Molecular mechanisms of human colorectal cancer metastasis – identification of novel biomarkers

A dissertation submitted for the degree of Ph.D

by

David W. Murray B.Sc.

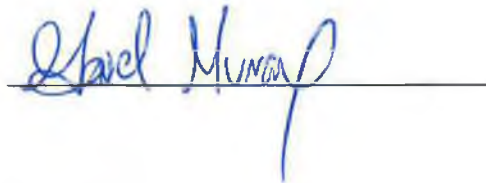Under the supervision of Dr Susan McDonnell

September 2004

School of Biotechnology, Dublin City University, Dublin 9, Ireland.

Declaration

'I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work'.

Signed: _David Munro_   I.D. Number: 96533439

Date: Friday, 17 September 2004

Acknowledgements

Of all the sections in this thesis, this one has proven to be one of the most difficult to write. How do I acknowledge all those who have been there over the past four years? What if I forget someone? And how do I make it an interesting read ?! I have also found it impossible to fully convey my thanks for the help and support received in only a few words. I would have to write for hours to adequately show how grateful I am to you all. Nevertheless, I will give it a bash!

Firstly, I wish to acknowledge Dr Susan McDonnell for her continuous guidance. I would not have wanted to do a PhD for anyone else. Researching under your supervision has been a huge learning experience and I know the skills and knowledge I have gained will always stand to me. Thank you for your enthusiasm, you kept my motivation up and thank you for allowing me to work in a great area of cancer research. I am also grateful for the opportunities to attend and present at many conferences and meetings, these were priceless experiences. Thanks also for your help with proof reading the thesis.

A whole chapters-worth of thank you's go to Dr Geraldine Grant and all at the Centre for Biomedical Genomics at George Mason University, Virginia, including, Amy, Luca, Mike and Francesco. Thanks for the help on the arrays and all the nights out in Manassas! I wish to thank Prof John Dalton and Peter Collins for helping with the cathepsin C study. A big thanks goes to Dr Mary Morrin who has been a great source of CD44 information, antibodies and primers, thanks for the help. I also wish to thank Mary McGrath for her help on the tissue work.

Thank you Áine O'Connor for all the craic in the early days. A thanks also goes to Dr Conor Lynch for all your help back then. Thank you Lynda and Aisling for making the lab a pleasant place to work, I wish both of you the best of luck with your projects which I'm sure you'll carry out in an alphabetically organised, colour-coded and labelled fashion! Thanks for the proof reading. Thanks to all the friends I have made in DCU during the PhD including prep room staff, fellow PhDers and lecturers, you made the school of Biotech a great place to work. I will always remember the BRS trips to Prague, Budapest and the equally as exotic Carlow, great craic, great memories. Thanks to my drinkin buddies, my friends both old and new, especially (and I hope I haven't forgotten anyone!) Nick, Roman, Eoin J, Aj, Dallas, Scully, Max, Lynda Magee, Rob, Deano, Fitzer, Greg, Frenchy, Iano, Johnny-boy, Ken, Shane O'Malley, Tara, Paul Ó Connor, Mark 'Stanton' Brennan and Pat. Thanks for great nights out, good laughs and thanks for the convenient and well-timed distractions! Thank you Mick for the proof reading and for being a lifelong and true friend.

Loving thanks to Sinéad. These words are not powerful enough to describe my gratitude towards you for being a wonderful girlfriend and for putting up with me and cheering me up when I was stressed or annoyed. Thank you for always being there and the joy of being near you and thanks for being great company. You have always put a smile on my face XXX

I would mostly like to thank my family; Stephen, Jennifer, Annette and John, your love and support have always been there for me. It is you who have been there from the start, I would not be where I am today without you. Thank you also for the financial aid.

This thesis is dedicated to my Mam and Dad, who have always encouraged me to follow this path, you are my heroes, and I will always look up to you. I admire your hard-working determination and am touched my how much you give, you have always been providers for your family. However, you both know how to enjoy life, relax and laugh, both of you having a great sense of humour (even if you do laugh at your own jokes Dad!). I am grateful to know, let alone be the son of two such great people.

David William Murray
September 2004

Abstract

Tumour metastasis represents the most lethal aspect of cancer. It involves initial attachment of cells to the extracellular matrix (ECM), degradation of the ECM and subsequent migration and detachment of cells from the ECM. This process depends on alterations in the expression of many key players including cell adhesion molecules (CAMs) and proteolytic enzymes such as the matrix metalloproteinases (MMPs). The MMPs are a large family of proteases collectively capable of degrading the entire ECM. CD44 is a multifunctional transmembrane adhesion molecule, which also plays a key role in cell signalling. The interactions between CAMs and MMPs are crucial therefore in controlling tumour cell invasion and metastasis and this study aimed to elucidate these interactions.

We have established a panel of human colon cancer cell lines including a primary human colon adenocarcinoma cell line SW480 and its lymph node metastatic derivative, SW620. The SW480 cells have been stably transfected with the cDNA for MMP-7 (SW480M7) and MMP-9 (SW480M9). When the *in vitro* invasive activities of the cell lines were compared, the SW620 cells and the MMP transfected cells were more invasive than their parental SW480. The SW480M7 cells were more migratory in comparison to SW480 cells and interestingly, the SW620 were less migratory. MMP-9 expression increased when the SW480M9 cells were grown on HA and collagen. In addition, SW480M9 cells cultured in the presence of a CD44 activating antibody resulted in a time dependent increase in MMP-9 activity and increased invasion. The SW480M9 cells were shown to be the most adherent to plastic, HA and fibronectin and the SW480M7 were slightly less adherent. The increased HA adhesion in the SW480M9 correlated with increased 80kDa and 120kDa CD44 expression.

cDNA microarray analysis was used to identify additional metastasis associated genes by examining and comparing differential gene expression profiles in these cell lines Analysis of gene expression data revealed that 233 genes were significantly up regulated and 208 genes were down regulated in the SW620 when compared to the SW480. Over expression of MMP-7 and MMP-9 in the SW480 cells also caused differential expression. Following literature searches on the gene lists, 5 genes of biological interest were chosen for further analysis and validation. Selected genes were the cysteine protease, cathepsin C, caveolin 1, the nuclear receptor, peroxisome proliferative activated protein gamma (PPARγ), the ECM component tenascin C and asparagine synthetase. These clones were sequenced from the cDNA library to verify their identities and real time PCR analysis confirmed the fold changes in expression. Using a fluorescent assay, enhanced cathepsin C enzymatic activity in the SW620 cells was confirmed. A PCR-based method was also used to monitor the expression profiles of particular alternatively spliced tenascin C isoforms in the cell lines studied. An initial pilot study was conducted to examine expression of these genes in matched samples of normal and tumour colon tissue. RNA was extracted from tissues of varying Dukes' classification and expression was examined by real time PCR.

This study identified potential new biomarkers of colorectal metastasis, and examined the involvement of many key players in this process. This study demonstrated that interactions between CD44 and MMPs were important in controlling tumour cell invasion. It is hoped that this work will help elucidate the exact mechanisms of such interactions as well as highlighting their importance in the metastatic process. Further study on these biomarkers and interactions would validate their use as novel markers and therapeutic targets in metastatic colorectal disease.

Abbreviations and Acronyms

A	Absorbance
aa	Aminoallyl
AEBSF	4-(2-aminoethyl)benzenesulfonyl fluoride
ALL	Acute Lymohoblastic Leukemia
AMC	7-Amino-4-methylcoumarin
AML	Acute Myeloid Leukaemia
AP-1	Activator Protein-1
APC	Adenomatous Polyposis Col
Arg	Arginine
aRNA	Amplified RNA
Asn	Asparagine
Asp	Aspartic Acid
ATP	Adenosine Triphosphate
BCA	Bicinchonic Acid
bFGF	Basic FGF
BHK	Baby Hamster Kidney
BSA	Bovine Serum Albumin
CA	Cysteine Array
CAM	Cell Adhesion Molecule
CAT	Catalytic Domain
cDNA	Complementary DNA
CM	Conditioned Media
CTP	Cytosine triphosphate
Cy	Cyanine
Cys	Cysteine
DAP	Dipeptidyl Aminopeptidase
DEPC	Diethyl Pyrocarbonate
DMEM	Dulbecco's Modified Eagle's Medium
DMF	Dimethylformamide
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
DNase	Deoxynuclease
dNTP	Deoxyribonucleotide Triphosphate

DPX	Distrene, Plasticiser, Xylene
dsDNA	Double stranded DNA
DTT	Dithiothreitol
ECL	Enhanced Chemoluminescence
ECM	Extracellular Matrix
EDTA	Ethylenediamine Tetraacetic Acid
EGF	Epidermal Growth Factor
EMMPRIN	Extracellular Matrix Metalloproteinase Inducer
ERM	Ezrin, Radixin, Moesin
FAK	Focal Adhesion Kinase
FAP	Familial Adenomatous Polyposis
FBS	Fetal Bovine Serum
FGF	Fibroblast Growth Factor
Glu	Glutamic Acid
Gly	Glycine
GMU	George Mason University
GPI	Glycosyl Phosphatidyl Inositol
GSEA	Gene set enrichment analysis
GTP	Guanosine triphosphate
HA	Hyaluronic Acid
HBEGF	Heparin Bound EGF
HCl	Hydrochloric Acid
HEM	Hemopexin
His	Histidine
HNPCC	Hereditary Non-Polyposis Colon Cancer
HRP	Horseradish Peroxidase
Ig	Immunoglobulin
I_{LBS}	Local Background Subtracted Intensity
LCM	Laser Capture Micro-dissection
LOWESS	Local Weighted Linear Regression Analysis
MAPK	Mitogen Activated Protein Kinase
MDCK	Madin-Darby Canine Kidney
MMLVRT	Murine Moloney Leukemia Virus Reverse Transcriptase
MMP	Matrix Metalloproteinase
MMPI	MMP Inhibitor

MOPS	3-Morpholinopropane Sulfonic Acid
mRNA	Messenger RNA
MT MMP	Membrane Type MMP
MTS	[3-(4,5-dimethylthiazol-2-yl)-5-(3- carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium]
MW	Molecular Weight
NCBI	National Centre for Biotechnology Information
NHR	Nuclear Hormone Receptor
p53	Protein 53
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PI-3-K	Phosphoinositol-3-Kinase
PPARγ	Peroxisome Proliferator Activated Receptor Gamma
Pro	Proline
RNA	Ribonucleic Acid
RNase	Ribonuclease
RT	Room Temperature (25 ^{o}C)
RT-PCR	Reverse Transcription PCR
SAGE	Serial Analysis of Gene Expression
SCLC	Squamous Cell Lung Carcinoma
SDS-PAGE	Sodium Dodecyl Sulphate PAGE
SSC	Saline Sodium Citrate
Taq	Thermus Aquaticus
TBE	Tris Borate EDTA
TBST	Tris Buffered Saline plus Tween
TEMED	NNN'N'-Tetramethylethylenediamine
TGF	Transforming Growth Factor
TIFF	Tagged Image File Format
TIGR	The Institute of Genomic Research
TIMP	Tissue Inhibitor of Metalloproteinases
TM	Transmembrane
TNFα	Tumour Necrosis Factor alpha
TNM	Tumour-Node-Metastasis
TPA	12-O-tetradecanoyl-phorbal-13-acetate

TRE	TPA responsive element
Triton-X-100	t-Octylphenoxypolyethoxyethanol
TTP	Thymidine Triphosphate
USA	United States of America
UTP	Uracil Triphosphate
UV	Ultraviolet
VA	Virginia
Val	Valine
VCAM	Vascular Cell Adhesion Molecule
VEGF	Vascular Endothelial Growth Factor

Units

bp	Base Pairs
cm^2	Centimetres Squared
g	Gram or gravities
hr	Hour(s)
kDa	Kilodalton
M	Molar
mA	Milliamps
mg	Milligram
min	Minute(s)
mJ	Millijoule
ml	Millilitre
mM	Millimolar
nm	Nanometre
^{o}C	Degrees Celsius
rpm	Revolutions Per Minute
sec	Second(s)
U	Units of enzymatic activity
V	Volts
v/v	By volume
W	Watts
w/v	By Weight
μg	Microgram
μl	Microlitre

Publications

Murray, D., Gorreta, F., Grant, G., Chandhoke, V. and McDonnell, S. (2004) The investigation of tumour metastasis using cDNA microarrays. In Appasani, K. (Ed) *Bioarrays: from basics to diagnostics.* (*in press*) Totowa, NJ: Humana Press.

Murray, D., Morrin, M. and McDonnell, S. (2004). Increased invasion and expression of MMP-9 in human colorectal cell lines by a CD44-dependent mechanism. *Anticancer Res* **24**: 489-494.

Murray, D., Goretta, F., Grant, G., Chandhoke, V. and McDonnell, S. (2004). cDNA array analysis of gene expression profiles in a metastatic model of human colorectal cancer. *Proc. Amer. Assoc. Cancer Res.* **45**: 1725.

Murray, D. and McDonnell S. (2002). Interactions between CD44, beta-1 integrin and the matrix metalloproteinases in regulating colorectal cancer cell adhesion and invasion. *Br. J. Cancer* **86**: S82.

Morrin, M., Murray, D., McDonnell, S. and Delaney, P. V. (2002). Characterisation of effect on colorectal tumour cells of CD44 activation. *Br. J. Cancer* **86**: S74.

Mc Donnell, S., O'Connor, Á., Murray, D. and Lynch, C. (2002). Matrix Metalloproteinases, Chapter 3, pp17-24. *Hormone Replacement Therapy and Cancer. The current status of research and practice.* Edited by A. R. Genazzani. Parthenon Publishing.

Morrin, M., Murray, D., McDonnell, S. and Delaney, P. V. (2001). Signal transduction via CD44 regulates CD44 variant isoform and β-3 integrin expression, and MMP-2 production. *National Institute of Health Sciences Research Bulletin* **1**: 18.

Oral Presentations

'CD44 and MMP interactions in colorectal cancer adhesion migration and *in vitro* invasion' Irish Association for Cancer Research (IACR), Annual Meeting, Kilkenny, Ireland, April 2003.

'Interactions between CD44, β-1 integrin and the matrix metalloproteinases in regulating colorectal cancer cell adhesion and invasion' DCU School of Biotechnology Seminar Series, April 2002.

Poster Presentations

Murray, D., Gorreta, F., Grant, G., Chandhoke, V. and McDonnell, S. Microarray based identification of novel potential biomarkers of colorectal metastasis. IACR Annual Meeting, Belfast, Northern Ireland, April 2004.

Murray, D., Gorreta, F., Grant, G., Chandhoke, V. and McDonnell, S. cDNA array analysis of gene expression profiles in a metastatic model of human colorectal cancer. American Association for Cancer Research Annual Meeting, Orlando, March 2004.

Murray, D. and McDonnell, S. Increased invasion and expression of matrix metalloproteinase-9 in colorectal cell lines by a CD44 dependent mechanism. Biotechnology, Cancer & Drug Resistance Conference, Dublin, July 2003.

Morrin, M., Murray, D., McDonnell, S. and Delaney, P. V. Characterisation of effect on colorectal tumour cells of CD44 activation. British Cancer Research Meeting, Glasgow, Scotland, June 2002.

Murray, D. and McDonnell, S. Interactions between CD44, β1 integrin and the matrix metalloproteinases in regulating colorectal cancer cell adhesion and invasion. British Cancer Research Meeting, Glasgow, Scotland, June 2002.

Murray, D., Lynch, C. and McDonnell, S. The regulation of MMPs/TIMPs in leukemia cell lines. IACR and Irish Society for Medical Oncology Joint Cancer Research Meeting Cork, Ireland, September 2001.

Morrin, M., Murray, D., McDonnell, S. and Delaney, P. V. Signal Transduction via CD44 regulates CD44 variant isoform and β3 Integrin expression, and MMP production. IACR and Irish Society for Medical Oncology Joint Cancer Research Meeting Cork, Ireland, September 2001.

Academic Awards

Irish Cancer Society Oncology Scholar Travel Award (2004).

Best Speaker Award IACR Annual Meeting (2003).

North Carolina State University Summer Institute in Statistical Genetics Scholarship (2002).

IACR/MWG Travel Scholarship (2002).

DCU School of Biotechnology Orla Benson Memorial Travel Award (2002).

Table of Contents

Chapter 1: An Introduction to Tumour Invasion and Metastasis.

Chapter 2: Materials and Methods

Chapter 3: The roles of MMPs and CD44 in colon cancer metastasis

Chapter 4: Gene expression profiling of a panel of colorectal cancer cell lines using cDNA microarrays

Chapter 5: Expression of novel markers of colorectal metastasis in paired normal and tumour colon samples

Chapter 6: Final conclusions and summary

Chapter 7: Bibliography

Appendix A: Genes with significant changes in expression

Appendix B: Sequence verification

Appendix C: Genes with common changes in expression over all experiments

Appendix D: Murray *et al.*, (2004) *Anticancer Res* **24**: 489-494.

Chapter 1

An introduction to tumour invasion and metastasis

1.1 Tumour invasion and metastasis

Metastasis represents the most lethal aspect of cancer and is the cause of 90 % of cancer deaths (Hanahan and Weinberg, 2000). This is usually because it is difficult to treat patients when the cancer is at this late stage. The metastatic spread of cancer cells from their site of origin enables them to escape the primary tumour mass and colonise new terrain in the body. Figure 1.1 gives a graphical representation of the overall process. Metastasis is a multi-step process that involves local area invasion at the primary site, followed by intravasation of tumour cells to either the blood or lymphatic vessels, thus accessing the general circulation. After circulating in the vascular systems, tumour cells arrest and extravasate at distant organs where they establish secondary tumours.
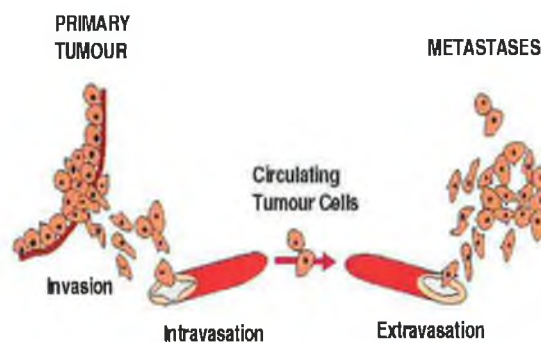
Figure 1.1 The multistage process of metastasis. (Figure adapted from www.ma.hw.ac.uk/~tea/ research.html, 17 September 2004)

The process of metastasis therefore involves the participation of numerous biomolecules in a variety of intricate cellular functions including altered cell adhesion, proteolysis and migration. It is not clear how the cancer cell coordinates these events but it is clear that the components of the tumour microenvironment are involved in contributing to tumour cell progression (Lynch and Matrisian, 2002). Apparently, normal bystanders such as fibroblasts and endothelial cells are conscripted by cancer cells to serve as active collaborators in their neoplastic agenda.

Groups of certain key players are involved in the main events of metastasis. These include the extracellular matrix (ECM), the cell adhesion molecules (CAMs) and proteolytic enzymes such as the matrix metalloproteinases (MMPs) (Bogenrieder and

Herlyn, 2003). Figure 1.2 illustrates how these molecules are employed in the molecular mechanisms of tumour metastasis. These key players are implicated at almost every stage in the metastatic cascade. It is clear these groups of biomolecules have roles in metastasis but quite what these roles are, remains unclear. This chapter will introduce these key players.
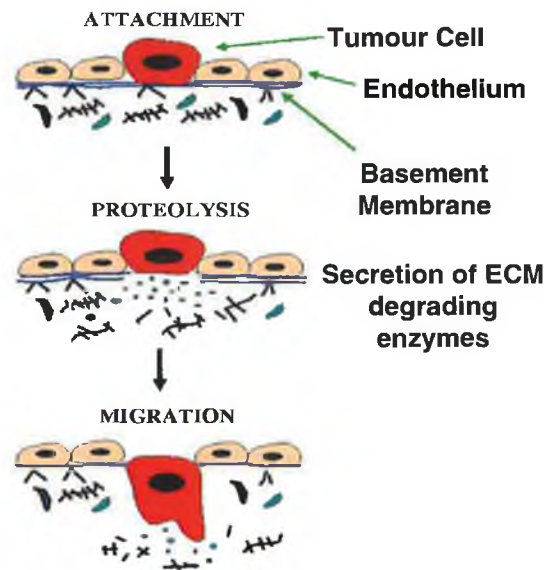
Figure 1.2 Key events in tumour cell invasion. Before tumour cells breech the basement membrane they must first adhere to and degrade the ECM prior to their migration and intravasation into the circulation. (Figure adapted from www.ma.hw.ac.uk/~tea/ research.html, 17 September 2004)

1.2 The extracellular matrix

Most cells in multi-cellular organisms are in intimate and often in permanent contact with either the surface of other cells or with the complex networks of structural macromolecules, termed the ECM. The ECM is an insoluble non-cellular material present between cells throughout the body of multi-cellular organisms. This extracellular superstructure is composed of a complex mixture of secreted molecules (Ruoslahti, 1996). The principal components of the ECM are; glycoproteins, such as collagen IV, laminin, fibronectin, proteoglycans and glycosaminoglycans, such as hyaluronic acid (HA). These molecules assemble into the extracellular matrix once secreted.

There are three major functions of the ECM. Firstly, it provides structural support and tensile strength. Secondly, it provides substrates for cell adhesion and cell migration. Thirdly, the ECM is a functional microenvironment that regulates cellular differentiation and metabolic function, for example by modulation of cell growth by binding growth factors (Park *et al.*, 2000). The ECM is therefore no longer solely thought of as a static structural support for cells. The ECM plays a key role in maintaining normal tissue homeostasis between epithelial cells and their microenvironment and neighbouring cells, such as fibroblasts and endothelial cells. ECM-cell interactions trigger signalling cascades with broad affects on a variety of cellular functions including survival, differentiation and proliferation (Giancotti and Ruoslahti, 1999). The ECM also provides mechanical support for the migration of tumour cells. Furthermore, during tumour invasion soluble ECM proteins released due to their cleavage by proteolytic enzymes can stimulate cell migration. Matrix proteins such as fibronectin, collagen (type I and IV) and vitronectin have been shown to induce motility (Woodhouse *et al.*, 1997).

1.3 Cell adhesion molecules in metastasis

A key group of biomolecules that participate in metastasis are the proteins involved in linking cells to each other and to their surroundings. The adhesive interactions of cells with their immediate environment are mediated by the receptor-ligand interactions of plasma membrane proteins termed CAMs. CAMs are a limited number of gene families including the integrins, the immunoglobulin superfamily, cadherins and selectins (Figure 1.3) (Rosales *et al.*, 1995). Additional membrane components include syndecans and the CD44 protein. For the purpose of this study, only the integrins and CD44 will be discussed further in this chapter.

Interactions between the ECM and CAMs are key to tumour invasion and metastasis. Normal ECM-CAM interactions become deregulated during the metastatic process (Zetter, 1993). Throughout the entire metastatic cascade the tumour cells must undergo changes in adhesive preferences (i.e. from epithelial cells to endothelial cells to fibroblasts etc). These changes are dictated by specific interactions with the ECM and are caused by changes in CAM expression. Changes in adhesion interactions result in changes in the regulatory signals normally conveyed to the cells by the affected adhesion molecules.

4

| Cadherins | Ig Superfamily | Selectins | Integrins |
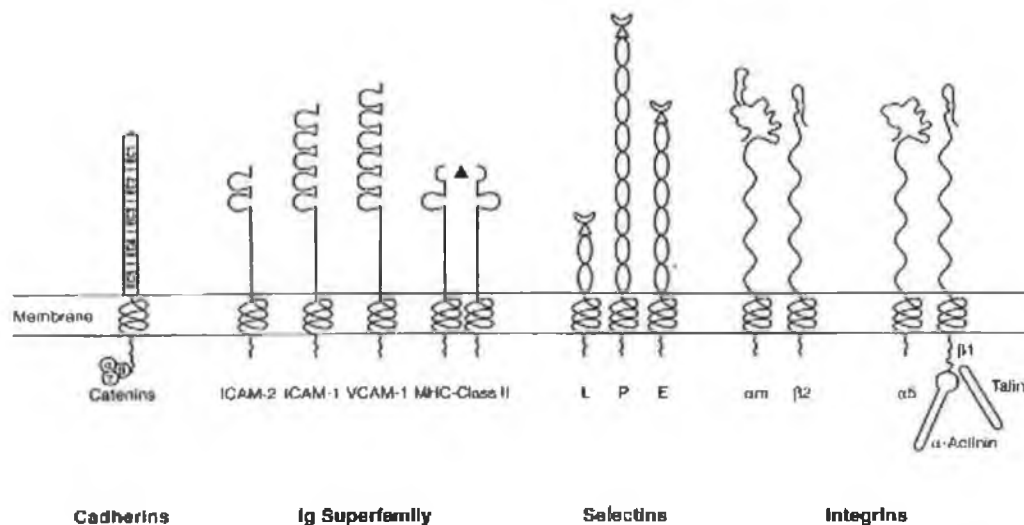
Figure 1.3 Structure of adhesion receptor families. This figure shows the general structures for the 4 major CAM families. Cadherins have 5 extracellular (EC) sites and they interact with catenins via their cytoplasmic tail. There are many members in the Ig superfamily, all of which share an immunoglobulin domain. Selectins are characterised by their short cytoplasmic tail and EGF-like repeats. Integrins are heterodimeric receptors consisting of one α and one β subunit linked non-covalently. As well as cell – cell interactions, integrins are the main CAM family involved in cell–ECM interactions. Integrins interact with talin, α-actinin and focal adhesion kinase (FAK) via their cytoplasmic tails. (Adapted from Rosales *et al.*, 1995)

1.3.1 Integrins

The integrins are one of the major CAM families and it is only 17 years since they have been recognised as such (Hynes, 1987). Integrins are a group of at least 24 bifunctional heterodimeric cell-surface glycoproteins formed by non-covalent binding of one of 18 α and one of 8 β subunits. Each subunit has a large extracellular domain (~1600 amino acids), a single membrane spanning region, and in most cases (apart from β4), a short cytoplasmic domain (20-50 amino acids) (Ruoslahti, 1991). Integrins act as receptors for ECM proteins, or for membrane bound counter-receptors on other cells.

Up to 24 integrin dimers have been found in nature, indicating that not all subunits can bind to each other (Hynes, 2002). Each integrin dimer has distinct ligand binding and signalling properties. The possible integrins are illustrated in figure 1.4. The β1 integrin subunit can heterodimerise with 12 α subunits each with different ligand

affinities, suggesting that the α subunits are involved with the ligand binding specificities of integrins. For example, in the β1-subunit 'sub'-family, four integrins (α1β1, α2β1, α11β1 and α10β1) bind collagens, five (α4β1, α5β1, α8β1, α9β1 and αVβ1) bind fibronectin and five (α1β1, α2β1, α3β1, α6β1, and α7β1) bind laminins (Humphries, 2000). Therefore, an individual integrin can recognize several distinct ECM proteins.
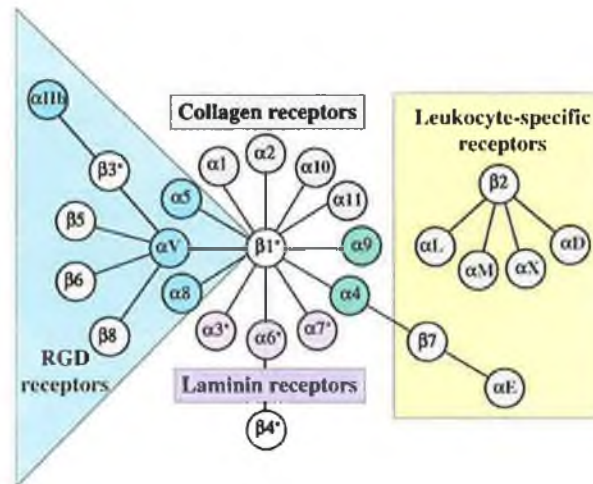
Figure 1.4 The Integrin Receptor Family. This figure illustrates integrin αβ subunit associations, where 8 β subunits can assort with 18 α subunits to form 24 distinct integrins that are considered in several subfamilies based on evolutionary relationships (colouring of α subunits), ligand specificity and, in the case of β2 and β7 integrins, restricted expression on white blood cells. This illustrates how the β4 subunit can only bind with the α6 subunit. On the other hand, the β1 subunit can form heterodimers with 12 different alpha subunits. Asterisks denote alternatively spliced cytoplasmic domains. Figure adapted from (Hynes, 2002).

Integrins are involved in cell-cell and cell-matrix adhesion interactions. It is mainly via integrins that cells bind to and respond to the ECM. Successful binding to specific moieties of the ECM enables integrin-mediated transduction of signals into the cytoplasm that influence cell behaviour ranging from quiescence in normal tissue to motility, resistance to apoptosis and entrance into the active cell cycle. As well as a variation in adhesive properties, each integrin dimer also has unique biological functions and transduces different signals. Failure of integrins to form extracellular links can impair cell motility, induce apoptosis or cause cell cycle arrest. The roles of integrins in tumour cell progression are therefore immense and they are therefore prime targets for the development of therapeutic agents (Juliano and Varner, 1993).

1.3.1.1 Integrin mediated signalling

In addition to their roles in adhesion to ECM ligands or counter-receptors on adjacent cells, integrins serve as transmembrane mechanical links from those extracellular contacts to the cytoskeleton inside cells. Integrin-ligand interactions are accompanied by the transduction of signals into intracellular signal transduction pathways mediating numerous intracellular events including proliferation, survival/apoptosis, changes in cell shape, motility, gene expression and differentiation (Schwartz, 1992). These signalling pathways are very similar to those emanating from growth factor receptors and some are coupled with them. In fact, the transduction of many growth factor and cytokine signals via their respective cell surface receptors is dependent on the anchorage of cells via the integrins (Schwartz *et al.*, 1995). Some integrin-mediated signals are essential for cell survival. For example, integrin signalling via the PI3-kinase pathway is essential to block apoptosis and β1-containing integrin-ligand binding results in activation of the mitogen activated protein kinase (MAPK) pathway; also integrin signalling via ERK is essential for the stimulation of cell cycle progression (Morino *et al.*, 1995). It is common in cancers for the control of these important pathways such as those that control cell proliferation to be lost, ultimately leading to uncontrolled growth.

1.3.2 CD44

Another cell adhesion molecule of interest to this study is CD44, a transmembrane glycoprotein expressed on the surface of a wide variety of cells that carries N- and O-linked sugars and glycosaminoglycan side chains. CD44 is the major cell surface receptor for HA in the ECM (Picker *et al.*, 1989 and Ponta *et al.*, 1998). CD44 is a multifunctional polymorphic cell adhesion molecule with proposed functions in adhesion, latent growth factor activation, invasion and migration (Lesley *et al.*, 1993).

1.3.2.1 CD44 structure

Ten exons of the CD44 mRNA are subject to alternative splicing (Figure 1.5), thus, generating different isoforms, containing inserts of varying sizes in the extracellular portion of the molecule (CD44 variants). It has been estimated that 45

CD44 isoforms exist as a result of this alternative splicing (van Weering *et al.*, 1993). The smallest and most common CD44 isoform, known as standard CD44 or CD44s, lacks any of the 10 variant exons and has a 248 amino acid extracellular domain with an apparent molecular weight of 80 kDa. The larger CD44 variants (CD44v) are similar to CD44s except they have sequences selected from the exons v1-v10. The 'stem' structure is therefore enlarged by sequences that are encoded by the alternatively spliced variant exons. CD44 isoforms are also subject to such post-translational modifications such as glycosylations with the addition of O- and N- linked carbohydrates. The sequence encoded by exon v3 includes a heparin-sulphate site, a modification to which several heparin-binding proteins such as growth factors can attach (Yu *et al.*, 2002).
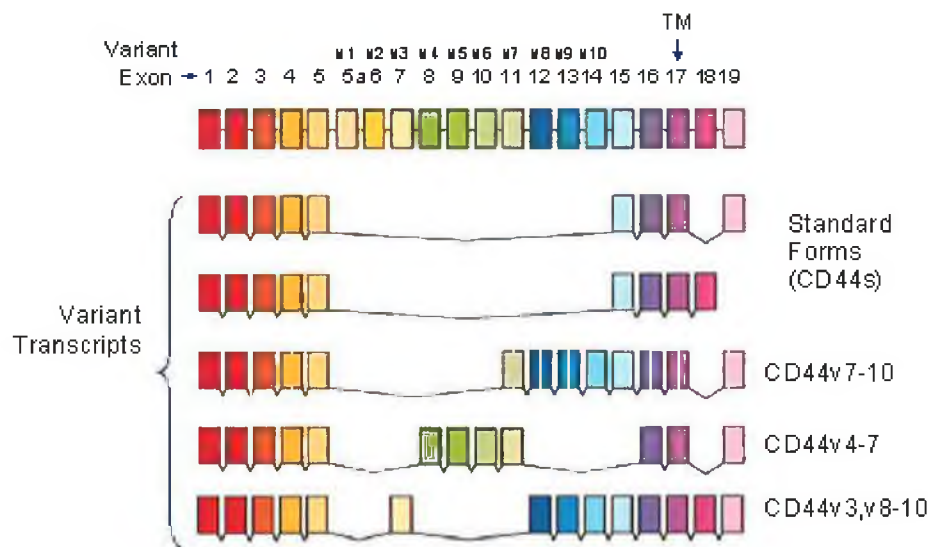
Figure 1.5 CD44 alternative splicing. Ten CD44 exons (v1 to v10) can be alternatively spliced giving rise to CD44 isoforms (CD44v) of varying structures and properties. None of these are expressed in standard CD44 (CD44s). TM - Transmembrane Domain. Figure adapted from Ponta *et al.*, 1998.

All CD44 isoforms share highly conserved transmembrane (23 hydrophobic amino acids) and cytoplasmic (72 amino acids) domains (Isacke and Yarwood, 2002). CD44 binds hyaluronic acid through an amino terminal globular domain (Bajorath *et al.*, 1998), which is separated from the plasma membrane by a 'stem' structure that contains proteolytic cleavage sites (Cichy and Pure, 2003 and Okamoto *et al.*, 1999) (Figure 1.6).

Little is known about the mechanisms controlling the alternative splicing of CD44, but they are precisely regulated, with particular isoforms being expressed in

particular cell types and quite often a variety of isoforms exist within the same cell. Abnormal patterns and levels of variant CD44 expression are frequently associated with malignancy and are thought to represent a breakdown in this regulation (Gunthert *et al.*, 1991 and Wielenga *et al.*, 1998). CD44 and its splice variants are therefore often used as prognostic markers of tumour progression.
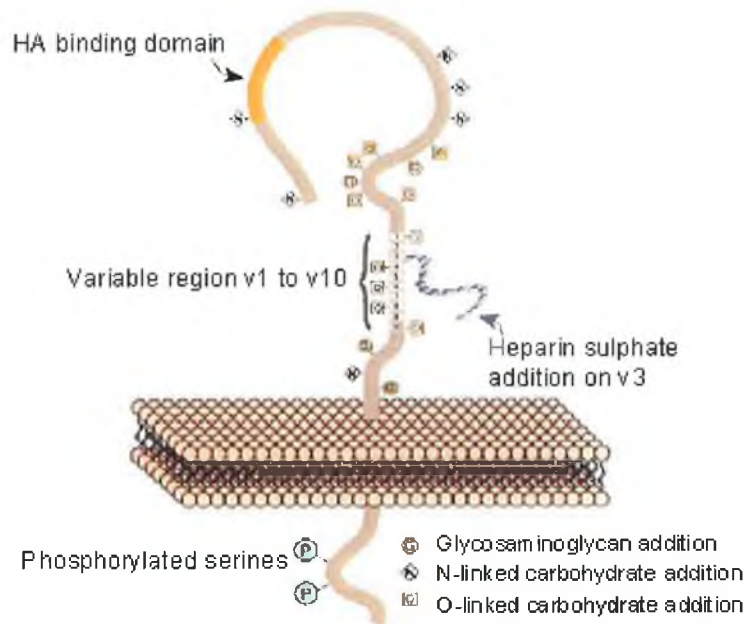
Figure 1.6 CD44 protein structure showing the HA binding domain and sites of glycosylation. Adapted from Ponta *et al.*, 1998.

1.3.2.2 CD44 functions and signalling

The functions of CD44 are exerted through the ligation of HA in the ECM. CD44 has a variety of other ligands due to its various isoforms, these include fibronectin, laminin and collagen type I and IV (Naor *et al.*, 1997). Phosphorylation at internal serine residues suggests a role for CD44 in signal transduction (Lesley *et al*, 1993). The cytoplasmic tail of CD44 has been shown to interact with the actin cytoskeleton through association with the ERM (ezrin, radixin, moesin) family of adaptor proteins (Tsukita *et al.*, 1994). ERM proteins function as molecular linkers between the plasma membrane and the actin cytoskeleton, and as such, have been implicated in a number of key physiological and pathological processes that involve cytoskeletal and membrane remodelling. These include cell-cell and cell-matrix adhesion interactions, cell motility, oncogenesis and apoptosis (Bretscher *et al.*, 2002 and Martin *et al.*, 2003).

1.4 The matrix metalloproteinases

The Matrix Metalloproteinases (MMPs) constitute a multigene family of over 23 structurally and functionally related secreted and membrane bound proteolytic enzymes. The MMPs are endopeptidases that process a wide range of pericellular substrates and they are primarily distinguished from other classes of proteinases by their dependence on a zinc ion for activity. Other distinguishing parameters associated with the MMPs include the requirement for pro-domain cleavage for enzyme activation and the inhibition of their enzymatic activity by the tissue inhibitors of metalloproteinases (TIMPs) (Nagase and Woesner, 1999 and McCawley and Matrisian, 2001). Inactive pro-MMPs are secreted as latent zymogens and their activation occurs extracellularly. Most MMPs are secreted but six contain transmembrane domains and are expressed as cell surface molecules. Included in the wide range of targets of MMP activity are ECM proteins, cell adhesion molecules and cell surface receptors. Collectively the MMPs are capable of degrading the components of the entire ECM.

ECM homeostasis is key to maintaining normal cellular function and the MMPs can therefore be described as modulators of cell function. ECM breakdown and remodelling are essential for many normal biological processes (e.g. embryonic development, blastocyte implantation, organ morphogenesis, nerve growth, ovulation, hair follicle cycling, bone remodelling, wound healing, apoptosis etc.) and pathological processes (e.g. arthritis and cancer). MMP degradation of ECM proteins regulates the behaviour of cells by allowing cells to interact with their immediate surroundings. Several other ways MMPs regulate cell behaviour include the activation of latent growth factors or other proteinases or through the cleavage of cell surface molecules (McCawley and Matrisian, 2001 and Lynch and Matrisian, 2002).

1.4.1 MMP structure and function

Each of the 23 members of the MMP family has distinct and often overlapping substrate specifities, however together the MMPs can degrade the entire ECM. Eight primary MMP structures exist all of which share several domain motifs. These shared stretches of sequence homology give the MMP family a somewhat conserved overall structure (Stocker *et al.*, 1995). These include a zinc-binding motif, a 'Met turn' and a

"Cysteine Switch". MMPs are therefore often grouped according to their modular domain structure (Figure 1.7).
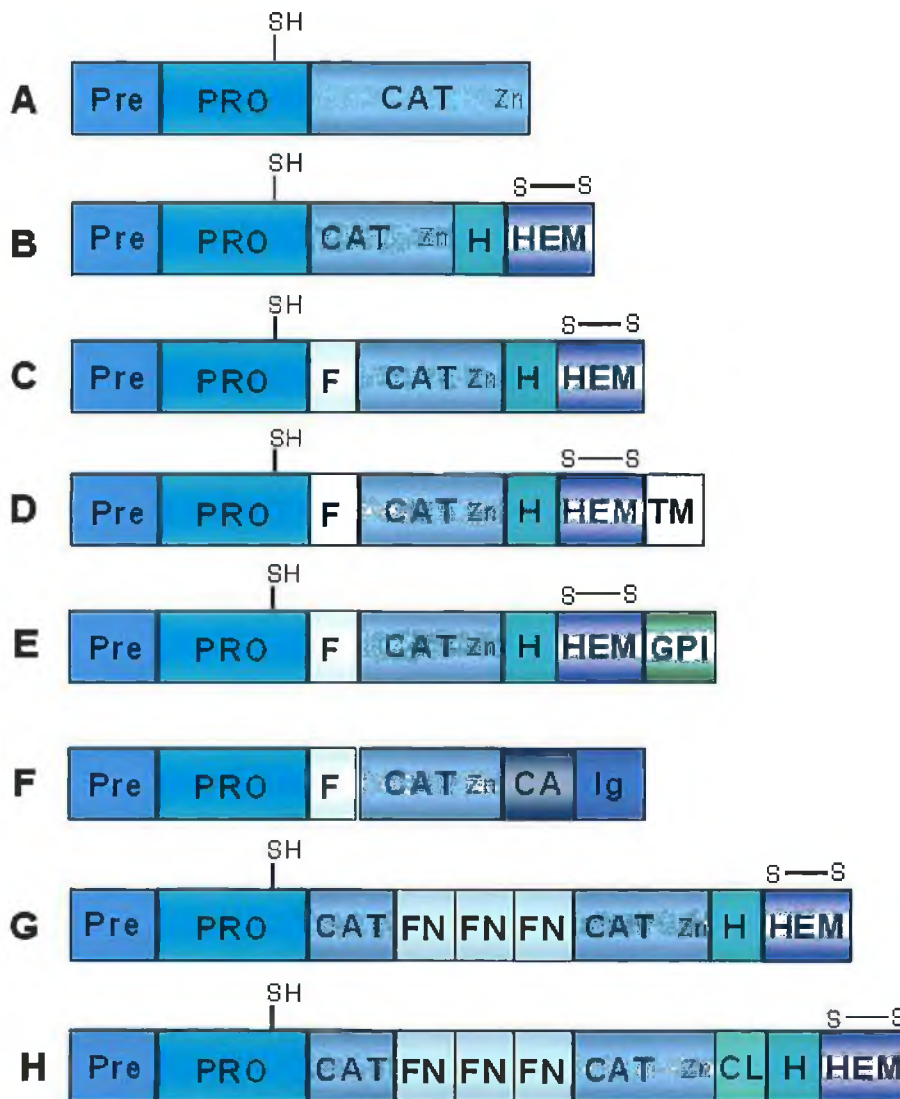
Figure 1.7 The MMP Domain Structures. (A) Minimal domain MMPs, e.g. MMP-7/matrilysin and MMP-26. (B) Simple hemopexin domain-containing MMPs, e.g. MMP-1, -8, -13, -3, -10, -12, -20, -19 and MMP–27. (C) Furin-activated MMPs, e.g. MMP-11 and MMP-28. (D) Transmembrane MMPs, e.g. MMP-14, -15, -16 and MMP-24. (E) GPI-linked MMPs, e.g. MMP-17 and MMP-25. (F) Immunoglobulin-like domain-containing MMPs, e.g. MMP-23. (G) and (H) Gelatin-binding MMPs, MMP-2/gelatinase A and MMP-9/gelatinase B respectively. Pre: leader/signal sequence, PRO: propeptide-domain with zinc-ligating thiol (SH) group, CA: cysteine array, CAT: catalytic domain, TM: transmembrane domain, Zn: zinc binding site, Ig: immunoglobulin-like domain, F: furin cleavage consensus sequence, FN: fibronectin type II inserts, H: hinge domain, CL: collagen-like domain, GPI: glycosyl phosphatidylinositol linkage signal, HEM: hemopexin domain. The hemopexin domain contains four repeats with the first and last joined by a disulfide bond (S-S). Adapted from (Brinckerhoff and Matrisian 2002).

11

MMP-7 (matrilysin) and MMP-26 (endometase, matrilysin-2) are the simplest and smallest members of the MMP family (Figure 1.7), containing only the minimum number of domains required for secretion and activity. It was originally thought that MMP-7 represented the primordial form of MMP and additional domains were added by exon shuffling (Matrisian, 1992). However, an extensive analysis of the sequences and structures of the MMP family indicates that the assembly of the catalytic and hemopexin domains into multidomain enzymes was an early evolutionary event, suggesting that the simpler MMPs such as MMP-7 might have arisen from larger complexes (Massova, *et al.*, 1998).

All MMPs have an N-terminal pre-domain, which directs their synthesis to the endoplasmic reticulum after which it is removed from the zymogen. Following the pre-domain is a pro-peptide domain which has a conserved and unique sequence; Pro-Arg-Cys-Gly-(Val/Asn)-Pro-Asp. The cysteine residue within this sequence ligates the catalytic zinc, maintaining latency in pro-MMPs. It is only when the Cys-Zn^{2+} interaction is disrupted and the pro domain is subsequently removed that the MMP becomes active. This mechanism of activation is known as the 'Cysteine Switch' (Becker *et al.*, 1995) and it produces active forms with lower molecular weights (Springman *et al.*, 1990 and Birkedal-Hansen *et al.*, 1993). The catalytic domain contains a zinc binding motif; His-Glu-X-X-His-X-X-Gly-X-X-His (where X represents any amino acid), and it also contains a conserved methionine residue which forms a distinct 'Met turn' structure (Nagase and Woessner, 1999 and Bode *et al.*, 1993). The binding of Zn^{2+} to the metal binding site in the catalytic domain is essential for proteolytic activity (and explains the word 'metallo' in the name). Substrate and cleavage site specificity are mostly dictated by the catalytic domain as well as binding sites termed exosites located outside the catalytic domain (Overall, 2002). The understanding of the catalytic site structure is used in the design of drugs that inhibit MMPs.

As seen in figure 1.7, MMP-2 and MMP–9 (gelatinase A and B, respectively) have three cysteine rich fibronectin-type II repeats within the catalytic domain, which interact with collagens and gelatins (Allan *et al.*, 1995 and Steffensen *et al.*, 1995). MMP-9 has a unique type V collagen like insert at the end of its hinge region and its function is unknown. Figure 1.7 also shows domain compositions B, C, G and H to have a C-terminal hemopexin domain, which is joined to the catalytic domain via a

hinge region. Though the function of this linker hinge region is unknown, it has been shown to be proline rich and using molecular modelling techniques, it has been hypothesised that it interacts with triple helical collagen (de Souza *et al.*, 1996). All MMPs, with the exception of MMP-7 and MMP-26 contain a hemopexin domain which is also an absolute requirement for collagenases to cleave triple helical interstitial collagens, otherwise proteolytic activity can only occur via the catalytic domain on non-collagen substrates (Clark and Cawston, 1989). The hemopexin domain also influences TIMP binding.

The membrane type matrix metalloproteinases (MT-MMPs) are comprised of 6 family members. The first MT-MMP cDNA (MT1-MMP/MMP-14) to be cloned encoded a protein of 63 kDa with a transmembrane domain and was later identified as a specific activator of pro-MMP-2 (Sato *et al.*, 1994). MT-MMPs contain a C-terminal hydrophobic stretch (transmembrane sequence or GPI anchor) that anchors the enzyme to the plasma membrane thus restricting its activity to the cell surface. MT-MMPs have varied substrate specificity and are capable of degrading fibronectin, laminins, and collagen types I, II and III. Expression of MT1-MMP cDNA in MDCK epithelial kidney cells gave rise to enhanced cellular invasion through reconstituted basement membrane *in vitro* (Hotary *et al.*, 2000). Five other MT-MMPs have since been identified and all six are listed in Table 1.1.

1.4.2 MMP substrates

Together, the MMP family can degrade all the components of the ECM. Each MMP has a unique yet slightly overlapping substrate specificity (Table 1.1). The MMPs are classified with respect to their protein domain characteristics but they were once classed into three subclasses according to their substrate specificity: 1) type I collagenases, 2) type IV collagenases and 3) stromelysins. Numerous MMP substrates have been tested and identified *in vitro* (Woessner and Nagase, 2000).

In addition to ECM components, MMPs can also cleave the pro-forms of other MMPs, and can activate and release latent and embedded growth factors. For example, MT1-MMP can activate pro-MMP-2 and MMP-9 has been shown to activate latent TGF-β (Sato *et al.*, 1994 and Yu and Stamankovic, 2000) thus stimulating angiogenesis and invasion. These wider roles for MMPs are often implicated in malignant tumours.

Table 1.1 Properties of the human matrix metalloproteinase family.

MMP Domain	Enzyme Name	MW kDa (latent)	MW kDa (active)	Substrates	Activation of other MMPs
MINIMAL					
MMP-7	Matrilysin, Pump-1	28	19	Proteoglycans, laminin, fibronectin, gelatins, collagen IV, elastin, entactin, tenascin HBEGF, TNFα, β4 integrin	MMP-1, MMP-2, MMP-9
MMP-26	Endometase	28	18	Collagen type IV, Fibronectin, Fibrin, Gelatin.	MMP-9
HEMOPEXIN					
MMP-1	Interstitial collagenase	55	45	Fibrillar collagens (types I,II,III,VII,X), gelatin, proteoglycans. Pro-TNFα	Unknown
MMP-8	Neutrophil collagenase	75	58	Collagens I, II, III.	Unknown
MMP-13	Collagenase-3	65	55	Collagen I, II, III, IV	Unknown
MMP-12	Metalloelastase	53	45/22	Elastin, fibronectin, collagen IV.	MMP-2, MMP-12
MMP-3	Stromelysin-1	57	45	Proteoglycans, laminin, fibronectin, collagen III,IV, V, X, gelatins. Plasminogen	MMP-1, MMP-8, MMP-9
MMP-10	Stromelysin-2	57	44	Proteoglycans, fibronectin, collagen III, IV, V, gelatins.	MMP-8
MMP-11	Stromelysin-3	51	44	Laminin, fibronectin (very weakly).	Unknown
FIBRONECTIN					
MMP-2	Gelatinase A	72	66	Gelatins, collagens IV, V, VII, X, elastin, fibronectin. TNFα	Unknown
MMP-9	Gelatinase B	92	86	Gelatins, collagens IV, V, elastin. TGFβ2	Unknown
TRANSMEMBRANE					
MMP-14	MT1-MMP	63	-	CD44	MMP-2
MMP-15	MT2-MMP	72	-	Unknown	Unknown
MMP-16	MT3-MMP	64	-	Unknown	MMP-2
MMP-17	MT4-MMP	70	-	Unknown, TNFα	Unknown
MMP-24	MT5-MMP	60	-	Proteoglycan ECM components.	Unknown
MMP-25	MT6-MMP	62	-	Unknown	MMP-2
MISCELLANEOUS					
MMP-19	RASI-1	-	-	Gelatins, Aggrecan, cartilage.	Unknown
MMP-20	Enamelysin	54	-	Amelogenin, Aggrecan, cartilage.	Unknown
MMP-21	Recently cloned MMP	-	-	Unknown	Unknown
MMP-22	Recently cloned MMP	-	-	Unknown	Unknown
MMP-23	Recently cloned MMP	28	-	Unknown	Unknown
MMP-27	Epilysin	59	-	Casein	Unknown

1.4.3 Regulation of MMP activity

MMP activity is tightly regulated. Constitutive levels of MMP expression are normally quite low, with the enzymes being induced under various normal physiological circumstances when ECM turnover is required, for example during wound healing or bone remodelling. To accomplish their normal functions MMPs must be present in the right cell type at the right time in the right amount and at the right extracellular location. In addition they must be activated or inhibited appropriately. Several MMPs are regulated separately and are tissue specific. Differential patterns of expression largely dictate the biological function of individual MMPs and differences in their temporal, spatial and inducible expression often indicate their unique roles. The complex and tight regulation of MMPs occurs at the transcriptional and post-transcriptional level and at a protein level by their inhibitors and activators (Sternlicht and Werb, 2001).

1.4.3.1 Transcriptional regulation of MMPs

Most MMPs are closely regulated at the level of transcription through various stimulatory and suppressive factors that influence numerous signalling pathways (Fini *et al.*, 1998). MMPs are regulated both positively and negatively at the transcriptional level by growth factors, cytokines, oncogenes, phorbol esters, integrin-derived signals, extracellular matrix proteins, cell stress and changes in cell shape (Fingleton and McDonnell, 1997, Birkedal-Hansen *et al.*, 1993, Crawford and Matrisian, 1996, Kheradmand *et al.*, 1998 and Sternlicht and Werb, 1999). Cytokines and growth factors that regulate MMP expression include interleukins, interferons, epidermal growth factor (EGF), basic fibroblast growth factor (bFGF), vascular endothelial growth factor (VEGF), tumour necrotic factor alpha (TNFα), TGFβ and the extracellular matrix metalloproteinase inducer (EMMPRIN) (Fini *et al.*, 1998). Many of these stimulate expression and/or activation of c-*jun* and c-*fos* proto-oncogene products, which heterodimerize and bind to activator protein-1 (AP-1) sites within the MMP promoters (Crawford and Matrisian, 1996).

1.4.3.2 Post-transcriptional regulation of MMP activity

Post-transcriptional control mechanisms are central to the regulation of MMP activity as all the soluble MMPs are secreted as latent zymogens that require activation through cleavage of the N-terminal pro-domain. The extracellular activation of most MMPs can be initiated by other already activated MMPs or by several serine proteinases that cleave peptide bonds within MMP pro-domains (Woessner and Nagase, 2000). Plasmin, which is produced by the action of plasminogen activator on plasminogen, converts both pro-MMP-1 and pro-MMP-3 to their active forms (He *et al.*, 1989). MT1-MMP has been shown to activate pro-MMP-2 using TIMP-2 as an activator by forming a trimolecular structure on the cell surface (Sato *et al.*, 1994 and Seiki, 1999). Sequences have been identified in the pro-peptide of MT1-MMP that are specifically cleaved by the pro-hormone convertase furin (Yana and Weiss, 2000). This intracellular cleavage represents the initiating event in a cascade that leads to the activation of MMP-2 (Itoh *et al.*, 2001). All these proteolytic events are tightly regulated in a cascade style system serving as a powerful and controlled mechanism co-ordinating complete breakdown of the multiple components of the ECM.

1.4.3.3 The role of TIMPs in the regulation of MMP activity

Proteolytic inhibitors play an important role in controlling the overall proteolytic activity in tissues. The proteolytic activity of MMPs is specifically inhibited by a family of low molecular weight proteins: the TIMPs (Brew *et al.*, 2000). There are four members in the TIMP family and they are produced by various cell types, including fibroblasts and endothelial cells. All TIMPs are secreted in soluble form with the exception of TIMP-3, which is bound to the ECM (Leco *et al.*, 1994). TIMPs inhibit MMP activity by reversibly binding them in a 1:1 stoichiometric manner. TIMPs also play a role in the activation of MMPs as exemplified above with the activation of pro-MMP-2. This activation is dependent on TIMP-2 concentration, and only occurs at low concentrations because MT1-MMP activity is also inhibited by TIMP-2 (Strongin *et al.*, 1995). Once generated, active MMP-2 is capable of auto-proteolysis, with further activation of the latent form. Such co-localisations may be a key mechanism for the regulation of MMP activity at the cell surface.

1.4.4 MMP activity in cancer

The roles of MMPs in cancer cells are discussed in detail in Chapter 3. Briefly, MMP expression is found more in malignant cancers than in normal, benign or premalignant tissues, with the highest expression occurring in areas of active invasion at the tumour-stroma interface (Sternlicht and Bergers, 2000). A correlation exists, in virtually all types of cancer between increased MMP levels and poor prognosis, therefore MMP levels are used routinely as prognostic markers for how aggressive a particular tumour is (Murray *et al.*, 1996 and Stetler-Stevenson *et al.*, 1996).

1.5 Summary

Advanced metastasis represents a major hurdle in cancer treatment mainly because its molecular mechanisms remain unclear. Tumour metastasis is a multistage process involving many participants including CAMs and proteolytic enzymes. CAMs have roles to play in adhesion to the ECM, cell signalling and ligation of growth factors. Proteolytic enzymes such as the MMPs are crucial to ECM turnover, growth factor activation and angiogenesis making them and the CAMs potential biomarkers of tumour progression. The interactions between CAMs such as integrins and CD44 with the MMPs remain unclear. New insights into these interactions and the definitive roles of CAMs and MMPs in metastasis are needed for the development of successful therapies.

1.6 Thesis overview and aims

The primary aim of this study was to elucidate the complexity of tumour cell metastasis and highlight the individual and collective importance of the many participating biomolecules in this process. The research presented in this thesis examined the molecular mechanisms involved in colorectal tumour invasion and metastasis. The thesis has been divided into three main chapters.

- Chapter 3:
  The aim of this study was to investigate the interactions between the cell adhesion molecule CD44 and the MMPs in regulating colon tumour cell invasion and migration. A panel of colorectal cancer cell lines were characterised based on their MMP and CAM expression profiles as well as their adhesive and *in vitro* invasive and migratory abilities. The effects of ECM components and the activation of CD44 on MMP activity and *in vitro* invasion were also examined.

- Chapter 4:
  The aim of this study was to identify novel biomarkers of metastasis using cDNA microarray analysis. Gene expression profiles were examined in a panel of colorectal cancer cell lines of varying invasive and metastatic abilities. Following data analysis, expression changes contributing to tumour cell metastasis could easily be identified. Several differentially expressed genes were selected for further validation and analysis.

- Chapter 5:
  The aim of this study was to further validate the importance of the genes identified in Chapter 4 by examining their expression in paired samples of normal and tumour colorectal tissues.

For convenience, this thesis has been divided into seven chapters. Chapter 1 serves as a general introduction to the process of tumour cell metastasis as well as highlighting the roles of its main players. There is a common materials and methods section (Chapter 2) and bibliography (Chapter 8). Chapters 3, 4 and 5 each have their own introduction, results and discussion sections, while chapter 6 provides an overall summary of the thesis.

Chapter 2

Materials and methods

2.1 Materials

All general-purpose chemicals and reagents used in experimental work were of analytical grade and were purchased from Sigma-Aldrich Inc., Dublin, Ireland.

All PCR primers were obtained from MWG Biotech, Ebersberg, Germany.

Anti-Actin antibody was obtained from Merck Biosciences, Nottingham, England.

Anti-CD44 antibodies, D2.1 and IM7.8.1 were obtained from Dr Aideen Long, Royal College of Surgeons, Dublin, Ireland and $F_{10-44-2}$ was obtained from Biodesign, Saco, Maine, USA.

Anti-MMP-7 polyclonal serum was obtained from Prof. Lynn Matrisian, Vanderbilt University, Nashville, TN, USA.

Anti-MMP-9 (MAC96) and MMP-2 (GL8) antibodies were obtained from Dr. Andrew Docherty, Celltech, Slough, England.

Anti-β1 integrin antibody (W1B10) was obtained from Sigma-Aldrich Inc., Dublin, Ireland.

Anti β-1 integrin antibody (AB1937) was obtained from Chemicon Inc., Hampshire, United Kingdom

BCA reagent for protein determination was obtained from Pierce Chemicals, Rockford, Illinois, USA.

BHK cells transfected with the human MMP-9 cDNA (BHK 92) were obtained from Prof. Dylan Edwards, University of East Anglia, Norwich, England.

Biocoat Matrigel invasion chambers were obtained from Becton Dickinson Labware, Bedford, MA, USA.

21

Blue/Orange Loading Dye (Cat # G1881) supplied by Promega, was distributed by Medical Supply Company, Dublin, Ireland.

Cell culture media and supplements including foetal bovine serum were supplied by Sigma-Aldrich Inc., Dublin, Ireland.

Cell freezing medium was obtained from Sigma-Alrdich Inc., Dublin, Ireland.

ColorBurst™ (Cat # C4105) electrophoresis marker was supplied by Sigma-Aldrich Inc., Dublin, Ireland.

Cy-3 and Cy-5 esters (Cat # PA23001, PA25001) were supplied by Amersham Biosciences Corp, NJ, USA.

DNA-free™ (Cat # 1906) was supplied by Ambion Inc., TX, USA.

G418 (geneticin) (Cat # US11379-1g) was supplied by Amersham Biosciences Corp, NJ, USA.

Glycogen (Cat # 9510) was supplied by Ambion Inc, Cambridgeshire, England.

L-glutamine, trypsin and penicillin/streptomycin were obtained from Sigma-Aldrich Inc., Dublin, Ireland.

MessageAmp™ aRNA system (Cat # 1750) including all the appropriate buffers and consumables, was supplied by Ambion Inc, TX, USA.

Microcon™ centrifugal filter devices were supplied by Amicon Ltd., Limerick, Ireland.

MMLV-RT supplied by Promega, was distributed by Medical Supply Company, Mulhuddart, Dublin, Ireland.

PBS tablets were obtained from Oxoid Ltd., Basingstoke, Hampshire, England.

Plastic consumables for cell culture, and 96-well plates were obtained from Sarstedt, Sinnottstown Lane, Drinagh, Co. Wexford, Ireland.

QIAquick PCR purification kit (Cat # 28106) including all the appropriate buffers and consumables was supplied by Qiagen Inc. CA, USA.

RNA Sample Loading Buffer (Cat # R4268) was supplied by Sigma-Aldrich Inc., Dublin, Ireland.

SuperScript[TM] II Reverse Transcriptase was supplied by Invitrogen, CA, USA.

Supersignal West ECL detection kit was obtained from Pierce Chemicals, Rockford, IL, USA.

SYBR green and all real time PCR reagents were obtained from Qiagen Inc. CA, USA.

Taq polymerase and other RT-PCR reaction components were purchased from Sigma-Aldrich Inc, Dublin, Ireland and Promega Corp, Southampton, England.

TRI-Reagent[TM] (Cat # T9424) was supplied by Sigma-Aldrich Inc., Dublin, Ireland.

X-ray film was obtained from Amersham, Little Chalfont, England.

2.2 Methods

2.2.1 Cell culture

All cells used in this study were adherent cell lines. The cell lines used were BHK92, SW480 (ATCC Cat # CCL-228), SW620 (ATCC Cat # CCL-227) and two transfected cell lines: SW480M7 (SW480 cells stably transfected with the cDNA for MMP-7) (Witty *et al.*, 1994) and SW480M9 (SW480 cells stably transfected with the cDNA for MMP-9) (Murray *et al.*, 2004). The SW480M7 cell line was received from Professor Matrisian's laboratory (Vanderbilt University, Nashville, TN). This cell line has been previously studied (Witty et al., 1994) but it was important to re-characterise these cells as they had been frozen down since 1995. The SW480M9 cell line was developed by Dr McDonnell whilst on sabbatical in Memorial Sloan Kettering Cancer Center, New York, NY, USA. However, the cells had not been extensively characterised. All cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM) supplemented with 5 % (v/v) foetal bovine serum (FBS) [DMEM S_5], 2 mM L-glutamine, 1 U/ml penicillin and 1 µg/ml streptomycin. Transfected cell lines were maintained in DMEM S_5 media containing 600 µg/ml G418. All cells were incubated in a humid 5 % CO_2 atmosphere at 37 °C in a Heraeus 6000 cell culture incubator. All cell culture techniques were performed in a sterile environment using a Holten HB225 laminar airflow cabinet. Cells were routinely grown in 25 cm^2 or 75 cm^2 tissue culture flasks. Cell growth and morphology were monitored using an Olympus CK2 inverted microscope.

2.2.1.1 Subculturing of cells

Once cells had grown to confluency, as judged microscopically, they were subcultured by trypsinisation. Firstly, all growth media was removed and cells were rinsed twice with 3 ml PBS to remove any residual FBS, which contains a trypsin inhibitor (α_2 macroglobulin). 3 ml of fresh trypsin-EDTA (0.025 % (w/v) trypsin with 0.02 % (w/v) EDTA in 0.15 M PBS, pH 7.4) was placed in each flask and incubated at 37 °C for 5-10 min or until all the cells had detached from the growth surface. 3 ml of DMEM S_5 growth medium was added to the cell suspension and the 5 ml was transferred to a sterile 25 ml universal container before being centrifuged at 2000 rpm for 5 min. Cells were resuspended in growth medium and an aliquot was kept for cell

counting (see section 2.2.1.2). Cells were seeded at $2 \times 10^5 - 1 \times 10^6$ cells/ml using 10 ml growth medium per 75 cm^2 culture flask and 5 ml per 25 cm^2 flask.

2.2.1.2 Cell counts

A Naubauer haemocytometer slide was used to perform cell counts. Trypan blue exclusion dye was routinely used to determine cell viability. 20 µl trypan blue was added to 100 µl cell suspension, mixed by vortexing, and the mixture left to incubate for 2 min at room temperature. A 20 µl sample of the mixture was applied to the counting chamber of the haemocytometer and the cells were visualised by light microscopy. Viable cells excluded the dye and remained clear while dead cells stained blue. Cells in the four outer quadrants were counted and an average obtained. The number of cells counted was determined as follows (average number of viable cells) x 1.2 (dilution factor) x 1×10^4 (area under coverslip) = viable cells/ml.

2.2.1.3 Cryopreservation: long term storage of cells

Long-term storage of cells was achieved by storing batches of cells in a Taylor-Wharton 35VHC cryofreezer (supplied by Cooper Cryoservice Ltd.) under liquid nitrogen at -196 °C. Cells to be stored were first grown until 80 % confluent, trypsinised, and resuspended at a concentration of 5×10^6 cells/ml in cell freezing medium (DMEM:FBS:DMSO; 45:45:10). 1 ml aliquots were transferred to cryovials and frozen at -20 °C for 30 min, -80 °C for 10 hr and finally under liquid nitrogen (-196 °C) for long term storage. 1 vial from each batch was randomly chosen and thawed to check the viability of the stocks. For recovery, cells were first removed from liquid nitrogen storage and thawed rapidly using a 37 °C waterbath. Each 1 ml aliquot was mixed with 5 ml of DMEM S_5 in a sterile 25 ml universal tube and centrifuged at 2000 rpm for 5 min. The pellet was resuspended in DMEM S_5 at 1×10^6 cells/ml and transferred to culture flasks and incubated at 37 °C and 5 % CO_2. An aliquot was also checked for cell viability as described in section 2.2.1.2.

2.2.2 Isolation of total RNA from cell lines

RNA was isolated from cells when 80 % confluent. The growth media was removed and cells were washed twice with PBS. Once all PBS was removed, 1 ml of

TRI-Reagent^{TM} was added to one 75 cm^2 flask of cells and left for 5 min at room temperature with occasional shaking. Once cells were lysed, as judged microscopically, the suspension formed was removed to a clean 1.5 ml microfuge tube, 200 μl chloroform was added and the mixture was shaken, left at room temperature for 15 min and centrifuged at 13,000 x g at 4 °C for 15 min. The upper aqueous layer was transferred to a fresh 1.5 ml tube. Care was taken not to transfer the DNA and protein-containing lower- and inter-phase. 0.5 ml ice-cold isopropanol was added to the aqueous phase, shaken and left to stand on ice for 10 min before it was centrifuged at 13,000 x g at 4 °C for 10 min. The supernatant was removed and 1 ml of sterile 75 % ethanol was added to wash the pellet by gentle vortexing and centrifugation at 7500 x g for 5 min. The ethanol was removed and the pellet was allowed to air-dry for 5 min. Pellets were resuspended in 50 μl 0.1 % DEPC treated H$_2$O by heating at 60 °C for 15 min. All RNA was stored at -80°C.

2.2.2.1 RNA quantitation and analysis

The absorbance values at 260 and 280 nm of a 1/100 dilution (5 μl in 500 μl 0.1 % DEPC treated H$_2$O) of total RNA were measured on a Shimadzu UV-160A spectrophotometer using a quartz cuvette. The absorbance value at 260 nm (A_{260}) was multiplied by 4, to give the RNA concentration in μg/μl (where 40 μg of RNA has an absorbance value of 1 at 260 nm and because of the dilution factor used). The A_{260}/A_{280} ratio gave an indication of RNA purity, where a value of 1.8 was desirable, and a value greater than 1.8 indicated protein contamination, a value lower than 1.8 was indicative of organic solvent contaminants.

RNA integrity was assessed by agarose electrophoresis. A 1 % agarose gel was prepared by boiling the agarose in 1 X Tris Borate EDTA (TBE) (0.08 M Tris-HCl pH 8.2; 0.04 M boric acid; 1 mM EDTA). After sufficient cooling, the gel was cast in a Hybaid Horizontal Gel Electrophoresis system. 1 μl RNA loading buffer (62.5 % (v/v) deionised formamide; 1.14 M formaldehyde; 200 μg/ml bromophenol blue; 200 μg/ml xylene cyanole; 1.25 X MOPS-EDTA-sodium acetate; 50 μg/ml ethidium bromide) was mixed with 2 μl RNA and heated at 65 °C for 10 min then cooled on ice prior to loading on the gel. The gel was run at 100 V in 1 X TBE. Before the dye front had run off the gel, the gel was stopped, removed and visualised on a UV transluminator. The presence of two and sometimes three strongly stained bands, representing the 28S, 18S and 6S

subunits, signified intact RNA. Degradation could be seen as a smear running down the length of the gel.

2.2.3 Reverse transcription (RT)

1 μg of total RNA (as prepared in section 2.2.2) was mixed with 1 μl oligo dT (500 μg/ml), the total volume was brought to 5 μl with sterile nuclease-free water, incubated at 70 °C for 10 min and then placed on ice. The following components were then added; 4 μl 5 X transcription buffer supplied with the enzyme (Promega), 1 μl RNasin (40 units/μl), 1 μl dNTP mix (each at 10 mM), 8 μl sterile water and 1 μl (10 units/μl) murine Moloney leukemia virus reverse transcriptase (MMLV-RT). Repeated pipetting was used to mix and the sample was briefly centrifuged (~5 sec) to collect the sample at the bottom of the PCR tube. The reaction mixture was incubated at 37 °C for 1 hr and heated to 95 °C for 2 min to inactivate the enzyme. The resulting cDNA was stored at 4 °C until required for PCR.

2.2.4 Real time PCR

Following reverse transcription, the following was added to 2 μl of cDNA template; 12.5 μl SYBR Green master mix that contained Taq and dNTPs (Qiagen), 8.5 μl Dnase-free water, and 1 μl each of forward and reverse primers. The sequences of the primers are listed in Table 2.1. Using a Rotor Gene™ 3000 multiplex system the reaction was held at 95 °C for 15 min followed by the following cycling conditions; 95 °C; 20 sec, annealing temperature (Table 2.1); 30 sec, 72 °C; 30 sec. This cycle was repeated 45 times before being held at 60 °C for 1 min. Following cycling, to ensure specificity, melt curve analysis was carried out starting at 50 °C and ramping to 90 °C. One peak in the melt curve indicated no secondary, non-specific products were formed.

Table 2.1 Details of the primers used for RT-PCR and real time PCR experiments. Primers were chosen so that they amplified the target mRNA specifically and not DNA.

Target	Primer Sequence	Product size	Annealing Temp (°C)	Reference
β Actin	5' ATC CTG CGT CTG GAC CTG GCT 3' 5' CTT GCT GAT CCA CAT CTG GTG 3'	682 bp	55	Murray *et al.*, 2004
Asparagine Synthetase	5' CGC CCA GAT TTT CTT CAA TC 3' 5' CCA AAG CAG CAG TTG GTG TA 3'	202 bp	56	Ma *et al.*, 2003
Cathepsin C	5' CCA CAG TCG AAA AAT CCC AAG G 3' 5' GTG GTG GTA GAT CCC CTT TTT G 3'	549 bp	60	Tokuriki, *et al.*, 2003
Caveolin 1	5' GTC AAC CGC GAC CCT AAA C 3' 5' TGT CCC TTC TGG TTC TGC AAG 3'	80 bp	60	Uray *et al.*, 2003
CD44 C13 - C2a	5' AAG ACA TCT ACC CCA GCA AC 3' 5' CCA AGA TGA TCA GCC ATT CTG G 3'	multiple	56	Van Weering *et al.*, 1993
CD44v6 v6 - C2a	5' CAG GCA ACT CCT AGT AGT AC 3' 5' CCA AGA TGA TCA GCC ATT CTG G 3'	multiple	56	Van Weering *et al.*, 1993
MMP-7	5' TGT ATC CAA CCT ATG GAA ATG 3' 5'CAT TTA TTG ACA TCT ACG CGC 3'	341 bp	52	Witty *et al.*, 1994
MMP-9	5 ATG AGT TCG GCC ACG CGC TGG GCT T 3' 5' T GCC GGT GAT GAC ACG GAA ACT CAG 3'	329 bp	54	Onisto *et al.*, 1993
PPAR γ	5' AGT GGG GAT GTC TCA TAA TGC C 3' 3' AGC TCA GCG GAC TCT GGA TTC 5'	113 bp	56	Pavan *et al.*, 2003
Tenascin C	5' GAG ATT TAG CCG TGT CTG AGG TTG 3' 5' GGT AGC CAT CCA GGA GAG ATT GAA G 3'	331 bp	70	Nagata *et al.*, 2003
Tenascin C 8-18	5' CAA TCC AGC GAC CAT CAA CG 3' 5' CGT CCA CAG TTA CCA TGG AG 3'	multiple	58	Adams *et al.*, 2002
Tenascin C 9/17-18	5' CAT CCA CTG CCA TGG GCT C 3' 5' CGT CCA CAG TTA CCA TGG AG 3'	149 bp	58	Adams *et al.*, 2002
Tenascin C 9/16-18	5' GCA TCC ACT GAA GCC GAA C 3' 5' CGT CCA CAG TTA CCA TGG AG 3'	431 bp	58	Adams *et al.*, 2002
Tenascin C 8-14/16	5' CAA TCC AGC GAC CAT CAA CG 3' 5' TTC GGC TTC TGT CGT GGC 3'	multiple	58	Adams *et al.*, 2002

2.2.5 Collection of conditioned media

When cells had grown to 80 % confluency, all media was removed and cells were rinsed twice in 3 ml PBS before 5 ml serum free media (S_0) was added. The cells were incubated under standard conditions for 3 hr, thus starving the cells. The S_0 media was then replaced with 5 ml fresh S_0 and the cells incubated for a further 72 hr. Conditioned media (CM) was removed and centrifuged at 2,000 rpm for 5 min to remove debris. CM was concentrated using Microcon^{TM} centrifugal filter devices with a 10 kDa cut-off. Briefly, 500 µl of CM was added to the filter and centrifuged at 10,000 rpm for 15 min. Following centrifugation, the retained media was recovered and protein concentration was determined using the bicinchoninic acid (BCA) assay (Pierce).

2.2.6 Bicinchonic acid (BCA) assay

The BCA assay was carried out to determine the total protein concentration in a sample. In this assay, Copper (Cu^{2+}) reacts with protein under alkaline conditions to give Cu^{1+}, which in turn interacts with BCA to produce a BCA-copper purple coloured complex that can be easily detected spectrophotometrically. Protein concentration was determined by preparing a standard curve from BSA standards ranging from 0.0 to 2.0 mg/ml at 0.2 mg increments in distilled water. A BCA working solution was prepared using a commercially available kit, by combining 50 parts reagent **A**, an alkaline bicarbonate solution with 1 part reagent **B**, a copper sulphate solution. 200 µl of this solution was added to 10 µl sample or standard in a 96-well microtitre plate in triplicate. The plate was covered with aluminium foil and incubated at 37 °C for 30 min. The absorbance of each well was read at 560 nm using a Rosys Anthos 2010 plate reader.

2.2.7 SDS-PAGE

Sodium dodecylsulphate polyacrylamide gel electrophoresis (SDS-PAGE) was performed using the discontinuous system described by Laemmli (1970) using 10 % polyacrylamide gels and 3 % stacking gels prepared as follows:

10 % Resolving Gel   :   3.3 ml 30 % (w/v) Acrylamide containing 0.8 % (w/v)
bisacrylamide

4 ml Distilled Water

2.5 ml 1.5 M Tris-HCl; pH 8.8

100 μl 10 % SDS

100 μl 10 % Ammonium persulphate (freshly prepared)

20 μl TEMED

The resolving gel was poured, covered with a layer of 100 % ethanol and allowed to set before the stacking gel was cast.

3 % Stacking Gel   :   0.8 ml (0.5 M Tris-HCl, pH 6.8; 0.4 % (w/v) SDS)

0.5 ml 30 % (w/v) Acrylamide containing 0.8 % (w/v)
bisacrylamide

2 ml Distilled water

33 μl 10 % Ammonium persulphate (freshly prepared)

20 μl TEMED

For western blot analysis 80 μg of total protein was loaded on the gel. Samples were mixed 4:1 with sample buffer (2 % (w/v) SDS; 0.08 M Tris-HCl, pH 6.8; 10 % (w/v) glycerol; 0.2 % (w/v) Coomassie Brilliant Blue; 1 M β-mercaptoethanol) and boiled at 100 °C for 10 min. Samples were loaded, together with 10 μl of coloured marker (see Table 2.2 for molecular weights) and the gel was run at a constant voltage of 100 V (~50 mA) in laemmli buffer (0.025 M Tris-HCl, pH 8.3; 0.19 M glycine; 0.1 % SDS) using an Atto vertical mini-electrophoresis system until the blue dye front reached the bottom of the gel.

2.2.8 Western blot analysis

Following electrophoresis as described in section 2.2.7 the gels were then transferred to nitrocellulose. For each gel, one sheet of nitrocellulose and 4 sheets of Whatman-1 filter paper were pre-soaked in cold transfer buffer (0.025 M Tris, pH 8.3; 0.192 M Glycine; 20 % (v/v) Methanol). Proteins were then transferred to nitrocellulose for 1 hr at 100 V at 4 °C. Transfer was confirmed by the presence of the colour marker on the nitrocellulose membrane and by the presence of protein bands

when the blot was stained with a 0.1 % Ponceau S solution in 5 % (v/v) acetic acid for 10 min at room temperature.

Table 2.2 Molecular weight distributions of protein marker used for western blots.

Band Colour	Molecular Weight (kDa)
Violet	220
Pink	100
Blue	60
Pink	45
Orange	30
Blue	20
Pink	12
Blue	8

Following transfer, the blot was first rinsed in TBST (10 mM Tris, pH 8.0; 150 mM NaCl; 0.05 % (v/v) Tween 20) to remove any Ponceau S, and then blocked using 5 % dried milk in TBST for 2 hr at room temperature with gentle shaking and then incubated overnight at 4 °C with primary antibody diluted in the same blocking solution. The primary antibody dilutions used are listed in Table 2.3. Blots were washed three times for 15 min in TBST at room temperature with gentle rocking. Blots were then incubated at room temperature with gentle rocking for 2 hr with secondary HRP conjugated antibody, diluted in TBST. Secondary antibodies used are listed in Table 2.3. Blots were washed three times for 15 min in TBST and ECL (Pierce) was used to develop them. 1 ml of 1:1 H_2O_2 : Luminol mixture was poured over the blot. The blot was covered with a sheet of transparent acetate and under red light a sheet of Hyperfilm^{TM} (Amersham) was placed over the blot for 5 min.

Table 2.3 Details of antibodies used for western blot analysis.

Target	Primary Antibody	Secondary Antibody	Tertiary Component
MMP-2	GL8 (Celltech) 1/2000	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a
MMP-7	Rat anti-human MMP-7 Polyclonal sera used neat	BA-4000 (Vector) Biotinylated anti-Rat IgG 1/15,000	SA-5004 (Vector) HRP-conjugated streptavidin 1/20,000
MMP-9	MAC96 (Celltech) 1/500	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a
CD44	D2.1 1/1000	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a
CD44v3	3G5 (R&D Systems) 1/500	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a
CD44v6	2F10 (R&D Systems) 1/100	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a
β1 Integrin	AB1937 (Chemicon) 1/10,000	P0448 (DAKO) Goat Anti-Rabbit 1/2000	n/a
β-Actin	JLA20 (Merck) 1/10,000	NA931 (Amersham) Anti-Mouse HRP 1/5000	n/a

The films were developed using a Hyper™ processor (Amersham). By lining up the film with the blot, bands were identified by molecular weight using the colour marker. Film was photographed using the Pharmacia Biotech Image Master system and images were analysed using Pharmacia Biotech Image Master 1D Prime densitometry package for Windows.

2.2.9 Zymography

Gelatinase activity was assessed in 50 µg total protein. 10 % acrylamide gels were prepared for MMP-2 and MMP-9 analysis. The following were the volumes used to prepare one gel;

Resolving Gel :   2.5 ml (1.5 M Tris-HCl, pH 8.8; 0.4 % (w/v) SDS)

2.5 ml 3 mg/ml Gelatin stock

3.3 ml 30 % (w/v) Acrylamide containing 0.8 % (w/v) bisacrylamide

1.7 ml Distilled water

33 µl 10 % Ammonium persulphate (freshly prepared)

20 µl TEMED

The resolving gel was poured, covered with a layer of 100 % ethanol and allowed to set before the stacking gel was cast.

Stacking Gel   :   0.8 ml (0.5 M Tris-HCl, pH 6.8; 0.4 % (w/v) SDS)

0.5 ml 30 % (w/v) Acrylamide containing 0.8 %(w/v) bisacrylamide

2 ml Distilled water

33 µl 10 % Ammonium persulphate (freshly prepared)

20 µl TEMED

Once set, sample wells were washed with laemmli buffer (0.025 M Tris-HCl, pH 8.3; 0.19 M glycine; 0.1 % SDS) to remove any excess ammonium persulphate. Samples containing 50 µg total protein were mixed 4:1 with sample buffer (15 % sucrose; 0.25 M Tris-HCl, pH 6.8; 0.1 % (w/v) bromophenol blue). Samples were loaded together with a 20 µl aliquot of conditioned BHK92 media used as a positive control. A molecular weight marker, the sizes of which are listed in Table 2.4 was also loaded. Gels were run at a constant voltage of 100 V (~50 mA) in laemmli buffer. Gels were run for ~2 hr or until the blue dye front had reached the bottom of the gel, during this time the current dropped to ~20 mA. The resolving gel was removed and washed twice in 2.5 % Triton-X 100 for 30 min and then rinsed in substrate buffer (50 mM Tris-HCl, pH 8.0; 5 mM $CaCl_2$) before being submerged in the same buffer and incubated at

37 °C for 48 hr in a sealed container. Following incubation the gels were stained with Coomassie blue (0.25 % in 6:3:1 water:propanol:glacial:acetic acid) for two hours with gentle shaking and destained using the same solvent system with gentle shaking until clear bands appeared. The gels were photographed using a Pharmacia Biotech Image Master system and images were analysed using Pharmacia Biotech Image Master 1D Prime densitometry package for Windows.

Table 2.4 Molecular weight distributions of protein marker used for zymography.

Proteins	Molecular Weight (kDa)
Myosin, rabbit muscle	205
β-Galactosidase, *E. coli*	116
Phosphorylase b, rabbit muscle	97
Fructose-6-phosphate Kinase, rabbit muscle	84
Albumin, bovine serum	66
Glutamic Dehydrogenase, bovine liver	55
Ovalbumin, chicken egg	45
Glyceraldehyde-3-phosphate Dehydrogenase, rabbit muscle	36
Carbonic Anhydrase, bovine erythrocytes	29
Trypsinogen, bovine pancreas	24
α-Lactalbumin, bovine milk	14.2
Aprotinin, bovine lung	6.5

2.2.10 *In vitro* invasion assays

Biocoat Matrigel invasion chambers were used to investigate and compare the *in vitro* invasive activities of the cell lines used in this study. The 24-well invasion chambers were removed from -20 °C storage and allowed to come to room temperature for 10 min. The inserts were rehydrated by adding 250 µl warm (37 °C) S_0 to each chamber and left for 2 hr at room temperature. After rehydration the media was removed from the inserts and replaced with 500 µl of a 5×10^5 cells/ml cell suspension in S_0. 750 µl of DMEM containing 20 % FBS (S_{20}) was added to the outer chambers and acted as a chemoattractant for the cells. The plates were then incubated for 24 hr in a 5 % CO_2 humidified 37 °C incubator.

To investigate the effect of HA and CD44 activation on cell invasion, cells were seeded in serum free media containing 0.25 mg/ml HA in both upper and lower

chambers. Cells were also allowed to invade in the presence of 50 µg/ml CD44 activating antibody $F_{10\text{-}44\text{-}2}$, CD44 blocking antibody IM7.8.1, or mouse IgG (Sigma) as a control.

Following incubation, non-invasive cells were removed from the upper chamber using cotton swabs soaked in PBS. Cells that had invaded through the membrane were fixed by soaking in methanol for 10 min. Cells were then stained in 0.75 % Mayers Haematoxylin solution for 5 min. The inserts were then washed several times in tap water to 'blue' the dye. The inserts were then dehydrated by soaking in a series of solutions of increasing ethanol concentration for two minutes each. 25 %, 50 %, 75 % and 100 % ethanol solution were used before the inserts were finally soaked in 100 % xylene solution. The membrane was carefully removed from the housing of the insert using a scalpel blade and mounted on a slide with the outside facing up using DPX mounting medium. Cells were visualised at 10 X magnification and the number of cells in 5 fields were counted and an average was obtained.

2.2.11 *In vitro* migration assays

The *in vitro* migratory capacities of cells was assessed using similar cell culture inserts to those used in section 2.2.10, with the exception that these inserts lacked the matrigel coating. Following trypsinisation and cell counts, 5×10^5 cells were added into the upper chamber of these inserts in 500 µl serum-free media in triplicate. The inserts were then inserted into 24-well culture plates containing 750 µl DMEM with 20 % FBS (S_{20}) and allowed to incubate at 37 °C and 5 % CO_2 for 24 hr. Following incubation, the migratory cells were enumerated by removing the inserts and staining and counting as outlined in section 2.2.10.

2.2.12 Cell adhesion assay

Separate wells of a sterile 96-well microtitre plate were coated overnight at 4 °C by adding 100 µl of the following ECM components; hyaluronic acid (HA) 1 mg/ml or fibronectin 50 µg/ml. Adhesion to plastic was also monitored using uncoated wells. All wells were washed twice with 100 µl PBS and blocked with 100 µl 1 mg/ml BSA for 2 hr at 4 °C. Following trypsinisation and counting of semi-confluent (~80 %) cells, 1 x 10^5 cells in 100 µl growth medium (S_5) were added to each well and allowed to adhere

for 90 min at 4 °C. Growth media and cells that had not adhered were removed by inverting and flicking the 96-well plate once. Plates were centrifuged at 150 x g for 5 min to collect the adherent cells at the bottom of the well. Residual media was aspirated and 120 μl MTS (1/6 dilution) in S_5 growth media was added to each well and incubated at 37 °C, 5 % CO_2 for 2 hr. For % adhesion calculations 120 μl of the above MTS dilution was added to wells containing 1 x 10^5 cells which had not been inverted, these acted as controls for 100 % adhesion. After 2 hr, when a colour change (yellow to red/brown) occurred, the absorbance at 492 nm was read using a Rosys Anthos 2010 plate reader. % adhesion was calculated and statistical analysis was carried out using Microsoft Excel.

2.2.13 Treatment of cells with antibodies and ECM components

Cells were passaged at least twice after thawing prior to any treatment. To investigate the effect of CD44 activation on MMP activity, cells were first grown until 80 % confluent and the media was replaced with serum free media for 3 hr. Cells were then incubated with fresh serum free media containing 50 μg/ml CD44 activating antibody $F_{10-44-2}$ (Biodesign) or mouse IgG (Sigma) as control. Replicates were prepared and media and RNA were harvested at times 0, 6, 24 and 48 hr. Cells were also treated with 0.25 mg/ml HA, 0.01 mg/ml Collagen type IV 10 mg/ml FN or 0.3 μg/ml anti-β_1-integrin stimulating antibody (W1B10, Sigma) in serum free media for 72 hr. All harvested media was centrifuged at 2000 rpm for 5 min to remove debris. Supernatant was stored at 4 °C prior to further analysis. RNA was also harvested as described in section 2.2.2 for further analysis.

2.2.14 Protein extraction from cells

Total cellular protein were extracted for the analysis of membrane bound cell adhesion molecules by western blot. Following trypsination of 1 x 10^7 cells, the cell pellet was resuspended in 100 μl lysis buffer (1 % Triton-X-100, 150 mM NaCl, 10 mM Tris-HCl; pH 7.0, 10 mM EDTA, 1 mM 4-(2-aminoethyl)-benzene-sulfonylfluoride (AEBSF), 0.8 μM Aprotinin, 20 μM Leupeptin, 40 μM Bestatin, 15 μM Pepstatin A, 14 μM E-64). The mixture was placed on ice for 1 hr with occasional mixing. Samples were then centrifuged at 12,000 x g for 20 min to remove cell debris. The supernatant

was aliquoted into clean microfuge tubes and protein concentration was determined using the BCA assay.

2.2.15 Polymerase chain reaction (PCR)

A 45 µl PCR mix was prepared by adding 34.5 µl sterile nuclease-free water, 5 µl 10 X $MgCl_2$-free reaction buffer (supplied with the Taq polymerase enzyme), 2 µl 25 mM $MgCl_2$, 1 µl dNTP mix (each at 10 mM), 0.5 µl Taq polymerase (5 U/µl), 1 µl forward primer and 1 µl reverse primer to a sterile 0.2 ml PCR tube. All primers used were at 100 pmol/µl concentration and are listed in Table 2.1. 5 µl cDNA as prepared in section 2.2.3 was added to this reaction mixture and the tube was briefly centrifuged to collect the sample at the bottom. Mineral oil was not required as the lid of the Eppendorf mastercycler used was heated to 105 °C to prevent condensation. The tube was placed in the thermocycler and programmed with an initial temperature of 94 °C for 3 min, followed by 30 cycles consisting of the following sequential steps: 94 °C for 90 sec (denaturation), annealing temperature, as listed in Table 2.1, for 3 min and 72 °C for 3 min (elongation). The reaction was finally held at 72 °C for a further 7 min before tubes were removed and stored at 4 °C until further analysis.

2.2.15.1 PCR analysis of CD44 splice products

To analyse CD44 variant 6 expression C13 and C2a primers that anneal to the constant region of CD44 were used to carry out an initial PCR using the conditions listed in Table 2.1. 2 µl of this reaction was then used as template for exon specific PCR, where the C2a primer was used as the reverse primer in conjunction with a primer specific for the variant 6 region.

2.2.15.2 Analysis of PCR products by agarose gel electrophoresis

A 1.5 % gel was prepared by boiling the required amount of agarose in TBE. The agarose was allowed to cool sufficiently and 1 µl ethidium bromide (10 mg/ml) was added before the gel was poured and cast in a Hybaid horizontal gel electrophoresis system. 10 µl of PCR product was mixed with 3 µl DNA loading dye (Promega) and loaded onto the gel together with a 100 bp ladder. The gel was run at 100 V until the dye front had reached the end of the gel. The gel was removed and bands were

visualised and photographed under UV using a Pharmacia Biotech Image Master system.

2.2.16 cDNA microarray analysis

cDNA microarray analysis was carried out at George Mason University (GMU), Virginia (VA), USA to simultaneously compare the expression of 40,000 genes in all cell lines used in this study.

2.2.16.1 Microarray fabrication

The microarrays used in this study were manufactured at GMU, VA, USA. Briefly, a microarray was constructed from 39,360 human cDNA clones (Research Genetics: ftp://ftp.resgen.com/pub/sv_libraries/RG_Hs_seq_ver_060101.txt, 17 September 2004). The complete list of genes with accession numbers is published at http://www.gmu.edu/centers/genomics/keys, 17 September 2004. The microarray was distributed over two slides because of the large number of cDNA clones. A Beckman Biomek 2000 Laboratory Automation Workstation was used to amplify cDNA inserts directly from clones in culture using GF200F (5'-CTGCAAGGCGATTAAGTTGGGTAAC) and GF200R (5'-GTGAGCGGATAACAATTTCACACAGGAAACAGC) universal primers. Selected aliquots of PCR amplifications were electrophoresed on agarose gels in order to monitor the yield and the specificity of the amplification reaction. Amplification products were purified using Multiscreen-PCR plates (Millipore), then dried and resuspended in 30 μl of 3 X saline sodium citrate (SSC). Selected aliquots were monitored by agarose gel electrophoresis after purification. The collection of amplified cDNAs were printed on poly-L-lysine coated slides in a single replicate using Gene Machines OGR-03 OmniGrid Microarrayer with SMP3 pins (Telechem International). Negative controls consisting of no-template PCR amplifications were also printed on the microarray as well as blank controls (areas on which nothing was printed but were quantified as normal spots). Blanks and negative controls were distributed in several sub-arrays to monitor different areas on the slide surface. Spotted slides were stored at room temperature under vacuum until further use.

2.2.16.2 DNase treatment of RNA

Following isolation of total RNA from cells, contaminating DNA was removed using the Ambion DNA-free™ system. 25 µl (~40 µg) of total RNA, 1 µl DNase 1 (2 units) and 2.5 µl 10 X DNase buffer were mixed gently and incubated at 37 °C for 30 min. 5 µl DNase inactivation reagent was added, mixed and incubated at room temperature for 2 min. The mixture was centrifuged at 10,000 x g for 2 min to pellet the DNase inactivation suspension. The RNA supernatant was removed to a fresh tube and stored at -80 °C until further use. RNA quantitation and analysis was carried out as described in section 2.2.2.1.

2.2.16.3 RNA amplification

RNA was amplified using the Ambion MessageAmp™ aRNA system. This contained all the reagents and consumables required including buffers and purification cartridges. This method consists of reverse transcription with an oligo(dT) primer bearing a T7 promoter followed by *in vitro* transcription of the resulting DNA with T7 RNA polymerase to generate hundreds to thousands of antisense RNA copies of each mRNA in a sample.

2.2.16.3.1 First strand synthesis

5 µg of DNase treated total RNA, isolated from cells as described in section 2.2.16.2 and quantitated as in section 2.2.2.1, was mixed with 1 µg of T7 oligo(dT) primer and the final volume brought to 12 µl with nuclease-free water. The mixture was vortexed and centrifuged briefly (~5 sec) to collect the sample at the bottom of the microfuge tube and heated at 70 °C for 10 min, then centrifuged briefly and placed on ice. The following were then added and mixed by gentle repeat pipetting; 2 µl 10 X first strand buffer, 1 µl ribonuclease inhibitor, 4 µl dNTP mix, and 1 µl reverse transcriptase to give a final volume of 20 µl. The mixture was then incubated at 42 °C for 2 hr and briefly centrifuged then placed on ice before immediately proceeding to the second strand synthesis step.

2.2.16.3.2 Second strand synthesis

The following reagents were added (in order) to the above first strand synthesis reaction with mixing by gentle pipetting up and down; 63 µl nuclease-free water, 10 µl 10 X second strand synthesis buffer, 4 µl dNTP mix, 2 µl DNA polymerase, and 1 µl RNaseH. The mixture was briefly centrifuged then incubated at 16 °C for 2 hr after which it was stored on ice or at -20 °C before the cDNA purification step.

2.2.16.3.3 cDNA purification

cDNA filter cartridges were equilibrated with 50 µl cDNA binding buffer for 5 min at room temperature. 250 µl cDNA binding buffer was added to the cDNA samples above and mixed by gentle vortexing. The cDNA sample/cDNA binding buffer mixture was applied to the equilibrated cartridge, which was firmly seated in its 2 ml wash tube and centrifuged for 2 min at 10,000 x g. The flow-through was discarded and 500 µl of ethanol-containing cDNA wash buffer was applied to the cartridge, which was replaced in the wash tube and centrifuged as above. The flow-through was discarded and the cartridge was replaced in the wash tube and centrifuged again to remove any trace ethanol. The cDNA filter cartridge was transferred to a 1.5 ml cDNA elution tube and 10 µl of nuclease-free water (preheated to 50 °C) was applied to the centre of the filter and left at room temperature for 2 min before being centrifuged as above. The elution step was repeated to give an eluate of ~16 µl containing dsDNA. The eluate was then placed on ice or stored at -20 °C before proceeding to the *in vitro* transcription step.

2.2.16.3.4 *In vitro* transcription to synthesise amplified RNA (aRNA)

The following was added with mixing by gentle pipetting to the 16 µl of the double stranded cDNA from the previous step; 4 µl ATP solution, 4 µl GTP solution, 4 µl TTP solution, 4 µl CTP solution, 4 µl UTP solution, 4 µl T7 10 X reaction buffer and 4 µl T7 enzyme mix. The mixture was centrifuged briefly to collect it at the bottom of the tube before being incubated at 37 °C for 12 hrs. 60 µl elution solution was added and mixed by gentle pipetting and vortexing.

2.2.16.3.5 aRNA purification

Following *in vitro* transcription, all enzymes, salts and unincorporated nucleotides were removed from the aRNA sample. 350 µl aRNA binding buffer was added and mixed thoroughly followed by the addition of 250 µl 100 % ethanol. The entire mixture was applied to an aRNA filter cartridge, which was placed in an aRNA collection tube and centrifuged for 2 min at 10,000 x g. The flow-through was discarded and 650 µl of ethanol-containing aRNA wash buffer was applied to the cartridge which was replaced in the collection tube and centrifuged at 10,000 x g for 2 min. The flow-through was discarded and the cartridge was centrifuged briefly to remove any trace ethanol. The cartridge was placed in a fresh collection tube. 50 µl elution solution (0.1 M EDTA in nuclease-free water) preheated to 50 °C was applied to the centre of the cartridge and left for 2 min at room temperature before being centrifuged for 2 min at 10,000 x g. The elution was repeated with a second 50 µl aliquot of elution solution to give an aRNA solution of ~100 µl. All aRNA was aliquoted and stored at -80 °C until further use. RNA quantitation and analysis was carried out as described in section 2.2.2.1.

2.2.16.4 Aminoallyl labelling of RNA

RNA was labelled with aminoallyl labelled nucleotides via first strand synthesis followed by a coupling of the aminoallyl groups to either Cyanine 3 or 5 (Cy 3/Cy5) fluorescent molecules.

2.2.16.4.1 Reagent preparation

A 50 X aminoallyl dNTP labelling mix was prepared with a 2:3 dUTP:dCTP ratio by mixing 5 µl dATP, dCTP, dGTP, 3 µl dTTP and 2 µl aa-dUTP (where each stock was 100 mM). The solution was stored at -20 °C.

2.2.16.4.2 First stand synthesis and aminoallyl labelling

2 µl random hexamer primers (3 mg/ml) was added to 4 µg of DNase treated aRNA (as prepared in sections 2.2.16.2 and 2.2.16.3 and quantitated as in section 2.2.2.1) and the final volume brought to 18.5 µl with RNase free water. The sample

was mixed well, incubated at 70 °C for 10 min, centrifuged briefly and then placed on ice. 6 μl 5 X first strand buffer, 3 μl 0.1M DTT, 0.6 μl 50 X aminoallyl-dNTP mix (as prepared above) and 2 μl SuperScript™ II RT (200 U/μl) were added, mixed and incubated at 42 °C for 15 hr. RNA was hydrolysed by the addition of 10 μl 1 M NaOH and 10 μl 0.5 M EDTA. Samples were mixed and incubated at 65 °C for 15 min. The pH was then neutralised by adding 10 μl 1 M HCl.

2.2.16.4.3 Reaction purification I: removal of unincorporated aa-dUTP

Amicon centrifugal filter devices were used to purify the aminoallyl labelled cDNA. 450 μl of sterile water was added to the contents of the reaction and the entire mixture was applied to the filter, which was placed in a 1.5 ml tube and centrifuged at 12,000 x g for 8 min. The flow-through was discarded and the above wash step was repeated. The cartridge was inverted and spun at 12,000 x g for 2 min in a fresh 1.5 ml tube to collect the retentate. The samples were then dried using a speed vac.

2.2.16.4.4 Coupling of aa-cDNA to Cy dye ester

Aminoallyl labelled cDNA samples were resuspended in 4.5 μl freshly prepared 0.05 M NaHCO$_3$, pH 9.0. 4.5 μl of the appropriate Cy dye ester was added (i.e. Cy 3 (green) to control and Cy 5 (red) to 'treated') and all tubes were wrapped in aluminium foil. The reaction was incubated for 4 hr in the dark at room temperature.

2.2.16.4.5 Reaction purification II: removal of uncoupled dye

A QIAquick PCR purification kit (Qiagen) was used to remove uncoupled dye. This kit included all buffers required for the procedure. 35 μl 100 mM NaOAc pH 5.2 was added followed by 250 μl PB (binding) buffer. The sample was applied to a QIAquick spin column in a 2 ml collection tube and centrifuged at 13,000 x g for 2 min. The collection tube was emptied and 0.75 ml of ethanol-containing PE (wash) buffer was added to the column and centrifuged for 2 min at 13,000 x g. The collection tube was emptied and the column was centrifuged for an additional minute at maximum speed. A coloured membrane (i.e. pink for Cy 3 or blue for Cy 5) at this stage indicated good labelling. The column was placed in a clean 1.5 ml microfuge tube and 30 μl EB (elution) buffer was added to the centre of the column membrane and allowed to

incubate for 1 min at room temperature. The labelled cDNA was eluted by centrifugation at 13,000 x g for 1 min. The elution step was repeated with a fresh 30 μl of EB buffer using the same collection tube. The Cy dye labelled probes were then dried using a speed vac prior to hybridisation.

2.2.16.5 UV-crosslinking and prehybridisation of slides

Prior to hybridisation the spotted cDNAs were first rehydrated in a humidity chamber with 1 X SSC (0.15 M $NaCl_2$; 15 mM $Na_3C_6H_5O_7.2H_2O$), denatured at 95 °C for 4 min then UV cross-linked to the slide surface at 120 mJ. Slides were incubated at 45 °C for 45 min in preheated prehybridisation buffer (5 X SSC, 0.1 % SDS, 1 % BSA). The slides were washed twice in Coplin jars containing MilliQ water by repeated dipping ensuring the entire slide was submerged. Slides were then submerged once in isopropanol, removed and allowed to air-dry. The water/water/isopropanol wash cycle was repeated with any slides that appeared smeared at this stage.

2.2.16.6 Probe hybridisation

Care was taken to expose Cy-labelled probes to as little light as possible. Labelled probes were resuspended in 23 μl hybridisation buffer (50 % formamide, 5 X SSC, and 0.1 % SDS). Probes that originated from the same RNA and that were labelled the same were pooled together at this stage. A single aliquot from the pool of Cy3 labelled cDNA was combined with an independently labelled Cy5 in a total volume of 45 μl for each slide, which was denatured at 95 °C for 3 minutes. A 22 mm x 60 mm microscope glass coverslip was placed on the prehybridised slide and positioned in a hybridisation chamber. The labelled probe mixture (45 μl) was pipetted onto the slide surface at one end of the printed area and allowed to cover the entire print area by capillary attraction. Gentle tapping on the cover slip surface removed any bubbles underneath. 10 μl water was added to the wells at each end of the chamber to ensure the atmosphere within remained humid during hybridisation. The chamber was sealed tightly and placed in a light proof incubator at 42 °C for 15 hr. Following hybridisation, the cover slips were removed and all slides were washed twice in 1 X SSC; 0.2 % SDS for 10 min at 45 °C followed by two washes in 0.1 X SSC; 0.1 % SDS for 10 min at 45 °C and then two washes in 0.1 X SSC for 10 min at 45 °C. Slides were finally rinsed in MilliQ water and dried by brief centrifugation. Slides were scanned as soon as possible.

2.2.16.7 Image acquisition and image processing

All hybridised slides were scanned using the confocal laser scanner ScanArray Express HT at 75 % photomultiplier tube, 75 % of laser power, and 10 μm of pixel resolution. The wavelengths used were 550 nm (Cy 3) and 675 nm (Cy 5). Images were acquired by ScanArray Express 2.0 software (GSI Lumonics) and processed with Quantarray 3.0 software (Packard Bioscience) in order to measure the intensity and the local background for both Cy3 and Cy5 channels of each spot. This data was then further analysed as discussed in Chapter 4.

2.2.17 Protein extraction for the analysis of cathepsin C activity

To maintain cathepsin C activity trypsin was not used on the cells. All growth media was removed and cells were then rinsed in PBS. 1 ml of PBS was added to the culture flask and a cell scraper was used to remove cells into a 1.5 ml tube. The cell suspensions were then frozen and thawed three times to shear the cells. Cell suspensions were kept on ice while they were sonicated at 200 W with 1.5 sec intervals using a Sonics and Materials Inc VibraCell^{TM} sonicator. The lysates were centrifuged at 13,000 x g for 10 min and supernatants were kept at –20 ^{o}C for further use.

2.2.18 Cathepsin C fluorescent substrate assay

This 96-well microtitre plate assay is based on the substrate specificity of proteolytic enzymes, where proteolysis of a specific synthetic substrate is monitored as fluorescence due to the liberation of a fluorescent leaving group. The leaving group used was the fluorescent molecule 7-Amino-4-methylcoumarin (AMC) which was conjugated to a Glycine-Arginine sequence. Gly-Arg is a specific substrate for cathepsin C.

In a white 96-well plate in triplicate 10 μl total cell lysis solution (as prepared in section 2.2.17) was added to 0.2 μl 10 mM Gly-Arg-AMC (in DMF), 4 μl 0.1 M DTT and the final volume brought to 200 μl with 0.1 M sodium phosphate buffer, pH 6.0. Using the same 96-well plate a standard curve ranging from 0 to 10 μM AMC was prepared by adding 195 μl standard to 5 μl 0.1 M DTT. The plate was covered and

incubated at 37 °C for 60 min. 50 µl 10 % (v/v) acetic acid was added to each well to terminate the reaction. The plate was read on a fluorometer with an excitation wavelength of 370 nm (slit width 10) and an emission wavelength of 440 nm (slit width 2.5).

2.2.19 Isolation of total RNA from tissue samples

Tissue samples were removed from - 80 °C storage and mounted on a CM 1900 Cryostat (Leica) set to - 20 °C. 8 micron sections were repeatedly cut into 1 ml TRI^{TM} reagent and a fine gauge needle was used to facilitate disintegration of the tissue. RNA was then isolated as described in section 2.2.2 except 1 µl 5 mg/ml glycogen and isopropanol were added to the aqueous phase to precipitate the RNA.

Chapter 3

The roles of MMPs and CD44 in colon cancer invasion

3.1 Introduction

3.1.1 MMPs in cancer and metastasis

ECM turnover and therefore proteolytic activity are key to tumour progression. Malignant cancers as opposed to benign neoplasms are characterised by their ability to cross tissue boundaries and spread to other organs of the body. Degradation and remodelling of the tumour microenvironment and the ECM, including the basement membrane through proteolysis of matrix proteins are essential steps in local invasion and metastasis. MMP degradation of ECM and cell associated proteins is essential for almost every step of metastasis; for example in aiding cells crossing ECM barriers and the epithelial basement membrane as well as facilitating intravasation and extravasation and colonisation at distant sites. Without the ECM-degrading MMPs, cancer cells would be unable to cross these matrix barriers that normally hinder their spread. MMPs are the main proteolytic enzymes associated with tumour progression and play a pivotal role in virtually all steps of cancer evolution, both early and late. The significant gaps in knowledge regarding the mechanisms underlying such contributions must be filled before the development of therapeutic MMP inhibitors (MMPIs) can proceed with a legitimate chance of success.

As well as physically removing the barriers for tumour progression, MMP activities have wider roles in tumour progression including growth factor and pro-MMP activation (Yu and Stamankovic, 2000 and Sato *et al.*, 1994). MMP overexpression in a tumour therefore, is associated with a more aggressive phenotype. The contribution of the microenviroment to tumour progression is stressed by the fact that in epithelial cancers most of the upregulated MMPs are expressed by the neighbouring supporting stromal cells as apposed to the carcinoma cells (Werb, 1997 and Nelson *et al.*, 2000). Malignant cells can consequently recruit this MMP activity and implement it in a variety of proteolytic tasks, i.e. from ECM turnover to the proteolytic release of latent ECM-bound growth factors. Increased levels of MMPs in a tumour occur due to their uncontrolled overexpression or because of diminished expression of the TIMPs. The TIMP family of secreted proteins inhibit the activity of MMPs through binding them selectively and reversibly in a 1:1 stoichiometric manner. The main MMPs associated with all aspects of tumour invasion (i.e. from intravasation to angiogenesis) are MMP-1,

-2, -3, -7, -9, -12, -13 and –14 (Sternlicht and Bergers, 2000), with other MMPs having more discrete roles in the metastatic process.

MMP activity is implicated at all stages of cancer and metastasis including the early stages of tumour development.   The exact roles of MMPs in early stage tumourigenesis are still unclear.   However, MMP activity at the stages preceding invasion is thought to contribute to tumour growth rather than tumour initiation.   One way for this to happen is through increased neovascularisation.   MMPs play a key role in angiogenesis, an important part of tumour development where new capillaries are formed in response to angiogenic stimuli such as FGF and EGF (Fisher et al., 1994). The formation of tumour vasculature is crucial for tumour development.   As the tumour grows in mass, a critical point is reached where further growth is prevented without angiogenesis occurring.   MMP activity is involved in ECM degradation, thus allowing vascular endothelial cells to advance through the basement membrane in order to lay down new blood vessels (Stetler-Stevenson, 1999).   Secreted MMP activity has been detected as a response of cultured endothelial cells to angiogenic stimuli (Fisher et al., 1994).   Studies using MMP inhibitors have shown that a reduction in MMP-2 activity results in reduced angiogenesis (Moses and Langer, 1991 and Brooks et al., 1998).   The requirement for MMP activity during neovascularisation has implications for the growth of both the initial tumour and distant metastases.

3.1.1.1 MMP-7 in colon cancer and metastasis

MMP-7 (also known as matrilysin) expression is found in 90 % of colonic adenocarcinomas (Newel et al., 1994), where clinically, a correlation exists between expression and distant metastasis (Adachi et al., 1999).   MMP-7 is sub-classed as a minimal domain MMP.   At 28 kDa it is the smallest member of the MMP family and its active form has a molecular weight of 19 kDa.   MMP-7 has the widest range of substrates, with the ability to degrade collagens, elastin, fibronectin, gelatin, laminin, tenascin, E-cadherin, proteoglycans and glycoproteins (Fingleton et al., 1999).   The most distinctive aspect of MMP-7 expression patterns is the fact that it is expressed in the epithelial cells of tumours of gastrointestinal, prostate or breast origin, where as the majority of other MMPs are produced in the stroma.   It is this aspect of MMP-7 that gives it its potentially novel functions (Wilson and Matrisian, 1996).   MMP-7 is expressed in a variety of tumours including those of the breast, colon, prostate and lung.

MMP activity is implicated at all stages of cancer even at the early stages of tumourigenesis (Fingleton *et al.*, 1999). MMP-7 is thought to play key roles in both the growth of early colonic adenomas as well as their transformation into invasive cancers (Wilson *et al.*, 1997). A significant amount of evidence has accumulated regarding the role of MMP-7 in the early stages of colorectal tumourigenesis. MMP-7 activity has been shown to be present in colorectal adenomas (Yamamoto *et al.*, 1994). MMP-7 has also been found localised to the cytoplasm of colon adenoma cells, giving it possible intracellular functions in adenomas (Fingleton *et al.*, 1999). In a model of colorectal tumour cell invasion, MMP-7 has been shown to be more important in the early stages of tumour formation, progression and growth as well as the later processes of metastasis (Witty *et al.*, 1994). This study of the MMP expression profiles in two genetically matched human colon adenocarcinoma derived cell lines showed that the lymph node metastasis derived SW620 cell line expressed MMP-7 where the SW480 primary tumour cell line, derived from the same patient did not. In the Witty *et al.* (1994) study, over expression of MMP-7 in the SW480 cell line by cDNA transfection caused increased tumourigenicity and metastases when injected into nude mice, while MMP-7 inhibition has also been shown to decrease metastasis (Hasegawa *et al.*, 1998). Transgenic mice studies have also revealed the effect of MMP-7 overexpression on initial tumour establishment and growth (Rudolph-Owen *et al.*, 1998). In this study, overexpression of MMP-7 led to enhanced tumourigenesis in a mouse breast cancer model. Furthermore, MMP-7 knockout mice that are null for MMP-7 have been shown to develop fewer tumours (Wilson *et al.*, 1997) thus supporting the role for MMP activity in tumour establishment and growth. MMP-7 can therefore be considered to function as an oncogene since colorectal tumourigenesis is suppressed in mice lacking MMP-7 and tumourigenic potential is enhanced in cell lines overexpressing MMP-7 (Ougolkov *et al.*, 2002). One important downstream cellular function targeted by active MMP-7 is the suppression of Fas-mediated apoptosis and the subsequent promotion of cell survival. This is due to the MMP-7 mediated release of soluble Fas ligands from the cell surface (Mitsiades *et al.*, 2001 and Powell *et al.*, 1999). The expression of MMP-7 in primary tumours can therefore be monitored as a marker for progression.

3.1.1.2 MMP-2 and MMP-9 in colon cancer and metastasis

Another group of MMPs that have extensive associations with colon cancer are the gelatinases. Gelatinase A (MMP-2) and gelatinase B (MMP-9) are the only members of the fibronectin domain sub-class of MMPs. MMP-2 and MMP-9 degrade denatured collagens (gelatins) and are specific for the degradation of type IV basement membrane collagen giving them essential roles in the processes of intravasation and extravasation. MMP-2 is secreted as a 72 kDa pro-enzyme and is activated to a 66 kDa form that degrades gelatin, collagens type I, III, IV, V, VII, X, VI and fibronectin. MMP-2 expression is widespread and is frequently elevated in malignant tumours. The expression of MMP-9 is also widespread with activity being described in malignant cells, neutrophils, and keratinocytes (Collier *et al.*, 1988 and Wilhelm *et al.*, 1989). MMP-9 is secreted as a 92 kDa zymogen and its active form has a molecular weight of 86 kDa. MMP-9 cleaves gelatins, collagens type IV, V, XI, XIV, elastin and laminin. Increased expression of MMP-9 has been shown in colon adenomas when compared with normal mucosa, possibly representing an early change in progression from adenoma to carcinoma (Parsons *et al.*, 1998). Another important observation is that this up-regulation was not seen in rectal tumours (Roeb *et al.*, 2001). The major roles of MMP-9 in colon cancer are in ECM degradation and the subsequent migration and invasion of tumours. Inhibition of MMP-9 leads to impaired invasion and proliferation of tumour cells in a three-dimensional collagen matrix (Agrez *et al.*, 1999).

3.1.1.3 MMPs and the ECM

The ECM plays a key role in regulating cell signalling and maintaining normal cell homeostasis. Degradation of the ECM at the site of adhesion can disturb these interactions thus modulating the signals to the cells. It has been shown that MMP-2 degradation of a collagen matrix reveals hidden integrin adhesion sites (Montgomery *et al.*, 1994). Integrins expressed on the cell surface of melanoma cells are then free to act as receptors for degraded collagen and signal for cell survival. MMP cleavage of matrix components releases soluble ECM fragments which have been shown to increase cell migration (Woodhouse *et al.*, 1997 and Giannelli *et al.*, 1997).

There is increasing evidence for a relationship between stromal cells and colorectal carcinoma cells in promoting MMP production in either tumour or stromal

cells (Ornstein *et al.*, 1999). It is well established that MMP-9 is produced in the stromal cells of colon cancers (Zeng and Guillem, 1996 and Nielsen *et al.*, 1996). The observation that MMPs are largely produced by the stromal compartment surrounding the tumour tissue suggests that the tumour, by communicating with its surrounding environment, can stimulate MMP expression. This is often due to the MMP-mediated release of growth factors by either direct cleavage of bound cell surface growth factors or by their release from the ECM as a result of ECM turnover. As well as speeding up tumour proliferation, the bioavailability of such mitogens induces MMP genes, for example, MMP expression is often increased in response to members of the EGF and FGF families (McCawley and Matrisian, 2001). Tumours are known to produce cytokines and growth factors, including those with MMP-inducing ability (Birkedal-Hansen *et al.*, 1993) and it is quite possible that these factors account for MMP production in the stromal compartment (Klein *et al.*, 1997). Human colorectal cancer cell lines that did not express MMPs were shown to induce host MMP expression when transplanted into nude mice (McDonnell *et al.*, 1999). MMP-7 has been shown to dock on the cell surface with CD44 where it can activate heparin bound epidermal growth factor (HBEGF), which also binds CD44, making it available to its cell surface receptor (Yu *et al.*, 2002). This highly orchestrated mechanism illustrates how MMPs can finely control the tumour environment by utilising the host surroundings.

3.1.2 CD44 in cancer

CD44 expression has been identified on a variety of different tumour cells, including colon and breast (Gunthert *et al.*, 1991, Choi *et al.*, 2000, Herrlich *et al.*, 1995 and Bourguignon *et al.*, 2002). Tumour progression is associated with abnormal CD44 variant expression and by a generalised dysregulation of the splicing process leading to the expression of a multitude of CD44 isoforms in tumours of increasing malignancies (Jothy, 2003). Of particular interest are the CD44 variants CD44v6 and CD44v4-v7, which confer metastatic properties when overexpressed *in vivo* (Gunthert *et al.*, 1991). Transfection of defined CD44 variants (v4 through v7, v6 and v7) converted non-metastatic cell lines to cells that spread and colonised the lung *in vivo* where transfection with standard truncated CD44 (CD44s) was negative for metastasis (Herrlich *et al.*, 1993). Colorectal tumour progression is associated with a decrease in CD44s expression and an increase in CD44v expression (Wielenga *et al.*, 1993). In a study of the different stages of colorectal carcinogenesis, CD44v6 expression has been

shown to be negative in normal colon epithelium and positive in 100 % of carcinomas and metastases (Mulder *et al.*, 1994 and Herrlich *et al.*, 1995). The functional roles for CD44v6 in the developed tumour are still unclear but one possibility is that it confers a growth advantage. Experimental animal studies have shown that an increase in CD44s expression prevents the establishment of metastasis (Pereira *et al.*, 2001). Overexpression of CD44v6 has been observed more often in advanced colon cancers (Dukes' stage C/D) compared with Dukes' A/B, and it can therefore be used as a prognostic marker (Wielenga *et al.*, 1993 and 1998). Interestingly, CD44 upregulation has been shown to be an early event, suggesting a role in tumourigenesis (Kim *et al.*, 1994).

The exact roles of CD44 in cancer have yet to be elucidated. The acquisition of a metastasising phenotype involves a large number of functions, but it is still unclear which function involves CD44. The remarkable and unusual property of CD44 in comparison with other adhesion molecules is that it mediates both cell-cell and cell-ECM interactions. Such varied adhesion interactions are required for the migration of tumour cells into the foreign stromal surroundings. Therefore, alterations in CD44 expression can hypothetically aid the metastatic spread.

3.1.3 Integrins: roles in the transformed cell

Integrins are essential for tumour progression due to their capability to mediate specific physical interactions with the ECM and their ability to control signalling pathways especially those involved in cell movement. Cancer cells can switch the types of integrins they express, favouring ones that transmit pro-growth signals. On their travels, invading and metastasising cells experience changing tissue microenvironments and are presented with different matrix components. Successful metastasis therefore requires adaptation. Therefore, changes in integrin subunit expression in transformed cells are important in tumour progression (Juliano and Varner, 1993). As discussed in chapter 1, integrin signals influence cellular proliferation and differentiation as well as cell survival. They can, therefore directly or indirectly (via promoting growth arrest, differentiation and/or apoptosis) influence the growth of a tumour *in vivo*. For example, overexpression of the αvβ6 integrin, a fibronectin receptor, has been shown to enhance the proliferative capacity of SW480 colon carcinoma cells *in vitro* and in nude mice (Agrez *et al.*, 1994).

3.1.4 MMP interactions with cell adhesion molecules

Both altered matrix degradation and cell adhesion are vital factors in promoting tumour invasion. MMPs are thought to participate in both these processes via interactions with cell adhesion molecules.

MMP proteolysis of the ECM effects cell adhesion via the integrin family of cell adhesion molecules by changing the way integrins interact with the ECM. Several studies have shown a functional link between integrins and the MMPs. It has been shown that ligation of collagen by α2β1 integrin on the surface of ovarian carcinoma cells results in tyrosine kinase signalling and the subsequent expression of MT1-MMP which activates pro-MMP-2 resulting in enhanced cell surface MMP activity (Ellerbroek *et al.*, 1999). Increased αvβ1 integrin expression leads to enhanced MMP-9 secretion (Agrez *et al.*, 1999). Individual integrins can bind several ECM proteins, with different interactions having different effects. For example, α4β1-mediated adhesion to fibronectin induced the expression of MMP-2 and MMP-9 but upon α4β1-mediated adhesion to vascular cell adhesion molecule-1 (VCAM-1), only MMP-2 was induced in the same cells (Yakubenko *et al.*, 2000). αvβ3 integrin will not bind native matrix collagen but it will bind proteolytically degraded collagen. αvβ3 expression and signalling confers an invasive advantage on cells (Cheresh, 1991). MT1-MMP can process the αv subunit of the αvβ3 integrin leading to a C-terminal truncated dimer that is more efficient at promoting cell migration than the full-length αvβ3 (Deryugina *et al.*, 2002 and Ratnikov *et al.*, 2002). Moreover, αvβ3 integrin can bind active MMP-2 at its C-terminus and target its activity on collagen (Brooks *et al.*, 1996). This is a fine example of symbiosis between a cell adhesion molecule involved in cellular motility and an enzyme capable of basement membrane degradation. Tumour cell motility is stimulated by TGF-β1 through direct transcriptional activation of α3 integrin (Giannelli *et al.*, 2002), which leads to acquired motility of tumour cells on a laminin-5-containing matrix. The presence of MMP-9 in this system leads to the cleavage and activation of TGF-β1, giving it a key role in the signalling machinery responsible for cell motility and invasion (Yu and Stamenkovic, 2000). Certain integrins affect the transcription of various MMP genes. α2β1 overexpression has been associated with an increase in MMP-1 (interstitial collagenase) expression (Jones and Walker, 1997) and αvβ6 integrin has been shown to induce MMP-9 in colon cancer cells (Niu *et al.*, 1998). The

β4 subunit of the α6β4 integrin is a specific receptor for laminin, and it is important in preventing tumour spread. Monoclonal antibody blocking of β4 integrin subunit activity resulted in enhanced MMP-2 expression and increased migration (Daemi *et al.*, 2000), providing evidence for the functional relationship between individual MMPs, cell adhesion, and cell invasion.

Both MMPs and CD44 are critical to the invasive process, and several studies have recently begun to define how these molecules interact with each other. Stimulation of CD44 has been shown to increase MMP-2 expression in human melanoma cells (Takahasi *et al.*, 1999) and in lung carcinoma cells (Zhang *et al.*, 2002). MMP-9 has also been shown to associate with CD44 on the surface of mouse mammary carcinoma and human melanoma cells thus enhancing invasion (Yu and Stamenkovic, 1999). Invading cells therefore have a mechanism by which they can direct MMP-9 activity to the area of invasion. MMP-9 is also free to activate other MMPs or growth factors in this tethered orientation (Yu and Stamenkovic, 2000). A recent paper has shown that CD44v3 can anchor MMP-7 on the cell surface and target its activity (Yu *et al.*, 2002). CD44 can form a complex with MT1-MMP and both molecules co-localise on the cell surface at the cell migration edge. In this arrangement CD44 functions as a molecular linker between MT1-MMP and the actin cytoskeleton, linking the factors involved in cell motility such as matrix degradation, cell adhesion, morphology with the cytoskeleton (Mori *et al.*, 2002). It has also been shown that MT1-MMP can directly cleave CD44 and promote cell migration in a pancreatic tumour cell line (Kajita *et al.*, 2001 and Cichy and Pure, 2003).

3.1.5 Therapeutic inhibition of MMPs

The ever growing information regarding the roles of MMPs in tumour development and progression have spurred the development of many synthetic matrix metalloproteinase inhibitors (MMPIs). Inhibition of MMP activity in tumours is desirable because it would ideally result in the inhibition of tumour growth, angiogenesis, invasion and metastasis. The key to the design of effective MMPIs is in understanding the roles and properties of individual MMPs. In this way, indiscriminate effects of such therapies on normal healthy tissue and subsequent side effects can be minimised. Originally, MMPI designs were based on small peptides that mimicked MMP substrates (Brinkerhoff and Matrisian, 2002). Based on the structure of collagen,

the broad spectrum MMPI Marimastat (British Biotech, Oxford, UK) has been shown to significantly increase survival when administered to patients with early-stage, metastasis free disease (Whittaker *et al.*, 1999 and Curran and Murray, 2000). Contributions of MMPs, especially MMP-7, to early stage tumourigenesis, suggest it may be beneficial to treat tumours at the early stages using MMPIs. Marimastat was not without side effects, with patients experiencing extreme joint pains. Trials on BAY 12-9566, an MMPI developed by Bayer Pharmaceuticals were halted because the placebo performed better than it (Coussens *et al.*, 2002). The fate of MMPI as a therapy lies in the hands of basic researchers. Further understanding of the exact roles of specific MMPs at different stages of tumour progression is needed. This would lead to more targeted therapies and would minimise side effects.

3.1.6 Summary

CD44 and MMPs are among the many participants in tumour cell invasion and metastasis. MMPs have well-established roles in tumour cell invasion where MMP activity is involved in ECM turnover, neovascularisation and growth factor activation. Increased MMP-7 and MMP-9 activity have been associated with colorectal tumour progression. CAMs such as CD44 also have specific roles in the metastatic process. As well as mediating cell – cell and cell – matrix interactions, CAMs are also involved in cell signalling. Recent evidence has suggested CD44 may be involved in colorectal tumour cell invasion. Of particular interest are the CD44 splice variants, with splice variant expression having prognostic potential in colorectal cancer. Collaborative interactions between these biomolecules in the metastatic process have been unearthed by recent research into their functional relationships. While both these biomolecules are well-established participants in the metastatic process, their collective interactive functions are still unclear. Although the interactions between CAMs and MMPs in regulating tumour metastasis have yet to be completely elucidated, it is no longer suitable to consider these groups of biomolecules as acting in isolation. In this chapter, interactions between CD44 and MMPs in regulating invasion and migration in colorectal tumour cell lines were examined. The elucidation of such collaborative roles would validate their potential as either potential therapeutic targets or prognostic markers of colorectal metastasis.

3.2 Results

3.2.1 Description of colon cancer cell lines used in this study

MMPs were overexpressed in colorectal cancer cells to investigate their roles in various aspects of tumour cell invasion. The main cell lines used in this study were the paired colon cancer cell lines SW480 and SW620. SW480 is a cell line established from a primary Dukes' stage B colon adenocarcinoma of a 50-year old male patient. 6 months later a recurrent cancer was discovered in the same patient, with liver and lymph metastases present. The SW620 cell line was derived from a lymph node metastasis (Leibovitz *et al.*, 1976). These isogenic cell lines have been validated as a model of colorectal metastasis and therefore comparisons between them could provide new insights into the metastatic process (Hewitt *et al.*, 2000). The non-metastatic SW480 cell line was stably transfected with the cDNAs for MMP-9 and MMP-7 (Murray *et al.*, 2004 and Witty *et al.*, 1994). The new cell lines have been termed SW480M9 and SW480M7 respectively.

3.2.2 Expression of MMPs in colon cancer cell lines

3.2.2.1 Expression of MMP-7 mRNA in colon cancer cell lines

The SW480M7 cell line was received from Professor Matrisian's laboratory (Vanderbilt University, Nashville, TN). This cell line has been previously studied (Witty *et al.*, 1994) and was referred to as SW480 (Mut6) in this paper, for convenience the name SW480M7 has been adopted for this thesis. It was important to re-characterise these cells as they had been frozen down since 1995. The expression of MMP-7 in these cells was monitored by western blot and real time PCR analysis.

Firstly, total cellular RNA was isolated as described in section 2.2.2 and its integrity was assessed prior to RT-PCR as described in section 2.2.2.1. RNA is easily degraded by RNases. Unlike most nucleases, RNases are resistant to autoclaving but can be inactivated by DEPC when it is added to solutions at a final concentration of 0.1 %. Therefore all solutions for RNA work were DEPC treated and gloves and disposable plastics were used at all times. To help prevent RNA degradation by endogenous RNases, all RNA procedures were carried out quickly and on ice. RNA was examined

for degradation by running an aliquot of each isolated RNA sample on a 1 % agarose gel (Figure 3.1). The presence of 28S, 18S and 6S RNA ribosomal subunits were typical of undegraded RNA. Isolated RNA was quantified by measuring its absorbance at 260 nm (40 µg RNA had an OD of 1). RNA purity was assessed using the absorbance value for each sample at 280 and 260 nm to calculate its A260/A280 ratio. RNA samples had ratio values between 1.6 and 1.8, which indicated they were free of DNA, protein and organic solvent contaminants.
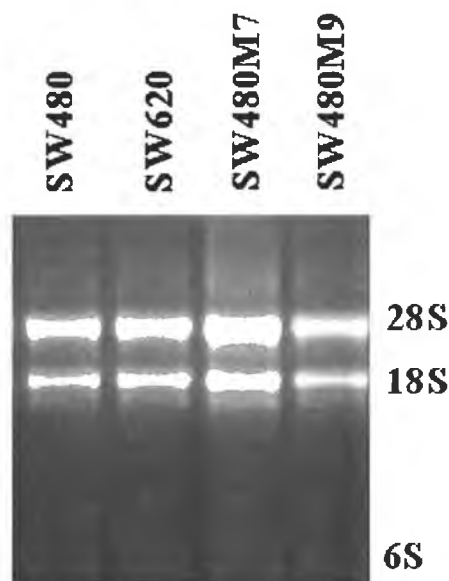
Figure 3.1 An aliquot of total RNA was analysed for degradation by agarose gel electrophoresis. The presence of 28S, 18S and 6S ribosomal sub-units was typical of undegraded RNA.

To show MMP-7 overexpression in the SW480M7 at the level of transcription, real time PCR as outlined in section 2.2.4 was carried out on RT reactions. Real time PCR is a sensitive and reproducible technique that gives a quantitative expression value. Real-time PCR quantitates the initial amount of template more specifically, sensitively and with greater reproducibility than normal quantitative PCR, which detects the amount of final amplified product. This is because real-time PCR monitors the fluorescence emitted during the reaction as an indicator of amplified product formed during each PCR cycle (i.e. in real time) as opposed to the endpoint detection by conventional quantitative PCR methods, thus allowing the real-time progress of the reaction, especially its exponential phase, to be viewed. Although real-time PCR removes the need for electrophoresis of the PCR products, it does not detect the size of the product. The system is based on the detection and quantitation of a fluorescent

reporter, where the signal increases in direct proportion to the amount of PCR product in a reaction. SYBR green, a double stranded DNA binding dye that binds minor grooves and will not bind ssDNA was used. The real-time system was used for comparative gene expression analysis, normalising with beta-actin levels.
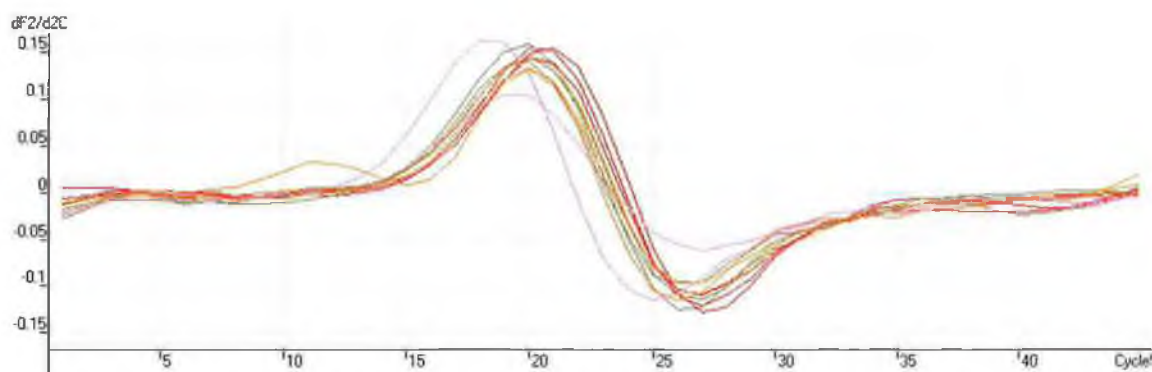
Figure 3.2 The real time PCR comparative concentration curve for β-actin expression in SW480, SW620, SW480M7 and SW480M9 cells. The data from this curve (Table 3.1) was used to normalise MMP-7 and MMP-9 expression data. The colours used are summarised in Table 3.1.

Table 3.1 β-actin real time PCR comparative concentration expression data. 'Take-off' refers to the cycle in which the reaction entered the exponential phase. The comparative concentration data was used to normalise MMP-7 and MMP-9 real time PCR expression data.

Colour	Name	Takeoff	Comparative Concentration
	SW480 Actin	16.6	1.00E+00
	SW480 Actin	16.5	1.06E+00
	SW480 Actin	16.7	9.46E-01
	SW620 Actin	13.0	7.25E+00
	SW620 Actin	15.2	2.16E+00
	SW620 Actin	14.7	2.84E+00
	SW480M7 Actin	15.8	1.55E+00
	SW480M7 Actin	15.9	1.47E+00
	SW480M7 Actin	15.8	1.55E+00
	SW480M9 Actin	16.1	1.32E+00
	SW480M9 Actin	15.7	1.64E+00
	SW480M9 Actin	17.0	8.02E-01
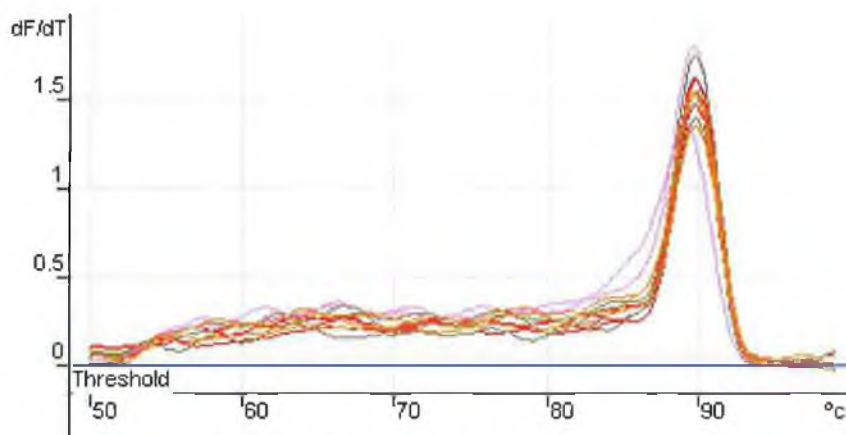
Figure 3.3: Melt curve for β-actin real time PCR reactions.

The constitutively expressed cytoskeletal protein β-actin was used as a control. The β-actin real time PCR data obtained from each cell line is summarised in Table 3.1 and was used to normalise MMP-7 and MMP-9 expression data. Following real time PCR, melt curve analysis was carried out on all reactions by heating from 50 °C to 100 °C while monitoring fluorescence. As can be seen from Figure 3.3 all β-actin reactions resulted in only one melt curve peak indicating that only one PCR product was formed in the reaction and that the primers and conditions used were specific for β-actin amplification.
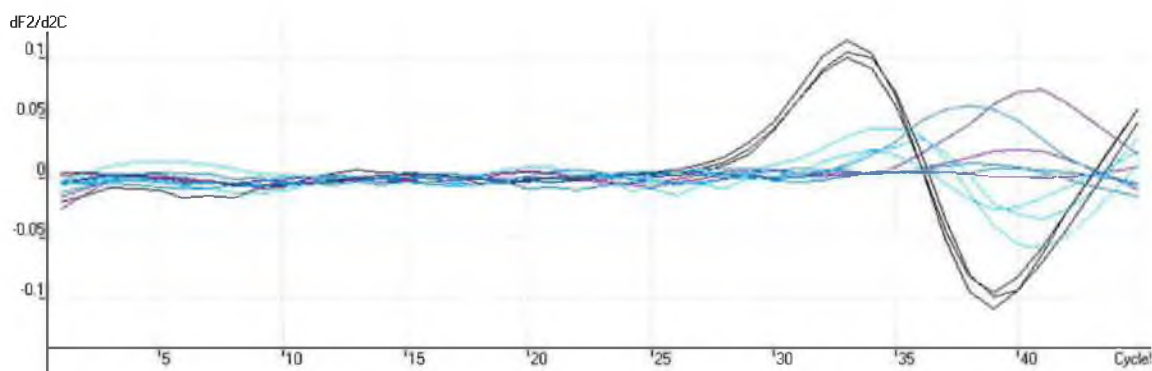
Figure 3.4 Real time PCR comparative concentration curve for MMP-7. The data obtained from this reaction (Table 3.2) was normalised using β-actin data. The colours used are also summarised in Table 3.2.

Real time PCR was carried on all samples for MMP-7 (Figure 3.4). The data obtained is listed in Table 3.2. As can be seen from Figure 3.5 only one peak was obtained for the melt curve analysis of all reactions, indicating specificity.

60

Table 3.2 Comparative concentration data obtained from the MMP-7 real time PCR reaction.

Colour	Name	Takeoff	Comparative Concentration
■	SW480 MMP7	35.8	1.00E+00
■	SW480 MMP7	24.1	9.34E-01
■	SW480 MMP7	36.2	8.56E-01
■	SW620 MMP7	30.0	9.48E+00
■	SW620 MMP7	31.1	6.19E+00
■	SW620 MMP7	32.4	3.74E+00
■	SW480M7 MMP7	29.5	1.15E+01
■	SW480M7 MMP7	29.2	1.29E+01
■	SW480M7 MMP7	29.4	1.20E+01
■	SW480M9 MMP7	33.8	2.17E+00
■	SW480M9 MMP7	28.8	1.51E+00
■	SW480M9 MMP7	33.4	2.54E+00
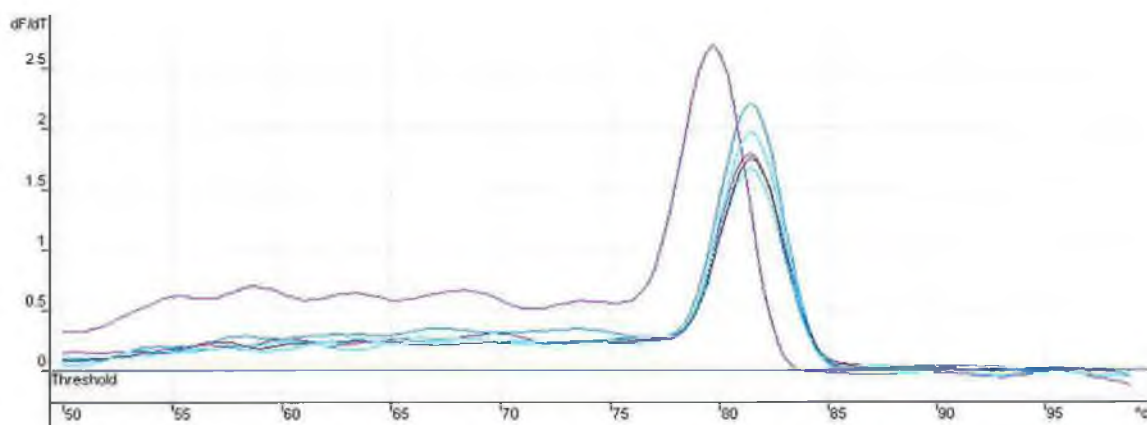
Figure 3.5 Melt curve analysis of MMP-7 real time PCR products. One peak for each reaction indicated the reaction was specific.

In order to make quantitative comparisons between the cell lines used the comparative concentration values obtained from the MMP-7 real time PCR reactions (Table 3.2) were normalised using the β-actin real time PCR data. It can be seen from Figure 3.6 that SW480M7 expressed the most MMP-7 at the mRNA level with the SW480 cells producing barely detectable levels. Interestingly, the SW620 cells and the SW480M9 cells produced similarly low amounts.
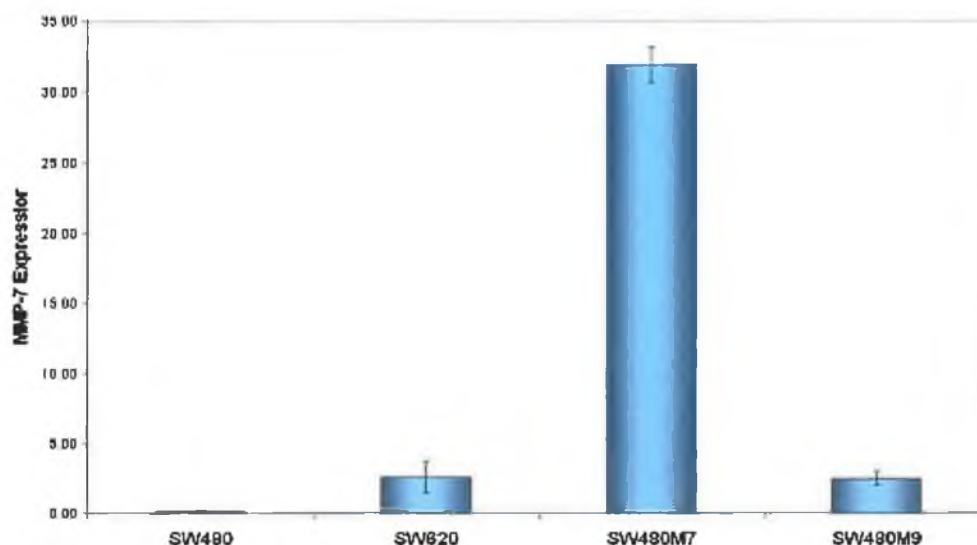
Figure 3.6 Overexpression of MMP-7 in SW480M7 cells. Real-time PCR was used to compare the levels of MMP-7 expression in cell lines. β-actin was used for normalisation.

3.2.2.2 Expression of MMP-7 protein in colon cancer cell lines

Conditioned media was collected from cells as described in section 2.2.5 and protein concentration was determined using the BCA assay as described in section 2.2.6. Using a specific anti-MMP-7 antibody, MMP-7 overexpression in SW480M7 cells was shown by western blot analysis of 50 μg total protein (Figure 3.7). Although no positive control was used, MMP-7 identity was confirmed by its 28 kDa molecular weight. MMP-7 expression was not detected in either the parental SW480 cells, the SW620 cells or the SW480M9 cells.

Figure 3.7 MMP-7 overexpression in SW480M7 cells was determined using western blot analysis.

3.2.2.3 Expression of MMP-9 mRNA in colon cancer cell lines

The SW480M9 cell line was developed by Dr McDonnell whilst on sabbatical in Memorial Sloan Kettering Cancer Center, New York, NY, USA. However, the cells had not been extensively characterised. MMP-9 expression in all four cell lines was assayed by real time PCR, gelatin zymography and western blot. Total cellular RNA was isolated and analysed. Figure 3.1 shows SW480M9 RNA to be intact. Real time PCR was carried out as outlined in section 2.2.4.
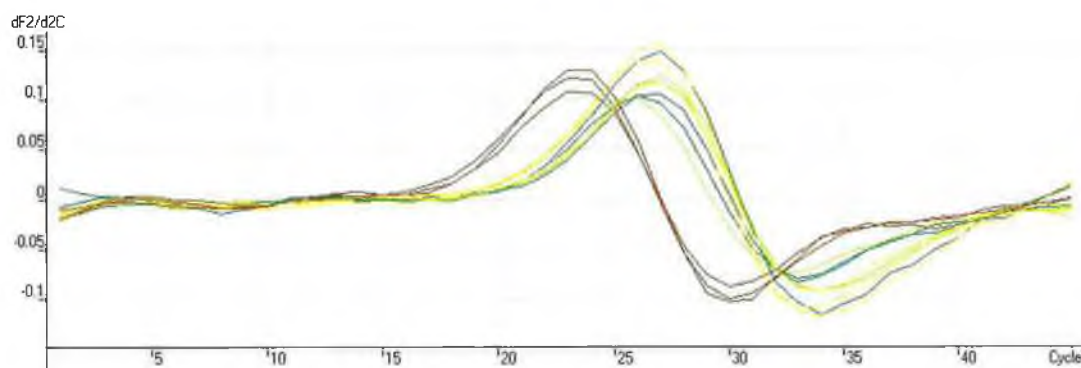
Figure 3.8 MMP-9 real time PCR data. The colours used are summarised in Table 3.3. The comparative concentration values obtained (Table 3.3) were normalised using β-actin real time PCR data (Table 3.1).

Table 3.3: MMP-9 real time PCR comparative concentration data.

Colour	Name	Takeoff	Comparative Concentration
	SW480 MMP9	22.5	1.00E+00
	SW480 MMP9	22.5	1.00E+00
	SW480 MMP9	21.3	1.84E+00
	SW620 MMP9	22.3	1.11E+00
	SW620 MMP9	22.4	1.05E+00
	SW620 MMP9	22.0	1.29E+00
	SW480M7 MMP9	22.1	1.23E+00
	SW480M7 MMP9	22.1	1.23E+00
	SW480M7 MMP9	21.5	1.67E+00
	SW480M9 MMP9	18.9	6.27E+00
	SW480M9 MMP9	19.2	5.38E+00
	SW480M9 MMP9	19.2	5.38E+00

As can be seen from Table 3.3 the SW480M9 cells had the earliest 'take-off' time and the highest comparative concentration values. A melt curve was carried out after the PCR reaction to conclude the reaction was specific (Figure 3.9). One major

peak at 94 °C indicated one product was formed. When normalised, the data clearly showed SW480M9 cells expressed the highest amounts of MMP-9 mRNA. MMP-9 expression in SW480M9 cells was five times higher than that of the parental SW480 cells at the mRNA level. There were no significant differences in MMP-9 expression between the SW480, SW620 and SW480M7 cells (Figure 3.10).
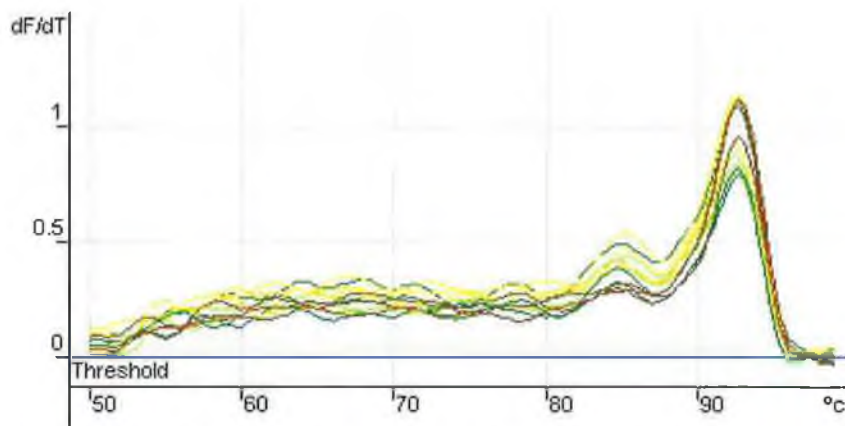
Figure 3.9 Melt curve analysis of MMP-9 real time PCR reactions.
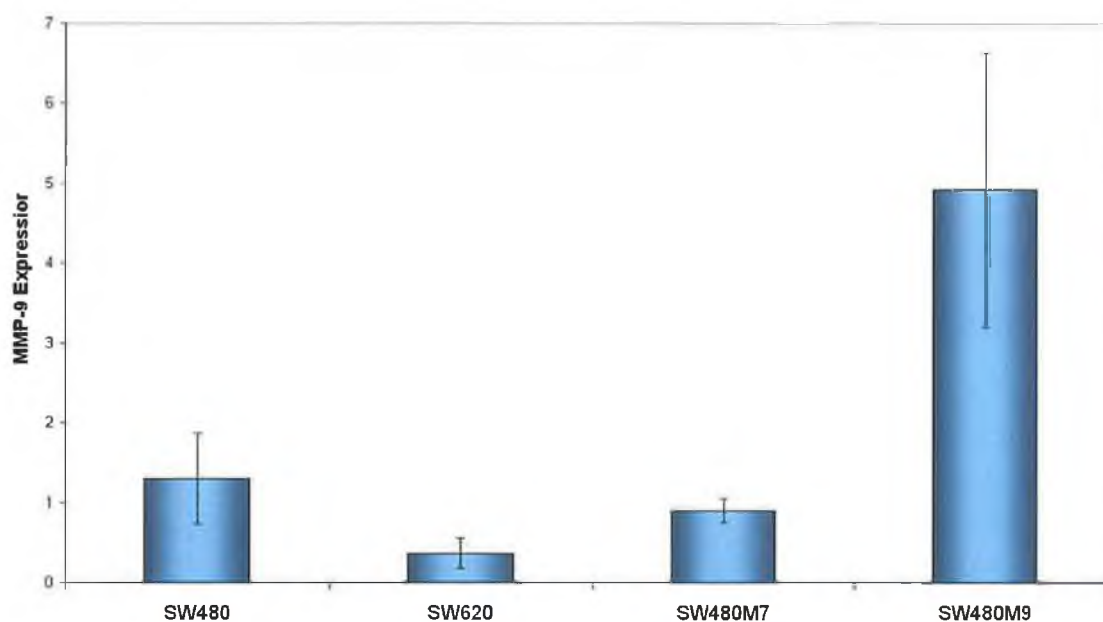
Figure 3.10 Overexpression of MMP-9 in SW480M9 cells. Real time PCR was used to determine levels of MMP-9 expression in the cell lines studied. MMP-9 expression data was normalised using β-actin real time PCR data.

3.2.2.4 Expression of MMP-9 protein in colon cancer cell lines

Gelatinase activity in colon cell lines was studied by analysing conditioned media (CM) using gelatin zymography, which is similar to sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) except the substrate, gelatin is incorporated into the polyacrylamide gel. Zymography is an extremely sensitive method and has the ability to detect picogram quantities of gelatinase activity (Kleiner and Stetler-Stevenson, 1994). Incorporation of gelatin into the gel allows detection of gelatinase activity specifically. Proteolytic enzymes such as the MMPs are easily characterised by digestion of the substrate and by their molecular weight. Maintenance of enzymatic activities is crucial and therefore care must be taken, e.g. precise buffer pH values are important and over-heating of gels must be avoided.

50 µg of total protein was assessed for gelatinase activity as described in section 2.2.9. It can be seen from Figure 3.11 that the SW480M9 cells secreted the highest levels of MMP-9 in comparison with the parental SW480 cells, which did not produce any measurable amounts of the enzyme. Low levels of MMP-9 activity were also detected in the SW620 cells. MMP-9 was identified by its molecular weight and its identity was also confirmed by western blot analysis. Using an MMP-9 specific antibody, MMP-9 was only detected in SW480M9 cells by western blot analysis (Figure3.12).

Figure 3.11 Gelatin zymography was used to show MMP-9 activity in the SW480M9 cells.

Figure 3.12 Western blot analysis was used to confirm MMP-9 overexpression in SW480M9 cells. MMP-9 was undetected in the other cell lines used.

3.2.3 The *in vitro* invasive activities of the cell lines studied

Invasion is a key process in tumour metastasis and therefore the differences in invasion between the metastatic SW620 cells and non-metastatic SW480 cells were monitored. Since ECM degradation is key to tumour cell invasion, the effect of MMP over-expression on *in vitro* invasion through matrigel coated membranes was also examined. The *in vitro* invasion assay mimics the behaviour of invasive cells as they migrate through the basement membrane. Membranes at the bottom of the 24 well cell culture inserts used contained 8 μm pores that were coated with a layer of matrigel artificial basement membrane complex. This layer occludes the pores of the membrane blocking non-invasive cells from migrating through the membrane. In contrast, invasive cells were able to migrate through the matrigel-coated membrane.
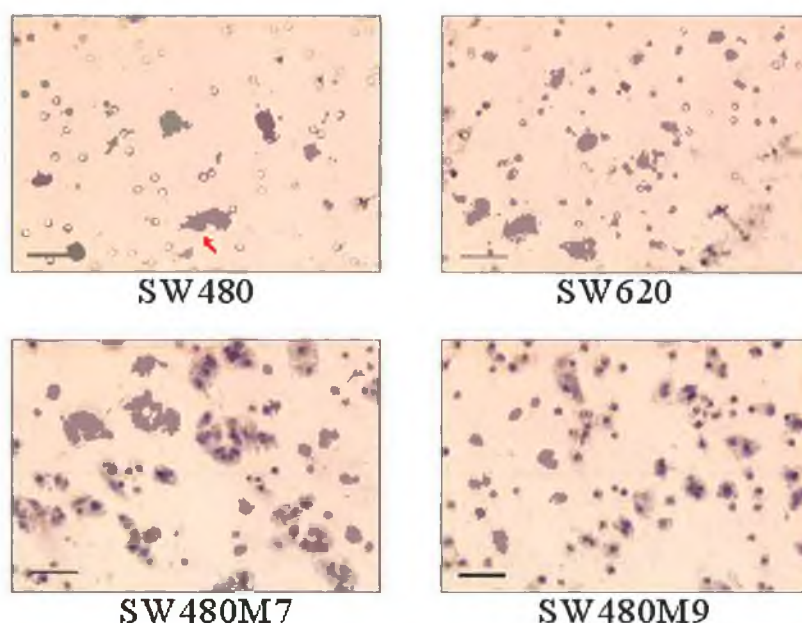
Figure 3.13 10X microscopy of invasive cells using the *in vitro* assay. This figure gives an example of what was observed when invasive cells (red arrow) that had invaded through the 8 μm pores (black arrow) were counted at 10X magnification. Line = 50 μm.
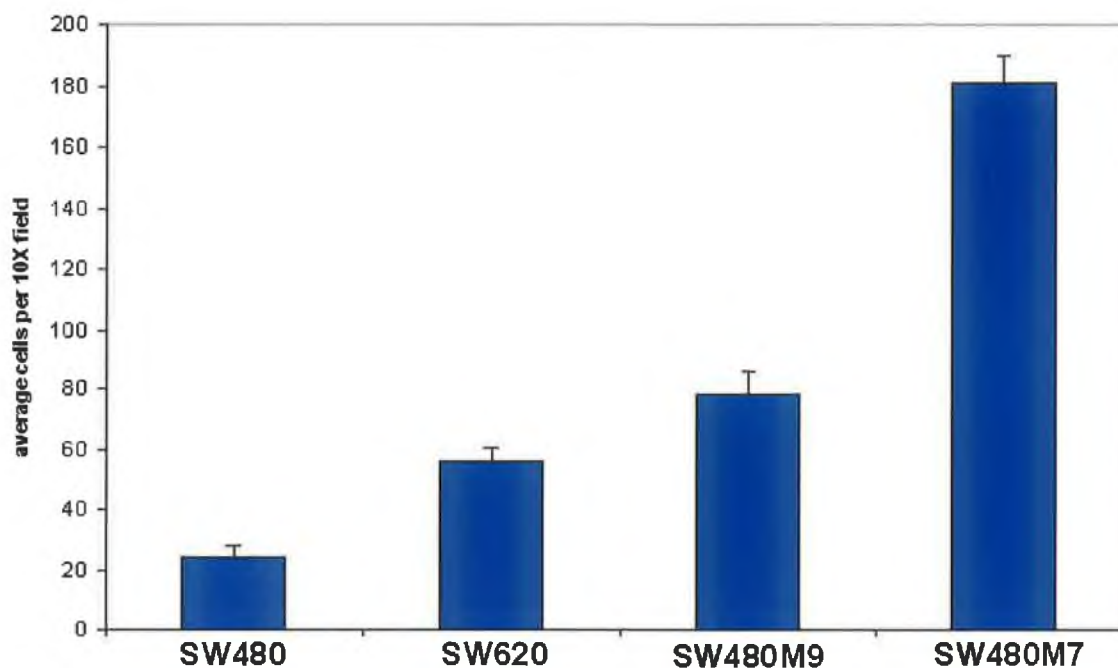
Figure 3.14 Comparison of *in vitro* invasion of cell lines. 5 x 10^5 cells were allowed to invade through matrigel for 24 hr and the invading cells were counted in five 10 X fields.

The assay was carried out as outlined in section 2.2.10. Invasive cells were counted at 10 X magnification (Figure 3.13) and invasive cells were expressed as an average cell count per 10 X magnification field (Figure 3.14). The SW480 cells showed little invasion in comparison to the SW620, SW480M7 and SW480M9 cells, which were significantly ($p < 0.005$) more invasive. The SW620 and SW480M9 were respectively twice and three times more invasive than the parental SW480. SW480M7 cells were 9 times more invasive than the parental SW480 cells.

3.2.4 The *in vitro* migratory capacities of the cell lines studied

For the determination of the *in vitro* migratory abilities of the cell lines studied, the same cell culture inserts as those used in the invasion assays were used without the matrigel coating. In this case, the ability to degrade the basement was not required and motile cells could pass freely through the membrane to the lower chemoattractant-containing chamber.
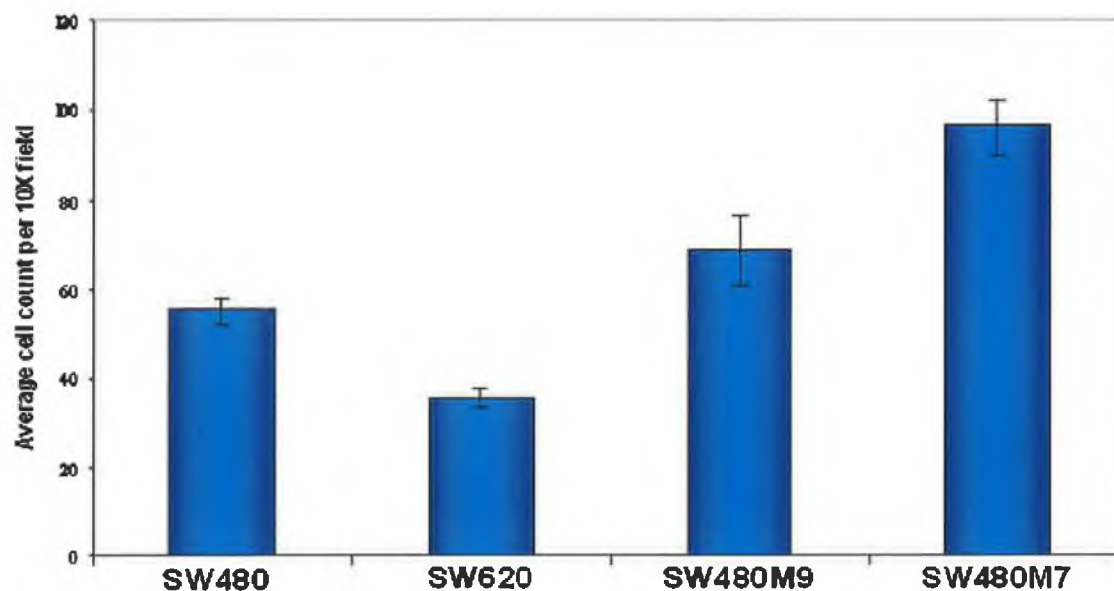
Figure 3.15 *In vitro* migratory capacities of cell lines. Cell migration was assessed by allowing 5×10^5 cells to migrate for 24 hr.

MMP overexpression increased the migratory capacity of the transfected cell lines in comparison to the parental SW480 cells (Figure 3.15). The migration of SW480 cells increased with MMP-9 overexpression and more so with MMP-7 overexpression. However, the parental SW480 cells were significantly ($p<0.005$) more migratory than the SW620.

3.2.5 Comparison of cell adhesion to ECM components

To investigate the adhesion of cells to ECM components, adhesion assays were carried out in the presence of HA and FN as described in section 2.2.10. Adhesion is a key event in the metastatic process where by cells must first adhere to the ECM prior to its degradation. Increased cell-cell signalling and contact is also mediated by increased expression of cell adhesion molecules.
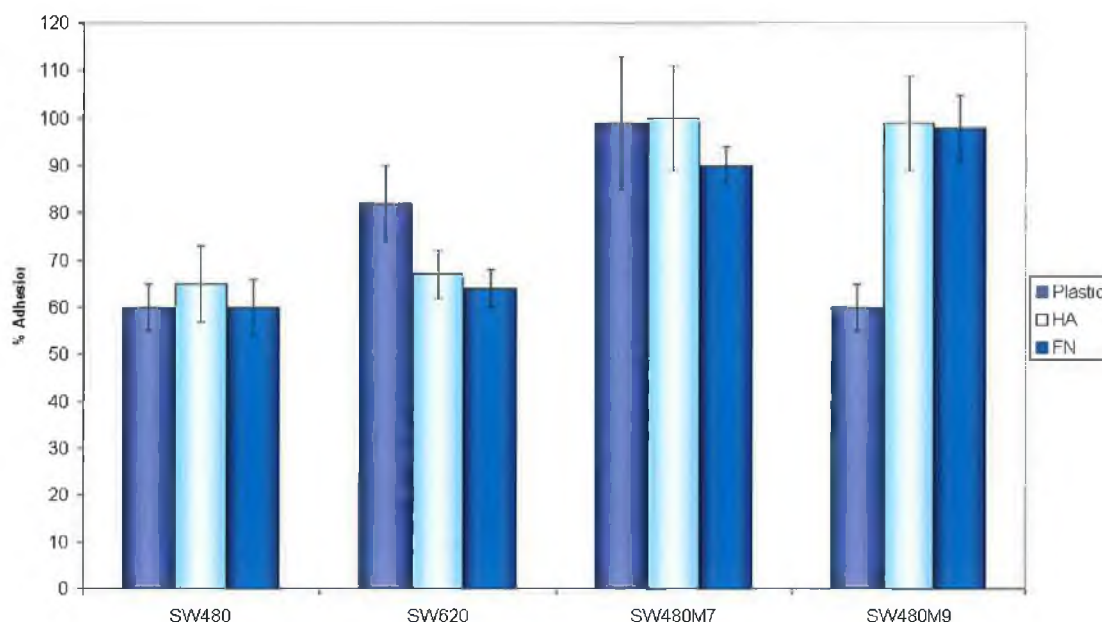
Figure 3.16 The adhesion of cells on plastic, hyaluronic acid (HA) and fibronectin (FN).

SW480M7 cells were overall the most adhesive cells, with no significant differences in their adhesion to plastic, fibronectin or hyaluronic acid. The SW480M9 cells showed the most adherence to FN and HA (Figure 3.16) when compared to the adhesion of the same cells on plastic. All percentage adhesion values were significant ($p < 0.005$).

3.2.6 Effect of ECM components on MMP protein expression

To investigate the effect of ECM components on MMP expression, cells were cultured in the presence of HA, FN, type IV collagen and a β1-integrin activating antibody as outlined in section 2.2.13.

After treatment, neither the SW480, SW620 or SW480M7 cells showed any increase in MMP-9 expression levels (Figure 3.17). The SW480M9 cells showed no increase in MMP-9 activity in the presence of FN or the β1-integrin-stimulating antibody and a slight increase when grown on type IV collagen. Growth of the SW480M9 cells on HA greatly increased the levels of MMP-9. Densitometric analysis of the gel clearly showed this increase (Figure 3.18).
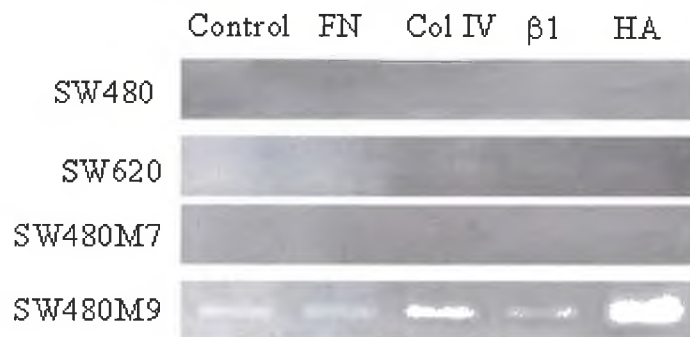
Figure 3.17 Effect of ECM components on MMP-9 activity. Cells were treated with fibronectin (FN), collagen type IV (Col IV), β1 integrin stimulating antibody and hyaluronic acid (HA). Untreated cells acted as a control.
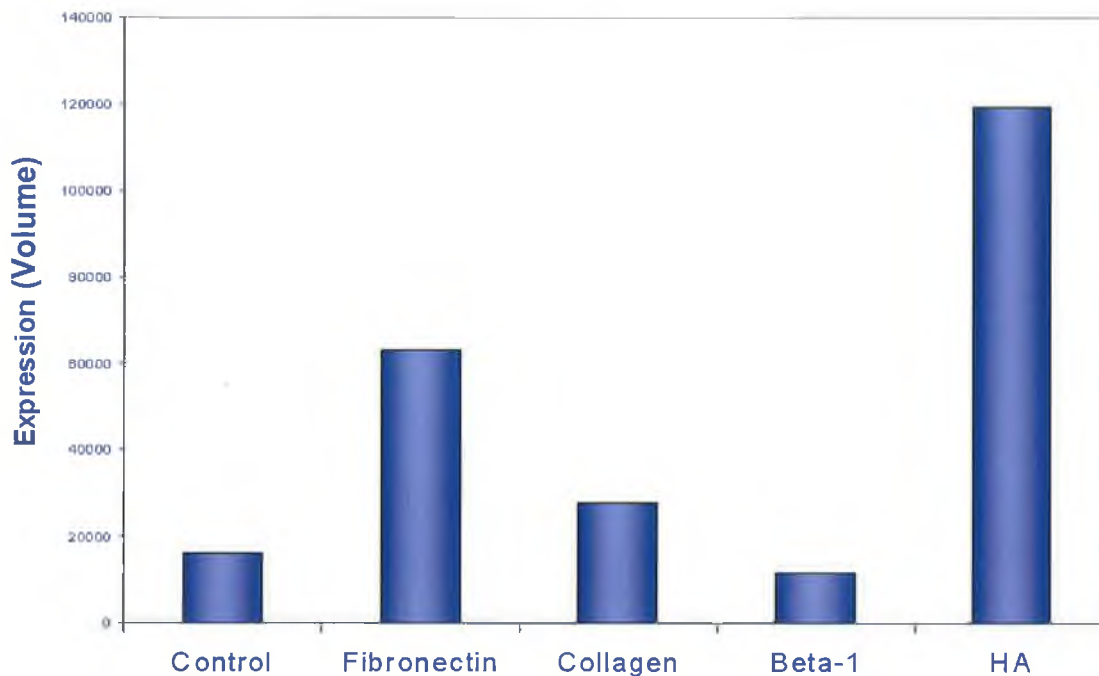
Figure 3.18 The effect of ECM components on MMP-9 activity in the SW480M9 cells. Densitometry was performed on the resulting zymography gel and it showed hyaluronic acid to have the greatest enhancing effect.

3.2.7 Effect of ECM components on MMP-9 mRNA expression

RNA from SW480M9 cells was also harvested and analysed for MMP-9 expression by real time PCR. Firstly, β-actin PCR was carried out.
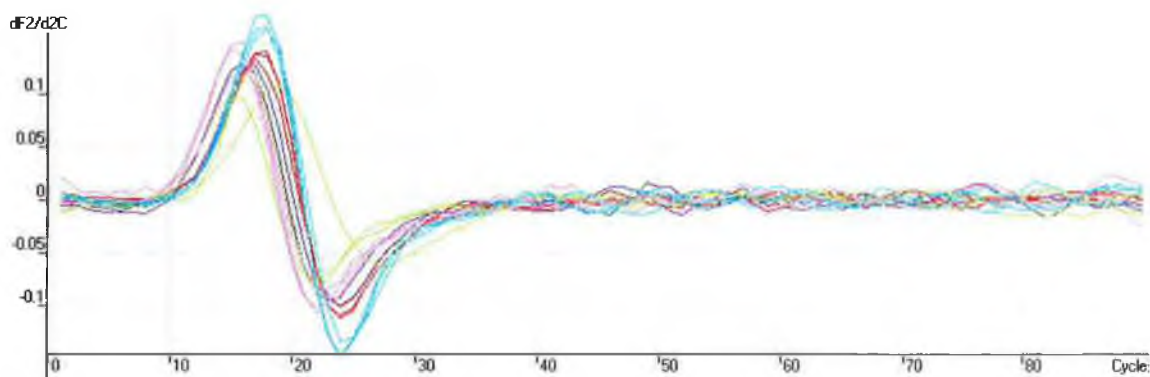
Figure 3.19 β-actin real time PCR comparative concentration curve. The data obtained and the colours used are summarised in Table 3.4.

Table 3.4: β-actin (Act) real time PCR data for SW480M9 cells treated with fibronectin (FN), hyaluronic acid (HA), β1 integrin antibody and collagen (Coll). This data was used to normalise the MMP-9 real time PCR data.

Colour	Name	Takeoff	Comparative Concentration
■	SW480M9 Plastic Act	16.0	1.11E+00
■	SW480M9 Plastic Act	16.1	1.11E+00
■	SW480M9 Plastic Act	15.8	2.02E+00
■	SW480M9 FN Act	16.4	1.15E+00
■	SW480M9 FN Act	16.7	1.14E+00
■	SW480M9 FN Act	16.8	1.35E+00
■	SW480M9 HA Act	15.6	1.00E+00
■	SW480M9 HA Act	15.6	1.00E+00
■	SW480M9 HA Act	15.6	1.00E+00
■	SW480M9 β-1 Act	15.4	1.13E+00
■	SW480M9 β-1 Act	15.3	1.20E+00
■	SW480M9 β-1 Act	15.3	1.20E+00
■	SW480M9 Coll Act	16.6	1.34E+00
■	SW480M9 Coll Act	16.7	1.24E+00
■	SW480M9 Coll Act	17.3	1.07E+00

The β-actin PCR reactions were tested for specificity by carrying out melt curve analysis. The one resulting peak (Figure 3.20) indicated that the conditions were specific.

71
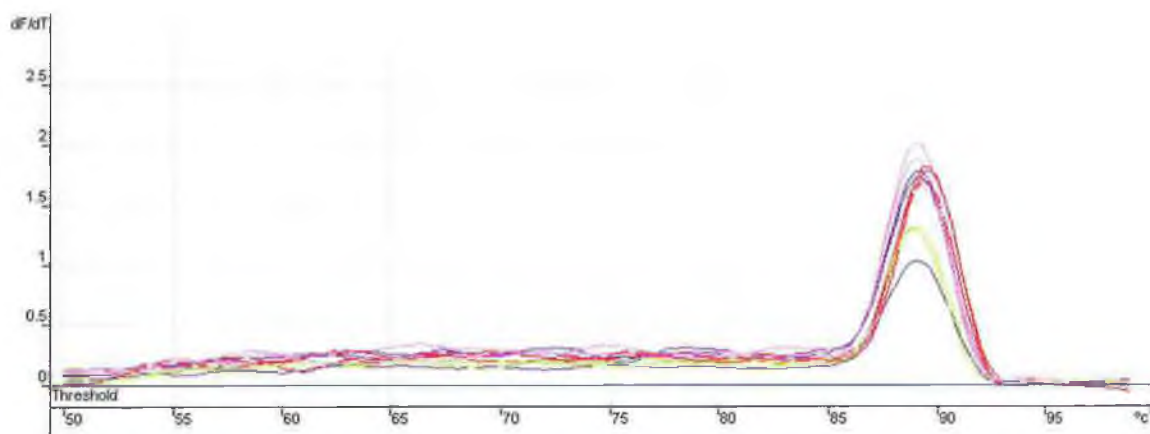
Figure 3.20 Melt curve analysis for β-actin real time PCR reactions yielded one peak at 89 °C that indicated one product was formed in the reaction.

The β-actin PCR data was used to normalise the comparative concentration data obtained from MMP-9 real time PCR reactions carried out on RNA harvested from the cells treated with various ECM components.
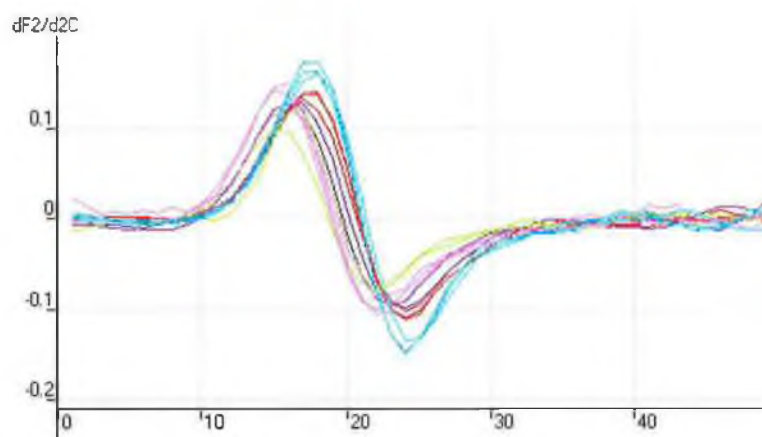
Figure 3.21 Comparative concentration of MMP-9 in SW480M9 cells treated with various ECM components. Real time PCR was carried out on RNA from SW480M9 cells. The data obtained and the raw data is listed in Table 3.5.

Figure 3.21 shows the comparative concentration curve for the MMP-9 reactions and Table 3.5 gives the raw data obtained.

Table 3.5 MMP-9 real time PCR raw comparative concentration data for SW480M9 cells treated with fibronectin (FN), hyaluronic acid (HA), β1 integrin antibody and collagen (Coll).

Colour	Name	Takeoff	Comparative Concentration
■	SW480M9 Plastic M9	16.0	1.00E+00
■	SW480M9 Plastic M9	16.1	1.00E+00
■	SW480M9 Plastic M9	15.8	1.26E+00
■	SW480M9 FN M9	16.4	1.26E+00
■	SW480M9 FN M9	15.7	1.34E+00
■	SW480M9 FN M9	16.8	1.34E+00
■	SW480M9 HA M9	15.6	9.43E-01
■	SW480M9 HA M9	16.6	1.00E+00
■	SW480M9 HA M9	15.6	1.00E+00
■	SW480 β-1 M9	16.4	1.00E+00
■	SW480 β-1 M9	15.3	1.43E+00
■	SW480 β-1 M9	15.3	1.56E+00
■	SW480M9 Coll M9	16.6	1.01E+00
■	SW480M9 Coll M9	16.7	1.22E+00
■	SW480M9 Coll M9	17.3	1.17E+00

Again, melt curve analysis was performed to ensure only one amplified product was obtained in the PCR reactions. One peak indicated the reaction was specific for MMP-9 amplification (Figure 3.22).
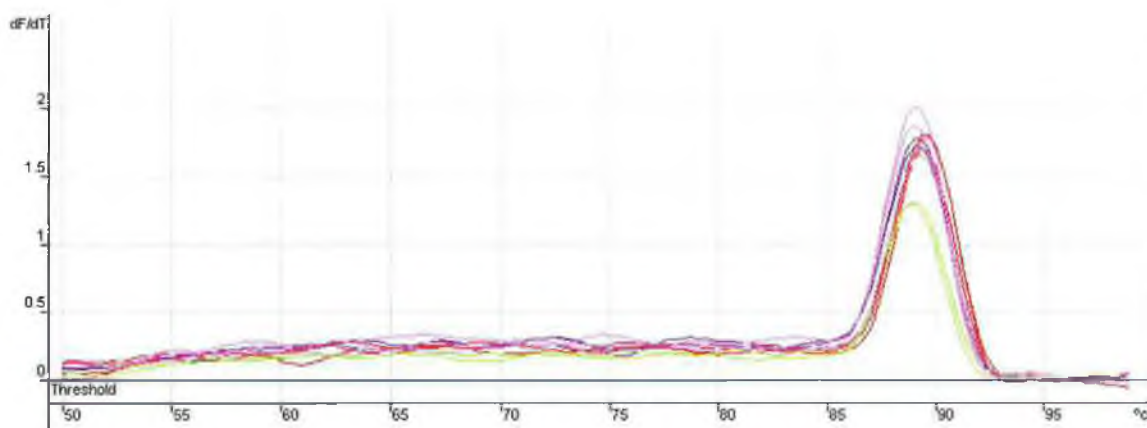
Figure 3.22 The above melt curve indicated that the MMP-9 real time PCR reaction was specific.

The plot of normalised expression showed MMP-9 levels to increase slightly upon treatment with ECM components (Figure 3.23). None of the treatments resulted in a significant increase in MMP-9 mRNA expression.

73
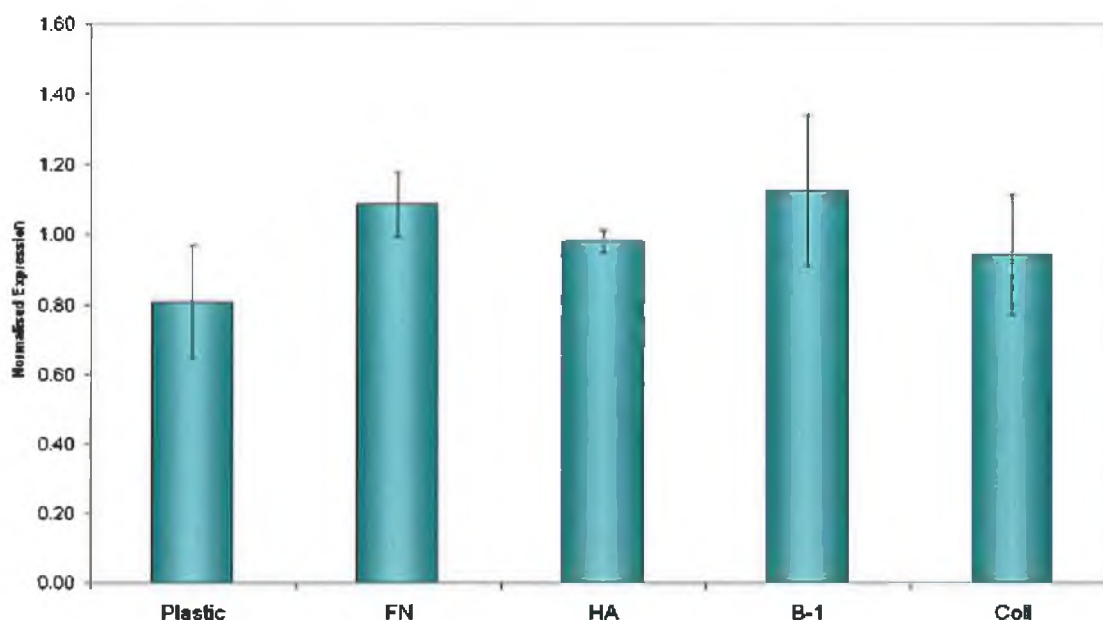
Figure 3.23 The effect of ECM components on MMP-9 mRNA expression as shown by normalised real time PCR expression data. β-actin data was used to normalise the MMP-9 data.

3.2.8 Effect of CD44 activation on MMP protein activity

Since CD44 is the major cell surface receptor for HA, the SW480M9 were treated with an anti-CD44 monoclonal antibody F10-44-2. Instead of blocking its activity, F10-44-2 activates CD44, mimicking the ligation of HA in the ECM and subsequently activating its associated intracellular signalling pathways. CM from antibody treated cells was collected at timed intervals and analysed for MMP expression by zymography and western blot analysis.
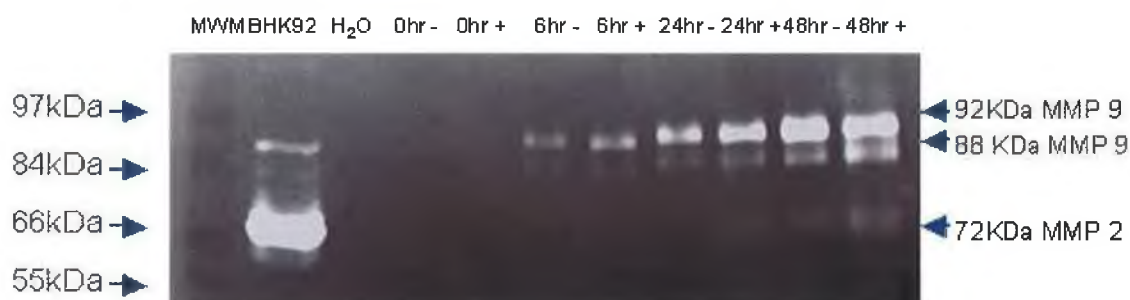
Figure 3.24 The effect of crosslinking CD44 on MMP-2 and MMP-9 activity. SW480M9 cells were treated with CD44 activating antibody (+) or with mouse IgG as a control (-). Conditioned media was analysed for gelatinases at times 0, 6, 24 and 48 hr by zymography.

74

Zymography showed that within 6 hr of treatment with the CD44-activating antibody, cells showed a slight increase in 92 kDa MMP-9 activity (Figure 3.24) when compared with cells treated with a control mouse IgG. This difference in activity was more apparent after 24 and 48 hr. Interestingly, activation of pro-MMP-9 to its active 88 kDa form was more apparent after 24 and 48 hr. 72 kDa MMP-2 activity was also detected after 24 and 48 hr of CD44 activation. These observations were confirmed by western blot analysis (Figure 3.25) with specific MMP-2 and MMP-9 antibodies.

BHK92 H₂O 0hr - 0hr +   6hr - 6hr + 24hr - 24hr + 48hr - 48hr +
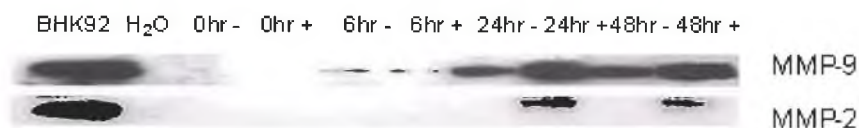
MMP-9

MMP-2

Figure 3.25 MMP-2 and MMP-9 western blot analysis of conditioned media from SW480M9 cells treated with CD44 activating antibody (+) or with mouse IgG as a control (-).

3.2.9 Effect of CD44 activation on MMP-9 mRNA expression

RNA from the cells treated for 0, 6 and 24 hours was also collected for real time MMP-9 PCR analysis. β-actin real time PCR was first carried out on these samples to allow the normalisation of later results.
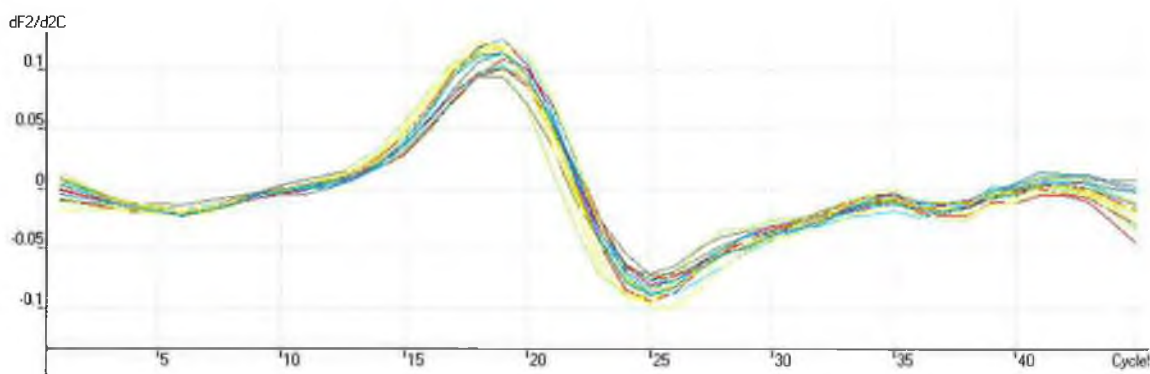
Figure 3.26 β-actin real time PCR comparative concentration curve for all samples from all time points. The data obtained and the colours used are summarised in Table 3.6.

As can be seen from the comparative concentration curve for β-actin expression (Figure 3.26) the curves for all samples fall on top of one another indicating very similar amounts of expression in these samples.

75

**Table 3.6** β-actin real time PCR was carried out to normalise the data obtained from MMP-9 PCR.

Colour	Name	Takeoff	Comparative Concentration
🟥	0- actin	15.3	1.00E+00
🟥	0- actin	14.6	1.40E+00
🟥	0- actin	14.6	1.40E+00
🟩	0+ actin	14.6	1.40E+00
🟩	0+ actin	14.2	1.70E+00
🟩	0+ actin	13.8	2.07E+00
🟦	6- actin	15.0	1.16E+00
🟦	6- actin	14.8	1.27E+00
🟦	6- actin	15.2	1.19E+00
🟫	6+ actin	15.0	1.16E+00
🟫	6+ actin	14.9	1.21E+00
🟫	6+ actin	14.6	1.40E+00
🟨	24- actin	14.9	1.21E+00
🟨	24- actin	14.3	1.62E+00
🟨	24- actin	14.4	1.55E+00
🟦	24+ actin	14.8	1.27E+00
🟦	24+ actin	14.7	1.22E+00
🟦	24+ actin	14.8	1.27E+00

Performing a melt curve tested the specificity of the β-actin real time PCR reaction. One peak indicated that only one product was formed and that no non-specific products were formed (Figure 3.27).
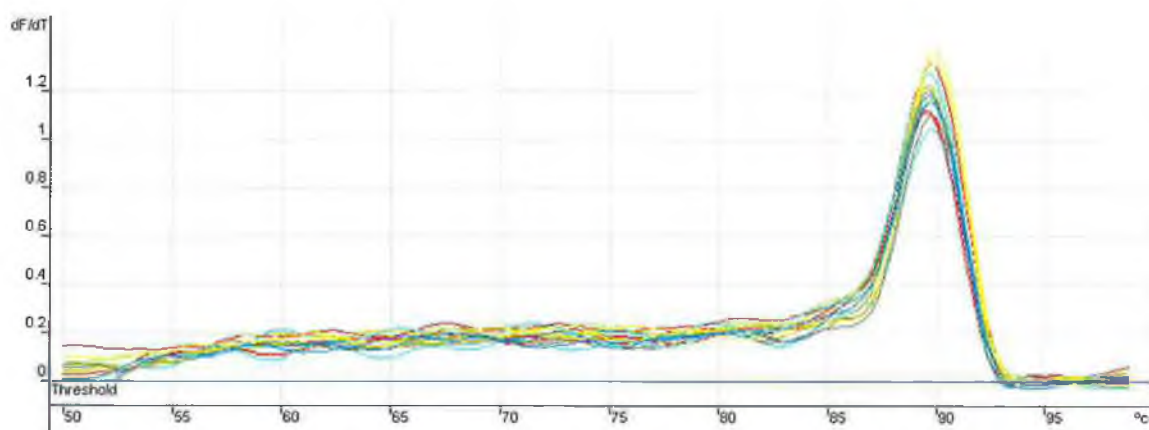
Figure 3.27 Melt curve analysis of the β-actin real time PCR reactions revealed only one product was formed with a denaturing temperature of 89 °C.
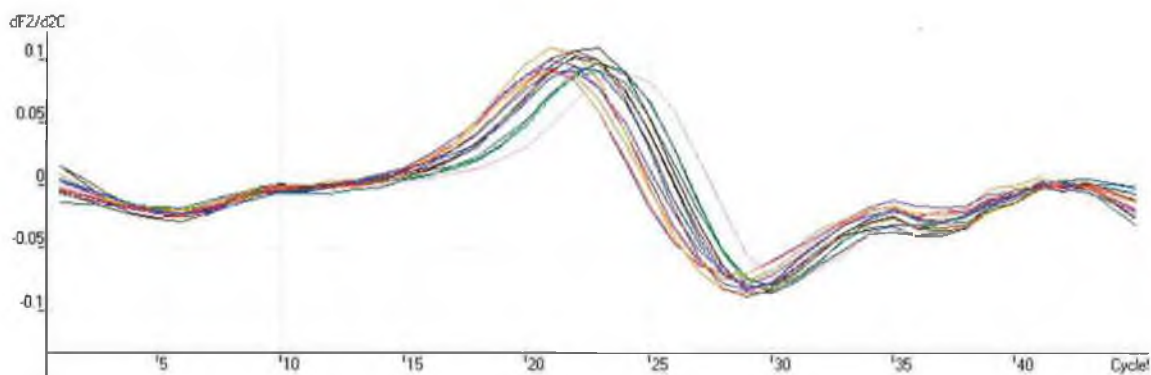
Figure 3.28 MMP-9 real time PCR for samples treated with $F_{10-44-2}$ antibody or with mouse IgG as a control. The data obtained and the colours used are summarised in Table 3.7.

Table 3.7 MMP-9 real time PCR data for samples treated with $F_{10-44-2}$ (+) or with mouse IgG (-) as a control over 0, 6 and 24 hr.

Colour	Name	Takeoff	Comparative Concentration
	0- M9	16.9	1.00E+00
	0- M9	15.8	1.58E+00
	0- M9	16.6	1.13E+00
	0+ M9	19.2	1.40E+00
	0+ M9	18.5	1.60E+00
	0+ M9	19.1	1.32E+00
	6- M9	18.6	9.94E-01
	6- M9	18.7	8.74E-01
	6- M9	18.1	9.08E-01
	6+ M9	18.0	9.03E-01
	6+ M9	18.0	8.33E-01
	6+ M9	17.7	7.17E-01
	24- M9	16.9	1.00E+00
	24- M9	16.8	1.04E+00
	24- M9	17.5	1.18E+00
	24+ M9	17.3	9.17E-01
	24+ M9	17.7	8.17E-01
	24+ M9	17.1	9.20E-01

Following β-actin PCR, MMP-9 PCR was carried out on the same samples. The resulting curve is shown in Figure 3.28 and the data obtained is listed in Table 3.7. The melt curve shows that the conditions were specific for MMP-9 amplification in all reactions because only one peak was obtained (Figure 3.29).

77
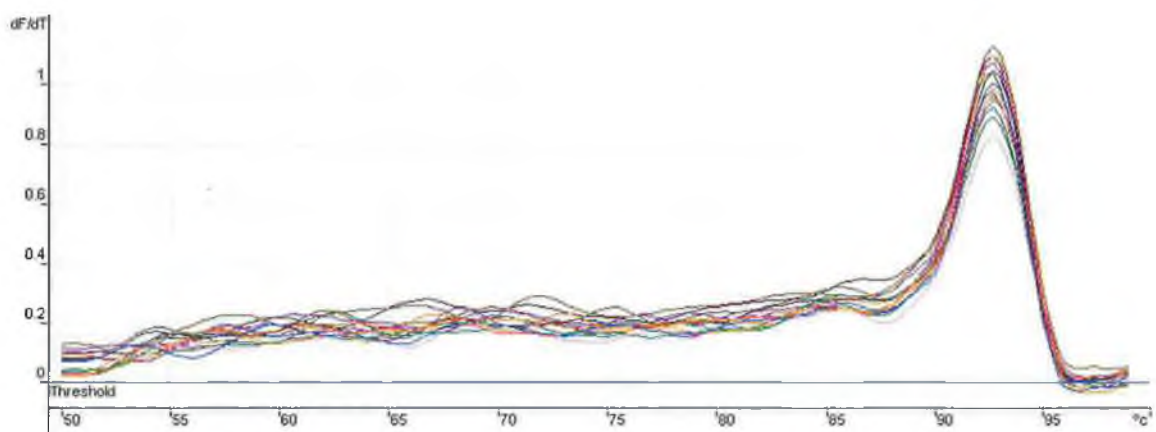
Figure 3.29 MMP-9 real time PCR melt curve analysis. One peak at 93 °C indicated that the conditions used in the reaction were specific.

The MMP-9 real time data was normalised with β-actin data and the results were plotted in Figure 3.30. There was no significant difference in MMP-9 expression between cells treated with $F_{10\text{-}44\text{-}2}$ antibody and the control.
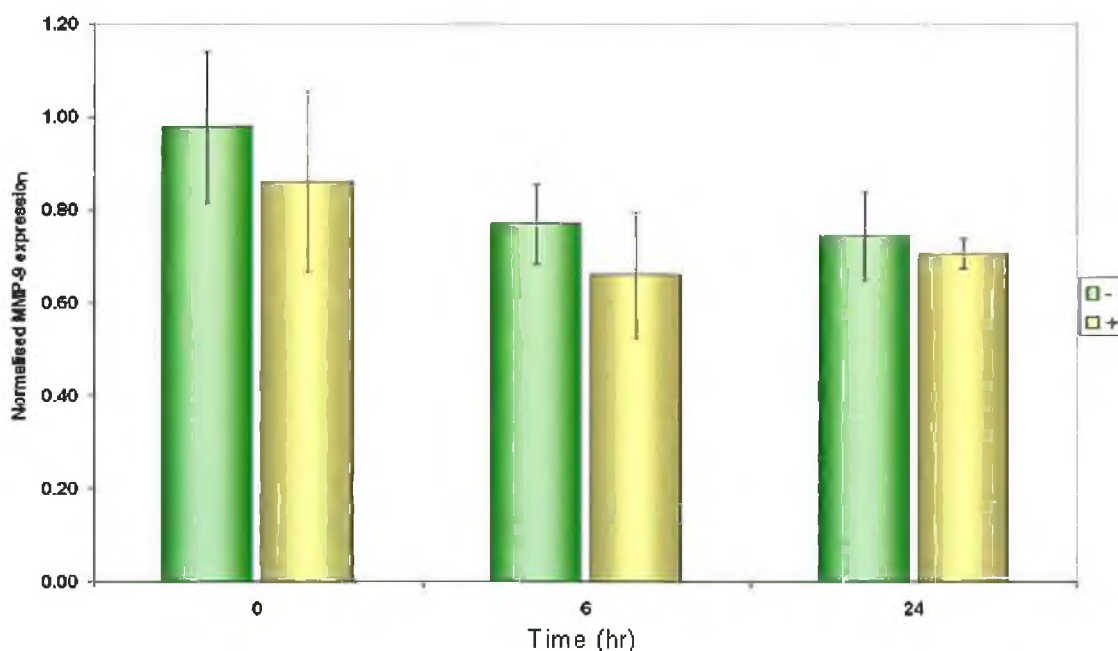

Figure 3.30 Normalised MMP-9 real time PCR expression data for SW480M9 cells treated with or without $F_{10\text{-}44\text{-}2}$ antibody over 0, 6 and 24 hours.

3.2.10 Analysis of CD44 expression in colorectal cancer cell lines

Due to the increased adhesion of the SW480M9 cells on HA and the enhancing effect of HA and CD44 stimulation on MMP-9 activity, the expression of CD44 protein

78

was examined. All cell lines studied expressed the 80kDa CD44 (Figure 3.31) often referred to as CD44s or standard CD44.
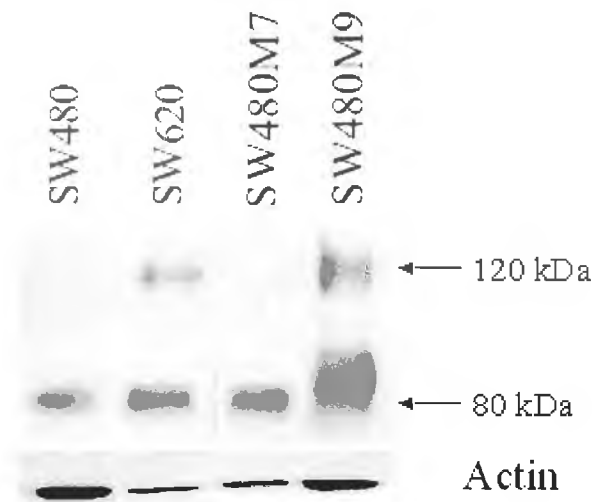
Figure 3.31 The expression of CD44 in the cell lines studied. β-actin was used as a loading control.

The SW480M9 cells expressed the highest levels of CD44s compared to the SW480, SW620 and SW480M7 cells. Interestingly, the metastatic SW620 and the more invasive SW480M9 showed expression of a 120kDa variant of CD44 with SW480M9 showing highest levels of expression.

Attempts to identify the 120 kDa CD44 variant by western blot analysis using antibodies that detect the V3 and V6 isoforms gave negative results.

3.2.11 Analysis of CD44 mRNA expression in colorectal cancer cell lines

RT-PCR was carried out as outlined in section 2.2.15.1 using primers that anneal to either end of the constant region of the CD44 gene (Figure 3.32). This PCR reaction resulted in multiple bands with different expression profiles in each sample. SW620 cells possessed a band at 516 bp that was not seen in any of the other samples.

831bp ⟶

516bp ⟶

396bp ⟶

Figure 3.32 PCR of the CD44 constant region resulted in multiple bands as a result of variant exon splicing.

RT-PCR was then carried out using primers specific for CD44 isoforms containing the V6 alternatively spliced exon. The results (Figure 3.33) showed that all cells except SW620 expressed the CD44 isoform containing v6 through v10. All cells also expressed the CD44 isoform containing v6 through v8.

SW480 SW620 SW480M9 SW480M7

821bp ⟶

512bp ⟶

⟵ v6-v7-v8-v9-v10

⟵ v6-v7-v8
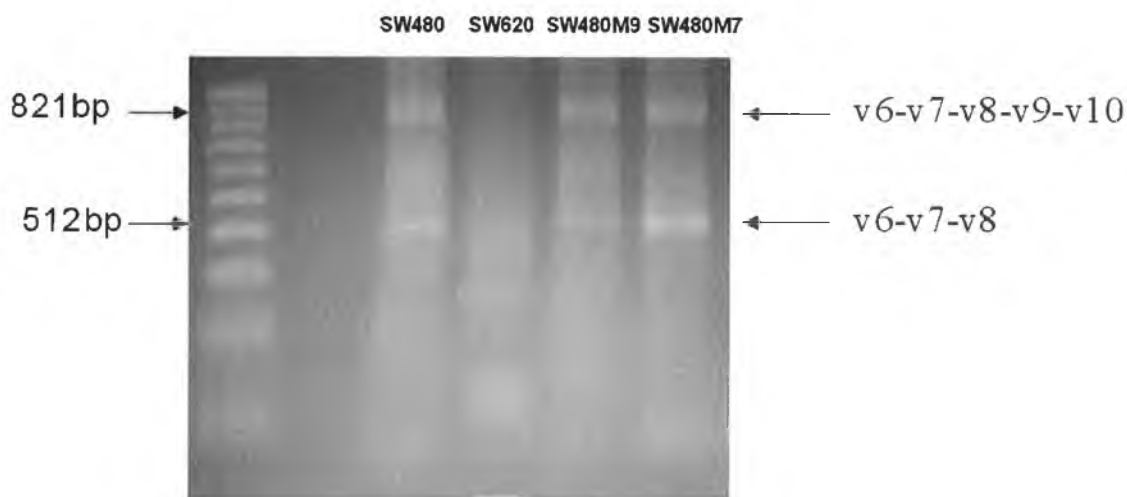
Figure 3.33 CD44 v6 RT-PCR revealed the SW620 cells to have a different v6 isoform expression profile in comparison to the other cell lines studied.

These differences in CD44 expression may help explain the increased adhesion of SW480M9 to HA as well as the effect HA had on MMP-9 activity in these cells.

3.2.12 Expression of β1 integrin in colorectal cancer cell lines

Because of its well established roles in cell signalling, the expression of β1 integrin was analysed in the cell lines used in this study. Western blot analysis (Figure

3.34) showed that all four cell lines expressed the 116 kDa $\beta 1$ subunit with the metastatic SW620 cells expressing higher levels than any other cell line studied.

Figure 3.34 Expression of $\beta 1$ integrin in cell lines used in this study. β-actin expression was used as a loading control.

3.2.13 Effect of CD44 activation on cell invasion

To show that CD44-mediated upregulation of MMP-9 expression had an effect on cell invasion the *in vitro* invasion of the cells in the presence of HA was assayed. The presence of HA increased invasion of the SW480, SW620 and SW480M9 cells (Figure 3.35). The most dramatic increase was seen with the SW480M9 cells where three times as many cells (p < 0.005) invaded when treated with HA. Therefore, the effect of F10-44-2, a CD44 activating antibody on invasion was studied. CD44 activation increased the *in vitro* invasion of SW620 and SW480M9 cells. SW480M9 invasion increased two fold (p < 0.005) (Figure 3.35) when treated with F10-44-2 antibody, in comparison with the invasion of the same cells treated with control mouse IgG. A CD44 blocking antibody (IM7.8.1) was then used to further confirm that the increased invasion was CD44-dependent. Blocking of CD44 with this antibody resulted in a significant suppression of invasion in all cells studied (p < 0.005) (Figure 3.35).
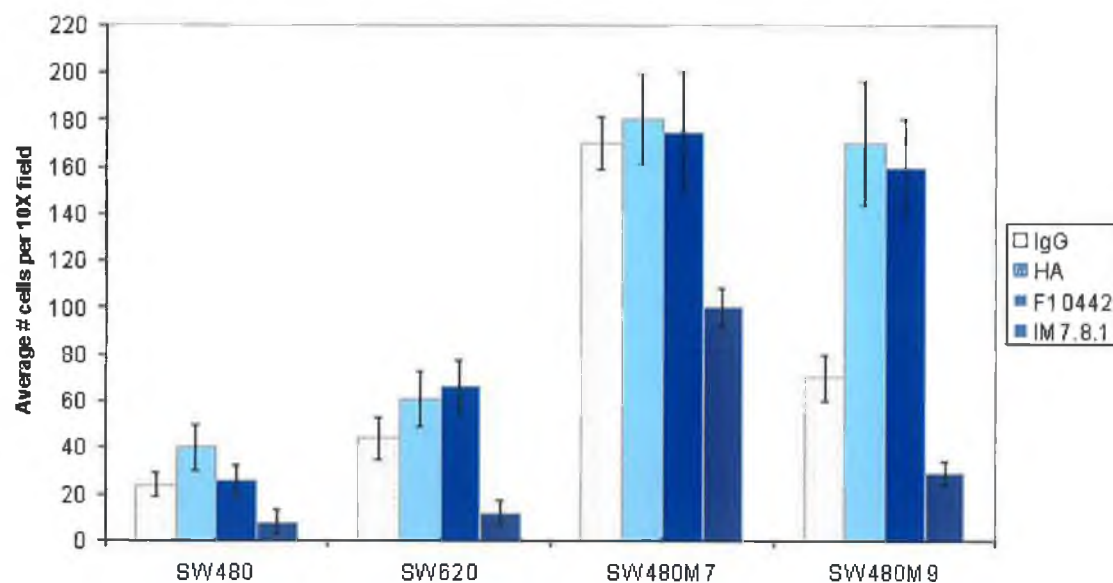
Figure 3.35 The effect of CD44 on *in vitro* invasion. Cells were allowed to invade in the presence of hyaluronic acid (HA), CD44-stimulating antibody (F10442), CD44 blocking antibody (IM7.8.1) or mouse IgG as a control.

3.3 Discussion

Modern therapeutic approaches to eliminate or control metastasis are based on defining its critical events and the specific targets that regulate these events. Therefore, the identification and understanding of novel cellular and molecular determinants of metastasis is crucial. Both CD44 and MMPs have established roles in the process but there is little insight into how they interact in metastasis. Uncovering these interactions would help validate their use in the clinic as potential targets.

In this chapter, the interactions between CD44 and MMPs in regulating invasion and migration in human colorectal cancer cell lines were investigated. The non-metastatic SW480 cell line established from a primary human colon adenocarcinoma was stably transfected with the cDNAs for MMP-7 and MMP-9. The effects of MMPs on cell invasion, adhesion and migration were assessed. Cells were grown on various ECM components to study the effect on MMP activity. CAM expression in these cell lines was analysed and the effects of CD44 activation through ligand binding and antibody treatment on MMP expression and invasion were investigated. This work aimed to demonstrate the importance of MMP and CD44 interactions in regulating colon tumour cell invasion.

MMP-7 and MMP-9 overexpression in SW480M7 and SW480M9 cells was shown by real time PCR and by western blot analysis. SW480 cells were used as a common control for both SW480M7 and SW480M9 experiments instead of SW480Neo cells which had been transfected with the empty pCMV vector used to transfect MMP-9 cDNA. SW480 cells have previously been shown to be equal to SW480Neo cells in terms of *in vitro* invasiveness. The parental SW480 cells do not express MMP-7 or MMP-9, and transfection of the cDNAs for MMP-7 and MMP-9 into these cells increased their *in vitro* invasion. MMP secretion caused increased cell invasion most likely due to ECM turnover. Matrigel, which normally hinders the progression of cells, was degraded, allowing cells to pass freely to the lower chamber. Interestingly, the SW620 cells which were established from a lymph node metastasis from the same patient as the SW480 also did not express MMP-9 but were more invasive than the parental SW480 cells. Overexpression of MMP-9 in a rat transformed embryo cell line has been shown previously to increase the invasion and migration of these cells (Bernhard *et al.*, 1994).

The MMP transfected cells also showed increased migratory capacity. The parental SW480 cells were more migratory than the SW620 which has also been shown by Hewitt *et al.* (2000). A similar result has also been obtained by Kubens and Zanker (1998) using a three-dimensional collagen matrix and time-lapse video recording assay. It is possible that ECM components and indeed matrigel are acting as inducers of cell migration and that in the absence of matrigel the SW620 cells do not migrate or invade. It is also possible that ECM turnover and therefore the generation of soluble ECM components acted as stimuli for cell progression.

Adhesion of these cells on various matrices including, plastic, FN and HA was also investigated. The MMP-9 transfected cells were more adherent to FN and HA than any of the other cell lines. Examination of the levels of CD44 protein expressed in the cell lines showed that all cell lines expressed the standard 80 kDa form of CD44. An additional higher molecular weight band was seen in the metastatic SW620 and the MMP-9 transfected cell line. This band had a molecular weight of ~120 kDa and could be CD44v6 which has been previously been shown to be involved in colon metastasis (Wielenga *et al.*, 1993). A 120 kDa CD44v6 isoform has previously been identified in the serum of patients with metastatic breast disease using western blot analysis (Lackner *et al.*, 1998). Although specific antibodies for CD44v3 and CD44v6 failed to identify this protein by western blot analysis, exon specific PCR was carried out to identify this band. CD44v6 specific PCR revealed that all cells expressed the CD44 isoform containing the v6 through v8 domains. Further analysis is therefore required to determine the identity of this 120 kDa isoform. Examination of $\beta 1$ integrin expression showed that the SW620 cells expressed the highest amounts in comparison with the other cells used in this study, all of which expressed $\beta 1$ integrin.

The effect of growing cells on various ECM components on MMP-9 activity was then investigated. Cells were grown in the presence of a $\beta 1$ integrin antibody, HA, fibronectin, and type IV collagen. MMP-9 activity was increased in SW480M9 cells grown on hyaluronic acid and on collagen. There was no effect in SW480, SW620 and SW480M7 cells. Real time MMP-9 PCR analysis of the SW480M9 cells showed that no difference occurred between any of the treatments at the mRNA level. The enhancing effect of the CD44-HA ligation interaction on MMP-9 activity was further investigated by the treatment of the SW480M9 cells with a CD44 activating monoclonal

antibody. Increased levels of MMP-9 activity were observed after 24 hr of antibody treatment and this increase in activity was greater after 48 hr. Treatment with this antibody also had the effect of increasing MMP-2 activity after 24 hr. These observations were confirmed by western blot analysis. Real time MMP-9 PCR analysis proved no changes occurred at the mRNA level.

Since growth of the SW480M9 cells on HA and direct stimulation of CD44 caused an increase in MMP activity, it was then investigated whether this increase in MMP-9 activity had a functional effect on cell invasion. Stimulation of CD44 by treatment with HA caused an increase in the invasion of the SW480, SW620 and SW480M9 cells when compared with control cells. Stimulation of CD44 using a monoclonal antibody increased the invasion of the SW620 and SW480M9 cells. SW480 cell invasion which was responsive to HA treatment showed no change when the cells were treated with the monoclonal antibody possibly due to the fact that HA has other receptors besides CD44. CD44 induced invasion was inhibited by treatment with a CD44 blocking antibody. SW480M7 cells, which were overall the most invasive cells, showed no difference in invasion upon CD44 ligation. Interestingly, the most responsive cells were the SW620 and SW480M9 cells both of which expressed higher levels of CD44 as well as the 120 kDa CD44 isoform. The presence of 120 kDa CD44 alone was not sufficient for a large HA-induced increase in invasion and endogenous MMP-9 expression was also required. This was well displayed by the difference in cell invasion between the SW620 and SW480M9 cells in response to CD44 activation. However, regardless of the MMP expression profile, inhibition of CD44 function significantly reduced the invasion of all cells used in this study. This result shows the importance of CD44 in cell invasion, while all cell lines used were not equally responsive to CD44 activation, in terms of cell invasion they all behaved similarly when CD44 function was blocked. CD44 binds many components of the ECM and blocking this interaction completely, would result in complete loss of function thus both explaining the decrease in invasion as well as highlighting the importance of CD44 in this process. These observations suggest a dual role for CD44 in cell invasion, one, which involves MMP expression, and one, which is independent of MMP expression. Upon CD44 activation, endogenous MMP activity was enhanced resulting in increased invasion. In addition, blocking of CD44 even in cells that did not express MMP-9 still resulted in a reduction in cell invasion.

Many studies have recently demonstrated the importance of MMP - CAM interactions in tumour cell invasion and metastasis. The importance of integrins in regulating MMP expression was highlighted by a study showing αVβ6 integrin to promote squamous carcinoma cell invasion through up-regulation of MMP-9 (Thomas *et al.*, 2001). α2β1 integrin-mediated ligation of collagen by ovarian carcinoma cells results in tyrosine kinase signalling and the subsequent expression of MT1-MMP which activates latent MMP-2 resulting in enhanced cell surface MMP activity (Ellerbroek *et al.*, 1999). Stimulation of CD44 has been shown to increase MMP-2 expression in human melanoma cells (Takahasi *et al.*, 1999) and in lung carcinoma cells (Zhang *et al.*, 2002). Treatment of melanoma cells with the same CD44 stimulating antibody used in this study has been shown to upregulate MMP-2 expression and enhance their migratory ability (Takahashi *et al.*, 1999). This type of use of monoclonal antibodies renders them superior to natural ligands, which often bind to multiple receptors. Blocking of CD44 with a monoclonal antibody has previously been shown to significantly decrease the migration and invasion of murine mammary carcinoma cells (Ladeda *et al.*, 1998). The CD44v3,8-10 isoform has been shown to be associated with MMP-9 at invadopodia of breast cancer cells and CD44v3 has been shown to recruit MMP-7 to the surface of neuroblastoma and lymphoma cells thus suggesting a collective role for MMPs and CD44 in tumour cell invasion and migration (Bourguignon *et al.*, 1998). Overexpression of this isoform has also been shown to result in enhanced metastasis in mice (Barbour *et al.*, 2003). The individual roles of CD44 isoforms in tumour metastasis are strengthened by a study showing that expression of antisense CD44v6 inhibited colorectal metastasis in nude mice (Reeder *et al.*, 1998). In contrast to this observation, CD44v6 expression has been shown to decrease in the advanced stages of colorectal metastasis suggesting a role for this variant in early colorectal neoplasia (Weg-Remers *et al.*, 1998). MT1-MMP mediated cleavage and subsequent release of CD44 has been shown to promote cell migration in breast carcinoma cells (Kajita *et al.*, 2001). Such proteolytic cleavage of CD44 results in the translocation of the CD44 intracellular domain to the nucleus where it activates transcription through the TPA responsive element (Okamoto *et al.*, 2001). Inhibition of CD44 has been shown to inhibit the invasion of human glioma cells (Okada *et al.*, 1996). Interestingly, Yu and Stamenkovic (2000) have shown that MMP-9 can associate with CD44 on the surface of mouse mammary carcinoma and human melanoma cells, mediating migration and invasion. This association may cause activation of signal transduction molecules, in the way that HA-CD44 binding has been shown to activate the ras-MEK1 and the PI3

kinase-Akt signalling pathways in a human lung carcinoma cell line (Zhang *et al.*, 2002).

3.4 Conclusion

In this chapter, it has been shown that culturing colon cancer cells on ECM components increased MMP-9 expression. HA, the major receptor for CD44 caused the greatest increase. In addition, direct stimulation of CD44 by treating with an activating antibody (F10-44-2) caused an increase in MMP-2 and MMP-9 expression and caused activation of MMP-9. This is the first report to show up-regulation of MMP-9 expression following CD44 stimulation in colon cancer cells (Murray *et al.*, 2004). This upregulation was shown to be post-transcriptional as the levels of MMP-9 mRNA were unaffected. It is possible that CD44 activation directly triggers intracellular signals that result in MMP induction. Indeed ligation of CD44 on T cells has been shown to lead to tyrosine phosphorylation of intracellular proteins suggesting a role in signalling (Taher *et al.*, 1996).

The interactions between CD44 and MMPs are diverse, and further studies would clarify exactly what mechanism underlies these observations. Figure 3.36 summarises some known CD44-MMP interactions. This summary highlights how numerous these interactions are. The absence of any difference at the transcriptional level suggests that CD44 activation cannot induce MMP mRNA expression but it may affect post-transcriptional events. This explains why cell lines that did not express MMPs were not responsive to CD44 activation.

This study also highlights a specific interaction between MMP-9 and CD44. SW480M7 cells, which were stably transfected with the cDNA for MMP-7, were not as adherent to HA, did not express high levels of CD44 and their invasion was not responsive to CD44 activation. Expression of the 120 kDa isoform was also absent in SW480M7 cells. SW480M9 cells which were stably transfected with the cDNA for MMP-9 expressed a 120 kDa CD44 isoform, so too did the metastatic SW620 cells. There was a correlation between the expression of the 120 kDa CD44 and a CD44-responsive increase in invasion. Why cells overexpressing MMP-9 have differential splicing of CD44 remains to be elucidated. These cells were more adherent to HA and this adhesion interaction also resulted in enhanced MMP-9 activity.
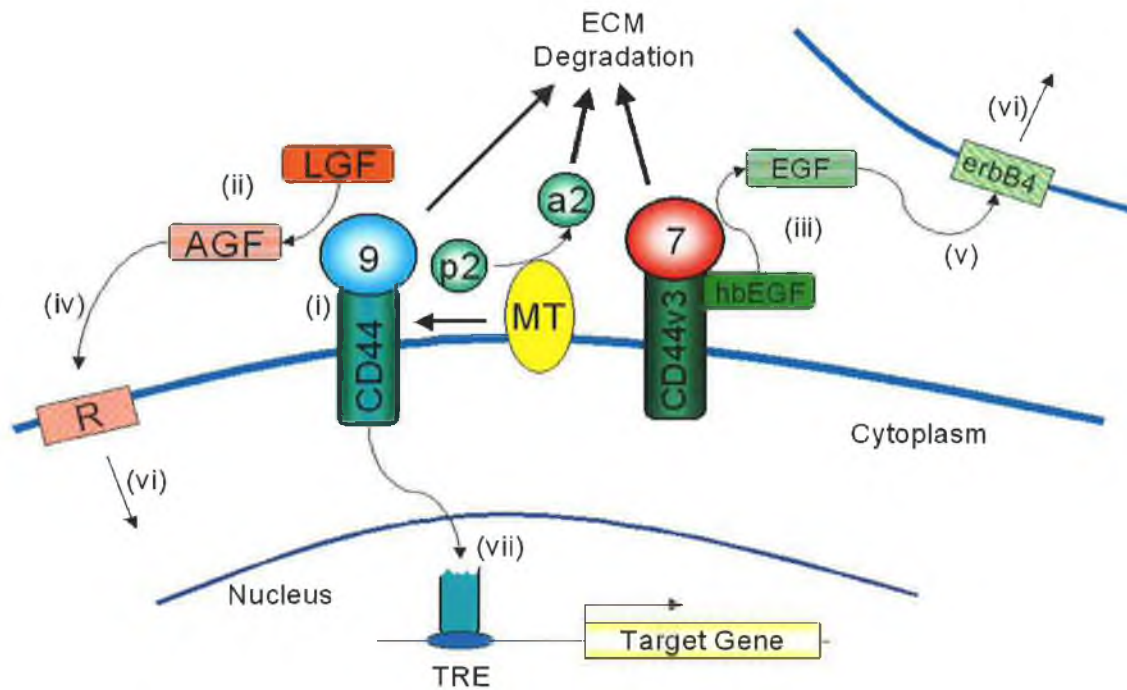
Figure 3.36 Summary of interactions between CD44 and MMPs. (i) CD44 can dock MMP-9 (9) on the cell surface thus directing its ECM degrading activity as well as (ii) the MMP-9 mediated activation of latent cytokines or growth factors (LGF) to active growth factors (AGF). CD44v3 can dock MMP-7 (7) and heparin bound epidermal growth factor (hbEGF) allowing MMP-7 to activate hbEGF (iii). Activated cytokines or growth factors can then bind to their respective receptors in an autocrine (iv) or paracrine (v) manner where they then transduce intercellular signals (vi). As well as activating the latent pro form of secreted MMP-2 (p2) to its active form (a2). MT1-MMP (MT) can process CD44. The resulting intracellular domain can translocate to the nucleus (vii) and promote transcription through TPA-responsive elements (TRE), including transcription of CD44 itself.

Although the mechanism remains to be elucidated, CD44 activation clearly induces MMP-9 production. These results suggest that CD44 activation can directly increase the levels of MMP-9 and cause an increase in invasion in colon cancer cell lines.

89

Chapter 4

Gene expression profiling of a panel of colorectal

cancer cell lines using cDNA microarrays

4.1 Introduction

4.1.1 An introduction to microarray technology

Microarray-based technologies have become the most widely used analytical technique for the study of gene-expression patterns on a genome-wide scale (Leung and Cavalieri, 2003 and Schena *et al.*, 1995). Because cellular processes are governed by the repertoire of expressed genes, and the levels and timing of expression, it is important to have tools to monitor a large number of mRNAs in parallel. The collection of genomic information from two types of biological states and the ability to determine differences in mRNA expression profiles between these states using DNA microarray technology has become a powerful and valuable genomic tool to elucidating the molecular basis of almost every aspect of cancer biology as well as many other diseases. DNA array based technologies allow one to study expression levels of many genes in parallel, thus providing static (i.e. in which sample a gene of interest is expressed) and dynamic (i.e. how the expression pattern of one gene relates to those of others) information about gene expression (Schena *et al.*, 1996 and Lockhart *et al.*, 1996). Essentially, one can compare the expression level either of the same gene in different samples, or of different genes in the same sample, thus giving an insight into gene function and regulation.

Like many other molecular techniques, DNA microarrays exploit the sequence complementarity preferential binding features of the DNA duplex (Southern *et al.*, 1999). Because of mutual selectivity between complementary strands of nucleic acids, these techniques provide high sensitivity and specificity of detection. In many ways the path to array based methods began by binding DNA to solid supports such as nitrocellulose in 'dot-blot' methods (Southern, 1975). Briefly, this method typically involved mixtures of nucleotides attached to a solid support followed by incubation with a labelled gene-specific target.

Most of this section concerns cDNA microarrays, however other array technologies including oligonucleotide arrays have been developed. Affymetrix have developed an approach based on hybridisation to small but high-density arrays containing thousands of synthetic oligonucleotides. These arrays are synthesised *in situ* using a combination of photolithography and oligonucleotide chemistry (Lockhart et *al.*,

1996 and Lipshutz, *et al.*, 1999). One advantage of Affymetrix chips over cDNA arrays is that every gene is represented at least 5 times with equally as many mismatch negative controls. Although these chips provide reliable and reproducible data, Affymetrix analysis is very expensive with one experiment costing at least €1,000.

With cDNA technology, the hybridisation strategy is altered in many ways (Figure 4.1). Firstly, with cDNA arrays impermeable rigid supports such as glass are used. As liquids cannot penetrate the support, the rate of hybridisation between target nucleic acids and the probes is enhanced. In array experiments, many gene-specific single-stranded polynucleotides are individually spotted on a single matrix. This matrix is then simultaneously probed with fluorescently tagged cDNA representations of total RNA pools from different samples, allowing the determination of the relative amount of transcript present in the pool by the intensity of fluorescent signal generated. Typical samples for analysis include cell lines and both clinical and laboratory animal derived-tissue biopsies. The quantitation of relative transcript abundance is based on the direct comparison between an 'experimental' cell state and a 'control' cell state.
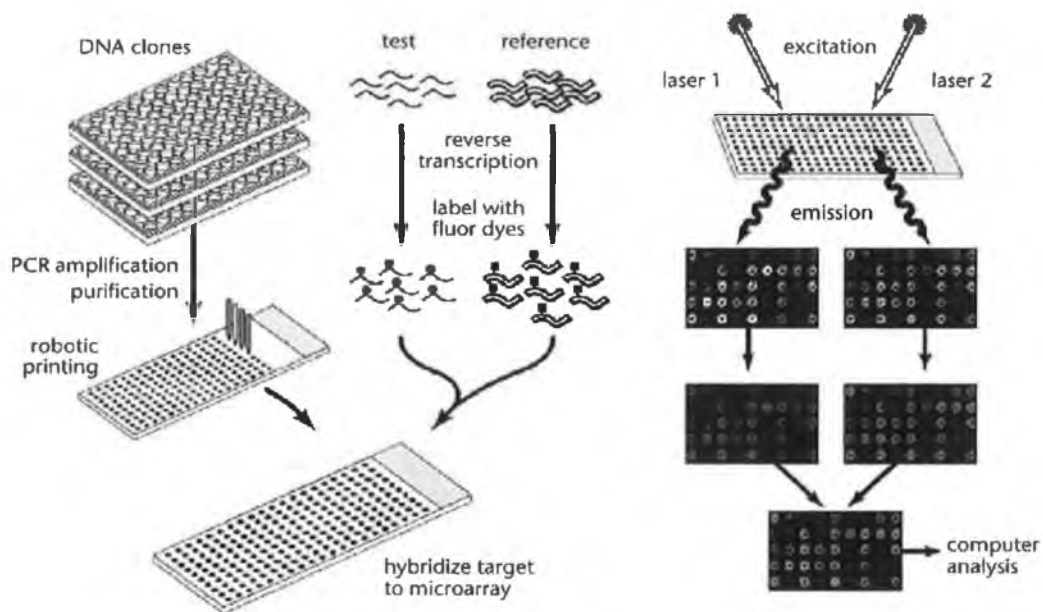
Figure 4.1 cDNA microarray basics. Templates of interest are amplified from cDNA libraries and arrayed on a glass slide. mRNA from 'test' and 'reference' samples are reverse transcribed and labelled with separate fluorescent dyes which are mixed and allowed to hybridise to the array. Slides are scanned by laser excitation of the incorporated targets and emission signals in both channels are compared. Figure adapted from Duggan *et al.*, 1999.

4.1.1.1 Array fabrication

Array fabrication begins with the selection of clones to be spotted onto the slide. These are often chosen from databases such as GenBank or UniGene (http://www.ncbi.nlm.nih.gov, 17 September 2004, and Bowtell, 1999). cDNA arrays are produced by robotically spotting or printing a few nanolitres (100 – 500 µg/ml) of purified PCR products of approximately 1 to 2 kb in size, representing specific genes onto a slide (Cheung *et al.*, 1999). Most cDNA arrays are spotted on glass microscope slides that have been treated with poly-L-lysine, which enhances the hydrophobicity of the slide and the adherence of the printed DNA, thus limiting the spread of the spotted DNA droplet on the slide. DNA is then crosslinked to the matrix by UV irradiation.

4.1.1.2 Experimental design

Proper planning of microarray experimental design is required to ensure that questions of interest can be answered accurately. Inadequate experimental design could result in either the loss of real biological data or the acquisition of false and untrustworthy information. Good microarray experimental designs should minimise the variation introduced at each stage of the experiment. The design of an array experiment has three layers (Figure 4.2).

Independent sources of variation are introduced at each of these layers. The top layer is subject to biological variance, which is unavoidable and intrinsic to all biological systems. Biological replication by using multiple animals or multiple flasks of cells in culture, is important in ensuring the end results are statistically significant, real and reproducible (Rosenbaum, 2001 and Lee *et al.*, 2000). Pooling of labelled replicate RNA prior to hybridisation can reduce the biological component of variation but not the technical error due to sample handling or measurement.
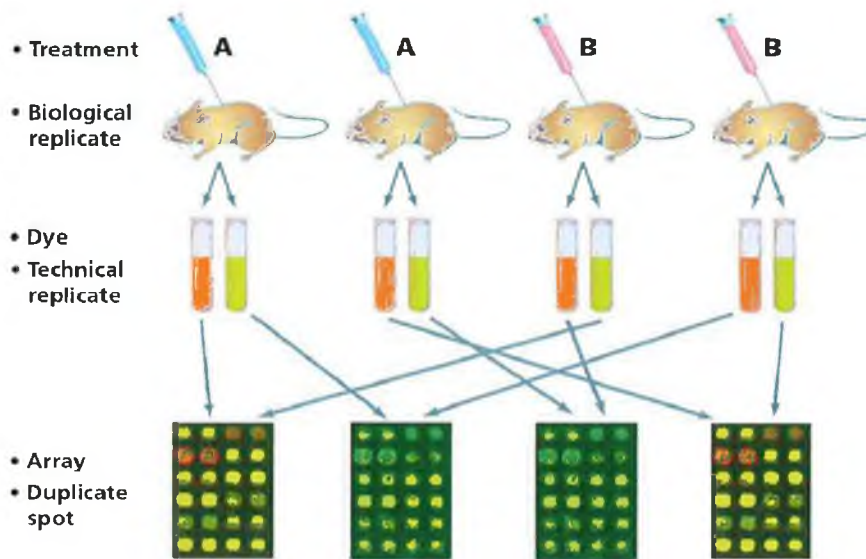
Figure 4.2 The three layers of two-colour microarray experimental design in a hypothetical experiment that compares the effects of two treatments, A and B on gene expression in mice. At the top layer are the experimental units, i.e. the two mice and assigning two mice to each treatment ensures biological replication. In the middle layer, obtaining two RNA samples from each experimental unit and assigning them two different colour dye labels ensures a technical replicate. These technical replicates are paired (one red with one green), mixed and co-hybridised to microarrays. The lower layer of a microarray experiment involves the array itself and the arrangement of the elements on the array. This layer of design includes deciding which cDNA probe sequences to print, how many replicates to print, which control sequences to print and where these should be printed. Figure adapted from Churchill (2002).

Technical variation is introduced in the middle layer by the processes of RNA extraction, labelling and hybridisation. Deciding which and how many technical replicates to pair and co-hybridise together can reduce the middle layer variability (Kerr and Churchill, 2001a and Yang and Speed, 2002). The dye-swap experiment is effective for the comparison of two samples (Kerr and Churchill 2001b). Using two arrays for the comparison of the two samples, this design also accounts for dye-bias with the dye assignments being reversed on the second array. Another measure used to minimise middle-layer variation is the use of a reference sample to which all samples in the array experiment are directly compared with, allowing subsequent sample to sample cross comparison. Reference samples are constructed using complex mixtures of tissue- or cell line -derived RNA (Quackenbush, 2002).

Error associated with the measurement of fluorescence is introduced at the lower layer. This can be a result of dust on the slide or uneven background across the slide due to non-specific hybridisation. An important part of lower-layer design with a great impact on array data analysis is the arrangement of spots on a slide (Yang *et al.*, 2002b). Important considerations here are which probes to spot on the slide, how many replicates of these should be spotted and where these replicates are distributed about the slide. Because array printing is subject to error and heterogeneous spot quality is common, replicate spot printing is advisable and increases precision. Negative controls (PCR mix with no template) and blanks (nothing printed) are often distributed about the slide to give an estimation of the background levels in the sub arrays.

4.1.1.3 Probe preparation

Microarrays assay differential gene expression by the co-hybridisation of fluorescently labelled probes prepared from different RNA sources. Labelled representations of cellular mRNA pools, usually reverse transcriptions, are used as targets for cDNA arrays. The purity and quality of the starting RNA is critical to the hybridisation performance, and therefore has a significant effect on the results of the assay. Contaminants such as cellular proteins, lipids and carbohydrates can mediate non-specific labelled cDNA binding. The commonly used fluorescent Cy dyes Cy3-dUTP and Cy5-dUTP are incorporated at the reverse transcription stage. Cy3 and Cy5 dyes have good photostability and are widely separated in their excitation (550 and 650 nm respectively) and emission spectra (570 and 670 nm respectively), which allows highly selective optical filtration.

4.1.1.4 Hybridisation

The main goal in the hybridisation step is to obtain high specificity while minimising background. The Institute of Genomic Research (TIGR) have developed protocols that yield reproducible, quality hybridisations while maximizing the measured fluorescence on the array (http://www.tigr.org/, 17 September 2004, and Hegde *et al.*, 2000). Pre-hybridisation in a 1 % BSA solution has the advantage of both blocking non-specific interactions between the probe and the slide as well as washing unbound DNA from the slide before the hybridisation step. This free DNA would otherwise compete with bound DNA for hybridisation.

95

4.1.1.5 Image acquisition and image processing

The next step in assessing differential gene expression is scanning the hybridised slide using a confocal laser scanner capable of detecting both the Cy3- and Cy5-labeled probes and producing separate 16-bit greyscale tagged image file format (TIFF) images for each. However, most scanners are capable of distinguishing and detecting many other additional fluorescent dyes (Bowtell, 1999). Grids specifying target locations are overlaid on the images that are then analysed to calculate the relative expression level of each individual gene and hence identify differentially regulated genes. The hypothesis underlying microarray analysis is that the measured intensity for each arrayed gene represents its relative expression level. Local background is often sampled and subtracted from the signal and can be used to specify a threshold which the true signal must exceed (Chen *et al.*, 1997). Local background sampling for each spot is preferred to global background sampling for the entire slide as uneven background across the array often arises during hybridisation. This can be due to non-specific hybridisations or artefacts such as dust on the slide surface. Local background-subtracted hybridisation intensities (I_{LBS}) for each spot in both Cy3 and Cy5 channels are calculated and then used to calculate the Cy3 to Cy5 ratio of I_{LBS} intensities for each individual spot (gene).

4.1.1.6 Data normalisation

Before biologically relevant patterns of expression are identified by comparing expression levels between samples on a gene-by-gene basis, it is appropriate to carry out normalisation on the data to eliminate questionable, low quality measurements and to adjust the measured intensities to make them comparable. The process of obtaining actual readings of expression levels from a cDNA microarray is subject to several sources of variability (biologic variability, printing of slides, preparation of fluorescent probes, measurement of the fluorescent light intensity etc) that will all be represented in the random fluctuations of observed expression levels (Schuchhardt *et al.*, 2000). True differentially expressed genes need to be distinguished from differences generated by fluctuations in the data. The many sources of systematic variation in microarray experiments that affect the measured expression levels are best demonstrated in an experiment where two identical mRNA samples are labelled with different dyes and

hybridised to the same slide (Yang *et al.*, 2002c). In this example, it was rare to have the dye intensities equal across all spots. Although such systematic differences may be small, they may also have a confounding effect when assaying for subtle biological differences. Normalisation is a process that attempts to remove this variation. Normalisation of the individual hybridisation intensities in each of the two scanned channels is essential to adjust for label-specific differences such as incorporation and detection efficiencies thus facilitating channel-to-channel comparisons (Quackenbush, 2002). Normalisation is also necessary to adjust for differences in the starting RNA in the samples. All these issues can shift the average Cy3/Cy5 ratio if these intensities are not rescaled. Only by balancing these individual intensities appropriately can meaningful biological comparisons be made.

Total intensity normalisation considers all the genes in an array experiment and is based on the simple assumption that the total amount of RNA labelled with either Cy3 or Cy5 is equal. This form of global normalisation is a useful tool in instances of closely related samples where the transcription level of many genes will remain unchanged. While the intensity for one spot may be higher in one channel than the other, when averaged over thousands of spots in the array, these differences should average out. Therefore, the total fluorescence across all the spots in the array should be equal for both channels and a scatter plot of the measured Cy5 versus Cy3 intensities should have a slope of one. Using this approach, a normalisation factor (N_{total}) is calculated by summing up the measured intensities in both channels;

$$N_{total} = \frac{\sum_{i=1}^{N_{array}} R_i}{\sum_{i=1}^{N_{array}} G_i}$$

where R_i and G_i are the measured intensities (for the red and green colours commonly used to represent array data) of the *i*th array element (gene) on an array that has from 1 to N_{array} distinct elements. The ratio for the *i*th gene (T_i) is then divided by this normalisation factor and becomes the normalised expression ratio;

$$T_i = \frac{R_i}{G_i} \frac{1}{N_{total}}$$

thus adjusting each ratio so that the mean ratio is equal to one.

Chen *et al.* (1997) have developed a normalisation method that makes use of a specific subset of genes as opposed to all elements on the slide. This method is useful for more divergent samples where the total intensity normalisation model would give a poorer estimation of normalisation. This method assumes that the distribution of transcription levels for a set of housekeeping genes that are expected to be unchanging have a mean value and standard deviation that are independent of the sample. In this case, the ratio of measured Cy5 to Cy3 ratios for these genes is used as a model to adjust the mean of ratios to one.

Locally weighted linear regression (LOWESS) analysis is a non-linear intensity dependent normalization procedure that is suited to two colour experiments where there are more than 100 elements on the chip (Cleveland, 1979 and Yang *et al.*, 2002b). LOWESS normalisation is a technique that is used to eliminate dye-related artefacts in two-colour experiments that cause the Cy5/Cy3 ratio to be affected by the total intensity of the spot. The artefacts that LOWESS attempts to correct for include non-linear rates of dye incorporation as well as inconsistencies in the relative fluorescence intensity between the two dyes used.

4.1.1.7 Data Mining: the identification of differentially expressed genes

The extraction of biologically relevant information from large array generated gene lists still proves to be a serious bottleneck (Butte, 2002). Regardless of the nature of the experiment, the major interest in performing microarray analysis is in the identification of genes that are differentially expressed between samples in the data set. Therefore simply generating the data is not enough, one must be able to extract from it meaningful information about the system being studied. This currently requires the combined efforts of biologists, bioinformaticians, computer scientists, statisticians and software engineers.

A general two-fold cut off is often used to identify those genes that are most variable between samples (Yang *et al.*, 2002a, Schena *et al.*, 1995 and Schena *et al.*, 1996). Therefore, genes are taken to be differentially expressed if their expression under one condition is over two-fold greater or less than that under the other condition. The 'fold' change test is not a statistical one and does not provide any information on the level of confidence in the designation of genes as differentially expressed or not differentially expressed.

In order to distinguish genes that are significantly differentially expressed from random changes, independent microarray assays are often conducted starting from independent mRNA isolations and used to define differential expression based on their statistical consensus. As discussed in section 4.1.1.2, for statistical purposes, biological replicates - RNA samples obtained from independent biological sources - are preferable to technical replicates - RNA samples from the one biological source, especially if conclusions are to be made regarding the significance of expression changes. Ideally, each biological condition for comparison should be represented by at least three independent biological samples (Lee *et al.*, 2000). Obviously, this depends on the expense of the microarray experiment and the availability of samples and therefore RNA.

The t-test is one of many statistical methods suitable for detecting differentially expressed genes (Cui and Churchill, 2003). The t-test uses the error variance for a given gene over replicated experiments to determine if that gene is differentially expressed and if this difference is significant or not (Callow *et al.*, 2000). A t-value for a given gene is computed as follows;

$$t = \frac{R_g}{SE_g}$$

where R_g is the mean ratio of expression levels for that gene and SE_g is its standard error. After a t-value is calculated it is converted to a p-value. Genes with p-values falling below a prescribed 'nominal' level may be regarded as significant.

Several problems exist with this 2-fold approach to identifying significantly differentially expressed genes as genes of importance in the biology under investigation. The first is that not only changing genes but also the genes which are unchanged from one state to another (i.e. from normal to tumour or from primary to metastatic) may understandably have an important role in both states. The rudimentary analysis of picking changing genes at the top of a list may draw attention away from the biological importance of maintaining transcription of a gene from normal to disease state. Also real biological data might be lost when alterations in gene expression are more modest than two-fold. To address this problem Mootha *et al* (2003) have devised a powerful analytical approach termed gene set enrichment analysis (GSEA). GSEA studies the behaviour of genes as a group termed a 'gene-set'. Biologically associated genes or genes involved in a similar cellular pathway or function are unified into gene-sets. The hypothesis being that subtle changes in these genes can have a profound effect on the biological pathway and when added together the relatively small changes in the genes of a 'gene-set' can reach extraordinary significance. This approach of data analysis aims to capture the real biology underlining the distinction of interest as apposed to subjectively cherry picking top ranking genes and creating hypotheses about pathway membership.

4.1.1.8 Post-analysis follow-up: validation of microarray data

Post-array verification of results using an independent laboratory approach provides experimental verification of gene-expression levels and ideally begins with the same biological samples that were compared in the array experiment. Common mRNA and protein methods used to validate array data include RT-PCR (Rajeevan *et al.*, 2001), northern analysis (Chaib *et al.*, 2001), ribonuclease protection assay (Taniguchi *et al.*, 2001), in *situ* hybridisation (Mousses *et al.*, 2002), immunoblot analysis (Al Moustafa *et al.*, 2002) and immunohistochemistry (Mousses *et al.*, 2002).

Before genes of interest are pursued for further investigation and validation, sequencing is required as a quality control measure. As many array experiments are carried out on the genome-wide scale the probability of error associated with spot identity and printing are also increasing. Array production is a high-throughput operation involving multiple steps of sample handling and processing, and is therefore

understandably subject to error. Consequently, sequence verification is required to identify and correct for such errors (Taylor *et al.*, 2001).

For post-array validation, real time RT-PCR is the most popular method for quantitatively measuring specific mRNA transcripts as it is reproducible, rapid, inexpensive and requires little starting template (Giulietti *et al.*, 2001 and Walker, 2002). However, Rajeevan *et al* (2001) have shown that for many genes studied there were significant quantitative differences between array and RT-based data.

As well as validating array results at the mRNA level it is also worth investigating the changes in expression at the level of the corresponding protein (Chuaqui *et al.*, 2002). The observed changes at the mRNA level may not always translate to the protein. This discrepancy may be a factor of either the laboratory test used to determine protein expression or of the biology of the protein, as protein function in the cell is affected by several factors besides abundance.

4.1.2 The application of microarray analysis in cancer studies

Cancer is a genetic disease and as a powerful genomic tool, microarray analysis holds great promise in the molecular medicine of cancer research. This is because cancer is a complex polygeneic and multifactorial disease, resulting from successive changes in the genome of cells and from the accumulation of molecular alterations in both tumour and host cells (Hanahan and Weinberg, 2000). Such genetic alterations include expression suppression or enhancement as well as deletions, mutations, insertions and rearrangements of genes controlling regulatory pathways and cellular processes such as proliferation, differentiation, cell cycle, DNA repair and apoptosis. These genetic changes lead to genetic instability, tumourigenesis, malignancy, and an invasive and drug-resistant phenotype. Therefore, an understanding of the molecular behaviour of tumours would aid their molecular classification and the therapeutic decisions made (Khan *et al.*, 1998a). New tools are required to predict the clinical behaviour and outcome of individual tumours and to ideally prescribe individual tailored treatments.

DNA microarray technology has the potential to identify novel genes that have key roles in mediating either sensitivity or resistance to cancer drugs and

chemotherapeutic agents. Such analysis of drug-responsive genes will greatly assist the identification of new biomarkers, novel therapeutic targets and the development of logical and effective drug combinations. Several studies have investigated the use of DNA microarrays to predict the response of cancer cells to chemotherapies (Scherf *et al.*, 2000, Zembutsu *et al.*, 2002 and Blower *et al.*, 2002). Such studies involve correlating gene expression with drug activity, thus providing the ability to group or cluster drugs based on their activity. Microarray-mediated identification of the downstream signalling pathways involved in tumour-cell response to drugs will be a key step in the future development of drug cytotoxicity. Most current cancer medications are relatively non-specific cytotoxic agents that exert their effect on both normal and tumour cells. New selective agents are therefore needed to further improve cure rates and reduce host toxicity. Stegmaier *et al.* (2004) have devised a novel approach for screening candidate small molecules for their ability to differentiate acute myeloid leukaemia (AML) cells. This method reduces the genes under scrutiny to a small subset, and assays the ability of the drugs to induce a characteristic signature of the transition from AML cells to their differentiated myeloid counterparts. Successful candidates can then be investigated on a genome wide scale, saving both time and money.

Microarray analysis has been used to investigate the underlying biology of many cancers including AML (Virtaneva *et al.*, 2001) breast (Sgroi *et al.*, 1999), ovarian (Ono *et al.*, 2000 and Welsh *et al.*, 2001), lung (Anbazhagan *et al.*, 1999), and colon cancer (Alon *et al.*, 1999). Array analysis is applicable to any area or stage (e.g. initiation, progression, drug-resistance etc) of cancer biology where it is used to elucidate the underlying molecular mechanisms (Bertucci *et al.*, 2001). A better understanding of the disease at a molecular level would therefore allow the design of therapies targeted to the specific molecular cause of the disease. Array technology provides an ideal tool for the investigation of metastasis as well as all areas of cancer, as the activity of many genes can be simultaneously monitored, thus giving an insight into the complexity of cancer genetics. The control of metastasis is an appealing target in cancer treatment as it is the main cause of mortality in cancer and it is commonly the reason for the failure of chemotherapy to eradicate cancer. The metastatic process has been described as an obstacle course in which rare cancer cells with appropriate combinations of attributes survive to form a metastatic colony (Poste and Fidler, 1980 and Fidler and Kripke, 1977). The challenge therefore, is to discover the minimal set of genes that are

functionally necessary for metastasis. Consequently, several studies using microarrays have focused on metastasis by comparing the expression profiles of highly metastatic cells with less or non-metastatic paired cells (Clark *et al.*, 2000, Hegde *et al.*, 2001, Bittner *et al.*, 2000 and Maniotis *et al.*, 1999). By using paired or genetically related cell lines, it can be presumed that the expression of the majority of genes is shared with the exception of those promoting or related to metastasis. This application of microarray technology is ideal for the identification of biomarkers of metastasis. Using array analysis, Weigelt *et al.* (2003) have shown that breast metastases often maintain very similar expression profiles to those of the primary breast tumours from the same patient. This study also showed an inability to establish a pattern in genetic expression changes between the primary tumours and metastases. This stands in contrast with the customary model where only a sub-set of primary tumour cells gain metastatic ability and that this is only a rare event in cancer progression (Fidler and Hart, 1982). These observations also differ from a recent study using multiple tumour types, which found a gene-expression signature that distinguished primary from metastatic adenocarcinomas (Ramaswamy *et al.*, 2003). However, the tissues used in the Ramaswamy study were not paired which might explain the differences nevertheless these gene-expression signatures may represent a composite of multiple, tissue-specific metastasis programs.

4.1.3 The future of microarrays

The applications of microarrays to all areas of investigative biology are endless. Microarrays have become an accessible and common oncology research tool that will have a great impact in the evolution of the molecular medicine of cancer. More and more cancers will be classified into new previously unrecognised subclasses based on their molecular signatures thus facilitating tailored treatments (Khan *et al.*, 1998a, Scherf *et al.*, 2000, Golub *et al.*, 1999 and Su *et al.*, 2001). For a young technique, microarray analysis is already a key one to the investigation of the complexity of cancer.

As well as having a great impact on the development of novel therapeutics, the recent completion of the human genome sequence will also spur a desire to carry out microarray analysis on a whole genome level (The Genome International Sequencing Consortium, 2001 and Austin, 2004). Such genome-wide analysis is physically hampered by the problems of miniaturising arrays, such as cost, data acquisition and

analysis. Modern methods of printing arrays aim to reduce the size of arrays. By reducing the size of the spots on a chip from 150 – 100 μm to 30 – 25 μm, an entire 100,000 spot array could fit into the a 4 cm^2 area and such miniaturisation may be commonplace in future array analysis (Okamoto *et al.*, 2000). However, the current approach of users selecting and printing sets of genes to produce customised arrays is likely to remain attractive.

Further advances and improvements in array analysis will concern the care of the biological samples under investigation. In particular, and of clinical relevance, is the care of tissue samples, which is important in avoiding RNA degradation. Stricter rules are needed for the collection, storage and processing of tissues. Another major issue in need of improvement is the purity of tissue samples. Most biopsies are heterogeneous, containing many different cell types including normal surrounding cells such as epithelial, endothelial, adipose and stromal cells, infiltrating lymphocytes as well as tumour cells (Ronnov-Jessen *et al.*, 1996). A pure homogenous sample is desirable in order to obtain representative results. One solution to obtaining a pure sample is to use laser capture microdissection (LCM). LCM is used to harvest subpopulations of cells, however, obtaining sufficient RNA from homogenous material for microarray analysis usually requires an amplification step (Emmert-Buck *et al.*, 1996). Although it introduces another variable into the process, RNA amplification allows array analysis to be carried out on 1 – 50 ng of mRNA (Wang *et al.*, 2000). Currently, the isolation of tumour cells from surrounding cells is not critical with many clinical studies on crude samples being published (Bertucci *et al.*, 2000, Perou *et al.*, 2000 and Bittner *et al.*, 2000). One approach to estimate the contribution of each cell type is through parallel analysis of the expression profiles of cell lines representing the different cell types (Perou *et al.*, 1999).

The future of arrays and the interpretation of the vast amount of data generated revolve around the synergy between different scientific cultures. Multidisciplinary collaborations between medical researchers, mathematicians, computer scientists and bioinformaticians are required to carry studies forward to the clinic. A synergy is also required between different types of information such as the transcriptional and translational data for the same samples. Such data should be comparable from one laboratory to another, thus providing a platform for connectivity.

4.2 Results: analysis of gene expression profiles in colon cancer cell lines

The expression profiles of the following colorectal cancer cell lines were compared by cDNA microarray analysis; SW480, SW620, SW480M7 and SW480M9. All array analysis was carried out at GMU, VA, USA. Together, the SW480 and SW620 cell lines represented a validated model of colorectal metastasis (Hewitt *et al.*, 2000). Both cell lines were derived from the same patient at different stages of tumour progression. The SW480 cell line was cultured from a primary lesion (Dukes stage B colon carcinoma) from a 50-year-old Caucasian male patient and the SW620 cell line was derived from a lymph node metastasis in the same patient a year later. The comparison of the expression profiles of these isogenic cell lines therefore provided an insight into tumour metastasis as the differences in expression most likely represented changes that occurred when the cells acquired metastatic potential. The increased invasion shown by the MMP transfected cell lines SW480M7 and SW480M9 was investigated by array analysis with the aim of determining the downstream effects of MMP-7 and MMP-9 overexpression and to identify MMP-specific transcriptional
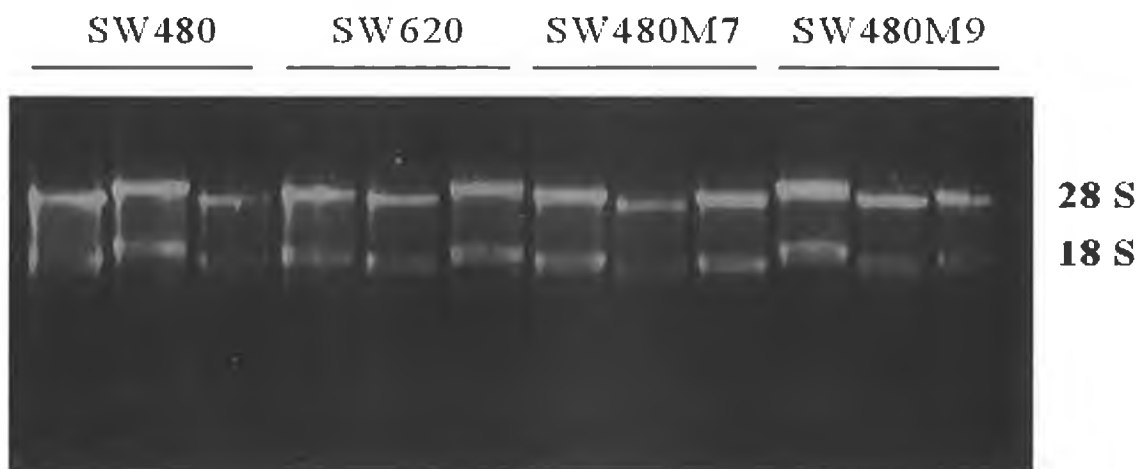
events.

Figure 4.3 The integrity of RNA used in array analysis. It was crucial that the starting RNA for analysis was not degraded. The prominent bands of the 28 S and 18 S ribosomal subunits indicated intact and good quality RNA. All RNA was prepared in triplicate.

4.2.1 RNA extraction, analysis and labelling

Prior to labelling and hybridisation (sections 2.2.16.4 and 2.2.14.6 respectively), RNA was DNase treated and its quality and quantity were then assessed by gel

electrophoresis and by spectroscopy respectively (section 2.2.2.1). It can be seen in Figure 4.3 that all RNA samples were intact and not degraded. As the 39,360 element experiment was a two chip one, 8 µg of DNase treated RNA was required as opposed to half that which would normally be required for a one chip experiment (section 2.2.16.4.2). RNA was therefore amplified to provide sufficient amounts for all array experiments including replicates (section 2.2.16.3). The initial concentrations of RNA samples plus the concentrations of DNase treated and amplified RNA are listed in table 4.1.

Table 4.1 Effects of DNase treatment and amplification on RNA concentration and A260/280 values.

RNA Sample	A260/280	Yield (µg in 25 µl)	Yield: µg DNase treated RNA in 25 µl	Yield: µg aRNA in 100µl	A260/280
SW480(1)	1.5	123.9	38.5	73	2.0
SW480(2)	1.7	87.9	25.8	67	1.6
SW480(3)	1.5	50.7	17.8	59	1.8
SW620(1)	1.5	95.9	37.5	46	1.4
SW620(2)	1.6	44.6	27.5	72	1.8
SW620(3)	1.4	121.2	38.5	106	1.6
SW480M7(1)	1.3	121.6	50.5	90	1.6
SW480M7(2)	1.6	90.8	34.8	67	1.7
SW480M7(3)	1.5	110.2	36.3	104	1.7
SW480M9(1)	1.5	132.8	39.0	91	1.6
SW480M9(2)	1.5	51.5	23.8	102	1.5
SW480M9(3)	1.5	96.8	35.5	84	1.6

4.2.2 Experimental design

Table 4.2 outlines the microarray experiments carried out. In all array experiments carried out the SW480 cell line was used as a control. Where possible, experiments were carried out in triplicate to ensure statistically significant expression changes could be identified. Experiment 2 represents a technical replicate of experiment 1, in which the same RNA preparations from SW480 and SW620 cells were compared in both experiments. Experiment 3 and 4 represent biological replicates, where independent RNA preparations of SW480 and SW620 were used in each experiment. Gene expression in two separate SW480 preparations were compared with

the same SW480M7 preparation in experiments 5 and 6. Experiment 7 represents a dye switch control for experiment 6. To investigate dye bias and uneven incorporation the SW480 control normally labelled with Cy3 was labelled with Cy5, where the SW480M7 was labelled with Cy3.

Table 4.2 Experimental design for the comparison of gene expression in colorectal cancer cell lines by cDNA microarray analysis. Each numbered line represents an individual experiment where one sample was labelled with Cy3 and one with Cy5. Numbers in brackets indicate the original RNA preparation batch. Colour shading indicates the aRNA amplification batch. Dotted borders indicate samples that were pooled before hybridisations. Experiment 7 and 10 represent switch dye controls for experiments six and nine respectively. Experiments 11 and 12 were 'self' vs. 'self' controls where experiment 12 compares the expression of two separate aRNA amplifications.

Experiment	Cy3	Cy5
1	SW480(1)	SW620(1)
2	SW480(1)	SW620(1)
3	SW480(2)	SW620(2)
4	SW480(3)	SW620(3)
5	SW480(2)	SW480M7(2)
6	SW480(2)	SW480M7(2)
7	SW480M7(2)	SW480(2)
8	SW480(2)	SW480M9(2)
9	SW480(2)	SW480M9(2)
10	SW480M9(2)	SW480(2)
11	SW480(2)	SW480(2)
12	SW480(2)	SW480(2)

The expression of SW480 and SW480M9 cells were compared in experiments 8, 9 and 10. Separate SW480 RNA preparations were compared with the same SW480M9

preparation in experiments 8 and 9. Experiment 10 acted as a dye switch control for experiment 9. The 'self' vs. 'self' control experiments, 11 and 12, were carried out in order to control the number of false positive that the technique and data analysis generated. Experiment 12 is a 'self' vs. 'self' comparison of two separate aRNA batches.

4.2.3 Hybridisation and scanning

Labelled cDNAs were hybridised and scanned according to TIGR protocols. The arrays used in this experiment were fabricated as outlined in section 2.2.16.1. Briefly, 39,360 elements were spotted over two slides and spotted slides were stored at room temperature under vacuum until further use. Slides were hybridised as outlined in section 2.2.16.6, this optimised TIGR protocol briefly involved allowing the labelled cDNAs to hybridise to the slide under humid conditions at 42 °C for 15 hours. Following washing, hybridised slides were then scanned using a confocal laser scanner and the generated TIFF images were interpreted using Quantarray software as outlined in section 2.2.16.7.

4.2.4 Data analysis

Following scanning, the raw data generated was analysed using Genespring 6.0 (Silicon Genetics) software, a powerful commercially available expression data visualization and analysis tool. The first step in the analysis of gene array data with this software was to install a gene key, a TXT file that describes the location of genes on the slide as well as a gene identifier (gene name or annotation number). The gene key used can be found at the web address: http://www.dcu.ie/~biotech/dmurray/COMBINED-GENE-KEY.xls, June 25, 2004. Data was then imported and replicate and dye swap experiments were identified. Certain data transformations were then carried out. The data was normalised using the LOWESS per-spot method, which compensates for non-linear dye-bias. Genespring 6.0 software was used as a data analysis tool, to simply identify genes with altered expression from one state to another. To carry the resulting lists of genes forward to the subsequent validation steps, detailed literature searches were required. This section describes the analysis of the data obtained from the array experiments.

4.2.4.1 Data Analysis: comparison of SW480 and SW620 gene expression profiles

In all the data analysis carried out, the colour bar in Figure 4.4 was used to indicate the fold changes in expression. Figure 4.5 is a scatter plot of the normalised data for the SW480 Vs SW620 experiment. This is a plot of the data obtained from all replicates in this experiment including the dye switch control. Each data point or spot represents a gene and the colour of that spot represents its expression ratio.
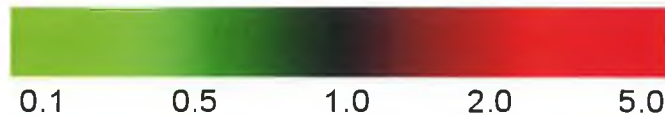
Figure 4.4 Colour bar used to display fold changes in expression. Data points with red colouring indicated upregulation and data points with green colouring indicated downregulation.

Genes that are black in colour and generally fall on the 45° line in the scatter plot are unchanged in expression from the SW480 to the SW620. Genes that are green and also generally fall below the lower 45° line have 2 fold more expression in the control SW480 cells than the SW620 cells. These genes are therefore down regulated at least two fold in the progression to metastasis.

X-axis: SW480 Vs SW620 (Apr 04) (Default Interpretation), SW480 V SW620   Colored by: SW480 Vs SW620 (Apr 04) (Default Interpretation)
Y-axis: SW480 Vs SW620 (Apr 04) (Default Interpretation), SW480 V SW620   Gene List: all genes (39933)
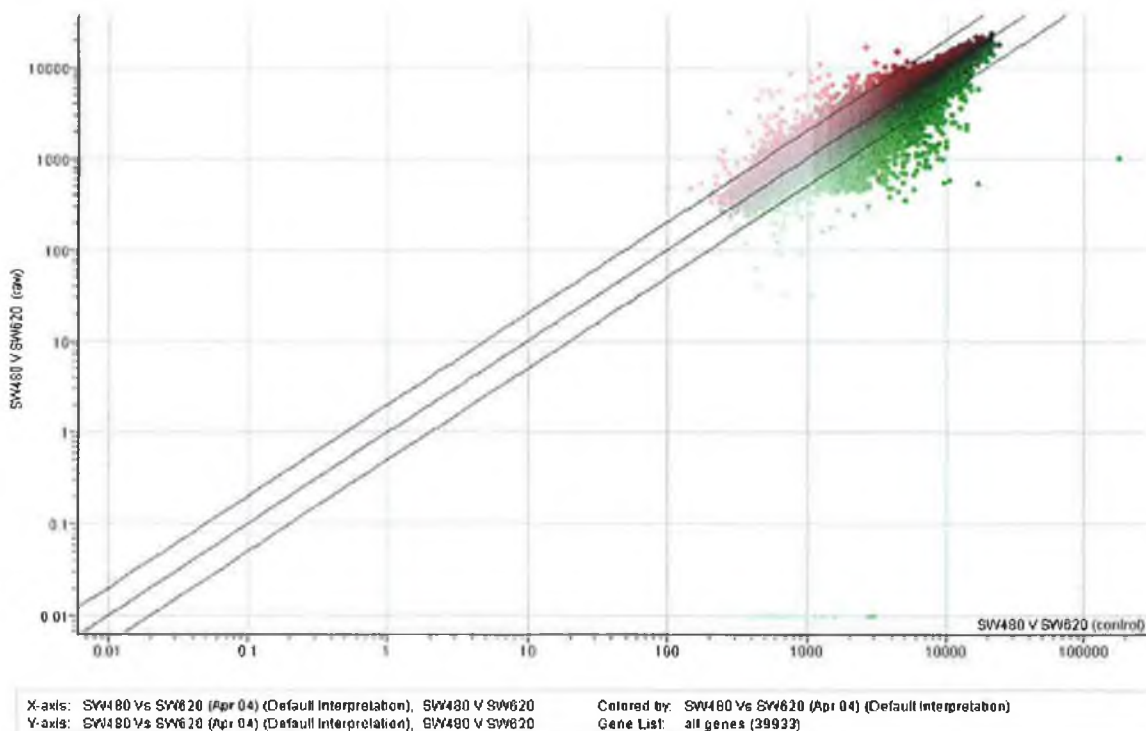
Figure 4.5 Scatter plot of normalised expression data for the comparison of SW480 with SW620 cells. This plot represents composite data from all replicates.

Genes that are red in colour and fall above the upper 45° line have a minimum of 2-fold increase in expression in the metastatic SW620 cells in comparison to the non-metastatic SW480 cells. It must also be noted that the darker the red or green colouring, the greater the fold difference, also, more trust is associated with the spots that are intensely coloured. These data points and therefore expression ratios are trustworthy because there is a consensus and therefore little variance for that ratio between replicates. Data points of little or unsaturated intensity have less trust associated with them and this is usually because data is not achieved in both channels for all replicates. Trust can also be lost if the ratios are not similar between replicates. This loss in trust can be due to bad hybridisation, dust on the array, high background, bad dye incorporation or uneven labelling. Considering this, real biological data can be lost as a result of one 'bad' experiment.

Nevertheless, the next step was to investigate the genes that were changing in expression. Figure 4.6 shows the selection of genes (black in colour) with expression in the SW620 at least two fold more than that in the SW480 cells.

X-axis: SW480 Vs SW620 (Apr 04) (Default Interpretation), SW480 V SW620   Colored by SW480 Vs SW620 (Apr 04) (Default Interpretation)
Y-axis: SW480 Vs SW620 (Apr 04) (Default Interpretation), SW480 V SW620   Gene List: all genes (39933), 418 genes selected
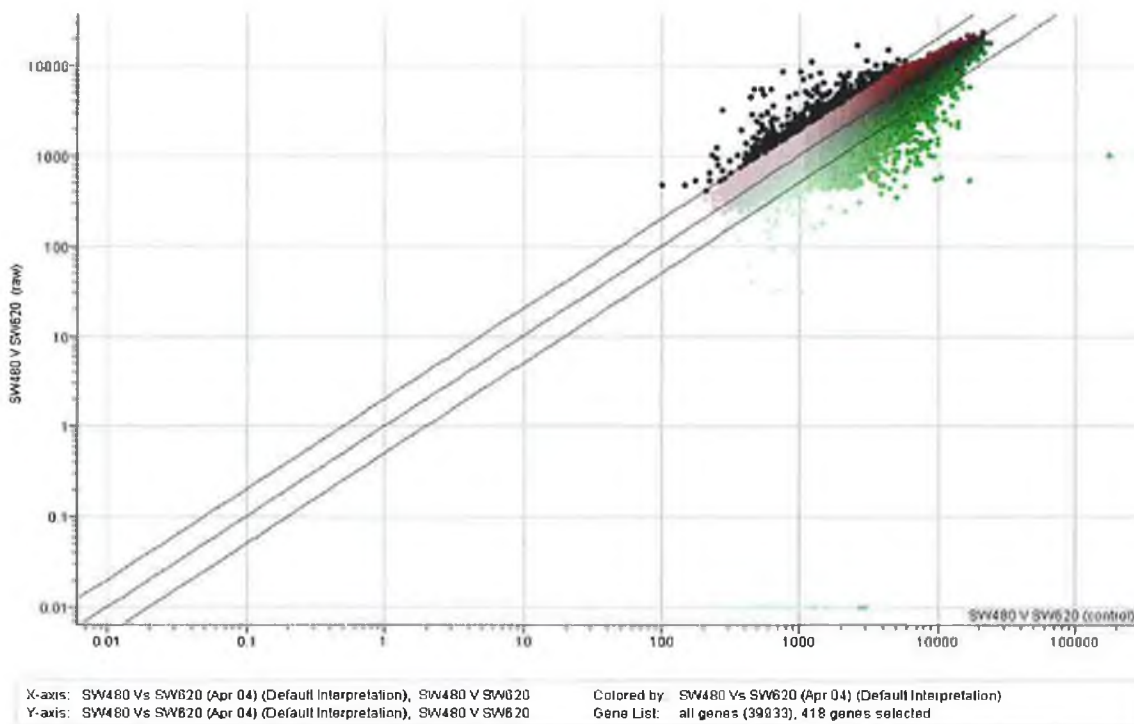
Figure 4.6 Selection of genes that are at least two fold up regulated in the SW620 cells when compared with the SW480 cells. Selected genes are coloured black.

It can be seen from Figure 4.6 that 418 genes were selected as having an increase in expression of at least 2-fold. These genes could be visualised as blocks, as demonstrated in Figure 4.7.

Colored by   SW480 Vs SW620 (Apr 04) (Default Interpretation)
Gene List   2 fold up (418)
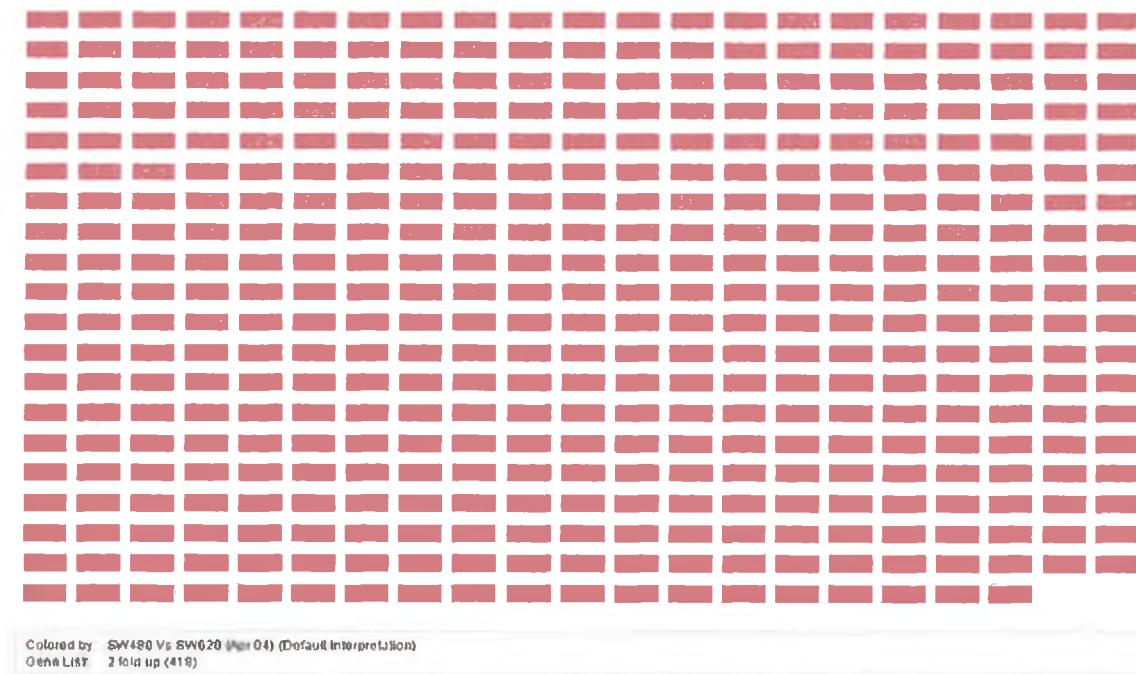
Figure 4.7 Block representation of the 418 genes with at least two-fold increase in expression in the SW620 cells when compared with SW480 cells.

This visualisation gives a good indication of the trust of these genes with the more trustworthy, intensely coloured genes towards the bottom. Genes represented by blocks towards the bottom of figure 4.7 are darker in colour indicating a high level of trust associated with them. The expression ratios for these genes are therefore similar over replicates. Blocks with unsaturated colouring represent genes with low levels of trust associated with their expression changes. The expression ratios that represented significant changes in expression were selected next. For this a t-test was performed on all of the 418 genes. Expression ratios with a p-value lower than 0.005 were regarded as significant and are represented in Figure 4.8.

As can be seen from Figure 4.8, 233 out of 418 genes were significantly upregulated at least 2-fold in the SW620 cells in comparison with the SW480 cells. Because this experiment was carried out in triplicate it can be concluded that these 233 genes have statistically significant changes of at least a two fold increase in expression in the SW620 cell line in comparison with their expression in the parental SW480 cell line.

Figure 4.8 Block representation of genes with 2-fold or more expression in SW620 in comparison with SW480 cells. (p<0.005)

A selection of these 233 genes is listed in table 4.3, and a full list can be found in appendix A of this thesis.

Table 4.3 Selection of genes with at least 2-fold increased expression in SW620 cells in comparison to SW480 cells (p<0.005)

Gene Number	Fold Change	Gene Identity
AA916325	21.624	aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II)
AA608575	17.343	propionyl Coenzyme A carboxylase, alpha polypeptide
AA135152	11.576	glutathione peroxidase 2 (gastrointestinal)
R37743	5.089	T54 protein
R39239	4.897	tenascin C (hexabrachion)
R16134	4.069	transmembrane 4 superfamily member 11 (plasmolipin)
N70463	3.958	B-cell translocation gene 1, anti-proliferative
AA088420	3.588	peroxisome proliferative activated receptor, gamma
H79534	3.549	hemoglobin, epsilon 1
W46900	3.436	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
AA779165	3.421	ADP-ribosylation factor-like 4
AA460463	3.416	cytokine-like protein C17
R60170	3.167	guanine deaminase
R98936	3.092	membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10)
AA478585	2.776	butyrophilin, subfamily 3, member A3
AA055835	2.725	caveolin 1, caveolae protein, 22kDa
AA644088	2.676	cathepsin C
AA486220	2.670	lysyl-tRNA synthetase

Also of interest in this experiment were the genes that were downregulated in the metastatic SW620 cells in comparison with the SW480 cells. Therefore, all the above post-array analysis was carried out on the genes that were at least 2 fold downregulated, i.e. selection of genes and a t-test to determine significance. The genes that were significantly at least two-fold downregulated in comparison with the SW480 cells are represented as blocks in Figure 4.9. Genes that are two fold down regulated are represented as green blocks and the 'more green' these blocks are the more downregulated these genes are in the SW620 cells. The more saturated in colour these blocks are also indicated more trust associated with their expression changes.
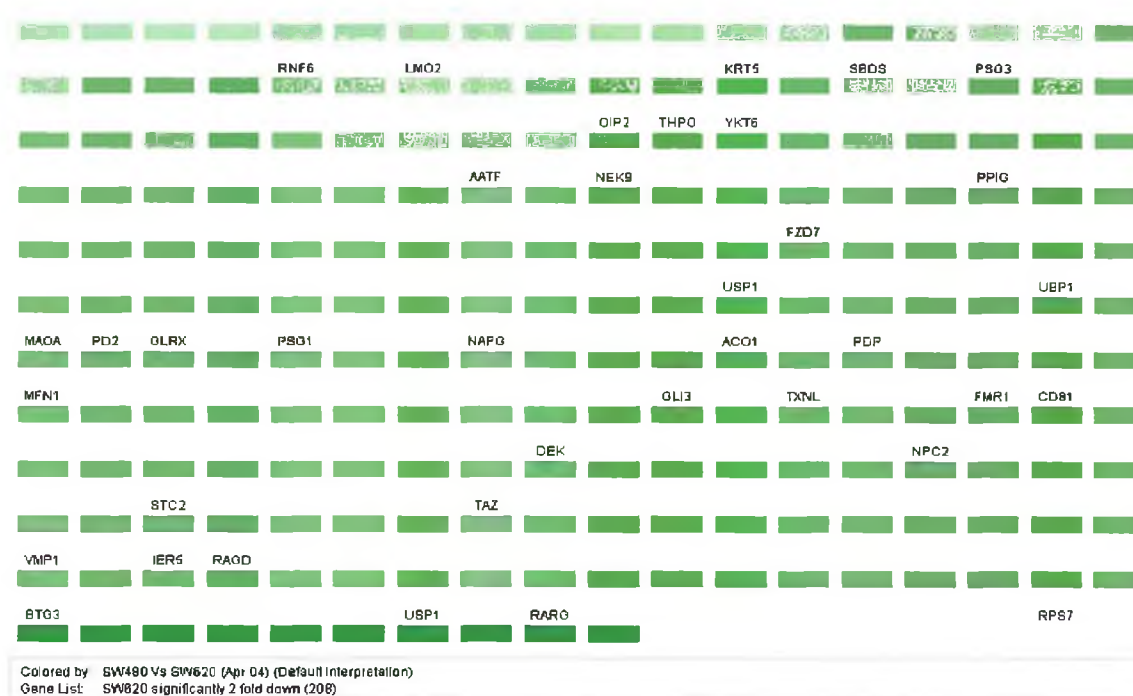
Figure 4.9 Block representation of genes down regulated in the SW620 cells in comparison with the SW480 cells (p<0.005).

It can be seen from figure 4.9 that the down regulation of 208 genes in SW620 cells was significant. Also visible in figure 4.9 next to each block is the abbreviation for the gene that this block represented. These abbreviations were not visible in figure 4.8 due to the larger number of genes. A selection of these genes can be found in table 4.4 and the full list can be found in appendix A of this thesis.

Table 4.4 Selection of genes with expression ratios below 0.5 for the comparison of SW620 with SW480 cells. (p<0.005)

Gene Number	Fold Change	Gene Identity
AA598601	0.148	insulin-like growth factor binding protein 3
AA406552	0.165	solute carrier family 2 (facilitated glucose transporter), member 3
AA487425	0.331	zinc finger protein 463
AA621315	0.344	catenin (cadherin-associated protein), alpha-like 1
AA291163	0.351	glutaredoxin (thioltransferase)
AA486556	0.383	CD81 antigen (target of antiproliferative antibody 1)
AI439571	0.384	apoptosis antagonizing transcription factor
AA683077	0.400	mitogen-activated protein kinase 1
N67766	0.406	acetyl-Coenzyme A synthetase 2 (AMP forming)-like
AA017526	0.421	collagen, type IX, alpha 3
AA487486	0.433	cyclin D1 (PRAD1: parathyroid adenomatosis 1)
AA827551	0.445	Notch homolog 2 (Drosophila)
AI688757	0.464	cytochrome c oxidase subunit Vb

4.2.4.2 Data analysis: expression profiling of MMP transfected cell lines

To investigate the downstream effects of individually overexpressing MMP-7 and MMP-9 on the expression profile of a colon cancer cell line, array analysis was used to compare the SW480M7 and SW480M9 cells with the parental SW480 cells. By studying the global expression of these cells, cDNA array technology provided an ideal tool in determining the transcriptional effects of MMP activity

4.2.4.2.1 Data analysis: profiling SW480M7 expression

Figure 4.10 is the scatter plot of normalised data comparing the expression profile of the MMP-7 transfected cell line (SW480M7) with that of the non-transfected cell line SW480. Again, the expression ratio of each gene is represented by a datapoint and the colour of that data point represents the nature of that ratio. Green data points represent genes that are downregulated in cells transfected with MMP-7 and red data points represent genes with up regulated expression in the same cells. Data points that are black in colour indicate that these genes are unchanged from one condition to another. Because altered gene expression resulting from MMP-7 overexpression was of interest, genes with at least two-fold increase or decrease in expression were selected for further statistical analysis. Genes with at least 2-fold enhanced expression in the MMP-7 transfected cell line SW480M7 in comparison with the SW480 control are depicted in

Figure 4.11. The changes in expression of these 147 genes reached significance with p values smaller than 0.005.

X-axis: SW480 Vs SW480M7 (Default Interpretation), SW480 Vs SW480M7   Colored by: SW480 Vs SW480M7 (Default Interpretation)
Y axis: SW480 Vs SW480M7 (Default Interpretation), SW480 Vs SW480M7   Gene List: all genes (39933)
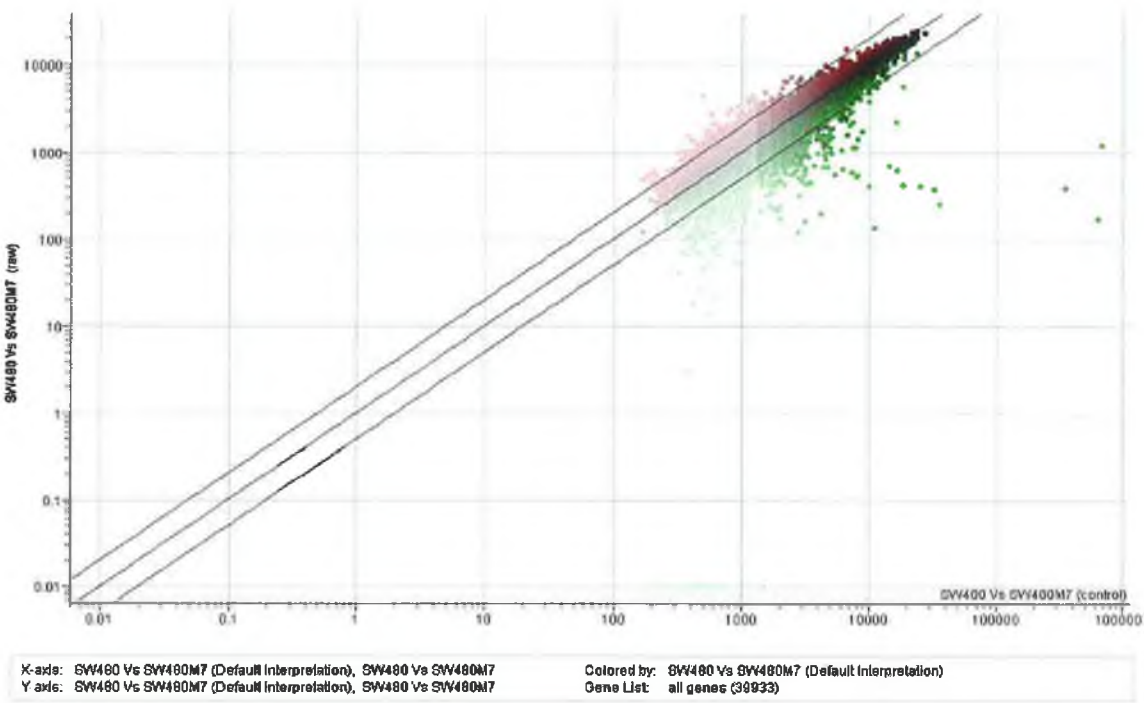
Figure 4.10 Scatter plot of normalised data for the comparison of SW480M7 cells with the control SW480 cells. This plot represents data obtained from all replicates of this experiment including the dye switch control.

Colored by: SW480 Vs SW480M7 (Default Interpretation)
Gene List: M7 significantly 2 fold up (147)
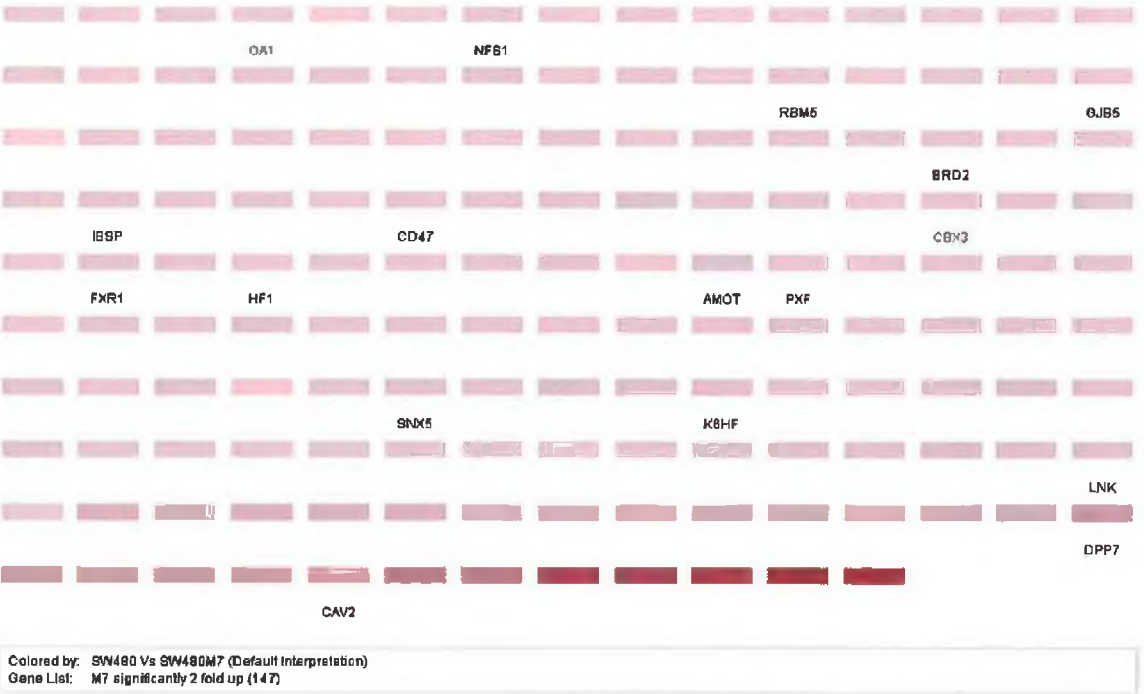
Figure 4.11 Significantly up regulated genes in a cell line over expressing MMP-7.

These 147 genes are listed in appendix A of this thesis, but a selection of them are listed in table 4.5.

Table 4.5 Genes with at least 2 fold increase in expression in cells transfected with MMP-7 (p<0.005). The full list can be found in appendix A of this thesis.

Gene Number	Fold Change	Gene Identity
AI339434	2.750	caveolin 2
AI250799	2.676	defensin, alpha 4, corticostatin
AI276487	2.617	cytokeratin type II
AA669218	2.540	HBS1-like (S. cerevisiae)
AA916562	2.494	vacuolar protein sorting 4B (yeast)
H21892	2.492	Homo sapiens cDNA FLJ36638 fis, clone TRACH2018950.
AI288845	2.421	chemokine (C-C motif) receptor-like 2
AA479344	2.401	gap junction protein, beta 5 (connexin 31.1)
AA463462	2.384	Toll-interacting protein
AA778851	2.347	Human clone 137308 mRNA, partial cds.
AA953249	2.323	H factor 1 (complement)
AA977836	2.310	peroxisomal farnesylated protein
AA410434	2.218	collagen triple helix repeat containing 1
AI198232	2.132	C-terminal PDZ domain ligand of neuronal nitric oxide synthase
AA868008	2.042	histone 1, H4c
AA424562	2.017	tumor necrosis factor (ligand) superfamily, member 13
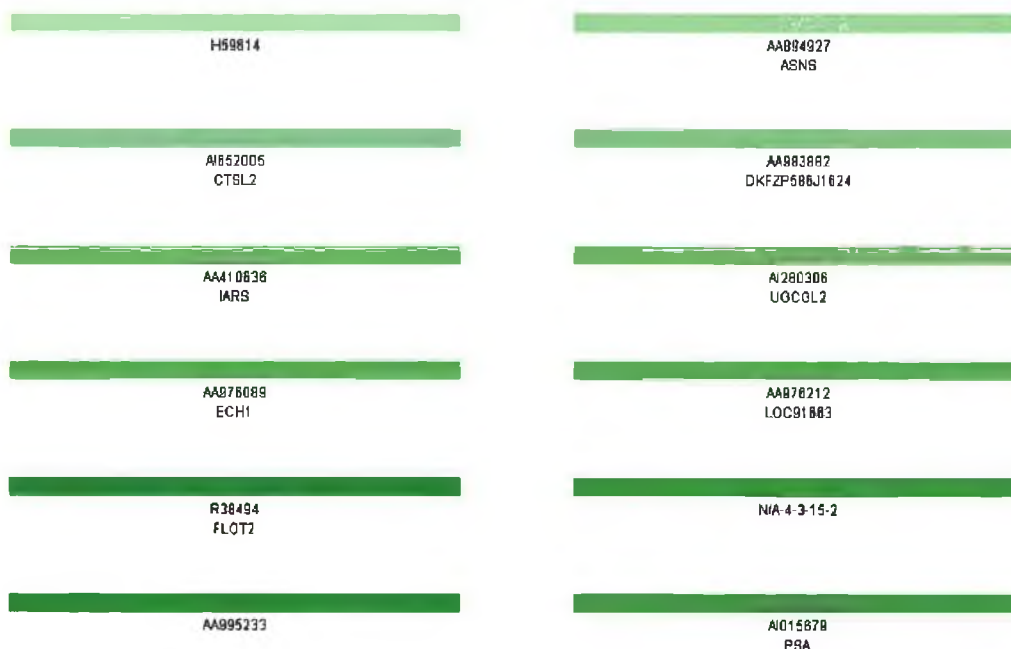
Colored by: SW480 Vs SW480M7 (Default Interpretation)
Gene List: M7 significantly 2 fold down (12)

Figure 4.12 Block representation of genes that were at least 2-fold downregulated in SW480M7 cells in comparison with SW480 cells.

Figure 4.12 represents the genes that were significantly downregulated in the SW480M7 cells in comparison with the parental SW480. When tested for significance only 12 genes were statistically downregulated due to MMP-7 overexpression. Listed in Table 4.6 is a selection of these genes as well as their fold change in expression.

Table 4.6 Selection of genes downregulated in cells overexpressing MMP-7 (p<0.005).

Accession Number	Fold Change	Gene Identity
AI015679	0.304	phosphoserine aminotransferase
AI280306	0.341	UDP-glucose ceramide glucosyltransferase-like 2
AA894927	0.358	asparagine synthetase
AA410636	0.487	isoleucine-tRNA synthetase
R38494	0.493	flotillin 2
AI652005	0.515	cathepsin L2

As with the SW620 experiment these genes were listed for further investigation and for literature searches.

4.2.4.2.2 Analysis of SW480M9 expression data

The data obtained from the array experiments in which the expression profile of the SW480M9 cells, which overexpress MMP-9, was compared with that of the parental, non-transfected SW480 cells was analysed for genes that were significantly changing. Figure 4.13 is a scatter plot representing the normalised data for the expression ratios for all genes in this experiment. Data from all replicate experiments including a dye switch control were considered in this analysis and Figure 4.14 represents genes with at least 2-fold more expression in the SW480M9 cells. The upregulation of these 193 genes also reached statistical significance. Highlighted in Figure 4.14 is MMP-9, which, as expected, was significantly up-regulated according to the array results.
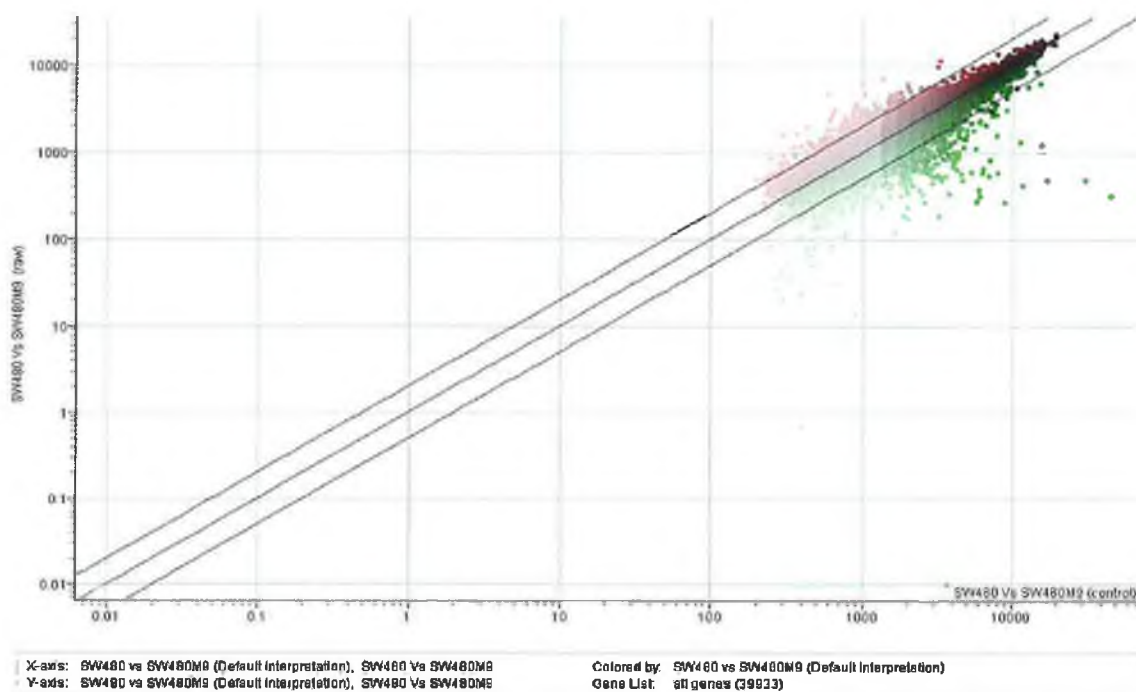
Figure 4.13 Scatter plot of data obtained from the array experiment comparing SW480 cells with MMP-9 overexpressing cells: SW480M9.

Figure 4.14 Overexpression of MMP-9 (highlighted in black) resulted in the upregulation of 193 genes (p<0.005).

The full list of all 193 genes that were significantly up regulated due to MMP-9 overexpression is included in appendix A of this thesis. Table 4.7 lists a selection of these and also includes the fold changes in expression for these genes.

As seen in Figure 4.15, 20 genes were significantly downregulated in SW480M9 cells when compared with SW480 cells. Each gene is identified by is accession number or its abbreviation.

Table 4.7 MMP-9 overexpression enhanced the expression of 193 genes (p<0.005). Above is a selection of those genes.

Gene Number	Fold Change	Gene Identity
AA394240	6.417	X-box binding protein 1
T72581	4.686	matrix metalloproteinase 9 (gelatinase B, 92kDa gelatinase, 92kDa type IV collagenase)
AA479344	3.822	gap junction protein, beta 5 (connexin 31.1)
AA775447	3.433	alpha-2-macroglobulin
AA676765	3.389	mitochondrial ribosomal protein 63
AA453796	3.2	tetraspanin similiar to uroplakin 1
AA733038	2.729	RNA polymerase I subunit
AA443193	2.698	homolog of yeast Sec5
AA479950	2.497	ATPase, Class II, type 9B
AA683578	2.377	Homo sapiens mRNA; cDNA DKFZp667G2419 (from clone DKFZp667G2419)
H29895	2.304	guanine nucleotide binding protein (G protein), gamma transducing activity polypeptide 2
AA700155	2.158	Crm, cramped-like (Drosophila)
AA706790	1.959	kinesin 2 60/70kDa

Colored by: SW480 vs SW480M9 (Default Interpretation)
Gene List: M9 significantly 2 fold down (20)

Figure 4.15 Genes downregulated in SW480M9 cells that met statistical significance (p<0.005).

Some of these 20 genes and their fold change in expression are listed in Table 4.8.

Table 4.8 Selection of genes downregulated as a result of MMP-9 overexpression in SW480 cells.

AI015679	0.298	phosphoserine aminotransferase
AA676466	0.393	argininosuccinate synthetase
AA894927	0.408	asparagine synthetase
AI080633	0.453	pyruvate dehydrogenase phosphatase
AA447746	0.460	HIF-1 responsive RTP801
AI359120	0.470	chromodomain helicase DNA binding protein 6
AA598496	0.473	IQ motif containing GTPase activating protein 1
AA278630	0.491	differential display and activated by p53

4.2.4.3 Data analysis of a 'self' vs. 'self' experiment

As outlined in Table 4.2, experiments 11 and 12 represented replicates of a 'self' Vs 'self' control experiment where the same sample was labelled differently and co-hybridised. The 'self' vs. 'self' experiment acts as a control to rule out false results that might occur due to dye bias or unequal incorporation.

X-axis: SW480 Vs SW480 (Default Interpretation), SW480 V SW480   Colored by: SW480 Vs SW480 (Default Interpretation)
Y-axis: SW480 Vs SW480 (Default Interpretation), SW480 V SW480   Gene List all genes (39933)
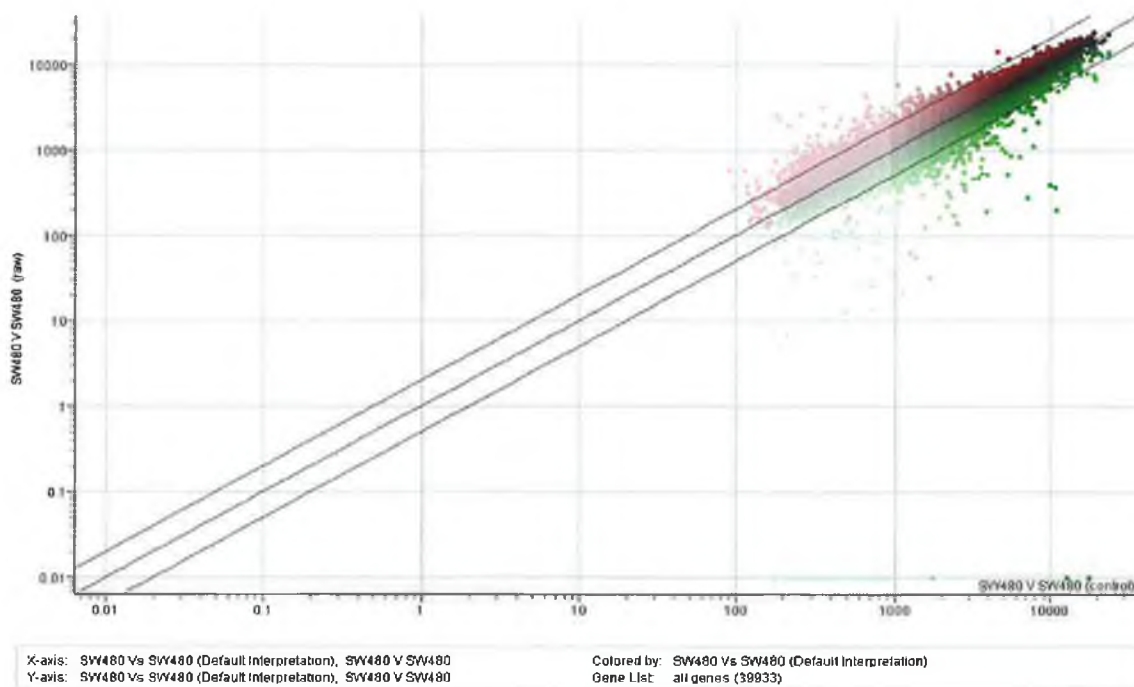
Figure 4.16 Normalised data for the 'self' vs. 'self' control experiment where SW480 cDNA was labelled differently and cohybridised on the same slide.

The normalised data obtained from this experiment is represented in Figure 4.16 and was analysed for genes that had changes in expression. Out of all 39,365 genes on

the array, no genes were significantly downregulated but 9 upregulated genes met statistical significance (figure 4.17).

Figure 4.17 In a 'self' vs. 'self' control experiment 9 upregulated genes met statistical significance (p<0.005).

These genes are listed in Table 4.9.

Table 4.9 Genes with 2 fold or more increased expression in a 'self' vs. 'self' experiment.

Gene Number	Fold Change	Gene Identity
AA057433	3.043	Data not found
AA878391	2.508	glypican 5
R26082	2.439	zinc finger protein 354A
AA789328	2.437	cyclin-dependent kinase (CDC2-like) 10
AA598665	2.374	SLIT-ROBO Rho GTPase activating protein 2
AA450351	2.290	protein inhibitor of activated STAT, 1
AA708605	2.196	Transcribed sequences
H74008	2.154	Data not found
H06282	2.014	suppressor of variegation 3-9 homolog 2 (Drosophila)

4.2.5 Post array analysis and validation of selected genes

After genes that were significantly up or down regulated were identified, the next step was to carry out detailed literature searches and select genes for further

validation and investigation. Table 4.10 lists 5 genes that had expression changes that met statistical significance in the array experiments. The decision to further investigate and validate these genes is discussed later in this chapter.

Table 4.10 Genes selected for further analysis and validation.

Accession #	Gene Name	Fold Change	Experiment
AA644088	Cathepsin C	3 up	SW620
AA055835	Caveolin 1	3 up	SW620
AA088420	Peroxisome Proliferative Activated Receptor	4 up	SW620
R39239	Tenascin C	5 up	SW620
AA894927	Asparagine Synthetase	3 down	SW480M7 and SW480M9

4.2.5.1 Sequence validation

As a quality control measure, the clones printed on the slides were also sequenced to confirm their identity. The sequences of the 5 genes chosen for further analysis were positively identified by their sequence using an ABI 377 DNA sequencing platform. The resulting sequences are included in appendix B of this thesis. Figure 4.18 is the resulting chromatogram for the sequencing of the first 100 bases of the tenascin C clone and is shown as a representation of the 5 genes chosen for further validation.
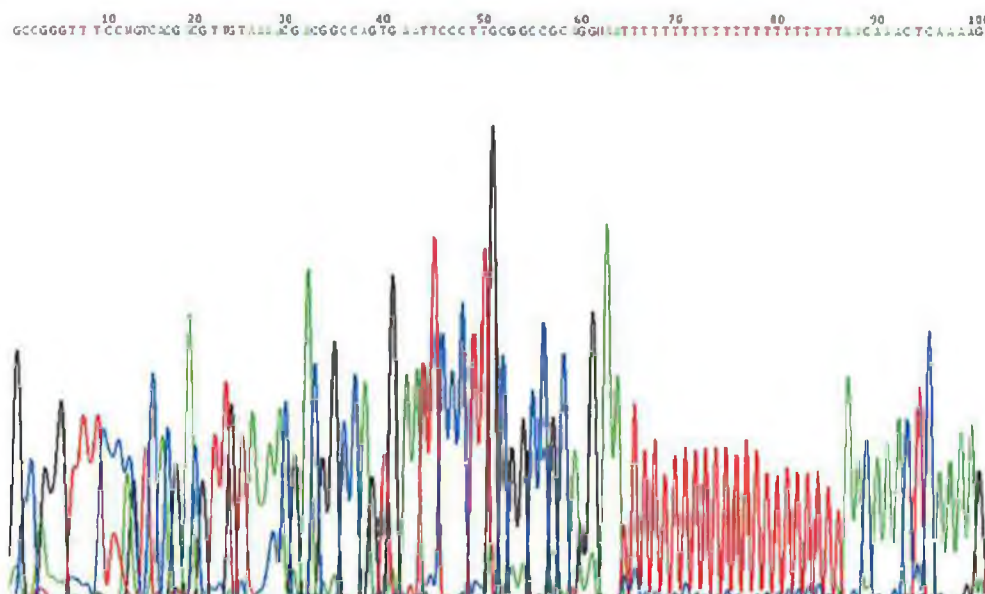
Figure 4.18 Sequencing chromatogram for the first 100 bases of the tenascin C clone. Each coloured line represents a different base, (black: guanine, blue: cytosine, red: thymine, green: adenine. 'N' is inserted when the identity of the base cannot be determined.

Using the BLAST tool on the National Centre for Biotechnology Information (NCBI) website (http://www.ncbi.nlm.nih.gov/blast/, 17 September 2004) the sequences of the clones matched their respective targets with the following accuracy: Tenascin C 95%, Caveolin 1 88%, PPARγ 96%, Asparagine Synthetase 98% and Cathepsin C 88%.

4.2.5.2 Quantitative real time PCR of target genes: validation of expression ratios

Quantitative real time PCR as outlined in section 2.2.4 was used to confirm the fold changes in expression. In comparison to conventional PCR, real time PCR provides a much more accurate and quantitative representation of the differences in the expression of a gene from one state to another.

4.2.5.2.1 Assessment of primers for use in real time PCR

Real Time PCR removes the need to run products on an agarose gel, nevertheless the primers used must be specific and therefore the products of control reactions were run on a gel to confirm the reaction yielded one product of the desired size. Standard RT-PCR was carried out as outlined in section 2.2.15.

RNA extracted from the SW620 cells was used to test the primers for cathepsin C, tenascin C, caveolin 1 and PPARγ as their expression was shown to be increased in this cell line. RNA from SW480 cells was used to test the primers for asparagine synthetase. As can be seen in Figure 4.19 the primers used in the real time PCR yielded only one product of the correct size when used in a standard PCR reaction.
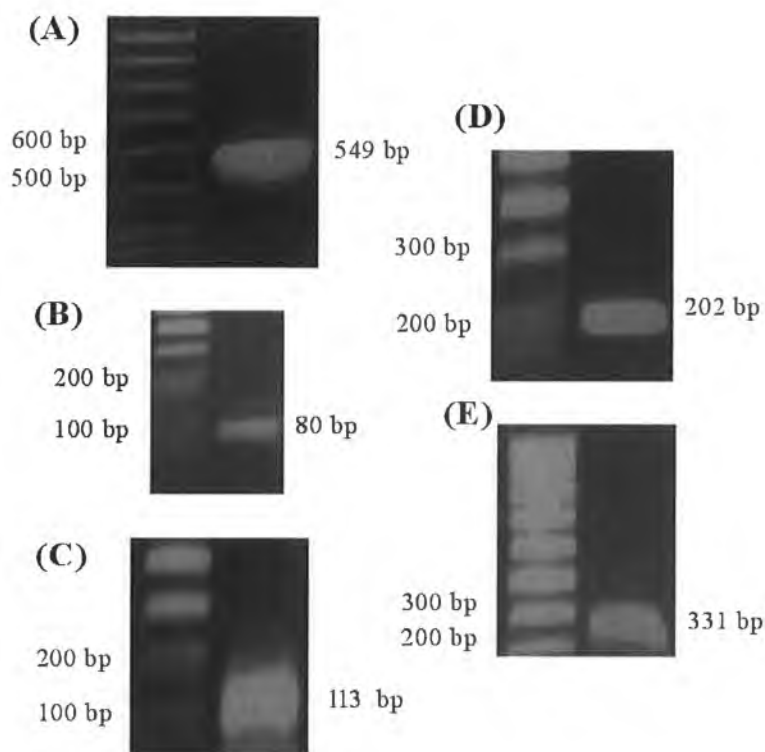
Figure 4.19 Specificity of primers used in real time PCR. Separate RT-PCR reactions using primers for the five selected genes yielded only one product, each of the following sizes: (A) Cathepsin C 549 bp, (B) Caveolin 1 80 bp, (C) PPARγ 113 bp, (D) Asparagine Synthetase 202 bp and (E) Tenascin C 331 bp.

4.2.5.2.2 Real time PCR analysis of β-actin expression

All real time PCR analysis was carried out in triplicate. The expression of β-actin was used to normalise the expression of the target genes, thus making inter-sample comparisons possible. For comparative analysis, the outputted real time fluorescent signal was differentiated to obtain the curve in Figure 4.20.
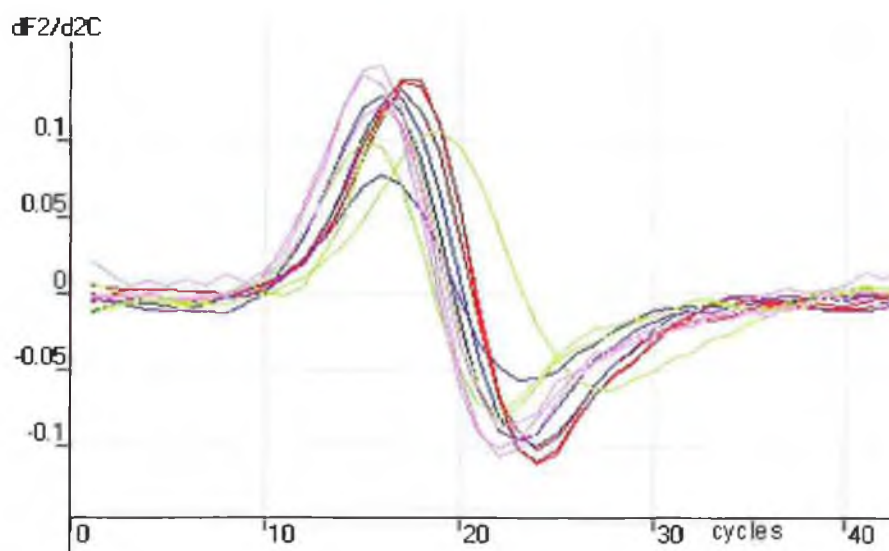
Figure 4.20 β-actin real time PCR analysis. The above curve was used to quantitate expression. Each colour represents a sample analysed in triplicate (Table 4.11).

The comparative quantitation curve for β-actin for all samples fall on one another, indicating, as expected, that there was no major difference in actin expression between the samples.

Table 4.11 Take off and comparative quantitation values for β-actin real time PCR for each sample in triplicate. The units for take-off are cycle number.

Colour	Name	Takeoff	Comparative Concentration
■	SW480 actin	13.1	1.00E+00
■	SW480 actin	13.3	9.07E-01
■	SW480 actin	12.7	1.22E+00
■	SW620 actin	11.8	1.88E+00
■	SW620 actin	12.9	1.10E+00
■	SW620 actin	11.7	1.98E+00
■	SW480M9 actin	11.3	2.40E+00
■	SW480M9 actin	13.5	8.23E-01
■	SW480M9 actin	13.5	8.23E-01
■	SW480M7 actin	11.2	2.52E+00
■	SW480M7 actin	11.1	2.65E+00
■	SW480M7 actin	11.9	1.79E+00

Comparative quantitation is based on the take off point and it calculates the relative concentration of each sample compared to the control sample chosen by the user. The number given is expressed in scientific notation.

To further confirm primer specificity, a melt curve was carried out at the end of each reaction. Fluorescence was measured as the completed reaction tubes were heated from 50 °C to 99 °C. A sharp drop in fluorescence indicated the melting or denaturing of the PCR product. To determine there was only one product, this output signal was differentiated (Figure 4.21). One peak indicated that there was only PCR product and that the primers used in the reaction were specific.
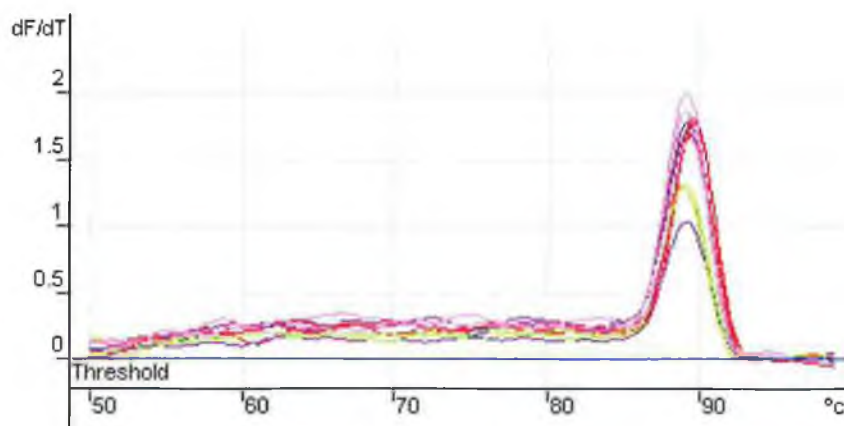
Figure 4.21 Melt curve analysis for β-actin PCR reaction. One peak at 89 °C indicated there was only one PCR product and that the primers used were therefore specific.

The above β-actin comparative quantitation values were used to compare the expression of the five target genes between these samples.

4.2.5.2.3 Quantitative real time PCR analysis of cathepsin C expression

Real time PCR analysis was used to investigate the levels of Cathepsin C expression in the SW480 and SW620 cells. The comparative quantitation curve for the expression of cathepsin C (Figure 4.22) shows that the curves for both samples are well resolved indicating a difference in expression between SW480 and SW620 cells as seen in Table 4.12.
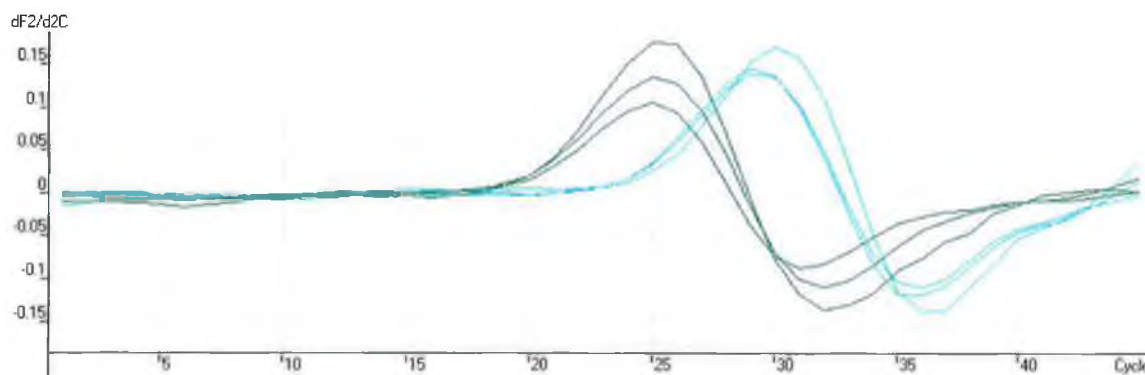
Figure 4.22 Comparison of cathepsin C expression in SW480 and SW620 cells. Each reaction was carried out in triplicate. Table 4.12 summarises colouring used.

Table 4.12 Comparative quantitation of cathepsin C (Cath C) expression in SW480 and SW620 cells.

Colour	Name	Takeoff	Comparative Concentration
▪	SW480 Cath C	25.3	1.00E+00
▪	SW480 Cath C	25.1	1.13E+00
▪	SW480 Cath C	25.9	6.98E-01
▪	SW620 Cath C	21.0	1.31E+01
▪	SW620 Cath C	21.4	1.03E+01
▪	SW620 Cath C	21.0	1.31E+01

A melt curve (Figure 4.23) was used to conclude that the cathepsin C primers used were specific. The one resulting peak indicated there was only one PCR product was formed. The PCR product was denatured at 85 °C.
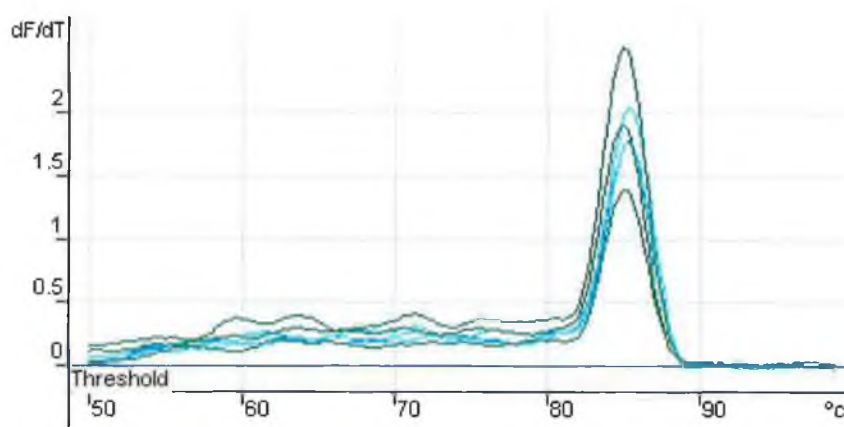
Figure 4.23 Melt curve analysis, to determine specificity of PCR primers used in cathepsin C reaction.

127

4.2.5.2.4 Caveolin 1 expression as determined by real time PCR

Figure 4.24 shows the real time data obtained for the comparison of caveolin 1 expression in the SW480 and SW620 cells. As can be seen from Figure 4.24 the curves obtained for caveolin 1 expression in the SW480 and SW620 cells did not fall on top of one another indicating differential expression.
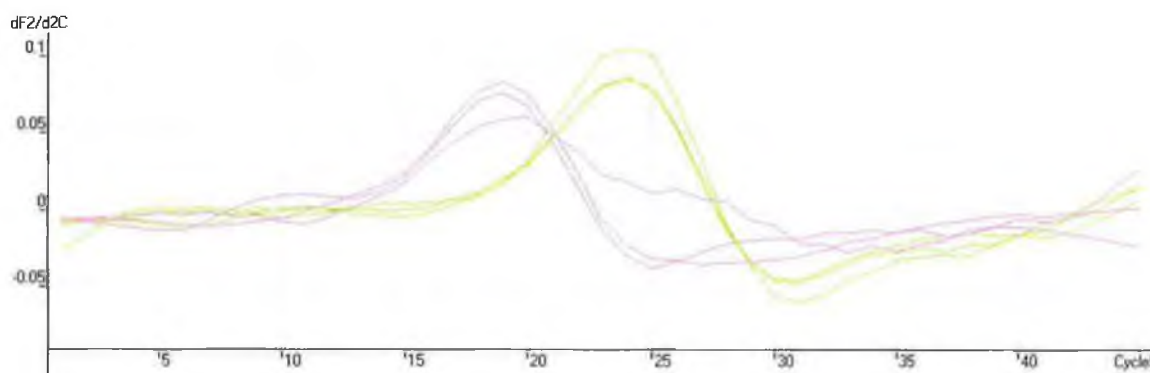
Figure 4.24 Real time data for the analysis of caveolin 1 expression in the SW480 and SW620 cell lines. Table 4.13 summarises colouring used.

Table 4.13 lists the comparative quantitation values obtained from the real time PCR reaction. It can be seen from the raw data in Table 4.13 that the SW620 caveolin 1 PCR reaction in the SW620 cells 'took-off' in the exponential phase earlier, indicating a higher level of template and therefore expression. All raw data was subsequently normalised using the β-actin data.

Table 4.13 Values obtained for the comparison of caveolin 1 expression in the SW480 and SW620 cells by real time PCR. Take off indicated the cycle in which the exponential phase was entered.

Colour	Name	Takeoff	Comparative Concentration
	SW480 Cav1	19.7	1.00E+00
	SW480 Cav1	19.8	9.44E-01
	SW480 Cav1	19.9	8.91E-01
	SW620 Cav1	15.1	1.42E+01
	SW620 Cav1	14.8	1.68E+01
	SW620 Cav1	15.2	1.34E+01

As a quality control step, melt curve analysis was carried out to highlight the specificity of the PCR primers used (Figure 4.25). One peak at 82 °C indicated that the

128

reaction yielded only one product as desired. The melting temperature of PCR products depends on the size of the product and the GC content.
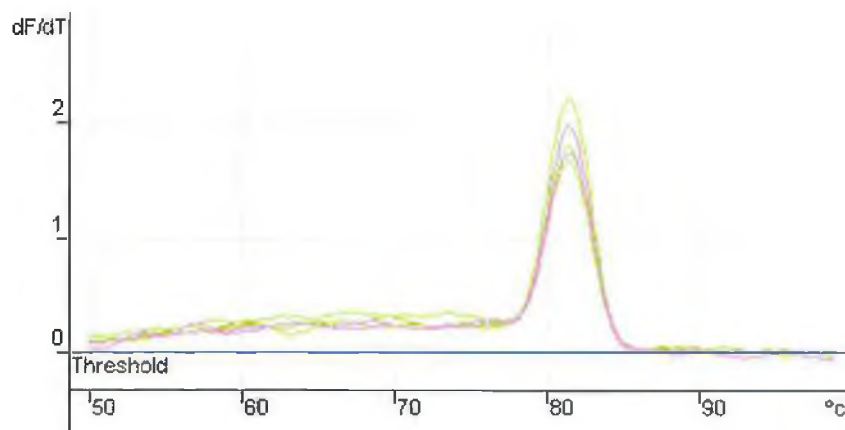
Figure 4.25 Melt curve analysis of the caveolin 1 PCR products. Table 4.13 summarises colouring used.

4.2.5.2.5 Real time PCR analysis of PPARγ expression

The expression of PPARγ in the SW480 and SW620 cells was analysed using real time PCR. The resulting comparative quantitation curve (Figure 4.26) revealed that the individual curves for both samples were well resolved, indicating a difference in PPARγ expression levels.
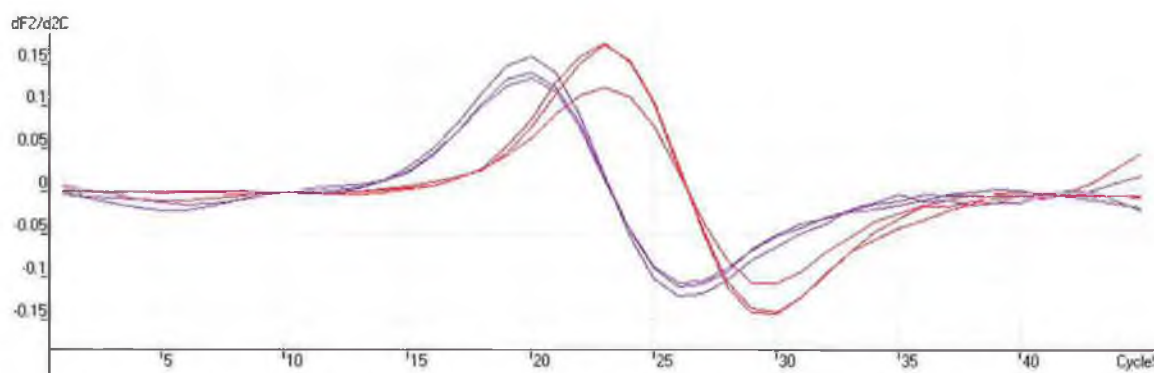
Figure 4.26 Real Time PCR comparative quantitation curve for the expression of PPARγ in the SW480 and SW620 cells. Table 4.14 summarises colouring used.

The raw data for this real time PCR reaction is listed in Table 4.14. Although this is raw un-normalised data, the comparative concentration of PPARγ was higher in the SW620 than in the SW480 cells. β-actin expression levels were later used to normalise the data, and this is discussed later in this chapter.

Table 4.14 Raw comparative concentration data obtained by real-time PCR for the expression of PPARγ in SW480 and SW620 cells.

Colour	Name	Takeoff	Comparative Concentration
■	SW480 PPAR	19.0	1.00E+00
■	SW480 PPAR	19.2	8.93E-01
■	SW480 PPAR	18.9	1.06E+00
■	SW620 PPAR	16.0	5.42E+00
■	SW620 PPAR	16.0	5.42E+00
■	SW620 PPAR	15.7	6.42E+00

Following PCR, a melt curve was carried out to assure primer specificity. The one resulting peak (Figure 4.27) indicated there was only one product formed in the reaction.
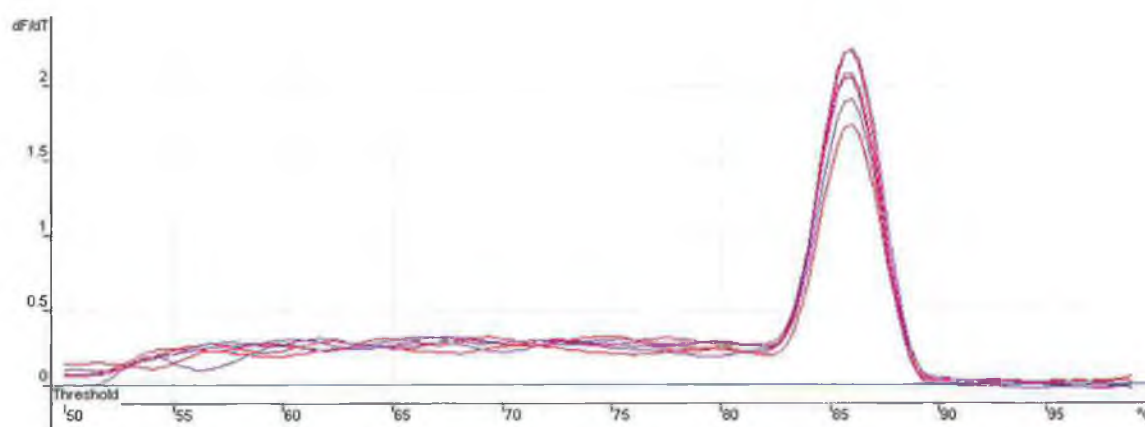
Figure 4.27 Melt curve analysis. One peak at 86 °C indicated there was only one product formed in all PPARγ real time PCR reactions.

4.2.5.2.6 Confirmation of tenascin C expression by quantitative real time PCR

Tenascin C expression was chosen for further investigation as it had increased expression in the SW620 cells in comparison with the SW480 cells. Figure 4.28 displays the tenascin C real time PCR data obtained.
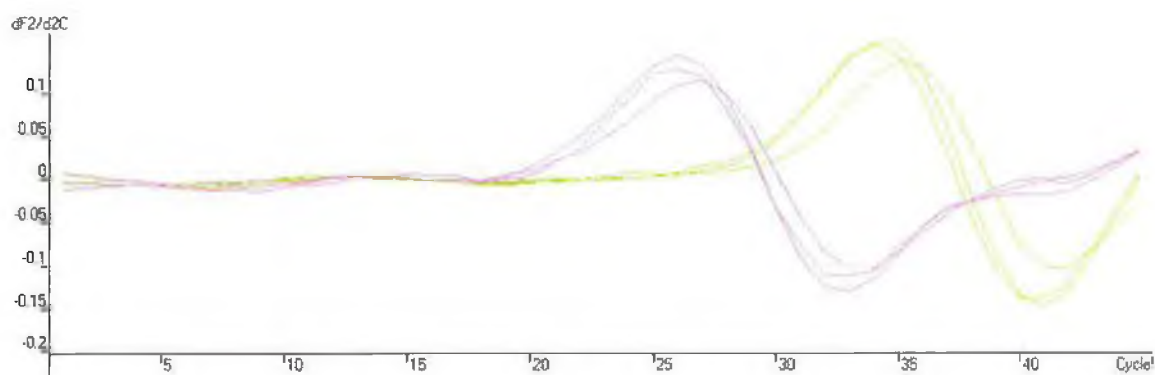
Figure 4.28 Tenascin C real time PCR comparative concentration data in the SW480 and SW620 cells. The raw data and colouring used are summarised in Table 4.15.

Table 4.15 Tenascin C expression in SW480 and SW620 cells.

Colour	Name	Takeoff	Comparative Concentration
	SW480 TenascinC	31.3	1.46E+00
	SW480 TenascinC	30.1	1.70E+00
	SW480 TenascinC	29.7	1.54E+00
	SW620 TenascinC	21.6	4.30E+00
	SW620 TenascinC	21.9	6.83E+00
	SW620 TenascinC	22.1	5.99E+00

The raw comparative concentration values for tenascin C are higher in the SW620 cells and the 'take off' time is earlier indicating higher expression levels. The melt curve carried out after the PCR reaction shown in Figure 4.29 had one peak indicating that only one product was formed in the reaction.
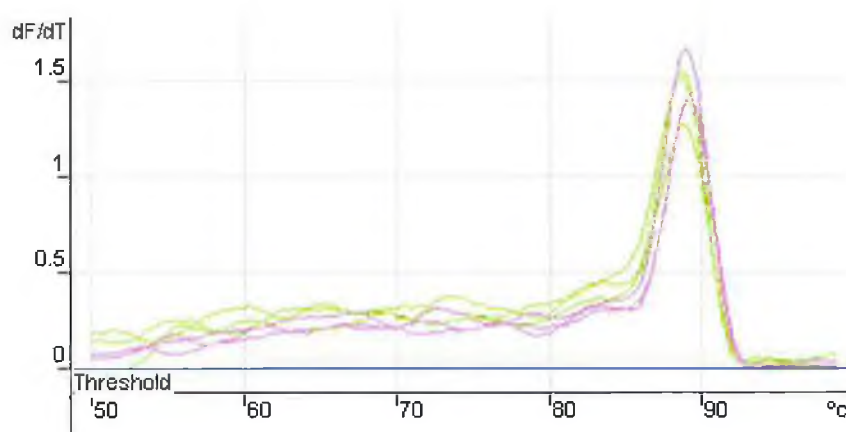
Figure 4.29 The melt curve for tenascin C real time PCR indicated that only one PCR product was formed in both SW480 and SW620 reactions.

131

4.2.5.2.7 Real time PCR analysis of asparagine synthetase expression

The expression of asparagine synthetase was investigated in the SW480, SW480M7 and SW480M9 cell lines as it was shown by microarray analysis to be down regulated in SW480M7 and SW480M9 in comparison with SW480.
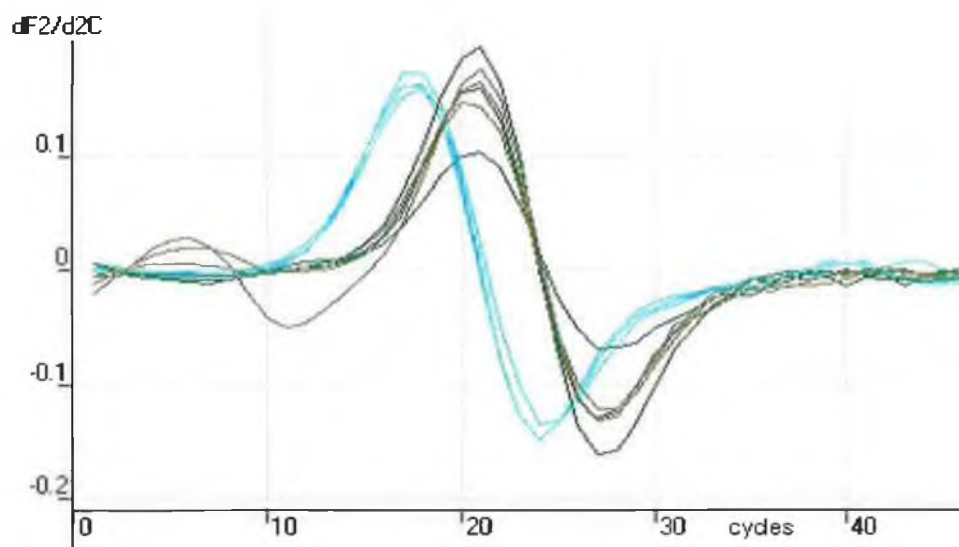
Figure 4.30 Real time PCR asparagine synthetase comparative concentration data for SW480, SW480M7 and SW480M9 cells. Raw data and colouring used is summarised in Table 4.16.

Table 4.16 Raw data obtained from the real time PCR analysis of asparagine synthetase expression. SW480 cells had an earlier take off time and higher comparative concentration values.

Colour	Name	Takeoff	Comparative Concentration
	SW480 AspSyn	13.6	1.00E+00
	SW480 AspSyn	13.4	1.13E+00
	SW480 AspSyn	13.3	1.20E+00
	SW480M7 AspSyn	16.4	2.03E-01
	SW480M7 AspSyn	16.7	2.77E-01
	SW480M7 AspSyn	16.8	2.11E-01
	SW480M9 AspSyn	16.6	1.06E-01
	SW480M9 AspSyn	16.7	1.25E-01
	SW480M9 AspSyn	17.3	1.45E-01

As can be seen from the comparative concentration curve (Figure 4.30) and the raw data obtained (Table 4.16) the SW480 cells expressed more asparagine synthetase than the

SW480M7 and SW480M9 cells. The melt curve displayed in Figure 4.31 showed only one peak, indicating the PCR reaction was specific.
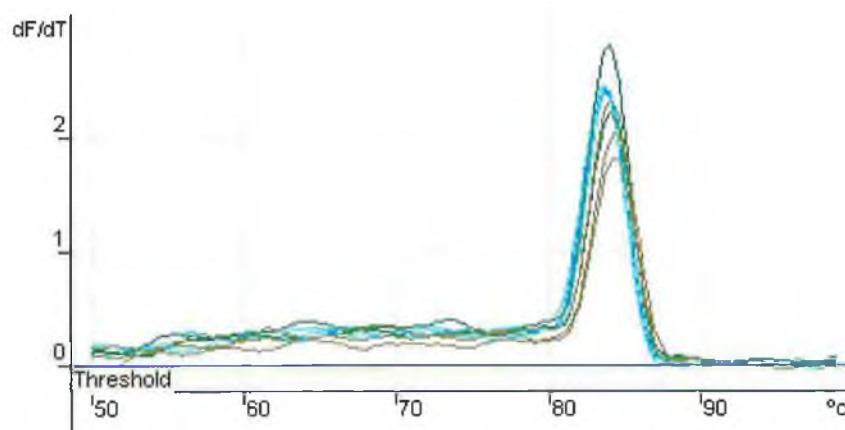
Figure 4.31 The melt curve for asparagine synthetase indicated that conditions used were specific for one PCR product.

4.2.5.2.8 Normalisation of real time PCR data

The raw data obtained from all PCR reactions was normalised using the β-actin expression levels. Figure 4.32 is the normalised real-time PCR data for the genes of interest. This data was obtained by analysing three separate RNA preparations for each cell line. This confirmed that the data obtained by cDNA microarray was real and could be validated using another technique.
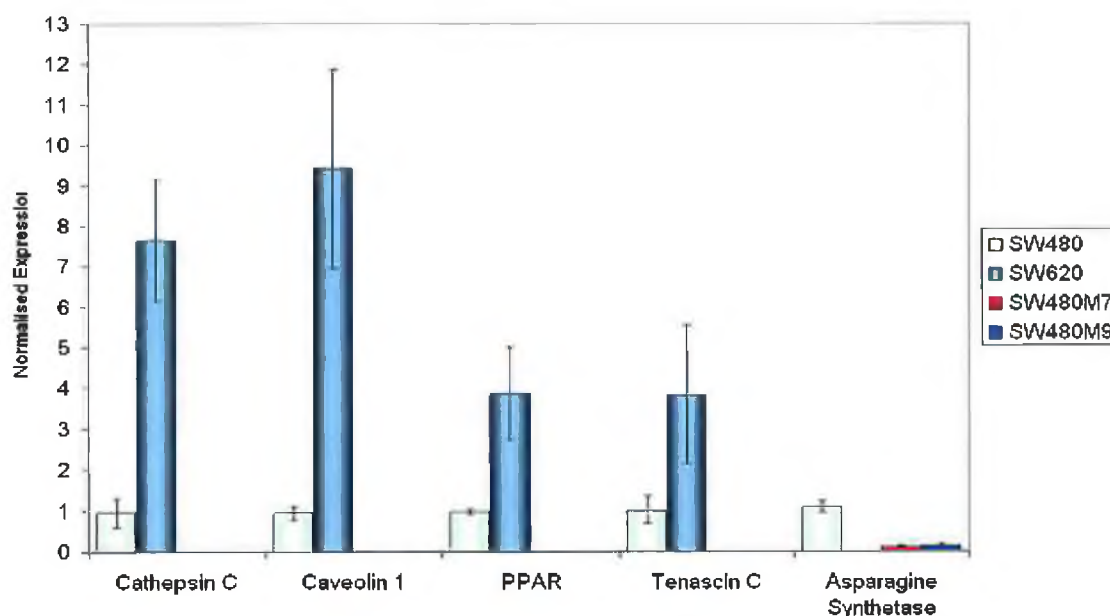
Figure 4.32 Real time PCR confirmation of fold changes in expression.

4.2.5.3 Cathepsin C: validation of protein activity

As Cathepsin C is a proteolytic enzyme, its activity was assayed at the protein level using a specific fluorescent substrate as outlined in section 2.2.18. A specific peptide substrate for cathepsin C was coupled to amino-methylcoumarin (AMC). This compound emitted fluorescence when the peptide was cleaved, thus allowing quantitation of cathepsin C activity by comparison to an AMC standard curve (Figure 4.33).
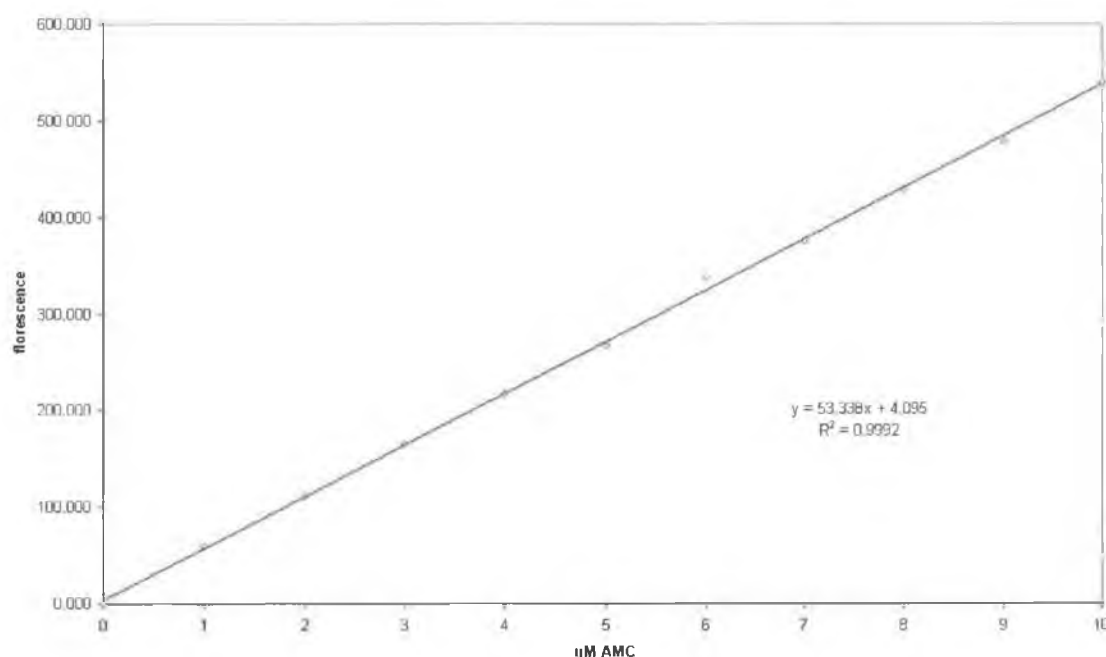
Figure 4.33 AMC standard curve.

To calculate the amount of AMC liberated per µg protein the total protein concentration of cell lysates was determined by BCA assay. The raw fluorescence and protein concentration data for all four samples is listed in Table 4.17.

Table 4.17 The amount of AMC liberated per µg total protein was calculated using the total protein concentration values. Note: 10 µl of each sample was used in each assay.

Sample	Protein (µg/µl)	µM AMC liberated	µM AMC liberated per minute per µg Protein
SW480	1.4	1.35	0.097
SW620	2.4	5.08	0.212
SW480M7	1.5	0.63	0.043
SW480M9	1.4	1.11	0.077

Figure 4.34 displays the data obtained for the amount of AMC liberated per µg total protein. This assay was carried out in triplicate.
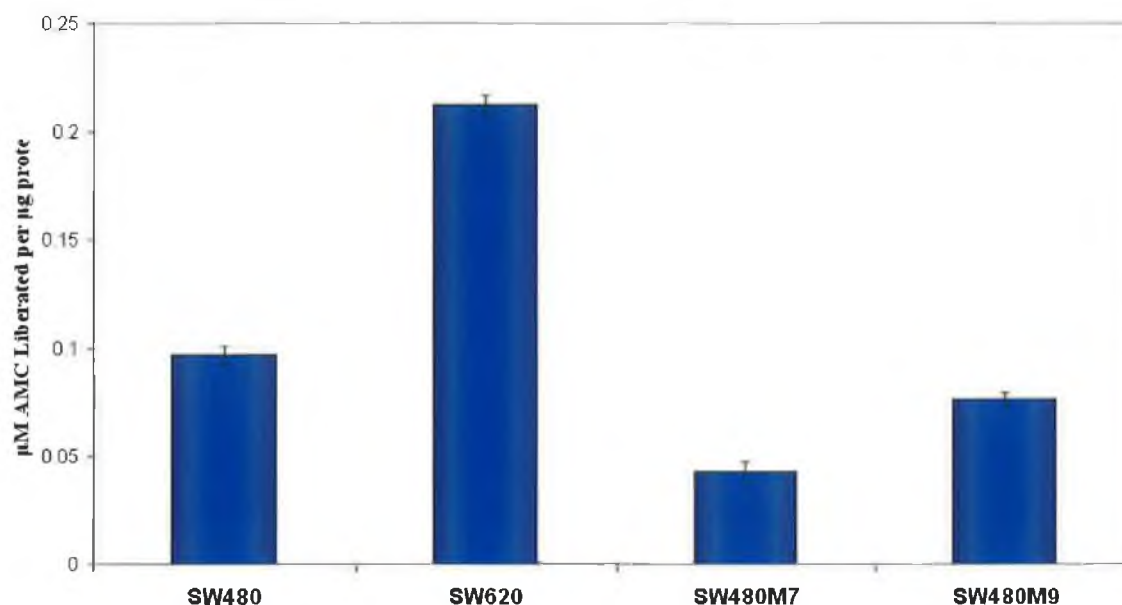
Figure 4.34 Cathepsin C activity in colorectal cancer cell lines. Cathepsin C expression was assayed in total cell lysates and expressed as the amount of AMC liberated per µg total protein.

4.2.5.4 Analysis of variant tenascin C isoform expression

To further analyse the expression of Tenascin C in the SW480 and SW620 cell lines, specific PCR primers were designed to investigate the multiple isoforms that exist due to alternative splicing (Figure 4.35).
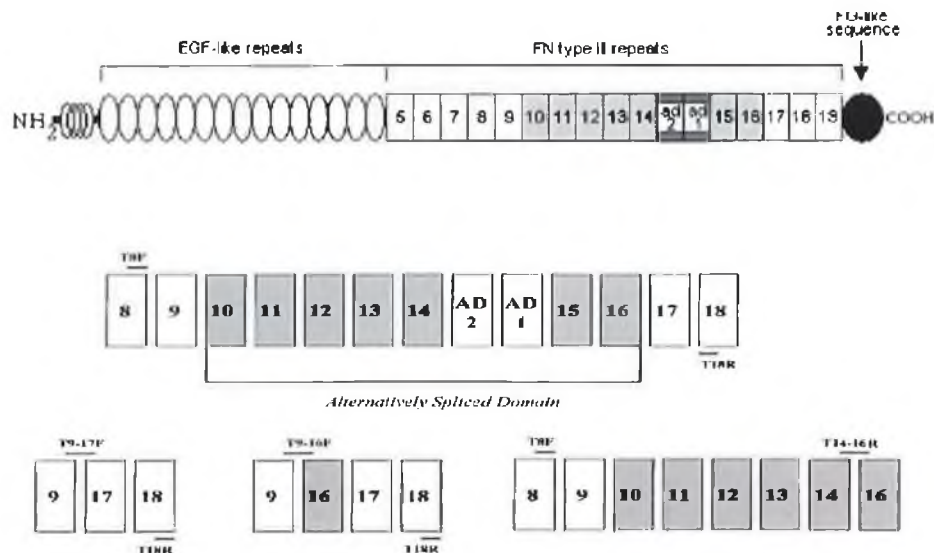
Figure 4.35 Design of primers used to investigate alternative splicing. The top part of this image is a graphic representation of the tenascin C protein. Exons 10 to 16 are subject to variant exon splicing. The 8F/18R primer cassette was used to amplify all variants. The T9-17F/18R primer cassette was used to amplify the truncated isoform of tenascin C only. The T9-16F/18R primer cassette was used to amplify tenascin C isoforms with exon 16 only inserted. The T8F/T-14-16R cassette was used to amplify all isoforms where exons 14 and 16 were spliced together. Figure adapted from Adams *et al.*, 2002.

Figure 4.36 shows the expression profile of tenascin C in both the SW480 and SW620 cell lines using an 8F/18R PCR primer cassette, which amplified all tenascin C isoforms.

Figure 4.36 Expression profile of all tenascin C isoforms in SW480 and SW620 cells as determined by RT-PCR using 8F/18R primer cassette. β-actin was used as a loading control.

As can be seen from figure 4.36 the SW480 and SW620 cells have different expression profiles and the SW620 cells expressed more overall tenascin C. The SW480 cells expressed less truncated tenasin C than the SW620 cells. Because all tenascin C variant exons are 272 bp in length, they cannot be identified using this PCR reaction. PCR using exon specific primers was therefore used.
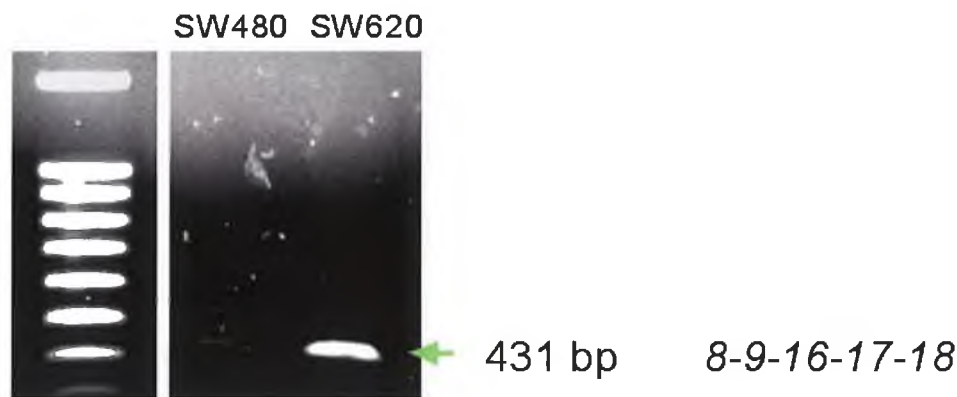
Figure 4.37 Exon-16 specific tenascin C PCR using specific primer set to amplify the isoform with exons 8, 9, 16, 17 and 18 expressed in sequence.

PCR using exon 16-specific primers (figure 4.37) was used to show that the SW620 cells expressed this isoform, where only exon 16 is spliced. SW480 cells did not express measurable levels of this exon.
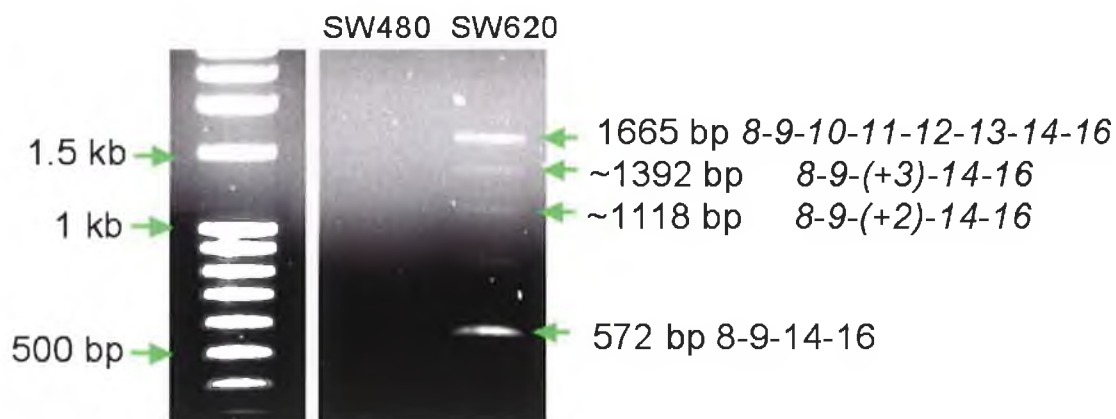
Figure 4.38 The expression of tenascin C isoforms where exons 14 and 16 are spliced together.

The SW620 cells were also shown to express exons 14 and 16 side by side (figure 4.38). The SW620 cells expressed 5 different variants all of which omit exon 15. SW480 cells did not express any 14/16 isoforms.

137

4.3 Discussion

In this chapter, the microarray analysis of a panel of colorectal cancer cell lines was described. cDNA arrays were used to compare and contrast the expression profiles of two cell lines which represented a model of human colorectal metastasis. The non-metastatic SW480 and metastatic SW620 cells lines were derived from primary and secondary tumours from a single patient respectively and therefore represent an ideal model for studying the later stages of colorectal cancer progression (Leibovitz *et al.*, 1976 and Hewitt *et al.*, 2000). Such analysis has the potential of identifying genes of importance in the metastatic process by comparison of expression profiles of these isogenic cell lines and identifying differentially expressed genes. Because these cell lines share a common genetic background, studying their genetic differences was therefore simplified because background genetic variation was minimised and expression changes most likely represented metastasis specific as opposed to individual specific changes.

The aim of this array study was to identify metastasis related genes or genes which conferred a metastatic potential or survival advantage on these cells and therefore gain an insight into this process. There was also the potential of identifying novel genes as possible therapeutic targets or prognostic markers. The expression profiles of cells stably transfected with the cDNAs for MMP-7 and MMP-9, which have been termed SW480M7 and SW480M9 respectively were also compared with that of the parental SW480 cells. The MMPs have established roles in the metastatic cascade (Curran and Murray, 1999) and the identification of downstream transcriptional events due to their activity could further clarify these roles. Cells were grown in culture under standard conditions and RNA was prepared from three independent culture flasks. This ensured biological replication.

Table 4.18 Summary of results from cDNA microarray experiments.

Experiment	≥ 2 fold up	≤ 2 fold down
SW480 vs SW620	233	208
SW480 vs SW480M7	147	12
SW480 vs SW480M9	193	20

Following the array experiments and subsequent data analysis, genes of interest were chosen. As table 4.18 summarises there were many genes that were significantly two fold up or down in each experiment (e.g. 233 genes 2 fold up and 208 genes 2 fold down in SW620 cells when compared with SW480 cells). This data alone is an important collection of quality information. Due to the experimental design and the subsequent data analysis, these gene lists have use in further analysis and data mining where the potential to answer many biological questions regarding metastasis is very real.

The overlapping genes with expression changes that are common in all experiments are represented by the Venn diagrams in Figure 4.39. The full list of these common genes is included in appendix C to this thesis. In all experiments the parental SW480 cells were used as the control. The gene changes represented in Figure 4.39 all reach statistical significance. 44 genes were significantly up regulated in both the SW480M7 and SW480M9 cells. Interestingly, there were no significant gene changes that were common to all experiments.
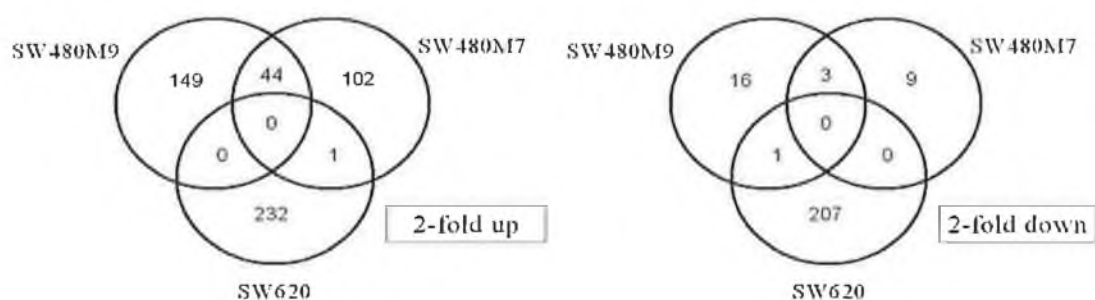
Figure 4.39 Number of genes with common changes in expression over all experiments.

Recently the same SW480 and SW620 cell model has been used in an array study into the mechanisms of drug resistance (Huerta *et al.*, 2003). Although different array technology to that described here was employed, there are several differentially expressed genes that are common to both studies including glutathione transferase, guanine nucleotide-binding protein and cyclin-D1. It must be noted that the Huerta *et al.* (2003) study reported genes that met two criteria: differentially expressed as a result of treatment with the chemotherapeutic agent cisplatin and differentially expressed in the untreated SW480 versus SW620 cells and this explains why so few genes are in common. The RNA was not prepared in triplicate in the Huerta *et al* (2003) study plus the array used had only 12,625 elements in comparison with the 39,365 used in this study. In this study glutathione transferase expression was 2.1 fold higher in the SW620 cells in comparison with the SW480 cells. Glutathione transferase levels are elevated in patients with a high risk of developing colon cancer and its over expression leads to chemotherapeutic resistance in ovarian cancer cell lines (Grubben *et al.*, 2001 and Zhang *et al.*, 2001).

Serial analysis of gene expression (SAGE) has also been used to study the same cell model to identify changes in gene expression during colorectal cancer progression (Parle-McDermott *et al.*, 2000). 5 genes with differential expression were reported in the Parle-McDermott *et al* (2000) study but then only 5,000 tagged clones were studied. SAGE is an effective technique but does not allow for the same statistical interpretation and analysis therefore it can be argued that the results are less reliable. Only one of these 5 differentially expressed genes is common to this study, keratin 5 (K5) was three fold down regulated in SW620 cells in comparison with SW480 cells in this study whereas K5 was expressed in SW480 cells and undetected in SW620 cells by SAGE analysis (Parle-McDermott *et al.*, 2000). K5 is a cytoskeletal filament protein known to play a major role in cell motility, invasion, proliferation, differentiation and in the transduction of extracellular signals (Fuchs and Weber, 1994). The same cell model was used, along with two other cell models of colorectal metastasis in an elegant 19,000 element array study (Hedge *et al.*, 2001). 1,569 genes were shown to be differentially regulated between the SW480 and SW620 cells and 176 of these were differentially regulated in all three models of colorectal metastasis. Of the 176 reported genes very few were common to those identified in this study. Interestingly caveolin 1 which was

upregulated in this study was detected as downregulated in the Hedge *et al* (2001) study although its validation was not reported.

For the sake of this study several genes of interest were chosen for further investigation and validation. These are listed in Table 4.10. The rationale for choosing these genes was based on detailed literature searches that revealed them as good candidates for further investigation. Genes were short listed because of the variety of functions of their respective protein products. These included a proteolytic enzyme (cathepsin C), an ECM component (tenascin C), a protein involved in cellular signalling (caveolin 1) and a transcription factor (PPARγ) all of which represented classes of molecules with known roles in metastasis. Asparagine synthetase was chosen for further investigation because overexpression of both MMP-7 and MMP-9 resulted in its downregulation. The differences in the expression of these genes in the cell lines studies was confirmed using real-time quantitative PCR (Figure 4.32) as it is more quantitative and sensitive in comparison to conventional RT-PCR. Although the exact fold changes as judged by microarray and real-time are not equal, the trends are the same. The reason the fold changes are not measured to be equal between both methods is possibly due to the differences in starting RNA. Both were carried out in triplicate but different RNA preparations were used.

Cathepsin C [EC 3.4.14.1] was up regulated 2.7 fold in the SW620 cells when compared to the parental SW480 cells. A member of the papain superfamily of cysteine proteases, cathepsin C is an oligomeric lysosomal protease capable of removing dipeptides from the amino terminus of its protein substrates (Paris *et al.*, 1995). As well as a dipeptidyl aminopeptidase (DAP) activity, cathepsin C has also been reported to have an endopeptidase activity (Kuribayashi *et al.*, 1993). Cathepsin C is expressed in a variety of tissues where its main functions are protein degradation and the activation of pro-enzymes including serine proteases and granzymes (Rao *et al.*, 1997). Elevated cathepsin C serum activity is associated with the pathogenesis of heart attacks, diabetes, and hepatitis (Dolenc *et al.*, 1995). Cathepsins have been particularly well described for their abnormal expression patterns during cancer development (Iacobuzio-Donahue *et al.*, 1997 and Leto *et al.*, 1997). Cysteine proteases have known roles in apoptosis, the cell cycle and in the activation of growth factors (Nitatori *et al.*, 1996, Fu *et al.*, 1998 and Mithofer *et al.*, 1998). The roles of cathepsin C in tumour progression have not been investigated thoroughly, making it a novel candidate. One study showed that

activity of cathepsin C was elevated in squamous cell lung carcinoma (SCLC) when compared to normal tissue (Krepela *et al.*, 1996). Other cathepsins have implicated roles in malignancy. In malignant colorectal tumours and premalignant lesions, the expression of cathepsin B was highly upregulated (Murnane *et al.*, 1991 and Emmert-Buck *et al.*, 1994). Cathepsin B activity is thought to be a part of a proteolytic cascade leading to degradation of extracellular matrices (Khan *et al.*, 1998b). Similarly, cathepsin H activity was increased in colorectal cancers, especially at the later stages involving invasion to lymph nodes (del Re *et al.*, 2000). The level of cathepsin D in primary breast cancer is a prognostic marker and correlates with the incidence of clinical metastasis and shorter survival (Foekens *et al.*, 1999). In addition, the down regulation of cathepsin D activity by antisense technology resulted in the inhibition of tumour growth and lung metastasis (Glondu *et al.*, 2002). The established roles of other cathepsins in cancer makes cathepsin C an interesting target for further analysis. The roles of proteolytic enzymes in metastasis are broad and varied but essential nonetheless (Liotta, 1990). Indeed, cathepsin C inhibitors have been designed with the aim of suppressing metastasis (Katunuma *et al.*, 2002).

Following array analysis and confirmation of fold change using real-time PCR, the differences in cathepsin C expression at the protein level were investigated. For this study, a fluorescent conjugate that acted as a cathepsin C specific substrate was used. Proteolysis of the substrate resulted in an increase in measurable fluorescence when excited at 370 nm. Levels of fluorescence were therefore proportional to the levels of specific cathepsin C activity. As displayed in Figure 4.34, cathepsin C activity in the SW620 cells was twice that in the SW480 cells. The differences observed at the transcription level, where SW620 cells expressed more cathepsin C mRNA than the SW480 cells, therefore translated to similar differences at the level of mature protein activity. Cathepsin C activity was not dramatically affected by overexpression of MMP-7 or MMP-9. This was expected as neither SW480M7 nor SW480M9 cells showed a statistical increase in cathepsin C expression by microarray analysis.

Caveolin 1 is a component of the plasma membrane. The caveolins are a three member family and are the major structural proteins of specialised inner foldings on the plasma membrane termed caveolae where they are involved in signal transduction and small molecule transport (Massimino *et al.*, 2002). Elevated levels of caveolin 1 have been observed in various cancers including pancreatic, oesophageal and colon cancer

(Suzuoki *et al.*, 2002, Kato *et al.*, 2002 and Fine *et al.*, 2001). Using experimental mice, increased caveolin 1 expression has been associated with increased metastatic ability and promotion of cell survival by protection from apoptosis (Thompson *et al.*, 1999). Human prostate cancers that are caveolin 1 positive have been shown to have a poor prognosis and increased metastases and serum caveolin 1 has been proposed as a potential prognostic marker of prostate cancer (Yang *et al.*, 1998 and Tahir *et al.*, 2003). Caveolin 1 has been shown to be overexpressed in colon adenocarcinoma where its expression was directly associated with tumour growth rate (Patlolla *et al.*, 2004). The role of caveolin 1 in malignancy is thought to be in the coordination of specific signalling molecules including the Src-family of tyrosine kinases, the ras family and the EGF receptor (Tahir *et al.*, 2001 and Li *et al.*, 2001). In this study the expression of caveolin 1, as determined by cDNA microarray analysis, was 2.7 times higher in the metastatic SW620 cell line than in the parental SW480 cell line. Quantitative real time PCR analysis showed a 9.5 fold increase in caveolin 1 expression in the SW620 cells in comparison with the SW480 cells thus supporting this finding.

Peroxisome Proliferative Activated Receptor gamma (PPARγ) is a nuclear hormone receptor (NHR) that regulates the transcription of various genes (Rosen and Spiegelman, 2001). PPARγ is a ligand-activated transcriptional factor that regulates cell proliferation, differentiation and apoptosis in both normal and cancer cells. The PPARγ receptor is targeted by endogenous prostaglandins (Fajas *et al.*, 2001) and it forms heterodimers with the retenoid X receptor, which binds PPARγ responsive elements thus modulating the transcription of the genes involved (Juge-Aubry *et al.*, 1997). High levels of PPARγ expression have been seen in many cancers including liposarcoma, a malignancy of the adipose lineage (Tontonoz *et al.*, 1997). PPARγ regulates the transcription of genes involved in epithelial differentiation (Rosen and Spiegelman, 2001) and has been shown to upregulate caveolin 1 expression (Burgermeister *et al.*, 2003). PPARγ expression has been found in various epithelial tissues including colonic mucosa (Mansen *et al.*, 1996). PPARγ expression is primarily localized in more differentiated epithelial cells of the colon and its expression and activation therefore are associated with the differentiation of intestinal cells (Lefebvre *et al.*, 1999). While PPARγ ligands have been shown to reduce cell growth in human colon cancers (Sarraf *et al.*, 1998), paradoxically, PPARγ ligands have also been shown to cause an increase in colon tumour number in mice suggesting the role of PPARγ in the biology of colon cancer may be complex (Saez *et al.*, 1998). PPARγ ligands have

recently been shown to inhibit the *in vitro* invasion of a highly metastatic human breast cell line (Liu *et al.*, 2003). Although the complete mechanisms of PPARγ are still unclear, the modulation of its activity may have therapeutic benefits. There are few reports examining the expression of PPARγ in metastasis making it a novel candidate for further investigation in this situation. In this study PPARγ was upregulated 3.9 fold in the SW620 cells in comparison with the SW480 cells as determined by cDNA array analysis. Real time PCR data confirmed its upregulation and showed that PPARγ expression in SW620 cells was 4 times that of SW480 cells.

Tenascin C is a large glycoprotein and a major component of the ECM. Tenascin C expression is usually limited to areas of cell proliferation or motility, and its expression is therefore transient. The protein is re-expressed in various malignant neoplasms, including colon and breast cancers where high expression has been associated with a poor prognostic outcome and a high risk of distant metastasis (Hauptmann *et al.*, 1995, Ishihara *et al.*, 1995 and Kressner *et al.*, 1997). Tenascin C expression is found in normal surrounding stromal and epithelial tissue as well as in the tumour itself where its expression is often concentrated at the invasion front (Yoshida *et al.*, 1997 and Jahkola *et al.*, 1998). The linear arrangement of one tenascin C polypeptide can be seen in Figure 4.35. Six tenascin C polypeptides can interact via their cysteine rich amino termini to form hexamers. Alternatively spliced fibronectin type III domains follow a series of EGF-like repeats. The structure and size of tenascin C varies as a result of this pre-mRNA alternative splicing which gives rise to functional diversity of the protein (Jones and Jones, 2000). The truncated form of tenascin C, in which no alternatively spliced exons are included, is generally found in static tissues, and large variants containing alternatively spliced FNIII domains in various combinations are found in developing tissues (Mackie and Tucker, 1992 and Joester and Faissner, 1999). Tenascin C can interact with cell surface receptors and with other ECM components for example fibronectin thus influencing cell behaviour (Jones and Jones, 2000 and Chiquet-Ehrismann *et al.*, 1991). Tenascin C can promote cell migration, angiogenesis, cell proliferation, and act as a cell survival factor (Phillips *et al.*, 1998, Schenk *et al.*, 1999, Swindle *et al.*, 2001 and Cowan *et al.*, 2000). Tenascin C can induce the expression of MMP-9, which has also been shown to promote tumour growth and progression (Tremble *et al.*, 1994 and Kalembeyi *et al.*, 2003). Tenascin C and MMP-2 are often coexpressed during normal processes such as wound healing and in the tumour stroma, suggesting their functions may be inter-related (Sternlicht *et al.*,

2000). Expression of larger tenascin C isoforms has also been associated with increased migration and invasion of carcinomas (Borsi *et al.*, 1992) and a switch from small to larger isoforms has been reported in colorectal (Dueck *et al.*, 1999) and breast cancers (Adams *et al.*, 2002). An increase in tenascin C isoforms containing exons 10-13 has been shown to be related to the stage of colorectal cancer progression while an increase in exon 16 is associated with a metastatic phenotype in colon cancer (Dueck *et al.*, 1999). An increase in the expression of two tenascin C isoforms has been associated with breast malignancies, one containing exon 16 only and another containing both exons 14 and 16 (Adams *et al.*, 2002). Expression of the 14/16 isoform correlated with a more invasive phenotype. In this study, cDNA array analysis showed that the expression of tenascin C in SW620 cells was increased 4.9 fold in comparison with the SW480 cells. Real time PCR showed tenascin C expression was upregulated 4 fold in the metastatic SW620 cells in comparison with the parental SW480 cells, thus confirming the cDNA microarray data.

To examine the differences in isoform expression between these cell lines, PCR with primers specific for various exons was carried out. Firstly, all isoforms present in both cell lines were revealed using primers located outside and at each end of the alternatively spliced domain as illustrated in Figure 4.35. These primers annealed to exon 8 (forward) and exon 18 (reverse) neither of which are subject to alternative splicing. This resulted in a range of PCR products which represented different tenascin C isoform transcripts that varied in length due to exon splicing (Figure 4.36). This PCR reaction was not definitive in terms of information on specific isoform expression as all the alternatively spliced tenascin C exons were 272 bp in size, therefore individual bands on a gel could represent any of the possible 5040 mathematical combinations. The only conclusion that could be drawn from this information was that a PCR product of 272 bp represented the truncated form of tenascin C with no exons spliced in and a band of 2345 represented the full length isoform with all 7 alternative exons spliced in. The bands in between these two products represented different combinations of exons, and while the exact combination could not be deciphered, the number of exons inserted could be determined by the product size. As can be seen from Figure 4.36 the overall total expression of tenascin C as shown by microarray and real time PCR analysis was higher in the SW620 in comparison with the SW480 cells. It can also be seen that the expression profiles of both cell lines were different. The cell lines expressed different isoforms and different amounts of isoforms. This was apparent from the presence of a

713 bp product in the SW620 cells that was absent in the SW480 cells. This was also apparent from the difference in the expression of truncated tenascin C and therefore the difference in the intensity of the 441 bp product with the SW620 cells expressing more. β-actin was used as a loading control (Figure 4.36, lower panel) allowing comparisons to be made between samples using the intensity of the 550 bp actin product to normalise. PCR using primers specific for different exons was then carried out. Using a forward primer that spanned the 3' end of exon 9 and the 5' start of exon 16 together with exon 18 specific reverse primer (Figure 4.35) the tenascin C transcript with exon 16 alone spliced in could be specifically amplified and detected. This PCR reaction therefore yielded only one product of 431 bp representing the expression of tenascin C with exon 16 alone spliced in. As can be seen from Figure 4.37 only the SW620 cells expressed this isoform. Exon 16 expression has been associated with pre-invasive breast disease and a metastatic phenotype in colorectal cancers (Adams *et al.*, 2002 and Dueck *et al.*, 1999). The absence of exon 16 tenascin C expression in SW480 cells and its presence in SW620 cells could indicate a key event in the progression to a metastatic phenotype. Another group of tenascin C isoforms associated with metastatic diseases are those in which exon 15 is not included and therefore exons 14 and 16 are spliced next to each other. 14/16 Tenascin C isoforms have been reported to have increased expression in metastatic breast and ovarian tumours (Adams *et al.*, 2002 and Wilson *et al.*, 1996). PCR using a reverse primer that spanned the 3' end of exon 14 and the 5' start of exon 16 and an exon 8 specific forward primer (Figure 4.35) was carried out to determine expression of 14/16 tenascin C isoforms in the SW480 and SW620 cells. This PCR reaction was specific for all isoforms containing exons 14 and 16 side by side, so multiple products were therefore obtained (Figure 4.38). As seen in Figure 4.38 the SW620 cells expressed a range of 14/16 tenascin C transcripts whereas the non-metastatic SW480 cells did not express any. As the alternatively spliced exons were the same size, only the identity of the 572 and 1665 bp products could be identified as the transcript into which only exons 14 and 16 were spliced and the transcript into which all exons except exon 15 were spliced respectively. As expected from the schematic of tenascin C in Figure 4.35, there are five bands on the gel, each one representing an additional exon.

The expression of several genes were enhanced by both MMP-9 and MMP-7 overexpression. Plasminogen expression was significantly increased 2.8 and 2.4 fold in the SW480 cells overexpressing MMP-9 and MMP-7 respectively. Aspargine

synthetase was chosen for further validation and investigation as it was 2.5 and 2.8 fold downregulated in both the SW480M9 and SW480M7 cell lines respectively. This gene was therefore of interest in terms of identifying genes that were responsive to MMP overexpression. Only one other common gene was downregulated as a result of MMP-7 and MMP-9 overexpression. Phosphoserine aminotransferase 1 was 5 and 3.3 fold downregulated in the MMP-9 and MMP-7 transfected cells respectively. Asparagine synthetase catalyses the biosynthesis of asparagine from aspartic acid. Asparagine is a nonessential amino acid that has been reported as regulating the expression of genes involved in cell growth (Barbosa-Tessmann *et al.*, 2000 and Leung-Pineda and Kilberg 2002). Using microarray analysis, asparagine synthetase has recently been shown to be a target for transactivation by p53 in cancer cells and is associated with an increase in the rate of tumour cell growth (Scian *et al.*, 2004). Asparagine depletion can induce apoptosis, and the enzyme asparaginase when administered in conjunction with chemotherapy has been reported to induce an apoptotic response due to asparagine depletion (Story *et al.*, 1993). Asparaginase has therefore long been a standard component of chemotherapy against childhood acute lymphoblastic leukaemia (ALL) (Oettgen *et al.*, 1967 and Sutow *et al.*, 1971). Considering this, it can be concluded that the expression of asparagine synthetase supports cell growth. These roles for asparagine synthetase would suggest that a decrease in its expression levels as seen in the microarray analysis of SW480M7 and SW480M9 cells would not be beneficial to the cancer cell as it would lead to increased apoptosis. However, it has been shown that asparagine synthetase expression is responsive to exogenous asparagine levels, and therefore forms part of a feedback mechanism where the end product of the enzymatic reaction regulates the levels of the enzyme itself, thus suggesting asparagine levels are not depleted in the SW480M7 and SW480M9 cells (Greco *et al.*, 1989). The levels of asparagine synthetase mRNA expression and enzymatic activity have been shown to be decreased in leukemic cells and this has been hypothesised as the explanation for the effectiveness of asparaginase as a chemotherapeutic agent in childhood leukaemias (Aslanian *et al.*, 2001). Asparagine synthetase can therefore be thought of as an agonist for asparaginase. Accumulation of asparagine synthetase mRNA is associated with a block and accumulation in the G1 phase of the cell cycle of asparagine deprived cells which can be reverted by its addition (Greco *et al.*, 1989).

4.4 Conclusion

This chapter focused primarily on the identification of metastasis associated genes using microarray analysis to compare the expression profiles of a non-metastastic cell line, SW480 with its metastatic derivative, SW620. 441 genes were shown to be differentially expressed between these two cell lines. In other experiments the expression profiles of two MMP overexpressing cell lines, SW480M7 and SW480M9, which have been stably transfected with the cDNAs for MMP-7 and MMP-9 respectively, were compared with that of the parental non-transfected SW480 cells, which did not express these MMPs. The aim of these experiments was to identify genes which are induced or repressed as a downstream effect of excess MMP production and activity and thus give a molecular insight into their effects in colon cancer. MMP-7 overexpression resulted in the differential expression of 159 genes and MMP-9 activity resulted in the differential expression of 213 genes. All analysis was carried out in triplicate allowing these expression changes to be identified as statistically significant. Several genes were chosen for further validation analysis. These genes were cathepsin C, caveolin 1, tenascin C, asparagine synthetase and PPARγ. As quality control steps, sequencing of the spotted cDNA and real-time PCR of the genes of interest confirmed their identity and changes in expression respectively. A proteolytic activity assay using a cathepsin C specific fluorescent peptide revealed its upregulation at the mRNA level translated into an increase in enzyme activity in the SW620 cells. Tenascin C PCR revealed differential expression of various alternatively spliced transcripts in the SW620 in comparison to the SW480 cells. Using exon specific PCR, the SW620 cells were shown to express the metastasis-associated tenascin C 14/16 and 16 isoforms.

This work validated the further investigation of these targets and other differentially regulated genes as players in the metastatic process. It must be noted that although the SW480 and SW620 provide a good model of colon to lymphatic system metastasis they only represent snapshots of this highly complex and multistage process. The occurrence of molecular alterations at stages between the establishment of these two cell lines is very realistic and may not be accounted for. Furthermore, contributions from normal stromal and surrounding tissue are also not accounted for and neither are the contribution of the various cells that form the heterogeneous population of the tumour.

Chapter 5

Expression of novel markers of colorectal metastasis
in paired normal and tumour colon samples

5.1 Introduction

5.1.1 The genetics of human colon cancer

Colon cancers are common and account for about 15 % of all cancers in western countries. In Ireland alone there are 2,000 new cases each year (National Cancer Registry of Ireland, 1997). Colon cancer occurs mainly between the ages of 55 and 70 and males are at a slightly higher risk of developing the disease. The 5-year survival rates for colon cancer patients is 80 % if the cancer has not metastasised and 40 % if it has. Therefore, patient survival depends on the degree of spread or the stage of the disease, which is discussed in more detail later in this section. Nearly 50 % of patients diagnosed with colon or rectal (referred to as colorectal) cancer develop metastasis within five years. Most commonly, these tumours metastasise to the liver and lungs. Colon cancer can be either sporadic or inherited. Hereditary forms of colon cancer account for approximately 20 % of cases and are further divided into two forms; familial adenomatous polyposis (FAP) and hereditary non-polyposis colon cancer (HNPCC). Both familial types occur before the age of 50, which is earlier than the sporadic cancers.

The origin of colon cancer follows a defined sequence of events (Figure 5.1) whereby a number of genetic mutations are required for the progression of the disease (Fearon and Vogelstein, 1999). This progression from normal intestinal epithelium to an invasive carcinoma is estimated to take 7 –12 years. These DNA mutations result in both the loss of function, usually of tumour suppressor genes and the gain of function, usually of oncogenes.
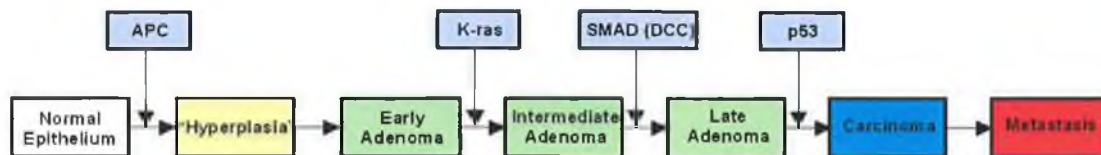
Figure 5.1 Gene changes occurring during colon carcinogenesis. The above flow diagram illustrates how colon tumourigenesis is a multistage process.

The adenomatous polyposis coli (APC) gene is located on chromosome 5 and codes for a protein that mediates signal transduction from E-cadherin. 70 % of colon cancers have APC mutations, which play an important role in both sporadic and

hereditary forms of colon cancer due to a loss in APC function. Where wild type APC normally targets β-catenin for ubiquitination, a loss of function therefore results in the translocation of excess β-catenin to the nucleus where it effects the transcription of several target genes including c-myc, cyclin D and MMP-7 (He *et al.*, 1998 and Crawford *et al.*, 1999).

In contrast, mutation of the K-Ras gene on chromosome 11 typically results in a gain of function. The Ras protein is a GTPase involved in the transduction of several cell signals in many signalling cascades including the MAPK pathway. The activation of the Ras protein is tightly regulated however Ras mutation results in its permanent activation. Therefore, the cell is constantly stimulated with mitogenic signals, thus enhancing tumourigenesis.

One of the most common regions of alleic loss is on chromosome 18, which is absent in 70 % of carcinomas and 50 % of late adenomas (Vogelstein *et al.*, 1989 and Delattre *et al.*, 1989). SMAD is located in this region and it is involved in the signalling pathway for TGFβ, a growth inhibitory, differentiation-inducing cytokine. SMAD deletion results in the cell ignoring TGFβ mediated cell proliferation inhibition and results in a loss of differentiation, increased proliferation and altered cell adhesion (Tarafa *et al.*, 2000).

The loss of large portions of chromosome 17 has been detected in 75 % of colorectal cancers (Risio *et al.*, 2003). This loss is not apparent in early adenomas, and it has been associated with the progression of the disease (Diep *et al.*, 2003 and Fearon *et al.*, 1987). This region contains the p53 gene, the product of which is a repressor protein that is essential for maintaining and regulating normal cell cycle function. p53 is the gene most frequently altered in human cancers (Levine *et al.*, 2004) and has been reported to be mutated in colorectal cancers (Bos *et al.*, 1987). p53 protects the cell from various insults such as radiation or drugs by co-ordinately blocking cell proliferation, stimulating DNA repair and promoting apoptotic cell death. In the proposed model, p53 mutations are postulated to function late in colorectal carcinogenesis (Fearon and Vogelstein, 1990).

The genetic mutations mentioned above take place during certain stages of colon tumourigenesis, thus leading to the hypothesis that colorectal carcinoma development is

a step wise procedure, however the accumulation of these mutations is of more importance than the order in which they occur.

5.1.2 Staging of colorectal cancer

Tumour staging is an extremely important prognostic parameter. In the staging of colorectal cancer, widely accepted criteria are used to categorize the extent to which the cancer has advanced thus allowing clinicians to discuss the cancer using the same language. Precise staging is critical for predicting disease free and overall survival outcomes and for planning treatments such as chemotherapy and radiation therapy. In colorectal cancers, stage is determined by: the extent to which a cancer has penetrated the bowel wall; the presence or absence of lymph node metastases; or by spread to distant organs such as the lung, liver, bone or other sites. Endoscopic methods are routinely used as a preoperative staging technique for rectal cancers and biopsies are quite often taken for analysis. Staging involves information obtained from microscopical pathological analysis of the surgically removed tumour sample where all clinical and pathologic findings are considered.

One of the most commonly used staging system is the Dukes' Classification (Sir Cuthbert Dukes), which has been in use for over 70 years (Dukes, 1932). Dukes' classification originally described three stages. Dukes' A tumours are defined as those limited to the mucosa wall of the colon. Dukes' B tumours are defined as those that have extended and penetrated through the wall and Dukes' C tumours are those with lymph node metastasis. Dukes' C tumours were later divided into two groups; C1 tumours are those where only regional lymph nodes are positive and C2 tumours are those where apical nodes are involved (Gabriel *et al.*, 1935). A fourth stage, Dukes' D was later added describing tumours with distant metastases, which were beyond the limits of surgical resection (Dukes 1949). These stages group together heterogeneous subsets of patients, therefore decreasing the ability of Dukes' classification as an indicator of prognosis for individual patients. The power of Dukes' classification to predict the behaviour of tumours in individual patients is therefore limited. To overcome these limitations additional staging systems have been developed incorporating more classification parameters. In an effort to add flexibility to staging, other systems such as the tumour-node-metastasis (TNM) staging (Greene and Sobin,

2002) have been used. However, the TNM staging system is only a modification of the Dukes classification system and is therefore not superior as a predictor of prognosis.

More detailed classification parameters have been explored to allow for more reliable prediction of tumour behaviour in individual patients. These include tumour cell differentiation (de Bruine *et al.*, 1993) and proliferation status (al-Sheneber *et al.*, 1993). Attention has also been paid to parameters that describe the invasive and metastatic behaviour of tumours such as, the expression of proteases (Adachi *et al.*, 2001 and Mulcahy *et al.*, 1994), the expression of adhesion molecules such as CD44 variant isoforms (Wielenga *et al.*, 1993), basement membrane deposits (Havenith *et al.*, 1988), venous invasion (Ouchi *et al.*, 1996) and lymph node micrometastasis (Greenson *et al.*, 1994). The exploration of these parameters aims to further sub classify colon tumours for more precise prognosis.

As colon cancers are the most extensively studied of all the solid tumours in terms of their genetic abnormalities, detailed molecular analysis of the underlying tumour cell genetics in search of prognostically relevant parameters could potentially yield highly valuable information which could be used in clinical management. Many studies have been based on the Fearon and Vogelstein (1990) model discussed earlier in this chapter. p53 and K-ras have been used as predictors of an unfavourable outcome in colorectal carcinomas (Bell *et al.*, 1994 and Moerkerk *et al.*, 1994). Colorectal tumours of increasing stage have been shown to have an increased frequency of chromosome 18 alleic losses, and it has therefore been proposed as a prognostic marker (Jen *et al.*, 1994).

5.1.3 Assessing differential expression in matched tissues

Pathologists, clinicians and all involved in cancer management are faced with a major problem because of the extensive heterogeneity of the disease. There is a lack of histological and clinical factors that can be used to reliably predict both the evolution of cancer and its sensitivity to therapies. As discussed earlier, new prognostic and predictive factors are required to allow effective individualised management of treatment strategies and therapies.

The availability of paired 'normal' and tumour tissues provides a simple model for investigating the complexity of tumour biology. The comparison of expression profiles has the potential to reveal important transcriptional differences underlying the development of the disease. As both samples originate from the same patient, the possibility of generating false results is greatly reduced and differences in expression are therefore most likely to represent changes that contribute to the disease state. Matched tissue samples also represent the biological state more so than a homogenous culture of cells. They therefore, provide a useful tool for the investigation of potential targets or markers of the disease. The expression of markers identified using cell lines can therefore be pursued in paired tissue samples to assess the significance of such expression.

5.2 Results

In this chapter, the expression of novel markers of colorectal metastasis was investigated in paired samples of normal and tumour colorectal tissue. As summarised in chapter 4, these genes were identified by cDNA expression profiling in two cell lines, one derived from a primary human colon adenocarcinoma and the other from its lymph node metastases. The genes identified were the ECM component, Tenascin C, the transcription factor, PPARγ, the membrane component, Caveolin 1 and the protease, Cathepsin C. The differences in expression of these genes were thought to be of importance in colorectal tumour metastasis.

The colorectal tissue bank used was established in 1992 in the Mid-Western Regional Hospital, Limerick, Ireland by Dr Mary Morrin. The bank is composed of matched tumour and normal specimens removed from patients during surgery for colorectal cancer. Tissues were frozen in liquid nitrogen and stored at -80 °C within 30 minutes of removal at surgery. The corresponding demographic, pathological and follow-up records for each patient were maintained using an Access database. The tissue bank contained 176 specimens which were composed of, 5 % Dukes' A, 54 % Dukes' B, 19 % Dukes' C and 22 % Dukes' D category tumours. This tissue bank with its associated records represented a unique tool for analysis of gene expression.

5.2.1 Analysis of tissue RNA

Overall, 25 randomly chosen samples, representing 12 pairs were cut, one pair also contained a polyp specimen. Information on these tissue samples was later collected and is listed in Table 5.1. As can be seen from Table 5.1 all Dukes' classifications were represented.

RNA was harvested as outlined in section 2.2.19 and the yields obtained are summarised in Table 5.2. As can be seen from Table 5.2, RNA was successfully extracted from all samples. There was a great variety in the yield obtained with some samples e.g. the tumour sample T7 and the normal samples N8 and N12 being very dilute. The 260/280 absorbance ratio values for all samples were close to the desired value of 1.8, although some samples were a lot higher, for example T9 and N12. Sample set 6 was unique in that it also included a polyp sample.

Table 5.1 Information on tissue samples used in this study, including patient gender, age and 5 year survival history. Information on the tumour includes its location and its Dukes' classification.

Sample No	Gender	Age	Location	Dukes' Classification	5-year survival
1	M	65	Left colon	D	No
2	F	69	Left colon	B	Yes
3	M	62	Rectal	B	Yes
4	M	59	Left colon	B	Yes
5	F	73	Right colon	B	No
6	F	76	Rectal	B	No
7	F	53	Right colon	C	Yes
8	M	59	Left colon	A	Yes
9	M	78	Left colon	A	Yes
10	M	57	Rectal	C	Yes
11	M	52	Left colon	D	Yes
12	M	31	Transverse	C	Yes

Table 5.2 RNA concentration and A260/280 values for paired normal (N), tumour (T) and polyp (P) specimens.

Sample	Concentration ($\mu g/\mu l$)	A260/280 ratio
N1	1.416	1.8
T1	4.748	1.6
N2	2.512	1.7
T2	3.880	1.6
N3	2.220	1.7
T3	5.800	1.3
N4	2.448	1.6
T4	4.888	1.6
N5	2.920	1.8
T5	4.348	1.7
N6	1.996	1.7
T6	3.500	1.6
P6	0.468	1.9
N7	1.348	1.6
T7	0.244	2.1
N8	0.226	1.7
T8	0.356	1.6
N9	0.856	1.6
T9	0.324	2.2
N10	1.432	1.6
T10	1.236	1.7
N11	1.672	1.6
T11	3.736	1.6
N12	0.252	2.5
T12	5.944	1.4

RNA integrity in one μl of each sample was assessed by agarose gel electrophoresis as previously outlined in section 2.2.2.1. Figure 5.2 shows the resulting gel. The gel was not loaded for equal amounts of RNA. The majority of samples had

visible 28 S and 18 S ribosomal bands. Although they were not dilute, 2N and 9T RNA were not visible on the gel. 5N RNA appeared to be degraded and 7T and 12N were not visible, although they were quite dilute.
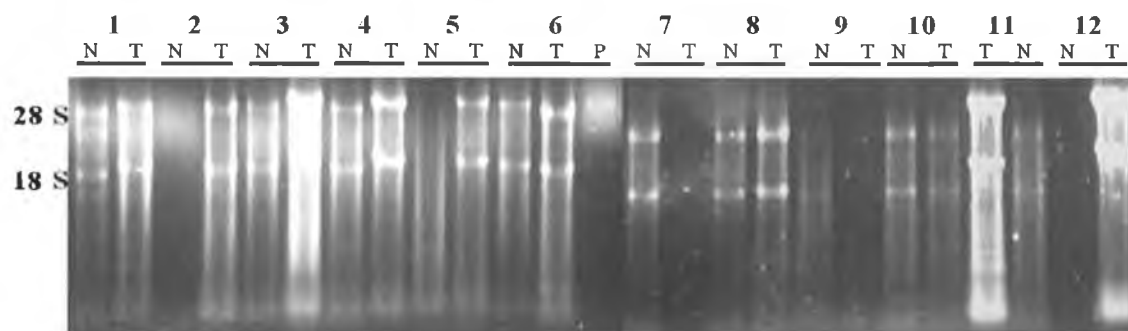
Figure 5.2 Analysis of RNA from paired normal (N) and tumour (T) tissue samples. The presence of 28 S and 18 S ribosomal bands was indicative of RNA integrity.

To further assess the quality of RNA and its usefulness in quantitative real time RT-PCR, a standard β-actin RT-PCR, as outlined in section 2.2.15, was carried out on all samples. A 550 bp PCR product was obtained in all samples except samples T7, T9 and N12. These samples were omitted from further analysis and comparisons, however the other samples in these pairings yielded PCR products, and were therefore included for real time PCR analysis. Although only faint bands were observed for the RNA from samples P6 and N9 (Figure 5.2), β-actin PCR on these samples yielded products (Figure 5.3).

Figure 5.3 β-Actin RT-PCR on paired normal (N) and tumour (T) samples. β-Actin expression was detected by the presence of a 550 bp product.

5.2.2 Real time PCR analysis

All real time PCR was carried out as outlined in section 2.2.4. As the Rotor-Gene 3000^{TM} (Corbett Research) multiplexing system used could only handle 36 samples at any one time, the study was divided in two: **Group one** included the matched normal and tumour paired samples 1, 2, 3, 4, and 5 and **Group two** included the matched normal and tumour paired samples 6, 8, 10, 11 plus N7, N9 and T12.

5.2.2.1 β-Actin expression

The comparative quantitation method was used to assess the levels of β-actin expression in all samples. β-actin levels were then used to normalise the expression levels of the genes studied in order to make inter sample comparisons. Figure 5.4 shows the comparative quantitation curve obtained for the β-actin real time PCR reaction for Group one tissue samples. The data obtained is listed in Table 5.3. To confirm the reaction conditions were specific for β-actin amplification a melt curve was performed following PCR. As can be seen from figure 5.5 each reaction resulted in only one peak indicating that only one product was formed.
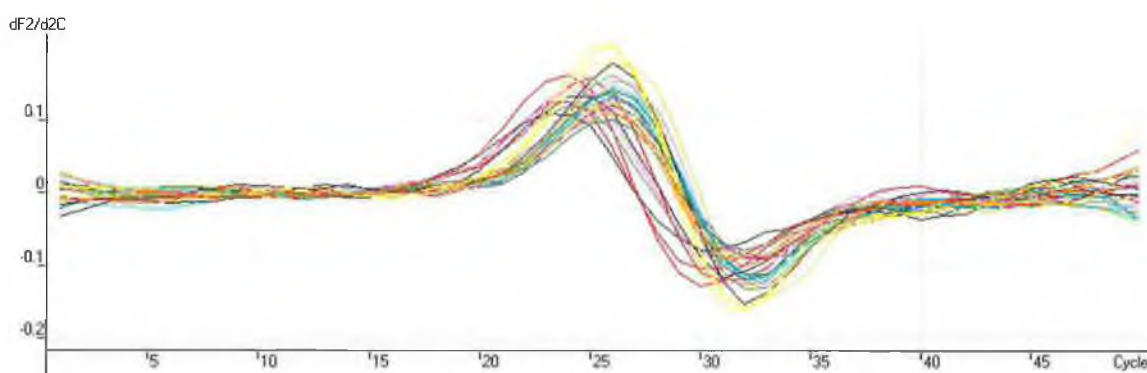
Figure 5.4 Comparative quantitation of β-actin expression in Group one paired tissue samples. The colours used and values obtained are summarised in Table 5.3.

Table 5.3 Raw real time PCR data obtained for β-actin expression in Group one normal (N) and tumour (T) tissue one samples.

Colour	Name	Takeoff	Comparative Concentration
	N1	20.6	1.00E+00
	N1	19.8	7.43E-01
	N1	19.5	7.88E-01
	T1	22.1	4.62E-01
	T1	22.3	5.52E-01
	T1	23.7	4.35E-01
	N2	22.3	7.31E-02
	N2	24.5	1.18E-01
	N2	22.7	8.73E-02
	T2	19.6	5.20E-01
	T2	20.0	2.87E-01
	T2	22.6	3.43E-01
	N3	22.0	5.11E-02
	N3	22.1	1.25E-01
	N3	22.3	2.99E-02
	T3	22.3	1.78E-01
	T3	21.9	5.43E-02
	T3	21.7	8.23E-02
	N4	22.1	2.87E-01
	N4	21.8	1.89E-01
	N4	20.0	1.49E-01
	T4	19.1	4.35E-01
	T4	22.1	5.52E-01
	T4	20.3	4.62E-01
	N5	22.1	6.01E-03
	N5	21.7	1.44E-03
	N5	21.5	3.73E-03
	T5	21.7	1.49E-01
	T5	22.0	9.84E-02
	T5	22.2	1.58E-01
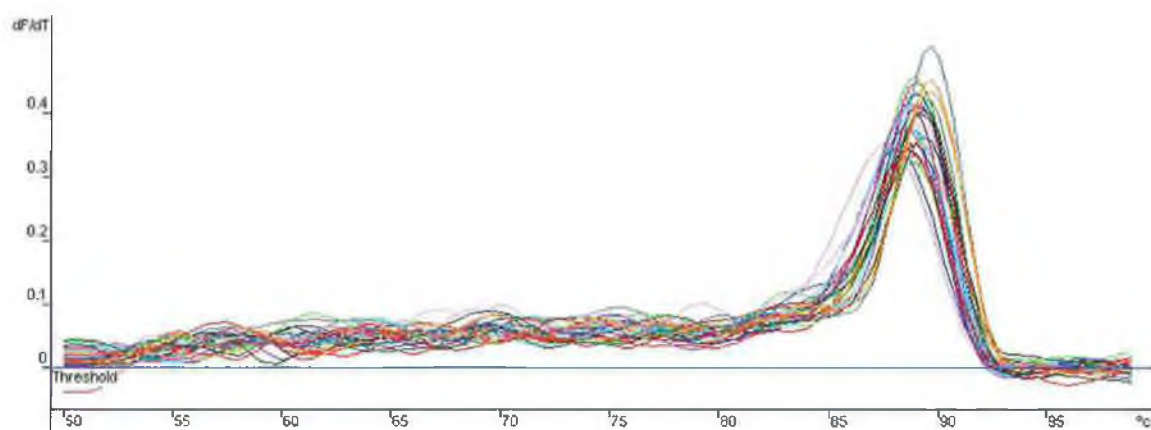
Figure 5.5 The melt curve for β-actin real time PCR of Group one tissue samples.

β-actin real time PCR was also carried out for Group two tissue samples. The comparative quantitation method was used and the resulting curve is shown in Figure 5.6.
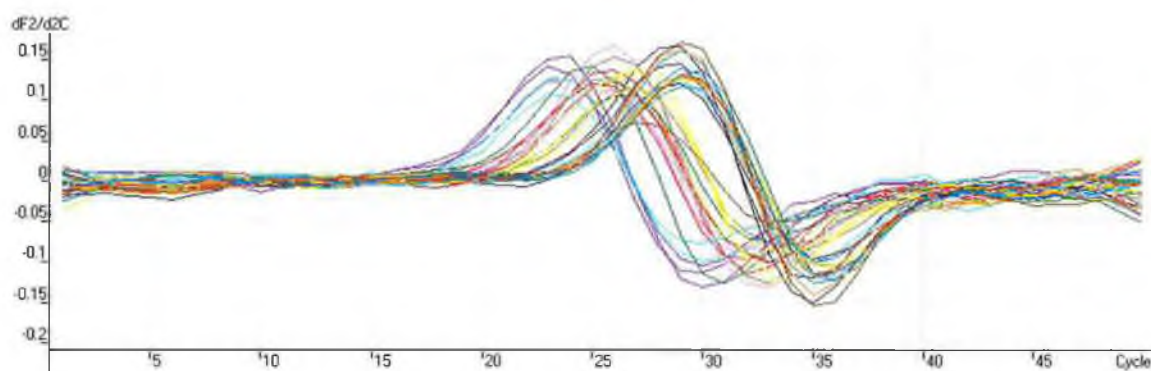
Figure 5.6 The resulting comparative quantitation curve for the real time β-actin PCR in Group two tissue samples. The data obtained is summarised in Table 5.4.
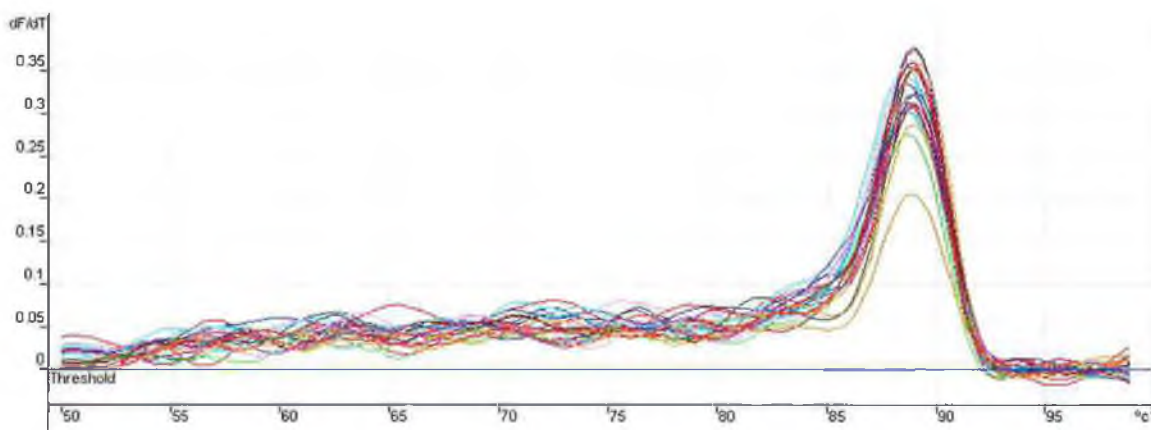
Figure 5.7 The melt curve for β-actin real time PCR of Group two tissue samples.

Table 5.4 Real time PCR data obtained for β-actin expression in Group two normal (N), polyp (P) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N6	21.4	1.00E+00
	N6	21.3	1.06E+00
	N6	22.2	6.15E-01
	T6	19.5	3.17E+00
	T6	18.9	4.56E+00
	T6	19.1	4.04E+00
	P6	24.9	6.15E-01
	P6	24.8	6.15E-01
	P6	22.2	6.15E-01
	N8	21.6	8.86E-01
	N8	22.3	5.79E-01
	N8	22.1	6.54E-01
	T8	24.9	2.99E+00
	T8	19.7	2.81E+00
	T8	19.5	3.17E+00
	N10	20.9	1.35E+00
	N10	22.6	1.13E+00
	N10	25.8	9.16E-01
	T10	20.9	1.35E+00
	T10	22.9	4.02E-01
	T10	22.5	5.13E-01
	N11	23.7	2.47E-01
	N11	24.8	1.27E-01
	N11	25.0	1.12E-01
	T11	22.6	4.82E-01
	T11	23.5	2.79E-01
	T11	23.8	2.33E-01
	N7	25.3	9.36E-02
	N7	24.7	1.35E-01
	N7	25.4	8.81E-02
	N9	25.5	8.29E-02
	N9	25.7	7.34E-02
	N9	25.6	7.80E-02
	T12	25.4	8.81E-02
	T12	25.4	8.81E-02
	T12	24.3	1.72E-01

As can be seen from the melt curve analysis of Group two tissue samples (Figure 5.7) only one peak occurred, this indicated that reaction was specific for β-actin amplification.

5.2.2.2 Cathepsin C expression

The expression of cathepsin C in the tissue samples was assessed using real time PCR analysis. As before the analysis was divided in two groups. Figure 5.8 shows the comparative quantitation curve obtained for the real time PCR analysis of cathepsin C expression in Group one tissue samples.
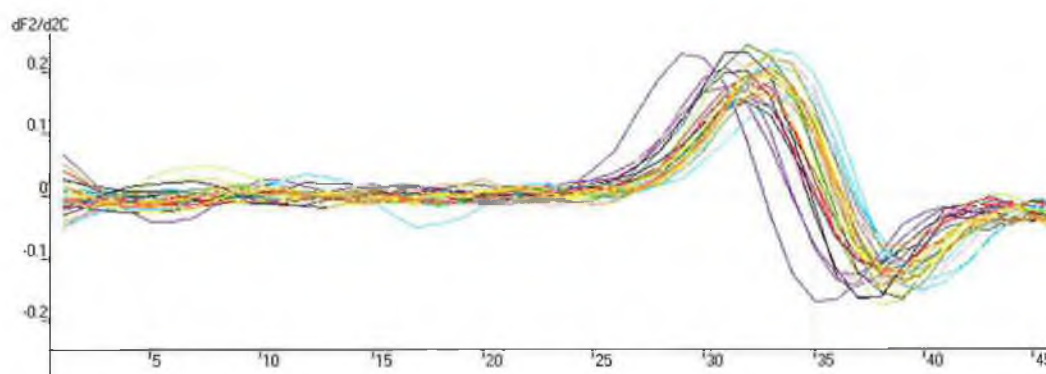
Figure 5.8 Comparative quantitation of cathepsin C expression in Group one paired tissue samples. The colours used and values obtained are summarised in Table 5.5.

The raw comparative quantitation data obtained for cathepsin C real time PCR is listed in Table 5.5. This data was later normalised using the real time β-actin data obtained in section 5.2.2.1.
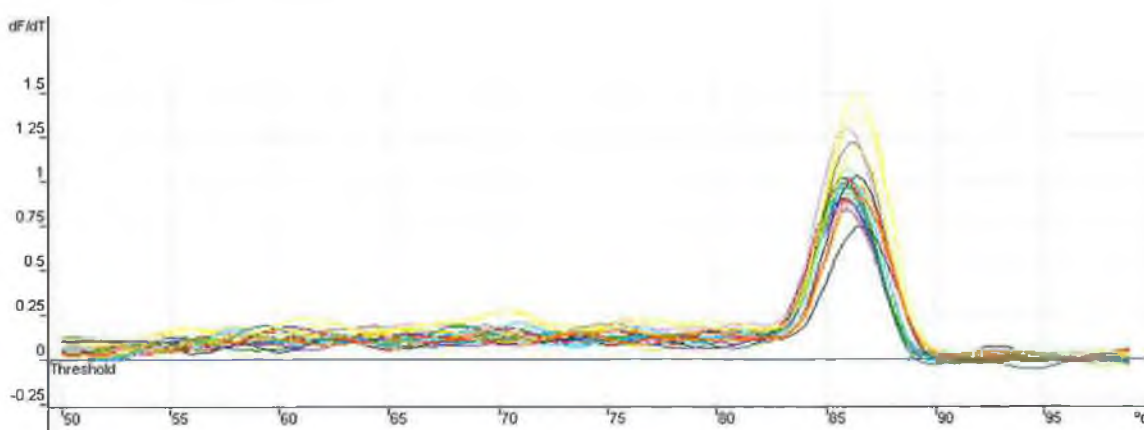
Figure 5.9 The melt curve for cathepsin C real time PCR of Group one tissue samples.

162

Following PCR, a melt curve was carried out to ensure specificity. The presence of only one peak indicated that no non-specific products were formed (Figure 5.9).

Table 5.5 Raw real time PCR data obtained for cathepsin C expression in Group one normal (N) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N1	28.5	1.00E+00
	N1	27.9	1.47E+00
	N1	28.2	1.21E+00
	T1	25.8	5.62E+00
	T1	27.5	4.20E+00
	T1	26.9	2.78E+00
	N2	28.5	7.74E-01
	N2	28.9	7.74E-01
	N2	28.4	7.74E-01
	T2	27.4	2.02E+00
	T2	28.8	8.26E-01
	T2	29.9	4.09E-01
	N3	29.8	4.36E-01
	N3	29.6	4.95E-01
	N3	29.5	5.28E-01
	T3	28.0	1.38E+00
	T3	28.4	1.07E+00
	T3	28.7	8.80E-01
	N4	29.3	6.00E-01
	N4	29.4	5.63E-01
	N4	28.8	8.26E-01
	T4	27.5	1.89E+00
	T4	27.6	1.78E+00
	T4	27.6	1.78E+00
	N5	29.3	6.00E-01
	N5	30.0	3.83E-01
	N5	28.4	1.07E+00
	T5	29.2	6.39E-01
	T5	28.9	7.74E-01
	T5	28.1	1.29E+00

Cathepsin C expression was also assessed in Group two tissue samples using real time PCR. The resulting comparative quantitation data is represented in Figure

5.10 and summarised in Table 5.6.   This data was later normalised using β-actin expression levels, thus allowing inter sample expression comparisons.
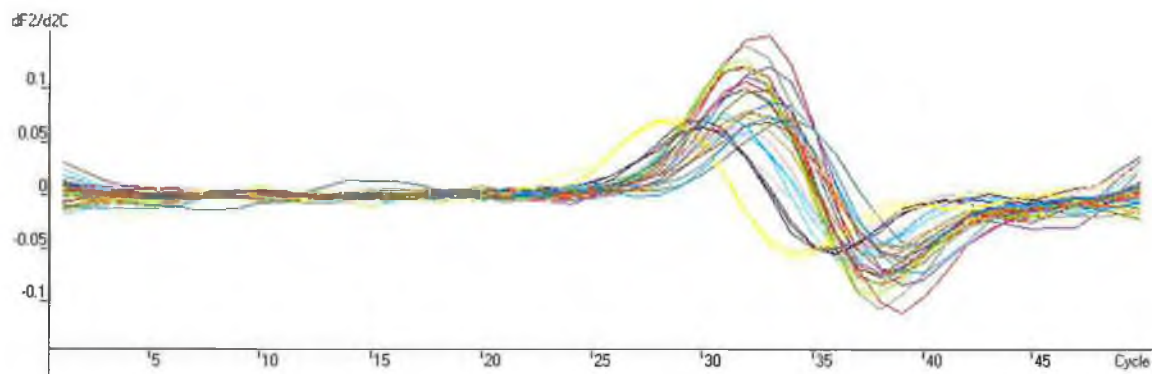
Figure 5.10 Comparative quantitation of cathepsin C expression in Group two paired tissue samples. The colours used and values obtained are summarised in Table 5.6.
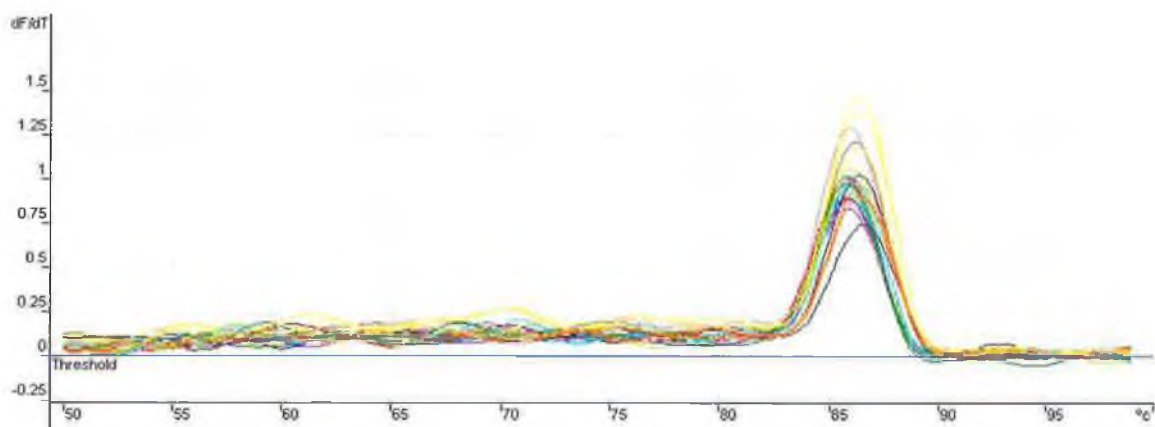
Figure 5.11 The melt curve for cathepsin C real time PCR of Group two tissue samples.

Only one peak was obtained from the melt curve analysis of the PCR reactions, indicating reaction specificity (5.11).

Table 5.6 Raw real time PCR data obtained for cathepsin C expression in Group two normal (N), polyp (P) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N6	28.0	1.00E+00
	N6	28.6	6.98E-01
	N6	28.9	5.83E-01
	T6	28.9	5.83E-01
	T6	28.2	5.17E-01
	T6	28.2	8.87E-01
	P6	28.0	1.00E+00
	P6	28.1	9.42E-01
	P6	27.9	1.06E+00
	N8	27.4	1.43E+00
	N8	26.7	2.18E+00
	N8	27.8	1.13E+00
	T8	27.2	1.62E+00
	T8	27.5	1.35E+00
	T8	27.3	1.52E+00
	N10	30.2	2.67E-01
	N10	27.9	3.83E-01
	N10	27.8	1.06E+00
	T10	28.4	1.13E+00
	T10	27.9	7.87E-01
	T10	27.9	1.06E+00
	N11	25.7	3.98E+00
	N11	26.6	2.32E+00
	N11	25.8	3.74E+00
	T11	25.2	5.37E+00
	T11	25.1	5.70E+00
	T11	24.5	8.17E+00
	N7	29.0	5.49E-01
	N7	27.4	1.43E+00
	N7	29.0	5.49E-01
	N9	29.3	4.58E-01
	N9	29.8	3.40E-01
	N9	30.1	2.84E-01
	T12	29.0	5.49E-01
	T12	28.4	7.87E-01
	T12	27.9	1.06E+00

β-actin expression levels were used to normalise the cathepsin C expression data. Figure 5.12 and 5.13 show the normalised cathepsin C expression values obtained for Group one and Group two tissue samples respectively.
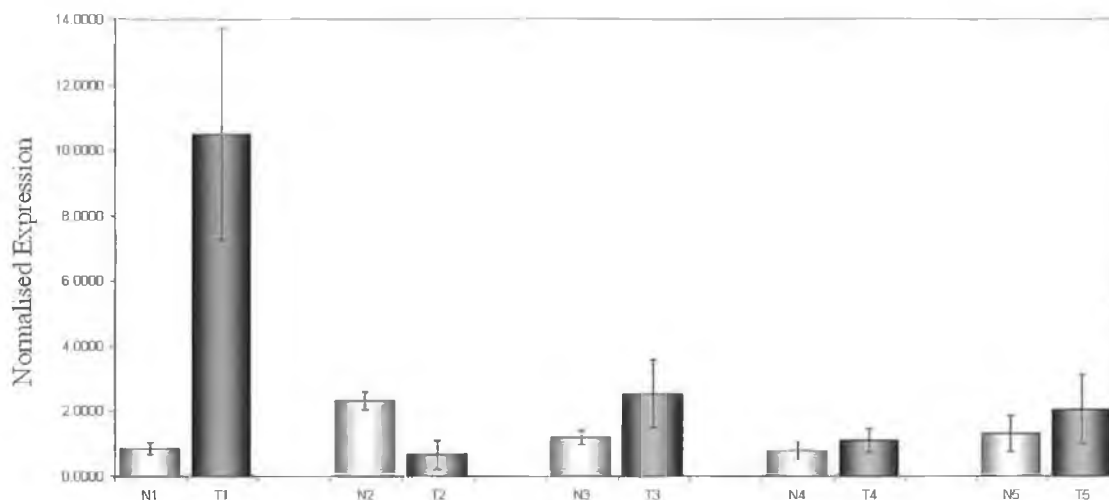
Figure 5.12 Normalised cathepsin C expression in Group one paired tissue samples.
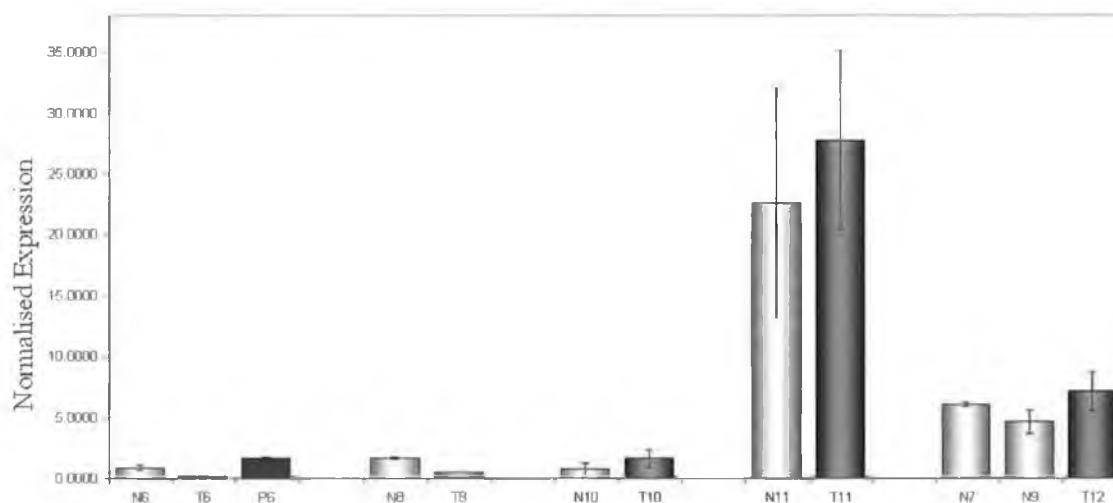
Figure 5.13 Normalised cathepsin C expression in Group two paired tissue samples.

Cathepsin C expression in tumours was compared with its expression in the normal samples. The comparative expression of cathepsin C is shown in Figure 5.14. Seven out of ten samples studied showed increased cathepsin C expression in tumour or

polyp tissue in comparison with normal tissue. Interestingly, the greatest increase in cathepsin C expression was observed in the Dukes' D tumour sample T1 although another Dukes' D tumour T11 did not show a large increase in expression.
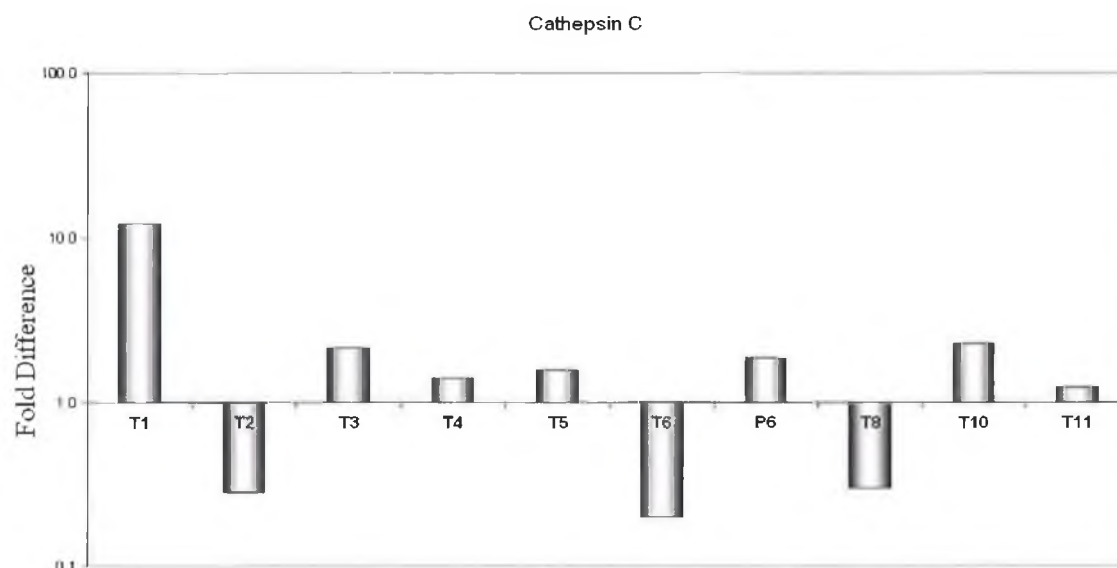
Figure 5.14 Expression of cathepsin C in tumours (T) and polyp (P) samples in comparison with normal samples. A log scale was used to display up regulation (bars above axis) and down regulation (bars below axis).

5.2.2.3 Caveolin 1 expression

The expression of caveolin 1 in the colon tissue samples was assessed using real time PCR. The analysis was again divided in two groups. Comparative quantitation was used to assess the expression of caveolin 1 in Group 1 tissue samples and the data obtained is shown in Figure 5.15 and summarised in Table 5.7.
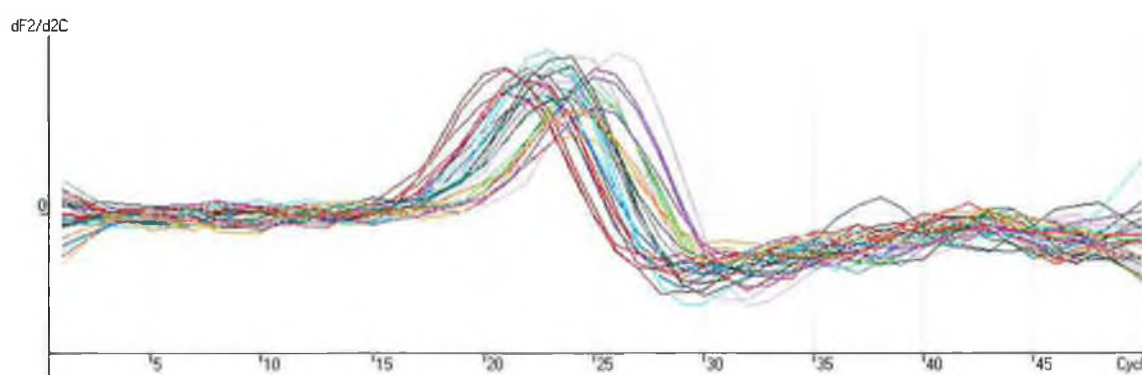
Figure 5.15 The expression of caveolin 1 in Group one paired tissue samples was determined using the comparative quantitation method. The colours used and values obtained are summarised in Table 5.7.

Table 5.7 Raw real time PCR data obtained for caveolin 1 expression in Group one normal (N) and tumour (T) tissue samples.

| Colour	Name	Takeoff	Comparative Concentration
	N1	17	1.00E+00
	N1	17.8	6.51E-01
	N1	17.1	9.48E-01
	T1	20.1	1.89E-01
	T1	20.5	1.53E-01
	T1	20.9	1.23E-01
	N2	18.4	4.72E-01
	N2	18.2	5.25E-01
	N2	17.7	6.87E-01
	T2	19.1	3.24E-01
	T2	18.3	4.98E-01
	T2	18.7	4.01E-01
	N3	22.1	6.47E-02
	N3	20.2	1.79E-01
	N3	18.1	1.22E-01
	T3	21.1	1.11E-01
	T3	21.1	1.11E-01
	T3	21	1.17E-01
	N4	17.4	8.07E-01
	N4	17.3	8.51E-01
	N4	16.6	1.24E+00
	T4	18.8	3.80E-01
	T4	19.5	2.61E-01
	T4	18.9	3.61E-01
	N5	19.7	2.35E-01
	N5	19.7	2.35E-01
	N5	20.2	1.79E-01
	T5	20.3	1.70E-01
	T5	20.8	1.30E-01
	T5	19.7	2.35E-01

Figure 5.16 shows that a melt curve analysis of the PCR reactions yielded only one peak indicating the reactions were specific.
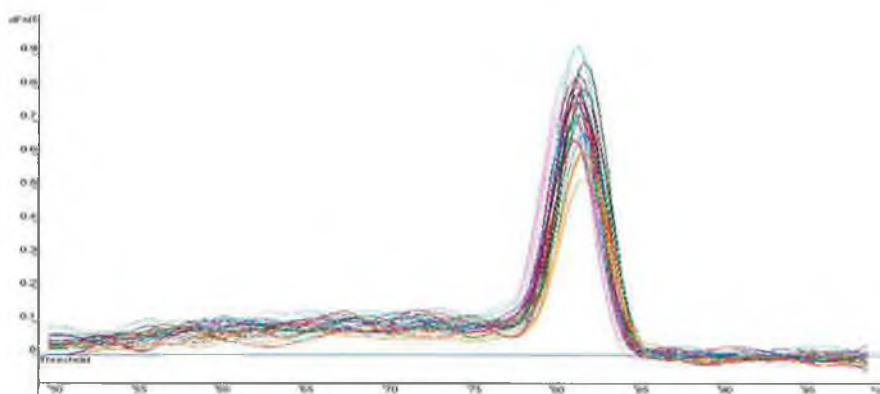
Figure 5.16 The melt curve for caveolin 1 real time PCR of Group one tissue samples.

Figure 5.17 shows the expression of caveolin 1 in Group two tissue samples determined using real time PCR. The data obtained is summarised in Table 5.8. The melt curve analysis of the PCR reactions yielded only one peak indicating the reactions were specific (Figure 5.16).
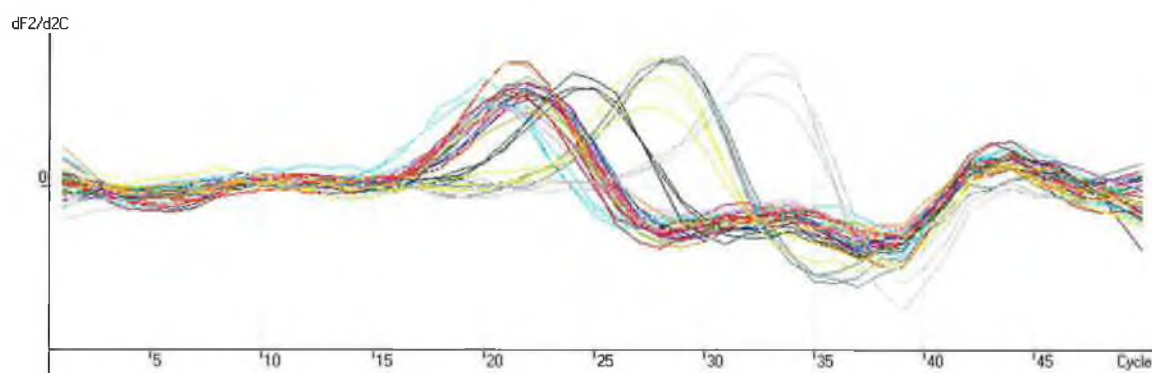
Figure 5.15 The expression of caveolin 1 in Group two paired tissue samples was determined using the comparative quantitation method. The colours used and values obtained are summarised in Table 5.8.

Table 5.8 Raw real time PCR data obtained for caveolin 1 expression in Group two normal (N), polyp (P) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N6	17.5	1.00E+00
	N6	17.3	1.11E+00
	N6	18.0	7.65E-01
	T6	17.9	8.07E-01
	T6	17.9	8.07E-01
	T6	18.2	6.87E-01
	P6	19.1	4.24E-01
	P6	18.4	6.17E-01
	P6	17.8	8.51E-01
	N8	15.4	3.08E+00
	N8	16.3	1.90E+00
	N8	16.2	2.01E+00
	T8	16.7	1.54E+00
	T8	17.6	9.48E-01
	T8	17.1	1.24E+00
	N10	18.3	6.51E-01
	N10	17.4	1.06E+00
	N10	17.5	1.00E+00
	T10	18.1	7.25E-01
	T10	17.5	1.00E+00
	T10	16.6	1.62E+00
	N11	21.0	1.53E-01
	N11	20.1	2.48E-01
	N11	20.5	2.00E-01
	T11	24.6	2.22E-02
	T11	24.8	1.99E-02
	T11	24.0	3.06E-02
	N7	17.7	8.98E-01
	N7	17.8	8.51E-01
	N7	18.3	6.51E-01
	N9	23.8	3.41E-02
	N9	24.0	3.06E-02
	N9	24.0	3.06E-02
	T12	28.7	2.46E-03
	T12	28.7	2.46E-03
	T12	26.9	6.47E-03
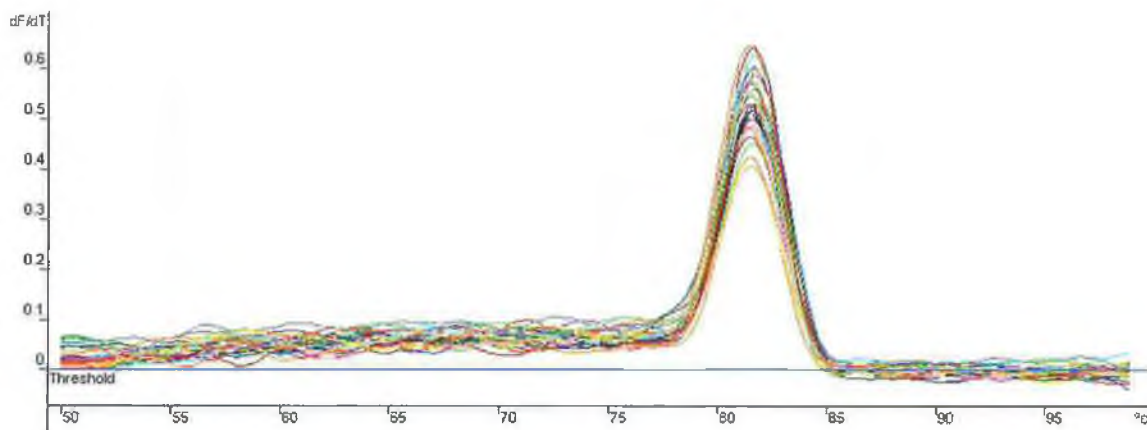
Figure 5.16 The melt curve for caveolin 1 real time PCR of Group two tissue samples.

The normalised levels of caveolin 1 expression in both groups of tissue samples were calculated using the relative β-actin expression levels as determined in section 5.2.2.1. Figures 5.17 and 5.18 show the normalised expression of caveolin 1 in Group one and two tissue samples respectively.
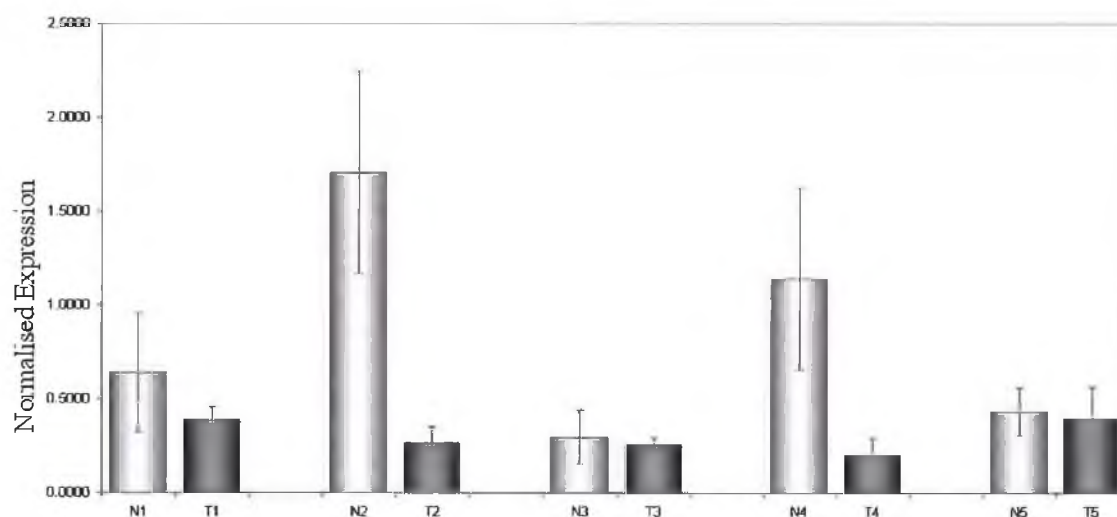
Figure 5.17 Normalised caveolin 1 expression in Group one paired normal (N) and tumour (T) tissue samples.
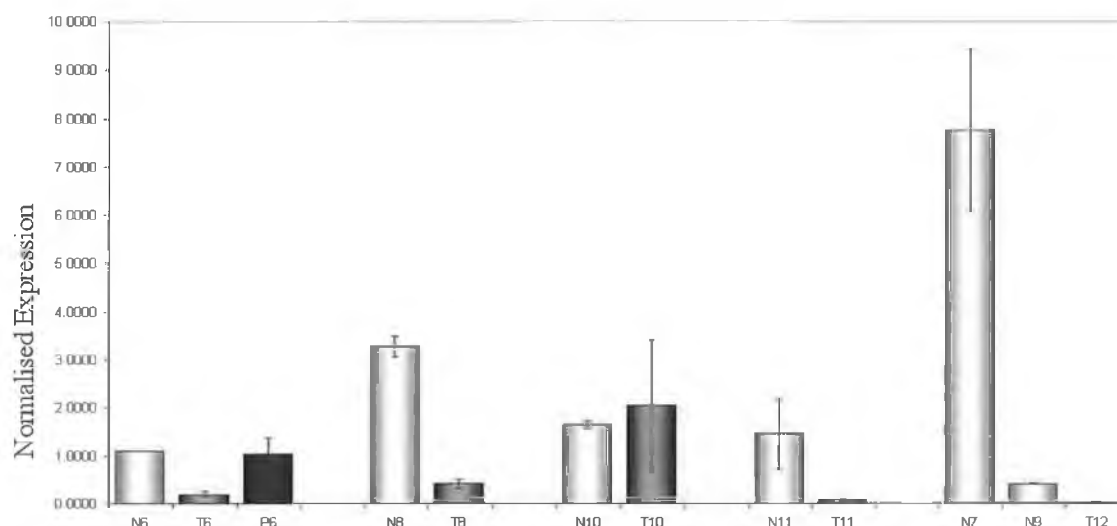
Figure 5.18 Normalised caveolin 1 expression in Group two paired normal (N), polyp (P) and tumour (T) tissue samples.

The level of caveolin 1 expression in tumours and polyp samples was compared with its expression in the normal tissue to which they were paired. Figure 5.19 shows the comparative expression of caveolin 1 in all paired tumour and polyp tissues studied. Interestingly, all tumour samples bar tumour sample 10, show decreased caveolin 1 expression in comparison with normal tissue.
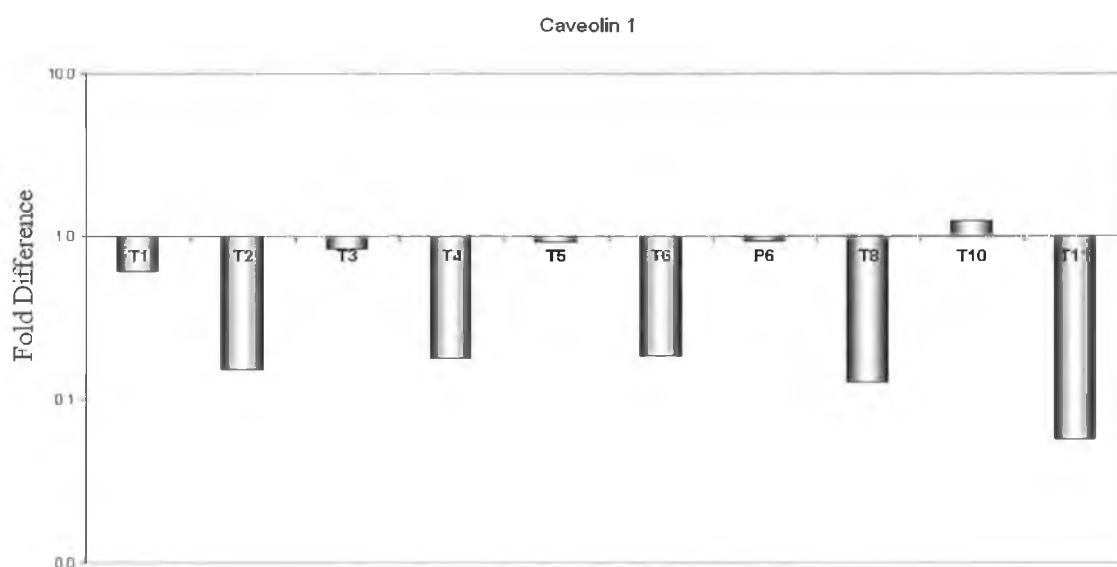
Figure 5.19 Expression of caveolin 1 in tumours (T) and polyp (P) samples in comparison with normal samples. A log scale was used to display up regulation (bars above axis) and down regulation (bars below axis).

5.2.2.4 Tenascin C expression

The expression of tenascin C in the colon tissue samples was determined using real time PCR. Figure 5.20 shows the comparative quantitation curve for tenascin C expression in Group one tissue samples. The raw data obtained is listed in Table 5.9.
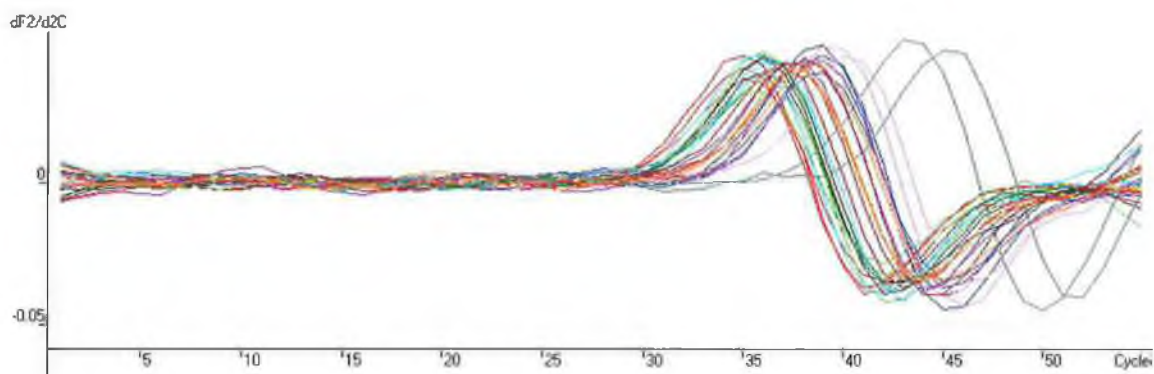
Figure 5.20 Comparative quantitation of tenascin C expression in Group one tissue samples. The colours used and values obtained are summarised in Table 5.9.

The melt curve analysis of the above PCR reactions yielded only one peak per reaction, thus indicating that the conditions were specific (Figure 5.21).
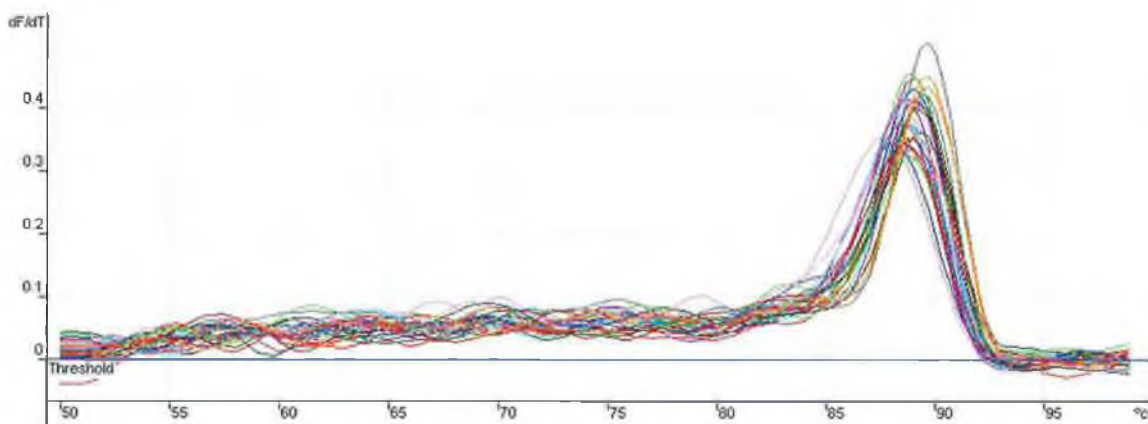
Figure 5.21 The melt curve for tenascin C real time PCR of Group one tissue samples.

Table 5.9 Raw comparative quantitation data obtained for tenascin C real time PCR in Group one normal (N) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N1	30.6	1.00E+00
	N1	31.1	7.43E-01
	N1	31.0	7.88E-01
	T1	31.9	4.62E-01
	T1	31.6	5.52E-01
	T1	32.0	4.35E-01
	N2	35.0	7.31E-02
	N2	34.2	1.18E-01
	N2	34.7	8.73E-02
	T2	31.7	5.20E-01
	T2	32.7	2.87E-01
	T2	32.4	3.43E-01
	N3	35.6	5.11E-02
	N3	34.1	1.25E-01
	N3	36.5	2.99E-02
	T3	33.5	1.78E-01
	T3	35.5	5.43E-02
	T3	34.8	8.23E-02
	N4	32.7	2.87E-01
	N4	33.4	1.89E-01
	N4	33.8	1.49E-01
	T4	32.0	4.35E-01
	T4	31.6	5.52E-01
	T4	31.9	4.62E-01
	N5	39.2	6.01E-03
	N5	41.6	1.44E-03
	N5	11.3	3.73E-03
	T5	33.8	1.49E-01
	T5	34.5	9.84E-02
	T5	33.7	1.58E-01

Similarly, real time PCR was used to examine the expression of tenascin C in Group two tissue samples. Figure 5.22 shows the resulting comparative quantitation curve and Table 5.10 summarises the raw data obtained. Melt curve analysis was carried out following PCR. One peak indicated that only one product was formed (Figure 5.23).
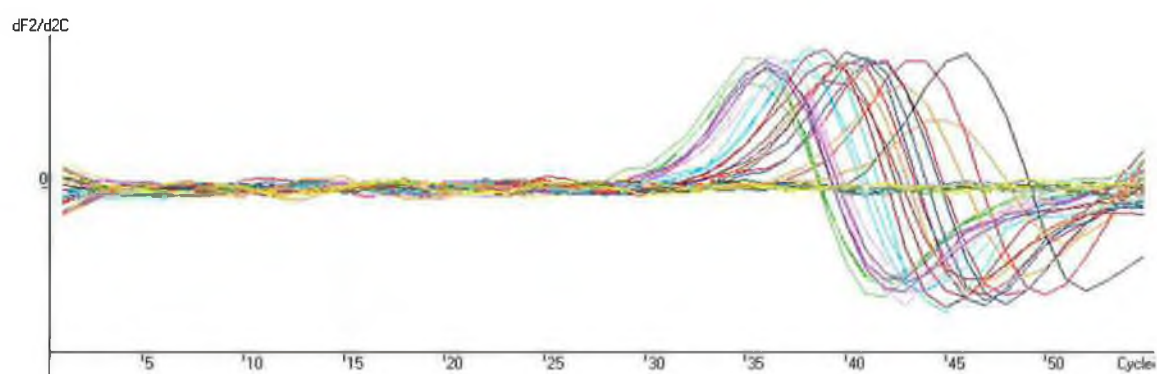
Figure 5.22 Comparative quantitation of tenascin C expression in Group two tissue samples. The colours used and values obtained are summarised in Table 5.10.
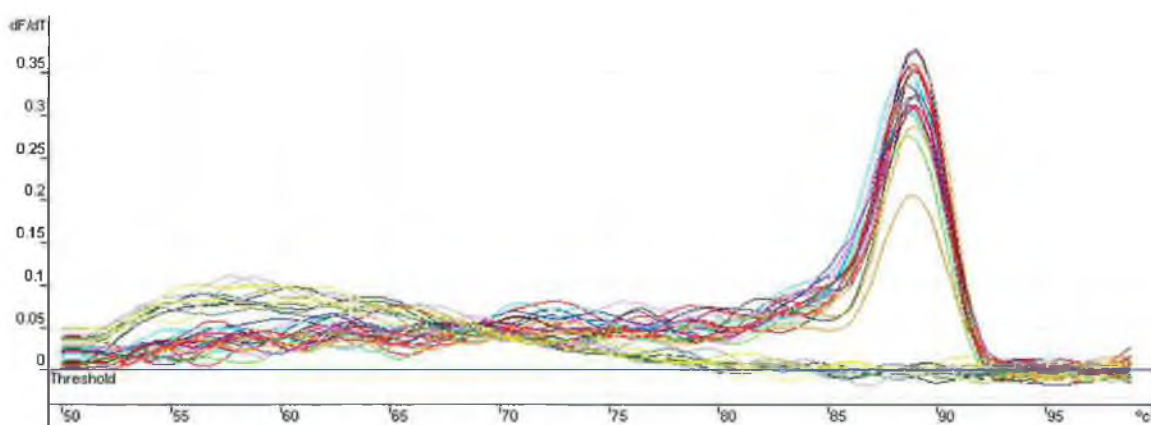
Figure 5.23 The melt curve for tenascin C real time PCR of Group two tissue samples.

Table 5.10 Raw comparative quantitation data obtained for tenascin C real time PCR in Group two normal (N), polyp (P) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N6	39.8	1.00E+00
	N6	37.5	1.60E+00
	N6	35.9	2.22E+00
	T6	31.3	5.69E+00
	T6	31.4	5.58E+00
	T6	31.9	5.03E+00
	P6	36.2	2.09E+00
	P6	37.6	1.57E+00
	P6	36.4	2.00E+00
	N8	34.0	3.28E+00
	N8	34.0	3.28E+00
	N8	34.4	3.02E+00
	T8	32.6	4.36E+00
	T8	32.8	4.19E+00
	T8	33.2	3.86E+00
	N10	32.5	4.45E+00
	N10	32.1	4.83E+00
	N10	32.2	4.73E+00
	T10	34.5	2.96E+00
	T10	34.9	2.72E+00
	T10	35.1	2.62E+00
	N11	44.1	6.78E-01
	N11	48.1	6.78E-01
	N11	41.7	6.78E-01
	T11	22.1	2.51E-01
	T11	44.9	3.52E-01
	T11	49.1	1.49E-01
	N7	38.6	1.28E+00
	N7	39.8	1.00E+00
	N7	36.0	2.18E+00
	N9	48.1	1.83E-01
	N9	45.9	2.87E-01
	N9	13.1	2.35E-01
	T12	48.1	1.83E-01
	T12	33.9	2.00E-01
	T12	24.3	1.92E-01

Tenascin C expression was normalised using β-actin expression levels. Figures 5.24 and 5.25 show the resulting normalised expression in Group one and Group two tissue samples respectively.
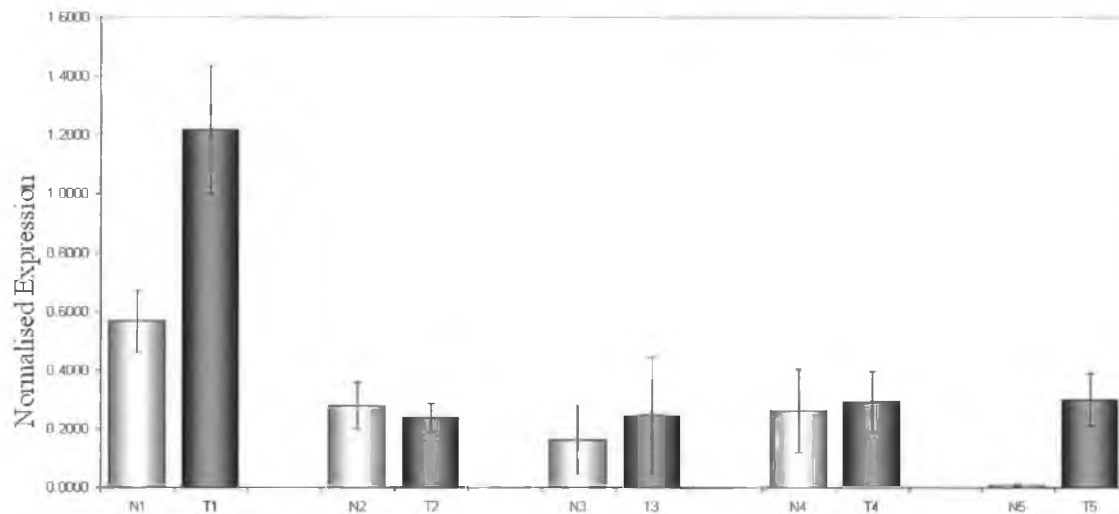
Figure 5.24 Normalised tenascin C expression in Group one paired normal (N) and tumour (T) tissue samples.
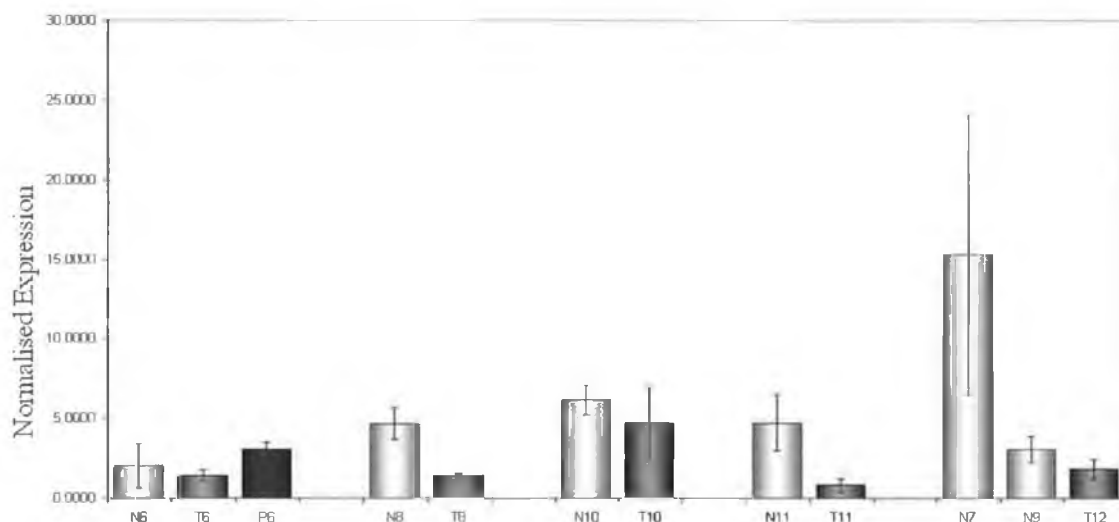
Figure 5.25 Normalised tenascin C expression in Group two paired normal (N), polyp (P) and tumour (T) tissue samples.

The expression of tenascin C in the tumour and polyp tissue samples was compared with its expression in the respective normal tissue sample. Figure 5.26 shows the comparative expression of tenascin C in all paired samples studied. There was no

clear pattern of expression. Five out the ten tumour samples show increased tenascin C expression and the other five tumour samples show decreased expression in comparison with their normal tissue samples. These changes in expression do not seem to be related to Dukes' staging as some Dukes' B tumour samples show increased tenascin C expression (T5) and other Dukes' B samples show decreased expression (T6).
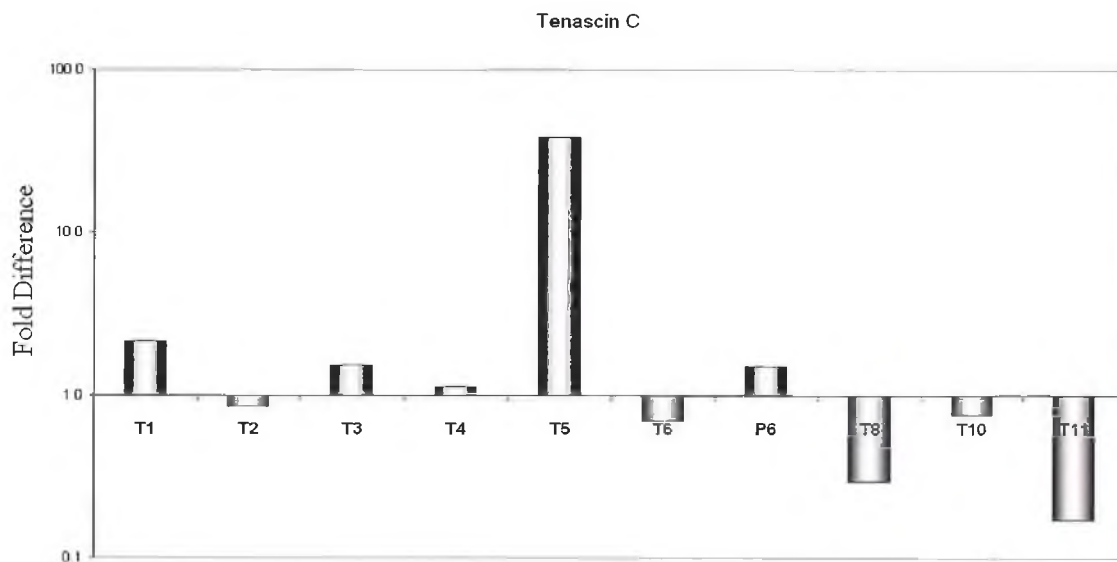
Figure 5.26 Expression of tenascin C in tumours (T) and polyp (P) samples in comparison with normal samples. A log scale was used to display up regulation (bars above axis) and down regulation (bars below axis).

5.2.2.5 PPARγ expression

The expression of PPARγ in paired colon tissue samples was analysed using quantitative real time PCR.
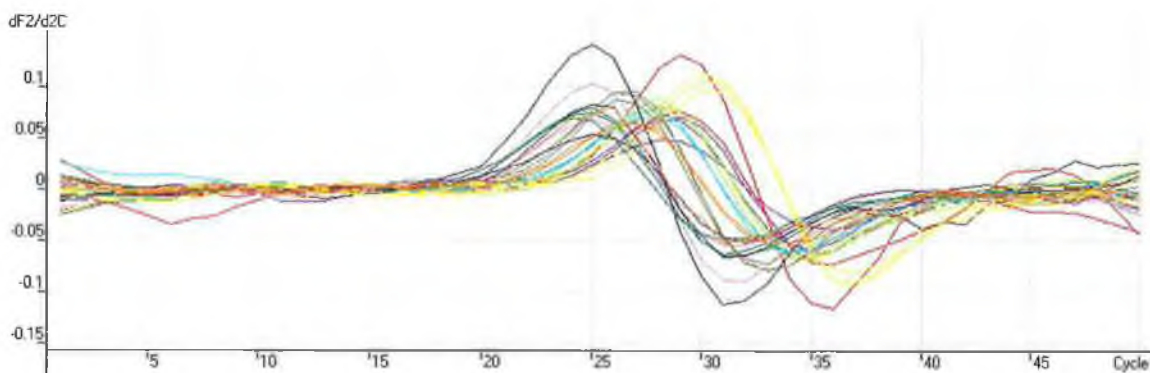
Figure 5.27 The expression of PPARγ in Group one tissue samples. The data obtained and colours used are summarised in Table 5.11.

178

PPARγ expression in Group one tissue samples was determined using the comparative quantitation method (Figure 5.27). The raw data obtained is summarised in Table 5.11.

Table 5.11 Raw comparative quantitation data obtained for PPARγ real time PCR in Group one normal (N), and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N1	24.8	1.00E+00
	N1	24.2	1.47E+00
	N1	25.0	8.79E-01
	T1	23.9	1.79E+00
	T1	24.9	9.37E-01
	T1	24.0	1.68E+00
	N2	23.9	1.79E+00
	N2	24.0	1.68E+00
	N2	24.3	1.38E+00
	T2	21.0	1.16E+01
	T2	21.1	1.09E+01
	T2	21.4	8.98E+00
	N3	23.5	2.31E+00
	N3	23.8	1.91E+00
	N3	23.7	2.03E+00
	T3	20.6	1.50E+01
	T3	20.3	1.83E+01
	T3	20.7	1.41E+01
	N4	22.8	3.64E+00
	N4	23.1	3.00E+00
	N4	21.9	6.50E+00
	T4	20.9	1.24E+01
	T4	20.6	1.50E+01
	T4	21.5	8.42E+00
	N5	25.7	5.59E-01
	N5	25.7	5.59E-01
	N5	26.7	2.93E-01
	T5	22.8	3.64E+00
	T5	22.5	4.41E+00
	T5	21.9	6.50E+00

Following PCR, melt curve analysis was carried out to ensure the reaction was specific, this was shown by the presence of only one peak in all reactions (Figure 5.28).
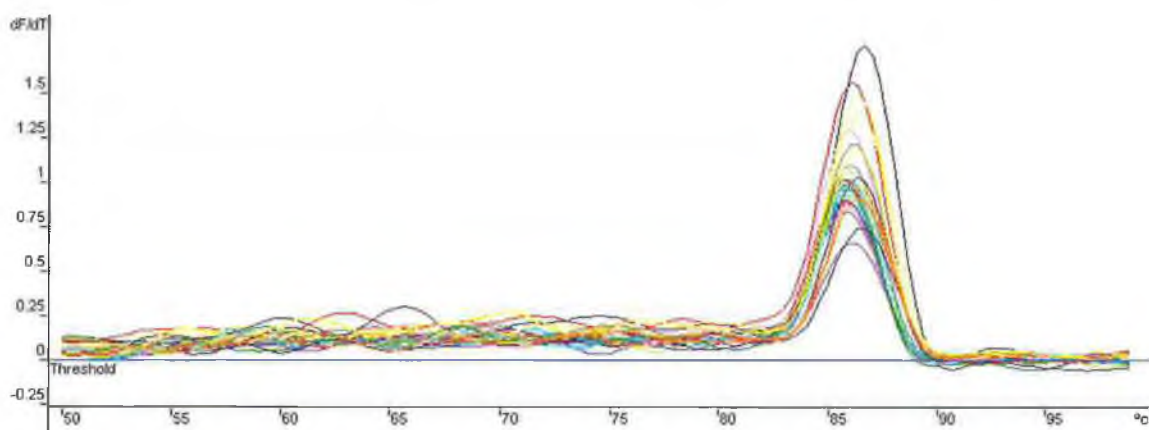
Figure 5.28 The melt curve for PPARγ real time PCR of Group one tissue samples.

The expression of PPARγ in Group two tissue samples was assessed using the same quantitative PCR (Figure 5.29). The raw data obtained is summarised in Table 5.12. Melt curve analysis was carried out to ensure the reaction was specific (Figure 5.30). No secondary peaks were detected.
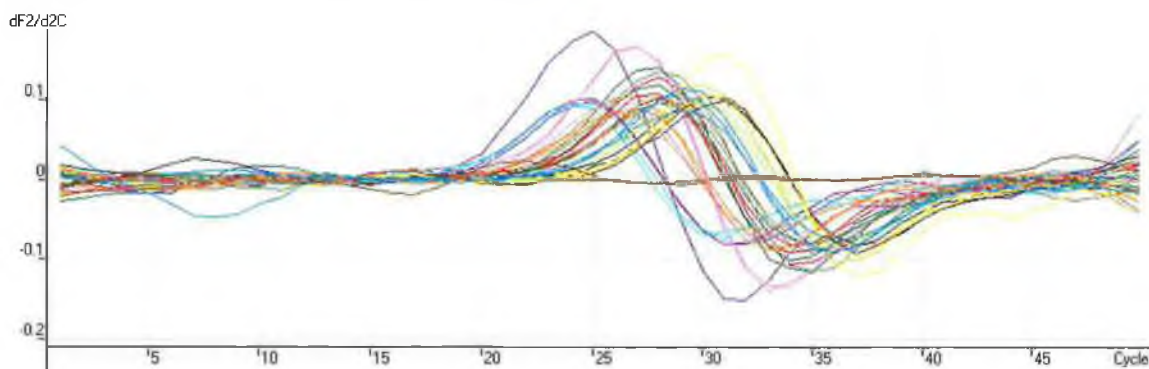
Figure 5.29 The expression of PPARγ in Group two tissue samples. The data obtained and colours used are summarised in Table 5.12.
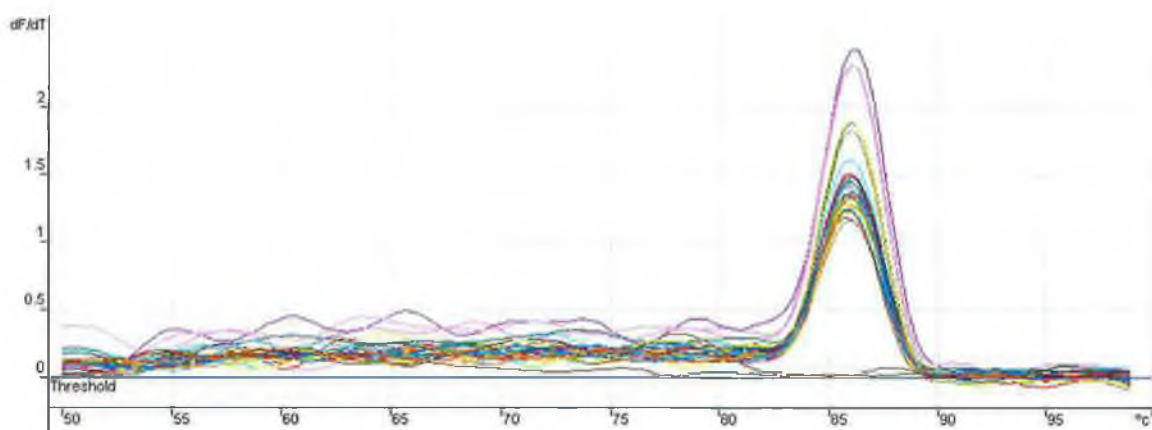
Figure 5.30 The melt curve for PPARγ real time PCR of Group two tissue samples.

Table 5.12 Raw comparative quantitation data obtained for PPARγ real time PCR in Group two normal (N), Polyp(P) and tumour (T) tissue samples.

Colour	Name	Takeoff	Comparative Concentration
	N6	24.1	1.00E+00
	N6	24.1	1.00E+00
	N6	23.5	1.51E+00
	T6	20.8	9.72E+00
	T6	20.7	1.04E+01
	T6	20.8	9.72E+00
	P6	24.9	5.76E-01
	P6	25.3	4.37E-01
	P6	25.6	3.56E-01
	N8	22.4	3.23E+00
	N8	22.5	3.01E+00
	N8	22.7	2.62E+00
	T8	20.4	1.28E+01
	T8	20.7	1.04E+01
	T8	19.6	1.16E+01
	N10	23.2	1.86E+00
	N10	24.1	1.00E+00
	N10	24.0	1.07E+00
	T10	24.0	1.07E+00
	T10	23.8	1.23E+00
	T10	24.2	9.33E-01
	N11	25.8	3.10E-01
	N11	26.5	1.91E-01
	N11	26.7	1.67E-01
	T11	26.5	1.91E-01
	T11	26.4	2.05E-01
	T11	26.9	1.45E-01
	N7	22.3	3.46E+00
	N7	21.8	4.88E+00
	N7	22.7	2.62E+00
	N9	25.1	5.02E-01
	N9	24.9	5.76E-01
	N9	24.4	8.13E-01
	T12	19.1	2.17E-03
	T12	33.0	2.17E-03
	T12	44.7	2.17E-03

The normalised expression of PPARγ in all tissues studied was calculated using the respective levels of β-actin expression determined in section 5.2.2.1. Figures 5.31 and 5.32 show the normalised PPARγ expression in Group one and Group two tissue samples respectively.
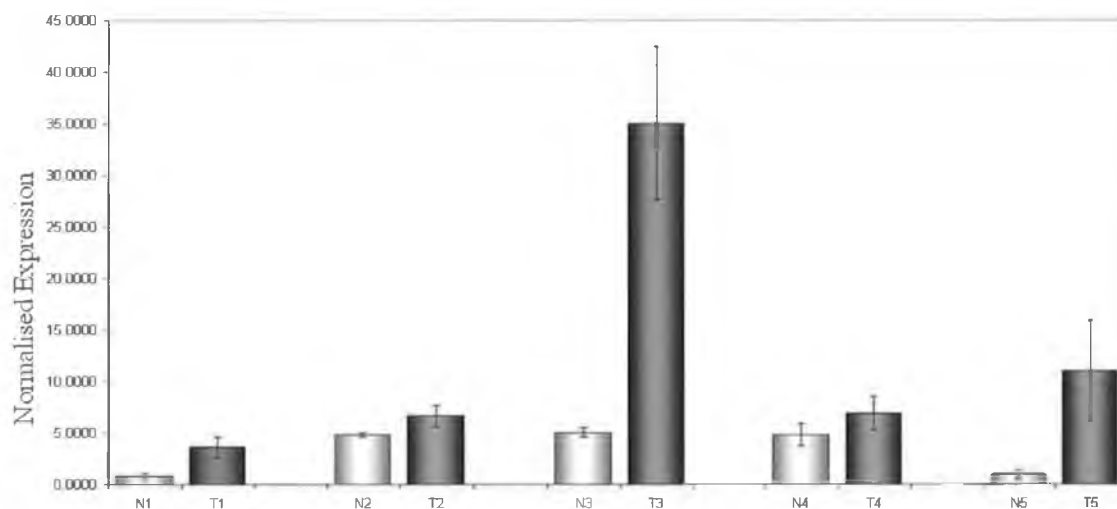
Figure 5.31 Normalised PPARγ expression in Group one paired normal (N) and tumour (T) tissue samples.
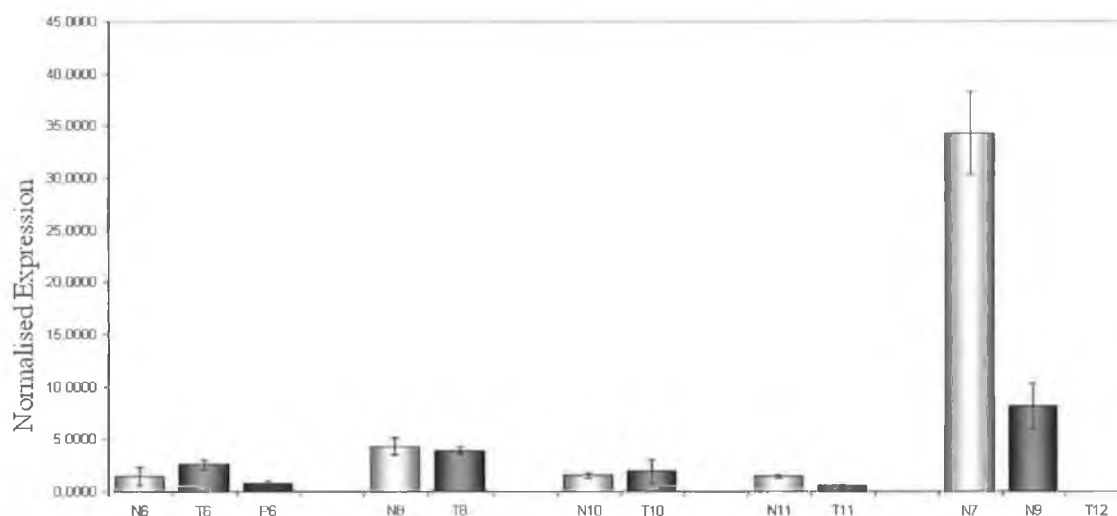
Figure 5.32 Normalised PPARγ expression in Group two paired normal (N), polyp (P) and tumour (T) tissue samples.

The expression of PPARγ in paired samples was compared and the comparative expression is shown in Figure 5.33. Interestingly, seven out of the ten tumour samples

studied showed increased PPARγ expression in comparison with normal tissue and three showed decreased expression. The differences in expression did not appear to be related to Dukes' staging as Dukes' D tumour samples showed both increased PPARγ expression (T1) and decreased expression (T11) in comparison with normal tissue.
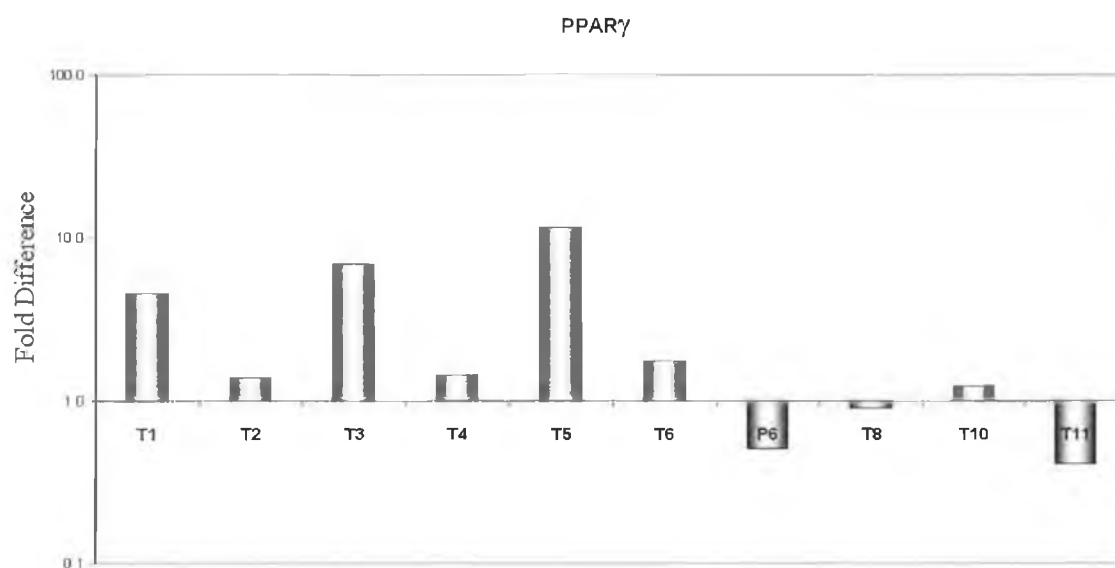
Figure 5.33 Comparative expression of PPARγ between tumour(T) or polyp (P) tissue and normal colon tissue samples. A log scale was used to display up regulation (bars above axis) and down regulation (bars below axis).

5.3 Discussion

The identification of novel biomarkers of colorectal tumour metastasis is crucial to both the understanding and treatment of this disease. As described in chapter 4, cDNA microarray analysis was used to identify genes with significantly increased expression in a lymph node metastases derived cell line when compared to the primary colon adenocarcinoma derived cell line. These selected genes have been identified as potential markers of colorectal metastasis. These genes were cathepsin C, tenascin C, PPARγ and caveolin 1. To further validate the importance of these genes in the process their expression was monitored in a pilot study of paired normal and tumour colorectal biopsy specimens. RNA was extracted from samples and analysed for expression of these genes using RT and quantitative real time PCR. Following RNA analysis, nine paired samples were chosen for further scrutiny. The unpaired samples N7, N9 and T12 were also analysed for expression of these genes.

Although the absolute levels of expression are of importance, the comparative expression of the above genes was also of interest. Quantitative real time PCR analysis allowed comparisons of gene expression to be made from one sample to another. Expression in tumours was compared with expression in the paired normal specimens.

Table 5.13 Expression ratios of selected genes in paired colorectal tissue. Increasing (red) or decreasing (green) expression of cathepsin C (CathC), PPAR, tenascin C (TenC) and caveolin 1 (Cav1) in tumour (T) or polyp (P) tissue samples in comparison with normal tissue.

Sample	Cath C	PPAR	TenC	Cav1
T1	12.2	4.6	2.2	0.6
T2	0.3	1.4	0.6	0.2
T3	2.1	7.0	1.5	0.8
T4	1.4	1.4	1.1	0.2
T5	1.6	11.5	38.7	0.9
T6	0.2	1.8	0.7	0.2
P6	1.9	0.5	1.5	0.9
T8	0.3	0.9	0.3	0.1
T10	2.3	1.2	0.8	1.3
T11	1.2	0.4	0.2	0.1

Table 5.13 summarises the trend in expression of all four genes in the nine paired samples. Sample set six also included a polyp tissue specimen. Table 5.13 shows the expression ratio of tumour/polyp : normal tissue. The most obvious trend is that caveolin 1 expression is decreased in all tumour samples in comparison with the normal tissue, with the exception of T10 where caveolin 1 expression is slightly increased in the tumour sample. Reduced caveolin 1 protein levels have been observed in a number of human tumours including ovarian (Bagnoli *et al.*, 2000), colon (Bender *et al.*, 2000) and lung (Racine *et al.*, 1999). A reduction in caveolin 1 expression has also been associated with increased breast tumour cell invasion (Lu *et al.*, 2004).

Cathepsin C expression was increased in tumour specimens T1, T3, T4, T5, T10 and T11 in comparison with normal tissues. Cathepsin C expression was also elevated in the polyp sample (P6) but not in the corresponding tumour sample (T6). Although the roles of cathepsin C in colorectal tumour progression have not been fully investigated, elevated cathepsin C activity has been associated with the tumourigenesis of pancreatic B-cell carcinoma in RIP1-Tag2 mice (Joyce *et al.*, 2004). Other cathepsins have been studied in colorectal cancer and increased cathepsin B expression has been associated with premalignant colorectal lesions where it is thought to be involved in ECM degradation (Murnane *et al.*, 1991, Emmert-Buck *et al.*, 1994 and Khan *et al.*, 1998b).

PPARγ expression was increased in tumour samples T1, T2, T3, T4, T5, T6 and T10 and decreased in T8 and T11 in comparison with their normal tissue samples. Interestingly, PPARγ expression was decreased in the polyp tissue sample P6 where its expression was increased in the tumour sample T6 from the same patient. While the expression of PPARγ in colonic mucosa has been previously reported, its roles in colon cancer remain undefined (Mansen *et al.*, 1996).

Tenascin C expression was enhanced in tumour samples T1, T3, T4 and T5 and polyp sample P6 in comparison with its expression in normal tissue. Tenascin C expression was decreased in tumour samples T2, T6, T8, T10 and T11. Increased tenascin C expression has been associated with poor prognosis in colorectal and breast cancer (Dueck *et al.*, 1999, Kressner *et al.*, 1997 and Ishihara *et al.*, 1995). In this study, tumour samples with increased tenascin C expression also had increased cathepsin C expression.

Tumour tissue T1, T3, T4 and T5 had the same trends in expression and while T1 was classified as Dukes' D, T3, T4 and T5 were classified as Dukes' B. T2 and T6 were classified as Dukes' B and have the same expression trends. Interestingly, tumour sample T8, which was classified as Dukes' A, was the only tumour sample in this study with decreased expression of all four genes.

5.4 Conclusion

In this study the expression of four previously identified potential markers of colorectal metastasis was investigated in paired normal and tumour colorectal specimens. All Dukes' classifications were represented and expression was monitored using real time PCR. The expression of PPARγ, tenascin C, cathepsin C and caveolin 1 was detected in all tissue samples studied. As this was a pilot study, expression in only a small number of paired tissue samples was investigated. No conclusive correlation could be made regarding expression trends, patient survival or tumour classification. The only notable trend was that caveolin 1 expression was decreased in all tumour tissue specimens in comparison with normal tissue, with the exception of tumour sample T10 where caveolin 1 expression was slightly elevated. Expression of all genes was both increased and decreased in tumour tissue in comparison with normal tissue. No correlation could be made between expression and Dukes' staging as tumours with the same Dukes' classification showed opposite trends in expression. Although this study is limited in terms of conclusions it may exemplify the limitations of the Dukes' staging system, which groups together heterogeneous subsets of patients with diseases that often behave differently to identical therapies. This highlights the need for the development of new and effective prognostic markers and parameters for colorectal cancer.

A larger number of samples would be required to conclusively determine expression patterns. This study does not give information on the location of expression. Analysis was carried out on crude heterogeneous tissue samples. In future, immunohistochemical studies would clarify which cells were producing the proteins and also in which cellular location these proteins were being expressed.

Chapter 6

Final conclusions and summary

6.1 Final conclusions and summary

While colorectal cancer is preventable if detected early, almost 50 % of patients with the disease develop metastasis within 5 years (Wagenaar-Miller *et al.*, 2004). The ability of tumour cells to invade and spread to distant sites of the body is the cause of 90 % of cancer deaths (Sporn, 1996). As metastasis is the cause of the high mortality rates associated with cancer, a better understanding of the process is required to either develop novel therapeutic strategies or improve on existing ones with the ultimate aim of ensuring longer patient survival. The key therefore, is to extend our knowledge of the metastatic process by identifying new interactions between existing markers and identifying novel biomarkers of the disease.

The cellular and molecular steps required for metastasis are generally similar for all solid tumour types and involve many key players including proteases and CAMs (Liotta and Kohn, 2003). The metastatic process involves intricate interplay between these molecules resulting in altered cell adhesion, proteolysis and migration (Bogenrieder and Herlyn, 2003). The ECM provides mechanical support for tumour cell migration and ECM remodelling is essential for tumour invasion (Chang and Werb, 2001). Included in the many groups of molecules involved in metastasis are the MMPs. Enhanced MMP activity and therefore increased ECM turnover have been shown to aid tumour locomotion (Liotta *et al.*, 1980 and Stetler-Stevenson *et al.*, 1996). As well as degrading the ECM, MMPs can cleave and therefore activate latent growth factors as well as activating other MMPs (Wagenaar-Miller *et al.*, 2004 and Yu and Stamenkovic, 2000). MMPs contribute to all stages of tumour progression, including local area invasion, intravasation, extravasation and growth at distant sites (Chambers and Matrisian, 1997 and Sternlicht and Werb, 2001).

To investigate the effects of enhanced MMP activity on tumour cell behaviour, a colon adenocarcinoma cell line, SW480, which did not produce detectable amounts of MMP-9 or MMP-7, was stably transfected with the cDNAs for these MMPs. MMP overexpression was detected using real time PCR, zymography and western blot analysis. The transfected cells showed increased *in vitro* invasion through reconstituted basement membrane in comparison with the non-transfected parental cell line. The MMP-7 transfected cell line, SW480M7 was more invasive than the MMP-9 transfected cell line SW480M9. MMP inhibition has previously been shown to suppress tumour

invasion (Agrez *et al.*, 1999). Using *in vitro* migration assays, the MMP transfected cell lines were also shown to be more migratory than the parental SW480 cell line. Interestingly, the SW620 cells derived from a lymph node metastasis of the same patient from which the SW480 cell line was established, were more invasive yet less migratory than the parental SW480 cell line. The SW620 produce low levels of MMP-9 and undetectable amounts of MMP-7 and it was not as invasive or migratory as the MMP transfected cells.

The function of CAMs are crucial to metastatis, as tumour cells must undergo changes in their adhesion properties throughout the process. CAMs with established roles in metastasis include the integrins and CD44. The integrins are a large family of adhesion molecules with roles in both cell-cell and cell-matrix interactions. Integrin-mediated signalling also occurs via these interactions. CD44 is another adhesion molecule with roles in both cell adhesion and signalling. CAM mediated signalling often requires successful adhesion to either ECM components or neighbouring cells (Rosales *et al.*, 1995).

In this study cell adhesion assays were carried out to compare the adhesive properties of the cell lines and also to investigate the role of cell adhesion in the invasive process. Cell adhesion to plastic, HA and fibronectin was compared. The SW480 cells were equally as adhesive to plastic, HA and fibronectin. The SW620 cells were more adhesive to plastic than the parental SW480 cells however their adhesion to HA and fibronectin was the same as the SW480 cells. Although the MMP-7 transfected cell line SW480M7 was equally as adhesive to all matrices, overall it was more adhesive than the parental SW480 cells. The adhesion of the MMP-9 transfected cell line SW480M9 to plastic was similar to that of the parental SW480 cells however its adhesion to both HA and fibronectin was significantly increased.

Considering the individual roles of both CAMs and MMPs, the interactions between these groups of molecules were of extreme interest to this study. CAMs and MMPs are no longer thought of as acting in isolation in the metastatic process and more focus is being put on their collective roles. Furthermore, MMP-CAM interactions group the key metastatic steps of cell adhesion and proteolysis together. Both cell-cell and cell-matrix adhesion interactions have the ability to enhance MMP activity (Ellerbroek *et al.*, 1999 and Yakubenko *et al.*, 2000). CAMs have been shown to bind

MMPs on the cell surface and target their proteolytic activity (Brooks *et al.*, 1996 and Yu and Stamenkovic, 1999). As a result of MMP-mediated ECM turnover, tumour cells must adapt to the new ECM by varying CAM-ECM interactions which also effects CAM-mediated signal transduction (Cheresh, 1991).

For an insight into the influence of extracellular conditions on MMP activity cells were grown in the presence of various ECM components. SW480, SW620 and SW480M7 cells did not secrete MMP-9 and the presence of HA, fibronectin, type IV collagen or β-1 integrin-stimulating antibody did not stimulate MMP-9 expression in these cells. SW480M9 cells showed enhanced MMP-9 activity when grown in the presence of ECM components and the greatest effect was observed when cells were grown in the presence of HA. This effect was shown to occur at the post-transcriptional level.

CD44 is the major cell surface receptor for HA in the ECM. CD44 mediates both cell-cell and cell-matrix adhesion interactions. Increased expression of alternatively spliced CD44 isoforms is associated with an invasive phenotype and has been observed in metastatic cancers including colorectal cancer (Wielenga *et al.*, 1993, Jothy, 2003 and Gunthert *et al.*, 1991). Expression of CD44 isoforms containing the v6 alternatively spliced exon was increased during the progression of colorectal cancer (Herrlich *et al.*, 1993 and Wielenga *et al.*, 1993 and 1998).

CD44 and MMPs are both of importance in the metastatic process and it is thought that their roles are closely related. Stimulation of CD44 has been shown to increase MMP-2 expression (Takahasi *et al.*, 1999 and Zhang *et al.*, 2002). CD44 can also bind to MMP-2 and MMP-9 on the cell surface, linking cell adhesion and signalling with proteolytic activity (Yu and Stamenkovic, 1999 and Yu *et al.*, 2002).

Since MMP-9 activity in the SW480M9 cells was responsive to treatment with HA, the cells were treated with a monoclonal antibody designed to stimulate CD44. CD44 activation enhanced MMP-9 activity in a time dependent manner and also resulted in the activation of pro-MMP-9. This increase in MMP-9 was shown to be post-transcriptional.

The expression of CD44 and β1 integrin was investigated in all four cell lines. All cells expressed similar amounts of the standard 80 kDa CD44 except SW480M9 cells which showed increased expression. The SW480M9 and SW620 cells also expressed a 120 kDa CD44 isoform. All cell lines expressed the CD44 isoform containing the v6 through v8 sequence and all cell lines except the SW620 cells expressed isoforms containing the v6 through v10 sequence. The cell lines also expressed β1 integrin with SW620 cells expressing the highest amount. These results may explain the differences in cell adhesion to various ECM components and the differences in response to ECM components in terms of MMP activity.

To show a functional role for CD44 crosslinking and activation, *in vitro* invasion assays were performed. Treatment of cell lines with HA increased their *in vitro* invasion. The most dramatic increase was observed in the SW480M9 cells. The activation of CD44 with a monoclonal antibody resulted in increased invasion of the SW620 and SW480M9 cells. Blocking of CD44 function resulted in decreased invasion in all cell lines studied.

Taken together, these results highlight the importance of CAM-MMP interactions in tumour cell invasion. The CD44-MMP-9 interactions reported here are of significance as they demonstrate the interplay between the ECM, CAMs and MMP activity in tumour cell invasion. CD44 activation was shown to enhance MMP-9 activity which resulted in increased cell invasion that could be inhibited by blocking CD44 function. Similar increases in invasion were not observed in cell lines that did not express MMP-9 however blocking CD44 did reduce their invasion. This suggests that CD44 operates in cell invasion independently of MMP activity and that the combination of both molecules results in a cumulative effect.

Allowing one to study the expression levels of many genes in parallel, microarray analysis has become a powerful tool in the investigation of the molecular basis of metastasis. To identify metastasis related genes, cDNA microarray analysis was used to compare the gene expression profiles of the colorectal cancer cell lines. Using a 39,365 element microarray, genes with increased or decreased expression in the SW620, SW480M7 and SW480M9 cell lines were identified. In all experiments, the SW480 cell line was used as the control. 233 genes had increased expression of at least 2 fold in the SW620 cells and 208 genes had decreased expression of at least 2 fold in

the SW620 cells when compared with the SW480 cells. 147 genes had increased expression of at least 2 fold in the SW480M7 cells and 208 genes had decreased expression of at least 2 fold in comparison to the SW480 cells. 193 genes had increased expression of at least 2 fold in the SW480M9 cells and 20 genes had decreased expression of at least 2 fold in the SW480M9 cells when compared with the SW480 cells. The following genes were chosen for validation and further analysis: cathepsin C, tenascin C, PPARγ and caveolin 1, which were all up regulated in the SW620 cell line. Asparagine synthetase was chosen for further analysis as its expression was decreased in both the SW480M7 and SW480M9 cell lines.

Cathepsin C is a cysteine protease, and its roles in colorectal metastasis have not yet been established (Rao et al., 1997). Caveolin 1 is a component of the plasma membrane and is involved in signal transduction (Massimino et al., 2002). Tenascin C is an ECM component involved in proliferation, and high expression has been reported in malignant colon cancers (Kressner et al., 1997). PPARγ is a transcription factor involved in regulating the transcription of various genes involved in proliferation and differentiation (Rosen and Spiegelman, 2001). Asparagine synthetase is involved in the synthesis of asparagine and its levels have been shown to be decreased in leukaemic cells (Aslanian et al., 2001).

Sequencing confirmed the identity of the respective spotted cDNAs and real time PCR confirmed the changes in expression. A fluorescent substrate assay confirmed cathepsin C activity at the protein level. PCR was used to investigate the alternative splicing of tenascin C. SW620 cells expressed more tenascin C than the SW480 cells and it also expressed a tenascin C isoform containing exon 16 that the SW480 cells did not express. The SW620 cells also expressed tenascin C isoforms containing exons 14 and 16 which the SW480 cells did not express. These isoforms have previously been associated with malignancy in breast and colon (Wilson et al., 1996, Adams et al., 2002 and Dueck et al., 1999). Further follow up investigation of these genes could potentially validate their use as prognostic markers of colorectal cancer or as therapeutic targets to control metastasis.

A pilot study investigating the expression of the four selected genes in paired samples of normal and tumour colon tissue was conducted. All Dukes' classifications were represented in the nine paired specimens. Expression of all four genes was

detected in all tissue samples although no conclusive trends in expression regarding Dukes' classification or patient survival were observed. Tumour samples that had the same classification showed opposite trends in expression. The only notable trend was that Caveolin 1 expression was decreased in all tumour tissue in comparison with normal tissue, except tumour sample T10 where its expression was only slightly elevated. A reduction in caveolin 1 levels has been associated with increased invasion in various cancers including ovarian (Bagnoli *et al.*, 2000), colon (Bender *et al.*, 2000), breast (Lu *et al.*, 2004) and lung (Racine *et al.*, 1999). This finding contradicts the previous observation that caveolin 1 expression was upregulated in the SW620 cells in comparison with the SW480 cells. It is possible that the increase in caveolin 1 expression in the SW620 cells is patient specific. The SW620 and SW480 cell lines are both of tumour origin and this study compared the gene expression of normal and tumour tissue. Therefore, the increase in caveolin 1 expression could be a key event in the development of metastasis whereas a decrease in caveolin 1 expression may represent an earlier event in tumour development. A larger sample set would be required to draw any conclusions regarding the expression of these genes in colorectal tumours of varying Dukes' classification.

In conclusion, this data highlighted the role of many of the participants involved in tumour invasion and the need to consider the interplay and interactions between them. The first part of this thesis described a mechanism by which colorectal tumour cells interacted with the ECM resulting in an increase in MMP-9 activity. Crosslinking of CD44 resulted in a post-transcriptional enhancement of MMP-9 activity together with increased cell invasion. This mechanism could potentially be exploited by invading tumour cells, by either expressing MMP-9 themselves or by triggering MMP-9 secretion in neighbouring stromal cells. Subsequently, invading tumour cells would benefit from the increased ECM turnover. Further investigation is required to elucidate the exact mechanism of activation. Nevertheless, this work described a mechanism, the intervention of which could lead to the development of new therapies.

Further insights were gained into the transcriptional events regulating colorectal metastasis using cDNA microarray analysis. Differentially regulated genes were chosen and their expression was validated and detected in matched colorectal tissue specimens. Further investigation of these genes will elucidate their underlying roles in colorectal metastasis and validate their use as prognostic markers of the disease. These genes

were chosen from a large list of genes that were differentially regulated in a metastatic cell line. The other genes in this list are included as an appendix to this thesis and their investigation in the process is equally valid.

Analysis of the expression of these genes in a greater number of tissue samples would help to elucidate their roles in colorectal cancer. A larger sample set would also help correlate expression patterns with tumour stage. Analysis of tenascin C isoform expression in these tissues would define which tenascin C isoform(s) are expressed in colorectal tissues of varying Dukes' classifications. Because the roles of these genes in colorectal cancer are undefined, loss of function and gain of function experiments would be required to elucidate them. The roles of these genes in *in vitro* invasion and migration could be monitored by overexpression in non-expressing cell lines and by diminishing their expression using small interfering RNA (siRNA) technology in expressing cell lines.

Chapter 7

Bibliography

Adachi, Y., Yamamoto, H., Itoh, F., Hinoda, Y., Okada, Y. and Imai, K (1999) Contribution of matrilysin (MMP-7) to the metastatic pathway of human colorectal cancers. *Gut* **45**:252 –258.

Adachi, Y., Yamamoto, H., Itoh, F., Arimura, Y., Nishi, M., Endo, T. and Imai, K. (2001) Clinicopathologic and prognostic significance of matrilysin expression at the invasive front in human colorectal cancers. *Int J Cancer.* **95**: 290-294.

Adams, M., Jones, J. L., Walker, R. A., Pringle, J. H. and Bell, S. C. (2002) Changes in tenascin-C isoform expression in invasive and preinvasive breast disease. *Cancer Res.* **62**: 3289-3297.

Agrez, M., Chen, A., Cone, R. I., Pytela, R. and Sheppard, D. (1994) The alpha v beta 6 integrin promotes proliferation of colon carcinoma cells through a unique region of the beta 6 cytoplasmic domain. *J. Cell. Biol* **127**: 547-556.

Agrez, M., Gu, X., Turton, J., Meldrum, C., Niu, J., Antalis, T. and Howard, E. W. (1999) The alpha v beta 6 integrin induces gelatinase B secretion in colon cancer cells. *Int J Cancer* **81**:90-97.

Al Moustafa, A. E., Alaoui-Jamali, M. A., Batist, G., Hernandez-Perez, M., Serruya, C., Alpert, L., Black, M. J., Sladek, R. and Foulkes, W. D. (2002) Identification of genes associated with head and neck carcinogenesis by cDNA microarray comparison between matched primary normal epithelial and squamous carcinoma cells. *Oncogene.* **21**: 2634-2640.

Allan, J. A., Docherty, A. J. P., Barker, P. J., Huckinsson, N. S., Reynolds, J. J. and Murphy, G. (1995) Binding of gelatinases A and B to type-I collagen and other matrix components. *Biochem. J.* **309**: 299-306.

Alon, U., Barkai, N., Notterman, D. A., Gish, K., Ybarra, S., Mack, D. and Levine, A. J. (1999) Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. *Proc Natl Acad Sci U S A.* **96**: 6745-6750.

Al-Sheneber, I. F., Shibata, H. R., Sampalis, J. and Jothy, S. (1993) Prognostic significance of proliferating cell nuclear antigen expression in colorectal cancer. *Cancer.* **71**: 1954-1959.

Anbazhagan, R., Tihan, T., Bornman, D. M., Johnston, J. C., Saltz, J. H., Weigering, A., Piantadosi, S. and Gabrielson, E. (1999) Classification of small cell lung cancer and pulmonary carcinoid by gene expression profiles. *Cancer Res.* **59**: 5119-5122.

Aslanian, A. M., Fletcher, B. S. and Kilberg, M. S. (2001) Asparagine synthetase expression alone is sufficient to induce l-asparaginase resistance in MOLT-4 human leukaemia cells. *Biochem J.* **357**: 321-328.

Austin, C. P. (2004) The impact of the completed human genome sequence on the development of novel therapeutics for human disease. *Annu Rev Med.* **55**: 1-13.

Bagnoli, M., Tomassetti, A., Figini, M., Flati, S., Dolo, V., Canevari, S. and Miotti, S. (2000) Downmodulation of caveolin-1 expression in human ovarian carcinoma is directly related to alpha-folate receptor overexpression. *Oncogene.* **19**: 4754-4763.

Bajorath, J., Greenfield, B., Munro, S. B., Day, A. J. and Aruffo, A. (1998) Identification of CD44 residues important for hyaluronan binding and delineation of the binding site. *J Biol Chem.* **273**: 338-343.

Barbosa-Tessmann, I. P., Chen, C., Zhong, C., Siu, F., Schuster, S. M., Nick, H. S. and Kilberg, M. S. (2000) Activation of the human asparagine synthetase gene by the amino acid response and the endoplasmic reticulum stress response pathways occurs by common genomic elements. *J Biol Chem.* **275**: 26976-26985.

Barbour, A. P., Reeder, J. A., Walsh, M. D., Fawcett, J. Antalis, T. M. and Gotley, D. C. (2003) Expression of the CD44v2-10 isoform confers a metastatic phenotype: Importance of the heparin sulfate attachment site CD44v4. *Cancer Res.* **63**: 887-892.

Becker, J. W., Marcy, A. I., Rokosz, L. L., Axel, M. G., Burbaum, J. J., Fitzgerald , P. M., Cameron, P. M., Esser, C. K., Hagmann, W. K., Hermes, J. D. and Springer, J. P. (1995) Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme. *Protein Sci.* **4**: 1966-1976.

Bell, S. M., Scott, N., Cross, D., Sagar, P., Lewis, F. A., Blair, G. E., Taylor, G. R., Dixon, M. F. and Quirke, P. (1994) Prognostic value of p53 overexpression and c-Ki-ras gene mutations in colorectal cancer. *Gastroenterology.* **104**: 57-64.

Bender, F. C., Reymond, M. A., Bron, C. and Quest, A. F. (2000) Caveolin-1 levels are down-regulated in human colon tumors, and ectopic expression of caveolin-1 in colon carcinoma cell lines reduces cell tumorigenicity. *Cancer Res.* **60**: 5870-5878.

Bernhard, E. J., Gruber, S. B. and Muschel, R. J. (1994) Direct evidence linking expression of matrix metalloproteinase 9 (92-kDa gelatinase/collagenase) to the metastatic phenotype in transformed rat embryo cells. *Proc Natl Acad Sci USA* **91**: 4293-4297.

Bertucci, F., Houlgatte, R., Benziane, A., Granjeaud, S., Adelaide, J., Tagett, R., Loriod, B., Jacquemier, J., Viens, P., Jordan, B., Birnbaum, D. and Nguyen, C. (2000) Gene expression profiling of primary breast carcinomas using arrays of candidate genes. *Hum Mol Genet.* **9**: 2981-2991.

Bertucci, F., Houlgatte, R., Nguyen, C., Viens, P., Jordan, B. R. and Birnbaum, D. (2001) Gene expression profiling of cancer by use of DNA arrays: how far from the clinic? *Lancet Oncol.* **2**: 674-682.

Birkedal-Hansen, H., Moore, W. G. I., Bodden, M. K., Windsor, L. J., Birkedal-Hansen, B., DeCarlo, A. and Engler, J. A. (1993) Matrix Metalloproteinases – A Review. *Crit. Rev. Oral Biol. Med.* **4**:197-250.

Bittner, M., Meltzer, P., Chen, Y., Jiang, Y., Seftor, E., Hendrix, M., Radmacher, M., Simon, R., Yakhini, Z., Ben-Dor, A., Sampas, N., Dougherty, E., Wang, E., Marincola, F., Gooden, C., Lueders, J., Glatfelter, A., Pollock, P., Carpten, J., Gillanders, E., Leja, D., Dietrich, K., Beaudry, C., Berens, M., Alberts, D. and Sondak, V. (2000) Molecular classification of cutaneous malignant melanoma by gene expression profiling. *Nature.* **406**: 536-540.

Blower, P. E., Yang, C., Fligner, M. A., Verducci, J. S., Yu, L., Richman, S. and Weinstein, J. N. (2002) Pharmacogenomic analysis: correlating molecular substructure classes with microarray gene expression data. *Pharmacogenomics J.* **2**: 259-271.

Bode, W., Gomis-Ruth, F. X. and Stocker, W. (1993) Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the 'metzincins'. *FEBS Lett.* **331**: 134-140.

Bogenrieder, T. and Herlyn, M. (2003) Axis of evil: molecular mechanisms of cancer metastasis. *Oncogene* **22**: 6524-6536.

Borsi, L., Carnemolla, B., Nicolo, G., Spina, B., Tanara, G. and Zardi, L. (1992) Expression of different tenascin isoforms in normal, hyperplastic and neoplastic human breast tissues. *Int J Cancer.* **52**: 688-692.

Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan-de Vries, M., van Boom, J. H., van der Eb, A. J. and Vogelstein, B. (1987) Prevalence of ras gene mutations in human colorectal cancers. *Nature.* **327**: 293-297.

Bourguignon, L. Y. W., Gunja-Smith, Z., Iida, N., Zhu, H. B., Young, L. J. T., Muller, W, J. and Cardiff, R. D. (1998) CD44v3,8-10 is involved in cytoskeletal-mediated tumour cell migration and matrix metalloproteinase (MMP-9) association in metastatic breast cancer cells. *J Cell Physiol* **176**: 206-215.

Bourguignon, L. Y., Singleton, P. A., Zhu, H. and Zhou, B. (2002) Hyaluronan promotes signaling interaction between CD44 and the transforming growth factor beta receptor I in metastatic breast tumor cells. *J Biol Chem.* **277**:39703-39712.

Bowtell, D. D. L. (1999) Options available – from start to finish- for obtaining expression data by microarray. *Nature Genet.* **21**: 25-32.

Bretscher, A., Edwards, K. and Fehon, R. G. (2002) ERM proteins and merlin: integrators at the cell cortex. *Nat Rev Mol Cell Biol.* **3**: 586-599.

Brew, K., Dinakarpandian, D. and Nagase, H. (2000) Tissue inhibitors of metalloproteinases: evolution, structure and function. *Biochim Biophys Acta.* **1477**: 267-283.

Brinkerhoff, C. E. and Matrisian, L. M. (2002) Matrix metalloproteinases: a tail of a frog that became a prince. *Nat Rev Mol Cell Biol.* **3**: 207-213 .

Brooks, P. C., Silletti, S., von Schalscha, T. L., Friedlander, M., and Cheresh, D. A. (1998) Disruption of angiogenesis by PEX, a noncatalytic metalloproteinase fragment with integrin binding activity. *Cell* **92**: 391-400.

Brooks, P. C., Stromblad, S., Sanders, L. C., von Schalscha, T. L., Aimes, R. T., Stetler-Stevenson, W. G., Quigley, J. P. and Cheresh, D. A. (1996) Localization of matrix metalloproteinase MMP-2 to the surface of invasive cells by interaction with integrin alpha v beta 3. *Cell* **85**: 683-693.

Burgermeister, E., Tencer, L. and Liscovitch, M. (2003) Peroxisome proliferator-activated receptor-gamma upregulates caveolin-1 and caveolin-2 expression in human carcinoma cells. *Oncogene.* **22**: 3888-3900.

Butte, A. (2002) The use and analysis of microarray data. *Nat Rev Drug Discov.* **1**: 951-960.

Callow, M. J., Dudoit, S., Gong, E. L., Speed, T. P. and Rubin, E. M. (2000) Microarray expression profiling identifies genes with altered expression in HDL-deficient mice. *Genome Res.* **10**:2022-2029.

Chaib, H., Cockrell, E. K., Rubin, M. A., Macoska, J. A. (2001) Profiling and verification of gene expression patterns in normal and malignant human prostate tissues by cDNA microarray analysis. *Neoplasia* **3**: 43-52.

Chambers, AF. and Matrisian, LM. (1997) Changing views of the role of matrix metalloproteinases in metastasis. *J. Nat. Cancer. Inst.* **89**: 1260-1270.

Chang, C. and Werb, Z. (2001). The many faces of metalloproteases: cell growth, invasion, angiogenesis and metastasis. *Trends Cell. Biol.* **11**:S37–S43.

Chen, Y., Dougherty, E. R. and Bittner, M. L. (1997) Ratio-Based Decisions and the Quantitative Analysis of cDNA Microarray Images. *J Biomed. Optics.* **2**: 364-374.

Cheresh, D. A. (1991) Structure, function and biological properties of integrin alpha v beta 3 on human melanoma cells. *Cancer Metast. Rev.* **10**: 3-10.

Cheung, V. G., Morley, M., Aguilar, F., Massimi, A., Kucherlapati, R. and Childs, G. (1999) Making and reading microarrays. *Nat Genet.* **21**: 15-19.

Chiquet-Ehrismann, R., Matsuoka, Y., Hofer, U., Spring, J., Bernasconi, C. and Chiquet, M. (1991) Tenascin variants: differential binding to fibronectin and distinct distribution in cell cultures and tissues. *Cell Regul.* **2**: 927-938.

Choi, S. H., Takahashi, K., Eto, H., Yoon, S. S. and Tanabe, K. K. (2000) CD44s expression in human colon carcinomas influences growth of liver metastases. *Int J Cancer.* **85**: 523-526.

Chuaqui, R. F., Bonner, R. F., Best, C. J., Gillespie, J. W., Flaig, M. J., Hewitt, S. M., Phillips, J. L., Krizman, D. B., Tangrea, M. A., Ahram, M., Linehan, W. M., Knezevic, V. and Emmert-Buck, M. R. (2002) Post-analysis follow-up and validation of microarray experiments. *Nat Genet.* **32**: 509-514.

Churchill, G. A. (2002) Fundamentals of experimental design for cDNA microarrays. *Nat. Genet.* **32**: 490-495.

Cichy, J. and Pure, E. (2003) The liberation of CD44. *J Cell Biol.* **161**: 839-843.

Clark, E. A., Golub, T. R., Lander, E. S. and Hynes, R. O. (2000) Genomic analysis of metastasis reveals an essential role for RhoC. *Nature.* **406**: 532-535.

Clark, I. M. and Cawston, T. E. (1989) Fragments of human fibroblast collagenase. Purification and charcterization. *Biochem. J.* **263**: 201-206.

Cleveland, W. S. (1979) Robust locally weighted regression and smoothing scatterplots. J Amer. Stat. Assoc. **74**: 829-835.

Collier, I. E., Wilhelm, S. M., Eisen, A. Z., Marmer, B. L., Grant, G. A., Seltzer, J. L., Kronberger, A., He, C. S., Bauer, E. A. and Goldberg, G. I. (1988) H-ras oncogene-transformed human bronchial epithelial cells (TBE-1) secrete a single metalloprotease capable of degrading basement membrane collagen. *J. Biol. Chem.* **263**: 6579-6587.

Coussens, L. M., Fingleton, B. and Matrisian, L. M. (2002) Matrix metalloproteinase inhibitors and cancer: trials and tribulations. *Science* **295**: 2387-2392.

Cowan, K. N., Jones, P. L. and Rabinovitch, M. (2000) Elastase and matrix metalloproteinase inhibitors induce regression, and tenascin-C antisense prevents progression of vascular disease. *J Clin Invest.* **105**: 21-34.

Crawford, H. C. and Matrisian, L. M. (1996) Mechanisms controlling the transcription of matrix metalloproteinase genes in normal and neoplastic cells. *Enzyme Prot.* **49**:20-37.

Crawford, H. C., Fingleton, B. M., Rudolph-Owen, L. A., Goss, K. J., Rubinfeld, B., Polakis, P. and Matrisian, L. M. (1999) The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. *Oncogene.* **18**: 2883-2891.

Cui, X. and Churchill, G. A. (2003) Statistical tests for differential expression in cDNA microarray experiments. *Genome Biol.* **4**: 210.

Curran, S. and Murray, G. I. (1999) Matrix metalloproteinases in tumour invasion and metastasis. *J Pathol* **189**: 300–308.

Curran, S. and Murray, G. I. (2000) Matrix metalloproteinases: molecular aspects of their roles in tumour invasion and metastasis. *Eur J Cancer.* **36**: 1621-1630.

Daemi, N., Thomasset, N., Lissitzky, J. C., Dumortier, J., Jacquier, M. F., Pourreyron, C., Rousselle, P., Chayvialle, J. A. and Remy, L. (2000) Anti-beta4 integrin antibodies enhance migratory and invasive abilities of human colon adenocarcinoma cells and their MMP-2 expression. *Int J Cancer.* **85**:850-856.

De Bruine, A. P., Wiggers, T., Beek, C., Volovics, A., von Meyenfeldt, M., Arends, J. W. and Bosman, F. T. (1993) Endocrine cells in colorectal adenocarcinomas: incidence, hormone profile and prognostic relevance. *Int J Cancer.* **54**: 765-771.

De Souza, S. J., Pereira, H. M., Jacchieri, S. and Brentani, R. R. (1996) Collagen/collagenase interaction: does the enzyme mimic the conformation of its own substrate? *FASEB J.* **10**: 927-930.

Del Re, E. C., Shuja, S., Cai, J. and Murnane, M. J. (2000) Alterations in cathepsin H activity and protein patterns in human colorectal carcinomas. *Br J Cancer.* **82**: 1317-1326 .

Delattre, O., Olschwang, S., Law, D. J., Melot, T., Remvikos, Y., Salmon, R. J., Sastre, X., Validire, P., Feinberg, A. P. and Thomas, G. (1989) Multiple genetic alterations in distal and proximal colorectal cancer. *Lancet.* **2**: 353-356.

Deryugina, E. I., Ratnikov, B. I., Postnova, T. I., Rozanov, D. V. and Strongin, A. Y. (2002) Processing of integrin alpha(v) subunit by membrane type 1 matrix metalloproteinase stimulates migration of breast carcinoma cells on vitronectin and enhances tyrosine phosphorylation of focal adhesion kinase. *J Biol Chem.* **277**: 9749-9756

Diep, C. B., Thorstensen, L., Meling, G. I., Skovlund, E., Rognum, T. O. and Lothe, R. A. (2003) Genetic tumor markers with prognostic impact in Dukes' stages B and C colorectal cancer patients. *J Clin Oncol.* **21**: 820-829.

Dolenc, I., Turk, B., Pungercic, G., Ritonja, A. and Turk, V. (1995) Oligomeric structure and substrate induced inhibition of human cathepsin C. *J. Biol. Chem.* **270**: 21626-21631.

Dueck, M., Riedl, S., Hinz, U., Tandara, A., Moller, P., Herfarth, C. and Faissner, A. (1999) Detection of tenascin-C isoforms in colorectal mucosa, ulcerative colitis, carcinomas and liver metastases. *Int J Cancer.* **82**: 477-483.

Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J. M. (1999) Expression profiling using cDNA microarrays. *Nat Genet.* **21**: 10-4.

Dukes, C. E. (1932) The classification of cancer of the rectum. *J Pathol Bacteriol.* **35**: 323-332.

Dukes, C. E. (1949) The surgical pathology of rectal cancer. *J Clin Pathol* **2**: 95-99.

Ellerbroek, S. M., Fishman, D. A., Kearns, A. S., Bafetti, L. M. and Stack, M. S. (1999) Ovarian carcinoma regulation of matrix metalloproteinase-2 and membrane type 1 matrix metalloproteinase through beta1 integrin. *Cancer Res.* **59**: 1635-1641.

Emmert-Buck, M. R., Roth, M. J., Zhuang, Z., Campo, E., Rozhin, J., Sloane, B. F., Liotta, L. A. and Stetler-Stevenson, W. G. (1994) Increased gelatinase A (MMP-2) and cathepsin B activity in invasive tumor regions of human colon cancer samples. *Am J Pathol.* **145**: 1285-1290.

Emmert-Buck, M. R., Bonner, R. F., Smith, P. D., Chuaqui, R. F., Zhuang, Z., Goldstein, S. R., Weiss, R. A. and Liotta, L. A. (1996) Laser capture microdissection. *Science.* **274**: 998-1001.

Fajas, L., Debril, M. B. and Auwerx, J. (2001) Peroxisome proliferator-activated receptor-gamma: from adipogenesis to carcinogenesis. *J Mol Endocrinol.* **27**: 1-9.

Fearon, E. R., Hamilton, S. R. and Vogelstein, B. (1987) Clonal analysis of human colorectal tumors. *Science.* **238**: 193-197.

Fearon, E. R. and Vogelstein, B. (1990) A genetic model for colorectal tumorigenesis. *Cell.* **61**: 759-767.

Fidler, I. J. and Kripke, M. L. (1977) Metastasis results from preexisting variant cells within a malignant tumor. *Science.* **197**: 893-895.

Fidler, I. J. and Hart, I. R. (1982) Biological diversity in metastatic neoplasms: origins and implications. *Science.* **217**: 998-1003.

Fine, S. W., Lisanti, M. P., Galbiati, F. and Li, M. (2001) Elevated expression of caveolin-1 in adenocarcinoma of the colon. *Am J Clin Pathol.* **115**: 719-724.

Fingleton, B. and McDonnell S. (1997) Cytokine regulation of matrilysin gene expression. *Biochem Soc Trans.* 25(2): 155-155.

Fingleton, BM., Goss, KJH., Crawford, HC. and Matrisian, LM. (1999) Matrilysin in early stage intestinal tumourigenesis. *APMIS.* **107**: 102-110 .

Fini, M. E., Cook, J. R., Mohan, R. and Brickerhoff, C. E. (1998) Regulation of matrix metalloproteinase gene expression. In *Matrix Metalloproteinases*, ed. W. C. Parks and R. P. Mecham. pp. 299-356. New York Academic.

Fisher, C., Gilbertson-Beadling, S., Powers, E. A., Petzold, G., Poorman, R. and Mitchell, M. A. (1994) Interstitial collagenase is required for angiogenesis *in vitro. Dev. Biol.* **162**: 499-510.

Foekens, J. A., Look, M. P., Bolt-de Vries, J., Meijer-van Gelder, M. E., van Putten, W. L. and Klijn, J. G. (1999) Cathepsin-D in primary breast cancer: prognostic evaluation involving 2810 patients. *Br J Cancer.* **79**: 300-307.

Fu, Y. H., Nishinaka, T., Yokoyama, K. and Chiu, R. (1998) A retinoblastoma susceptibility gene product, RB, targeting protease is regulated through the cell cycle. *FEBS Lett.* **421**: 89-93.

Fuchs, E. and Weber, K. (1994) Intermediate filaments: structure, dynamics, function, and disease. *Annu Rev Biochem.* **63**: 345-382

Gabriel, W. B., Dukes, C. and Bussey, H. J. R. (1935) Lymphatic spread in cancer of the rectum. Br J Surg **23**: 395-413.

Giancotti, F. G. and Ruoslahti, E. (1999) Integrin signaling.*Science* **285**: 1028-1032.

Giannelli, G., Falk-Marzillier, J., Schiraldi, O., Stetler-Stevenson, WG. and Quaranta, V. (1997) Induction of cell migration by matrix metalloprotease-2 cleavage of laminin-5. *Science* **277**: 225-228.

Giannelli, G., Fransvea, E., Marinosci, F., Bergamini, C., Colucci, S., Schiraldi, O. and Antonaci, S. (2002) Transforming growth factor-beta1 triggers hepatocellular carcinoma invasiveness via alpha3beta1 integrin. *Am J Pathol.* **161**: 183-193.

Giulietti, A., Overbergh, L., Valckx, D., Decallonne, B., Bouillon, R. and Mathieu, C. (2001) An overview of real-time quantitative PCR: applications to quantify cytokine gene expression. *Methods* **25**: 386-401.

Glondu, M., Liaudet-Coopman, E., Derocq, D., Platet, N., Rochefort, H. and Garcia, M. (2002) Down-regulation of cathepsin-D expression by antisense gene transfer inhibits tumor growth and experimental lung metastasis of human breast cancer cells. *Oncogene.* **21**: 5127-5134.

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., Bloomfield, C. D. and Lander, E. S. (1999) Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science.* **286**: 531-537.

Greco, A., Gong, S. S., Ittmann, M. and Basilico, C. (1989) Organization and expression of the cell cycle gene, ts11, that encodes asparagine synthetase. *Mol Cell Biol.* **9**: 2350-2359.

Greene, F. L. and Sobin, L. H. (2002) The TNM system: our language for cancer care. *J Surg Oncol.* **80**: 119-120.

Greenson, J. K., Isenhart, C. E., Rice, R., Mojzisik, C., Houchens, D. and Martin, E. W. Jr. (1994) Identification of occult micrometastases in pericolic lymph nodes of Duke's B colorectal cancer patients using monoclonal antibodies against cytokeratin and CC49. Correlation with long-term survival. *Cancer.* **73**: 563-569.

Grubben, M. J., Nagengast, F. M., Katan, M. B. and Peters, W. H. (2001) The glutathione biotransformation system and colorectal cancer risk in humans. *Scand J Gastroenterol Suppl.* 234: 68-76.

Gunthert, U., Hofmann, M., Rudy, W., Reber, S., Zoller, M., Haussmann, I., Matzku, S., Wenzel, A., Ponta, H. and Herrlich, P. (1991) A new variant of glycoprotein CD44 confers metastatic potential to rat carcinoma cells. *Cell* **65**: 13-24.

Hanahan, D and Weinberg, R. A. (2000) The hallmarks of cancer. *Cell* **100**: 57-70.

Hasegawa, S., Koshikawa, N., Momiyama, N., Moriyama, K., Ichikawa, Y., Ishikawa, T., Mitsuhashi, M., Shimada, H. and Miyazaki, K. (1998) Matrilysin-specific antisense oligonucleotide inhibits liver metastasis of human colon cancer cells in a nude mouse model. *Int J Cancer* **76**:812 –816.

Hauptmann, S., Zardi, L., Siri, A., Carnemolla, B., Borsi, L., Castellucci, M., Klosterhalfen, B., Hartung, P., Weis, J., Stocker, G., *et al.* (1995) Extracellular matrix proteins in colorectal carcinomas. Expression of tenascin and fibronectin isoforms. *Lab Invest.* **73**: 172-182.

Havenith, M. G., Arends, J. W., Simon, R., Volovics, A., Wiggers, T. and Bosman, F. T. (1988) Type IV collagen immunoreactivity in colorectal cancer. Prognostic value of basement membrane deposition. *Cancer.* **62**: 2207-2211.

He, C. S., Wilhelm, S. M., Pentland, A. P., Marmer, B. L., Grant, G. A. Eisen, A. Z. and Goldberg, G. I. (1989) Tissue cooperation in a proteolytic cascade involving human interstitial collagenase. *Proc. Natl. Acad. Sci. USA.* **86**:2632-2636.

He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., da Costa, L. T., Morin, P. J., Vogelstein, B. and Kinzler, K. W. (1998) Identification of c-MYC as a target of the APC pathway. *Science.* **281**: 1509-1512.

Hegde, P., Qi, R., Abernathy, K., Gay, C., Dharap, S., Gaspard, R., Earle-Hughes, J., Snesrud, E., Lee, N. and Quackenbush, J. (2000) A Concise Guide to cDNA Microarray Analysis. *Biotechniques*. **20**: 548-562.

Hegde, P., Qi, R., Gaspard, R., Abernathy, K., Dharap, S., Earle-Hughes, J., Gay, C., Nwokekeh, N. U., Chen, T., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and Quackenbush, J. (2001) Identification of tumor markers in models of human colorectal cancer using a 19,200-element complementary DNA microarray. Cancer Res. **61**: 7792-7797.

Herrlich, P., Rudy, W., Hofman, M. Arch, R., Zöller, M., Zawadzki, V., Tölg, C., Hekele, A., Koopman, G., Pals, S., Heider, K. H., Sleeman, J. and Ponta, H. (1993) CD44 and Splice Variants of CD44 in Normal Differentiation and Tumour Progression. In *Cell Adhesion Molecules,* ed. M. E. Hemler and E. Mihich, pp 265-278. New York, USA. Plenum Press.

Herrlich, P., Pals, S. and Ponta, H. (1995) CD44 in colon cancer. *Eur J Cancer*. **31**: 1110-1112.

Hewitt, R. E., McMarlin, A., Kleiner, D., Wersto, R., Martin, P., Tsokos, M., Stamp, G. W., Stetler-Stevenson, W. G. and Tsoskos, M. (2000) Validation of a model of colon cancer progression. *J Pathol.* **192**: 446-454.

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.

Huerta, S., Harris, D. M., Jazirehi, A., Bonavida, B., Elashoff, D., Livingston, E. H. and Heber, D. (2003) Gene expression profile of metastatic colon cancer cells resistant to cisplatin-induced apoptosis. *Int J Oncol.* **22**: 663-670.

Humphries, M. J. (2000) Integrin cell adhesion receptors and the concept of agonism. *Trends Pharmacol. Sci.* **21**: 29-32.

Hynes, R. O. (1987) Integrins: a family of cell surface receptors. *Cell* **48**: 549-554.

Hynes, R. O. (2002) Integrins: bidirectional, allosteric signaling machines. *Cell* **110**: 673-687.

Iacobuzio-Donahue, C. A., Shuja, S., Cai, J., Peng, P. and Murnane, M. J. (1997) Elevations in cathepsin B protein content and enzyme activity occur independently of glycosylation during colorectal tumor progression. *J Biol Chem.* **272**: 29190-29199.

Isacke, C. M. and Yarwood, H. (2002) The hyaluronan receptor, CD44. *Int J Biochem Cell Biol.* **34**: 718-721.

Ishihara, A., Yoshida, T., Tamaki, H. and Sakakura, T. (1995) Tenascin expression in cancer cells and stroma of human breast cancer and its prognostic significance. *Clin Cancer Res.* **1**: 1035-1041.

Itoh, Y., Takamura, A., Ito, N., Maru, Y., Sato, H., Suenaga, N., Aoki, T. and Seiki, M. (2001) Homophilic complex formation of MT1-MMP facilitates proMMP-2 activation on the cell surface and promotes tumour cell invasion. *EMBO J.* **20**: 4782-4793.

Jahkola, T., Toivonen, T., Virtanen, I., von Smitten, K., Nordling, S., von Boguslawski, K., Haglund, C., Nevanlinna, H. and Blomqvist, C. (1997) Tenascin-C expression in invasion border of early breast cancer: a predictor of local and distant recurrence. *Br J Cancer.* **78**: 1507-1513.

Jen, J., Kim, H., Piantadosi, S., Liu, Z. F., Levitt, R. C., Sistonen, P., Kinzler, K. W., Vogelstein, B. and Hamilton, S. R. (1994) Allelic loss of chromosome 18q and prognosis in colorectal cancer. *N Engl J Med.* **331**: 213-221.

Joester, A. and Faissner, A. (1999) Evidence for combinatorial variability of tenascin-C isoforms and developmental regulation in the mouse central nervous system. *J Biol Chem.* **274**: 17144-17151.

Jones, J. L. and Walker, R. A. (1997) Control of matrix metalloproteinase activity in cancer. *J. Pathol.* **183**: 377-379.

Jones, P. L. and Jones, F. S. (2000) Tenascin-C in development and disease: gene regulation and cell function. *Matrix Biol.* **19**: 581-596.

Jothy, S. (2003) CD44 and its partners in metastasis. *Clin Exp Metastasis.* **20**: 195-201.

Joyce, J. A., Baruch, A., Chehade, K., Meyer-Morse, N., Giraudo, E., Tsai, F. Y., Greenbaum, D. C., Hager, J. H., Bogyo, M. and Hanahan, D. (2004) Cathepsin cysteine proteases are effectors of invasive growth and angiogenesis during multistage tumorigenesis. *Cancer Cell.* **5**: 443-453.

Juge-Aubry, C., Pernin, A., Favez, T., Burger, A. G., Wahli, W., Meier, C. A. and Desvergne, B. (1997) DNA binding properties of peroxisome proliferator-activated receptor subtypes on various natural peroxisome proliferator response elements. Importance of the 5'-flanking region. *J Biol Chem.* **272**: 25252-25259.

Juliano, R. L. and Varner, J. A. (1993) Adhesion molecules in cancer: the role of integrins. *Curr Opin Cell Biol.* **5**: 812-818.

Kajita, M., Itoh, Y., Chiba, T., Mori, H., Okada, A., Kinoh, H. and Seiki, M. (2001) Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. *J Cell Biol.* **153**: 893-904.

Kalembeyi, I., Inada, H., Nishiura, R., Imanaka-Yoshida, K., Sakakura, T. and Yoshida, T. (2003) Tenascin-C upregulates matrix metalloproteinase-9 in breast cancer cells: direct and synergistic effects with transforming growth factor beta1. *Int J Cancer.* **105**: 53-60.

Kato, K., Hida, Y., Miyamoto, M., Hashida, H., Shinohara, T., Itoh, T., Okushiba, S., Kondo, S. and Katoh, H. (2002) Overexpression of caveolin-1 in esophageal squamous cell carcinoma correlates with lymph node metastasis and pathologic stage. *Cancer.* **94**: 929-933.

Katunuma, N., Tsuge, H., Nukatsuka, M. and Fukushima, M. (2002) Structure-based development of cathepsin L inhibitors and therapeutic applications for prevention of cancer metastasis and cancer-induced osteoporosis. *Adv Enzyme Regul.* **42**: 159-172.

Kerr, M. K. and Churchill, G. A. (2001a) Experimental design for gene expression microarrays. *Biostatistics.* **2**: 183-201.

Kerr, M. K. and Churchill, G. A. (2001b) Statistical design and the analysis of gene expression microarray data. *Genet Res.* **77**: 123-128.

Khan, J., Simon, R., Bittner, M., Chen, Y., Leighton, S. B., Pohida, T., Smith, P. D., Jiang, Y., Gooden, G. C., Trent, J. M. and Meltzer, P. S. (1998a) Gene expression profiling of alveolar rhabdomyosarcoma with cDNA microarrays. *Cancer Res.* **58**: 5009-5013.

Khan A, Krishna M, Baker SP, Malhothra R, Banner BF. (1998b) Cathepsin B expression and its correlation with tumor-associated laminin and tumor progression in gastric cancer. *Arch Pathol Lab Med.* **122**: 172-177.

Kheradmand F., Werner, E., Tremble, P., Symons, M. and Werb, Z. (1998) Role of Rac 1 and oxygen radicals in collagenase-1 expression induced by cell shape change. *Science.* **280**: 898-902.

Kim, H., Yang, X. L., Rosada, C., Hamilton, S. R. and August, J. T. (1994) CD44 expression in colorectal adenomas is an early event occurring prior to K-ras and p53 gene mutation. *Arch Biochem Biophys.* **310**: 504-507

Klein, R. D., Borchers, A. H., Sundareshan, P., Bougelet, C., Berkman, M. R., Nagle, R. B. and Bowden, G. T. (1997) Interleukin-1beta secreted from monocytic cells induces the expression of matrilysin in the prostatic cell line LNCaP. *J Biol Chem.* 272: **14**188-14192.

Kleiner, D. E. and Stetler-Stevenson, W. G. (1994) Quantitative zymography: detection of picogram quantities of gelatinases. *Anal Biochem.* **218**: 325-329.

Krepela, E., Prochazka, J., Mynarikova, H., Karova, B., Cermak, J. and Roubkova, H. (1996) Lysosomal dipeptidyl-peptidases I and II in human squamous cell lung carcinoma and lung parenchyma. *Neoplasma.* **43**: 171-178.

Kressner, U., Lindmark, G., Tomasini-Johansson, B., Bergstrom, R., Gerdin, B., Pahlman, L. and Glimelius, B. (1997) Stromal tenascin distribution as a prognostic marker in colorectal cancer. *Br J Cancer.* **76**: 526-530.

Kubens, B. S. and Zanker, K. S. (1998) Differences in the migration capacity of primary human colon carcinoma cells (SW480) and their lymph node metastatic derivatives (SW620). *Cancer Lett* **131**: 55-64.

Kuribayashi, M., Yamada, H., Ohmori, T., Yanai, M. and Imoto, T. (1993) Endopeptidase activity of cathepsin C, dipeptidyl aminopeptidase I, from bovine spleen. *J Biochem (Tokyo)*. 113: 441-449.

Lackner, C., Moser, R., Bauernhofer, T., Wilders-Truschnig, M., Samonigg, H., Berghold, A. and Zatloukal, K. (1998) Soluble CD44 v5 and v6 in serum of patients with breast cancer. Correlation with expression of CD44 v5 and v6 variants in primary tumors and location of distant metastasis. *Breast Cancer Res Treat*. **47**: 29-40.

Ladeda, V., Aguirre Ghiso, J. A. and de Kier Joffé, E. B. (1998) Function and expression of CD44 during spreading migration and invasion of murine carcinoma cells. *Exp Cell Res*. **242**: 515-527.

Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**: 680–685 .

Leco, K. J., Khokha, R., Pavloff, N., Hawkes, S. P. and Edwards, D. R. (1994) Tissue inhibitor of metalloproteinases-3 (TIMP-3) is an extracellular matrix-associated protein with a distinctive pattern of expression in mouse cells and tissues. *J Biol Chem*. **269**: 9352-9360.

Lee, M. L., Kuo, F. C., Whitmore, G. A. and Sklar, J. (2000) Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. *Proc Natl Acad Sci U S A*. **97**: 9834-9839.

Lefebvre, M., Paulweber, B., Fajas, L., Woods, J., McCrary, C., Colombel, J. F., Najib, J., Fruchart, J. C., Datz, C., Vidal, H., Desreumaux, P. and Auwerx, J. (1999) Peroxisome proliferator-activated receptor gamma is induced during differentiation of colon epithelium cells. *J Endocrinol*. **162**: 331-340.

Leibovitz, A., Stinson, J. C., McCombs, W. B. 3rd., McCoy, C. E., Mazur, K. C. and Mabry, N. D. (1976) Classification of human colorectal adenocarcinoma cell lines. *Cancer Res*. **36**: 4562-4569.

Lesley, J., Hyman, R. and Kincade, P. W. (1993) CD44 and its interaction with extracellular matrix. *Adv. Immunol*. **54**: 271-335.

Leto, G., Tumminello, F. M., Pizzolanti, G., Montalto, G., Soresi, M., Carroccio, A., Ippolito, S. and Gebbia, N. (1997) Lysosomal aspartic and cysteine proteinases serum levels in patients with pancreatic cancer or pancreatitis. *Pancreas.* **14**: 22-27.

Leung, Y. F. and Cavalieri, D. (2003) Fundamentals of cDNA microarray data analysis. *Trends Genet.* **19**: 649-659.

Leung-Pineda, V. and Kilberg, M. S. (2002) Role of Sp1 and Sp3 in the nutrient-regulated expression of the human asparagine synthetase gene. *J Biol Chem.* **277**: 16585-16591.

Levine, A. J., Finlay, C. A. and Hinds, P. W. (2004) P53 is a tumor suppressor gene. *Cell.* **116**: 67-69.

Li, L., Yang, G., Ebara, S., Satoh, T., Nasu, Y., Timme, T. L., Ren, C., Wang, J., Tahir, S. A. and Thompson, T. C. (2001) Caveolin-1 mediates testosterone-stimulated survival/clonal growth and promotes metastatic activities in prostate cancer cells. *Cancer Res* **61**: 4386-4392.

Liotta, L. A., Tryggvason, K., Garbisa, S., Hart, I., Foltz, C. M. and Shafie, S. (1980) Metastatic potential correlates with enzymatic degradation of basement membrane collagen. *Nature* **284**: 67–68.

Liotta, L. (1990) The role of cellular proteases and their inhibitors in invasion and metastasis. Introductionary overview. *Cancer Metastasis Rev.* **9**: 285-287.

Liotta, L. A. and Kohn, E. C. (2003) Cancer's deadly signature. *Nat Genet.* **33**: 10-11.

Lipshutz, R. J., Fodor, S. P., Gingeras, T. R., Lockhart, D. J. (1999) High density synthetic oligonucleotide arrays. *Nat Genet.* **21**: 20-24.

Liu, H., Zang, C., Fenner, M. H., Possinger, K. and Elstner, E. (2003) PPARgamma ligands and ATRA inhibit the invasion of human breast cancer cells *in vitro*. *Breast Cancer Res Treat.* **79**: 63-74.

Lockhart, D. J., Dong, H., Byrne, M. C., Follettie, M. T., Gallo, M. V., Chee, M. S., Mittmann, M., Wang, C., Kobayashi, M., Horton, H. and Brown, E. L. (1996) Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat Biotechnol.* **14**:1675-1680.

Lu, Z., Ghosh, S., Wang, Z. and Hunter, T. (2004) Downregulation of caveolin-1 function by EGF leads to the loss of E-cadherin, increased transcriptional activity of beta-catenin, and enhanced tumor cell invasion. *Cancer Cell.* **4**: 499-515.

Lynch, C. S. and Matrisian, L. M. (2002) Matrix Metalloproteinases in tumour-host cell communication. *Differentiation.* **70**: 561-573.

Mackie, E. J. and Tucker, R. P. (1992) Tenascin in bone morphogenesis: expression by osteoblasts and cell type-specific expression of splice variants. *J Cell Sci.* **103**: 765-771.

Maniotis, A. J., Folberg, R., Hess, A., Seftor, E. A., Gardner, L. M., Pe'er, J., Trent, J. M., Meltzer, P. S., and Hendrix, M. J. (1999) Vascular channel formation by human melanoma cells *in vivo* and *in vitro*: vasculogenic mimicry. *Am. J. Pathol.* **155**: 739–752.

Mansen, A., Guardiola-Diaz, H., Rafter, J., Branting, C. and Gustafsson, J. A. (1996) Expression of the peroxisome proliferator-activated receptor (PPAR) in the mouse colonic mucosa. *Biochem Biophys Res Commun.* **222**: 844-851.

Martin, T. A., Harrison, G., Mansel, R. E. and Jiang, W. G. (2003) The role of the CD44/ezrin complex in cancer metastasis. *Crit Rev Oncol Hematol.* **46**: 165-186.

Massimino, M. L., Griffoni, C., Spisni, E., Toni, M. and Tomasi, V. (2002) Involvement of caveolae and caveolae-like domains in signalling, cell survival and angiogenesis. *Cell Signal.* **14**: 93-98.

Massova, I., Kotra, L. P., Fridman, R. and Mobashery, S. (1998) Matrix metalloproteinases: structures, evolution and diversification. *FASEB J.* **12**: 1075-1095.

Matrisian, L. M. (1992) The matrix-degrading metalloproteinases. *BioEssays* **14**: 455-463.

Mc Donnell, S., Chaudhry, V., Mansilla-Soto, J., Zeng, ZS., Shu, WP. and Guilem, JG. (1999) Metastatic and non-metastatic colorectal cancer (CRC) cells induce host metalloproteinase production *in vivo*. *Clin. Exp. Metastasis.* **17**: 341-349.

McCawley, L. J. and Matrisian, L. M. (2001) Matrix metalloproteinases: they're not just for matrix anymore! *Curr Opin Cell Biol.* **13**: 534-540.

Mithofer, K., Fernandez-del Castillo, C., Rattner, D. and Warshaw, A. L. (1998) Subcellular kinetics of early trypsinogen activation in acute rodent pancreatitis. *Am J Physiol.* **274**: 71-79.

Mitsiades, N., Yu, W. H., Poulaki, V., Tsokos, M. and Stamenkovic, I. (2001) Matrix metalloproteinase-7-mediated cleavage of Fas ligand protects tumor cells from chemotherapeutic drug cytotoxicity. *Cancer Res.* 61: 577-581.

Moerkerk, P., Arends, J. W., van Driel, M., de Bruine, A., de Goeij, A. and ten Kate, J. (1994) Type and number of Ki-ras point mutations relate to stage of human colorectal cancer. *Cancer Res.* **54**: 3376-3378.

Montgomery, A. M., Reisfeld, R. A. and Cheresh, D. A. (1994) Integrin alpha v beta 3 rescues melanoma cells from apoptosis in three-dimensional dermal collagen. *Proc. Natl. Acad. Sci. USA.* **91**: 8856-8860.

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

Mori, H., Tomari, T., Koshikawa, N., Kajita, M., Itoh, Y., Sato, H., Tojo, H., Yana, I. and Seiki, M. (2002) CD44 directs membrane-type 1 matrix metalloproteinase to lamellipodia by associating with its hemopexin-like domain. *EMBO J.* 21: 3949-3959.

Morino, N., Mimura, T., Hamasaki, K., Tobe, K., Ueki, K., Kikuchi, K., Takehara, K., Kadowaki, T., Yazaki, Y. and Nojima, Y. (1995) Matrix/integrin interaction activates the mitogen-activated protein kinase, p44erk-1 and p42erk-2. *J Biol Chem.* **270**: 269-273.

Moses, M. A. and Langer, R. A. (1991) A metalloproteinase inhibitor as an inhibitor of neovascularisation. *J. Cell. Biochem.* **47**: 230-235.

215

Mousses, S., Bubendorf, L., Wagner, U., Hostetter, G., Kononen, J., Cornelison, R., Goldberger, N., Elkahloun, A. G., Willi, N., Koivisto, P., Ferhle, W., Raffeld, M., Sauter, G. and Kallioniemi, O. P. (2002) Clinical validation of candidate genes associated with prostate cancer progression in the CWR22 model system using tissue microarrays. *Cancer Res.* **62**: 1256-1260.

Mulcahy, H. E., Duffy, M. J., Gibbons, D., McCarthy, P., Parfrey, N. A., O'Donoghue, D. P. and Sheahan, K. (1994) Urokinase-type plasminogen activator and outcome in Dukes' B colorectal cancer. *Lancet.* **344**: 583-584.

Mulder, J. W., Kruyt, P. M., Sewnath, M., Oosting, J., Seldenrijk, C. A., Weidema, W. F., Offerhaus, G. J. and Pals, S. T. (1994) Colorectal cancer prognosis and expression of exon-v6-containing CD44 proteins. *Lancet.* **344**: 1470-1472

Murnane, M. J., Sheahan, K., Ozdemirli, M. and Shuja, S. (1991) Stage-specific increases in cathepsin B messenger RNA content in human colorectal carcinoma. *Cancer Res.* **51**: 1137-1142.

Murray, D., Morrin, M. and McDonnell, S. (2004) Increased Invasion and Expression of MMP-9 in Human Colorectal Cell Lines by a CD44-dependent Mechanism. *Anticancer Res.* **24**: 489-494.

Murray, G. I., Duncan, M. E., O'Neill, P., Melvin, W. T. and Fothergill, J. E. (1996) Matrix metalloproteinase-1 is associated with poor prognosis in colorectal cancer. *Nat. Medicine.* **2**: 461-462.

Nagase, H. and Woesner, J. F. (1999) Matrix Metalloproteinases. *J Biol. Chem.* **274**: 21491-21494.

Nagata, M., Fujita, H., Ida, H., Hoshina, H., Inoue, T., Seki, Y., Ohnishi, M., Ohyama, T., Shingaki, S., Kaji, M., Saku, T. and Takagi, R. (2003) Identification of potential biomarkers of lymph node metastasis in oral squamous cell carcinoma by cDNA microarray analysis. *Int J Cancer.* **106**: 683-689.

Naor, D., Sionov, R. V. and IshShalom, D. (1997) CD44: Structure, function, and association with the malignant process. *Adv. Cancer Res.* **71**: 241-319.

Nelson, A. R., Fingleton, B., Rothenberg, M. L. and Matrisian, L. M. (2000) Matrix metalloproteinases: biologic activity and clinical implications. *J Clin Oncol.* **18**: 1135-1149

Newell, K. J.,Witty, J. P., Rodgers, W. H. and Matrisian, L. M (1994) Expression and localization of matrix-degrading metalloproteinases during colorectal tumorigenesis. *Molec Carcinogenesis* **10**:199 –206.

Nielsen, B. S., Timshel, S. and Kjeldsen, L. (1996) 92 kDa type IV collagenase (MMP-9) is expressed in neutrophils and macrophages but not in malignant epithelial cells in human colon cancer. *Int. J Cancer* **65**: 57-62.

Nitatori, T., Sato, N., Kominami, E. and Uchiyama, Y. (1996) Participation of cathepsins B, H, and L in perikaryal condensation of CA1 pyramidal neurons undergoing apoptosis after brief ischemia. *Adv Exp Med Biol.* **389**: 177-185.

Niu, J., Gu, X., Turton, J., Meldrum, C., Howard, E. W. and Agrez, M. (1998) Integrin-mediated signalling of gelatinase B secretion in colon cancer cells. Biochem *Biophys Res Commun.* **249**: 287-291.

Oettgen, H. F., Old, L. J., Boyse, E. A., Campbell, H. A., Philips, F. S., Clarkson, B. D., Tallal, L., Leeper, R. D., Schwartz, M. K. and Kim, J. H. (1967) Inhibition of leukemias in man by L-asparaginase. *Cancer Res.* **27**: 2619-2631.

Okada, H., Yoshida, J., Sokabe, M., Wakabayashi, T. and Hagiwara, M. (1996) Suppression of CD44 expression decreases migration and invasion of human glioma cells. *Int J Cancer.* **66**:255-60.

Okamoto, I., Kawano, Y., Tsuiki, H., Sasaki, J., Nakao, M., Matsumoto, M., Suga, M., Ando, M., Nakajima, M. and Saya, H. (1999) CD44 cleavage induced by a membrane-associated metalloprotease plays a critical role in tumor cell migration. *Oncogene* **18**: 1435-1446.

Okamoto, I., Kawano, Y., Murakami, D., Sasayama, T., Araki, N., Miki, T., Wong, A. J. and Saya, H. (2001) Proteolytic release of CD44 intracellular domain and its role in the CD44 signalling pathway. *J Cell Biol.* **155**: 755-762.

Okamoto, T., Suzuki, T. and Yamamoto, N. (2000) Microarray fabrication with covalent attachment of DNA using bubble jet technology. *Nat. Biotechnol.* **18**: 438–441.

Onisto, M., Garbisa, S., Caenazzo, C., Freda, M. P., Di Francesco, C., Nitti, D., Liotta, L. A. and Stetler-Stevenson, W. G. (1993) Reverse transcription-polymerase chain reaction phenotyping of metalloproteinases and inhibitors involved in tumor matrix invasion. *Diagn Mol Pathol.* **2**:74-80.

Ono, K., Tanaka, T., Tsunoda, T., Kitahara, O., Kihara, C., Okamoto, A., Ochiai, K., Takagi, T. and Nakamura, Y. (2000) Identification by cDNA microarray of genes involved in ovarian carcinogenesis. *Cancer Res.* **60**: 5007-5011.

Ornstein, D. L., MacNab, J. and Cohn, K. H. (1999) Evidence for tumor-host cooperation in regulating MMP-2 expression in human colon cancer. *Clin Exp Metastasis.* **17**: 205-212.

Ouchi, K., Sugawara, T., Ono, H., Fujiya, T., Kamiyama, Y., Kakugawa, Y., Mikuni, J. and Tateno, H. (1996) Histologic features and clinical significance of venous invasion in colorectal carcinoma with hepatic metastasis. *Cancer.* **78**: 2313-2317.

Ougolkov, A. V., Yamashita, K., Mai, M. and Minamoto, T. (2002) Oncogenic beta-catenin and MMP-7 (matrilysin) co segregate in late-stage clinical colon cancer. *Gastroenterology* **122**: 60-71.

Overall, C. M. (2002) Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites. *Mol. Biotechnol.* **22**(1): 51-86.

Paris, A., Strukelj, B., Pungercar, J., Renko, M., Dolenc, I. and Turk, V. (1995) Molecular cloning and sequence analysis of human preprocathepsin C. *FEBS Lett.* **369**: 326-330.

Park, C. C., Bissell, M. J. and Barcellos-Hoff, M. H. (2000) The influence of the microenvironment on the malignant phenotype. *Mol. Med. Today* **6**: 324-329.

Parle-McDermott, A., McWilliam, P., Tighe, O., Dunican, D. and Croke, D. T. (2000) Serial analysis of gene expression identifies putative metastasis-associated transcripts in colon tumour cell lines. *Br J Cancer.* **83**: 725-728.

Parsons, S. L., Watson, S. A., Collins, H. M., Grifffin, N. R., Clarke, P. A. and Steele R. J. C. (1998) Gelatinase (MMP-2 and -9) expression in gastrointestinal malignancy. *Brit J Cancer* **78**:1495 –1502.

Patlolla, J. M., Swamy, M. V., Raju, J. and Rao, C. V. (2004) Overexpression of caveolin-1 in experimental colon adenocarcinomas and human colon cancer cell lines. *Oncol Rep*. **11**: 957-963.

Pavan, L., Tarrade, A., Hermouet, A., Delouis, C., Titeux, M., Vidaud, M., Therond, P., Evain-Brion, D. and Fournier, T. (2003) Human invasive trophoblasts transformed with simian virus 40 provide a new tool to study the role of PPARgamma in cell invasion process. *Carcinogenesis*. **24**: 1325-1336.

Pereira, P. A., Rubenthiran, U., Kaneko, M., Jothy, S. and Smith, A. J. (2001) CD44s expression mitigates the phenotype of human colorectal cancer hepatic metastases. *Anticancer Res*. **21**: 2713-2717.

Perou, C. M., Jeffrey, S. S., van de Rijn, M., Rees, C. A., Eisen, M. B., Ross, D. T., Pergamenschikov, A., Williams, C. F., Zhu, S. X., Lee, J. C., Lashkari, D., Shalon, D., Brown, P. O. and Botstein, D. (1999) Distinctive gene expression patterns in human mammary epithelial cells and breast cancers. *Proc Natl Acad Sci U S A*. **96**: 9212-9217.

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

Phillips, G. R., Krushel, L. A. and Crossin, K. L. (1998) Domains of tenascin involved in glioma migration. *Cell Sci*. **111**: 1095-1104.

Picker, L. J., Nakache, M. and Butcher, E. C. (1989) Monoclonal antibodies to human lymphocyte homing receptors define a novel class of adhesion molecules on diverse cell types. *J Cell Biol*. **109**:927-937.

Ponta, H., Wainwright, D. and Herrlich, P. (1998) The CD44 protein family. *Int J Biochem Cell Biol*. **30**: 299-305.

Poste, G. and Fidler, I. J. (1980) The pathogenesis of cancer metastasis. *Nature*. **283**: 139-146.

Powell, W. C., Fingleton, B., Wilson, C. L., Boothby, M. and Matrisian, L. M. (1999) The metalloproteinase matrilysin proteolytically generates active soluble Fas ligand and potentiates epithelial cell apoptosis. *Curr Biol.* **9**: 1441-1447.

Quackenbush, J. (2000) Microarray data normalization and transformation. *Nature Genet.* **32**: 496-501.

Racine, C., Belanger, M., Hirabayashi, H., Boucher, M., Chakir, J. and Couet, J. (1999) Reduction of caveolin 1 gene expression in lung carcinoma cell lines. *Biochem Biophys Res Commun.* **255**: 580-586.

Rajeevan, M. S., Vernon, S. D., Taysavang, N. and Unger, E. R. (2001) Validation of array-based gene expression profiles by real-time (kinetic) RT-PCR. *J Mol Diagn.* **3**: 26-31.

Ramaswamy, S., Ross, K. N., Lander, E. S. and Golub, T. R. (2003) A molecular signature of metastasis in primary solid tumors. *Nat Genet.* **33**: 49-54.

Rao, N. V., Rao, G.V. and Hoidal, J. R. (1997) Human dipeptidyl-peptidase I. Gene characterization, localization, and expression. *J Biol Chem.* 272: 10260-10265.

Ratnikov, B. I., Rozanov, D. V., Postnova, T. I., Baciu, P. G., Zhang, H., DiScipio, R. G., Chestukhina, G. G., Smith, J. W., Deryugina, E. I., Strongin, A. Y. (2002) An alternative processing of integrin alpha(v) subunit in tumor cells by membrane type-1 matrix metalloproteinase. *J Biol Chem.* **277**: 7377-7385.

Reeder, J. A., Gotley, D. C., Walsh, M. D., Fawcett, J. and Antalis, T. M. (1998) Expression of antisense CD44 variant 6 inhibits colorectal tumour metastasis and tumour growth in a wound environment. *Cancer Res.* **58**: 3719-3726.

Risio, M., Casorzo, L., Chiecchio, L., De Rosa, G. and Rossini, F. P. (2003) Deletions of 17p are associated with transition from early to advanced colorectal cancer. *Cancer Genet Cytogenet.* **147**: 44-49.

Roeb, E., Dietrich, C. G., Winograd, R., Arndt, M., Breuer, B., Fass, J., Schumpelick, V. and Matern, S. (2001) Activity and cellular origin of gelatinases in patients with colon and rectalcarcinoma differential activity of matrix metalloproteinase-9. *Cancer* **92**: 2680 – 2691.

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

Rosales, C., O'Brien, V., Kornberg, L. and Juliano, R. (1995) Signal transduction by cell adhesion receptors. *Biochim Biophys Acta.* **1242**: 77-98.

Rosen, E. D and Spiegelman, B. M. (2001) PPARgamma : a nuclear regulator of metabolism, differentiation, and cell growth. *J Biol Chem.* **276**: 37731-37734.

Rosenbaum, P.R. (2001) Replicating effects and biases. *Am. Stat.* **55**: 223–227.

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

Ruoslahti, E. (1991) Integrins. *J. Clin. Invest.* **87**: 1-5.

Ruoslahti, E. (1996) How cancer spreads. *Sci. Am.* **275**: 72-77.

Saez, E., Tontonoz, P., Nelson, M. C., Alvarez, J. G., Ming, U. T., Baird, S. M., Thomazy, V. A., Evans, R. M. (1998) Activators of the nuclear receptor PPARgamma enhance colon polyp formation. *Nat Med.* **4**: 1058-1061.

Sarraf, P., Mueller, E., Jones, D., King, F. J., DeAngelo, D. J., Partridge, J. B., Holden, S. A., Chen, L. B., Singer, S., Fletcher, C. and Spiegelman, B. M. (1998) Differentiation and reversal of malignant changes in colon cancer through PPARgamma. *Nat Med.* **4**: 1046-1052.

Sato, H., Takino, T., Okada, Y., Cao, J., Shinagawa, A., Yamamoto, E. and Seiki, M. (1994) A matrix metaloproteinase expressed on the cell surface of invasive tumour cells. *Nature* **370**: 61-65.

Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467-470.

Schena, M., Shalon, D., Heller R., Chai, A., Brown, P. O. and Davis, R. W. (1996) Parallel human genome analysis: microarray-based expression monitoring of 1000 genes. *Proc Natl Acad Sci U S A*. **93**: 10614-10619.

Schenk, S., Chiquet-Ehrismann, R. and Battegay, E. J. (1999) The fibrinogen globe of tenascin-C promotes basic fibroblast growth factor-induced endothelial cell elongation. *Mol Biol Cell*. **10**: 2933-2943.

Scherf, U., Ross, D. T., Waltham, M., Smith, L. H., Lee, J. K., Tanabe, L., Kohn, K. W., Reinhold, W. C., Myers, T. G., Andrews, D. T., Scudiero, D. A., Eisen, M. B., Sausville, E. A., Pommier, Y., Botstein, D., Brown, P. O. and Weinstein, J. N. (2000) A gene expression database for the molecular pharmacology of cancer. *Nat Genet*. **24**: 236-244.

Schuchhardt, J., Beule, D., Malik, A., Wolski, E., Eickhoff, H., Lehrach, H. and Herzel, H. (2000) Normalization strategies for cDNA microarrays. *Nucleic Acids Res*. **28**: e47.

Schwartz, M. A. (1992) Transmembrane signalling by integrins. *Trends Cell Biol*. **2**: 304-308.

Schwartz, M. A., Schaller, M. D., and Ginsberg, M. H. (1995) Integrins: emerging paradigms of signal transduction. *Annu. Rev. Cell Dev. Biol*. **11**: 549-599.

Scian, M. J., Stagliano, K. E., Deb, D., Ellis, M. A., Carchman, E. H., Das, A., Valerie, K., Deb, S. P. and Deb, S. (2004) Tumor-derived p53 mutants induce oncogenesis by transactivating growth-promoting genes. *Oncogene*. **23**: 4430-4443

Seiki, M. (1999) Membrane type matrix metalloproteinases, *Apmis* **107**: 137-143.

Sgroi, D. C., Teng, S., Robinson, G., LeVangie, R., Hudson, J. R. Jr. and Elkahloun, A. G. (1999) *In vivo* gene expression profile analysis of human breast cancer progression. *Cancer Res*. **59**: 5656-5661.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Biotechnology* **24**:122-39.

Southern, E. Mir, K. and Shchepinov, M. (1999) Molecular interactions on microarrays. *Nat Genet*. 1999 **1**: 5-9

Sporn, M. B. (1996) The war on cancer. *Lancet* **347**: 1377-1381.

Springman, E. B., Angleton, E. L., Birkedal-Hansen, H. and Van Wart, H. (1990) Multiple modes of activation of latent human fibroblast collagenase: evidence for the role of a cyc 73 active-site zinc complex in latency and a "cysteine switch" mechanism for activation, *Proc Natl. Acad. Sci. USA* **87**: 364-368.

Steffensen, B., Wallon, U. M. and Overall, C. M. (1995) Extracellular matrix binding properties of recombinant fibronectin type II-like modules of human 72-kDa gelatinase/type IV collagenase. High affinity binding to native type I collagen but not native type IV collagen. *J. Biol. Chem.* **270**: 11555-11566.

Stegmaier, K., Ross, K. N., Colavito, S. A., O'Malley, S., Stockwell, B. R. and Golub, T. R. (2004) Gene expression-based high-throughput screening (GE-HTS) and application to leukemia differentiation. *Nat Genet.* **36**: 257-263.

Sternlicht, M. D. and Werb, Z. (1999) ECM proteinases. In *Guidebook to the extracellular matrix, anchor and adhesion proteins*, ed. T. Kries and R. Vale, pp. 503-562. Oxford, UK: Oxford Univ. Press.

Sternlicht, M. D. and Bergers, G. (2000) Matrix metalloproteinases as emerging targets in anti-cancer therapy: status and prospects. *Emerging Ther. Targets* **4**: 609-33.

Sternlicht, M. D., Bissell, M. J. and Werb, Z. (2000) The matrix metalloproteinase stromelysin-1 acts as a natural mammary tumor promoter. *Oncogene.* **19**: 1102-1113.

Sternlicht, M. D and Werb, Z. (2001) How matrix metalloproteinases regulate cell behavior. *Annu. Rev. Cell. Dev. Biol.* **17**: 463-516.

Stetler-Stevenson, W. G., Hewitt, R. and Corcoran, M. (1996) Matrix metalloproteinases and tumor invasion: from correlation and causality to the clinic. *Semin Cancer Biol.* **7**: 147-154.

Stetler-Stevenson, W. G. (1999) Matrix metalloproteinases in angiogenesis: a moving target for therapeutic intervention. *J. Clin. Invest.* **103**: 1237-1241.

Stocker, W., Grams, F., Baumann, U., Reinemer, P., Gomis-Ruth, F. X., McKay, D. B. and Bode, W. (1995) The metzincins-topological and sequential relations between the astacins, adamalysins, serralysins, and matrixins (collagenases) define a superfamily of zinc-peptidases. *Protein Sci.* **4**(5): 823-840.

Story, M. D., Voehringer, D. W., Stephens, L. C. and Meyn, R. E. (1993) L-asparaginase kills lymphoma cells by apoptosis. *Cancer Chemother Pharmacol.* **32**: 129-133.

Strongin, A. Y., Collier, I., Bannikov, G., Marmer, B. L., Grant, G. A. and Goldberg, G. I. (1995) Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J Biol Chem.* **270**: 5331-5338.

Su, A. I., Welsh, J. B., Sapinoso, L. M., Kern, S. G., Dimitrov, P., Lapp, H., Schultz, P. G., Powell, S. M., Moskaluk, C. A., Frierson, H. F. Jr. and Hampton, G. M. (2001) Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res.* **61**: 7388-7393.

Sutow, W. W., Garcia, F., Starling, K. A., Williams, T. E., Lane, D. M. and Gehan, E. A. (1971) L-asparaginase therapy in children with advanced leukemia. The Southwest Cancer Chemotherapy Study Group. *Cancer.* **28**: 819-824.

Suzuoki, M., Miyamoto, M., Kato, K., Hiraoka, K., Oshikiri, T., Nakakubo, Y., Fukunaga, A., Shichinohe, T., Shinohara, T., Itoh, T., Kondo, S. and Katoh, H. (2002) Impact of caveolin-1 expression on prognosis of pancreatic ductal adenocarcinoma. *Br J Cancer.* **87**: 1140-1144.

Swindle, C. S., Tran, K. T., Johnson, T. D., Banerjee, P., Mayes, A. M., Griffith, L. and Wells, A. (2001) Epidermal growth factor (EGF)-like repeats of human tenascin-C as ligands for EGF receptor. *J Cell Biol.* **154**: 459-468.

Taher, T. E. I., Smit, L., Griffioen, A. W., Schider-Tol, E, J. M., Borst, J. and Pals, S. T. (1996) Signalling through CD44 is mediated by tyrosine kinases. *J Biol Chem.* **271**: 2863-2867.

Tahir, S. A., Yang, G., Ebara, S., Timme, T. L., Satoh, T., Li, L., Goltsov, A., Ittmann, M., Morrisett, J. D. and Thompson, T. C. (2001) Secreted caveolin-1 stimulates cell survival/clonal growth and contributes to metastasis in androgen-insensitive prostate cancer. *Cancer Res.* **61**: 3882-3885.

Tahir, S. A., Ren, C., Timme, T. L., Gdor, Y., Hoogeveen, R., Morrisett, J. D., Frolov, A., Ayala, G., Wheeler, T. M., Thompson, T. C. (2003) Development of an immunoassay for serum caveolin-1: a novel biomarker for prostate cancer. *Clin Cancer Res.* **9**: 3653-3659.

Takahasi, K., Eto, H. and Tanabe, K. (1999) Involvement of CD44 in matrix metalloproteinase-2 regulation in human melanoma cells. *Int. J Cancer* **80**: 387-395.

Taniguchi, M., Miura, K., Iwao, H. and Yamanaka, S. (2001) Quantitative assessment of DNA microarrays-comparison with Northern blot analyses. *Genomics.* **71**: 34-39.

Tarafa, G., Villanueva, A., Farre, L., Rodriguez, J., Musulen, E., Reyes, G., Seminago, R., Olmedo, E., Paules, A. B., Peinado, M. A., Bachs, O. and Capella, G. (2000) DCC and SMAD4 alterations in human colorectal and pancreatic tumor dissemination. *Oncogene.* **19**: 546-555.

Taylor, E., Cogdell, D., Coombes, K., Hu, L., Ramdas, L., Tabor, A., Hamilton, S. and Zhang, W. (2001) Sequence verification as quality-control step for production of cDNA microarrays. *Biotechniques.* **31**: 62-65.

The Genome International Sequencing Consortium. (2001) Initial sequencing and analysis of the human genome. *Nature* **409**: 860-921.

Thomas, G. J., Lewis, M. P., Whawell, S. A., Russell, A., Sheppard, D., Hart, I. R., Speight, P. M. and Marshall, J. F. (2001) Expression of the alphavbeta6 integrin promotes migration and invasion in squamous carcinoma cells. *J Invest Dermatol* **117**: 67-73.

Thompson, T. C., Timme, T. L., Li, L. and Goltsov, A. (1999) Caveolin-1, a metastasis-related gene that promotes cell survival in prostate cancer. *Apoptosis.* **4**: 233-237.

Tokuriki, M., Noda, I., Saito, T., Narita, N., Sunaga, H., Tsuzuki, H., Ohtsubo, T., Fujieda, S. and Saito, H. (2003) Gene expression analysis of human middle ear cholesteatoma using complementary DNA arrays. Laryngoscope. **113**: 808-814.

Tontonoz, P., Singer, S., Forman, B. M., Sarraf, P., Fletcher, J. A., Fletcher, C. D., Brun, R. P., Mueller, E., Altiok, S., Oppenheim, H., Evans, R. M. and Spiegelman, B. M. (1997) Terminal differentiation of human liposarcoma cells induced by ligands for peroxisome proliferator-activated receptor gamma and the retinoid X receptor. *Proc Natl Acad Sci U S A*. **94**: 237-241.

Tremble, P., Chiquet-Ehrismann, R. and Werb, Z. (1994) The extracellular matrix ligands fibronectin and tenascin collaborate in regulating collagenase gene expression in fibroblasts. *Mol Biol Cell*. **5**: 439-453.

Tsukita, S., Oishi, K., Sato, N., Sagara, J., Kawai, A. and Tsukita, S. (1994) ERM family members as molecular linkers between the cell surface glycoprotein CD44 and actin-based cytoskeletons. *J Cell Biol*. **126**: 391-401.

Uray, I. P., Connelly, J. H., Frazier, O. H., Taegtmeyer, H. and Davies, P. J. (2003) Mechanical unloading increases caveolin expression in the failing human heart. *Cardiovasc Res* **59**: 57-66.

Van Weering, D. H. J., Baas, P. D. and Bos, J. L. (1993) A PCR-based method for the analysis of human CD44 splice products. *PCR Methods Appl* **3**: 100-106.

Virtaneva, K., Wright, F. A., Tanner, S. M., Yuan, B., Lemon, W. J., Caligiuri, M. A., Bloomfield, C. D., de La Chapelle, A. and Krahe, R. (2001) Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc Natl Acad Sci U S A*. **98**: 1124-1129.

Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura, Y., White, R., Smits, A. M. and Bos, J. L. (1989) Genetic alterations during colorectal-tumor development. *N Engl J Med*. **319**: 525-532.

Wagenaar-Miller, R. A., Gorden, L. and Matrisian, L. M. (2004) Matrix metalloproteinases in colorectal cancer: Is it worth talking about? *Cancer Metastasis Rev*.**23**: 119–135.

Walker, N. J. (2002) A Technique Whose Time Has Come. *Science*. **296**: 557-558.

Wang, E., Miller, L. D., Ohnmacht, G. A., Liu, E. T., Marincola, F. M. (2000) High-fidelity mRNA amplification for gene profiling. *Nat Biotechnol*. 18: 457-459.

Weg-Remers, S., Anders, M. von Lampe, B. Riecken, E. O. Schuder, G., Feifel, G., Zeitz, M. and Stallmach, A. (1998) Decreased expression of CD44 splicing variants in advanced colorectal carcinomas. *Eur J Cancer.* **34**: 1607-1611.

Weigelt, B., Glas, A. M., Wessels, L. F., Witteveen, A. T., Peterse, J. L and, van't Veer, L. J. (2003) Gene expression profiles of primary breast tumors maintained in distant metastases. *Proc Natl Acad Sci U S A.* **100**: 15901-15905.

Welsh, J. B., Zarrinkar, P. P., Sapinoso, L. M., Kern, S. G., Behling, C. A., Monk, B. J., Lockhart, D. J., Burger, R. A. and Hampton, G. M. (2001) Analysis of gene expression profiles in normal and neoplastic ovarian tissue samples identifies candidate molecular markers of epithelial ovarian cancer. *Proc Natl Acad Sci U S A.* **98**: 1176-1181.

Werb, Z. (1997) ECM and cell surface proteolysis: regulating cellular ecology. *Cell* **91**: 439-442.

Whittaker, M., Floyd, C. D., Brown, P. and Gearing, A. J. (1999) Design and therapeutic application of matrix metalloproteinase inhibitors. *Chem Rev.* **99**: 2735-2776.

Wielenga, V. J., Heider, K. H., Offerhaus, G. J., Adolf, G. R., van den Berg, F. M., Ponta, H., Herrlich, P. and Pals, S. T. (1993) Expression of CD44 variant proteins in human colorectal cancer is related to tumor progression. *Cancer Res.* **53**: 4754-4756.

Wielenga, V. J., van der Voort, R., Mulder, J. W., Kruyt, P. M., Weidema, W. F.. Oosting, J., Seldenrijk, C. A., van Krimpen, C., Offerhaus, G. J. and Pals, S. T. (1998) CD44 splice variants as prognostic markers in colorectal cancer. *Scand J Gastroenterol.* **33**: 82-87.

Wilhelm, S. M., Collier, I. E., Marmer, B. L., Eisen, A. Z., Grant, G. A. and Goldberg, G. I. (1989) SV40-transformed human lung fibroblasts secrete a 92-kDa type IV collagenase which is identical to that secreted by normal human macrophages. *J. Biol. Chem.* **264**: 17213-17221.

Wilson C. L. and Matrisian L. M. (1996) Matrilysin: An epithelial matrix metalloproteinase with potentially novel functions. *Int. J. Biochem. Cell. Biol.* **28**: 123-136.

Wilson, C. L., Heppner, K. J., Labosky, P. A., Hogan, B. L. and Matrisian, L. M. (1997) Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. *Proc. Natl. Acad. Sci. USA.* **94**: 1402-1407.

Wilson, K. E., Langdon, S. P., Lessells, A. M. and Miller, W. R. (1996) Expression of the extracellular matrix protein tenascin in malignant and benign ovarian tumours. *Br J Cancer.* **74**: 999-1004.

Witty, J. P., McDonnell, S., Newell, K. J., Cannon, P., Navre, M., Tressler, R. J. and Matrisian. L. M. (1994) Modulation of matrilysin levels in colon carcinoma cell lines affects tumorigenicity *in vivo. Cancer Res.* **54**(17): 4805-4812.

Woessner, J. F. and Nagase, H. (2000) *Matrix Metalloproteinases and TIMPs.* New York: Oxford Univ. Press.

Woodhouse, E. C., Chuaqui, R. F. and Liotta, L. A. (1997) General mechanisms of metastasis. *Cancer* **80**: 1529-1537.

Yakubenko, V. P., Lobb, R. R., Plow, E. F. and Ugarova, T. P. (2000) Differential induction of gelatinase B (MMP-9) and gelatinase A (MMP-2) in T lymphocytes upon alpha(4)beta(1)-mediated adhesion to VCAM-1 and the CS-1 peptide of fibronectin. *Exp Cell Res.* **260**: 73-84.

Yamamoto, H., Itoh, F., Hinoda, Y., Senota, A., Yoshimoto, M., Nakamura, H., Imai, K. and Yachi, A. (1994) Expression of matrilysin mRNA in colorectal adenomas and its induction by truncated fibronectin. *Biochem. Biophys. Res. Commun.* 201: 657-664.

Yana, I and Weiss, S. J. (2000) Regulation of membrane type-1 matrix metalloproteinase activation by proprotein convertases. *Mol Biol Cell.* **11**: 2387-2401.

Yang, G., Truong, L. D., Timme, T. L., Ren, C., Wheeler, T. M., Park, S. H., Nasu, Y., Bangma, C. H., Kattan, M. W., Scardino, P. T. and Thompson, T. C. (1998) Elevated expression of caveolin is associated with prostate and breast cancer. *Clin Cancer Res.* **4**: 1873-1880.

Yang, I. V., Chen, E., Hasseman, J. P., Liang, W., Frank, B. C., Wang, S., Sharov, V., Saeed, A. I., White, J., Li, J., Lee, N. H., Yeatman, T. J. and Quackenbush, J. (2002a) Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol.* **3**: 62.1-62.13.

Yang, Y. H. and Speed, T. P. (2002) Design issues for cDNA microarray experiments. *Nat Rev Genet.* **3**: 579-588.

Yang, Y. H., Dudoit, S., Luu, P., Lin, D. M., Peng, V., Ngai, J. and Speed, T. P. (2002b) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res.* **30**: e15.

Yang, Y. H., Dudoit, S., Speed, T. P. and Callow, M. J. (2002c) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Statistica Sinica,* **12**: 111-139.

Yoshida, T., Matsumoto, E., Hanamura, N., Kalembeyi, I., Katsuta, K., Ishihara, A. and Sakakura, T. (1997) Co-expression of tenascin and fibronectin in epithelial and stromal cells of benign lesions and ductal carcinomas in the human breast. *J Pathol.* **182**: 421-428.

Yu, Q and Stamenkovic, I. (1999) Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion. *Genes Dev.* **13**: 35-48.

Yu, Q. and Stamenkovic, I. (2000) Cell surface localised matrix metalloproteinase-9 proteolytically activates TGF-β and promotes tumour invasion and angiogenesis. *Genes. Dev.* **14**: 163-176.

Yu, W. H., Woessner, J. F. Jr., McNeish, J. D. and Stamenkovic, I. (2002) CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulates female reproductive organ remodeling. *Genes Dev.* **16**: 307-323.

Zembutsu, H., Ohnishi, Y., Tsunoda, T., Furukawa, Y., Katagiri, T., Ueyama, Y., Tamaoki, N., Nomura, T., Kitahara, O., Yanagawa, R., Hirata, K. and Nakamura, Y. (2002) Genome-wide cDNA microarray screening to correlate gene expression profiles with sensitivity of 85 human cancer xenografts to anticancer drugs. *Cancer Res.* **62**:518-527.

Zeng, Z. S. and Guillem, J. G. (1996) Colocalisation of matrix metalloproteinase-9 mRNA and protein in human colorectal cancer stromal cells. *Br. J Cancer* **74**: 1161-1167.

Zetter, B. R. (1993) Adhesion molecules in tumor metastasis. *Semin Cancer Biol.* **4**: 219-229.

Zhang, F., Qi, L. and Chen, H. (2001) Value of P-glycoprotein and glutathione S-transferase-pi as chemoresistant indicators in ovarian cancers. *Zhonghua Zhong Liu Za Zhi.*: 313-316.

Zhang, Y., Thant, A. A., Machida, K., Ichigotani, Y., Naito, Y., Hiraiwa, Y., Senga, T., Sohara, Y., Matsuda, S. and Hamaguchi, M. (2002) Hyaluron-CD44s signalling regulates matrix metalloproteinase-2 secretion in a human lung carcinoma cell line QG90. *Cancer Res.* **62**(14): 3962-3965.

Appendix A

Genes with significant changes in expression

Genes with significant 2 fold or more increased expression in SW620 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
R32457	9.781 (4.585 to 45.874)	hypothetical protein FLJ23516
R93124	9.323 (4.377 to 20.716)	aldo-keto reductase family 1, member C1
AA912032	6.749 (2.825 to 14.293)	ESTs, Weakly similar to 2108276A ssDNA-binding protein
AA459407	6.015 (2.115 to 62.53)	SWI/SNF related, matrix associated, actin dependent regulator of chromatin
AA862465	5.822 (5.355 to 6.681)	alpha-2-glycoprotein 1, zinc
AA913480	5.708 (0.889 to 28.721)	killer cell lectin-like receptor subfamily C, member 1
R37743	5.089 (3.208 to 7.518)	T54 protein
H22922	5.041 (2.732 to 7.143)	manic fringe homolog (Drosophila)
R39239	4.897 (3.667 to 9.456)	tenascin C (hexabrachion)
AA495936	4.365 (3.02 to 6.401)	microsomal glutathione S-transferase 1
R28294	4.097 (2.911 to 5.627)	glycine cleavage system protein H (aminomethyl carrier)
R16134	4.069 (2.521 to 6.342)	transmembrane 4 superfamily member 11 (plasmolipin)
H94063	4.052 (2.565 to 6.329)	chromosome 13 open reading frame 1
AA451851	3.999 (2.551 to 8.188)	single-stranded DNA binding protein 2
N70463	3.958 (2.664 to 6.885)	B-cell translocation gene 1, anti-proliferative
AA436565	3.757 (2.196 to 8.12)	pleckstrin homology-like domain, family A, member 1
R51331	3.74 (1.798 to 5.821)	hypothetical protein DKFZp761B0514
R60981	3.6 (2.223 to 5.532)	ESTs, Highly similar to T00343 hypothetical protein KIAA0584 - human (fragment)
AA088420	3.588 (2.81 to 5.428)	peroxisome proliferative activated receptor, gamma
AA421048	3.578 (2.763 to 4.118)	ESTs
H79534	3.549 (2.038 to 6.782)	hemoglobin, epsilon 1
AA173755	3.526 (0.541 to 8.139)	roundabout, axon guidance receptor, homolog 1 (Drosophila)
R59936	3.497 (2.325 to 4.735)	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 43101.
N48804	3.479 (2.663 to 6.234)	ESTs
W46900	3.436 (2.319 to 4.711)	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
AA055052	3.427 (2.11 to 4.889)	Homo sapiens cDNA FLJ11639 fis, clone HEMBA1004327.
AA779165	3.421 (2.237 to 5.184)	ADP-ribosylation factor-like 4
AA460463	3.416 (1.328 to 7.632)	cytokine-like protein C17
R15922	3.345 (1.963 to 6.949)	ESTs
AA036952	3.302 (1.275 to 11.652)	hypothetical protein FLJ30973
AA701963	3.288 (1.709 to 5.365)	aldo-keto reductase family 1, member B1 (aldose reductase)
AA935273	3.279 (1.052 to 12.46)	chemokine (C-X-C motif) ligand 3
AA487893	3.243 (1.567 to 5.931)	transmembrane 4 superfamily member 1
AA113347	3.203 (2.328 to 5.949)	tousled-like kinase 1
R68272	3.203 (1.475 to 6.679)	KIAA1796 protein
R60170	3.167 (1.944 to 4.153)	guanine deaminase
R98936	3.092 (2.104 to 4.458)	membrane metallo-endopeptidase (neutral endopeptidase, enkephalinase, CALLA, CD10)
H16537	3.09 (2.88 to 3.328)	ESTs
W42723	3.088 (2.247 to 4.818)	chemokine (C-X-C motif) ligand 1 (melanoma growth stimulating activity, alpha)
N26928	3.072 (2.407 to 3.901)	hypothetical protein from EUROIMAGE 588495
AA916325	21.624 (5.818 to 537.576)	aldo-keto reductase family 1, member C3 (3-alpha hydroxysteroid dehydrogenase, type II)
AA521339	2.859 (2.332 to 3.427)	Homo sapiens, clone IMAGE:4347271, mRNA
H56655	2.853 (1.961 to 3.73)	KIAA1727 protein
N57754	2.819 (1.989 to 4.267)	Data not found
R61847	2.785 (1.975 to 5.641)	glutamate receptor, ionotropic, N-methyl-D-aspartate 3A
R67081	2.783 (2.318 to 3.899)	hypothetical protein LOC115749
AA478585	2.776 (1.219 to 4.503)	butyrophilin, subfamily 3, member A3
T62546	2.767 (1.857 to 4.538)	KIAA0174 gene product
AA453759	2.739 (1.91 to 3.806)	sprouty homolog 2 (Drosophila)
AA055835	2.725 (1.605 to 3.346)	caveolin 1, caveolae protein, 22kDa
T74192	2.71 (2.272 to 4.094)	protein S (alpha)
AA464605	2.701 (1.156 to 5.293)	kidney ankyrin repeat-containing protein
AA434373	2.694 (1.871 to 5.425)	E74-like factor 3 (ets domain transcription factor, epithelial-specific)
W37905	2.678 (1.333 to 6.587)	hypothetical protein dJ462O23.2

Appendix A

AA644088	2.676 (1.991 to 3.645)	cathepsin C
AA486220	2.67 (2.118 to 3.791)	lysyl-tRNA synthetase
N28268	2.665 (1.361 to 5.217)	Homo sapiens mRNA; cDNA DKFZp564F112 (from clone DKFZp564F112)
T66320	2.648 (1.263 to 5.092)	glutathione S-transferase A4
H69335	2.635 (1.411 to 5.18)	Pirin
AA452270	2.626 (1.894 to 4.892)	ESTs
AA598640	2.622 (1.783 to 3.324)	midline 1 (Opitz/BBB syndrome)
AA931725	2.622 (1.552 to 4.029)	SPARC related modular calcium binding 2
R39446	2.605 (1.903 to 3.566)	ATP-binding cassette, sub-family G (WHITE), member 1
W80482	2.602 (1.44 to 3.725)	Homo sapiens cDNA FLJ13598 fis, clone PLACE1009921.
AA115761	2.601 (1.501 to 3.791)	Data not found
AA621155	2.569 (1.748 to 3.852)	mutS homolog 5 (E. coll)
W79301	2.569 (1.029 to 5.977)	arylhydrocarbon receptor repressor
AA018151	2.564 (1.734 to 3.353)	ESTs
R63811	2.561 (1.621 to 4.576)	LIM and cysteine-rich domains 1
AA910218	2.544 (1.008 to 4.668)	monogenic, audiogenic seizure susceptibility 1 homolog (mouse)
AA707922	2.533 (1.435 to 4.141)	phosphodiesterase 6H, cGMP-specific, cone, gamma
AA857364	2.521 (1.375 to 4.721)	Homo sapiens cDNA: FLJ22864 fis, clone KAT02164.
AA454990	2.504 (2.082 to 2.767)	BTB (POZ) domain containing 3
R40946	2.49 (1.953 to 5.041)	crystallin, zeta (quinone reductase)
AA427719	2.489 (1.469 to 4.285)	KIAA1201 protein
AI014877	2.483 (1.298 to 8.294)	ESTs
N49284	2.48 (1.02 to 4.333)	v-myb myeloblastosis viral oncogene homolog (avian)
N71028	2.465 (2.153 to 2.919)	membrane-spanning 4-domains, subfamily A, member 6A
H08862	2.456 (1.21 to 4.691)	chromosome 1 open reading frame 17
R95691	2.45 (2.25 to 2.836)	Data not found
AA150263	2.445 (0.916 to 3.947)	placenta-specific 8
AA131421	2.441 (1.624 to 3.536)	Homo sapiens mRNA, 3'UTR, up-regulated by BCG-CWS.
R71627	2.434 (1.202 to 3.545)	chromosome 6 open reading frame 37
H65988	2.387 (1.723 to 3.173)	ESTs
R61187	2.381 (1.678 to 3.726)	ESTs
AA625979	2.378 (1.721 to 2.955)	hypothetical protein FLJ20607
H13623	2.374 (1.331 to 4.696)	epidermal growth factor receptor pathway substrate 8
N23606	2.37 (1.739 to 3.168)	Homo sapiens mRNA full length insert cDNA clone EUROIMAGE 1977059
AA449821	2.369 (1.776 to 2.996)	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]
N75595	2.367 (1.72 to 3.181)	nuclear transport factor 2
H93050	2.357 (1.68 to 3.567)	ESTs, Weakly similar to hypothetical protein FLJ20378 [Homo sapiens] [H.sapiens]
H22566	2.352 (1.566 to 4.145)	dachshund homolog (Drosophila)
W37780	2.35 (1.66 to 2.941)	RNA-binding protein
AA620867	2.341 (1.442 to 3.993)	forkhead box P1
R15911	2.34 (1.503 to 5.621)	Homo sapiens clone 24461 mRNA sequence
AA778448	2.336 (1.724 to 3.307)	protein kinase, X-linked
N57998	2.326 (1.256 to 3.432)	ESTs
AA452113	2.323 (1.656 to 3.216)	KIAA1917 protein
AA449837	2.322 (1.849 to 3.589)	HTPAP protein
AA450336	2.318 (1.139 to 3.749)	neurexin 3
AA679150	2.298 (1.502 to 2.943)	BTB (POZ) domain containing 3
R42736	2.298 (1.064 to 3.238)	myelin transcription factor 1
R41565	2.292 (1.32 to 3.185)	ESTs
AA479978	2.289 (1.648 to 2.95)	Data not found
W52386	2.281 (1.217 to 3.742)	Data not found
R16241	2.28 (1.669 to 4.622)	ESTs
AA469965	2.271 (1.378 to 4.575)	lymphocyte-specific protein tyrosine kinase
H16824	2.27 (1.552 to 3.099)	kinesin-associated protein 3
AI240593	2.255 (1.968 to 2.618)	ESTs
AA283699	2.246 (1.315 to 4.034)	Homo sapiens, clone IMAGE:5264268, mRNA
AA933890 | AI822026	2.24 (1.491 to 3.072)	Data not found
AA666418	2.239 (1.682 to 2.734)	ALEX1 protein

2

Appendix A

AA937029	2.239 (1.326 to 3.92)	nudix (nucleoside diphosphate linked moiety X)-type motif 4
AA152347	2.232 (1.644 to 2.765)	glutathione S-transferase A4
AA452566	2.23 (1.801 to 2.491)	peroxisomal membrane protein 3, 35kDa (Zellweger syndrome)
R98070	2.225 (1.301 to 3.276)	ESTs
T50041	2.218 (1.409 to 3.794)	DKFZP586A0522 protein
AA676945	2.206 (1.904 to 2.691)	NPC1 (Niemann-Pick disease, type C1, gene)-like 1
R51524	2.202 (1.348 to 3.173)	kelch-like 3 (Drosophila)
AA626156	2.201 (1.848 to 2.538)	chromosome 1 open reading frame 28
AI023541	2.201 (1.402 to 3.18)	carbonic anhydrase IX
H11987	2.189 (1.968 to 2.421)	ESTs
N69672	2.189 (1.413 to 4.312)	glycosylphosphatidylinositol specific phospholipase D1
H07920	2.188 (1.489 to 3.22)	mitogen-activated protein kinase kinase 6
N34287	2.183 (1.956 to 2.425)	unc-5 homolog B (C. elegans)
H25551	2.181 (1.215 to 3.788)	Homo sapiens clone 161455 breast expressed mRNA from chromosome X.
AA452125	2.18 (1.672 to 2.902)	LIM and cysteine-rich domains 1
H62473	2.178 (1.624 to 3.249)	transforming growth factor, beta receptor III (betaglycan, 300kDa)
AA504777	2.177 (2.046 to 2.348)	ESTs
R51494	2.175 (1.76 to 2.534)	copine VIII
H12277 | H12278	2.175 (1.28 to 4.744)	Data not found
R41600	2.17 (1.597 to 2.574)	hypothetical protein FLJ12688
AA448653	2.17 (1.39 to 3.037)	hypothetical protein LOC155435
AA620415	2.167 (1.482 to 2.964)	vacuolar protein sorting 35 (yeast)
AA291486	2.167 (1.435 to 3.985)	PTK2 protein tyrosine kinase 2
AA598548 | AI732161	2.164 (1.391 to 4.172)	Data not found
AA457137	2.161 (1.571 to 3.026)	hypothetical protein MGC2555
AA504618	2.156 (1.791 to 2.691)	tubulin-specific chaperone e
H05445	2.154 (1.233 to 3.273)	growth associated protein 43
AA418907	2.152 (1.757 to 2.524)	cytochrome P450, family 1, subfamily A, polypeptide 1
AA406320	2.152 (1.752 to 2.887)	poliovirus receptor-related 3
T74688	2.15 (1.894 to 2.402)	sulfide quinone reductase-like (yeast)
AA705069	2.146 (1.569 to 2.783)	retinoic acid receptor, alpha
AA489069	2.146 (1.55 to 2.647)	ESTs
R50775	2.133 (1.824 to 2.809)	likely ortholog of mouse variant polyadenylation protein CSTF-64
AA663981	2.13 (1.518 to 3.783)	immunoglobulin heavy constant gamma 3 (G3m marker)
AA704255	2.129 (1.335 to 2.708)	spinocerebellar ataxia 7 (olivopontocerebellar atrophy with retinal degeneration)
AA428139	2.124 (1.895 to 2.339)	intestinal cell kinase
AI347748	2.124 (1.744 to 2.372)	DEAD/H (Asp-Glu-Ala-Asp/His) box polypeptide 19 (DBP5 homolog, yeast)
N56995	2.123 (1.194 to 3.678)	KIAA1706 protein
AA878391	2.121 (1.65 to 3.467)	glypican 5
N47484	2.116 (1.125 to 3.185)	glutathione S-transferase A4
N76084	2.111 (1.792 to 2.49)	Homo sapiens cDNA FLJ12807 fis, clone NT2RP2002316.
AA706929	2.083 (1.395 to 2.709)	protein phosphatase 1F (PP2C domain containing)
H26271	2.083 (1.056 to 4.079)	pleckstrin homology-like domain, family A, member 1
AA446251	2.077 (1.392 to 3.085)	laminin, beta 1
R60053	2.069 (1.68 to 2.44)	hypothetical protein LOC155036
AA885311	2.069 (1.294 to 3.066)	butyrylcholinesterase
AA457671	2.068 (1.343 to 3.09)	procollagen-proline, 2-oxoglutarate 4-dioxygenase, alpha polypeptide I
AA454082	2.067 (1.447 to 2.95)	Data not found
W84751	2.062 (1.548 to 3.26)	KIAA1357 protein
R15832	2.059 (1.521 to 2.524)	p21/Cdc42/Rac1-activated kinase 1 (STE20 homolog, yeast)
H17022	2.057 (1.255 to 3.252)	myosin VIIA and Rab interacting protein
AA521067	2.053 (1.681 to 2.616)	inositol polyphosphate-5-phosphatase, 145kDa
R39765	2.043 (1.791 to 2.392)	DC6 protein
AA425128	2.04 (1.63 to 2.469)	ESTs
AA459956	2.039 (1.611 to 2.388)	putative ribonuclease III
R64146	2.037 (1.703 to 2.31)	Homo sapiens, clone IMAGE:5552108, mRNA
R73909	2.034 (1.83 to 2.178)	pregnancy specific beta-1-glycoprotein 2
R63647	2.033 (1.378 to 4.442)	prolactin receptor

3

Appendix A

N53436	2.032 (1.758 to 2.306)	homolog of rat orphan transporter v7-3
AA040742	2.03 (1.451 to 2.776)	poly(A) binding protein, nuclear 1
AA418984	2.028 (1.687 to 2.499)	sentrin/SUMO-specific protease SENP8
AA609149	2.028 (1.463 to 2.809)	Homo sapiens, Similar to hypothetical protein FLJ10058, clone MGC:34306
H04230	2.028 (1.32 to 2.682)	fukutin-related protein
AA449773	2.025 (1.25 to 2.899)	CDC42 effector protein (Rho GTPase binding) 4
T95670	2.015 (1.011 to 2.765)	Werner helicase interacting protein
R06840	2.013 (1.497 to 3.075)	gamma-glutamyltransferase-like 3
R55786	2.005 (1.579 to 2.479)	A kinase (PRKA) anchor protein 6
AA608575	17.343 (4.64 to 293.5)	propionyl Coenzyme A carboxylase, alpha polypeptide
AA135152	11.576 (7.197 to 20.898)	glutathione peroxidase 2 (gastrointestinal)
R40057	10.929 (3.612 to 32.266)	prominin-like 1 (mouse)
N66008	10.31 (5.53 to 16.207)	Putative prostate cancer tumor suppressor
AA007597	1.997 (1.102 to 3.397)	Homo sapiens, clone IMAGE:4347271, mRNA
N58065	1.994 (1.614 to 3.138)	ESTs, Weakly similar to RIKEN cDNA 1110001O19 [Mus musculus] [M.musculus]
AA479155	1.994 (1.286 to 2.374)	Homo sapiens, clone IMAGE:4826905, mRNA, partial cds
AA453750	1.99 (1.347 to 2.35)	hypothetical protein MGC14480
R43604	1.989 (0.995 to 3.909)	Homo sapiens mRNA; cDNA DKFZp547E184 (from clone DKFZp547E184)
AA709271	1.98 (1.844 to 2.422)	neural cell adhesion molecule 2
AA075722	1.957 (1.225 to 2.556)	nuclear transport factor 2
T57241	1.954 (1.701 to 2.293)	high mobility group nucleosomal binding domain 3
AA452138	1.939 (1.362 to 2.964)	leucine-rich repeat protein, neuronal 3
AA434090	1.938 (1.215 to 3.653)	fibroblast growth factor receptor-like 1
AA779520	1.934 (1.543 to 2.359)	ESTs
AA670390	1.928 (1.751 to 2.373)	ESTs
R01211	1.915 (1.286 to 2.574)	microfibrillar-associated protein 1
AA775831	1.903 (1.387 to 2.381)	ESTs
AA284528	1.901 (1.204 to 3.154)	protease, serine, 2 (trypsin 2)
AI033662	1.901 (1.093 to 2.759)	Homo sapiens mRNA; cDNA DKFZp586O0724 (from clone DKFZp586O0724)
AA625673	1.896 (1.632 to 2.157)	amyotrophic lateral sclerosis 2 (juvenile) chromosome region, candidate 12
AA417921	1.891 (1.221 to 3.19)	C-type (calcium dependent, carbohydrate-recognition domain) lectin, superfamily member 2
R24530	1.891 (1.17 to 2.533)	folate receptor 1 (adult)
AA705229	1.884 (1.172 to 2.616)	hypothetical protein FLJ12270
R70685	1.883 (1.232 to 2.345)	jagged 1 (Alagille syndrome)
R52703	1.878 (1.525 to 2.359)	thymidine kinase 2, mitochondrial
AA455043	1.877 (1.553 to 2.114)	holocarboxylase synthetase
R40855	1.874 (1.377 to 2.407)	ESTs
W15339	1.874 (1.011 to 3.264)	hypothetical protein LOC161291
T74023	1.869 (0.91 to 3)	tumor endothelial marker 6
AA193381	1.866 (0.911 to 2.904)	Homo sapiens cDNA FLJ37217 fis, clone BRALZ2008763.
R33363	1.865 (1.463 to 2.867)	decidual protein induced by progesterone
AA610000	1.864 (1.567 to 2.229)	ESTs
AA628230	1.86 (1.333 to 2.335)	Data not found
AA399216	1.86 (1.125 to 3.605)	ESTs
AA456642	1.859 (1.283 to 2.629)	small proline-rich protein 2A
AA449118	1.858 (1.13 to 2.53)	transcription factor A, mitochondrial
H73731	1.857 (1.369 to 2.389)	KIAA0601 protein
AI004245	1.855 (1.513 to 2.32)	hypothetical protein FLJ20449
AA478476	1.854 (1.375 to 3.742)	ESTs
N66028	1.849 (1.323 to 2.176)	FKBP-associated protein
AA863471	1.844 (1.232 to 2.38)	KIAA1107 protein
R60014	1.827 (1.429 to 2.916)	hypothetical protein DKFZp761N09121
AA701655	1.826 (1 to 3.09)	Data not found
H05769	1.821 (1.096 to 3.451)	hypothetical protein MGC14254
AA890663	1.82 (1.298 to 2.905)	p21/Cdc42/Rac1-activated kinase 1 (STE20 homolog, yeast)
AA777123	1.812 (1.002 to 2.247)	Homo sapiens, Similar to RIKEN cDNA 1110001O19 gene, clone MGC:21689
AA464196	1.807 (1.256 to 2.538)	hyaluronoglucosaminidase 1
AA873089	1.803 (1.192 to 2.786)	cytochrome P450, family 3, subfamily A, polypeptide 5
H18949	1.79 (1.156 to 3.339)	chromosome 6 open reading frame 33

4

Appendix A

R49650	1.773 (1.349 to 2.756)	Fas apoptotic inhibitory molecule 2
AA486738	1.77 (1.276 to 2.636)	Bardet-Biedl syndrome 2
R20755	1.748 (1.133 to 2.36)	Data not found
R51210	1.708 (1.17 to 2.239)	hypothetical protein FLJ22353

Genes with significant 2 fold or more decreased expression in SW620 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
AA452095	0.548 (0.401 to 0.862)	SMC4 structural maintenance of chromosomes 4-like 1 (yeast)
AA190339	0.547 (0.408 to 0.693)	dual specificity phosphatase 3 (vaccinia virus phosphatase VH1-related)
AA488177	0.543 (0.462 to 0.67)	dynactin 4 (p62)
T46878	0.54 (0.377 to 0.683)	eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa
AA443830	0.538 (0.388 to 0.873)	chromosome 14 open reading frame 9
AA446021	0.537 (0.375 to 0.74)	Homo sapiens cDNA FLJ11801 fis, clone HEMBA1006253,
AA454015	0.535 (0.41 to 0.644)	HSPC126 protein
N95752	0.529 (0.385 to 0.657)	Homo sapiens, Similar to RIKEN cDNA 1500009M05 gene, clone MGC:40370
AA431967	0.524 (0.375 to 0.703)	LATS, large tumor suppressor, homolog 2 (Drosophila)
R51362	0.518 (0.364 to 0.825)	mitochondrial ribosomal protein S31
AA485151	0.51 (0.362 to 0.756)	heat shock 105kDa/110kDa protein 1
T64933	0.508 (0.384 to 0.8)	KIAA1724 protein
AA437140	0.507 (0.393 to 0.694)	hypothetical protein DKFZp434L0718
AA045083	0.506 (0.371 to 0.735)	Homo sapiens, clone IMAGE:4402981, mRNA
AA609463	0.505 (0.428 to 0.594)	hypothetical protein MGC19556
AA456968	0.505 (0.388 to 0.974)	neurotrophic tyrosine kinase, receptor, type 2
R42699	0.505 (0.366 to 0.764)	KIAA1387 protein
T55337	0.505 (0.343 to 0.617)	Shwachman-Bodian-Diamond syndrome protein
AA398365	0.505 (0.329 to 0.935)	Homo sapiens cDNA FLJ35589 fis, clone SPLEN2007498.
N74362	0.504 (0.344 to 0.623)	serologically defined colon cancer antigen 3
AI333397	0.502 (0.354 to 0.703)	Data not found
T67278	0.5 (0.43 to 0.573)	hypothetical protein MGC46235
AA971738	0.499 (0.386 to 0.636)	ESTs, Weakly similar to RTN2_HUMAN Reticulon protein 2
AA496796	0.498 (0.343 to 0.702)	FERM, RhoGEF (ARHGEF) and pleckstrin domain protein 1 (chondrocyte-derived)
AA134696	0.498 (0.32 to 0.726)	Homo sapiens mRNA; cDNA DKFZp586E1120 (from clone DKFZp586E1120)
AA432085	0.497 (0.438 to 0.635)	tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein
AA457731	0.497 (0.435 to 0.67)	SNARE protein Ykt6
AA011211	0.496 (0.452 to 0.609)	hypothetical protein FLJ10726
T70031	0.496 (0.35 to 0.66)	solute carrier family 1 (neutral amino acid transporter), member 5
AA278402	0.495 (0.451 to 0.556)	microtubule-actin crosslinking factor 1
AA115400	0.494 (0.365 to 0.694)	hHDC for homolog of Drosophila headcase
H96654	0.494 (0.336 to 0.893)	pp21 homolog
H96908	0.493 (0.4 to 0.752)	hypothetical protein MGC10067
N73536	0.493 (0.341 to 0.713)	hypothetical protein BM-002
T57834	0.49 (0.423 to 0.662)	chromosome 20 open reading frame 108
N49581	0.488 (0.384 to 0.662)	similar to CG6405 gene product
AA099034	0.487 (0.349 to 0.808)	ubiquitin specific protease 1
AI289851	0.483 (0.393 to 0.668)	Data not found
AA676749	0.483 (0.361 to 0.561)	dual-specificity tyrosine-(Y)-phosphorylation regulated kinase 1A
AA608632	0.482 (0.363 to 0.686)	chromosome 14 open reading frame 35
AA598625	0.482 (0.314 to 0.607)	solute carrier family 1 (glutamate/neutral amino acid transporter), member 4
T73780	0.481 (0.314 to 0.815)	ESTs, Similar to hypothetical protein FLJ21617; erythroid differentiation-related factor 1
AA156946	0.481 (0.286 to 0.754)	Data not found
AA126799	0.48 (0.395 to 0.582)	uncharacterized hematopoietic stem/progenitor cells protein MDS031
H07934	0.478 (0.352 to 0.599)	DKFZP564I1171 protein
N72265	0.476 (0.37 to 0.588)	casein kinase 1, alpha 1
N51837	0.475 (0.369 to 0.65)	DKFZP434A043 protein
H48097	0.475 (0.347 to 0.817)	hypothetical protein FLJ20986
R94947	0.475 (0.321 to 0.871)	Rho-associated, coiled-coil containing protein kinase 2
AA496884	0.474 (0.42 to 0.548)	hypothetical protein FLJ14564
AA195449	0.473 (0.375 to 0.534)	hypothetical protein LOC130074
AI675707	0.472 (0.256 to 0.622)	ribosomal protein S7
AA450334	0.471 (0.311 to 0.882)	Homo sapiens cDNA FLJ13209 fis, clone NT2RP4000424.
AA417825	0.469 (0.315 to 0.577)	hypothetical protein FLJ20643
N64494	0.468 (0.429 to 0.511)	Homo sapiens cDNA FLJ37850 fis, clone BRSSN2013733,

Appendix A

N74189	0.468 (0.37 to 0.685)	WD domain, G-beta repeat-containing protein
AA446316	0.467 (0.383 to 0.534)	ESTs
AA598504	0.467 (0.324 to 0.732)	beta-1,3-glucuronyltransferase 1 (glucuronosyltransferase P)
R26732	0.466 (0.318 to 0.594)	peripheral myelin protein 22
AI688757	0.464 (0.395 to 0.51)	cytochrome c oxidase subunit Vb
AA293653	0.462 (0.331 to 0.608)	phosphoprotein enriched in astrocytes 15
AA420992	0.46 (0.305 to 0.82)	ESTs
AA630374	0.459 (0.36 to 0.634)	dual specificity phosphatase 6
W76339	0.459 (0.354 to 0.637)	nuclear factor (erythroid-derived 2)-like 3
R23738	0.459 (0.277 to 0.649)	Homo sapiens cDNA FLJ30233 fis, clone BRACE2001971.
N32611	0.458 (0.408 to 0.571)	hypothetical protein FLJ20481
AA970066	0.458 (0.288 to 0.726)	ubiquitin specific protease 1
AA136133	0.455 (0.363 to 0.558)	hypothetical protein FLJ14735
AA464694	0.454 (0.344 to 0.684)	Human HeLa mRNA isolated as a false positive in a two-hybrid-screen.
AA629542	0.454 (0.3 to 0.683)	WAS protein family, member 3
W49633	0.453 (0.352 to 0.571)	mitofusin 1
N62924	0.452 (0.296 to 0.6)	mitochondrial ribosomal protein S22
AA456437	0.452 (0.274 to 0.661)	MKI67 (FHA domain) interacting nucleolar phosphoprotein
N62766	0.449 (0.359 to 0.583)	neutral sphingomyelinase (N-SMase) activation associated factor
AA489626	0.448 (0.327 to 0.619)	FYVE and coiled-coil domain containing 1
AA630449	0.448 (0.325 to 0.873)	Niemann-Pick disease, type C2
AI340883	0.447 (0.352 to 0.613)	tumor-associated calcium signal transducer 1
AA469953	0.446 (0.345 to 0.566)	nucleoporin 50kDa
AA973097	0.446 (0.307 to 0.77)	EST
AA827551	0.445 (0.327 to 0.899)	Notch homolog 2 (Drosophila)
AA497040	0.445 (0.277 to 0.87)	stanniocalcin 2
AA459690	0.445 (0.236 to 0.724)	brain protein 44-like
N48319	0.444 (0.308 to 0.702)	breast cancer anti-estrogen resistance 3
AI343879	0.442 (0.389 to 0.538)	hypothetical protein FLJ11838
AA460731	0.442 (0.363 to 0.72)	G protein-coupled receptor 107
AA293671	0.441 (0.297 to 0.584)	CD8 antigen, beta polypeptide 1 (p37)
R42174	0.441 (0.286 to 0.678)	sin3-associated polypeptide, 18kDa
AI263603	0.44 (0.398 to 0.478)	EST
AA029312	0.439 (0.368 to 0.501)	NIMA (never in mitosis gene a)- related kinase 9
H09322	0.438 (0.362 to 0.592)	ESTs
AA447992	0.437 (0.379 to 0.583)	ESTs
AA011096	0.437 (0.305 to 0.56)	monoamine oxidase A
AA078976	0.435 (0.316 to 0.558)	thioredoxin-like, 32kDa
W90381	0.434 (0.363 to 0.519)	hypothetical protein LOC145622
AA487486	0.433 (0.382 to 0.497)	cyclin D1 (PRAD1: parathyroid adenomatosis 1)
AA485349	0.433 (0.38 to 0.528)	N-ethylmaleimide-sensitive factor attachment protein, gamma
N36176	0.433 (0.298 to 0.521)	chromosome 1 open reading frame 9
AA004525	0.433 (0.267 to 0.601)	hypothetical protein F23149_1
N80741	0.432 (0.235 to 0.741)	ATP binding protein associated with cell differentiation
AA598573	0.429 (0.383 to 0.496)	unc-84 homolog A (C. elegans)
H00589	0.427 (0.298 to 0.571)	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 1
H15085	0.426 (0.304 to 0.584)	ADP-ribosylation factor 4-like
AI199321	0.426 (0.247 to 0.613)	eukaryotic translation initiation factor 2B, subunit 4 delta, 67kDa
W72005	0.425 (0.322 to 0.557)	adenomatosis polyposis coli down-regulated 1
AA443722	0.425 (0.321 to 0.52)	upstream binding protein 1 (LBP-1a)
AA460716	0.425 (0.318 to 0.483)	Data not found
R56870	0.425 (0.265 to 0.55)	similar to RIKEN cDNA 2700047N05
AI056063	0.424 (0.305 to 0.545)	ESTs
AA455235	0.423 (0.361 to 0.502)	aldehyde dehydrogenase 1 family, member A3
AA488610	0.421 (0.393 to 0.439)	MCM4 minichromosome maintenance deficient 4 (S. cerevisiae)
AA017526	0.421 (0.353 to 0.453)	collagen, type IX, alpha 3
H10335	0.416 (0.292 to 0.636)	Homo sapiens mRNA; cDNA DKFZp564L222 (from clone DKFZp564L222)
AA486445	0.415 (0.348 to 0.504)	KIAA0563 gene product
N34466	0.415 (0.278 to 0.606)	hypothetical protein DKFZp434H0820
N57005	0.412 (0.345 to 0.5)	ring finger protein (C3H2C3 type) 6

7

W46944	0.411 (0.283 to 0.526)	angiomotin like 2
H20809	0.41 (0.341 to 0.485)	Homo sapiens cDNA FLJ33540 fis, clone BRAMY2007613.
T97762	0.41 (0.34 to 0.6)	interferon-related developmental regulator 1
AA777187	0.41 (0.298 to 0.522)	cysteine-rich, angiogenic inducer, 61
AA398356	0.41 (0.28 to 0.564)	chromosome 11 open reading frame 14
N67766	0.406 (0.349 to 0.491)	acetyl-Coenzyme A synthetase 2 (AMP forming)-like
H15549	0.405 (0.265 to 0.599)	hypothetical protein FLJ32122
AI080633	0.405 (0.253 to 0.689)	pyruvate dehydrogenase phosphatase
AA458502	0.404 (0.334 to 0.503)	peptidyl-prolyl isomerase G (cyclophilin G)
AA608534	0.404 (0.288 to 0.523)	Homo sapiens mRNA; cDNA DKFZp586C1019 (from clone DKFZp586C1019)
AA683077	0.4 (0.378 to 0.419)	mitogen-activated protein kinase 1
AA283090	0.398 (0.324 to 0.456)	Homo sapiens CD44 isoform RC (CD44) mRNA, complete cds
N51225	0.396 (0.287 to 0.597)	Homo sapiens, clone IMAGE:4798592, mRNA
AA188789	0.395 (0.275 to 0.539)	Homo sapiens, clone IMAGE:6057297, mRNA
AI215610	0.394 (0.337 to 0.506)	ESTs
N93403	0.394 (0.333 to 0.523)	hypothetical protein FLJ20707
R12808	0.39 (0.31 to 0.525)	ESTs, Moderately similar to hypothetical protein, MGC:7199 [Mus musculus] [M.musculus
W65461	0.39 (0.229 to 0.553)	dual specificity phosphatase 5
AI095082	0.389 (0.249 to 0.605)	Rag D protein
N29844	0.386 (0.32 to 0.655)	peptidase (mitochondrial processing) beta
AI439571	0.384 (0.286 to 0.553)	apoptosis antagonizing transcription factor
AI146478	0.384 (0.273 to 0.482)	hypothetical protein DKFZp761H0421
AA024832	0.384 (0.245 to 0.587)	Homo sapiens mRNA; cDNA DKFZp586I0324 (from clone DKFZp586I0324)
AA486556	0.383 (0.374 to 0.392)	CD81 antigen (target of antiproliferative antibody 1)
N32768	0.383 (0.271 to 0.656)	pregnancy specific beta-1-glycoprotein 3
N52496	0.382 (0.33 to 0.452)	BTG family, member 3
AA479058	0.382 (0.318 to 0.488)	thrombopoietin
AA460286	0.381 (0.207 to 0.563)	guanine nucleotide binding protein (G protein), gamma 10
N46446	0.38 (0.229 to 0.564)	Data not found
R55046	0.376 (0.303 to 0.505)	MpV17 transgene, murine homolog, glomerulosclerosis
AI248758	0.376 (0.275 to 0.627)	EST
R59697	0.373 (0.292 to 0.509)	Homo sapiens mRNA fragment.
AA464544	0.368 (0.309 to 0.462)	LIM domain only 2 (rhombotin-like 1)
R78530 | R78490	0.368 (0.22 to 0.636)	Data not found
R67355	0.367 (0.201 to 0.715)	angiomotin like 2
AA927761	0.365 (0.302 to 0.471)	hypothetical protein LOC128977
R69622	0.362 (0.272 to 0.475)	CGI-141 protein
N72878	0.361 (0.219 to 0.631)	development and differentiation enhancing factor 1
AA459401	0.358 (0.246 to 0.542)	kallikrein 10
AA972654	0.353 (0.266 to 0.504)	hypothetical protein MGC48972
N25987	0.352 (0.339 to 0.384)	disrupted in renal carcinoma 2
AA723035	0.352 (0.265 to 0.582)	zinc finger protein 36, C3H type-like 1
AA490044	0.351 (0.286 to 0.504)	solute carrier family 5 (inositol transporters), member 3
AA291163	0.351 (0.251 to 0.521)	glutaredoxin (thioltransferase)
AA479276	0.347 (0.331 to 0.375)	ESTs
R66605	0.347 (0.287 to 0.412)	hypothetical protein MGC20576
AA621315	0.344 (0.305 to 0.397)	catenin (cadherin-associated protein), alpha-like 1
R25377	0.344 (0.29 to 0.412)	DEK oncogene (DNA binding)
AA929006	0.343 (0.28 to 0.431)	ESTs
AI347472	0.338 (0.235 to 0.573)	ATP binding protein associated with cell differentiation
N48355	0.338 (0.221 to 0.485)	fragile X mental retardation 1
N50079	0.336 (0.215 to 0.453)	Opa-interacting protein 2
AI088175	0.332 (0.255 to 0.557)	hypothetical protein FLJ10120
AA487425	0.331 (0.225 to 0.44)	zinc finger protein 463
AA135001	0.329 (0.246 to 0.437)	hypothetical protein BC016153
AI383171	0.328 (0.191 to 0.799)	Data not found
AA160507	0.327 (0.235 to 0.677)	keratin 5 (epidermolysis bullosa simplex, Dowling-Meara/Kobner/Weber-Cockayne types)
N66178	0.326 (0.216 to 0.603)	Homo sapiens cDNA FLJ23705 fis, clone HEP11066.
AI375330	0.325 (0.242 to 0.5)	Homo sapiens cDNA: FLJ21243 fis, clone COL01164.
N64734	0.324 (0.274 to 0.391)	Data not found

Appendix A

N25352	0.324 (0.228 to 0.42)	coxsackie virus and adenovirus receptor
AA159669	0.323 (0.265 to 0.359)	likely ortholog of rat vacuole membrane protein 1
AA496438	0.322 (0.155 to 0.499)	retinoic acid receptor, gamma
AA165678	0.319 (0.275 to 0.448)	ATP-binding cassette, sub-family C (CFTR/MRP), member 4
AA134595	0.317 (0.215 to 0.453)	eukaryotic translation initiation factor 5A2
R43270	0.308 (0.281 to 0.346)	protein phosphatase 1, regulatory (inhibitor) subunit 14C
AI264351	0.308 (0.233 to 0.414)	Homo sapiens cDNA FLJ40926 fis, clone UTERU2006524.
H29268	0.302 (0.191 to 0.542)	Homo sapiens, Similar to nuclear localization signals binding protein 1,
AA190626	0.298 (0.24 to 0.333)	sentrin-specific protease
W37447	0.296 (0.199 to 0.501)	ESTs
W69211	0.294 (0.211 to 0.364)	chemokine (C-C motif) ligand 11
T62844	0.294 (0.174 to 0.545)	hypothetical protein dJ473B4
AI334917	0.29 (0.148 to 0.91)	LIM domain protein
AA187933	0.286 (0.224 to 0.376)	transcriptional co-activator with PDZ-binding motif (TAZ)
AI096800	0.285 (0.251 to 0.33)	immediate early response 5
AA010818	0.274 (0.22 to 0.315)	myeloid/lymphoid or mixed-lineage leukemia
AA504265	0.273 (0.23 to 0.307)	LIM and senescent cell antigen-like domains 1
N34799	0.271 (0.231 to 0.315)	hypothetical protein FLJ23306
N54794	0.27 (0.241 to 0.304)	Data not found
AI271884	0.258 (0.227 to 0.31)	Wiskott-Aldrich syndrome-like
AA431753	0.252 (0.153 to 0.387)	CED-6 protein
AA485677	0.252 (0.152 to 0.496)	thyroid hormone receptor interactor 6
AA156988	0.249 (0.151 to 0.325)	aconitase 1, soluble
N70078	0.248 (0.201 to 0.362)	zizimin1
R32440	0.244 (0.152 to 0.357)	Homo sapiens clone 23698 mRNA sequence
N63598	0.242 (0.164 to 0.317)	Homo sapiens cDNA FLJ32453 fis, clone SKMUS2001703.
R96522	0.237 (0.137 to 0.329)	pregnancy specific beta-1-glycoprotein 1
N69049	0.235 (0.158 to 0.295)	frizzled homolog 7 (Drosophila)
AA457501	0.219 (0.12 to 0.288)	duodenal cytochrome b
AA418036	0.211 (0.103 to 0.347)	GLI-Kruppel family member GLI3 (Greig cephalopolysyndactyly syndrome)
AA406552	0.165 (0.118 to 0.204)	solute carrier family 2 (facilitated glucose transporter), member 3
AA609556	0.162 (0.101 to 0.292)	pinch-2
AA598601	0.148 (0.107 to 0.286)	insulin-like growth factor binding protein 3

Genes with significant 2 fold or more increased expression in SW480M7 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
AI350438	9.704 (0.745 to 76.349)	KIAA1966 protein
AA620619	3.29 (2.819 to 4.258)	ESTs
AI192052	3.055 (2.643 to 3.302)	chromosome 14 open reading frame 28
H62199	2.985 (1.43 to 6.286)	ESTs
AA912874	2.813 (1.568 to 4.216)	Homo sapiens, clone IMAGE:5266192, mRNA
AI339434	2.75 (1.972 to 4.564)	caveolin 2
AA418675	2.728 (1.508 to 7.132)	ADP-ribosylarginine hydrolase
AA424568	2.721 (1.238 to 4.992)	ESTs
AI250799	2.676 (2.047 to 3.591)	defensin, alpha 4, corticostatin
AI125237	2.664 (2.136 to 3.003)	ESTs
AA933773	2.628 (1.67 to 4.887)	chromobox homolog 3 (HP1 gamma homolog, Drosophila)
AI276487	2.617 (2.046 to 3.711)	cytokeratin type II
AA609884	2.579 (1.386 to 5.051)	Data not found
AI188068	2.565 (1.208 to 4.913)	parkin co-regulated gene protein
AI125898	2.554 (1.622 to 3.723)	ESTs
AA464583	2.552 (1.643 to 3.963)	ESTs
AA437243	2.549 (1.325 to 4.2)	ESTs
R99080	2.541 (1.819 to 4.839)	Data not found
AA669218	2.54 (2.072 to 3.354)	HBS1-like (S. cerevisiae)
AI028163	2.534 (1.689 to 3.592)	ESTs
AA620642	2.518 (2.31 to 2.942)	hypothetical protein LOC51319
AI140695	2.509 (1.827 to 3.304)	Data not found
AA916562	2.494 (2.212 to 2.691)	vacuolar protein sorting 4B (yeast)
N/A-2-1-15-4	2.493 (2.224 to 2.961)	Data not found
H21892	2.492 (1.624 to 5.384)	Homo sapiens cDNA FLJ36638 fis, clone TRACH2018950.
AA779228	2.476 (2.031 to 3.043)	ESTs
AI140566	2.433 (1.888 to 2.968)	ESTs,
AI288845	2.421 (1.818 to 3.089)	chemokine (C-C motif) receptor-like 2
AA634483	2.408 (1.883 to 2.976)	Homo sapiens, clone IMAGE:5266541, mRNA
AI204136	2.405 (1.667 to 3.392)	bromodomain containing 2
AI085068	2.404 (1.802 to 2.929)	ESTs
AA479344	2.401 (1.679 to 3.437)	gap junction protein, beta 5 (connexin 31.1)
AI187832	2.398 (1.629 to 3.185)	ESTs
AA868748	2.393 (1.966 to 2.676)	ESTs
AA463462	2.384 (1.852 to 3.85)	Toll-interacting protein
AA676548	2.376 (1.906 to 3.037)	ESTs, Moderately similar to hypothetical protein FLJ20234
AA778851	2.347 (1.881 to 2.748)	Human clone 137308 mRNA, partial cds.
R38885	2.345 (2.094 to 2.743)	ESTs
AI191761	2.339 (1.823 to 2.875)	serine protease inhibitor, Kazal type 4
AA953249	2.323 (1.855 to 3.2)	H factor 1 (complement)
AI022259	2.316 (1.462 to 4.611)	zinc finger protein 255
AA977836	2.31 (1.848 to 2.791)	peroxisomal farnesylated protein
AA620670	2.307 (1.392 to 4.403)	ESTs
AI025808	2.274 (1.735 to 3.098)	Data not found
AA100957	2.266 (1.086 to 3.754)	DKFZP434C171 protein
AA609630	2.26 (1.921 to 2.529)	Data not found
AA777779	2.255 (2.125 to 2.373)	ESTs
N/A-3-1-14-6	2.253 (1.869 to 2.744)	Data not found
AA669367	AI732124	2.251 (1.903 to 2.916)	Data not found
AA921818	2.228 (1.946 to 2.421)	ESTs
AI125351	2.228 (1.498 to 3.926)	ESTs
AA454949	2.227 (2.177 to 2.312)	hypothetical protein FLJ23468
AA410434	2.218 (1.952 to 2.752)	collagen triple helix repeat containing 1
AI188006	2.218 (1.775 to 3.075)	ESTs
AI125753	2.217 (1.74 to 2.709)	Homo sapiens, clone IMAGE:4822830, mRNA

Appendix A

AA621004	2.217 (1.374 to 3.394)	ESTs
AA707094	2.211 (1.617 to 4.113)	ESTs
AI187831	2.208 (1.723 to 3.587)	ESTs
AA504631	2.207 (1.737 to 2.872)	CDA14
AA868802	2.204 (1.634 to 3.536)	ESTs
AA677215	2.195 (1.447 to 3.83)	ESTs, Moderately similar to hypothetical protein FLJ20234
AA953254	2.181 (1.629 to 2.658)	collagen, type IV, alpha 5 (Alport syndrome)
AA778617	2.179 (1.81 to 2.44)	ESTs
AA694497	2.174 (1.856 to 2.569)	Data not found
AI032679	2.17 (1.587 to 2.749)	ESTs
AI139912	2.165 (1.708 to 2.781)	Data not found
AA418896	2.165 (1.489 to 4.332)	ESTs, Weakly similar to hypothetical protein FLJ20489 [Homo sapiens] [H.sapiens]
AA706094	2.163 (1.793 to 2.482)	ESTs, Weakly similar to 2109260A B cell growth factor [Homo sapiens] [H.sapiens]
AI138368	2.152 (1.467 to 3.578)	cytochrome c oxidase subunit VIIb2
AI139861	2.144 (1.457 to 3.314)	Data not found
AA437210	2.137 (1.559 to 2.546)	ESTs
AA609768	2.136 (1.99 to 2.36)	ESTs
AI198232	2.132 (1.945 to 2.41)	C-terminal PDZ domain ligand of neuronal nitric oxide synthase
AA626942	2.131 (1.647 to 2.883)	ESTs
AA778577	2.13 (1.603 to 2.982)	EST,
N63479	2.127 (1.473 to 3)	Data not found
AA682719	2.125 (1.777 to 2.664)	EST
AA921805	2.124 (1.316 to 3.28)	ESTs
AI278813	2.123 (1.424 to 3.049)	ESTs
H83116	2.117 (1.574 to 3.141)	origin recognition complex, subunit 6 homolog-like (yeast)
AI028784	2.107 (1.602 to 2.755)	Homo sapiens cDNA FLJ40501 fis, clone TESTI2045171.
AA465242	2.097 (1.963 to 2.345)	hypothetical protein MGC45400
AA629039	2.095 (1.213 to 3.349)	RNA binding motif protein 5
AI081823	2.093 (1.908 to 2.376)	chromosome 20 open reading frame 104
AI206318	2.089 (1.814 to 2.446)	gamma-aminobutyric acid (GABA) A receptor, beta 3
AI125829	2.086 (1.8 to 2.459)	EST
AA773190 | AI791127	2.086 (1.712 to 2.432)	Data not found
AA917956	2.085 (1.866 to 2.334)	Homo sapiens, clone IMAGE:3878708, mRNA
AI142095	2.084 (1.778 to 2.84)	ESTs
AA418728	2.083 (1.532 to 3.688)	palmdelphin
AA479512	2.074 (1.666 to 2.978)	ESTs,
AI082726	2.073 (1.639 to 2.45)	Data not found
AA909246	2.072 (1.221 to 2.986)	Homo sapiens, clone IMAGE:4824710, mRNA
AA778839	2.068 (1.624 to 2.66)	ESTs
AA633768	2.068 (1.454 to 4.049)	dipeptidylpeptidase 7
R16676	2.067 (1.663 to 2.569)	NFS1 nitrogen fixation 1 (S. cerevisiae)
AA232895	2.065 (1.3 to 3.606)	Data not found
AA889865	2.063 (1.75 to 2.712)	EST
H77461	2.063 (1.665 to 2.422)	lymphocyte adaptor protein
AI183541	2.063 (1.241 to 3.607)	Data not found
AI081607	2.061 (1.822 to 2.615)	ESTs
AA865729	2.059 (1.191 to 3.18)	ocular albinism 1 (Nettleship-Falls)
AA629924	2.056 (1.763 to 2.414)	Data not found
R07012	2.055 (1.592 to 2.702)	ESTs
AA425128	2.055 (1.191 to 3.588)	ESTs
AA625856	2.047 (1.641 to 3.028)	sorting nexin 5
AA626716	2.044 (1.77 to 2.374)	Data not found
AA868008	2.042 (1.716 to 2.822)	histone 1, H4c
R10885	2.04 (1.565 to 3.124)	aspartoacylase-3
N22685	2.039 (1.809 to 2.348)	plexin C1
AI018280	2.039 (1.742 to 2.746)	Homo sapiens, clone MGC:34040 IMAGE:4829245, mRNA, complete cds
AI140549	2.038 (1.734 to 2.7)	ESTs
AA677641	2.034 (1.841 to 2.306)	KIAA1910 protein
AI204635	2.034 (1.644 to 2.418)	Homo sapiens, clone IMAGE:5267179, mRNA
AA452255	2.034 (1.224 to 3.395)	integrin-binding sialoprotein (bone sialoprotein, bone sialoprotein II)

11

Appendix A

AA993474	2.031 (1.692 to 2.456)	Data not found
AA912886	2.029 (1.57 to 2.433)	ESTs
AA437245	2.028 (1.968 to 2.125)	Homo sapiens, clone IMAGE:5744200, mRNA
AI253652 | AI793206	2.028 (1.767 to 2.251)	Data not found
AA625858	2.025 (1.753 to 2.281)	ESTs
AI124607	2.021 (1.658 to 2.414)	Data not found
AI371323	2.019 (1.612 to 2.558)	leukocyte immunoglobulin-like receptor, subfamily B
AA424562	2.017 (1.111 to 2.734)	tumor necrosis factor (ligand) superfamily, member 13
AA505111	2.014 (1.327 to 2.823)	copine III
AA455935	2.012 (1.335 to 2.47)	Homo sapiens mRNA; cDNA DKFZp564H1916 (from clone DKFZp564H1916)
AA927182	2.009 (1.55 to 2.32)	similar to rat nuclear ubiquitous casein kinase 2
AA778611	2.006 (1.749 to 2.195)	KIAA1086 protein
AA463483	2.003 (1.354 to 3.717)	CD47 antigen (Rh-related antigen, integrin-associated signal transducer)
AA719374	2.002 (1.722 to 2.238)	EST
AA890153	1.994 (1.55 to 2.342)	ESTs
AA479950	1.984 (1.263 to 3.378)	ATPase, Class II, type 9B
AA927951	1.974 (1.126 to 3.346)	fragile X mental retardation, autosomal homolog 1
AA778609	1.973 (1.576 to 2.235)	testis development protein NYD-SP26
H45289	1.955 (1.336 to 2.429)	ESTs
AA936776	1.954 (1.527 to 2.877)	hypothetical protein FLJ38359
AI341099	1.953 (1.52 to 2.688)	Data not found
AA709162	1.952 (1.5 to 3.161)	Homo sapiens cDNA: FLJ22133 fis, clone HEP20529.
AA923743	1.934 (1.178 to 3.197)	ESTs
AA478606	1.932 (1.423 to 2.419)	x 006 protein
H67514	1.93 (1.529 to 2.946)	ESTs
AA398121	1.905 (1.496 to 2.452)	ESTs
AI190411	1.884 (1.243 to 2.593)	tubby like protein 3
AA911697	1.85 (1.488 to 2.733)	hypothetical protein FLJ31349
AI375330	1.846 (1.302 to 2.765)	Homo sapiens cDNA: FLJ21243 fis, clone COL01164.
AI131266	1.835 (1.34 to 2.397)	EST
AI000726	1.798 (1.194 to 2.707)	likely ortholog of mouse myocytic induction/differentiation originator
AA496984	1.773 (1.34 to 2.898)	angiomotin

Genes with significant 2 fold or more decreased expression in SW480M7 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
AI652005	0.515 (0.42 to 0.615)	cathepsin L2
R38494	0.493 (0.468 to 0.515)	flotillin 2
AA976089	0.488 (0.468 to 0.52)	enoyl Coenzyme A hydratase 1, peroxisomal
AA410636	0.487 (0.384 to 0.656)	Isoleucine-tRNA synthetase
AA995233	0.477 (0.43 to 0.506)	ESTs
AA983882	0.472 (0.356 to 0.668)	DKFZP586J1624 protein
AA976212	0.439 (0.315 to 0.595)	hypothetical protein BC013995
H59614	0.391 (0.286 to 0.62)	Homo sapiens, clone MGC:52263 IMAGE:4123447, mRNA, complete cds
N/A-4-3-15-2	0.368 (0.251 to 0.481)	Data not found
AA894927	0.358 (0.047 to 1.656)	asparagine synthetase
AI280306	0.341 (0.225 to 0.584)	UDP-glucose ceramide glucosyltransferase-like 2
AI015679	0.304 (0.275 to 0.324)	phosphoserine aminotransferase

Genes with significant 2 fold or more increased expression in SW480M9 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
H05775	7.081 (2.551 to 28.636)	monoamine oxidase A
AA394240	6.417 (4.968 to 8.801)	X-box binding protein 1
T72581	4.686 (2.546 to 7.611)	matrix metalloproteinase 9
N/A-2-1-15-4	4.46 (2.382 to 13.241)	Data not found
AA778916	4.134 (2.217 to 5.881)	ESTs
W80739	4.038 (1.84 to 13.195)	Homo sapiens cDNA FLJ25106 fis, clone CBR01467.
AA479344	3.822 (2.559 to 6.547)	gap junction protein, beta 5 (connexin 31.1)
AA598549	3.624 (2.02 to 10.901)	SMC2 structural maintenance of chromosomes 2-like 1 (yeast)
AA906569	3.482 (1.878 to 10.053)	Data not found
AA934753	AI733146	3.449 (1.207 to 8.188)	Data not found
AA775447	3.433 (1.27 to 8.108)	alpha-2-macroglobulin
AI253652	AI793206	3.426 (1.935 to 5.64)	Data not found
AA676765	3.389 (3.031 to 3.83)	mitochondrial ribosomal protein 63
AA496949	3.318 (2.061 to 8.036)	ESTs
AI288845	3.265 (1.584 to 8.376)	chemokine (C-C motif) receptor-like 2
AA453796	3.2 (1.98 to 6.645)	tetraspanin similiar to uroplakin 1
AA634483	3.163 (2.347 to 5.209)	Homo sapiens, clone IMAGE:5266541, mRNA
AI276487	3.045 (1.767 to 5.447)	cytokeratin type II
AI285947	3.034 (2.024 to 4.721)	outer dense fiber of sperm tails 1
AA700141	3.012 (1.644 to 4.481)	ESTs
AA777779	2.984 (1.804 to 7.067)	ESTs
AA707545	2.984 (1.433 to 7.845)	transition protein 1 (during histone to protamine replacement)
AA937734	2.866 (1.949 to 3.643)	Homo sapiens full length insert cDNA clone ZD16H11
AA620619	2.859 (2.287 to 3.893)	ESTs
AA933721	2.858 (1.432 to 5.858)	myotubularin related protein 2
AI125237	2.85 (2.088 to 4.203)	ESTs
AI250799	2.778 (1.507 to 4.783)	defensin, alpha 4, corticostatin
AA626316	2.771 (1.682 to 3.918)	likely ortholog of mouse kinesin light chain 2
AI140566	2.767 (1.248 to 5.303)	ESTs,
AI347622	2.76 (2.068 to 3.556)	tumor necrosis factor (ligand) superfamily, member 7
AI340031	2.74 (1.454 to 4.207)	ankyrin-like with transmembrane domains 1
AA777308	2.737 (2.14 to 3.727)	hypothetical protein FLJ13942
AA733038	2.729 (2.364 to 3.14)	RNA polymerase I subunit
AA418698	2.721 (1.74 to 5.203)	Homo sapiens, similar to SHC, clone MGC:34023 IMAGE:4827789
AA495818	2.699 (1.734 to 4.912)	dudulin 2
AA443193	2.698 (2.098 to 3.381)	homolog of yeast Sec5
AA912793	2.696 (1.741 to 5.399)	ESTs
AI357236	2.694 (2.044 to 3.962)	protamine 1
AA778851	2.692 (1.996 to 3.672)	Human clone 137308 mRNA, partial cds.
AA402915	2.678 (1.677 to 6.066)	aminoacylase 1
AI024634	2.667 (1.613 to 3.988)	ESTs
AI125868	2.661 (1.414 to 6.057)	Homo sapiens cDNA FLJ12032 fis, clone HEMBB1001880.
AA934618	2.66 (1.75 to 4.326)	Homo sapiens mRNA for tMDC II, isoform [a]
AI243367	2.652 (2.527 to 2.83)	ESTs
AA779199	2.647 (1.864 to 4.009)	ESTs
AI188006	2.628 (1.624 to 5.214)	ESTs
AA609630	2.619 (2.039 to 3.199)	Data not found
AI206318	2.614 (1.717 to 5.806)	gamma-aminobutyric acid (GABA) A receptor, beta 3
AI219893	2.601 (1.453 to 5.182)	ESTs

Appendix A

AA016001	2.589 (1.572 to 4.222)	transposon-derived Buster1 transposase-like protein
AI039422	2.577 (1.885 to 4.27)	hypothetical protein FLJ10770
AA907547	2.577 (1.713 to 3.297)	ESTs
AI140549	2.556 (1.329 to 6.064)	ESTs
AI023853	2.554 (1.558 to 4.084)	ESTs, Weakly similar to hypothetical protein FLJ11531
AA927721	2.548 (1.607 to 3.283)	ESTs
AI269119	2.547 (1.5 to 4.634)	peroxisome receptor 1
AA778590	2.522 (1.698 to 4.587)	ESTs
AA778611	2.521 (2.071 to 3.079)	KIAA1086 protein
N22685	2.517 (1.935 to 3.09)	plexin C1
T91254	2.51 (1.803 to 3.846)	ESTs, Weakly similar to neuronal thread protein
AA626716	2.509 (1.973 to 2.985)	Data not found
AA479950	2.497 (2.11 to 2.989)	ATPase, Class II, type 9B
AI140695	2.474 (2.078 to 3.144)	Data not found
AA669367 | AI732124	2.474 (1.743 to 4.704)	Data not found
AA777837	2.47 (1.882 to 4.193)	ESTs
H45289	2.458 (1.724 to 2.988)	ESTs
AA917956	2.455 (2.312 to 2.671)	Homo sapiens, clone IMAGE:3878708, mRNA
N/A-3-1-14-6	2.449 (2.29 to 2.575)	Data not found
AI002154	2.42 (1.61 to 4.38)	zinc finger protein 106
AI125425	2.419 (1.804 to 3.5)	Homo sapiens, clone IMAGE:5284350, mRNA
AA873143	2.409 (2.162 to 2.897)	pleckstrin homology domain containing, family B (evectins) member 1
AA704794	2.385 (1.883 to 2.753)	ESTs
AA705112	2.381 (1.862 to 3.783)	molybdenum cofactor synthesis 1
AA683578	2.377 (1.962 to 2.859)	Homo sapiens mRNA; cDNA DKFZp667G2419
AA884636	2.375 (1.592 to 3.218)	ESTs
AA455904	2.372 (1.489 to 5.143)	ESTs
AA707671	2.37 (1.643 to 4.191)	Homo sapiens cDNA FLJ31598 fis, clone NT2RI2002549.
AA778617	2.366 (1.929 to 2.81)	ESTs
AI198232	2.365 (1.75 to 2.781)	C-terminal PDZ domain ligand of neuronal nitric oxide synthase
AI091540	2.361 (1.858 to 3.41)	KIAA1376 protein
AA912886	2.357 (1.745 to 3.057)	ESTs
AA894577	2.35 (1.762 to 4.007)	nucleolar protein 5A (56kDa with KKE/D repeat)
AA732931	2.35 (1.69 to 3.792)	syntaxin 4A (placental)
AA279283	2.346 (1.943 to 2.909)	hypothetical protein FLJ20487
AA459690	2.335 (1.355 to 4.084)	brain protein 44-like
AA195276	2.334 (1.837 to 3.437)	ESTs, Weakly similar to L1 repeat, Tf subfamily, member 30
AA909002	2.333 (1.628 to 2.974)	ESTs
AI221641	2.331 (1.498 to 4.45)	potassium large conductance calcium-activated channel, subfamily M,
AA701664	2.329 (1.992 to 2.667)	ESTs
AA621132	2.327 (1.479 to 3.875)	aquaporin 9
AA437210	2.324 (1.687 to 3.485)	ESTs
AA669218	2.322 (2.14 to 2.478)	HBS1-like (S. cerevisiae)
AA778759	2.318 (1.729 to 2.718)	ESTs
H29895	2.304 (1.753 to 3.113)	guanine nucleotide binding protein (G protein)
AI339400	2.298 (1.261 to 3.984)	anti-Mullerian hormone receptor, type II
AA609992	2.297 (1.822 to 2.916)	NADP-dependent retinol dehydrogenase/reductase
AA479512	2.296 (1.996 to 2.842)	ESTs,
AI140628	2.294 (1.471 to 3.05)	EDAR-associated death domain
AA664210	2.287 (1.735 to 3.352)	protein kinase (cAMP-dependent, catalytic) inhibitor gamma
AA719285	2.283 (1.249 to 3.821)	Data not found
R10890	2.278 (1.518 to 4.861)	ESTs
AA699307	2.275 (1.281 to 4.157)	ESTs,

15

Appendix A

H73239 | AI821460	2.271 (1.356 to 3.454)	Data not found
W70293	2.271 (1.089 to 3.787)	Homo sapiens full length insert cDNA clone ZD58F01
AA884051	2.268 (1.877 to 2.904)	Data not found
AI142095	2.266 (1.608 to 3.355)	ESTs
AI140706	2.262 (1.45 to 4.045)	Homo sapiens cDNA FLJ11801 fis, clone HEMBA1006253
AA778646	2.257 (1.641 to 2.669)	hypothetical protein FLJ25955
AA463454	2.247 (1.955 to 2.842)	hypothetical protein FLJ10849
N59790	2.246 (1.752 to 3.124)	nuclear transcription factor, X-box binding 1
AA625858	2.245 (1.837 to 3.165)	ESTs,
AA490464	2.237 (1.766 to 3.111)	DKFZP564D116 protein
AA923758	2.233 (1.715 to 2.705)	hypothetical protein MGC10796
R91397	2.231 (1.425 to 4.259)	ESTs
H80879	2.225 (1.743 to 2.533)	Data not found
AA936710	2.225 (1.649 to 3.421)	COX15 homolog, cytochrome c oxidase assembly protein (yeast)
AA620642	2.224 (1.831 to 2.477)	hypothetical protein LOC51319
AI139861	2.221 (1.609 to 2.98)	Data not found
AA621332	2.22 (1.55 to 2.89)	ESTs
AA913206	2.208 (1.37 to 4.681)	G antigen 7B
AA706094	2.207 (1.747 to 2.87)	ESTs, Weakly similar to 2109260A B cell growth factor
AA883722	2.207 (1.556 to 3.355)	ESTs
AA668178	2.204 (1.785 to 2.812)	karyopherin alpha 3 (importin alpha 4)
R89855	2.2 (1.85 to 2.707)	ESTs, Highly similar to plakophilin 4 [Homo sapiens] [H.sapiens]
AI125753	2.196 (1.655 to 3.297)	Homo sapiens, clone IMAGE:4822830, mRNA
AA910876	2.196 (1.495 to 2.967)	KIAA0543 protein
AA625955	2.195 (1.812 to 3.188)	EST
AA905044	2.192 (1.326 to 3.108)	homeo box D8
AA485055	2.189 (1.972 to 2.338)	sperm associated antigen 6
H77461	2.186 (1.627 to 2.732)	lymphocyte adaptor protein
AA626152	2.178 (1.673 to 3.052)	ESTs
AI125490	2.176 (1.579 to 2.961)	EST, Moderately similar to T42689 hypothetical protein DKFZp434J1027.1
N63479	2.168 (1.807 to 2.504)	Data not found
H78407	2.167 (1.899 to 2.482)	ESTs
AA707468	2.163 (1.694 to 2.685)	olfactory receptor, family 7, subfamily E, member 104 pseudogene
H84154	2.161 (1.883 to 2.549)	cyclin D2
AA700155	2.158 (1.476 to 3.678)	Crm, cramped-like (Drosophila)
AA906879	2.153 (1.525 to 2.826)	hypothetical protein FLJ39441
AA897338	2.151 (1.298 to 2.936)	Homo sapiens, clone IMAGE:5271578, mRNA
AA777043	2.15 (1.567 to 3.207)	Data not found
AA598927	2.141 (1.747 to 2.7)	Data not found
AA969184	2.136 (1.972 to 2.44)	cardiac ankyrin repeat protein
AA877082	2.132 (1.98 to 2.284)	Homo sapiens, Similar to hypothetical protein FLJ20378
AA707672	2.131 (1.921 to 2.255)	membrane component, chromosome 11, surface marker 1
AA921818	2.129 (1.934 to 2.412)	ESTs
AA677206	2.129 (1.859 to 2.582)	ESTs
AA629344	2.123 (1.596 to 2.463)	Data not found
R02826	2.119 (1.662 to 2.537)	Data not found
AA634158	2.119 (1.643 to 3.018)	secretogranin III
AA909319	2.117 (1.506 to 3.332)	ESTs
AI205822	2.114 (1.095 to 3.093)	ESTs
AA935694	2.112 (1.803 to 2.867)	heparan sulfate (glucosamine) 3-O-sulfotransferase 2
AA158032	2.108 (1.757 to 2.675)	KIAA0783 gene product
H74017	2.107 (1.581 to 3.228)	ESTs
AI339376	2.103 (1.542 to 3.323)	crystallin, beta B1

16

Appendix A

H01504	2.101 (1.653 to 2.511)	Data not found
AA634425	2.1 (1.285 to 3.82)	ESTs, Weakly similar to AD20_HUMAN ADAM 20 precursor
R45255	2.099 (1.764 to 2.515)	EBNA1 binding protein 2
AA919087 | AI733140	2.097 (1.7 to 2.574)	Data not found
AI018179	2.094 (1.968 to 2.166)	ESTs
AA707503	2.092 (1.347 to 2.729)	DKFZP434I092 protein
W68220	2.087 (1.524 to 2.511)	KIAA0101 gene product
AA626249	2.085 (1.42 to 2.861)	ESTs
AA773608	2.082 (1.373 to 3.544)	Homo sapiens, Similar to hypothetical protein FLJ14840
AI339089	2.081 (1.788 to 2.522)	syntaxin 6
AA456027	2.081 (1.603 to 3.254)	hypothetical protein FLJ13081
AA694497	2.076 (1.739 to 2.408)	Data not found
AA923730	2.065 (1.583 to 2.909)	ESTs
AA608845	2.061 (1.445 to 3.729)	ESTs
R42174	2.057 (1.529 to 2.453)	sin3-associated polypeptide, 18kDa
AA778577	2.056 (1.72 to 2.671)	EST,
AA055114	2.046 (1.695 to 2.585)	phosphoinositide-3-kinase, regulatory subunit, polypeptide 1 (p85 alpha)
AA702678	2.046 (1.626 to 2.618)	hypothetical protein FLJ10204
AA454963	2.036 (1.686 to 2.248)	mitochondrial ribosomal protein L36
AI187980	2.033 (1.755 to 2.585)	Data not found
AA676548	2.017 (1.641 to 2.244)	ESTs, Moderately similar to hypothetical protein FLJ20234
AA115742	2.013 (1.864 to 2.193)	procollagen C-endopeptidase enhancer 2
AA437245	2.01 (1.776 to 2.224)	Homo sapiens, clone IMAGE:5744200, mRNA
N36948	2.005 (1.286 to 3.023)	ESTs, Weakly similar to hypothetical protein FLJ20294
W85710	2.002 (1.808 to 2.261)	carnitine palmitoyltransferase 1B (muscle)
AA126919	2 (1.413 to 2.442)	tyrosylprotein sulfotransferase 1
AA496113	1.996 (1.619 to 2.893)	ESTs, Moderately similar to PRO0478 protein
AA904562	1.988 (1.719 to 2.278)	Data not found
AA918288	1.987 (1.591 to 2.758)	Data not found
AI262978	1.981 (1.61 to 2.372)	methylenetetrahydrofolate dehydrogenase (NADP+ dependent)
AA706790	1.959 (1.3 to 2.794)	kinesin 2 60/70kDa
AI034168	1.955 (1.308 to 3.552)	ESTs
AA773190 | AI791127	1.95 (1.713 to 2.214)	Data not found
AA490843	1.931 (1.637 to 2.1)	anaphase promoting complex subunit 5
AA071122	1.92 (1.379 to 3.312)	Data not found
AA921738	1.869 (1.48 to 2.487)	Homo sapiens, clone IMAGE:5271140, mRNA
AA406388	1.863 (1.622 to 2.32)	hypothetical protein MGC10820
AA421300	1.806 (1.451 to 2.562)	zinc finger protein 261

Genes with significant 2 fold or more decreased expression in SW480M9 cells in comparison with SW480 cells

Gene ID	Fold change (range)	Gene Name
AA449832	0.531 (0.392 to 0.655)	adaptor-related protein complex 1, sigma 2 subunit
AI061186	0.495 (0.35 to 0.683)	Homo sapiens cDNA FLJ39845 fis, clone SPLEN2014452.
R51218	0.494 (0.397 to 0.685)	KIAA0092 gene product
AI286236	0.492 (0.456 to 0.563)	ESTs, Weakly similar to neuronal thread protein [Homo sapiens] [H.sapiens]
AA278630	0.491 (0.472 to 0.51)	differential display and activated by p53
AI675465	0.473 (0.418 to 0.589)	homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain
AA598496	0.473 (0.393 to 0.665)	IQ motif containing GTPase activating protein 1
AI359120	0.47 (0.442 to 0.501)	chromodomain helicase DNA binding protein 6
AI018218	0.464 (0.425 to 0.505)	ESTs
AA447746	0.46 (0.367 to 0.517)	HIF-1 responsive RTP801
AI080633	0.453 (0.42 to 0.497)	pyruvate dehydrogenase phosphatase
AI053778	0.443 (0.344 to 0.505)	ESTs, Weakly similar to hypothetical protein FLJ11267 [Homo sapiens] [H.sapiens]
H68542	0.442 (0.306 to 0.638)	ESTs
AI300286	0.44 (0.31 to 0.629)	Data not found
AI269079	0.439 (0.321 to 0.596)	ESTs
AA894927	0.408 (0.133 to 1.378)	asparagine synthetase
AA676405	0.402 (0.384 to 0.421)	argininosuccinate synthetase
AA676466	0.393 (0.354 to 0.437)	argininosuccinate synthetase
AA995233	0.365 (0.334 to 0.432)	ESTs, Highly similar to ASSY_HUMAN Argininosuccinate synthase
AI015679	0.298 (0.21 to 0.409)	phosphoserine aminotransferase

Appendix B

Sequence verification

APPENDIX II: Results obtained for the sequencing of chosen clones.

TENASCIN C

```
GCCGGGTTTCCNGTCACGACGTTGTAAAACGACGGCCAGTGAATTCCCTTGCGGCCGCAG
GNAATTTTTTTTTTTTTTTTTTTTTTTTAACAAACTCAAAAGTACTTGTGCTTTTATTTAAA
AAAAAAATACAATCAGGTACTGTCCAGAAATGTTTTGGAAAGAAAGATCTCTTGAAAAAT
CCTTAGTTTTCATCATCATCATCATCATTATTATATTAATAATATTAATCATATCCTTAA
AATGGAAACAGTATTGCTTTTCTGGTTTCTGTTGTATGAAATGTAAAAAAGGGATGGCT
TCCAATGACACATTTAATCTTTGCTAACAAAAATAATGACAATTAATTATACAGCTTCAT
GTAAAATACGGCTGGTTCTAAAACAAACTACCCCTGTACATCCTACCCCTCTCCCATTCC
CAGAGCCACCTAANAGAAGTAAAAAACTATTGCGATGTTGTCACTGGGAGATTTTTGCTG
AATCAAACAACAAAACAGAAACATGTTGGAGACTGATGTCTTTGGTGCAAAGAAAGAAAT
CNCAGAGGAGGTGAGGCCCATGCTGTTGCCGTTGGCCCCANGGATCCATGGTCAGCTTTG
ACTCTCACCAAATGCCCAGGTGTGGACCGAGGGTTGGGCTGGTTGTATTGAGCTTTGGTA
AAACCTTTCCTCGCTCTGGGCCTTATTCCTCTCTCCCCAGNGGCCCTGGAATTTATGCCC
GTTTGCGCCTGCCTTCAAAAATTCCGNAAATTGCTTGGCCNCANCTTCCTCCTCANCAAA
ANTGGGATTGAANGGTTCNNGGGCCCTTCCCCGNGGGAAACNNNTTTAACNCCCTGGACN
GGGGGTTATTGGCCCCCNTATCCCCCCCNTCNNGGTTGAACCNGGGGGACNNNTTCCNN
```

CAVEOLIN 1

```
GCCGGGTTTCCAGTCACGACGTTGTAAACGACGGCCAGTGCCAAGCTAAAATTAACCCTC
ACTAAAGGGAATAAGCTTGCGGCCGCATCTTTTTTTTTTTTTTTTTTCCTCCTCAGTTC
TTCATATTTCTCTCAAATTACCTTTCAACAATATGTTCATATCTCAGCCAATAAAGCGAT
GGTTGATCCACATGCAAATAAAAAGGTAAGGAAAAAAATTTAAATGGGTTTTAAAAATTC
AGACTGCCAAAAATAGATGAAATAGCTCAGAAGAGACATAATACTTATGAAAACTGAGAA
AGTGATCTTTTTATTTTCAACCAAGATAATTCAATTCATAAATTGTTGCTGTATTAGCAA
CTTGGAACTTGAAATTGGCACCAGGAAAATTAAAAGGAAAAAAAATCAGGTATACTTCTA
TCCTTGAAATGTCATTTATATTTCTTTCTGCAAGTTGATGCGGACATTGCTGAATATTTT
CCCAACAGCTTCAAANAGTGGGTCACAGACGGTGTGGACGTANATGGAATANACACGGCT
GATGCACTGAATCTCAATCAGGAAGCTCTTAATGCATGGTACAACTGCCCANATGTGCAG
GAAAGAGAGAATGGCGAAGTAAATGCCCCANATGAGTGCCATCGGGATGCCAAAGAGGGC
AGACAGCAAGCGGTAAAACCAGTATTTCGTCACAGTGAAGGTGGTGAAGCTGGCCTTCCA
AATGCCGTCAAAACTGTGTGTCCCTTCTGGTTCTGCAATCACATCTTCAAAGTCAATCTT
GACCACGTCCTCCTTGAAGTGTTTAAGGNCCCGGTTGGACCAANNCNATCCCTTGGGGGG
GNGCGTCNNAAACTTGGCTTCNCGCTCAACTNCGGCTGGNCANGGGCCN
```

PPAR-gamma

```
GCAGGGTTTCTCNTCACGGACGTTGTAANACGACGGCCAGTGNAATTGTAATACGACTCA
CTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGTTTTTTTTTTTTTTTTTTCATAAT
ATGGTAATTTTTAATATTAAAAGTAAATTGTAAATGTGTCTTTATAAACAATATGCATAA
AATAGATCATATTCTAAAACCTTTTCTTTTTAAAATGCTTTTTCACAGTAAATTTCTTAG
GTGTCANATTTTCCCTCANAATAGTGCAACTGGAANAAGGGAAATGTTGGCAGTGGCTCA
GGACTCTCTGCTAGTACAAGTCCTTGTANATCTCCTGCAGGAGCGGGTGAANACTCATGT
CTGTCTCCGTCTTCTTGATCACCTGCAGTAGCTGCACGTGTTCCGTGACAATCTGTCTGA
GGTCTGTCATTTTCTGGAGCAGCTTGGCAAACAGCTGTGAGGACTCAGGGTGGTTCAGCT
TCAGCTGGAGCTCCAGGGCTTGTAGCAGGTTGTCTTGAATGTCTTCAATGGGCTTCACAT
TCAGCAAACCTGGGCGGTCTCCACTGAGAATAATGACAGCAATAAATATTGCCAAGTCGC
TGTCATCTAATTCCAGNGCATTGAACTTCACAGCAAACTCAAACTTGGGCTCCATAAAGT
CACCAAAAGGCTTTCGCAGGCTCTTTAAAAACCCCTTGTCATGAANCCTTGGCCCCCGGA
TATGAAAACCCCCATCTTTTATTCNTCAANGANGCCCGCNTTGGTGTAAATGACCTCGNG
GACCCCCCTNTTTGGANGAAAAGTTNCCTTGGCCCTTTCCAGTCNAGAATTNCNAAACCC
CGNAANGGCTTTNTGGCANAAANCCGGGAACNCCCCGGCNCNNCCCTCCCCCGGAANNGA
N
```

ASPARAGINE SYNTHETASE

NNNNCNNNAANCATCATGACATGATTANGNATNTAATACGACTCACTATANGGAATNTGG
CCCTCGAGGCCAAGAATTCGGCACGAGGTTAATTTCCAAGTATATTCGGAAGAACACAGA
TAGCGTGGTGATCTTCTCTGGAGAAGGATCAGATGAACTTACGCAGGGTTACATATATTT
TCACAAGGCTCCTTCTCCTGAAAAAGCCGAGGAGGAGAGTGAGAGGCTTCTGAGGGAACT
CTATTTGTTTGATGTTCTCCGCGCAGATCGAACTACTGCTGCCCATGGTCTTGAACTGAG
AGTCCCATTTCTAGATCATCGATTTTCTTCCTATTACTTGTCTCTGCCACCAGAAATGAG
AATTCCAAAGAATGGGATAGAAAAACATCTCCTGAGAGAGACGTTTGAGGATTCCAATCT
GATACCCAAAGAGATTCTCTGGCGACCAAAAGAAGCCTTCAGTGATGGAATAACTTCAGT
TAAGAATTCCTGGTTTAAGATTTTACAGGAATACGTTGAACATCAGGTTGATGATGCAAT
GATGGCAAATGCAGCCCAGAAATTTCCCTTCAATACTCCTAAAACCAAAGAAGGATATTA
CTACCGTCCAAGTCTTTGAACGCCATTACCCAGGCCGGGCTGACTGGCTGANCCATTACT
GGATGCCCAAGTGGATCAATGCCACTGGACCTTCTGCCCGCACGCTGGACCACTACAAGT
CAGCCTGTCAAAGCTTAAGTGGGCTTTTTGCTGGTAATGNNGAAAGCCAAATATTTCCTC
CTGGTTGGGATGGGGGGGACTGGGGGGGCANATAAGGGGNAACAATGGNNAGTCAACCTCC
AGGCTAACCTTGGGGNGNGAAAAAAAATNAAAGTCCCTAATCTAAAAAAAAAAAAAAAAA
ANCATTGGGGGCCCCAGCTTNTCCCTTNAAGGAGGGTAATTTACCTGGNCCTNGCCCCNT
TTCCACGCCNGGNNGGGAAACCTGGGNNTACCANTTTCCCCTTGGACCNC

CATHEPSIN C

GCCGGGTTTCCCTCACGACGTTGTAANACGACGGCCAGTGNAATTGTAATACGACTCACT
ATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGTTTTTTTTTTTTTTTTTTTTATTTTT
AAAACTTTATTTTAAAAATATGAGCATCTATTTTAAAAGTTTTGATAATTATTGCCATTA
TTTTCTTGTGATTGGTACAATTTAAAAATAAGTCTATGTTTTCACATTGATTTTAAAAAA
TATAGCATGTTTGAATTACAAATGATTAAGCAAACTCTATTACTTCATAGCTGACCATCT
TCCAGAAAATTCCCACTTAATTGAATACTTAGAAAAAAATGGCCAGTGGCCGATTGAAAG
GTATATTAAAATTAAGGGCAGTTTTAATTCTGAAGACAAATATCTTCATGGAAATCTATT
TGTAAGCTTCTGANATTGCTGCTGAAAGTCTACAGTCTGTGAATATACCAATTCCCCTTT
ACAACTGATGCAGATCATTATGAAATACTGGAAGGCATACCCTACAATTTAGGAATTGGT
GTGGCTGCCACTGCTATGCTCTCAATTGCACACTCATCAGTTCCTCTGCGGATCCGGAAG
TAGCCATTCTCACCCCAGCCGGTGCCCCAGCTGTTTTTAACAATCCAGTAATCCATCCCA
GANGCTGANTCAGTGCCNTAGCCCACAAGCANAACAGCATGATTAGTCAGCTCCAAGGGG
TTGAAAGGGTCTCTAAACCNGTGNGGNGGTANATCCCCCTTTTGTNGNGGGAGNAAGTCN
TCCTATACTTCCAAAAGCCACNGCCCTNGGGCCCATGAAGGGACCCAACCTCCAGGCTTC
CNTCAAGGNCTTCATTGGCNGCCCCCCCCAAAAAAACCCNCCTNCNATAGGGGGNAACC

Appendix C

Genes with common changes in expression over all experiments

Genes with significant 2 fold or more increased expression in SW480M7 and SW480M9 cells in comparison with SW480 cells

Gene #	Gene id
AA437210	ESTs
AA437245	Homo sapiens, clone IMAGE:5744200, mRNA
AA479344	gap junction protein, beta 5 (connexin 31.1)
AA479512	ESTs, Weakly similar to B36298 proline-rich protein PRB3S (cys) - human (fragment) [H.sapiens]
AA479950	ATPase, Class II, type 9B
AA609630	Data not found
AA620619	ESTs
AA620642	hypothetical protein LOC51319
AA625858	ESTs,
AA626716	Data not found
AA634483	Homo sapiens, clone IMAGE:5266541, mRNA
AA669218	HBS1-like (S. cerevisiae)
AA669367	AI732124	Data not found
AA676548	ESTs, Moderately similar to hypothetical protein FLJ20234 [Homo sapiens] [H.sapiens]
AA694497	Data not found
AA706094	ESTs, Weakly similar to 2109260A B cell growth factor [Homo sapiens] [H.sapiens]
AA773190	AI791127	Data not found
AA777779	ESTs
AA778577	EST
AA778611	KIAA1086 protein
AA778617	ESTs
AA778851	Human clone 137308 mRNA, partial cds.
AA912886	ESTs
AA917956	Homo sapiens, clone IMAGE:3878708, mRNA
AA921818	ESTs
AI125237	ESTs
AI125753	Homo sapiens, clone IMAGE:4822830, mRNA
AI139861	Data not found
AI140549	ESTs
AI140566	ESTs
AI140695	Data not found
AI142095	ESTs
AI188006	ESTs
AI198232	C-terminal PDZ domain ligand of neuronal nitric oxide synthase
AI206318	gamma-aminobutyric acid (GABA) A receptor, beta 3
AI250799	defensin, alpha 4, corticostatin
AI253652	AI793206	Data not found
AI276487	cytokeratin type II
AI288845	chemokine (C-C motif) receptor-like 2
H45289	ESTs
H77461	lymphocyte adaptor protein
N/A-2-1-15-4	Data not found
N/A-3-1-14-6	Data not found
N22685	plexin C1
N63479	Data not found

Genes with significant 2 fold or more decreased expression in SW480M7 and SW480M9 cells in comparison with SW480 cells

Gene #	Gene id
AA894927	asparagine synthetase
AA995233	ESTs, Highly similar to ASSY_HUMAN Argininosuccinate synthase (Citrulline--aspartate ligase) [H.sapiens]
AI015679	phosphoserine aminotransferase

Genes with significant 2 fold or more increased expression in SW480M7 and SW620 cells in comparison with SW480 cells

Gene #   Gene id

AA425128   ESTs

Genes with significant 2 fold or more decreased expression in SW480M9 and SW620 cells in comparison with SW480 cells

Gene #   Gene id

AI080633   pyruvate dehydrogenase phosphatase

Increased Invasion and Expression of MMP-9 in Human Colorectal Cell Lines by a CD44-dependent Mechanism

DAVID MURRAY[1], MARY MORRIN[2] and SUSAN MCDONNELL[1]

[1]*School of Biotechnology, National Institute for Cellular Biotechnology, Dublin City University, Dublin 9;*
and [2]*Limerick Institute of Technology, Limerick, Ireland*

bstract. *Background: In order to investigate the interactions tween MMPs and CD44 we stably transfected a non-invasive lon cell line, SW480 with the cDNA for MMP-9 and vestigated the effect on CD44 expression and in vitro vasion and migration. Materials and Methods: In vitro vasion and migration assays were carried out using Biocoat atrigel invasion chambers. MMP and CD44 expression was termined using zymography, Western blot analysis and RT-CR. Results: Transfection of the parental SW480 cells with e cDNA for MMP-9 (SW480M9) caused increased invasion d migration. MMP-9 expression increased when the V480M9 cells were grown on HA and collagen and cultured the presence of a CD44-activating antibody. Treatment of e cells with HA and a CD44-activating antibody also resulted increased invasion, migration and attachment. Conclusion: iese results demonstrated that CD44 and MMP interactions e important in controlling tumour cell invasion and igration.*

ie matrix metalloproteinases (MMPs) are a family of highly nserved zinc-dependent endopeptidases which collectively e capable of degrading components of the extracellular atrix (ECM). The MMP family currently contains 23 embers that can be divided into at least 5 classes based on bstrate specificity: gelatinases, collagenases, stromelysins, embrane-type MMPs and others (1). They have been shown rough numerous studies to be key role players in both irmal and disease processes, but it is their role in tumour vasion and metastasis which has been most extensively udied (2). Metastasis is the spread of tumour cells from a imary tumour to distant sites in the body and is the most thal aspect of cancer.

orrespondence to: Susan McDonnell, PhD, School of Biotechnology, ublin City University, Glasnevin, Dublin 9, Ireland. Tel: 353-1-)05244, e-mail: susan.mcdonnell@dcu.ie

MMP activity has been implicated at almost every stage of the metastatic cascade from local invasion at the primary tumour site to growth and development of a tumour at a secondary site (2). The substrate specificities of the two gelatinases, MMP-2 and MMP-9, differ from other MMPs in that they can degrade type IV collagen which is present in the basement membrane (1). They have been shown in several studies to function as independent predictors of disease outcome and recurrence (3).

CD44, the major cell surface receptor for hyaluronic acid (HA), is a multifunctional transmembrane adhesion molecule that also plays a key role in cell signalling (4). CD44 mediates cell-cell and cell-matrix interactions and is involved in adhesion, migration and invasion. It has been estimated that 45 CD44 variants exist as a result of alternative mRNA splicing (5). Abnormal CD44 variant expression has been associated with tumour progression and over-expression of CD44v6 has been observed more often in advanced colon cancers (Dukes' stage C/D) compared with Dukes' A/B, and it can therefore be used as a prognostic marker (6).

Both MMPs and CD44 are critical to the invasive process and several studies have recently begun to define how these molecules interact with each other. Stimulation of CD44 has been shown to increase MMP-2 expression in human melanoma cells (7) and in lung carcinoma cells (8). MMP-9 has been shown to associate with CD44 on the surface of mouse mammary carcinoma and human melanoma cells (9). CD44v3 can anchor MMP-7 on the cell surface and target its activity (10). A recent paper has also shown that MT1-MMP can cleave CD44 and promote cell migration in a pancreatic tumour cell line (11).

Appreciating the importance of both attachment to and degradation of the ECM in tumour cell invasion, it is no longer appropriate to consider CD44 and MMPs as acting in isolation. In this study the interaction between CD44 and MMP-9 in regulating invasion and migration in colorectal tumour cell lines was examined. The non-metastatic SW480 cell line was stably transfected with the cDNA for MMP-9. The effect of MMP-9 on cell adhesion and migration was

assessed and cells were grown on various ECM components to study the effect on MMP activity. Activation of CD44 through ligand binding and antibody treatment caused increased MMP expression and invasion. Inhibition of CD44 activity by antibody treatment resulted in decreased invasion. These results demonstrated the importance of MMP and CD44 interactions in regulating colon tumour cell invasion.

Materials and Methods

Cell culture and transfection of SW480 cells. The cell lines used in this study were SW480, (ATCC# CCL-228) established from a primary human colon adenocarcinoma, SW620 (ATCC# CCL-227), established from a lymph node metastasis from the same patient as the SW480 and BHK92 (BHK cells transfected with the cDNA for MMP-9, a gift from Dylan Edwards, University of East Anglia). All cell lines were maintained in Dulbecco's modification of Eagles' medium (DMEM) supplemented with 5% foetal calf serum (FCS) in a humid 5% CO_2 atmosphere at 37°C. The SW480 cells were transfected with 10µg pCMV92 (human MMP-9 cDNA sub-cloned into pCMV expression vector) as previously described (12) and 2 µg pSV2neo using Lipofectamine (Invitrogen Paisley, UK) following the manufacturer's instructions. Cells were selected in 600 µg/ml Geneticin (G-418) for 4 weeks. Individual colonies were selected using cloning cylinders (Sigma, Dublin, Ireland).

Zymography. Conditioned media (CM) was collected by incubating 80% confluent cultures for 72 h in serum-free media. CM was concentrated using Microcon™ centrifugal filter devices (Amicon, Cork, Ireland) and protein concentration determined using the bicinchoninic acid (BCA) assay (Pierce, Tattenhall, UK). Gelatinase activity was assessed in 30 µg total protein as previously described (13). To investigate the effect of CD44 stimulation on MMP activity, cells were incubated with 50 µg/ml CD44 activating antibody $F_{10-44-2}$ (Biodesign, Saco, ME, USA) or with mouse IgG (Sigma) as control in serum-free media. Cells were also treated with 0.25mg/ml HA, 0.01mg/ml Collagen type IV or 0.3 µg/ml anti-β1-integrin stimulating antibody (Sigma) in serum-free media for 72 h.

Western blot analysis. Western blot analysis of concentrated CM (as prepared for zymography) was carried out for detection of MMP-2 and MMP-9. Western blot analysis for detection of CD44 was carried out on total cell lysates prepared as previously described (14). Equal amounts of protein (50µg) were separated on 10% sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, transferred and blocked in 5% dried milk in TBST (0.15M NaCl, 0.01M Tris, 0.05% (v/v) Tween-20). Primary antibodies were diluted; D2.1 mouse pan specific anti-CD44 (a gift from Aideen Long, Royal College of Surgeons in Ireland) 1/1000, MAC96 mouse anti-MMP-9 (a gift from Andrew Docherty, Celltech) 1/500, GL8 mouse anti-MMP-2 (a gift from Andrew Docherty, Celltech, Dublin, Ireland) 1/2000 and mouse anti-actin (Merck Biosciences, Nottingham, UK) 1/10,000 in the same blocking solution and incubated overnight at 4°C. The blots were washed, incubated with anti-mouse-horseradish peroxidase (HRP) (Amersham, Little Chalfont, UK) 1/5000 in TBST for 2 h at room temperature and then developed using the enhanced chemoluminesence (ECL) system (Pierce).

RNA isolation and reverse transcription polymerase chain reaction (RT-PCR). Total RNA was isolated from equal cell numbers using TRI REAGENT (Sigma) following the manufacturer's instructions. Reverse transcription was as previously described and 5µl of the resultant cDNA was used as template for PCR (13). The sequence for primers with annealing temperatures indicated in brackets were as follows: MMP-9-5' ATG AGT TCG GCC ACG CGC TG GCT T 3': 3' T GCC GGT GAT GAC ACG GAA ACT CAG (54 °C); β-Actin-5' TCA GGA GGA GCA ATG ATC TTG A 3' TC GAA GAG GAA TTA CAG TGC GTG CTA AAG (55°C). The number of PCR cycles used was 30 and the expected product size after primer amplification was as follows: MMP-9: 3 base pairs (bp) and β-actin: 682bp.

In vitro invasion and migration assays. Biocoat™ matrigel invasion chambers (Becton Dickinson Labware, UK) were used to test the *in vitro* invasive activity of the cell lines. 5×10^5 cells were seeded into the upper chamber and incubated for 24 h at 37°. Membranes were fixed and stained as previously described (13) and invasive cells counted in five 10X magnification fields. Results were expressed as average cell count per field. To investigate the effect of HA on cell invasion, cells were seeded in serum-free media containing 0.25mg/ml HA in both chambers. SW480M9 cells were allowed to invade in the presence of 50 µg/ml, CD44-activating antibody $F_{10-44-2}$ (Biodesign), CD44-blocking antibody IM7.8.1 (gift from Aideen Long, Royal College of Surgeons, in Ireland), or mouse IgG (Sigma) as a control. Migration assays were carried out under the same conditions except in the absence of matrigel. All experiments were done in duplicate and were repeated three times.

Adhesion assays. To investigate the adhesion of cells to various ECM components, a flat-bottomed 96-well plate was coated overnight at 4°C with either 1mg/ml HA or 50 µg/ml fibronectin (FN). The plate was washed twice with phosphate-buffered saline (PBS), blocked with 1mg/ml bovine serum albumin (BSA) for 2 at 4°C and then 1×10^5 cells were added and allowed to adhere for 90 min at 4°C. Residual media was aspirated and the adherent cells were incubated in media containing [3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS reagent) (Promega, Southampton, UK). Absorbance was read at 490nm after incubation at 37°C for 3 h. The absorbance value obtained by seeding untreated wells (*i.e.* media was not aspirated) represented 100% adhesion and all other values were divided by this to calculate percentage adhesion. Experiments were performed in triplicate on three separate occasions and T-test performed to ensure significance.

Results

Over-expression of MMP-9 in the SW480 cells and characterisation of colon cell lines. Following transfection of the SW480 cells with the cDNA for MMP-9, a total of 2 clones were selected and characterized. Clone 11 (referred to as SW480M9) was chosen for further analysis as secreted the highest levels of MMP-9 protein as shown by gelatin zymography (Figure 1A). SW480M9 showed marked increase in MMP-9 levels in comparison to the parental SW480 and its lymph node metastatic derivative SW620. The identity of MMP-9 was confirmed by Western
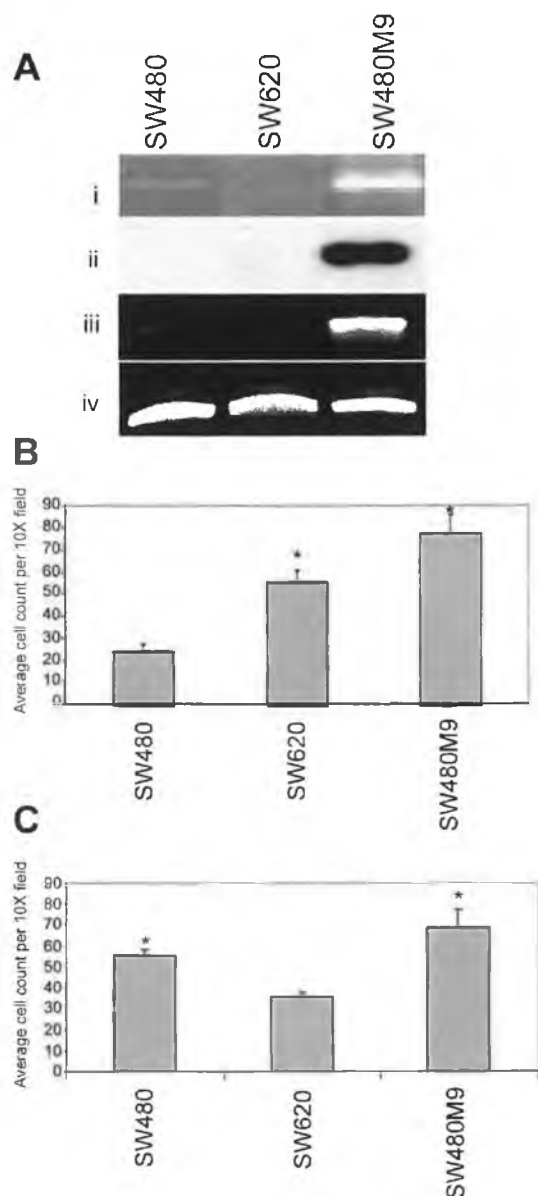
A

B

C

Figure 1. *Characterisation of colon cell lines. (A) MMP-9 expression in SW480, SW620 and SW480M9 cells was analysed by; (i) gelatin zymography, (ii) Western blot, (iii) RT-PCR with (iv) RT-PCR for β actin as a control. (B) Comparison of in vitro invasion of the SW480, SW620 and SW480M9 cells. (C) Comparison of in vitro migration of cell lines. The asterisk denotes that p<0.005.*
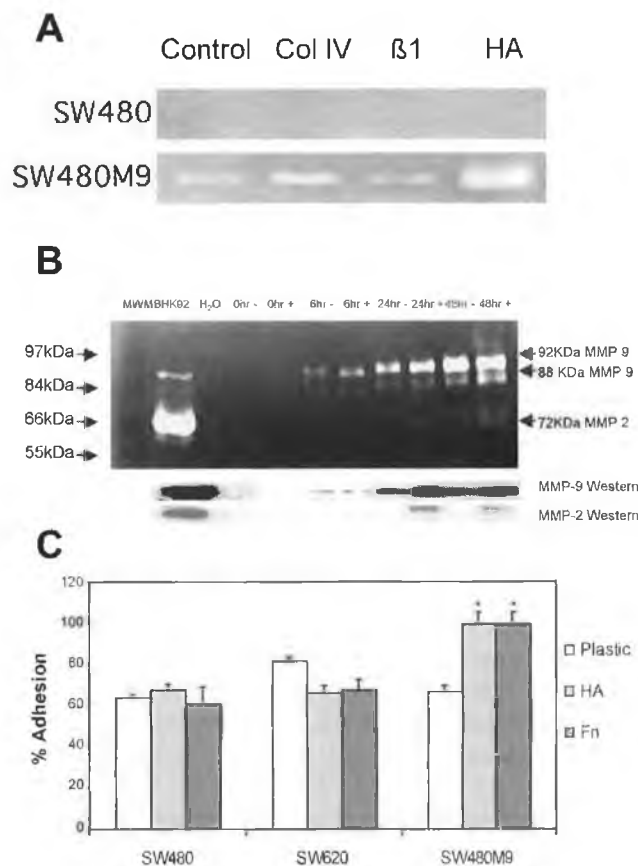
A

B

C

Figure 2. *The effect of CD44 stimulation on MMP-9 expression. (A) MMP-9 expression in SW480 and SW480M9 cells grown in the presence of type IV collagen, HA and β_1-integrin-activating antibody. (B) MMP-9 and MMP-2 expression in the SW480M9 cells treated with CD44-activating antibody over 48 h by zymography and Western blot using BHK92 as a positive control. (C) Comparison of cell adhesion to plastic, HA and FN. The asterisk denotes that p<0.005.*

lot analysis (Figure 1a) using an MMP-9 specific antibody. MMP-9 was not detected in either the SW480 or SW620 cells by Western blot analysis. RT-PCR data (Figure 1A) confirmed over-expression of MMP-9 in the SW480M9 cells at the mRNA level when compared to the SW480 and the SW620, which only showed faint bands. β-actin was used as an equal loading control for RT-PCR.

Since ECM degradation is key to tumour cell invasion, the *in vitro* invasiveness of these cell lines through matrigel-coated membranes was compared. The SW480 cells showed little invasion (Figure 1B) in comparison to the SW620 and SW480M9 cells, which were significantly ($p<0.005$) more invasive. The SW620 and SW480M9 were respectively twice and three times more invasive. The migratory capacities of the cells were determined using the same procedure as the *in vitro* invasion assays in the absence of matrigel. Over-expression of MMP-9 in the SW480M9 cells increased their migratory capacity in comparison to the parental SW480 cells (Figure 1C). However, the parental SW480 cells were significantly ($p<0.005$) more migratory than the SW620, which were established from a lymph node metastasis from the same patient.
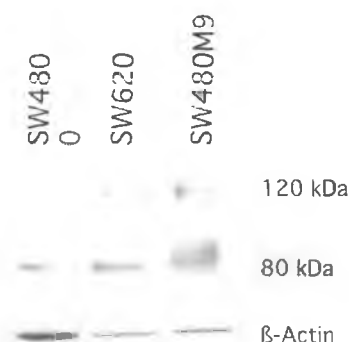
Figure 3. *CD44 expression in colon cell lines. CD44 expression profiles of the SW480, SW620 and SW480M9 cell lines by Western blot analysis using β-actin as an equal loading control.*

Effect of ECM components and CD44 activation on MMP-9 expression and cell adhesion. To investigate the effect of ECM components and integrin signalling on MMP-9 expression, cells were cultured in the presence of HA, type IV collagen and a β_1-integrin-activating antibody. After treatment, neither the SW480 or SW620 cells, which did not originally secrete measurable levels of MMP-9, showed any increase in MMP-9 expression levels (Figure 2A and data not shown). The SW480M9 cells showed no increase in MMP-9 activity in the presence of the β_1-integrin-stimulating antibody and a slight increase when grown on type IV collagen. Growth of the SW480M9 cells on HA greatly increased the levels of MMP-9. Since CD44 is the major cell surface receptor for HA, the SW480M9 were treated with an anti-CD44 monoclonal antibody ($F_{10-44-2}$). Instead of blocking its activity, $F_{10-44-2}$ activates CD44, mimicking its ligation of HA in the ECM and subsequently activating its associated intracellular signalling pathways. CM from antibody-treated cells was collected at timed intervals and analysed for MMP-9 expression. Within 6 h of treatment with the CD44-activating antibody, cells showed a slight increase in MMP-9 activity (Figure 2B) when compared with cells treated with a control mouse IgG. This difference in activity was more obvious after 24 and 48 h. Interestingly, activation of pro-MMP-9 to its active 88 kDa form occurred after 24 and 48 h. MMP-2 activity was also detected after 24 and 48 h of CD44 activation. These observations were confirmed by Western blot analysis (Figure 2B) with specific MMP-2 and MMP-9 antibodies.

To further investigate the behaviour of cells in the presence of ECM components, adhesion assays were carried out in the presence of HA and FN. Adhesion is a key event in the metastatic process where cells must first adhere to the ECM prior to its degradation. Increased cell-cell signalling and contact is also mediated by increased expression of cell
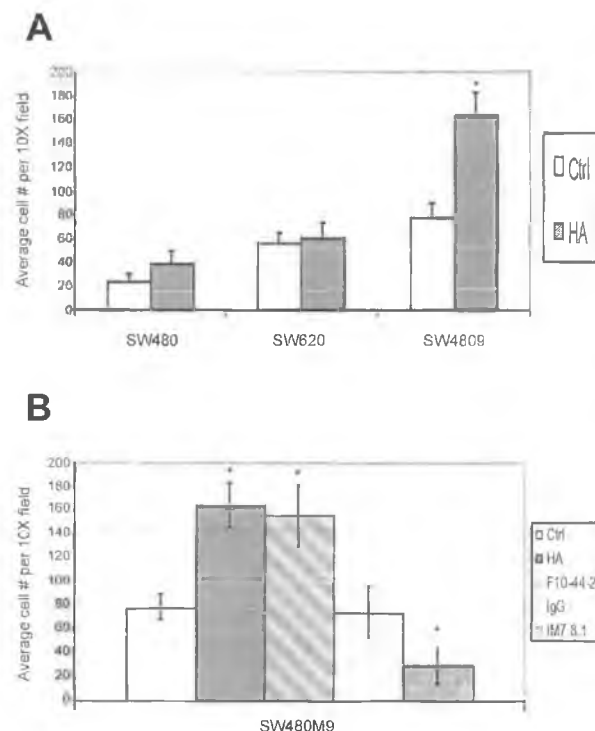

Figure 4. *The effect of enhanced expression of MMP-9 via CD44 on i vitro invasion. (A) In vitro invasion of the SW480, SW620 and SW480M cells in the presence of HA. (B) The in vitro invasion of the SW480M cells in the presence of HA, CD44-activating antibody ($F_{10-44-2}$), mous IgG (as a control) and CD44-blocking antibody (IM7.8.1). The asteris denotes that $p < 0.005$.*

adhesion molecules. The SW480M9 cells, which were th most responsive to HA treatment, also showed the mos adherence to FN and HA (Figure 2C). All percentag adhesion values were significant ($p < 0.005$).

Expression and up-regulation of CD44 in MMP-9 transfecte cells. Due to the increased adhesion of the SW480M9 cell on HA and the enhancing effect of HA and CD4 stimulation on MMP-9 activity, the expression of CD4 protein was examined. All cell lines studied expressed th 80kDa CD44 (Figure 3), often referred to as CD44s o standard CD44. The SW480M9 cells expressed the highes levels of CD44s compared to the SW480 and SW620 cells Interestingly, the metastatic SW620 and the more invasiv SW480M9 showed expression of a 120kDa variant of CD4 with SW480M9 showing highest levels of expression. Thes differences in CD44 expression may help explain th increased adhesion of SW480M9 to HA as well as the effec HA had on MMP-9 activity in these cells.

Effect of CD44 activation on in vitro invasion. To show that CD44-mediated up-regulation of MMP-9 expression had an effect on cell invasion, the *in vitro* invasion of the cells in the presence of HA was assayed. The presence of HA increased invasion in all cell lines tested (Figure 4A). The most dramatic increase was seen with the SW480M9 cells where twice as many cells ($p < 0.005$) invaded when treated with HA. Therefore, the effect of a CD44-activating antibody on invasion was studied. CD44 activation increased the *in vitro* invasion of SW480M9 cells two-fold ($p < 0.005$) (Figure 4B) when compared with the invasion of the same cells treated with control mouse IgG. A CD44-blocking antibody (IM7.8.1) was used to further confirm that the increased invasion was CD44-dependent. Blocking of CD44 with this antibody resulted in a significant reduction ($p < 0.005$) in invasion (Figure 4B).

Discussion

In this study the interactions between CD44 and MMP-9 in regulating invasion and migration in human colorectal cancer cell lines was investigated. The non-metastatic SW480 cell line, established from a primary human colon adenocarcinoma, was stably transfected with the cDNA for MMP-9. Over-expression of MMP-9 in a rat transformed embryo cell line has previously been shown to increase the invasion and migration of the cells (15). This paper was the first report of MMP-9 over-expression in a human colon cell line. The parental SW480 cells do not express MMP-9 and transfection of the cDNA for MMP-9 into these cells increased their invasion and migration. Interestingly, the SW620 cells, which were established from a lymph node metastasis from the same patient as the SW480, also did not express MMP-9. The MMP-9-transfected cells also showed increased migratory capacity. Surprisingly, the parental SW480 cells were more migratory than the SW620. A similar result was also obtained by Kubens and Zanker using a three-dimensional collagen matrix and time-lapse video recording assay (16).

Within the past few years, many researchers have demonstrated the importance of MMP interactions with cell adhesion molecules in regulating tumour cell invasion and metastasis. Several investigators have shown the importance of integrins in regulating MMP expression *e.g.* the $\alpha V\beta 6$ integrin promotes invasion of squamous carcinoma cells through up-regulation of MMP-9 (17). Stimulation of CD44 has been shown to increase MMP-2 expression in human melanoma cells (7) and in lung carcinoma cells (8). In this study we showed that culturing colon cells on ECM components increased MMP-9 expression. HA, the major receptor for CD44, caused the greatest increase. In addition, direct stimulation of CD44 by treating with an activating antibody ($F_{10-44-2}$) caused an increase in MMP-2 and MMP-

9 expression and also caused activation of MMP-9 to its active form. This is the first report to show up-regulation of MMP-9 expression following CD44 stimulation in colon cells.

We also looked at adhesion of the cells on various matrices including, plastic, FN and HA. The MMP-9-transfected cells attached better to FN and HA than any of the other cell lines. Examination of the levels of CD44 protein expressed in the cell lines showed that all three cell lines expressed the standard form of CD44. An additional higher molecular weight band was seen in the metastatic SW620 and the MMP-9-transfected cell line. This band had a molecular weight of ~ 120kDa and could be CD44v6, which has been previously been shown to be involved in colon metastasis (6). Exon-specific PCR would be needed to confirm the identity of this band.

Since we had shown that growth of the SW480M9 cells on HA and direct stimulation of CD44 caused an increase in MMP activity, we then looked at whether this increase in MMP-9 activity had a functional effect on cell invasion and migration. Stimulation of CD44 by treatment with HA and a CD44-activating antibody caused an increase in the invasion of the SW480M9 cells, which could be inhibited by treatment with a CD44-blocking antibody.

These results suggest that CD44 activation can directly increase the levels of MMP-9 and cause an increase in cell invasion in colon cell lines. Interestingly, Yu and Stamenkovic (9) have shown that MMP-9 can associate with CD44 on the surface of mouse mammary carcinoma and human melanoma cells. This association may cause activation of signal transduction molecules, *e.g.* HA-CD44 binding activates the ras-MEK1 and the PI3 kinase-Akt signalling pathways in a human lung carcinoma cell line (8). Future experiments will determine the signalling mechanisms involved in the up-regulation of MMP-9 expression in these cells. In addition, using MMP-9 promoter constructs, we will attempt to elucidate the transcriptional mechanisms involved in regulating MMP-9 gene expression in these cells.

References

1   Somerville, RPT, Oblander, SA and Apte SS: Matrix metalloproteinases: old dogs with new tricks. Genome Biol *4*: 216-226, 2003.
2   Coussens LM, Fingleton B and Matrisian LM: Matrix metalloproteinase inhibitors and cancer: trials and tribulations. Science *295*: 2387-2392, 2002.
3   Vihinen P and Kahari VM: Matrix metalloproteinases in cancer: prognostic markers and therapeutic targets. Int J Cancer *99*: 157-166, 2002.
4   Lesley J, Hyman R and Kincade PW: CD44 and its interaction with extracellular matrix. Adv Immunol *54*: 271-335, 1993.
5   van Weering DHJ, Baas PD and Bos JL: A PCR-based method for the analysis of human CD44 splice products. PCR Methods Appl *3*: 100-106, 1993.

6 Wielenga VJM, Heider KH, Offerhaus JA, Adolf GR, van den Berg FM, Ponta H, Herrlich P and Pals ST: Expression of CD44 variant proteins in human colorectal cancer is related to tumor progression. Cancer Res 53: 4754-4756, 1993.

7 Takahasi K, Eto H and Tanabe K: Involvement of CD44 in matrix metalloproteinase-2 regulation in human melanoma cells. Int J Cancer 80: 387-395, 1999.

8 Zhang Y, Thant AA, Machida K, Ichigotani Y, Naito Y, Hiraiwa Y, Senga T, Sohara Y, Matusuda S and Hamaguchi M: Hyaluronan-CD44s signalling regulates matrix metalloproteinase-2 secretion in a human lung carcinoma cell line QG90. Cancer Res 62: 3962-3965, 2002.

9 Yu Q and Stamenkovic I: Localization of matrix metalloproteinase 9 to the cell surface provides a mechanism for CD44-mediated tumor invasion. Genes Dev 13: 35-48, 1999.

10 Yu WH, Woessner JF, McNeish JD and Stamenkovic I: CD44 anchors the assembly of matrilysin/MMP-7 with heparin-binding epidermal growth factor precursor and ErbB4 and regulates female reproductive organ remodelling. Genes Dev 16: 307-323, 2002.

11 Kajita M Itoh Y, Chiba T, Mori H, Okada A, Kinoh H and Seiki M: Membrane-type 1 matrix metalloproteinase cleaves CD44 and promotes cell migration. J Cell Biol 153: 893-904, 2001.

12 Witty JP, McDonnell S, Newell KJ et al: Modulation of matrilysin levels in colon-carcinoma cell-lines affects tumorigenicity in vivo. Cancer Res 54: 4805-4812, 1994.

13 Lynch CC and McDonnell S: The role of matrilysin (MMP-7) in leukaemia cell invasion. Clin Exp Metastasis 18: 401-406, 2001.

14 Harada N, Mizoi T, Kinouchi M et al: Introduction of antisense CD44s cDNA down-regulates expression of overall CD44 isoforms and inhibits tumor growth and metastasis in highly metastatic colon carcinoma cells. Int J Cancer 91: 67-75 2001.

15 Bernhard EJ, Gruber SB and Muschel RJ: Direct evidence linking expression of matrix metalloproteinase 9 (92-kDa gelatinase/collagenase) to the metastatic phenotype in transformed rat embryo cells. Proc Natl Acad Sci USA 91: 4293-4297, 1994.

16 Kubens BS and Zanker KS: Differences in the migration capacity of primary human colon carcinoma cells (SW480) and their lymph node metastatic derivatives (SW620). Cancer Lett 131: 55-64, 1998.

17 Thomas GJ, Lewis MP, Whawell SA et al: Expression of the alphavbeta6 integrin promotes migration and invasion in squamous carcinoma cells. J Invest Dermatol 117: 67-73, 2001.