., 2005; McCaig et al., 2006), and yet here met with limited success, the main drawback being the low success rate in identifying proteins. Figures 5.1 and 5.2 show there are a large number of protein spots to which this method was unable to positively assign identifications. There are, however, benefits to using this approach. The ability to visualise the separation of proteins based on their isoelectric point and size allows the researcher to make inferences about the proteins which would be difficult using other proteomics methods. For example, in this study a number of proteins were identified which are represented by multiple spots which have similar or identical sizes, but different isoelectric points, possibly indicating post-translational modifications. A drawback of the visualisation of 2D gels is that proteins can potentially overlap on the gel, making it difficult to distinguish between proteins, particularly with complex samples.

SILAC and iTRAQ have several advantages over 2D-PAGE, the most significant being their ability to assign accurate measurements of relative abundance to peptides and proteins. They also require less starting sample material since 2D-PAGE relies on visualisation of proteins on a gel, while iTRAQ and SILAC methods identify proteins by sensitive LC-MS/MS. All of these factors are possible explanations for the improved coverage observed in this chapter by SILAC and iTRAQ over 2D-PAGE. The differences between SILAC and iTRAQ are smaller, though there are

still crucial factors which should be considered. Both methods utilise stable isotopic tagging to distinguish between proteins from different samples. SILAC uses tagged amino acids which are incorporated into proteins as they are synthesised (Ong et al., 2002). This means that there are no labelling steps once the protein samples have been collected, thereby reducing the number of processing steps between collection and LC-MS/MS analysis. This labelling can, however, be a time consuming process, particularly with slow growing cells. By contrast, iTRAQ uses amino-reactive tags which react with the free amino groups on peptides generated by trypsinisation of the samples (Wiese, 2007). One important advantage of these tags is that they are isobaric, meaning that they have identical masses. This means that during the first round of MS peptides will appear as a single peak, as opposed to the multiple peaks seen in a SILAC labelled sample. These peaks will also be larger since each of the pooled samples will contribute to it, making it more likely to be identified by the MS. It is not until the second round of MS that the tags which distinguish between the different samples are generated, with each of the isobaric tags producing an ion of different mass during fragmentation.

5.4.2 Technical improvements by COFRADIC

Combined fractional diagonal chromatography (COFRADIC) was used to try to further improve secretome analysis. Care must be taken when comparing the COFRADIC results with those generated by iTRAQ and SILAC, since the COFRADIC experiments used a different MS platform and different software. The 265% increase in secretome coverage is nonetheless impressive, and exceeds the largest coverage of any of the myofibroblast secretomes achieved by iTRAQ. It was also the only experiment to identify all of the growth factors commonly associated

with gastrointestinal myofibroblasts; hepatocyte growth factor, keratinocyte growth factor and transforming growth factor β1 (Powell et al., 1999). The greatest advantage of COFRADIC is the ability to identify post-translational modifications, in this case proteolytic processing. In this respect it exceeds 2D-PAGE, which can also be used to infer information about post-translational modifications, since COFRADIC offers greater coverage and identifies the cleavage site. Despite the greater time and equipment requirement of COFRADIC, it is therefore a valuable technique which is complementary to iTRAQ.

Recent studies have achieved significant coverage of secretomes by proteomic analysis of media samples from cultured cell lines (Faca et al., 2008; Kulasingam and Diamandis, 2007; Lawlor et al., 2009). In each case a large proportion of the proteins identified were intracellular or nuclear, with less than half of the proteins known to be extracellular of membrane-associated. Perhaps the largest coverage achieved to date was the identification of up to 1203 proteins in one cell line media sample, with roughly 20% being extracellular and 25% membrane proteins (Faca et al., 2008). The software and identification stringencies which were applied, however, varied greatly between these studies, as did the databases used to assign subcellular localisations. The coverage of secreted and membrane proteins achieved by methionyl-COFRADIC compares favourably with these studies, particularly since primary human cells were used, meaning that the amount of starting material was more limited.

N-terminal COFRADIC has previously been used to study the substrate degradomes of various proteases (Gevaert et al., 2003; Staes et al., 2008; Van Damme et al., 2005), as well as mechanisms of drug-induced cell death (Impens et al., 2008). In

each of these cases a single cell type was studied under different conditions, such as the presence or absence of a particular protease or drug. This study is the first to compare the substrate degradomes of two different populations of cells.

This chapter offers a unique comparison of a range of proteomics techniques. The data presented here show that iTRAQ is an effective method for efficiently comparing secretomes over a large number of samples, whilst COFRADIC can provide more detailed information and greater coverage of the secretome.

5.4.3 Using functional networks to interpret secretome data

Pathway analysis tools such as MetaCore attempt to tackle the complex problem of interpreting '-omics' scale data in terms of signalling networks. The approach taken in chapter 3 and 4 was to examine the enrichment of signalling pathways by experimental data generated by gene expression arrays. In this case, however, this approach was not suitable. The proteomic data generated here does not represent the entire proteome, and this is particularly true of the secretome experiments which are obviously strongly enriched for secreted proteins. A pathway map analysis is therefore difficult to interpret, since these usually assume unbiased representation of the proteome or genome in the data. In this chapter MetaCore was used to identify receptors which are likely to be targeted by the secreted proteins identified by iTRAQ, based on the number of curated interactions between the receptors and the dataset.

5.4.4 Analysis of myofibroblast secretomes and proteomes

The data presented here give a detailed view of the proteome and secretome of gastric cancer derived myofibroblasts and the differences between CAMs and

ANMs. It is clear that the greatest differences occur in the secreted rather than cellular proteins, with 59% of the secreted proteins being differentially abundant, compared with 21% of the cellular proteins. The differences in the secretome are likely to explain much of the functional difference observed between CAMs and ANMs in chapter 4, although the possibility remains that there are also functionally important changes in other secreted factors such as lipid and gaseous signalling molecules, as well as proteins and peptides that were not detected.

The iTRAQ analysis of myofibroblast media identified some potentially interesting proteins which are differentially abundant in CAM media compared with ANM media. Many of these have well established roles in cancer. The MMPs and their inhibitors the TIMPS (Egeblad and Werb, 2002) and members of the uPA system (Andreasen et al., 2000), which here include uPA inhibitors PAI-1 and glia-derived nexin, all have well established roles in tissue remodelling and tumorigenesis. Also identified were a number of extracellular proteins less commonly associated with cancer. Decorin, for example, is a matrix proteogylcan which may be involved in regulating the bioavailability of TGFβ (Iozzo, 1998). It is worth noting that no large changes were observed in growth factors, although these were only identified in the methionyl-COFRADIC experiment. The majority of differentially abundant proteins appear to be ECM constituents, proteases and protease inhibitors.

A particularly interesting protein identified here is βig-h3. This is an ECM-associated, TGFβ-induced protein (Skonier et al., 1992) for which a novel role in cancer has very recently been identified (Zhang et al., 2009). It appears that βig-h3 could be acting as a tumour suppressor by promoting cell adhesion and restricting

migration (Lebaron et al., 1995; Thapa et al., 2007). These features make βig-h3 a particularly attractive candidate for follow-up investigation.

With the exception of glia-derived nexin, the differences in protein abundance in the myofibroblast media are not reflected by the intracellular results, and even in this case the magnitude of the difference is much smaller for intracellular glia-derived nexin. This suggests that, as well as possible differences in transcription or translation, other mechanisms such as the cellular secretory machinery and extracellular degradation may be responsible for these differences.

The analysis of these data could be extended to include statistical tests to help to define the proteins which were differentially abundant. ProteinPilot can be used to calculate p-values which indicate the probability that a protein is differentially expressed based on the strength of evidence at the peptide level. The stringency applied to protein identification confidence could also be increased by estimating the false discovery rate using ProteinPilot.

5.4.5 Effects on important functional networks

The heterogeneity between the myofibroblast pairs might appear to be at odds with the consistent functional data presented in Chapter 4. One might expect to find a secreted protein or proteins which can explain these functional changes in each pair. Given the very consistent functional data it seems likely that the same pathways are being affected by the changes in myofibroblast secretome to produce the same net result, though the individual proteins responsible do not necessarily need to be the same in each case. If there is a group of functionally similar proteins in the myofibroblast secretome, it might only require a change in abundance of any one of

these to produce the same effect on cancer cells. One example from the data presented here are glia-derived nexin and PAI-1, both of which are potent inhibitors of uPA. Glia-derived nexin and PAI-1 are less abundant in CAM than AN in 7 out of 11 and 4 out of 11 secretome pairs respectively. Taken together, however, at least one of these proteins are less abundant in CAM secretomes of 10 out of 11 pairs.

Analysis of the secreted and membrane proteins which were differentially abundant in CAM secretomes compared with their ANM counterparts using MetaCore highlighted integrins, EGF receptor and A2M receptor as proteins which could be important in mediating cellular responses to myofibroblast secretomes. The importance of integrin signalling in mediating cancer cell responses to the microenvironment was first demonstrated by Bissell using an *in vitro* model of breast carcinoma (Howlett et al., 1995). It was shown that inhibition of different integrins can be either pro- or anit-tumourigenic (Weaver et al., 1997). Changes in myofibroblast secreted proteins which interact with integrins could therefore have important functional consequences in the communication between stroma and cancer cells. MetaCore identified more interactions between the differentially abundant proteins of myofibroblast secretomes and integrins than any other class of receptor, as can be seen in Figure 5.8. Of these proteins, several also have established roles in TGFβ signalling, including βig-h3 (Skonier et al., 1992), decorin (Iozzo, 1998) and PAI-1 (Sawdey and Loskutoff, 1991).

Epidermal growth factor (EGF) receptor and other members of the ErbB family of receptor tyrosine kinases have long been known to be important in various cancers (Vogelstein and Kinzler, 2004), as highlighted by the success of ErbB2 directed drug Herceptin (Shawver et al., 2002). Furthermore, there is a bidirectional relationship

between EGFR and integrins, mediated by the MAPK pathway (Wang et al., 1998). A2M receptor, or prolow-density lipoprotein receptor-related protein 1 (LRP-1) as it is also known, is less well studied than integrins and EGFR, but there is an emerging role for this receptor in cancer. It has been suggested to play important roles in regulating apoptosis (Campana et al., 2006) and cancer cell survival (Montel et al., 2007), and so might represent a further mechanism by which myofibroblast CM influences cancer cells.

5.4.6 N-terminal COFRADIC

Analysis of myofibroblast secretomes by N-terminal COFRADIC revealed a number of proteins which are proteolytically processed. Those proteins which are cleaved to a greater or lesser extent in CAM media compared with ANM media are of particular interest, since these may represent important differences between the secretomes of the two cell populations. Ten proteins were identified as being cleaved to a greater extent (i.e. a greater proportion of molecules cleaved) in CAM media, and two as being cleaved to a lesser extent in CAM media, with one, pro-neuregulin, being identified in a cleaved form in CAM media by N-terminal COFRADIC, but not at all in ANM media or at all by methionyl-COFRADIC. Those which are cleaved to a greater extent include three different collagen chains, IGFBP5, MMP2, MMP3, protein-lysine 6-oxidase, lysyl oxidase-like 2, major prion protein and latent-transforming growth factor beta binding protein.

The cleavage of both MMPs occurs within the pro-peptide regions of the proteins, indicating conversion from an inactive zymogen to active protease. In the case of MMP2, there was no difference in the overall abundance of the protein between CAM and ANM media. The results of the N-terminal COFRADIC therefore

highlight the importance of performing this type of analysis, since they reveal greater activation of MMP2. Cleavage of IGFBP5 has previously been identified as an important mechanism by which IGF-II bioavailability is increased by *H. pylori* infection (Hemers et al., 2005). It seems likely, therefore, that a similar mechanism is at work here. Proteolysis of latent transforming growth factor-binding protein-1 has been shown to be a physiological mechanism for TGFβ release (Dallas et al., 2002). It was also shown to be a substrate for MMP2, MMP9, plasmin and elastase. Its cleavage in this system could therefore be related to the increase in MMP2 activation in CAM media. The cleavage of uPA within the kringle domain might indicate conversion of the receptor-binding high molecular weight form of uPA to the soluble low molecular weight form (Kasai et al., 1985a; Kasai et al., 1985b). The data indicate that in CAM media, a slightly smaller proportion of the overall uPA is cleaved here than in ANM media. There is, however, a much greater total amount of uPA in CAM media than ANM, and so the overall amount of both the high and low molecular weight forms is greater. A proteolytically processed form of lysyl oxidase-like 2 has been shown to promote migration in non-invasive breast cancer cells, but not in normal breast epithelial cells (Hollosi et al., 2009). These all therefore represent differences between CAM and ANM secretomes which could have important functional significance.

5.4.7 Summary

This study is the first to define the secretome of a myofibroblast population from any cancer, as well as the first to compare CAM and ANM secreted factors using a proteomics approach. Previous proteomic comparisons of CAMs and ANMs have largely focussed on breast cancer, and none have studied gastric myofibroblasts.

Furthermore, it is the first to attempt a comparison of the substrate degradomes of two populations of primary human cells. The overall trend seems to be that many proteins, particularly ECM proteins and protease inhibitors, are less abundant in CAM media than ANM media, with many of the more abundant proteins being proteases. As with the gene array analyses from Chapter 3, many of the identified proteins are involved in TGFβ signalling and related functional networks. It is now apparent that these proteins are in part regulated by proteolytic processing, although the functional significance of these proteins and their regulation by proteases must still be investigated.

5.5 Conclusions

- iTRAQ is an appropriate method for comparing the secretomes of CAM and ANM myofibroblasts.
- There are significant differences in the abundance of ECM proteins, proteases and protease inhibitors between CAMs and ANMs.
- Many of the differentially abundant proteins are known to be involved in integrin and EGF receptor signalling.
- There are significant differences in the substrate degradomes of CAMs and ANMs, with a number of proteins which have been implicated in cancer progression being cleaved.

Chapter 6

The role of βig-h3 in gastric cancer myofibroblasts and regulation by the uPA system

6.1 Introduction

In chapter 3, TGFβ signalling was identified as a functional network which is likely to be affected by the differences in gene expression between CAMs and ANMs. In chapter 4 it was identified as a network likely to be affected by the changes in gene expression induced in AGS cells by CAM and ANM media treatment. Chapter 4 also revealed that CAM media induces more AGS cell migration and invasion than ANM media. Genomic and proteomic approaches are valuable tools in defining the molecular mechanisms of disease, as has been demonstrated so far in this thesis. There still remains, however, the need to validate the functional significance of genes and proteins identified by these methods. This chapter will focus on the adhesion molecule transforming growth factor-beta-induced protein ig-h3 (βig-h3), one of the proteins identified by SILAC and iTRAQ as less abundant in CAM CM.

Beta ig-h3 is a particularly attractive protein to follow up on, since it could provide a link between gene array, migration, invasion and secretome data. Furthermore, beta ig-h3 had been previously identified as a possible target of urokinase plasminogen activator (uPA) in gastric myofibroblasts by a previous proteomic study in our group (Cédric Duval doctoral thesis). Since members of the uPA system have been shown to be upregulated in preneoplastic lesions of the stomach (Kenny et al., 2008), and inhibitors of uPA (PAI-1 and glia-derived nexin) were consistently less abundant in CAM media (Chapter 5), the uPA system might also be involved here.

This extracellular matrix protein is known by several names, including transforming growth factor beta-induced-protein ig-h3 (beta ig-h3), kerato-epithelin and RGD-containing collagen-associated protein (RGD-CAP), as well as the parent gene name TGFBI or BIGH3. Beta ig-h3 was first identified as a TGFβ responsive gene in the

lung adenocarcinoma cell line A549 (Skonier et al., 1992). It has since been characterised as a 683 amino acid protein containing a secretory signal (residues 1-24), an EMI domain, four fascilin 1-like (FAS1) domains and a carboxy-terminal RGD sequence (Thapa et al., 2007). It has been implicated as a regulator of differentiation (Dieudonne et al., 1999; Kim et al., 2000) and wound healing (Rawe et al., 1997). It is its role as an adhesion molecule, however, that has been documented more fully (Lebaron et al., 1995; Ohno et al., 1999), with interactions having been demonstrated between beta ig-h3 and integrin receptors $\alpha 3\beta 1$, $\alpha v\beta 3$ and $\alpha v\beta 5$ (Hanssen et al., 2003; Kim et al., 2002; Nam et al., 2003).

A role for beta ig-h3 in cancer has recently been identified, with *in vitro* studies proposing a role as a tumour suppressor (Shao et al., 2006; Zhao et al., 2004). It was recently shown by Zhang et al. (2009) that mice lacking the TGFBI gene show spontaneous tumour development, with embryonic fibroblasts derived from the strain exhibiting increased proliferation and early S-phase entry compared with wild-type. These findings suggest that beta ig-h3 acts as a tumour suppressor, though the findings of Ma et al. (2008) indicate that it can also promote the metastasis of colon cancer cells via the process of extravasation.

It would seem that the role of beta ig-h3 is a complex one, and likely to be context dependant. It is important, therefore, to investigate what function it has in this situation, and whether it has a role in mediating the functional differences between CAMs and ANMs. Given its previous identification as a target of uPA in gastric myofibroblasts, the role of proteolysis as a potential mechanism for its regulation in this system should also be explored.

6.1.1 Aims

The aims of this chapter are to investigate the involvement of βig-h3 in myofibroblast media-induced AGS cell migration, and how it might be regulated, particularly with regards to the uPA system. The specific objectives are;

1. To use Western blotting to confirm changes in βig-h3 abundance identified by iTRAQ.
2. To determine the role of βig-h3 in myofibroblast CM-stimulated AGS cell migration.
3. To investigate how βig-h3 functionality might be regulated by uPA proteolytic activity.

6.2 Materials and Methods

6.2.1 Western blotting

Western blotting was performed as described in Chapter 2. For concentrated media samples 2μg total protein was loaded per lane, and 10μg total protein per lane for cell extract samples. The primary antibody used was rabbit anti-human βig-h3 IgG (R&D), which was used at a dilution of 1 in 1000, and the secondary antibody was goat anti-rabbit horseradish peroxidise-conjugated IgG (Sigma), which was used at a dilution of 1 in 10,000. The masses of peptide bands were calculated by first plotting a standard curve of distance travelled by PageRuler Plus standard markers (Fermentas).

6.2.2 Migration assays

Modified Boyden chamber migration assays (BD Control Cell Culture Inserts, BD Biosciences, Massachusetts, USA) were performed as described in Chapter 2. AGS

cells were added at a density of 20,000 cells per well in serum free DMEM supplemented with 1% v/v penicillin/streptomycin solution (Sigma). Where indicated the media in the top and bottom wells were supplemented at the beginning of the migration assay with either 5μg/mL recombinant human βig-h3 (R&D), 2μg/mL uPA inhibitory antibody (American Diagnostica), or 10μM synthetic uPA inhibitor uPA STOP (American Diagnostica).

6.3 Results

6.3.1 βig-h3 can be downregulated or degraded in cancer-derived myofibroblast media

Western blotting was used to confirm the finding from Chapter 5 that βig-h3 is less abundant in CAM media than ANM media. In all 11 pairs the ANM media blots showed expression of βig-h3, which ran as two bands with apparent molecular masses of 70kDa and 75kDa. CAM media produced varying results, showing either similar βig-h3 expression to their ANM counterparts, decreased βig-h3, or an increase in apparent degradation of βig-h3 (examples shown in Figure 6.1 A, B, and C respectively). The two most prominent degradation products observed had molecular masses of 56kDa and 50kDa. Table 6.1 summarises the data for βig-h3 in myofibroblast media samples from SILAC, iTRAQ and Western blot experiments. There is good agreement between the proteomic data and Western blot for patient 3, which was the only one to show a decrease in βig-h3 abundance in CAM by Western blotting, and also shows the greatest decrease by iTRAQ. There is also good agreement for patients 6 and 7, where Western shows no changes and the iTRAQ fold change is less than 1.3. The samples which showed an increase in βig-h3 degradation do not appear to correlate with the iTRAQ data, with both increases and

decreases in abundance indicated by iTRAQ for these samples. With proteomic and Western blot data taken together there is evidence for either reduced abundance or increased degradation of βig-h3 in 9 out of 11 CAMs compared with their matching ANMs.

Cell extracts from all 11 pairs of myofibroblasts revealed no difference between CAMs and ANMs, with the 70kDa band representing the majority of the βig-h3, as seen in the media samples (Figure 6.1). Additional bands were apparent both above and below 70kDa, suggesting the presence of both degradation products, with molecular masses of 48kDa and 42kDa, and possible evidence of post-translation modification in the cell extract samples.

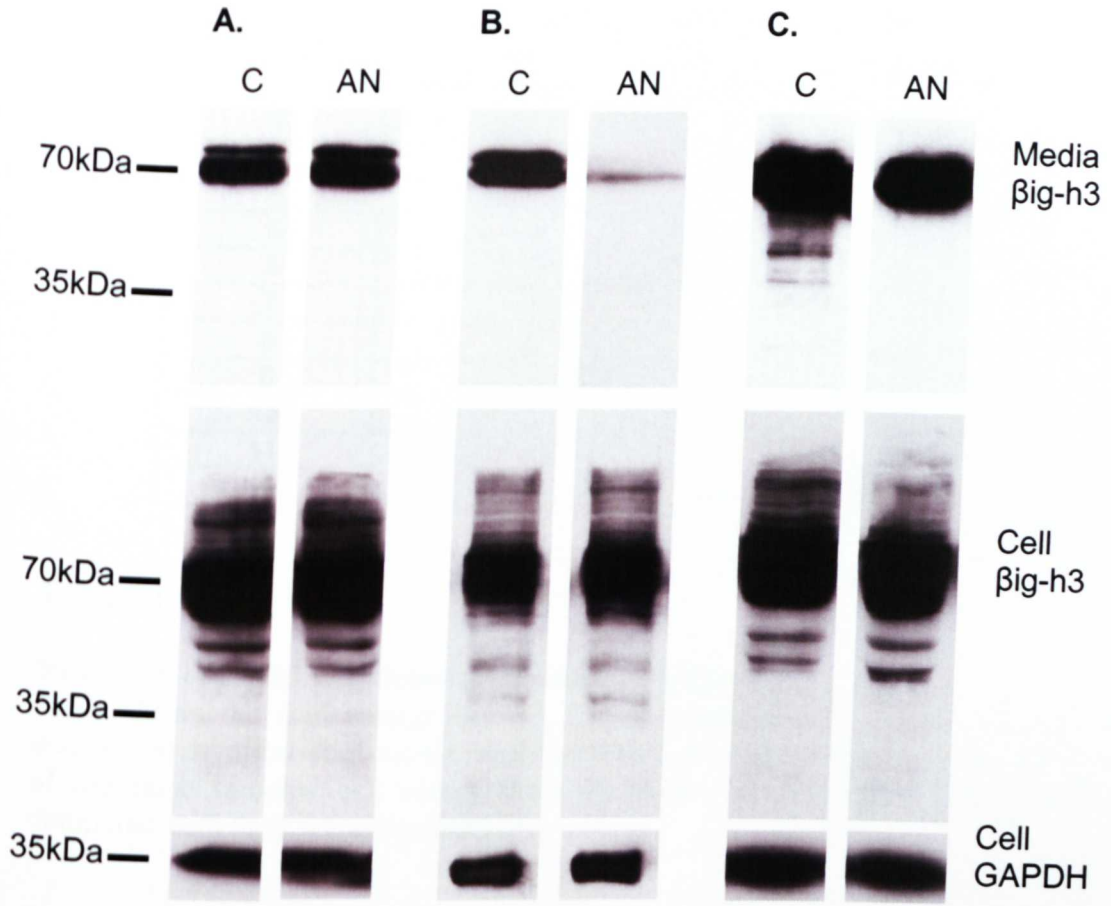


Figure 6.1

βig-h3 is either less abundant or degraded in different cancer derived myofibroblast media.
Western blots for βig-h3 in myofibroblast media (top panels) and cell extracts (middle and lower panels). Examples of A, CAM media with similar abundance of βig-h3 to the corresponding ANM media, B, CAM media with a lower abundance of βig-h3 and, C, with an increase in degradation of βig-h3. Corresponding cell extracts are shown directly beneath CM blots, showing no difference between CAM and ANM myofibroblasts. GAPDH loading control for cell extracts are shown in the lower panels.

Patient	Change in CAM relative to ANM		
	SILAC (fold change)	iTRAQ (fold change)	Western blot
1	-1.31	-1.28	Degradation
2	-2.43	-1.37	No change
3		-2.65	Less abundant
5		1.21	Degradation
6		-1.10	No change
7		1.17	No change
8		-1.94	Degradation
9		1.37	Degradation
10		1.09	Degradation
11		1.33	Degradation
12		-1.34	No change

Table 6.1

Summary of βig-h3 data from myofibroblast media
Table shows the fold change data for βig-h3 generated by SILAC and iTRAQ analysis of myofibroblast media, together with a summary of Western blot analysis of the same samples. All comparisons are shown as changes in CAM samples compared with their ANM counterparts.

6.3.2 βig-h3 inhibits myofibroblast-induced AGS cell migration

Migration assays of AGS cells were performed in the presence and absence of recombinant human βig-h3 in order to determine whether the changes in βig-h3 could be a mediator of the functional effects of myofibroblast CM on AGS cells. AGS cells show little or no basal migration in a modified Boyden chamber assay of migration (data not shown), but migrated when stimulated with myofibroblast CM. CAM CM from patient 1, which showed an increase in βig-h3 degradation compared with ANM, was used to stimulate AGS cell migration, with and without the addition of 5μg/mL exogenous βig-h3. There was a significant decrease in AGS migration in the wells with βig-h3 (Figure 6.2).

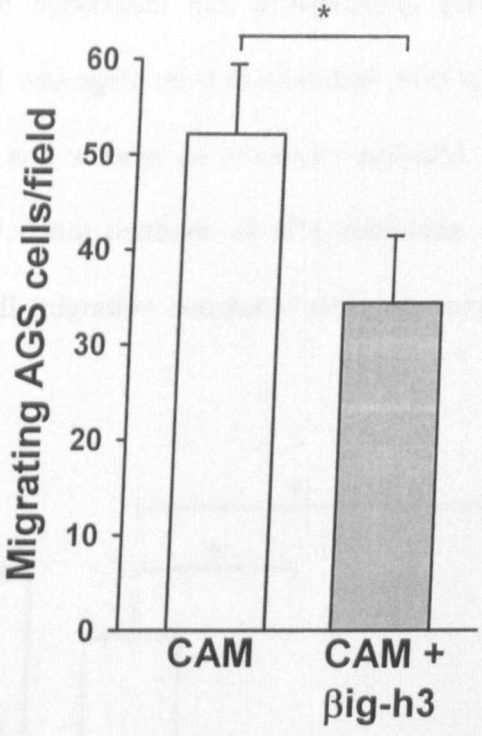


Figure 6.2

βig-h3 inhibits myofibroblast CM-induced AGS cell migration.
CAM CM stimulates migration of AGS cells. Addition of 5μg/mL exogenous βig-h3 to the myofibroblast CM results in a significant inhibition of induced AGS migration ($p<0.05$).

6.3.3 The uPA/Plasmin system influences AGS cell migration via degradation of βig-h3

The finding that βig-h3 is decreased or degraded in CAM CM raises the question; which protease or proteases are responsible for its degradation? As already mentioned, previous work has suggested that βig-h3 could be a target for uPA in gastric myofibroblasts, either directly or indirectly. Since the main role of uPA is the conversion of plasminogen to plasmin, both uPA and plasmin were investigated as possible candidates for cleaving βig-h3. Given that βig-h3 can inhibit AGS cell migration, it is possible that AGS cell migration may be influenced by proteolytic activity.

In order to test the hypothesis that uPA/plasmin activity promotes AGS cell migration, CAM CM was again used to stimulate AGS cells in a modified Boyden chamber assay, with and without an inhibitory anti-uPA antibody or the synthetic inhibitor, uPA STOP. Both methods of uPA inhibition resulting in a significant reduction in AGS cell migration compared with the control CAM CM treatment (Figure 6.3).

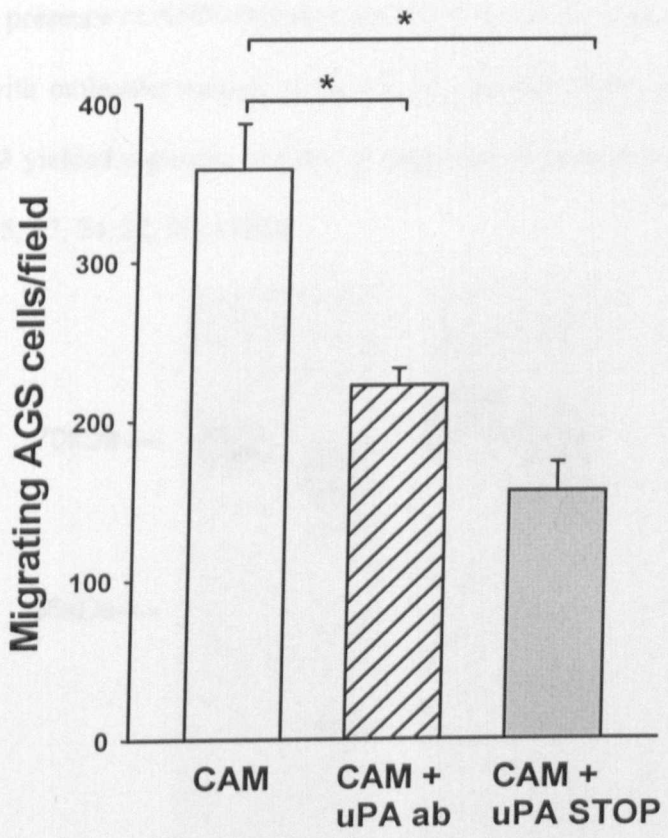


Figure 6.3

Inhibiting uPA reduces the stimulatory effect of myofibroblast CM on AGS cell migration.
CAM CM stimulates the migration of AGS cells. This effect is significantly inhibited ($p<0.05$) by either an anti-uPA inhibitory antibody or a synthetic uPA inhibitor (uPA STOP).

It was previously shown that uPA causes cleavage of βig-h3 in the presence of myofibroblasts, but that it does not cleave recombinant βig-h3 (Cedric Duval doctoral thesis). Plasmin was therefore investigated as a possible mediator of uPA-induced cleavage. Recombinant βig-h3 was incubated for 24 hours at 37°C in the presence and absence of plasmin. In addition, ANMs from patient 12 were treated with plasmin and media collected. Western blot analysis of these samples revealed that plasmin treatment yielded several cleaved fragments of both rhβigh3 and native βigh3 in the presence of ANM (Figure 6.4). The cleaved recombinant βig-h3 yielded fragments with molecular masses of 66, 58, 56, 50 and 23kDa, whilst the cleaved native βig-h3 yielded a greater number of fragments, with molecular masses of 76, 63, 56, 50, 35, 27, 24, 22, 20, 17kDa.

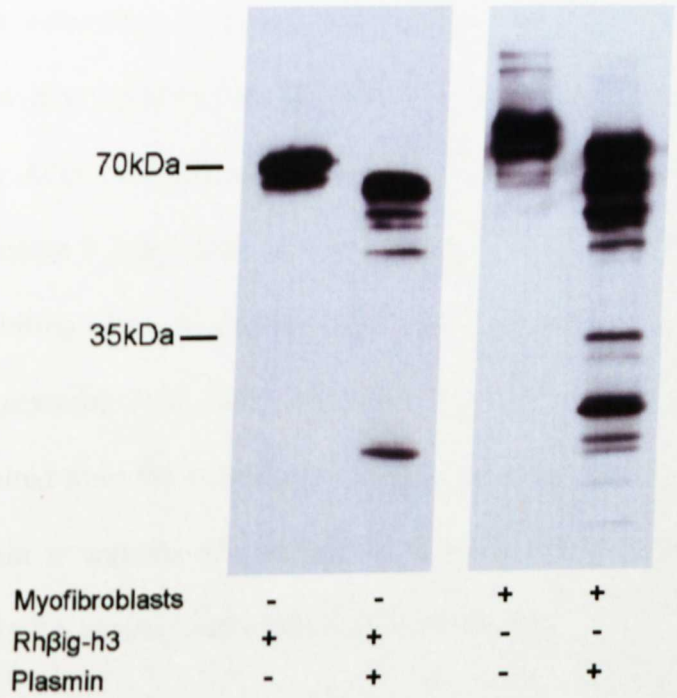


Figure 6.4

Plasmin degrades βig-h3
Western blots showing that plasmin is able to degrade recombinant human βig-h3 (left). Addition of plasmin to serum-free ANM culture media also results in degradation of secreted βig-h3 (right).

6.4 Discussion

The data presented here demonstrate a role for the adhesion molecule βig-h3 in mediating some of the functional differences between CAMs and ANMs. In Chapter 4 it was shown that CAMs promote the migration of gastric cancer cells, and it now appears that this may be in part via degradation of βig-h3 by plasmin. It was shown that βig-h3 is an effective inhibitor of AGS cell migration, and so it is loss of βig-h3 function, either by reduced abundance or degradation, that can mediate the effect of CAMs on AGS cell migration. It is possible that the increase in CAM migration compared with ANM shown in Chapter 3 might also be in part the result of increased βig-h3 degradation, since βig-h3 has previously been shown to also inhibit gastric myofibroblast migration (Cedric Duval, doctoral thesis). The data presented here show that inhibition of uPA activity in CAM media reduced the effect of the media on AGS cell migration. The use of uPA inhibiting antibody and synthetic inhibitor make it impossible to distinguish between the action of uPA and plasmin, since inhibiting uPA will also have the effect of reducing conversion of plasminogen to active plasmin. It is likely, however, that any βig-h3-mediated portion of this effect resulted from the reduction in plasmin activity, since it was demonstrated here that plasmin is capable of cleaving recombinant βig-h3, whilst it was previously shown that uPA is not (Cedric Duval doctoral thesis).

Western blot analysis of βig-h3 degradation by plasmin revealed the liberation of approximately five peptides from rhβig-h3, but several more from native βig-h3 in myofibroblast media. The obvious explanation for this difference is the presence of additional proteases in the myofibroblast media which could cleave the βig-h3 fragments following the initial cleavage by plasmin. In both cases, however, fragments with molecular masses of 56kDa and 50kDa were observed, as was also

seen in the untreated CAM media which showed βig-h3 degradation. These fragments might therefore be the result of plasmin cleavage alone. It was previously shown that adding exogenous uPA to myofibroblast media generates βig-h3 fragments with molecular masses of 55kDa and 27kDa (Cedric Duval, doctoral thesis). Similar fragments of 56kDa and 27kDa were observed in the plasmin treated myofibroblast media. Although the addition of uPA was likely to increase levels of active plasmin by conversion of plasminogen, uPA is known to also regulate the activity of several other proteases (Baramova et al., 1997; Keskioja et al., 1992), which might account for the differences in βig-h3 cleavage compared with the plasmin treatment.

The 27kDa fragment from the previous uPA experiment was identified as residues 377-590 of βig-h3. This fragment contained all but 2 residues of the third FAS1 domain, and most of the fourth FAS1 domain. Liberation of intact FAS1 domains could potentially be a mechanism by which proteolytic processing of βig-h3 might confer novel functionality. It was previously reported that the fourth FAS1 domain of βig-h3, designated Fastatin, could inhibit the adhesion and migration of endothelial cells by inhibiting binding to βig-h3, fibronectin and vitronectin (Nam et al., 2005). Furthermore, Fastatin was shown to inhibit angiogenesis and tumour growth *in vivo*, and to induce apoptosis of endothelial cells. The inhibition of adhesion and migration by Fastatin was also shown to be RGD-dependant. It is possible that an RGD-containing peptide is liberated from βig-h3 by proteolytic processing, since the generation of the 27kDa fragment required cleavage of the protein on the N-terminal side of the RGD domain (Figure 6.5). It was previously suggested that βig-h3 normally undergoes carboxy-terminal processing, resulting in the liberation of an RGD-containing peptide (Skonier et al., 1994). Study of βig-h3-

derived peptides containing the RGD domain indicated that they could be important in mediating TGFβ-induced apoptosis (Kim et al., 2003). It is possible, therefore, that proteolytic cleavage of βig-h3 in the myofibroblast secretome might not result simply in a loss of βig-h3 function, but could also generate peptides with novel functions. Further investigations to define the sequence of the peptides generated by plasmin cleavage and their possible functions are therefore required.

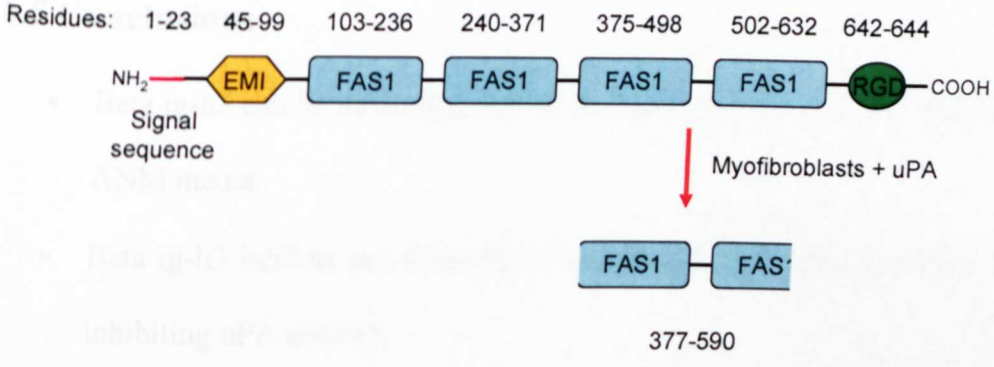


Figure 6.5

Domain structure and cleavage sites of βig-h3
The extracellular protein βig-h3 contains an EMILIN-like domain (EMI), four fasciclin-like domains (FAS1) and an RGD containing domain. One of the peptide fragments which are liberated from native βig-h3 when myofibroblasts are incubated with exogenous uPA is shown.

The findings here are consistent with previous observations that βig-h3 is a target of uPA in gastric myofibroblasts, and that uPA increases myofibroblast migration (Cedric Duval doctoral thesis). Given that data presented here, it now seems likely that the uPA-mediated degradation of βig-h3 occurs via conversion of plasminogen to plasmin, and that it is plasmin that directly cleaves βig-h3. The previous study also demonstrated that uPA activity is greater in CAMs than ANMs. This again fits well with the observation in this chapter that βig-h3 is often degraded in CAM, but not ANM media. These data are also consistent with the suggested role of βig-h3 as

a tumour suppressor (Kim et al., 2000; Skonier et al., 1994; Zhang et al., 2009), and with the idea that increased levels of uPA activity can be tumorigenic (Andreasen et al., 2000; Kenny et al., 2008). The model proposed by this thesis is that uPA activity is greater in CAM media than ANM media, and that this leads to an increase in βig-h3 degradation via plasmin, which is a novel mechanism by which uPA and βig-h3 can promote cancer cell migration.

6.5 Conclusions

- Beta ig-h3 can be downregulated or degraded in CAM media compared with ANM media.
- Beta ig-h3 inhibits myofibroblast CM-induced AGS cell migration, as does inhibiting uPA activity.
- Beta ig-h3 is cleaved by plasmin, which represents a possible mechanism by which it is degraded in CAM media.

Chapter 7

Discussion

7.1 Main findings and conceptual advances

The main findings of this thesis are:

(a) Gastric CAMs are distinct from their ANM counterparts, with a more migratory phenotype and differences in gene expression.

(b) CAM media induces a more migratory, proliferative and invasive phenotype in a gastric cancer cell line, with associated changes in gene expression.

(c) There are significant differences in the secretomes of CAMs and ANMs, some of which are the result of proteolytic processing.

(d) TGFβ signalling is an important functional network in both the differences between CAMs and ANMs (Figure 7.1), and in the effects of secreted factors on AGS cells.

(e) The different effects of CAMs and ANMs may, in part, be mediated by decreased abundance or increased degradation of βig-h3 in CAM secretomes (Figure 7.2).

7.1.1 CAMs are functionally distinct from ANMs

Whilst both the CAMs and ANMs used in this study were positively identified as myofibroblasts, there are clear differences between the two cell populations. It was previously shown that the myofibroblast population is expanded in the stroma of gastric adenocarcinoma compared with normal gastric tissue (Jiang et al., 2008), and it is now clear that the myofibroblasts from these two microenvironments are also functionally distinct. Many of the differences between CAMs and ANMs observed here could translate *in vivo* to CAMs defining a microenvironment which contributes to oncogenic disease. This would be consistent with studies of other cancers which have shown that CAMs promote neoplastic development and progression (Bhowmick et al., 2004; Olumi et al., 1999).

Previous work with gastric myofibroblasts has shown evidence for global hypomethylation of DNA in CAMs compared with adjacent non-cancer tissue myofibroblasts (ANMs) (Jiang et al., 2008). This thesis has further defined the differences between CAMs and ANMs, identifying differences in gene expression and a more migratory phenotype in CAMs. The increased migration of CAMs compared with ANMs, both unstimulated and in response to IGF-II, could be an important feature *in vivo*. In invasive carcinomas it has been observed that a large proportion of myofibroblasts are found at the invasive front (Elenbaas and Weinberg, 2001; Schurch, 1997), which suggests migration of myofibroblasts in response to factors produced by the invasive cancer cells. Expression of proteases such as the cell surface MMP14 and soluble MMP2 by the myofibroblasts enhances invasiveness by degrading the ECM (Bisson et al., 2003; Deryugina et al., 1997; Hotary et al., 2000). Perhaps more important with regards to their role in disease is the effect of myofibroblasts on cancer cells, which again revealed differences between CAMs and ANMs, with CAMs inducing greater migratory, invasive and proliferative cancer cell phenotypes with accompanying changes in gene expression. These findings demonstrated the role of myofibroblasts in paracrine signalling (Powell et al., 1999a).

The development of cancer can be thought of as an evolutionary process (Merlo et al., 2006), and as such the microenvironment can provide selective pressure which influences this evolution. It is now believed that cancer cells co-evolve with their microenvironment (Littlepage et al., 2005; Polyak et al., 2009), and the observation that CAMs induce different gene expression changes in AGS cells than their ANM counterparts would seem to agree with this concept. It is not necessarily the case, however, that CAMs evolve from ANMs. Myofibroblasts may derive from cancer

cells through EMT, differentiation from fibroblasts in response to factors such as TGFβ (Hinz et al., 2007), or differentiation from bone marrow-derived mesenchymal stem cells (Pittenger et al., 1999). Gastric CAMs might therefore derive either from ANMs in response to cancer cell-derived factors, or the two myofibroblast populations could originate from different cellular pools.

7.1.2 Changes in functional networks

Gene array analysis, both of CAMs and ANMs, and of AGS cells treated with myofibroblast CM, revealed that TGFβ signalling is one of the most significantly affected functional networks by the differences between CAMs and ANMs. This is a particularly interesting growth factor in the context of cancer. Early work on the TGFβ family of growth factors suggested a role as a tumour suppressor, since it inhibits the proliferation of many epithelial cell types (Tucker et al., 1984). It can also, however, have tumour promoting functions, such as mediating loss of adherens junctions in metastasis (Akhurst and Derynck, 2001). It has since been demonstrated to be important in mediating the tumour-promoting ability of CAMs in pancreas (Ohuchida et al., 2004), breast (Barcellos-Hoff, 1998; Kuperwasser et al., 2004) and other tissues (Bhowmick et al., 2004).

The binding of TGFβ to its receptors can induce the expression of a number of genes through activation of the SMAD signalling pathway. The gene for βig-h3 is one such example (Skonier et al., 1992). The role of βig-h3 in mediating the effects of gastric myofibroblasts on AGS cancer cells is therefore an example of the importance of TGFβ-associated signals in this system.

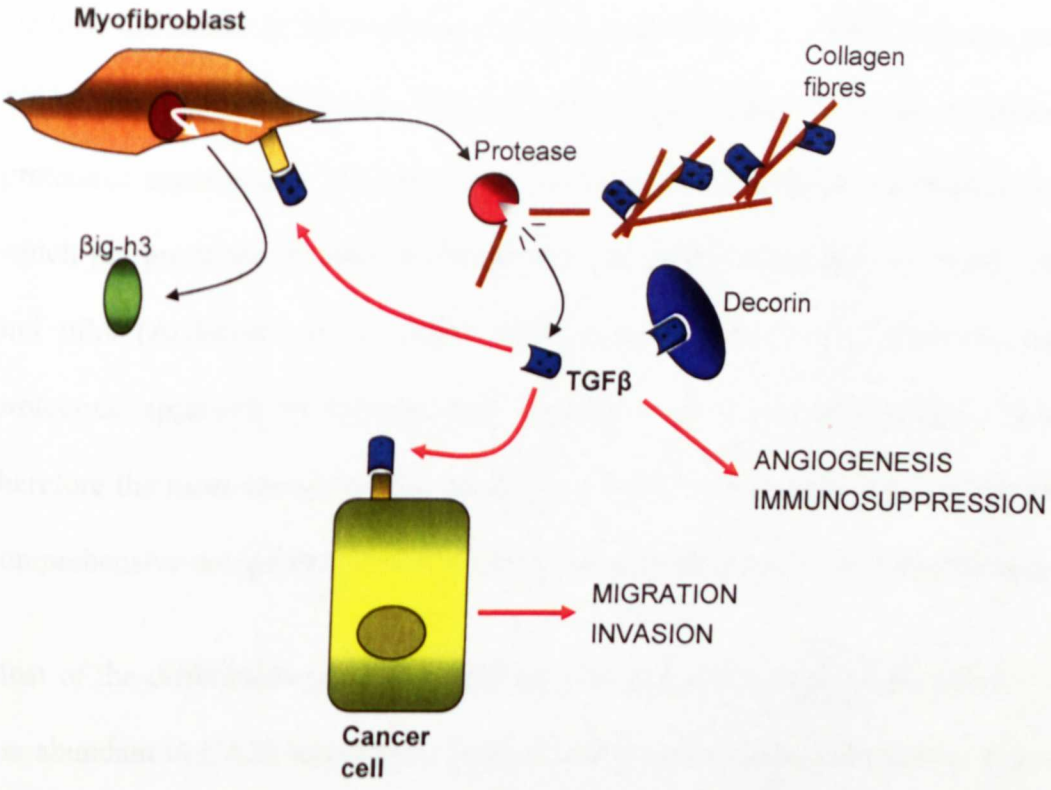


Figure 7.1

Possible roles of TGFβ in the gastric cancer microenvironment
The extracellular matrix stores TGFβ secreted by both myofibroblasts and cancer cells, through binding to proteins including collagen and decorin. Bioavailability of TGFβ could be increased by cleavage of these proteins, allowing it to bind receptors on cancer cells, myofibroblasts and other cells. This could induce production of proteins such as βig-h3 in myofibroblasts, induce an invasive phenotype in cancer cells, and influence other components of the microenvironment such as blood vessels and leukocytes.

7.1.3 The key differences are in the secretome

The different effects of CAM and ANM CM on AGS cell behaviour can only be explained by differences in the myofibroblast secretomes. Studies in breast revealed elevated production of SDF-1 (CXCL12) by CAMs, which contributed to tumour growth and angiogenesis (Orimo et al., 2005). Other studies of the role of CAMs has focussed on growth factors which are known to be produced by fibroblasts and myofibroblasts and with established roles in cancer, such as TGFβ (Bhowmick et al., 2004) and HGF (De Wever et al., 2004). In contrast to these studies, there appears to

be little difference in the secretion of factors such as TGFβ or HGF between gastric CAMs and ANMs, although these growth factors were only rarely identified in proteomic experiments. Yet others have focussed on the roles of proteases, many of which are produced by myofibroblasts, such as MMPs (Egeblad and Werb, 2002) and uPA (Andreasen et al., 2000). This study, however, is the first to use a proteomic approach to compare the secretomes of CAMs and ANMs. This is therefore the most complete description of a myofibroblast secretome, and the most comprehensive comparison of CAM and ANM secretomes and cellular proteomes.

Most of the differences between CAM and ANM media were proteins which were less abundant in CAM secretomes. Many of these were extracellular matrix proteins such as collagens, fibronectin and laminins, as well as protease inhibitors such as serpins and TIMPs. These results challenge the widely held belief that myofibroblasts are the main contributors to ECM deposition and are responsible for the observed increased ECM in cancer stroma (Powell et al., 1999a; Powell et al., 1999b). Of the proteins which were more abundant in CAM media than ANM media, many were proteases, particularly MMPs. Together these observations suggest greater proteolytic activity in CAM media. This might explain the apparent paradox that ECM protein levels are lower in CAM media than ANM media. It could also be the case that the greater number of myofibroblasts found in cancer stroma overcomes the lower levels of ECM proteins in the secretome to maintain an expanded ECM in cancer.

It is worth noting that no large changes were observed in growth factors, although these proteins were only identified in the methionyl-COFRADIC experiment. A number of the differences between CAM and ANM secretomes could, however,

have effects on the bioavailability of several growth factors. These include decorin, which binds TGFβ, and several of the insulin-like growth factor-binding protein (IGFBPs). It has previously been shown that IGFBP5 is an important regulator of IGF-II in the gastric mucosa, and that cleavage of IGFBP5 by MMP7 can increase IGF-II bioavailability (Hemers et al., 2005; McCaig et al., 2006). Downregulation or cleavage of growth factor-binding proteins might therefore represent an important mechanism by which functional networks such as TGFβ signalling are altered in the CAMs.

Identification by microarray of extracellular signalling pathways which differ between CAMs and ANMs might suggest that the differences in their secretomes are the result of differences in gene expression. When the secretome is compared with the cellular proteome, however, it appears that many of the differences in extracellular protein abundance are not reflected intracellularly. These observations highlight the importance of studying the secretome directly. Clearly gene expression and translation of mRNA can influence the abundance of extracellular proteins, but other factors such as the secretory machinery and protein degradation can also be significant. It is for these reasons that the secretome has become a popular topic within the proteomics research community (Gronborg et al., 2004; May, 2009; Paulitschke et al., 2009).

7.1.4 Proteases are important in defining the myofibroblast secretome

High protease activity in cancer stroma has been associated with a more aggressive phenotype, facilitating invasion, metastasis and liberation of growth factors by degrading extracellular matrix components (Andreasen et al., 2000; Egeblad and Werb, 2002; McCaig et al., 2006). It has recently been recognised that proteases can

also play complex roles in signalling events by targeted proteolytic processing (Lopez-Otin and Overall, 2002). In the context of gastric myofibroblasts it seems likely that proteases help to define the secretome, both through complete degradation of proteins and by proteolytic processing.

Using N-terminal COFRADIC a number of proteins were identified which, proportionally, were proteolytically processed to a greater extent in CAM media than in ANM media. A number of these differences in the substrate degradomes of CAM and ANM media were found to be in proteins which have been implicated in cancer progression, such as MMPs, IGFBP5 and uPA. These data also suggest the presence of protease cascades, since MMP2 and MMP3 were cleaved within their pro-peptide regions, which is necessary for their activation.

It is possible that some proteins are less abundant in CAM media than ANM media due to an increase in proteolytic degradation. Collagens, for example, were consistently identified as being less abundant in CAM media. This would clearly fit with the general increase in abundance of MMPs and decrease in protease inhibitors, since many of the MMPs, especially collagenases such as MMP1, degrade collagen (Woessner, 1991). Proteins which were completely degraded by protease activity are unlikely to be detected by N-terminal COFRADIC, since this technique relies on the presence of intact peptides which are sufficiently large as to be detected by MS (Gevaert et al., 2003). If, however, it is the case that a significant proportion of the differences between CAM and ANM secretomes are the result of protein degradation, and that there is a cascade of protease activation as suggested by N-terminal COFRADIC, then these differences may be precipitated by changes in abundance of a small number of proteases or protease inhibitors.

TGFβig-h3 is a good example of a protein which is regulated by protease activity. The data presented in this thesis indicate that cleavage of βig-h3 by plasmin reduces its ability to inhibit AGS cell migration. It seems likely that the lower abundance of uPA inhibitors in CAM media compared with ANM media contributes to higher uPA activity, in turn resulting in an increase in the conversion of plasminogen to plasmin, and thus an increase in βig-h3 degradation. Given that both uPA and plasmin are known to activate other protease such as MMP2 (Baramova et al., 1997; Kazes et al., 1998; Keskioja et al., 1992), an increase in uPA activity could be a key step in a protease cascade within the myofibroblast secretome.

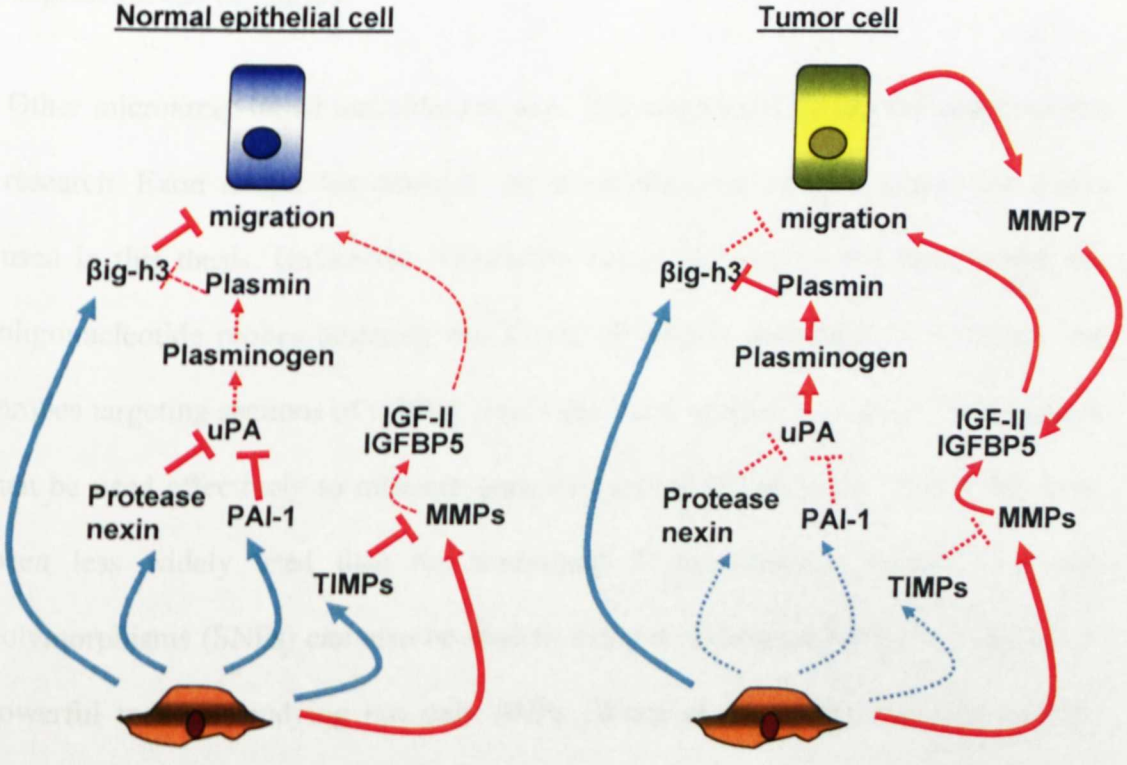


Figure 7.2

Factors mediating the effect of myofibroblasts on gastric cancer cell migration Schematic shows some of the mechanisms by which proteins in the myofibroblast secretome influence the migration of normal gastric epithelial cells and gastric cancer cells. Increased abundance of proteases and decreased proteases inhibitors in the cancer-associated myofibroblast secretome leads to cleavage of factors such as βig-h3 and IGFBP-5, resulting in an increase in cancer cell migration.

7.2 Methodology

The methods used in this thesis to characterise the differences between CAMs and ANMs are by no means exhaustive. For example, there has been great interest in recent years in the importance of micro RNAs in cancer. These small RNA molecules can regulate gene expression, and there is now strong evidence for altered levels of various micro RNAs in cancer (Ma et al., 2007; Meltzer, 2005). Micro RNAs can be profiled using microarrays similar to those used in this thesis to profile gene expression (Liu et al., 2008). An approach such as this might be able to identify some of the mechanisms by which gene expression is altered in CAMs compared with ANMs, although changes in micro RNAs are not the only means by which gene expression can be altered.

Other microarray-based technologies were also considered during the course of this research. Exon arrays, for example, are a variation on the gene expression arrays used in this thesis. Unlike the Affymetrix GeneChip arrays used here, which use oligonucleotide probes targeting the 3' end of mRNA molecules, exon arrays use probes targeting sections of mRNA from each exon of the parent gene. These arrays can be used effectively to measure gene expression (Kapur et al., 2007), but have been less widely used than the traditional 3' microarrays. Single nucleotide polymorphisms (SNPs) can also be studied using a microarray platform, which is a powerful tool for studying not only SNPs (Wang et al., 1998), but also genetic aberrations including allelic imbalances such as loss of heterozygosity (LOH) (Bignell et al., 2004). Single nucleotide polymorphisms can be important in cancer, including a familial form of gastric cancer (Guilford et al., 1998), whilst LOH has been identified both in gastric cancer cells (Tahara, 1993), and in the stromal

compartment of mammary carcinoma (Moinfar et al., 2000). This technique would be suited to a study comparing the genomes of CAMs and ANMs.

The COFRADIC experiments could have been expanded, not only to analyse more samples, but also to use other variations on the COFRADIC method to enrich for other classes of peptides. For example, COFRADIC can be used to study N-glycosylation (Ghesquiere et al., 2007; Ghesquiere et al., 2006), a post-translational modification which can influence protein function and with important roles in cancer (Hakomori, 2002).

7.3 Future work

As well as offering advances in the understanding of the role of myofibroblasts and their secretomes in gastric cancer, the work presented in this thesis has also identified a number of areas in which further progress could be made.

An important question which has not yet been addressed is whether CAMs and ANMs share a common origin, or are derived from separate pools of cells. A good way to begin this investigation might be to study the effects of cancer cell-derived factors on ANM gene expression. If CAMs and ANMs do indeed share a common origin, a possible explanation for the differences between them might be conditioning by cancer cells, resulting in permanent changes. It would therefore be interesting to see whether ANMs can be made to adopt a CAM-like pattern of gene expression, or indeed a more migratory phenotype.

The proteomic data presented here offer a wealth of potential targets for follow-up study. Following on from the investigations into the role of βig-h3, it would be interesting to look at the significance of other proteins which are related to TGFβ

signalling. Latent TGFβ-binding protein 1S, for example, was shown to be proteolytically processed to a greater extent in CAM media than ANM media by N-terminal COFRADIC. This could represent a mechanism by which TGFβ bioavailability is increased in CAM media (Dallas et al., 2002; Yee et al., 1993). Similarly decorin, another TGFβ-binding protein (Iozzo, 1998), was shown to often be less abundant in CAM media than ANM.

It could also be important to identify which specific proteases are responsible for the degradation or proteolytic processing of various members of the myofibroblast secretome. A study such as this could use combinations of protease inhibitors, gene knock downs and overexpression to dissect protease substrate repertoires. In this way it might be possible to identify the protease or proteases with the most significant role in defining the differences between CAM and ANM secretomes. This would be important for the development of therapeutic strategies, since previous attempts to inhibit broad ranges of proteases such as MMPs have proven to be ineffective in cancer treatment (Overall and Kleifeld, 2006; Turk, 2006).

Finally, it could be useful to investigate whether gastric myofibroblasts could promote tumour growth in an *in vivo* xenograft model, and whether CAMs and ANM differ in this respect. Such a model could also provide an opportunity to study the functional significance of some of the proteins identified in this thesis, either through gene knock-down or overexpression in the myofibroblasts prior to injection. These *in vivo* models could be expanded to include transgenic models. For example, murine models are available with over-expression or knock-down of members of the uPA system such as PAI-1. Similarly, genetic manipulation of myofibroblasts prior to xenograft, targeting genes of interest such as βig-h3, TGFβ or members of the

uPA system, could provide insight to the importance of particular functional networks *in vivo*.

Appendix I

Mass Spectrometry

Appendix I

Mass spectometers

Mass spectrometers comprise three principal components; an ion source, a mass analyser which sorts ions by electrical and/or magnetic fields, and a detector. There are many variations on each of these three components, and there are a large number of possible combinations which can comprise a mass spectrometry system.

Ion Sources

Ion source technologies vary widely and have been developed for use with a range of analytes. Electro and chemical ionisation, for example, are used particularly in gas and vapour applications, where the analyte is ionised by chemical ion-molecule reactions during collisions in the source. When the application is analysing complex biological molecules, a so-called 'soft-ionisation' option is often preferable. These include electrospray ionisation (EI) (Fenn et al., 1989) and matrix-assisted laser desorption-ionisation (MALDI) (Karas and Hillenkamp, 1988).

EI works by forcing the analyte, dissolved in a volatile solvent, through a fine, charged capillary. The analyte exists in either anionic or cationic form, and expelled from the capillary as a results of charge interactions. Manipulation of the analyte often involves the addition/subtraction of H+, or the addition of Na+. The resulting aerosol then undergoes Coulombic fission; as the solvent evaporates, the charged analyte molecules are forced into ever closer proximity within the aerosol droplets, causing them to fragment until eventually lone ions occur. The development of EI represented a significant progression in the analysis of biological samples, as it provides a liquid to gas interface, allowing the in-line use of LC systems with an EI

equipped mass spectrometer. EI has the additional advantage that it overcomes the propensity of macromolecules to fragment when ionised.

MALDI uses a crystalline matrix to assist the sublimation and ionisation of the analyte. In proteomics application, MALDI is generally used with relatively simple peptide mixtures, whilst LC-EI is appropriate with more complex mixtures. MALDI can be used in conjunction with LC systems, though not in-line with MS, as the analyte fractions can be mixed with matrix and spotted onto a plate as they elute from the column. It does possess certain advantages over EI, however, as MALDI is less sensitive to salt concentrations and does not rely on a polar sample.

Mass Analysers

Mass analysers separate ions on the basis of their mass and charge, as described by the equation; $(m/z)a = E + (v \times B)$, where m = mass, z = charge, a = acceleration, E = electric field and v x B = the vector cross product of ion velocity and magnetic field. The manner in which they do this can vary widely between different classes of mass analyser.

Time of flight (TOF) mass analysers, as their name implies, use an electric field to accelerate the ions through the same potential and measure the time taken to reach the detector (Cotter, 1999). Quadrupole instruments function by oscillating electrical fields between four rods to selectively stabilise or destabilise ions with a particular m/z value (Yost and Enke, 1979). Thus only selected ions are able to pass through the quadrupole. Related to the quadrupole is the quadrupole ion trap in which electrical field are applied in all three dimensions, thus 'trapping' the selected ions with the analyser (March, 1997). Selected ions are trapped and sequentially ejected

in either a destructive or non-destructive manner. For example, in mass instability mode, the RF potential within the Ion Trap is ramped so that the orbit of ions with a mass greater than a particular value are stable, whilst other ions become unstable and ejected along the z-axis onto the detector. Ion Traps suffer from a low mass sensitivity due to the limited space at the centre of the instrument, which results in a space-charging distortion of trapped ions. A variation on this system, the Linear Ion Trap, largely overcomes this problem by trapping ions in a larger cylindrical volume rather than a 3D space, resulting in increased sensitivity, resolution and mass accuracy (Mayya et al., 2005; Schwartz et al., 2002). Fourier transform ion cyclotron resonance MS (FT-MS) is also an ion trapping MS, which occurs under high vacuum in a high magnetic field (Comisarow and Marshall, 1975). Ions are injected into a static electric/magnetic ion trap, and their electrical signal is measured by fixed detector as they cycle within the trap, producing a periodic signal whose frequency is directly related to the ions' mass and charge. FT-MS systems have very high sensitivity, mass accuracy, resolution and dynamic range, but unfortunately their application in proteomics is limited due to low peptide-fragmentation efficiency, high costs and high operational demands (Amster, 1996).

Detectors

MS detectors typically record either the charge induced or current produced when ions pass by or hit their surface. In most applications some form of electron multiplier is used since the number of ions entering the detector at any one time is usually very small.

Tandem MS Systems

Tandem MS refers to multiple rounds of MS, usually separated by some form of fragmentation. This can occur within a single mass analyser, such as in a Quadrupole Ion Trap, or multiple mass analyser of the same or different classes.

Quadrupole-time-of-flight (qTOF) instruments, known generically as qTOF or QqTOF, consist of a series of multipoles though which ions are focused, stored and selected prior to TOF analysis (Chernushevich et al., 2001). This has become a classic set up with a number of companies offering qTOF instruments. They benefit from the sensitivity and resolution of the TOF, whilst also being able to be run in-line with LC as the quadrupole makes them easily compatible with an ESI source, making them an excellent choice for discovery proteomics experiments.

Ion trap mass spectrometers have also found wide application in the field of proteomics. A major advantage of this technology is that the ion storing capability of the machine allows multiple rounds of MS to be performed (March, 1997). MS^3 experiments using ion-trap MS instruments have become an important tool in the analysis of protein phosphorylation. The ability to store ions within the trap also means that MS/MS spectra are of a high quality, giving good peptide sequence data. Historically, trap instruments have suffered from the handicap of poor resolution in the low mass range (Brancia, 2006). Although there have been significant improvements in this area, they are still fall short of TOF or triple quadrupole in this respect.

The orbitrap is a recently developed form of ion trap MS, but with fundamental differences from conventional ion traps and a certain commonalities with FT-MS (Hu et al., 2005). Unlike a conventional ion trap, moving ions are trapped in an

orbitrap by an electrostatic field (Hardman and Makarov, 2003). The initial tangential velocity of the ions results in them moving in an orbital pattern around the central electrode. The orbitrap utilises a Fourier transform to detect oscillation frequencies of ions, allowing for calculation of m/z values with a resolution that rivals FT-MS. The commercially available LTQ Orbitrap system (Thermo) features a linear ion trap, a C-trap, which is an RF-only quadrupole which accumulates and stores ions, and the orbitrap itself. This instrument allows the detection of ions with high resolution, low-ppm mass accuracy, and compatability with in-line nano-LC separation of samples (Scigelova and Makarov, 2006)

Appendix II

N-terminal COFRADIC

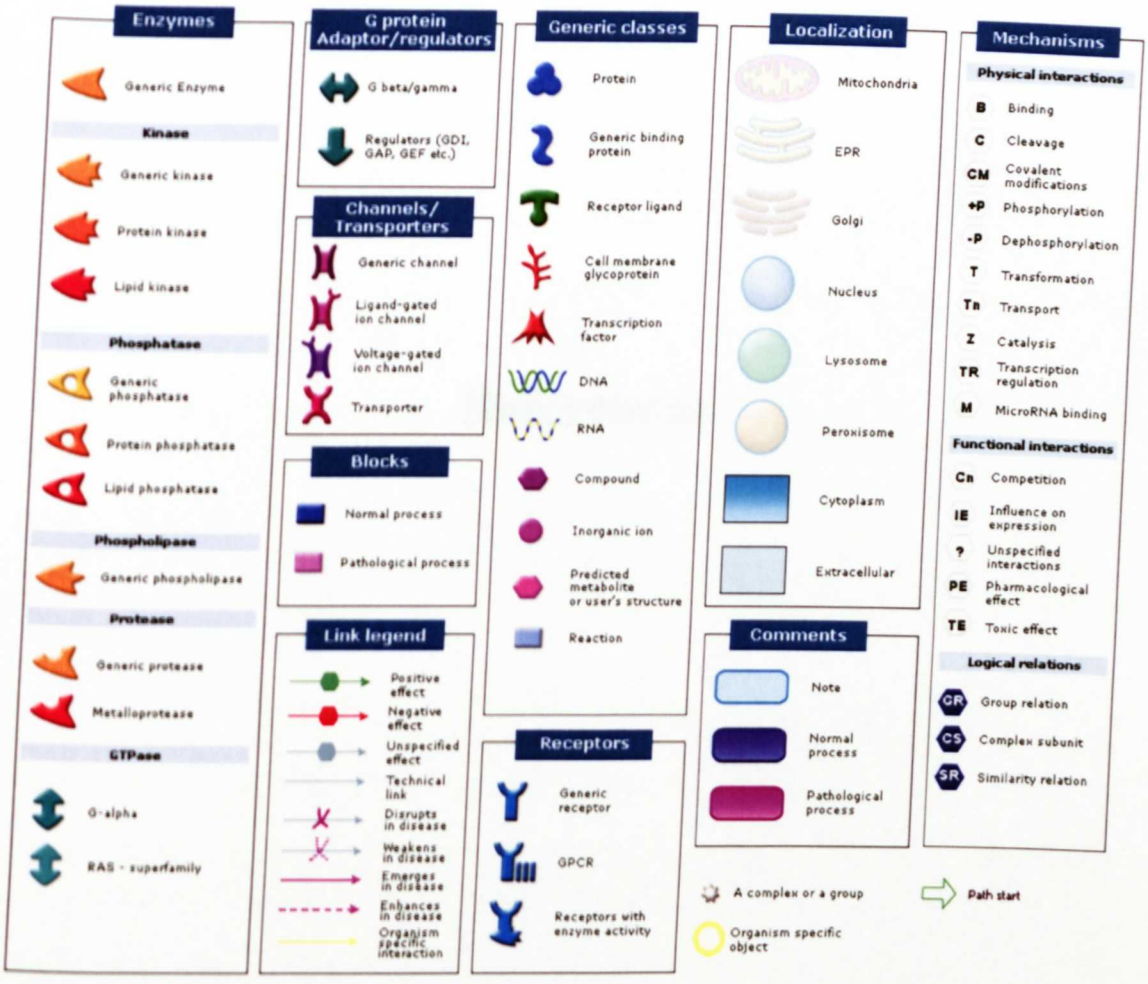
Appendix II

N-terminal COFRADIC

N-terminal COFRADIC uses conjugation of 2,4,6-trinitrobenzenesulphonic acid (TNBS) to peptides to increase their hydrophobicity, thus shifting their elution pattern in RP-LC (Gevaert et al., 2003). The N-termini of proteins are first blocked by an acetylation step in which acetyl groups are conjugated to free α- and ε-amines. This is followed by trypsinisation of the protein mixture. Due to the acetylation step trypsin will only cleave after arginine residues since the ε-amines of lysine residues also become acetylated, blocking the action of trypsin. The sample is then fractionated by the first RP-LC step. The collected peptides are then reacted with TNBS, with the result that all of the tryptic peptides become conjugated to TNBS, with the exception of the original N-terminal peptides which were blocked by acetlyation. The second RP-LC reveals the shift in hydrophobicity of the internal peptides, allowing the N-teminal peptides to be collected separately from the other peptides. In this way the complexity of the peptide mixture is greatly reduced, whilst ensuring that all of the original proteins are represented in the final enriched mixture. More interesting, however, is the potential to detect sites of proteolytic cleavage. Proteins which have been processed by proteases will not only have free N-termini at their first amino acid residue, but also at the site of each proteolytic cleavage.

The N-terminal COFRADIC protocol has been refined to improve enrichment and analytical power. One such improvement was the introduction of a second N-terminal peptide enrichment step prior to diagonal chromatography using strong cation exchange (SCX) (Staes et al., 2008). This SCX processing enriches the sample for N- and C-terminal peptides, since at an acidic pH the internal peptides which will have a charge of +1. N-terminal peptides will have a charge of 0 due to

the acetylated amino-terminus, +ve arginine residue and –ve COO^-. C-terminal peptides will also have a net charge of 0 due to the +ve NH_3 and the –ve COO^-. Thus the positively charged internal peptides will adhere to the SCX column and the N- and C-terminal peptides will pass through. In order to allow for accurate identification of N-terminal peptides, a further refinement was to use tri-deutero acetylation to block the α-amines, since this modification is easily distinguished from other acetyl groups which might exist in the sample.



Supplementary Figure 1.
Key for GeneGo interaction maps

This key describes the various components contained in GeneGo interaction maps, such as the one shown in Figure 4.6. Proteins are grouped into several generic classes, each of which are represented by different coloured objects. Interactions are described by arrows, and the nature of the interactions are indicated by the colour of the arrows.

References

References

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. Nature *422*, 198-207.

Akhurst, R. J., and Derynck, R. (2001). TGF-beta signaling in cancer - a double-edged sword. Trends Cell Biol *11*, S44-S51.

Albini, A., and Sporn, M. B. (2007). The tumour microenvironment as a target for chemoprevention. Nat Rev Cancer *7*, 139-147.

Alfano, D., Franco, P., Vocca, I., Gambi, N., Pisa, V., Mancini, A., Caputi, M., Carriero, M. V., Iaccarino, I., and Stoppelli, M. P. (2005). The urokinase plasminogen activator and its receptor - Role in cell growth and apoptosis. Thromb Haemost *93*, 205-211.

Alizadeh, A. A., Eisen, M. B., Davis, R. E., Ma, C., Lossos, I. S., Rosenwald, A., Boldrick, J. G., Sabet, H., Tran, T., Yu, X., *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature *403*, 503-511.

Allison, D. B., Cui, X. Q., Page, G. P., and Sabripour, M. (2006). Microarray data analysis: from disarray to consolidation and consensus. Nat Rev Genet *7*, 55-65.

Amster, I. J. (1996). Fourier transform mass spectrometry. J Mass Spectrom *31*, 1325-1337.

Anderson, A. R. A., Weaver, A. M., Cummings, P. T., and Quaranta, V. (2006). Tumor morphology and phenotypic evolution driven by selective pressure from the microenvironment. Cell *127*, 905-915.

Andreasen, P. A., Egelund, R., and Petersen, H. H. (2000). The plasminogen activation system in tumor growth, invasion, and metastasis. Cell Mol Life Sci *57*, 25-40.

Andreasen, P. A., Kjoller, L., Christensen, L., and Duffy, M. J. (1997). The urokinase-type plasminogen activator system in cancer metastasis: A review. Int J Cancer *72*, 1-22.

Anzano, M. A., Roberts, A. B., Meyers, C. A., Komoriya, A., Lamb, L. C., Smith, J. M., and Sporn, M. B. (1982). Synergistic interaction of 2 classes of transforming growth-factors from murine sarcoma-cells. Cancer Research *42*, 4776-4778.

Artavanis-Tsakonas, S., Rand, M. D., and Lake, R. J. (1999). Notch signaling: Cell fate control and signal integration in development. Science *284*, 770-776.

Balkwill, F., and Mantovani, A. (2001). Inflammation and cancer: back to Virchow? Lancet *357*, 539-545.

Baramova, E. N., Bajou, K., Remacle, A., Lhoir, C., Krell, H. W., Weidle, U. H., Noel, A., and Foidart, J. M. (1997). Involvement of PA/plasmin system in the processing of pro-MMP-9 and in the second step of pro-MMP-2 activation. FEBS Lett *405*, 157-162.

Barcellos-Hoff, M. H. (1998). The potential influence of radiation-induced microenvironments in neoplastic progression. J Mammary Gland Biol Neoplasia *3*, 165-175.

Barranco, S. C., Townsend, C. M., Casartelli, C., Macik, B. G., Burger, N. L., Boerwinkle, W. R., and Gourley, W. K. (1983). Establishment and characterization of an invitro model system for human adenocarcinoma of the stomach. Cancer Research *43*, 1703-1709.

Bayliss, W. M., and Starling, E. H. (1902). The mechanism of pancreatic secretion. J Physiol-London *28*, 325-353.

Becker, A. J., Till, J. E., and McCulloch, E. A. (1963). Cytological demonstration of clonal nature of spleen colonies derived from transplanted mouse marrow cells. Nature *197*, 452-&.

Bergers, G., Javaherian, K., Lo, K. M., Folkman, J., and Hanahan, D. (1999). Effects of angiogenesis inhibitors on multistage carcinogenesis in mice. Science *284*, 808-812.

Bhowmick, N. A., Chytil, A., Plieth, D., Gorska, A. E., Dumont, N., Shappell, S., Washington, M. K., Neilson, E. G., and Moses, H. L. (2004). TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. Science *303*, 848-851.

Bhowmick, N. A., Neilson, E. G., and Moses, H. L. (2004). Stromal fibroblasts in cancer initiation and progression. Nature *432*, 332-337.

Bierie, B., and Moses, H. L. (2006). TGF-beta and cancer. Cytokine Growth Factor Rev *17*, 29-40.

Bignell, G. R., Huang, J., Greshock, J., Watt, S., Butler, A., West, S., Grigorova, M., Jones, K. W., Wei, W., Stratton, M. R., *et al.* (2004). High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res *14*, 287-295.

Billings, P. C., Whitbeck, J. C., Adams, C. S., Abrams, W. R., Cohen, A. J., Engelsberg, B. N., Howard, P. S., and Rosenbloom, J. (2002). The transforming growth factor-beta-inducibie matrix protein beta ig-h3 interacts with fibronectin. J Biol Chem *277*, 28003-28009.

Bissell, M. J., Hall, H. G., and Parry, G. (1982). How does the extracellular-matrix direct gene-expression. J Theor Biol *99*, 31-68.

Bisson, C., Blacher, S., Polette, M., Blanc, J. F., Kebers, F., Desreux, J., Tetu, B., Rosenbaum, J., Foidart, J. M., Birembaut, P., and Noel, A. (2003). Restricted expression of membrane type 1-matrix metalloproteinase by myofibroblasts adjacent to human breast cancer cells. Int J Cancer *105*, 7-13.

Bjellqvist, B., Ek, K., Righetti, P. G., Gianazza, E., Gorg, A., Westermeier, R., and Postel, W. (1982). Isoelectric-focusing in immobilized ph gradients - principle, methodology and some applications. J Biochem Biophys Methods *6*, 317-339.

Black, R. A., Rauch, C. T., Kozlosky, C. J., Peschon, J. J., Slack, J. L., Wolfson, M. F., Castner, B. J., Stocking, K. L., Reddy, P., Srinivasan, S., *et al.* (1997). A metalloproteinase disintegrin that releases tumour-necrosis factor-alpha from cells. Nature *385*, 729-733.

Blagosklonny, M. V. (2002). Oncogenic resistance to growth-limiting conditions. Nat Rev Cancer *2*, 221-225.

Blasi, F. (1993). Urokinase and urokinase receptor - a paracrine autocrine system regulating cell-migration and invasiveness. Bioessays *15*, 105-111.

Boussioutas, A., Li, H., Liu, J., Waring, P., Lade, S., Holloway, A. J., Taupin, D., Gorringe, K., Haviv, I., Desmond, P. V., and Bowtell, D. D. L. (2003). Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. Cancer Research *63*, 2569-2577.

Brancia, F. L. (2006). Recent developments in ion-trap mass spectrometry and related technologies. Expert Rev Proteomics *3*, 143-151.

Breitling, R. (2006). Biological microarray interpretation: The rules of engagement. Biochim Biophys Acta-Gene Struct Expression *1759*, 319-327.

Brittan, M., and Wright, N. A. (2002). Gastrointestinal stem cells. J Pathol *197*, 492-509.

Buckley, C. D., Pilling, D., Henriquez, N. V., Parsonage, G., Threlfall, K., Scheel-Toellner, D., Simmons, D. L., Akbar, A. N., Lord, J. M., and Salmon, M. (1999). RGD peptides induce apoptosis by direct caspase-3 activation. Nature *397*, 534-539.

Campana, W. M., Li, X. Q., Dragojlovic, N., Janes, J., Gaultier, A., and Gonias, S. L. (2006). The low-density lipoprotein receptor-related protein is a pro-survival receptor in Schwann cells: Possible implications in peripheral nerve injury. J Neurosci *26*, 11197-11207.

Camps, J. L., Chang, S. M., Hsu, T. C., Freeman, M. R., Hong, S. J., Zhau, H. E., Voneschenbach, A. C., and Chung, L. W. K. (1990). Fibroblast-mediated acceleration of human epithelial tumor-growth invivo. Proc Natl Acad Sci U S A *87*, 75-79.

Carmeliet, P., and Jain, R. K. (2000). Angiogenesis in cancer and other diseases. Nature *407*, 249-257.

Carmeliet, P., Kieckens, L., Schoonjans, L., Ream, B., Vannuffelen, A., Prendergast, G., Cole, M., Bronson, R., Collen, D., and Mulligan, R. C. (1993). Plasminogen-activator inhibitor-1 gene deficient mice .1. generation by homologous recombination and characterization. J Clin Invest *92*, 2746-2755.

Carmeliet, P., Stassen, J. M., Schoonjans, L., Ream, B., Vandenoord, J. J., Demol, M., Mulligan, R. C., and Collen, D. (1993). Plasminogen-activator inhibitor-1 gene deficient mice .2. effects on hemostasis, thrombosis, and thrombolysis. J Clin Invest *92*, 2756-2760.

Carpenter, G., and Cohen, S. (1979). Epidermal growth-factor. Annu Rev Biochem *48*, 193-216.

Chari, N. S., and McDonnell, T. J. (2007). The sonic hedgehog signaling network in development and neoplasia. Adv Anat Pathol *14*, 344-352.

Chedid, M., Rubin, J. S., Csaky, K. G., and Aaronson, S. A. (1994). Regulation of keratinocyte growth-factor gene-expression by interleukin-1. J Biol Chem *269*, 10753-10757.

Chernushevich, I. V., Loboda, A. V., and Thomson, B. A. (2001). An introduction to quadrupole-time-of-flight mass spectrometry. J Mass Spectrom *36*, 849-865.

Chey, W. Y., and Chang, T. M. (2003). Secretin, 100 years later. J Gastroenterol *38*, 1025-1035.

Cohen, S. (1959). Purification and metabolic effects of a nerve growth-promoting protein from snake venom. J Biol Chem *234*, 1129-1137.

Cohen, S. (1962). Isolation of a mouse submaxillary gland protein accelerating incisor eruption and eyelid opening in new-born animal. J Biol Chem *237*, 1555-&.

Comisarow, M. B., and Marshall, A. G. (1975). Resolution-enhanced fourier-transform ion-cyclotron resonance spectroscopy. J Chem Phys *62*, 293-295.

Correa, P. (2004). The biological model of gastric carcinogenesis. IARC Sci Publ, 301-310.

Cotter, R. J. (1999). The new time-of-flight mass spectrometry. Anal Chem *71*, 445A-451A.

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

Cubellis, M. V., Wun, T. C., and Blasi, F. (1990). Receptor-mediated internalization and degradation of urokinase is caused by its specific inhibitor PAI-1. Embo J *9*, 1079-1085.

Cunha, G. R., Fujii, H., Neubauer, B. L., Shannon, J. M., Sawyer, L., and Reese, B. A. (1983). Epithelial-mesenchymal interactions in prostatic development .1. morphological observations of prostatic induction by urogenital sinus mesenchyme in epithelium of the adult rodent urinary-bladder. J Cell Biol *96*, 1662-1670.

Cunha, G. R., Hayward, S. W., and Wang, Y. Z. (2002). Role of stroma in carcinogenesis of the prostate. Differentiation *70*, 473-485.

Dale, H. H. (1914). The action of certain esters and ethers of choline, and their relation to muscarine. J Pharmacol Exp Ther *6*, 147-190.

Dallas, S. L., Rosser, J. L., Mundy, G. R., and Bonewald, L. F. (2002). Proteolysis of latent transforming growth factor-beta (TGF-beta)-binding protein-1 by osteoclasts - A cellular mechanism for release of TGF-beta from bone matrix. J Biol Chem *277*, 21352-21360.

Dano, K., Andreasen, P. A., Grondahlhansen, J., Kristensen, P., Nielsen, L. S., and Skriver, L. (1985). Plasminogen activators, tissue degradation, and cancer. Advances in Cancer Research *44*, 139-266.

Darmiento, J., Dicolandrea, T., Dalal, S. S., Okada, Y., Huang, M. T., Conney, A. H., and Chada, K. (1995). Collagenase expression in transgenic mouse skin causes hyperkeratosis and acanthosis and increases susceptibility to tumorigenesis. Mol Cell Biol *15*, 5732-5739.

Das, M., Miyakawa, T., Fox, C. F., Pruss, R. M., Aharonov, A., and Herschman, H. R. (1977). Specific radiolabeling of a cell-surface receptor for epidermal growth-factor. Proc Natl Acad Sci U S A *74*, 2790-2794.

Davie, E. W., and Neurath, H. (1955). Identification of a peptide released during autocatalytic activation of trypsinogen. J Biol Chem *212*, 515-529.

Davie, E. W., and Ratnoff, O. D. (1964). Waterfall sequence for intrinsic blood clotting. Science *145*, 1310-&.

De Wever, O., Nguyen, Q. D., Van Hoorde, L., Bracke, M., Bruyneel, E., Gespach, C., and Mareel, M. (2004). Tenascin-C and SF/HGF produced by myofibroblasts in vitro provide convergent proinvasive signals to human colon cancer cells through RhoA and Rac. Faseb J *18*, 1016-+.

De Wever, O., Westbroek, W., Verloes, A., Bloemen, N., Bracke, M., Gespach, C., Bruyneel, E., and Mareel, M. (2004). Critical role of N-cadherin in myofibroblast invasion and migration in vitro stimulated by colon-cancer-cell-derived TGF-beta or wounding. J Cell Sci *117*, 4691-4703.

Deryugina, E. I., Luo, G. X., Reisfeld, R. A., Bourdon, M. A., and Strongin, A. (1997). Tumor cell invasion through Matrigel is regulated by activated matrix metalloproteinase-2. Anticancer Res *17*, 3201-3210.

Dieudonne, S. C., Kerr, J. M., Xu, T., Sommer, B., DeRubeis, A. R., Kuznetsov, S. A., Kim, I. S., Gehron Robey, P., and Young, M. F. (1999). Differential display of human marrow stromal cells reveals unique mRNA expression patterns in response to dexamethasone. J Cell Biochem *76*, 231-243.

Dixon, M. F., Genta, R. M., Yardley, J. H., Correa, P., Batts, K. P., Dahms, B. B., Filipe, M. I., Haggitt, R. C., Haot, J., Hui, P. K., *et al.* (1996). Classification and grading of gastritis - The updated Sydney System. Am J Surg Pathol *20*, 1161-1181.

Dockray, G. J. (1999). Gastrin and gastric epithelial physiology. J Physiol-London *518*, 315-324.

Doucet, A., Butler, G. S., Rodriguez, D., Prudova, A., and Overall, C. M. (2008). Metadegradomics toward in vivo quantitative degradomics of proteolytic post-translational modifications of the cancer proteome. Mol Cell Proteomics *7*, 1925-1951.

Dvorak, H. F., Flier, J., and Frank, H. (1986). Tumors - wounds that do not heal - similarities between tumor stroma generation and wound-healing. N Engl J Med *315*, 1650-1659.

Egeblad, M., and Werb, Z. (2002). New functions for the matrix metalloproteinases in cancer progression. Nat Rev Cancer *2*, 161-174.

Elenbaas, B., and Weinberg, R. A. (2001). Heterotypic signaling between epithelial tumor cells and fibroblasts in carcinoma formation. Exp Cell Res *264*, 169-184.

Ellis, V., Scully, M. F., and Kakkar, V. V. (1989). Plasminogen activation initiated by single-chain urokinase-type plasminogen-activator - potentiation by u937 monocytes. J Biol Chem *264*, 2185-2188.

El-Omar, E. M., Carrington, M., Chow, W. H., McColl, K. E. L., Bream, J. H., Young, H. A., Herrera, J., Lissowska, J., Yuan, C. C., Rothman, N., *et al.* (2001). The role of interleukin-1 polymorphisms in the pathogenesis of gastric cancer (vol 404, pg 398, 2000). Nature *412*, 99-+.

Estep, J. M., O'Reilly, L., Grant, G., Piper, J., Jonsson, J., Afendy, A., Chandhoke, V., and Younossi, Z. M. (2009). Hepatic Stellate Cell and Myofibroblast-Like Cell Gene Expression in the Explanted Cirrhotic Livers of Patients Undergoing Liver Transplantation. Digestive diseases and sciences.

Faca, V. M., Ventura, A. P., Fitzgibbon, M. P., Pereira-Faca, S. R., Pitteri, S. J., Green, A. E., Ireton, R. C., Zhang, Q., Wang, H., O'Briant, K. C., *et al.* (2008). Proteomic analysis of ovarian cancer cells reveals dynamic processes of protein secretion and shedding of extra-cellular domains. PLoS One *3*, e2425.

Feiken, E., Romer, J., Eriksen, J., and Lund, L. R. (1995). Neutrophils express tumor necrosis factor-alpha during mouse skin wound healing. J Invest Dermatol *105*, 120-123.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989). Electrospray ionization for mass-spectrometry of large biomolecules. Science *246*, 64-71.

Figueiredo, C., Machado, J. C., Pharoah, P., Seruca, R., Sousa, S., Carvalho, R., Capelinha, A. F., Quint, W., Caldas, C., van Doorn, L. J., *et al.* (2002). Helicobacter pylori and interleukin 1 genotyping: An opportunity to identify high-risk individuals for gastric carcinoma. J Natl Cancer Inst *94*, 1680-1687.

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

Flanders, K. C., Ludecke, G., Engels, S., Cissel, D. S., Roberts, A. B., Kondaiah, P., Lafyatis, R., Sporn, M. B., and Unsicker, K. (1991). Localization and actions of transforming growth factor-beta-s in the embryonic nervous-system. Development *113*, 183-&.

Folkman, J. (1995). Angiogenesis in cancer, vascular, rheumatoid and other disease. Nat Med *1*, 27-31.

Folkman, J., and Klagsbrun, M. (1987). Angiogenic factors. Science *235*, 442-447.

Folkman, J., and Shing, Y. (1992). Angiogenesis. J Biol Chem *267*, 10931-10934.

Forman, D., Newell, D. G., Fullerton, F., Yarnell, J. W. G., Stacey, A. R., Wald, N., and Sitas, F. (1991). Association between infection with helicobacter-pylori and risk of gastric-cancer - evidence from a prospective investigation. Br Med J *302*, 1302-1305.

Fox, J. G., and Wang, T. C. (2001). Helicobacter pylori - not a good bug after all. N Engl J Med *345*, 829-832.

Fox, J. G., and Wang, T. C. (2007). Inflammation, atrophy, and gastric cancer. J Clin Invest *117*, 60-69.

Francisco, I. F. S., DeWolf, W. C., Peehl, D. M., and Olumi, A. F. (2004). Expression of transforming growth factor-beta 1 and growth in soft agar differentiate prostate carcinoma-associated fibroblasts from normal prostate fibroblasts. Int J Cancer *112*, 213-218.

Franke, W. W., Schmid, E., Osborn, M., and Weber, K. (1978). Different intermediate-sized filaments distinguished by immunofluorescence microscopy. Proc Natl Acad Sci U S A *75*, 5034-5038.

Frisch, S. M., and Francis, H. (1994). Disruption of epithelial cell-matrix interactions induces apoptosis. J Cell Biol *124*, 619-626.

Fuchs, E., Tumbar, T., and Guasch, G. (2004). Socializing with the neighbors: Stem cells and their niche. Cell *116*, 769-778.

Gabbiani, G. (1994). The cellular derivation and the life span of the myofibroblast. Paper presented at: Meeting of IIIrd EuroCellPath Workshop on Molecular Pathology of Tumors (Dalfsen, Netherlands, Gustav Fischer Verlag).

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res *31*, 3784-3788.

Gevaert, K., Ghesquiere, B., Staes, A., Martens, L., Van Damme, J., Thomas, G. R., and Vandekerckhove, J. (2004). Reversible labeling of cysteine-containing peptides allows their specific chromatographic isolation for non-gel proteome studies. Proteomics *4*, 897-908.

Gevaert, K., Goethals, M., Martens, L., Van Damme, J., Staes, A., Thomas, G. R., and Vandekerckhove, J. (2003). Exploring proteomes and analyzing protein processing by mass spectrometric identification of sorted N-terminal peptides. Nature Biotechnology *21*, 566-569.

Gevaert, K., Impens, F., Ghesquiere, B., Van Damme, P., Lambrechts, A., and Vandekerckhove, J. (2008). Stable isotopic labeling in proteomics. Proteomics *8*, 4873-4885.

Gevaert, K., Pinxteren, J., Demol, H., Hugelier, K., Staes, A., Van Damme, J., Martens, L., and Vandekerckhove, J. (2006). Four stage liquid chromatographic selection of methionyl peptides for peptide-centric proteome analysis: The proteome of human multipotent adult progenitor cells. Journal of Proteome Research *5*, 1415-1428.

Gevaert, K., Staes, A., Van Damme, J., De Groot, S., Hugelier, K., Demol, H., Martens, L., Goethals, M., and Vandekerckhove, J. (2005). Global phosphoproteome analysis on human HepG2 hepatocytes using reversed-phase diagonal LC. Proteomics *5*, 3589-3599.

Gevaert, K., Van Damme, J., Goethals, M., Thomas, G. R., Hoorelbeke, B., Demol, H., Martens, L., Puype, M., Staes, A., and Vandekerckhove, J. (2002). Chromatographic isolation of methionine-containing peptides for gel-free proteome analysis - Identification of more than 800 Escherichia coli proteins. Mol Cell Proteomics *1*, 896-903.

Gevaert, K., Van Damme, P., Ghesquiere, B., and Vandekerckhove, J. (2006). Protein processing and other modifications analyzed by diagonal peptide chromatography. Biochimica Et Biophysica Acta-Proteins and Proteomics *1764*, 1801-1810.

Ghesquiere, B., Buyl, L., Demol, H., Van Damme, J., Staes, A., Timmerman, E., Vandekerckhove, J., and Gevaert, K. (2007). A new approach for mapping sialylated N-glycosites in serum proteomes. Journal of Proteome Research *6*, 4304-4312.

Ghesquiere, B., Van Damme, J., Martens, L., Vandekerckhove, J., and Gevaert, K. (2006). Proteome-wide characterization of N-glycosylation events by diagonal chromatography. Journal of Proteome Research *5*, 2438-2447.

Gleave, M., Hsieh, J. T., Gao, C., Voneschenbach, A. C., and Chung, L. W. K. (1991). Acceleration of human prostate-cancer growth-invivo by factors produced by prostate and bone fibroblasts. Cancer Research *51*, 3753-3761.

Goldfarb, R. H., Murano, G., Brundage, R., Siegal, G. P., Terranova, V., Garbisa, S., and Liotta, L. A. (1986). Degradation of glycoprotein and collagenous components of the basement-membrane - studies with urokinase-type plasminogen-activator, alpha-thrombin, and plasmin. Semin Thromb Hemost *12*, 335-336.

Golub, T. R., Slonim, D. K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J. P., Coller, H., Loh, M. L., Downing, J. R., Caligiuri, M. A., *et al.* (1999). Molecular classification of cancer: Class discovery and class prediction by gene expression monitoring. Science *286*, 531-537.

Gorg, A., Weiss, W., and Dunn, M. J. (2004). Current two-dimensional electrophoresis technology for proteomics. Proteomics *4*, 3665-3685.

Gronborg, M., Kristiansen, T. Z., Iwahori, A., Chang, R., Reddy, R., Sato, N., Molina, H., Jensen, O. N., Hruban, R. H., Goggins, M. G., *et al.* (2004). Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. Paper presented at: 1st Annual Symposium on Proteome Fractionation (Cambridge, MA, Amer Soc Biochemistry Molecular Biology Inc).

Guilford, P., Hopkins, J., Harraway, J., McLeod, M., McLeod, N., Harawira, P., Taite, H., Scoular, R., Miller, A., and Reeve, A. E. (1998). E-cadherin germline mutations in familial gastric cancer. Nature *392*, 402-405.

Guo, Y. S., Beauchamp, R. D., Jin, G. F., Townsend, C. M., and Thompson, J. C. (1993). Insulin-like growth-factor binding-protein modulates the growth-response to insulin-like growth factor-i by human gastric-cancer cells. Gastroenterology *104*, 1595-1604.

Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. Proc Natl Acad Sci U S A *97*, 9390-9395.

Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., Gelb, M. H., and Aebersold, R. (1999). Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. Nature Biotechnology *17*, 994-999.

Hakomori, S. (2002). Glycosylation defining cancer malignancy: New wine in an old bottle. Proc Natl Acad Sci U S A *99*, 10231-10233.

Hanssen, E., Reinboth, B., and Gibson, M. A. (2003). Covalent and non-covalent interactions of beta ig-h3 with collagen VI - beta ig-h3 is covalently attached to the amino-terminal region of collagen VI in tissue microfibrils. J Biol Chem *278*, 24334-24341.

Hardman, M., and Makarov, A. A. (2003). Interfacing the orbitrap mass analyzer to an electrospray ion source. Anal Chem *75*, 1699-1705.

Hawsawi, N. M., Ghebeh, H., Hendrayani, S. F., Tulbah, A., Al-Eid, M., Al-Tweigeri, T., Ajarim, D., Alaiya, A., Dermime, S., and Aboussekhra, A. (2008). Breast carcinoma - Associated fibroblasts and their counterparts display neoplastic-specific changes. Cancer Research *68*, 2717-2725.

Hayward, S. W., Wang, Y. H., Cao, M., Hom, Y. K., Zhang, B. H., Grossfeld, G. D., Sudilovsky, D., and Cunha, G. R. (2001). Malignant transformation in a nontumorigenic human prostatic epithelial cell line. Cancer Research *61*, 8135-8142.

Hemers, E., Duval, C., McCaig, C., Handley, M., Dockray, G. J., and Varro, A. (2005). Insulin-like growth factor binding protein-5 is a target of matrix metalloproteinase-7: Implications for epithelial-mesenchymal signaling. Cancer Research *65*, 7363-7369.

Hintermann, E., and Quaranta, V. (2004). Epithelial cell motility on laminin-5: regulation by matrix assembly, proteolysis, integrins and erbB receptors. Matrix Biol *23*, 75-85.

Hinz, B., Phan, S. H., Thannickal, V. J., Galli, A., Bochaton-Piallat, M. L., and Gabbiani, G. (2007). The myofibroblast - One function, multiple origins. Am J Pathol *170*, 1807-1816.

Hollosi, P., Yakushij, J. K., Fong, K. S. K., Csiszar, K., and Fong, S. F. T. (2009). Lysyl oxidase-like 2 promotes migration in noninvasive breast cancer cells but not in normal breast epithelial cells. Int J Cancer *125*, 318-327.

Horvath, C., Melander, W., and Molnar, I. (1976). Solvophobic interactions in liquid-chromatography with nonpolar stationary phases. Journal of Chromatography *125*, 129-156.

Hotary, K., Allen, E., Punturieri, A., Yana, I., and Weiss, S. J. (2000). Regulation of cell invasion and morphogenesis in a three-dimensional type I collagen matrix by membrane-type matrix metalloproteinases 1, 2, and 3. J Cell Biol *149*, 1309-1323.

Houghton, J., Fox, J. G., and Wang, T. C. (2001). Gastric cancer: Laboratory bench to clinic. Paper presented at: 1st Asia Pacific Digestive Week (Sydney, Australia, Blackwell Publishing Asia).

Houghton, J., Morozov, A., Smirnova, I., and Wang, T. C. (2007). Stem cells and cancer. Semin Cancer Biol *17*, 191-203.

Houghton, J., Stoicov, C., Nomura, S., Rogers, A. B., Carlson, J., Li, H. C., Cai, X., Fox, J. G., Goldenring, J. R., and Wang, T. C. (2004). Gastric cancer originating from bone marrow-derived cells. Science *306*, 1568-1571.

Howlett, A. R., Bailey, N., Damsky, C., Petersen, O. W., and Bissell, M. J. (1995). Cellular growth and survival are mediated by beta-1 integrins in normal human breast epithelium but not in breast-carcinoma. J Cell Sci *108*, 1945-1957.

Hu, M., Yao, J., Carroll, D. K., Weremowicz, S., Chen, H., Carrasco, D., Richardson, A., Violette, S., Nikolskaya, T., Nikolsky, Y., *et al.* (2008). Regulation of in situ to invasive breast carcinoma transition. Cancer Cell *13*, 394-406.

Hu, Q. Z., Noll, R. J., Li, H. Y., Makarov, A., Hardman, M., and Cooks, R. G. (2005). The Orbitrap: a new mass spectrometer. J Mass Spectrom *40*, 430-443.

Hubner, G., Brauchle, M., Smola, H., Madlener, M., Fassler, R., and Werner, S. (1996). Differential regulation of pro-inflammatory cytokines during wound healing in normal and glucocorticoid-treated mice. Cytokine *8*, 548-556.

Hynes, R. O., and Yamada, K. M. (1982). Fibronectins - multifunctional modular glycoproteins. J Cell Biol *95*, 369-377.

Ichinose, A., Fujikawa, K., and Suyama, T. (1986). The activation of pro-urokinase by plasma kallikrein and its inactivation by thrombin. J Biol Chem *261*, 3486-3489.

Ijichi, H., Ikenoue, T., Kato, N., Mitsuno, Y., Togo, G., Kato, J., Kanai, F., Shiratori, Y., and Omata, M. (2001). Systematic analysis of the TGF-beta-Smad signaling pathway in gastrointestinal cancer cells. Biochem Biophys Res Commun *289*, 350-357.

Impens, F., Van Damme, P., Demol, H., Van Damme, J., Vandekerckhove, J., and Gevaert, K. (2008). Mechanistic insight into taxol-induced cell death. Oncogene *27*, 4580-4591.

Iozzo, R. V. (1998). Matrix proteoglycans: From molecular design to cellular function. Annu Rev Biochem *67*, 609-652.

Iozzo, R. V. (1999). The biology of the small leucine-rich proteoglycans - Functional network of interactive proteins. J Biol Chem *274*, 18843-18846.

Iozzo, R. V., and Murdoch, A. D. (1996). Proteoglycans of the extracellular environment: Clues from the gene and protein side offer novel perspectives in molecular diversity and function. Faseb J *10*, 598-614.

Irigoyen, J. P., Munoz-Canoves, P., Montero, L., Koziczak, M., and Nagamine, Y. (1999). The plasminogen activator system: biology and regulation. Cell Mol Life Sci *56*, 104-132.

Isaacson, P. G. (1994). Gastrointestinal lymphoma. Hum Pathol *25*, 1020-1029.

Ishizuka, J., Martinez, J., Townsend, C. M., and Thompson, J. C. (1991). The effect of gastrin on growth of human stomach-cancer cells. Paper presented at: 103rd Annual Scientific Session of the Southern Surgical Assoc (Hot Springs, Va, Lippincott-Raven Publ).

Jiang, L., Gonda, T. A., Gamble, M. V., Salas, M., Seshan, V., Tu, S., Twaddell, W. S., Hegyi, P., Lazar, G., Steele, I., *et al.* (2008). Global Hypomethylation of Genomic DNA in Cancer-Associated Myofibroblasts. Cancer Research *68*, 9900-9908.

Jones, S., Zhang, X. S., Parsons, D. W., Lin, J. C. H., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Kamiyama, H., Jimeno, A., *et al.* (2008). Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science *321*, 1801-1806.

Jorissen, R. N., Walker, F., Pouliot, N., Garrett, T. P. J., Ward, C. W., and Burgess, A. W. (2003). Epidermal growth factor receptor: mechanisms of activation and signalling. Exp Cell Res *284*, 31-53.

Kadler, K. E., Baldock, C., Bella, J., and Boot-Handford, R. P. (2007). Collagens at a glance. J Cell Sci *120*, 1955-1958.

Kalluri, R. (2003). Basement membranes: Structure, assembly and role in tumour angiogenesis. Nat Rev Cancer *3*, 422-433.

Kapur, K., Xing, Y., Ouyang, Z., and Wong, W. H. (2007). Exon arrays provide accurate assessments of gene expression. Genome Biol *8*, 25.

Karam, S. M., and Leblond, C. P. (1993). Dynamics of epithelial-cells in the corpus of the mouse stomach .1. Identification of proliferative cell-types and pinpointing of the stem-cell. Anat Rec *236*, 259-279.

Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons. Anal Chem *60*, 2299-2301.

Karnes, W. E., Samloff, I. M., Siurala, M., Kekki, M., Sipponen, P., Kim, S. W. R., and Walsh, J. H. (1991). Positive serum antibody and negative tissue-staining for helicobacter-pylori in subjects with atrophic body gastritis. Gastroenterology *101*, 167-174.

Kasai, S., Arimura, H., Nishida, M., and Suyama, T. (1985). Primary structure of single-chain pro-urokinase. J Biol Chem *260*, 2382-2389.

Kasai, S., Arimura, H., Nishida, M., and Suyama, T. (1985). Proteolytic cleavage of single-chain pro-urokinase induces conformational change which follows activation of the zymogen and reduction of its high-affinity for fibrin. J Biol Chem *260*, 2377-2381.

Katuri, V., Tang, Y., Marshall, B., Rashid, A., Jogunoori, W., Volpe, E. A., Sidawy, A. N., Evans, S., Blay, J., Gallicano, G. I., *et al.* (2005). Inactivation of ELF/TGF-beta signaling in human gastrointestinal cancer. Oncogene *24*, 8012-8024.

Kazes, I., Delarue, F., Hagege, J., Bouzhir-Sima, L., Rondeau, E., Sraer, J. D., and Nguyen, G. (1998). Soluble latent membrane-type 1 matrix metalloprotease secreted by human mesangial cells is activated by urokinase. Kidney Int *54*, 1976-1984.

Kenny, S., Duval, C., Sammut, S. J., Steele, I., Pritchard, D. M., Atherton, J. C., Argent, R. H., Dimaline, R., Dockray, G. J., and Varro, A. (2008). Increased expression of the urokinase plasminogen activator system by Helicobacter pylori in gastric epithelial cells. American Journal of Physiology-Gastrointestinal and Liver Physiology *295*, G431-G441.

Keskioja, J., Lohi, J., Tuuttila, A., Tryggvason, K., and Vartio, T. (1992). Proteolytic processing of the 72,000-da type-iv collagenase by urokinase plasminogen-activator. Exp Cell Res *202*, 471-476.

Khwaja, F. W., Svoboda, P., Reed, M., Pohl, J., Pyrzynska, B., and Van Meir, E. G. (2006). Proteomic identification of the wt-p53-regulated tumor cell secretome. Oncogene *25*, 7650-7661.

Kim, J. E., Jeong, H. W., Nam, J. O., Lee, B. H., Choi, J. Y., Park, R. W., Park, J. Y., and Kim, I. S. (2002). Identification of motifs in the fasciclin domains of the transforming growth factor-beta-induced matrix protein beta ig-h3 that interact with the alpha v65 integrin. J Biol Chem *277*, 46159-46165.

Kim, J. E., Kim, E. H., Han, E. H., Park, R. W., Park, I. H., Jun, S. H., Kim, J. C., Young, M. F., and Kim, I. S. (2000). A TGF-beta-inducible cell adhesion molecule, beta ig-h3, is downregulated in melorheostosis and involved in osteogenesis. J Cell Biochem *77*, 169-178.

Kim, J. E., Kim, S. J., Jeong, H. W., Lee, B. H., Choi, J. Y., Park, R. W., Park, J. Y., and Kim, I. S. (2003). RGD peptides released from beta ig-h3, a TGF-beta-induced cell-adhesive molecule, mediate apoptosis. Oncogene *22*, 2045-2053.

Kim, J. E., Kim, S. J., Lee, B. H., Park, R. W., Kim, K. S., and Kim, I. S. (2000). Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-beta-induced gene, beta ig-h3. J Biol Chem *275*, 30907-30915.

Kobayashi, H., Schmitt, M., Goretzki, L., Chucholowski, N., Calvete, J., Kramer, M., Gunzler, W. A., Janicke, F., and Graeff, H. (1991). Cathepsin-b efficiently activates the soluble and the tumor-cell receptor-bound form of the proenzyme urokinase-type plasminogen-activator (pro-uPA). J Biol Chem *266*, 5147-5152.

Krtolica, A., Parrinello, S., Lockett, S., Desprez, P. Y., and Campisi, J. (2001). Senescent fibroblasts promote epithelial cell growth and tumorigenesis: A link between cancer and aging. Proc Natl Acad Sci U S A *98*, 12072-12077.

Kulasingam, V., and Diamandis, E. P. (2007). Proteomics analysis of conditioned media from three breast cancer cell lines. Mol Cell Proteomics *6*, 1997-2011.

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. Proc Natl Acad Sci U S A *101*, 4966-4971.

Lagergren, J., Bergstrom, R., Lindgren, A., and Nyren, O. (1999). Symptomatic gastroesophageal reflux as a risk factor for esophageal adenocarcinoma. N Engl J Med *340*, 825-831.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., *et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Larsson, O., Diebold, D., Fan, D., Peterson, M., Nho, R. S., Bitterman, P. B., and Henke, C. A. (2008). Fibrotic myofibroblasts manifest genome-wide derangements of translational control. PLoS One *3*, e3220.

Lauren, P. (1965). 2 histological main types of gastric carcinoma - diffuse and so-called intestinal-type carcinoma - an attempt at a histo-clinical classification. Acta Pathologica Et Microbiologica Scandinavica *64*, 31-&.

Lawlor, K., Nazarlan, A., Lacomis, L., Tempst, P., and Villanueva, J. (2009). Pathway-Based Biomarker Search by High-Throughput Proteomics Profiling of Secretomes. Journal of Proteome Research *8*, 1489-1503.

Lebaron, R. G., Bezverkov, K. I., Zimber, M. P., Pavelec, R., Skonier, J., and Purchio, A. F. (1995). Beta-ig-h3, a novel secretory protein inducible by transforming growth-factor-beta, is present in normal skin and promotes the adhesion and spreading of dermal fibroblasts in-vitro. J Invest Dermatol *104*, 844-849.

Lee, S. L., Dickson, R. B., and Lin, C. Y. (2000). Activation of hepatocyte growth factor and urokinase/plasminogen activator by matriptase, an epithelial membrane serine protease. J Biol Chem *275*, 36720-36725.

Leedham, S. J., Brittan, M., Preston, S. L., McDonald, S. A. C., and Wright, N. A. (2006). The stomach periglandular fibroblast sheath: all present and correct. Gut *55*, 295-296.

Lengauer, C., Kinzler, K. W., and Vogelstein, B. (1998). Genetic instabilities in human cancers. Nature *396*, 643-649.

Levimontalcini, R. (1987). The nerve growth-factor 35 years later. Science *237*, 1154-1162.

Levimontalcini, R., and Hamburger, V. (1951). Selective growth stimulating effects of mouse sarcoma on the sensory and sympathetic nervous system of the chick embryo. J Exp Zool *116*, 321-361.

Levimontalcini, R., Meyer, H., and Hamburger, V. (1954). Invitro experiments on the effects of mouse sarcomas-180 and 37 on the spinal and sympathetic ganglia of the chick embryo. Cancer Research *14*, 49-&.

Liotta, L. A., Goldfarb, R. H., Brundage, R., Siegal, G. P., Terranova, V., and Garbisa, S. (1981). Effect of plasminogen-activator (urokinase), plasmin, and thrombin on glycoprotein and collagenous components of basement-membrane. Cancer Research *41*, 4629-4636.

Liotta, L. A., Goldfarb, R. H., and Terranova, V. P. (1981). Cleavage of laminin by thrombin and plasmin - alpha-thrombin selectively cleaves the beta-chain of laminin. Thromb Res *21*, 663-673.

Littlepage, L. E., Egeblad, M., and Werb, Z. (2005). Coevolution of cancer and stromal cellular responses. Cancer Cell *7*, 499-500.

Liu, C. G., Calin, G. A., Volinia, S., and Croce, C. M. (2008). MicroRNA expression profiling using microarrays. Nat Protoc *3*, 563-578.

Liu, D., Ghiso, J. A. A., Estrada, Y., and Ossowski, L. (2002). EGFR is a transducer of the urokinase receptor initiated signal that is required for in vivo growth of a human carcinoma. Cancer Cell *1*, 445-457.

Loewi, O. (1921). Humoral transferability of the heart nerve effect. I. Announcement. Pflugers Archiv Fur Die Gesamte Physiologie Des Menschen Und Der Tiere *189*, 239-242.

Lopez-Otin, C., and Overall, C. M. (2002). Protease degradomics: A new challenge for proteomics. Nat Rev Mol Cell Biol *3*, 509-519.

Lund, L. R., Riccio, A., Andreasen, P. A., Nielsen, L. S., Kristensen, P., Laiho, M., Saksela, O., Blasi, F., and Dano, K. (1987). Transforming growth-factor-beta is a strong and fast acting positive regulator of the level of type-1 plasminogen-activator inhibitor messenger-rna in wi-38 human-lung fibroblasts. Embo J *6*, 1281-1286.

Lund, L. R., Romer, J., Bugge, T. H., Nielsen, B. S., Frandsen, T. L., Degen, J. L., Stephens, R. W., and Dano, K. (1999). Functional overlap between two classes of matrix-degrading proteases in wound healing. Embo J *18*, 4645-4656.

Luster, A. D. (1998). Chemokines - Chemotactic cytokines that mediate inflammation. N Engl J Med *338*, 436-445.

Ma, C., Rong, Y., Radiloff, D. R., Datto, M. B., Centeno, B., Bao, S., Cheng, A. W. M., Lin, F., Jiang, S., Yeatman, T. J., and Wang, X. F. (2008). Extracellular matrix protein beta ig-h3/TGFBI promotes metastasis of colon cancer by enhancing cell extravasation. Genes Dev *22*, 308-321.

Ma, L., Teruya-Feldstein, J., and Weinberg, R. A. (2007). Tumour invasion and metastasis initiated by microRNA 10b in breast cancer. Nature *449*, 682-U682.

Macfarlane, R. G. (1964). Enzyme cascade in blood clotting mechanism + its function as biochemical amplifier. Nature *202*, 498-&.

Mann, M., Hendrickson, R. C., and Pandey, A. (2001). Analysis of proteins and proteomes by mass spectrometry. Annu Rev Biochem *70*, 437-473.

Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. Nature Biotechnology *21*, 255-261.

March, R. E. (1997). An introduction to quadrupole ion trap mass spectrometry. J Mass Spectrom *32*, 351-369.

Mars, W. M., Zarnegar, R., and Michalopoulos, G. K. (1993). Activation of hepatocyte growth-factor by the plasminogen activators uPA and tPA. Am J Pathol *143*, 949-958.

Marshall, B. J., and Warren, J. R. (1984). Unidentified curved bacilli in the stomach of patients with gastritis and peptic-ulceration. Lancet *1*, 1311-1315.

Martens, L., Van Damme, P., Van Damme, J., Staes, A., Timmerman, E., Ghesquiere, B., Thomas, G. R., Vandekerckhove, J., and Gevaert, K. (2005). The human platelet proteome mapped by peptide-centric proteomics: A functional protein profile. Proteomics *5*, 3193-3204.

Marusyk, A., and DeGregori, J. (2008). Declining cellular fitness with age promotes cancer initiation by selecting for adaptive oncogenic mutations. Biochim Biophys Acta-Rev Cancer *1785*, 1-11.

Matsukura, N., Suzuki, K., Kawachi, T., Aoyagi, M., Sugimura, T., Kitaoka, H., Numajiri, H., Shirota, A., Itabashi, M., and Hirota, T. (1980). Distribution of marker enzymes and mucin in intestinal metaplasia in human stomach and relation of complete and incomplete types of intestinal metaplasia to minute gastric carcinomas. J Natl Cancer Inst *65*, 231-240.

May, M. (2009). From cells, secrets of the secretome leak out. Nat Med *15*, 828.

Mayya, V., Rezaul, K., Cong, Y. S., and Han, D. (2005). Systematic comparison of a two-dimensional ion trap and a three-dimensional ion trap mass spectrometer in proteomics. Mol Cell Proteomics *4*, 214-223.

**McCaig, C., Duval, C., Hemers, E., Steele, I., Pritchard, D. M., Przemeck, S., Dimaline, R., Ahmed, S., Bodger, K., Kerrigan, D. D., *et al.* (2006). The role of matrix metalloproteinase-7 in redefining the gastric microenvironment in response to Helicobacter pylori. Gastroenterology *130*, 1754-1763.

**McDonald, S. A. C., Greaves, L. C., Gutierrez-Gonzalez, L., Rodriguez-Justo, M., Deheragoda, M., Leedham, S. J., Taylor, R. W., Lee, C. Y., Preston, S. L., Lovell, M., *et al.* (2008). Mechanisms of field cancerization in the human stomach: The expansion and spread of mutated gastric stem cells. Gastroenterology *134*, 500-510.

McQuibban, G. A., Gong, J. H., Tam, E. M., McCulloch, C. A. G., Clark-Lewis, I., and Overall, C. M. (2000). Inflammation dampened by gelatinase A cleavage of monocyte chemoattractant protein-3. Science *289*, 1202-1206.

Meltzer, P. S. (2005). Cancer genomics - Small RNAs with big impacts. Nature *435*, 745-746.

Merlo, L. M., Pepper, J. W., Reid, B. J., and Maley, C. C. (2006). Cancer as an evolutionary and ecological process. Nature reviews *6*, 924-935.

Miettinen, M., and Lasota, J. (2001). Gastrointestinal stromal tumors - definition, clinical, histological, immunohistochemical, and molecular genetic features and differential diagnosis. Virchows Arch Int J Pathol *438*, 1-12.

Mirgorodskaya, O. A., Kozmin, Y. P., Titov, M. I., Korner, R., Sonksen, C. P., and Roepstorff, P. (2000). Quantitation of peptides and proteins by matrix-assisted laser desorption/ionization mass spectrometry using O-18-labeled internal standards. Rapid Commun Mass Spectrom *14*, 1226-1232.

Mochan, E., and Keler, T. (1984). Plasmin degradation of cartilage proteoglycan. Biochimica Et Biophysica Acta *800*, 312-315.

Modlin, I. M., Kidd, M., and Farhadi, J. (2000). Bayliss and Starling and the nascence of endocrinology. Regul Pept *93*, 109-123.

Modlin, I. M., Kidd, M., Lye, K. D., and Wright, N. A. (2003). Gastric stem cells: an update. Keio J Med *52*, 134-137.

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 Research *60*, 2562-2566.

Montel, V., Gaultier, A., Lester, R. D., Campana, W. M., and Gonias, S. L. (2007). The low-density lipoprotein receptor-related protein regulates cancer cell survival and metastasis development. Cancer Research *67*, 9817-9824.

Moore, K. A., and Lemischka, I. R. (2006). Stem cells and their niches. Science *311*, 1880-1885.

Moss, M. L., Jin, S. L. C., Milla, M. E., Burkhart, W., Carter, H. L., Chen, W. J., Clay, W. C., Didsbury, J. R., Hassler, D., Hoffman, C. R., *et al.* (1997). Cloning of a disintegrin metalloproteinase that processes precursor tumour-necrosis factor-alpha. Nature *385*, 733-736.

Munier, F. L., Korvatska, E., Djemai, A., LePaslier, D., Zografos, L., Pescia, G., and Schorderet, D. F. (1997). Kerato-epithelin mutations in four 5q31-linked corneal dystrophies. Nature Genet *15*, 247-251.

Nakayama, H., Enzan, H., Miyazaki, E., Kuroda, N., Naruse, K., Kiyoku, H., and Hiroi, M. (2000). Myofibroblasts at the tumor border of invasive gastric carcinomas: With special reference to histological type and tumor depth. Oncol Rep *7*, 1011-1015.

Naldini, L., Tamagnone, L., Vigna, E., Sachs, M., Hartmann, G., Birchmeier, W., Daikuhara, Y., Tsubouchi, H., Blasi, F., and Comoglio, P. M. (1992). Extracellular proteolytic cleavage by urokinase is required for activation of hepatocyte growth-factor scatter factor. Embo J *11*, 4825-4833.

Nam, E. J., Sa, K. H., You, D. W., Cho, J. H., Seo, J. S., Han, S. W., Park, J. Y., Kim, S. I., Kyung, H. S., Kim, I. S., and Kang, Y. M. (2006). Up-regulated transforming growth factor beta-inducible gene h3 in rheumatoid arthritis mediates adhesion and migration of synoviocytes through alpha v beta 3 integrin - Regulation by cytokines. Arthritis Rheum *54*, 2734-2744.

Nam, J. O., Jeong, H. W., Lee, B. H., Park, R. W., and Kim, I. S. (2005). Regulation of tumor angiogenesis by fastatin, the fourth FAS1 domain of beta ig-h3 via alpha v beta 3 integrin. Cancer Research *65*, 4153-4161.

Nam, J. O., Kim, J. E., Jeong, H. W., Lee, S. J., Lee, B. H., Choi, J. Y., Park, R. W., Park, J. Y., and Kim, I. S. (2003). Identification of the alpha(v)beta(3) integrin-interacting motif of beta ig-h3 and its anti-angiogenic effect. J Biol Chem *278*, 25902-25909.

Nielsen, L. S., Hansen, J. G., Skriver, L., Wilson, E. L., Kaltoft, K., Zeuthen, J., and Dano, K. (1982). Purification of zymogen to plasminogen-activator from human glioblastoma cells by affinity-chromatography with monoclonal-antibody. Biochemistry *21*, 6410-6415.

Noble, D. (2008). Claude Bernard, the first systems biologist, and the future of physiology. Exp Physiol *93*, 16-26.

Nomura, A., Stemmermann, G. N., Chyou, P. H., Kato, I., Perezperez, G. I., and Blaser, M. J. (1991). Helicobacter-pylori infection and gastric-carcinoma among japanese-americans in hawaii. N Engl J Med *325*, 1132-1136.

Ofarrell, P. H. (1975). High-resolution 2-dimensional electrophoresis of proteins. J Biol Chem *250*, 4007-4021.

Ohno, S., Noshiro, M., Makihira, S., Kawamoto, T., Shen, M., Yan, W. Q., Kawashima-Ohya, Y., Fujimoto, K., Tanne, K., and Kato, Y. (1999). RGD-CAP (beta ig-h3) enhances the spreading of chondrocytes and fibroblasts via integrin alpha(1)beta(1). Biochim Biophys Acta-Mol Cell Res *1451*, 196-205.

Ohuchida, K., Mizumoto, K., Murakami, M., Qian, L. W., Sato, N., Nagai, E., Matsumoto, K., Nakamura, T., and Tanaka, M. (2004). Radiation to stromal fibroblasts increases invasiveness of pancreatic cancer cells through tumor-stromal interactions. Cancer Research *64*, 3215-3222.

Olsen, J. V., de Godoy, L. M. F., Li, G. Q., Macek, B., Mortensen, P., Pesch, R., Makarov, A., Lange, O., Horning, S., and Mann, M. (2005). Parts per million mass accuracy on an orbitrap mass spectrometer via lock mass injection into a C-trap. Mol Cell Proteomics *4*, 2010-2021.

Olson, D., Pollanen, J., Hoyerhansen, G., Ronne, E., Sakaguchi, K., Wun, T. C., Appella, E., Dano, K., and Blasi, F. (1992). Internalization of the urokinase-plasminogen activator inhibitor type-1 complex is mediated by the urokinase receptor. J Biol Chem *267*, 9129-9133.

Olumi, A. F., Grossfeld, G. D., Hayward, S. W., Carroll, P. R., Tisty, T. D., and Cunha, G. R. (1999). Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. Cancer Research *59*, 5002-5011.

Olumi, A. F., Hayward, S. W., Grossfeld, G. D., Carroll, P. R., Cunha, G. R., and Tlsty, T. D. (1998). Human prostatic carcinoma-associated fibroblasts enhance proliferation and inhibit death of prostatic epithelial cells. J Urol *159*, 12.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., and Mann, M. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics *1*, 376-386.

Oreilly, M. S., Holmgren, L., Chen, C., and Folkman, J. (1996). Angiostatin induces and sustains dormancy of human primary tumors in mice. Nat Med *2*, 689-692.

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.

Overall, C. M. (2004). Dilating the degradome: matrix metalloproteinase 2 (MMP-2) cuts to the heart of the matter. Biochem J *383*, e5-7.

Overall, C. M., and Kleifeld, O. (2006). Tumour microenvironment - Opinion - Validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy. Nat Rev Cancer *6*, 227-239.

Overall, C. M., and Lopez-Otin, C. (2002). Strategies for MMP inhibition in cancer: Innovations for the post-trial era. Nat Rev Cancer *2*, 657-672.

Pankov, R., and Yamada, K. M. (2002). Fibronectin at a glance. J Cell Sci *115*, 3861-3863.

Parks, W. C., Wilson, C. L., and Lopez-Boado, Y. S. (2004). Matrix metalloproteinases as modulators of inflammation and innate immunity. Nat Rev Immunol *4*, 617-629.

Parrinello, S., Coppe, J. P., Krtolica, A., and Campisi, J. (2005). Stromal-epithelial interactions in aging and cancer: senescent fibroblasts alter epithelial cell differentiation. J Cell Sci *118*, 485-496.

Parsonnet, J., Friedman, G. D., Vandersteen, D. P., Chang, Y., Vogelman, J. H., Orentreich, N., and Sibley, R. K. (1991). Helicobacter-pylori infection and the risk of gastric-carcinoma. N Engl J Med *325*, 1127-1131.

Parsons, D. W., Jones, S., Zhang, X. S., Lin, J. C. H., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., *et al.* (2008). An integrated genomic analysis of human glioblastoma Multiforme. Science *321*, 1807-1812.

Paulitschke, V., Kunstfeld, R., Mohr, T., Slany, A., Micksche, M., Drach, J., Zielinski, C., Pehamberger, H., and Gerner, C. (2009). Entering a New Era of Rational Biomarker Discovery for Early Detection of Melanoma Metastases: Secretome Analysis of Associated Stroma Cells. Journal of Proteome Research *8*, 2501-2510.

Pedersen, M. O., Larsen, A., Stoltenberg, M., and Penkowa, M. (2009). The role of metallothionein in oncogenesis and cancer prognosis. Prog Histochem Cytochem *44*, 29-64.

Peek, R. M., and Blaser, M. J. (2002). Helicobacter pylori and gastrointestinal tract adenocarcinomas. Nat Rev Cancer *2*, 28-37.

Pepper, M. S. (2001). Role of the matrix metalloproteinase and plasminogen activator-plasmin systems in angiogenesis. Arterioscler Thromb Vasc Biol *21*, 1104-1117.

Peschon, J. J., Slack, J. L., Reddy, P., Stocking, K. L., Sunnarborg, S. W., Lee, D. C., Russell, W. E., Castner, B. J., Johnson, R. S., Fitzner, J. N., *et al.* (1998). An essential role for ectodomain shedding in mammalian development. Science *282*, 1281-1284.

Petersen, L. C., Johannessen, M., Foster, D., Kumar, A., and Mulvihill, E. (1988). The effect of polymerized fibrin on the catalytic activities of one-chain tissue-type plasminogen-activator as revealed by an analog resistant to plasmin cleavage. Biochimica Et Biophysica Acta *952*, 245-254.

Pittenger, M. F., Mackay, A. M., Beck, S. C., Jaiswal, R. K., Douglas, R., Mosca, J. D., Moorman, M. A., Simonetti, D. W., Craig, S., and Marshak, D. R. (1999). Multilineage potential of adult human mesenchymal stem cells. Science *284*, 143-147.

Polyak, K., Haviv, I., and Campbell, I. G. (2009). Co-evolution of tumor cells and their microenvironment. Trends Genet *25*, 30-38.

Powell, D. W., Adegboyega, P. A., Di Mari, J. F., and Mifflin, R. C. (2005). Epithelial cells and their neighbors I. Role of intestinal myofibroblasts in development, repair, and cancer. American Journal of Physiology-Gastrointestinal and Liver Physiology *289*, G2-G7.

Powell, D. W., Mifflin, R. C., Valentich, J. D., Crowe, S. E., Saada, J. I., and West, A. B. (1999). Myofibroblasts. I. Paracrine cells important in health and disease. American Journal of Physiology-Cell Physiology *277*, C1-C19.

Powell, D. W., Mifflin, R. C., Valentich, J. D., Crowe, S. E., Saada, J. I., and West, A. B. (1999). Myofibroblasts. II. Intestinal subepithelial myofibroblasts. American Journal of Physiology-Cell Physiology *277*, C183-C201.

Prager, G. W., Breuss, J. M., Steurer, S., Mihaly, J., and Binder, B. R. (2004). Vascular endothelial growth factor (VEGF) induces rapid prourokinase (pro-uPA) activation on the surface of endothelial cells. Blood *103*, 955-962.

Prockop, D. J., and Kivirikko, K. I. (1995). Collagens - molecular-biology, diseases, and potentials for therapy. Annu Rev Biochem *64*, 403-434.

Pruss, R. M., and Herschman, H. R. (1977). Variants of 3t3 cells lacking mitogenic response to epidermal growth-factor - (receptors mitogens colchicine selection). Proc Natl Acad Sci U S A *74*, 3918-3921.

Pyke, C., Kristensen, P., Ralfkiaer, E., Grondahlhansen, J., Eriksen, J., Blasi, F., and Dano, K. (1991). Urokinase-type plasminogen-activator is expressed in stromal cells and its receptor in cancer-cells at invasive foci in human colon adenocarcinomas. Am J Pathol *138*, 1059-1067.

Quante, M., and Wang, T. C. (2008). Inflammation and Stem Cells in Gastrointestinal Carcinogenesis. Physiology *23*, 350-359.

Quax, P. H. A., Pedersen, N., Masucci, M. T., Weeningverhoeff, E. J. D., Dano, K., Verheijen, J. H., and Blasi, I. F. (1991). Complementation between urokinase-producing and receptor-producing cells in extracellular-matrix degradation. Cell Regulation *2*, 793-803.

Rawe, I. M., Zhan, Q. A., Burrows, R., Bennett, K., and Cintron, C. (1997). Beta-ig - Molecular cloning and in situ hybridization in corneal tissues. Invest Ophthalmol Vis Sci *38*, 893-900.

Rawlings, N. D., Morton, F. R., and Barrett, A. J. (2006). MEROPS: the peptidase database. Nucleic Acids Res *34*, D270-D272.

Reinboth, B., Thomas, J., Hanssen, E., and Gibson, M. A. (2006). beta ig-h3 interacts directly with biglycan and decorin, promotes collagen VI aggregation, and participates in ternary complexing with these macromolecules. J Biol Chem *281*, 7816-7824.

RemacleBonnet, M. M., Garrouste, F. L., and Pommier, G. J. (1997). Surface-bound plasmin induces selective proteolysis of insulin-like-growth-factor (IGF)-binding protein-4 (IGFBP-4) and promotes autocrine IGF-II bio-availability in human colon-carcinoma cells. Int J Cancer *72*, 835-843.

Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R., and Weissman, I. L. (2003). A role for Wnt signalling in self-renewal of haematopoietic stem cells. Nature *423*, 409-414.

Rhodes, D. R., Kalyana-Sundaram, S., Mahavisno, V., Varambally, R., Yu, J. 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.

Roberts, A. B., Anzano, M. A., Lamb, L. C., Smith, J. M., and Sporn, M. B. (1981). New class of transforming growth-factors potentiated by epidermal growth-factor - isolation from non-neoplastic tissues. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences *78*, 5339-5343.

RonnovJessen, 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.

Ronnovjessen, 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. J Clin Invest *95*, 859-873.

Rossi, D., and Zlotnik, A. (2000). The biology of chemokines and their receptors. Annu Rev Immunol *18*, 217-243.

Rudolph-Owen, L. A., Chan, R., Muller, W. J., and Matrisian, L. M. (1998). The matrix metalloproteinase matrilysin influences early-stage mammary tumorigenesis. Cancer Research *58*, 5500-5506.

Sakakura, T., Nishizuka, Y., and Dawe, C. J. (1976). Mesenchyme-dependent morphogenesis and epithelium-specific cytodifferentiation in mouse mammary-gland. Science *194*, 1439-1441.

Sandler, A. B., Johnson, D. H., and Herbst, R. S. (2003). Anti-vascular endothelial growth factor monoclonals in non-small cell lung cancer. Paper presented at: 1st International Conference on Novel Agents in the Treatment of Lung Cancer (Cambridge, MA, Amer Assoc Cancer Research).

Sawdey, M. S., and Loskutoff, D. J. (1991). Regulation of murine type-1 plasminogen-activator inhibitor gene-expression invivo - tissue-specificity and induction by lipopolysaccharide, tumor-necrosis-factor-alpha, and transforming growth-factor-beta. J Clin Invest *88*, 1346-1353.

Schafer, M., and Werner, S. (2008). Cancer as an overhealing wound: an old hypothesis revisited. Nat Rev Mol Cell Biol *9*, 628-638.

Schmidt, D. R., and Kao, W. J. (2007). The interrelated role of fibronectin and interleukin-1 in biomaterial-modulated macrophage function. Biomaterials *28*, 371-382.

Schmitt, M., Janicke, F., Moniwa, N., Chucholowski, N., Pache, L., and Graeff, H. (1991). Tumor-associated urokinase-type plasminogen-activator - biological and clinical-significance. Paper presented at: 3rd International Symp on Proteinase Inhibitors and Biological Control (Brno, Czechoslovakia, Walter De Gruyter & Co).

Schofield, R. (1978). Relationship between spleen colony-forming cell and hematopoietic stem-cell - hypothesis. Blood Cells *4*, 7-25.

Schor, S. L., Schor, A. M., and Rushton, G. (1988). Fibroblasts from cancer-patients display a mixture of both fetal and adult-like phenotypic characteristics. J Cell Sci *90*, 401-407.

Schor, S. L., Schor, A. M., Rushton, G., and Smith, L. (1985). Adult, fetal and transformed fibroblasts display different migratory phenotypes on collagen gels - evidence for an isoformic transition during fetal development. J Cell Sci *73*, 221-234.

Schulze, A., and Downward, J. (2001). Navigating gene expression using microarrays - a technology review. Nat Cell Biol *3*, E190-E195.

Schurch, W. (1997). The myofibroblast in neoplasia. Paper presented at: Meeting on Mechanisms Involved in Tissue Repair and Fibrosis - Role of the Myofibroblast (Differentiation and Apoptosis) (Lyon, France, Springer-Verlag Berlin).

Schwartz, J. C., Senko, M. W., and Syka, J. E. P. (2002). A two-dimensional quadrupole ion trap mass spectrometer. J Am Soc Mass Spectrom *13*, 659-669.

Scigelova, M., and Makarov, A. (2006). Orbitrap mass analyzer - Overview and applications in proteomics. Proteomics, 16-21.

Shao, G. Z., Berenguer, J., Borczuk, A. C., Powell, C. A., Hei, T. K., and Zhao, Y. L. (2006). Epigenetic inactivation of Betaig-h3 gene in human cancer cells. Cancer Research *66*, 4566-4573.

Shawver, L. K., Slamon, D., and Ullrich, A. (2002). Smart drugs: Tyrosine kinase inhibitors in cancer therapy. Cancer Cell *1*, 117-123.

Shipitsin, M., Campbell, L. L., Argani, P., Werernowicz, S., Bloushtain-Qimron, N., Yao, J., Nikolskaya, T., Serebryiskaya, T., Beroukhim, R., Hu, M., *et al.* (2007). Molecular definition of breast tumor heterogeneity. Cancer Cell *11*, 259-273.

Singer, A. J., and Clark, R. A. F. (1999). Mechanisms of disease - Cutaneous wound healing. N Engl J Med *341*, 738-746.

Sipponen, P., and Marshall, B. J. (2000). Gastritis and gastric cancer - Western countries. Gastroenterol Clin North Am *29*, 579-+.

Skonier, J., Bennett, K., Rothwell, V., Kosowski, S., Plowman, G., Wallace, P., Edelhoff, S., Disteche, C., Neubauer, M., Marquardt, H., *et al.* (1994). Beta-ig-h3 - a transforming growth factor-beta-responsive gene encoding a secreted protein that inhibits cell attachment in-vitro and suppresses the growth of cho cells in nude-mice. DNA Cell Biol *13*, 571-584.

Skonier, J., Neubauer, M., Madisen, L., Bennett, K., Plowman, G. D., and Purchio, A. F. (1992). cDNA cloning and sequence-analysis of beta-ig-h3, a novel gene induced in a human adenocarcinoma cell-line after treatment with transforming growth-factor-beta. DNA Cell Biol *11*, 511-522.

Small, H., Stevens, T. S., and Bauman, W. C. (1975). Novel ion-exchange chromatographic method using conductimetric detection. Anal Chem *47*, 1801-1809.

Smith, A. G. (2001). Embryo-derived stem cells: Of mice and men. Annu Rev Cell Dev Biol *17*, 435-462.

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. Proc Natl Acad Sci U S A *98*, 10869-10874.

Sotiriou, C., and Piccart, M. J. (2007). Opinion - Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? Nat Rev Cancer *7*, 545-553.

Sporn, M. B., and Todaro, G. J. (1980). Autocrine secretion and malignant transformation of cells. N Engl J Med *303*, 878-880.

Stack, M. S., and Johnson, D. A. (1994). Human mast-cell tryptase activates single-chain urinary-type plasminogen-activator (prourokinase). J Biol Chem *269*, 9416-9419.

Staes, A., Demol, H., Van Damme, J., Martens, L., Vandekerckhove, J., and Gevaert, K. (2004). Global differential non-gel proteomics by quantitative and stable labeling of tryptic peptides with oxygen-18. Journal of Proteome Research *3*, 786-791.

Staes, A., Van Damme, P., Helsens, K., Demol, H., Vandekerckhove, J., and Gevaert, K. (2008). Improved recovery of proteome-informative, protein N-terminal peptides by combined fractional diagonal chromatography (COFRADIC). Proteomics *8*, 1362-1370.

Stefansson, S., and Lawrence, D. A. (1996). The serpin PAI-1 inhibits cell migration by blocking integrin alpha(v)beta(3) binding to vitronectin. Nature *383*, 441-443.

Sternlicht, M. D., Lochter, A., Sympson, C. J., Huey, B., Rougler, J. P., Gray, J. W., Pinkel, D., Bissell, M. J., and Werb, Z. (1999). The stromal proteinase MMP3/stromelysin-1 promotes mammary carcinogenesis. Cell *98*, 137-146.

Sympson, C. J., Bissell, M. J., and Werb, Z. (1995). Mammary-gland tumor-formation in transgenic mice overexpressing stromelysin-1. Semin Cancer Biol *6*, 159-163.

Tahara, E. (1993). Molecular mechanism of stomach carcinogenesis. J Cancer Res Clin Oncol *119*, 265-272.

Taipale, J., and Beachy, P. A. (2001). The Hedgehog and Wnt signaling pathways in cancer. Nature *411*, 349-354.

Taylor, S. W., Fahy, E., and Ghosh, S. S. (2003). Global organellar proteomics. Trends Biotechnol *21*, 82-88.

Thapa, N., Lee, B. H., and Kim, I. S. (2007). TGFBIp/beta ig-h3 protein: A versatile matrix molecule induced by TGF-beta. Int J Biochem Cell Biol *39*, 2183-2194.

Thomasset, N., Lochter, A., Sympson, C. J., Lund, L. R., Williams, D. R., Behrendtsen, O., Werb, Z., and Bissell, M. J. (1998). Expression of autoactivated stromelysin-1 in mammary glands of transgenic mice leads to a reactive stroma during early development. Am J Pathol *153*, 457-467.

Thompson, N. L., Flanders, K. C., Smith, J. M., Ellingsworth, L. R., Roberts, A. B., and Sporn, M. B. (1989). Expression of transforming growth factor-beta-1 in specific cells and tissues of adult and neonatal mice. J Cell Biol *108*, 661-669.

Tlsty, T. D., and Hein, P. W. (2001). Know thy neighbor: stromal cells can contribute oncogenic signals. Curr Opin Genet Dev *11*, 54-59.

Toledo-Pereyra, L. H. (2009). Introduction to the study of the Experimental Medicine Surgical Revolution Part I. J Invest Surg *22*, 157-161.

Tu, S., Bhagat, G., Cui, G., Takaishi, S., Kurt-Jones, E. A., Rickman, B., Betz, K. S., Penz-Oesterreicher, M., Bjorkdahl, O., Fox, J. G., and Wang, T. C. (2008). Overexpression of Interleukin-1 beta Induces Gastric Inflammation and Cancer and Mobilizes Myeloid-Derived Suppressor Cells in Mice. Cancer Cell *14*, 408-419.

Tucker, R. F., Shipley, G. D., Moses, H. L., and Holley, R. W. (1984). Growth inhibitor from bsc-1 cells closely related to platelet type-beta transforming growth-factor. Science *226*, 705-707.

Turk, B. (2006). Targeting proteases: successes, failures and future prospects. Nat Rev Drug Discov *5*, 785-799.

Van Damme, P., Martens, L., Van Damme, J., Hugelier, K., Staes, A., Vandekerckhove, J., and Gevaert, K. (2005). Caspase-specific and nonspecific in vivo protein processing during Fas-induced apoptosis. Nature Methods *2*, 771-777.

van de Vijver, M. J., He, Y. D., van 't Veer, L. J., Dai, H., Hart, A. A. M., 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.

Vandekerckhove, J., and Gevaert, K. (2005). Combined fractional diagonal chromatography (Cofradic); a multipurpose tool for peptide-centric proteomics. Mol Cell Proteomics *4*, S16-S16.

Vandenhooff, A. (1988). Stromal involvement in malignant growth. Advances in Cancer Research *50*, 159-196.

van't Veer, L. J., Dai, H. Y., van de Vijver, M. J., He, Y. D. D., Hart, A. A. M., 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.

Varro, A., Noble, P. J., Wroblewski, L. E., Bishop, L., and Dockray, G. J. (2002). Gastrin-cholecystokinin(B) receptor expression in AGS cells is associated with direct inhibition and indirect stimulation of cell proliferation via paracrine activation of the epidermal growth factor receptor. Gut *50*, 827-833.

Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., *et al.* (2001). The sequence of the human genome. Science *291*, 1304-+.

Vogelstein, B., and Kinzler, K. W. (2004). Cancer genes and the pathways they control. Nat Med *10*, 789-799.

Volmer, M. W., Stuhler, K., Zapatka, M., Schoneck, A., Klein-Scory, S., Schmiegel, W., Meyer, H. E., and Schwarte-Waldhoff, I. (2004). Differential proteome analysis of conditioned media to detect Smad4 regulated secreted biomarkers in colon cancer. Paper presented at: 6th Siena Meeting on From Genome to Proteome (Siena, ITALY, Wiley-V C H Verlag Gmbh).

Wahl, S. M. (1994). Transforming growth-factor-beta - the good, the bad, and the ugly. J Exp Med *180*, 1587-1590.

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spencer, J., *et al.* (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science *280*, 1077-1082.

Wang, F., Weaver, V. M., Petersen, O. W., Larabell, C. A., Dedhar, S., Briand, P., Lupu, R., and Bissell, M. J. (1998). Reciprocal interactions between beta 1-integrin and epidermal growth factor receptor in three-dimensional basement membrane breast cultures: A different perspective in epithelial biology. Proc Natl Acad Sci U S A *95*, 14821-14826.

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

Weaver, V. M., Petersen, O. W., Wang, F., Larabell, C. A., Briand, P., Damsky, C., and Bissell, M. J. (1997). Reversion of the malignant phenotype of human breast cells in three-dimensional culture and in vivo by integrin blocking antibodies. J Cell Biol *137*, 231-245.

Werb, Z., and Chin, J. R. (1997). Extracellular matrix remodeling during morphogenesis. Paper presented at: Conference on Morphogenesis - Cellular Interactions (Bethesda, Maryland, New York Acad Sciences).

Wiese, S. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research (vol 7, pg 340, 2007). Proteomics *7*, 1004-1004.

Wilson, C. L., Heppner, K. J., Labosky, P. A., Hogan, B. L. M., and Matrisian, L. M. (1997). Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. Proc Natl Acad Sci U S A *94*, 1402-1407.

Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. Nature Methods *6*, 359-U360.

Woessner, J. F. (1991). Matrix metalloproteinases and their inhibitors in connective-tissue remodeling. Faseb J *5*, 2145-2154.

Wolf, B. B., Vasudevan, J., Henkin, J., and Gonias, S. L. (1993). Nerve growth factor-gamma activates soluble and receptor-bound single-chain urokinase-type plasminogen-activator. J Biol Chem *268*, 16327-16331.

Wood, L. D., Parsons, D. W., Jones, S., Lin, J., Sjoblom, T., Leary, R. J., Shen, D., Boca, S. M., Barber, T., Ptak, J., *et al.* (2007). The genomic landscapes of human breast and colorectal cancers. Science *318*, 1108-1113.

Wu, K. C., Jackson, L. M., Galvin, A. M., Gray, T., Hawkey, C. J., and Mahida, Y. R. (1999). Phenotypic and functional characterisation of myofibroblasts, macrophages, and lymphocytes migrating out of the human gastric lamina propria following the loss of epithelial cells. Gut *44*, 323-330.

Wu, W. W., Wang, G. H., Baek, S. J., and Shen, R. F. (2006). Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. Journal of Proteome Research *5*, 651-658.

Wun, T. C., Ossowski, L., and Reich, E. (1982). A proenzyme form of human urokinase. J Biol Chem *257*, 7262-7268.

Xin, X. P., Hou, Y. T., Li, L. N., Schmiedlin-Ren, P., Christmas, G. M., Cheng, H. L., Bitar, K. N., and Zimmermann, E. M. (2004). IGF-I increases IGFBP-5 and collagen alpha(1)(I) mRNAs by the MAPK pathway in rat intestinal smooth muscle cells. American Journal of Physiology-Gastrointestinal and Liver Physiology *286*, G777-G783.

Xu, J., Lamouille, S., and Derynck, R. (2009). TGF-beta-induced epithelial to mesenchymal transition. Cell Res *19*, 156-172.

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

Yang, L., and Moses, H. L. (2008). Transforming Growth Factor beta: Tumor Suppressor or Promoter? Are Host Immune Cells the Answer? Cancer Research *68*, 9107-9111.

Yang, Y. N., Pan, X. C., Lei, W. W., Wang, J., Shi, J., Li, F. Q., and Song, J. G. (2006). Regulation of transforming growth factor-beta 1-induced apoptosis and epithelial-to-mesenchymal transition by protein kinase A and signal transducers and activators of transcription 3. Cancer Research *66*, 8617-8624.

Yauch, R. L., Gould, S. E., Scales, S. J., Tang, T., Tian, H., Ahn, C. P., Marshall, D., Fu, L., Januario, T., Kallop, D., *et al.* (2008). A paracrine requirement for hedgehog signalling in cancer. Nature *455*, 406-U461.

Yee, J. A., Yan, L., Dominguez, J. C., Allan, E. H., and Martin, T. J. (1993). Plasminogen-dependent activation of latent transforming growth-factor-beta (tgf-beta) by growing cultures of osteoblast-like cells. J Cell Physiol *157*, 528-534.

Yost, R. A., and Enke, C. G. (1979). Triple quadrupole mass-spectrometry for direct mixture analysis and structure elucidation. Anal Chem *51*, 1251-&.

Zavadil, J., and Bottinger, E. P. (2005). TGF-beta and epithelial-to-mesenchymal transitions. Oncogene *24*, 5764-5774.

Zhang, Y., Wen, G. Y., Shao, G. Z., Wang, C. D., Lin, C. S., Fang, H. B., Balajee, A. S., Bhagat, G., Hei, T. K., and Zhao, Y. L. (2009). TGFBI Deficiency Predisposes Mice to Spontaneous Tumor Development. Cancer Research *69*, 37-44.

Zhao, Y. L., Piao, C. Q., and Hei, T. K. (2002). Downregulation of Betaig-h3 gene is causally linked to tumorigenic phenotype in asbestos treated immortalized human bronchial epithelial cells. Oncogene *21*, 7471-7477.

Zhao, Y. L., Shao, G. Z., Piao, C. Q., Berenguer, J., and Hei, T. K. (2004). Down-regulation of Betaig-h3 gene is involved in the tumorigenesis in human bronchial epithelial cells induced by heavy-ion radiation. Radiat Res *162*, 655-659.

Zhong, L., Roybal, J., Chaerkady, R., Zhang, W., Choi, K., Alvarez, C. A., Tran, H., Creighton, C. J., Yan, S., Strieter, R. M., *et al.* (2008). Identification of secreted proteins that mediate cell-cell interactions in an in vitro model of the lung cancer microenvironment. Cancer Research *68*, 7237-7245.