Clinical Applications of Molecular Profiling in Colorectal Cancer

A thesis submitted to the
National University of Ireland, Galway
for the degree of
Doctor of Philosophy

by

Dr. Elrasheid Ahmed Hassan Kheirelseid
MBBS, MRCS

Dicipline of Surgery
College of Medicine, Nursing & Health Sciences
National University of Ireland, Galway

Under the supervision of
Dr. Nicola Miller PhD
and the direction of
Professor Michael Kerin MCh, FRCSI

2011

This thesis is dedicated to:

My wife Reem and
my children Haneen, Ahmed & Nadia
for their
understanding, encouragement and support

My parents
Ahmed & Zainab
in acknowledgement of their scarifications

&

The soul of my grandmother Bit Karar
who taught us early
how to be in the use of meaningful life

TABLE OF CONTENTS

Acknowledgements

First and foremost I offer my sincerest gratitude to Professor Michael Kerin, who has supported me throughout, with his patience, knowledge and crucial contribution, which made him a backbone of this research and so to this thesis. Professor Kerin enthusiasm for the practice of surgery, patient care and translational research has been inspirational and his energy and insight has been highly motivational.

I am heartily thankful to Dr. Nicola Miller whose encouragement, guidance and support from the initial to the final level enabled me to develop an understanding of the subject. Nicola was always available for discussions regarding experiments and meticulously appraised manuscripts, presentations and thesis drafts, and this thesis would not have been possible without her input.

I convey special acknowledgement to Dr. Margate Sheehan, Department of pathology for her indispensable help in colorectal samples bio-banking, immunohistochemistry analysis and MMR study.

I would like to show my gratitude to Ms. Catherine Curran and Ms. Emer Hennessy who always facilitated the acquisition of data and tissue specimens and again for their technical assistance with experiments.

I would also acknowledge Dr. Kah Hoong Chang and Dr Mary Nugent “the colorectal team” for their advice and their willingness to share their bright thoughts with me, which was very fruitful for shaping up my ideas and research. In addition; I owe my deepest gratitude to all the researchers in Room 108 and in the Department of Surgery laboratory who were a pleasure to work with over the last few years

I would like to acknowledge the funding from the National Breast Cancer Research Institute (NBCRI) which made this study possible.

Finally, it is a pleasure to thank those who made this thesis possible, all colorectal cancer patients who have agreed to be involved in this research.

List of Figures

List of Tables

Communications originating from this work

Peer Reviewed Published Manuscripts

1- Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer.
EAH Kheirelseid, KH Chang, J Newell, MJ Kerin, N Miller.
BMC Molecular Biology 2010 feb; 11:12

2- MicroRNA-21 and PDCD4 expression in colorectal cancer.
Chang KH, Miller N, **Kheirelseid EAH**, Ingoldsby H, Hennessy E, Curran CE, Curran S, Smith MJ, Regan M, McAnena OJ, Kerin MJ
Eur J Surg Oncol. 2011 May; 37(7):597-603

3- MicroRNA signature analysis in colorectal cancer: identification of expression profiles in stage II tumors associated with aggressive disease.
Chang KH, Miller N, **Kheirelseid EA**, Lemetre C, Ball GR, Smith MJ, Regan M, McAnena OJ, Kerin MJ
Int J Colorectal Dis. 2011 Nov; (11):1415-22

Published Abstracts

1-Molecular profiling of colorectal cancer: Does it apply clinically?
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
British Journal of Surgery (2011). 98 (S2):11

2-Micrornas are differentially expressed in the circulation of colorectal cancer patients compared to healthy controls
M Nugent, KH Chang, **EAH Kheirelseid**, N Miller, OJ McAnena, M. Regan, MJ Kerin
British Journal of Surgery (2011). 98 (S2):12

3- Molecular Profiling of Colorectal Cancer
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
Irish Journal of Medical Science (2010).179(S9):360

4- Circulating MiRNAs in Colorectal Cancer
M Nugent, **EAH Kheirelseid**, HM Heneghan, KH Chang, N Miller, MJ Kerin
Irish Journal of Medical Science (2010).179(S9): 345

5- ANN Analysis to Identify MiRNA Expression Patterns in Colorectal Cancer
K Chang, N Miller, **E Kheirelseid**, G Ball, MJ Smith, M Regan, OJ McAnena, MJ. Kerin
Annals of Surgical Oncology (2010).17(1):S28

6- Clinical applications of molecular profiling of colorectal cancer
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
European Journal of Cancer (2010) .Sup.8 (5):197

7- Liver-intestinal cadherin (*CDH17*) is down-regulated in colorectal cancer and associated with tumour progression
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
British Journal of surgery (2010). 97(S6): 27-28

8- Identification of dysregulated oncomir expression and association with clinicopathological data in stage II colorectal cancer
KH Chang, N Miller, **EAH Kheirelseid**, G Ball, MJ Smith, M Regan, OJ McAnena, MJ Kerin
Irish Journal of Medical Science (2009).178(S7):250

9- *CXCL12* and *CXCR4* genes: markers of tumour status and disease progression in colorectal cancer
EAH Kheirelseid, N Miller, KH Chang, RM Dwyer, OJ McAnena, M Regan, MJ Kerin
Irish Journal of Medical Science (2009).178(S7): 231

10- Down-regulation of *PDCD4* induced by *miRNA-21* correlates with tumour aggressiveness in colorectal cancer
Chang KH; **Kheirelseid, EAK**; Miller N, Smith MJ, Regan M, McAnena OJ, Kerin MJ
Colorectal Disease (2009). 11(1):2

11- Association of dysregulated chemokine expression with tumour status and progression in colorectal cancer
EAH Kheirelseid, KH Chang, N Miller, RM Dwyer, O McAnena, M Regan, MJ Kerin
British Journal of surgery (2009). 96(2): 14-15

12- Down-regulation of *PDCD4* induced by *miRNA-21* correlates with tumour aggressiveness in colorectal cancer
KH Chang, **EAK Kheirelseid**, N Miller, MJ Smith, M Regan, OJ McAnena, MJ Kerin
British Journal of surgery (2009). 96(2): 1

Presentations to Learned Societies

2011

1- Molecular Profiling of Colorectal Cancer: Does it Apply Clinically?
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
Society of Academic and Research Surgery (SARS) meeting, Dublin, Ireland

2- MiRNAs are differentially Expressed in the Circulation of Colorectal cancer Patients Compared to Healthy Controls
M Nugent, KH Chang, **EAH Kheirelseid**, N Miller, OJ McAnena, M Regan, MJ Kerin
Society of Academic and Research Surgery (SARS) meeting, Dublin, Ireland

2010

3- Molecular Profiling of Colorectal Cancer
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
The 35th Sir Peter Freyer Surgical Symposium, NUI, Galway, Ireland

4- Circulating MiRNAs in Colorectal Cancer
M Nugent, **EAH Kheirelseid**, HM Heneghan, KH Chang, N Miller, MJ Kerin
The 35th Sir Peter Freyer Surgical Symposium, NUI, Galway, Ireland

5- ANN Analysis to Identify MiRNA Expression Patterns in Colorectal Cancer
K Chang, N Miller, **E Kheirelseid**, G Ball, MJ Smith, M Regan, OJ McAnena, MJ. Kerin
Society of Surgical Oncology (SSO) 63rd Annual Cancer Symposium, St.Louis, MO, USA.

6- Clinical applications of molecular profiling of colorectal cancer
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin
21st Meeting of the European Association for Cancer Research (EACR), Oslo, Norway.

7- Liver-intestinal cadherin (CDH17) is down-regulated in colorectal cancer and associated with tumour progression
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin.
Society of Academic and Research Surgery (SARS) meeting, London, UK

8- Circulating miRNA in colorectal cancer
M Nugent, **EAH Kheirelseid**, KH Chang, N Miller, MJ Kerin..
Society of Academic and Research Surgery (SARS) meeting, London, UK

9- Analysis of gene expression patterns as markers of disease classification in colorectal cancer
EAH Kheirelseid, N Miller, KH Chang, OJ McAnena, M Regan, MJ Kerin.
Gastrointestinal cancer symposium, Orlando, Florida, USA

2009

10- Association of dysregulated chemokine expression with tumour status and progression in colorectal cancer
EAH Kheirelseid, KH Chang, N Miller, RM Dwyer, O McAnena, M Regan, MJ Kerin.
Society of Academic and Research Surgery (SARS) meeting, Bristol, UK

11- Down-regulation of *PDCD4* induced by *miRNA-21* correlates with tumour aggressiveness in colorectal cancer
KH Chang, **EAK Kheirelseid**, N Miller, MJ Smith, M Regan, OJ McAnena, MJ Kerin.
Society of Academic and Research Surgery meeting (SARS), Bristol, UK

12- Down-regulation of *PDCD4* induced by *miRNA-21* correlates with tumour aggressiveness in colorectal cancer
KH Chang, **EAK Kheirelseid**, N Miller, MJ Smith, M Regan, OJ McAnena, MJ Kerin.
Association of Coloproctology of Great Britain and Ireland (ACGBI) Annual Meeting 2009, Harrogate, UK

13- *CXCL12* and *CXCR4* genes: markers of tumour status and disease progression in colorectal cancer
EAH Kheirelseid, N Miller, KH Chang, RM Dwyer, OJ McAnena, M Regan, MJ Kerin.
The 34th Sir Peter Freyer Surgical Symposium, NUI, Galway, Ireland

14- Identification of dysregulated oncomir expression and association with clinicopathological data in stage II colorectal cancer.
KH Chang, N Miller, **EAH Kheirelseid**, G Ball, MJ Smith, M Regan, OJ McAnena, MJ Kerin.
The 34th Sir Peter Freyer Surgical Symposium, NUI, Galway, Ireland

Grants and Awards related to this research

National Breast Cancer Research Institute (NBCRI)	2008
National Breast Cancer Research Institute (NBCRI)	2009
National Breast Cancer Research Institute (NBCRI)	2010

ABSTRACT

Despite developments in diagnosis and treatment, 20% of colorectal cancer (CRC) patients present with metastatic disease and 30% of cases recur after curative surgery. Furthermore, the molecular factors involved in prognosis and response to therapy in CRC are poorly understood.

The objectives of this study were to examine the expression patterns of candidate miRNAs and mRNAs and proteins in CRC in order to identify molecular biomarkers for disease classification and prognostication.

Expression patterns of miRNA and mRNA species were determined in FFPE and fresh colorectal tissues by microarray analysis and real-time quantitative PCR. Protein expression was determined by immunohistochemistry. Statistical analysis and correlation with clinicopathological data was performed using SPSS software.

MiRNA signatures predictive of response of rectal cancer to neoadjuvant therapy were determined. Expression levels of *CXCL12* (p=0.000), *CDH17* (P=0.026), *MUC2* (p=0.000), *L-FABP* (p=0.000) and *PDCD4* (p=0.000) were down regulated and *IL8* (p=0.000) was upregulated in tumours compared to normal colorectal tissues. Moreover, significant associations of gene expression levels and clinicopathological variables such as tumour size, grade, invasion and lymph node status were identified. Of further interest, significant miRNA/mRNA correlations were also determined in this study.

A comprehensive list of biomarkers with highly differential expression patterns in colorectal cancer that could serve as molecular markers to complement existing histopathological factors in diagnosis, follow up and therapeutic strategies for individualised care of patients. Furthermore, the significant miRNA/mRNA correlations highlight important novel mechanisms in CRC initiation and progression.

Chapter 1:
Introduction

1.1 Colorectal cancer

1.1.1 Epidemiology

Colorectal carcinoma (CRC) is one of the most common types of cancer worldwide with increasing incidence especially in developed countries [1]. Despite several advances in diagnosis and treatment, this disease remains a threat to life for a large number of people and approximately 20% of patients present with metastatic disease, and 30% of colorectal cancers recur [2]. In general, colorectal carcinoma is classified into three categories, based on increasing hereditary influence and cancer risk [3]. Sporadic CRC accounts for approximately 60% of patients and comprises patients with no notable family history and, by definition, with no identifiable inherited gene mutation that accelerates cancer development. Familial CRC accounts for approximately 30% of cases and refers to patients who have at least one blood relative with CRC or an adenoma, but with no specific germline mutation or clear pattern of inheritance. True hereditary CRC syndromes, accounting for approximately 10% of cases, originate from inheritance of single gene mutations in highly penetrant cancer susceptibility genes. Although the latter group of cancers occurs with the lowest frequency, due to the clear patterns of inheritance and identification of key pathogenic genes, it has helped to elucidate the molecular mechanisms of carcinogenesis applicable to sporadic CRC.

1.1.2 Pathology

From initial diagnosis through to definitive treatment, pathological evaluation plays a central role in the care of patients with colorectal cancer. Pathological stage of disease is widely recognised as the most accurate predictor of survival and is used to determine the appropriate treatment. Many other pathological factors have been shown to have prognostic significance that are independent of stage, and they may help to further sub-stratify tumours.

Histological types:

For consistency and uniformity in pathological reporting, the histological classification of CRC proposed by the World health Organisation (WHO) [4] is internationally accepted (table 1):

Adenocarcinoma:

Adenocarcinoma is the most common tumour type (95%). Most are moderately differentiated and lack specific histological features, although colorectal tumours tend to show cribriform patterns with central necrosis, a feature that is useful if a metastatic tumour is encountered when no colorectal primary has been diagnosed. Dysplasia in adjacent mucosa may be seen, but frequently the invasive tumour obliterates any pre-existing polyp from which it may have arisen.

Mucinous adenocarcinoma:

This is a subtype of adenocarcinoma that secretes extracellular mucin. At least 50% of the tumour must be mucinous in order to make this diagnosis. Mucinous adenocarcinomas are associated with microsatellite instability. Mucinous change may also be seen in ordinary adenocarcinomas treated with neoadjuvant chemoradiotherapy. Whether or not mucinous adenocarcinomas have a better prognosis is uncertain [5].

Medullary carcinoma

This is an important subtype of colorectal cancer, added to the World Health Organisation classification in 2000. It has a characteristic phenotype - right-sided tumours with sheets of cells and numerous tumour-infiltrating lymphocytes on microscopy [6]. This phenotype is associated with the Lynch cancer family syndrome (hereditary non-polyposis colorectal cancer). These colorectal tumours show a loss in expression of DNA mismatch repair proteins such as MSH1 (60% of cases) or MLH2 (30%).

Other tumours

Two specific tumours with a poor prognosis are signet ring cell and small cell carcinoma. Signet ring cell carcinoma is composed of at least 50% cells with intracytoplasmic mucin, resembling gastric signet ring cell tumours. Small cell carcinoma is a poorly differentiated neuroendocrine carcinoma. Occasionally, tumours from other sites involve the colorectum, and the pathologist should be aware, in particular, of direct extension into the rectum of prostate and bladder tumours. In most cases, morphology will distinguish these tumours from a

primary colorectal neoplasm, although in some immunohistochemical stains may be necessary to identify a tumour of non-colorectal origin

Table 1.1: **World Health Organisation (WHO) classification of CRC.**

Adenocarcinoma
Mucinous Adenocarcinoma (>50% mucinous)
Signet-ring-cell carcinoma (> 50% signet-ring cells)
Squamous cell carcinoma
Adenosquamous carcinoma
Small-cell (Oat-cell) carcinoma
Medullary carcinoma
Undifferentiated carcinoma
Other (e.g., papillary carcinoma)

Tumour grade:

Since the Broder's [7] and Duke's [8, 9] classification schemes were reported, the representative criterion of tumour grade employed for colorectal cancer has been the degree of tumour differentiation, as gauged primarily by architectural features. Tumour grade is defined based on the tumour differentiation in the TNM classification i.e., grade 1 is defined as well-differentiated, grade 2 as moderately differentiated, grade 3 as poorly differentiated, and grade 4 as undifferentiated. In the World Health Organization (WHO) classification, tumour grade is assessed based on the least differentiated component, with both well- and moderately differentiated adenocarcinomas being considered low-grade, and both poorly differentiated adenocarcinomas and undifferentiated carcinomas as high-grade. Poorly differentiated tumours can be identified by the absence of tubular formation and is an independent prognostic factor as it increases the risk of lymphatic spread from early stage tumours [10]. Although the relationship between histological grading based on tumour differentiation and disease prognosis has been well-recognized [11, 12], the existing grading systems have been criticized regarding the difficulty of making objective judgments. There are two causes of this problem. First, it is difficult to clarify the distinctions among

individual categories because tumour differentiation is a continuum parameter and an apparent break does not exist, especially between well-differentiated and moderately differentiated adenocarcinoma. Second, the extent of the component that examiners judge to be the least differentiated component has not been standardized.

Tumour staging:

The colon and rectum are unique among organs in that invasion of the lamina propria (that is the part of the mucosa surrounding the colorectal crypts) is considered to be *in situ* disease [13]. Thus, invasion of the submucosa is required to make a diagnosis of invasive carcinoma. The rationale for this is that because the colorectal lamina propria lacks lymphatics, tumours that are limited to the lamina propria have no means by which to spread. This is supported by evidence of a lack of malignant potential for such tumours [14].

Accurate and consistent pathological staging of colorectal cancer is vital to correct management. The central factor in T staging is the extent of invasion of the tumour through the bowel wall and it is still the most accurate predictor of prognosis in colorectal cancer patients. Table 1 summarises the two staging systems currently in use in Ireland and indicates the relationship between them.

Historically, the Duke's system has been valuable in clearly identifying patients who would benefit from postoperative chemotherapy (Duke's C). It has always been apparent that the Duke's B category is heterogeneous and includes patients who would also benefit from chemotherapy, especially with the advent of less toxic drug regimens. For these patients, the TNM staging system has advantages, as it identifies pT4 cases with a higher risk of local recurrence.

Table 1.2: **Staging of CRC and disease prognosis.**

1- Duke's' classification:

A	Tumour confined to the intestinal wall
B	Tumour invading through the intestinal wall
C	Lymph nodes involvement
D	With distant metastasis

2- AJCC and TNM classifications

AJCC stage	TNM stage	TNM stage criteria for colorectal cancer
		Tx: Primary tumour cannot be assessed
		T0: No evidence of primary tumour
Stage 0	Tis N0 M0	Tis: Tumour confined to mucosa; cancer-*in-situ*
Stage I	T1 N0 M0	T1: Tumour invades submucosa
Stage I	T2 N0 M0	T2: Tumour invades muscularis propria
Stage II-A	T3 N0 M0	T3: Tumour invades subserosa or beyond (without other organs involved)
Stage II-B	T4 N0 M0	T4: Tumour invades adjacent organs or perforates the visceral peritoneum
Stage III-A	T1-2 N1 M0	N1: Metastasis to 1 to 3 regional lymph nodes. T1 or T2.
Stage III-B	T3-4 N1 M0	N1: Metastasis to 1 to 3 regional lymph nodes. T3 or T4.
Stage III-C	any T, N2 M0	N2: Metastasis to 4 or more regional lymph nodes. Any T.
Stage IV	any T, any N, M1	M1: Distant metastases present. Any T, any N.

3- Correlation between CRC stage and 5-year survival:

Stage 0 (Tis, T1; N0;M0)	**>90%**
Stage I (T2;N0;M0)	**80-85%**
Stage II (T3;N0;M0)	**70-75%**
Stage III (T2;N1-3;M0)	**70-75%**
Stage III (T3;N1-3;M0)	**50-65%**
Stage III (T4;N1-3;M0)	**25-45%**
Stage VI (M1)	**<3%**

Tumour Markers:

Carcinoembryonic antigen (CEA) is a glycoprotein involved in cell adhesion. It is normally produced during fetal development, but the production of CEA stops before birth. Therefore, it is not usually present in the blood of healthy adults. It was found that serum from individuals with colorectal carcinoma, gastric carcinoma, pancreatic carcinoma, lung carcinoma and breast carcinoma, as well as individuals with medullary thyroid carcinoma, had higher levels of CEA than healthy individuals. CEA levels may also be raised in some non-neoplastic conditions like ulcerative colitis, pancreatitis, cirrhosis, COPD, Crohn's disease as well as in smokers. CEA measurement is mainly used as a tumor marker to identify recurrences after surgical resection, or localize cancer spread though dosage of biological fluids. The CEA blood test is not reliable for diagnosing cancer or as a screening test for early detection of cancer as most types of cancer do not produce a high CEA. Elevated CEA levels should return to normal after successful surgical resection or within 6 weeks of starting treatment if cancer treatment is successful. CEA and related genes make up the CEA family belonging to the immunoglobulin superfamily. In humans, the carcinoembryonic antigen family consists of 29 genes, 18 of which are normally expressed [15].

Figure 1.1: CEA as a marker of CRC.

CEA was first identified in 1965 by Phil Gold and Samuel O Freedman in human colon cancer tissue extracts [16].

DEMONSTRATION OF TUMOR-SPECIFIC ANTIGENS IN HUMAN COLONIC CARCINOMATA BY IMMUNOLOGICAL TOLERANCE AND ABSORPTION TECHNIQUES*

BY PHIL GOLD,‡ M.D., AND SAMUEL O. FREEDMAN, M.D.

(*From the McGill University Medical Clinic, Montreal General Hospital, and the Department of Physiology, McGill University, Montreal, Canada*)

PLATES 35 TO 39

(Received for publication, November 16, 1964)

Numerous attempts have been made by previous workers to demonstrate the presence of tumor-specific antibodies in sera obtained from animals immunized with preparations of human cancers (1–8). Such demonstrations, if consistently reproducible, would indicate the existence in human cancer tissue of unique homologous antigens not present in normal tissue, and might thus lead to a better understanding of the nature of the neoplastic process.

One of the techniques frequently employed in this type of investigation is the absorption of antitumor antiserum with normal tissue in order to remove antinormal components of the antiserum (1, 9, 10). Any residual antibody activity in the absorbed antiserum which is directed against tumor material is then considered to be tumor-specific. However, in interpreting the results of absorption experiments, little consideration has been given to the possibility that tumor-specific antibodies may have been removed or inactivated by normal tissue components similar to, but not identical with, the tumor antigens which initially stimulated the antibody production.

Another, more recent, technique for the demonstration of tumor-specific antibodies is the utilization of animals rendered immunologically tolerant to normal tissues during neonatal life (11). These tolerant animals are then immunized with tumor preparations of the same donor species. In those cases where adequate suppression of the immune response to normal tissue components has been achieved, immunization of the tolerant animals with tumor preparations during adult life has led to the development of antibodies apparently specific for the tumor (12–21). Nevertheless, the reported studies of human cancer employing acquired immunological tolerance (17–21) have left certain points unclear. In particular, the greatest cause for possible misinterpretation of results has been the failure to use the same donors as the source of normal and tumor material. Instead, the normal tissue has come from non-cancerous individuals

* This work was supported by a research grant from the Cancer Research Society, Montreal, Canada.

‡ Medical Research Fellow, Medical Research Council of Canada.

439

1.1.3 Management and role of Neoadjuvant CRT

By 1940, pathological analysis of rectal cancer resection specimens had identified penetration of the primary tumour through the bowel wall and involved lymph nodes as factors associated with poor outcome [8]. In 1954, Astler and Coller [17] confirmed the prognostic significance of direct cancer extension out side the

bowel wall. In 1970s, areas of failure found at re-operation following an initial curative resection for CRC were investigated, and the results showed that survival and disease recurrence rates are significantly related to the degree of bowel wall penetration and the extent of nodal disease [18]. This early work paved the way for the identification of those patients with high-risk disease.

Before the adoption of total mesenteric excision (TME) [19, 20], surgery alone for transmural or node positive rectal cancer was associated with local recurrence rates of up to 50% [21, 22]. This provided the rationale for exploration of management plans to improve outcomes following resection. The first trial was conducted by the Gastrointestinal Tumour Study Group, which randomised patients to surgery alone vs. chemotherapy vs. pelvic radiation vs. chemoradiation [23]. The arm that combined chemotherapy and radiotherapy showed a significant improvement in local control and survival [24]. Following this, investigators at Mayo Clinic/North Central Cancer Treatment group explored the postoperative radiotherapy alone vs. postoperative chemoradiation and found a significant reduction in local recurrence and cancer-related deaths in the chemoradiation group. Both trials set new standards for the postoperative management of high-risk rectal cancer. Once this new standard of care was established, the ongoing studies sought to determine the best regimen [25-27].

While optimising the treatment regimen and rationale for postoperative adjuvant therapy, researchers were also questioning whether preoperative therapy would be even more beneficial. Many reasons were proposed to demonstrate why treatment in the preoperative settings would be more efficacious [28]. The advantages of neoadjuvant therapy utilizing radiation are thought to be due to improved responsiveness of tissue not rendered hypoxic by previous surgery. Theoretically, ionizing radiation is more effective in irradiation of virgin tissue due to the increased oxygen tension in this tissue. Accordingly, preoperative radiation and chemotherapy are more effective in producing tumour necrosis in the non-disturbed pre-surgical tumour bed and cancer cells of the tumour periphery compared to the hypoxic post-surgical bed. Several other advantages with neoadjuvant therapy include less radiation-induced small bowel injury in the

pelvis, which has not been repaired by previous surgery, and the ability to excise the irradiated rectal segment and perform an anastomosis to healthy, non-irradiated colon, resulting in improved postoperative function [29]. In addition, studies have shown chemoradiation therapy, in the preoperative setting, results in less acute grades 3 and 4 toxic side effects and long-term toxic effects compared to giving it postoperatively [30]. Not surprisingly, there is less patient compliance with chemotherapy regimens provided in the postoperative period compared to giving it preoperatively [30, 31]. Taken together, the amalgamation of these modern regimens, including improved imaging, better chemotherapy, and more accurate and focused radiation, have resulted in an increased frequency of tumour down-staging, a higher likelihood of complete clinical and pathologic responses, and decreased local recurrence rates in stage II and III rectal cancer [30]. In addition, the utilization of neoadjuvant therapy in the management of stage IV disease has shown potential for prolonged survival.

In the 1990s, several institutions began evaluating the integration of preoperative radiotherapy approach. In Europe, investigators focused on the delivery of a short course of higher-dose radiation therapy alone followed in 1 week by resection. The Swedish Rectal trial reported improvement in survival adopting such an approach in 1997 [32]. This study randomised patients with respectable cancer to surgery alone or to surgery following a 1-week course of pelvic radiotherapy delivering 25 Gy in 5 daily fractions and their results showed that both local recurrence and 5-year survival were significantly improved. Moreover, the Dutch TME trial reported in 2001 [33] showed a higher local recurrence rate with TME alone without preoperative radiotherapy. In both Swedish and Dutch trials the interval from the end of pelvic radiation to surgical intervention was 1 week. Lyon R90-01 [34] studied the influence of this interval on down-staging and sphincter preservation and their results demonstrated that a longer interval between completion of radiotherapy and surgery was associated with increased tumour down-staging (26% vs. 10.3%, $p = 0.005$) and clinical tumour response (71.7% vs. 53.1%, $p = 0.007$). However; no significant differences were identified regarding morbidity, local recurrence or sphincter preservation.

Data from the Memorial Sloan-Kettering Cancer Center and the MD Anderson cancer Center supported the benefits of combining a total dose of 50.4 Gy of pelvic radiotherapy fractionated over 5.5 weeks in conjunction with concurrent chemotherapy [35]. Results from these series advocated for an improvement in sphincter preservation rates. In addition, patients with low-lying T2 lesions who would otherwise be offered abdomino-perineal resection (APR) were shown to benefit from such therapy [28]. Moreover; the German Rectal cancer group [36] confirmed the efficiency of a preoperative combined modality approach over the traditional strategy of providing subsequent postoperative adjuvant therapy. Additional trials have now shown that with longer course preoperative CRT significantly improves local control, tumour down-staging and down-sizing compared to radiotherapy alone [31, 37-39].

Quantification of tumour response to neoadjuvant therapy:

Assessment of response after preoperative CRT is essential in detecting patients who obtained a complete pathological response and can therefore be considered for a less aggressive surgical approach.

- *Pathological assessment of tumour response:*

Pathological complete response rates of 10-23% following neoadjuvant therapy have been reported. Although conflicting data exist, this suggests that good outcomes can be expected for patients with pathologic complete or near-complete response. Neoadjuvant CRT leads to characteristic histopathological changes in colorectal cancer. Grossly the tumour may be difficult to see, with in some cases no gross tumour visible in the mucosa. An area of scarring may be present in the bowel wall, or in surrounding fat, indicating treated tumour. To ensure consistency in reporting a complete response, in accordance with the protocol used in the CORE trial [40], the pathologist should extensively sample any areas of fibrosis seen in order to find any residual tumour.

Microscopically, these tumours display variable reduction in the volume of malignant cells, and an increase in the amount of stroma. The tumour cells may show phenotypic changes, such as mucinous metaplasia; the stroma may show

fibrosis, atypical fibroblasts or calcification. The degree of fibrosis correlates with outcome - recently it has been shown to be prognostic in R0 cases [41] and can be quantified with the tumour regression grade. Simplifications of the original five grades have been proposed based on the inter-observer variation when using five categories [42, 43]. Table 1.3 describes the two most commonly used tumour regression grading systems.

Table 1.3: **Tumour regression grades**

Mandard tumour regression grade [44, 45]	**Wheeler rectal cancer regression grade** [42]
1. No residual tumour, only fibrosis **2.** Rare residual tumour cells scattered throughout fibrosis **3.** Residual tumour cells outgrown by fibrosis **4.** Residual tumour cells outgrowing fibrosis **5.** No tumour regression	1. **Good response**: sterilisation or only microscopic foci of adenocarcinoma remaining, with marked fibrosis 2. **Moderate response**: marked fibrosis but macroscopic disease present 3. **Poor response**: little or no fibrosis, with abundant macroscopic disease

- *Clinical assessment of tumour response:*

This is where the most accurate T and N stage before and after treatment determined clinically e.g. by magnetic resonance imaging (MRI) or trans-rectal ultrasound (TRUS), is compared with the pathological T- and N-stage in the resected specimen [46, 47]. This is a commonly used means of assessing response, but the accuracy of this technique may be flawed by limitations in these imaging modalities.

Several studies have examined the accuracy of different imaging techniques in assessing rectal cancer response and lymph node involvement after preoperative CRT. The overall accuracy of endorectal ultrasound (EUS) ranges from 62-92% for initial assessment of T-stage, compared to 66-88% for initial assessment of N-stage. Following CRT however, it accurately identifies only 10 of 16 (63%)

patients with pathological complete response [48]. Moreover, EUS is far more likely to accurately stage non-responders than good responders (82% vs. 29%) [49]. The limitations of EUS following preoperative CRT are probably attributable, in part, to its inability to differentiate between tumour and radiation-induced inflammation [48]. Other imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) play a role in initial staging of rectal cancer patients. MRI performed with endorectal coil seems to be the most useful technique, with sensitivity and specificity similar to that of EUS in assessing wall penetration, and comparatively greater accuracy in assessing nodal involvement [50]. The accuracy of MRI declines in terms of response quantification following CRT, mostly due to overstating [51, 52].

Due to the limited ability of conventional imaging modalities to differentiate between scar, inflammation or fibrosis and tumour after CRT, new functional imaging techniques, such as fluorodeoxyglucose-positron emission tomography (FDG-PET), which is capable of assessing tumour viability and metabolism, have been examined. A number of studies have reported encouraging data on FDG-PET and its role in predicting response to neoadjuvant CRT [53-56]. Comparative studies have reported that FDG-PET is superior to EUS, CT and MRI in assessing the response after neoadjuvant CRT [53, 56]. In these studies, standardised uptake values (SUV) of tumour FDG were measured 2-4 weeks after completion of neoadjuvant therapy and compared to PDG-PET scans done before preoperative CRT. PDG-PET achieved a sensitivity of 100% and specificity ranging from 60-86% in predicting histopathological tumour response.

Pathological and molecular predictors of response to neoadjuvant CRT:

It is not known why such large differences in rectal cancer response to neoadjuvant CRT occur between patients. In order to elucidate factors that may allow for response prediction, existing research has focussed primarily on histological and molecular assessment of pre-treatment tumour biopsy specimens.

Clinical and histological indices:

Thusfar, this has not been systematically analysed in any single study. It has been indirectly addressed by multivariate analysis in four studies assessing molecular response predictors. These studies all concluded that pre-treatment T stage, N stage, grade, differentiation, age and gender could not predict histological response to RCT [46, 57]. Whilst conventional factors may have no influence over tumour radiosensitivity, they may, however, influence rates of local recurrence. Myerson *et al*. identified that tumour location <5 cm from the anal verge, circumferential lesions, obstruction and tethered/fixed tumours were all independent risk factors for local recurrence [58].

Molecular factors:

Colorectal cancer associated molecules *P53[46, 47, 57, 59, 60], P21[59, 61-63], EGFR [64], BCL2* and *BAX [57, 60, 61, 65]*, and *COX2 [66, 67]* have been extensively investigated as predictors of response to neoadjuvant CRT. Other molecules studied include: markers of tumour hypoxia (VEGF) [61, 68], spontenious apoptosis [57, 65, 66] , mismatch repair proteins [61, 63, 69] and proliferation [57, 59, 60, 63]. Nevertheless, the majority of markers assessed have yielded disappointing results.

1.2 Molecular biology of CRC

1.2.1 Cancer genetics

Oncogene and tumour-suppressor gene mutations all operate similarly at the physiological level: they drive the neoplastic process by increasing tumour cell number through the stimulation of cell birth or the inhibition of cell death or cell-cycle arrest. The increase can be caused by activating genes that drive the cell cycle, by inhibiting normal apoptotic processes or by facilitating the provision of nutrients through enhanced angiogenesis. A third class of cancer genes, called stability genes, promotes tumourigenesis in a completely different way when mutated. This class includes the mismatch repair (MMR), nucleotide-excision repair (NER) and base-excision repair (BER) genes responsible for repairing subtle mistakes made during normal DNA replication or induced by exposure to mutagens. Other stability genes control processes involving large portions of

chromosomes, such as those responsible for mitotic recombination and chromosomal segregation (e.g., *BRCA1*, *BLM* and *ATM*). Stability genes keep genetic alterations to a minimum, and thus when they are inactivated, mutations in other genes occur at a higher rate [70]. All genes are potentially affected by the resultant increased rate of mutation, but only mutations in oncogenes and tumour-suppressor genes affect net cell growth and can thereby confer a selective growth advantage to the mutant cell. As with tumour-suppressor genes, both alleles of stability genes generally must be inactivated for a physiologic effect to result.

Mutations in these three classes of genes can occur in the germline, resulting in hereditary predispositions to cancer, or in single somatic cells, resulting in sporadic tumours. It is important to point out that a mutation is defined as any change in the sequence of the genome. These changes include those affecting single base pairs as well as those creating large or small deletions or insertions, amplifications or translocations. In the germline, the most common mutations are subtle (point mutations or small deletions or insertions), whereas all types of mutation can be found in tumour cells. In fact, cancers represent one of the few disease types in which somatic mutations occurring after birth are pathogenic. The first somatic mutation in an oncogene or tumour-suppressor gene that causes a clonal expansion initiates the neoplastic process [71]. Subsequent somatic mutations result in additional rounds of clonal expansion and thus in tumour progression [72]. Germline mutations of these genes cause cancer predisposition, not cancer *per se*. Such individuals therefore often develop multiple tumours that occur at an earlier age than in individuals whose cancer-gene mutations have all occurred somatically [73].

1.2.2 Genetic and epigenetic alterations in CRC

Colorectal cancer results from the progressive accumulation of genetic and epigenetic alterations that lead to the transformation of normal colonic epithelium to colon adenocarcinoma. From the analysis of the molecular genesis of colon cancer, four central tenets concerning the pathogenesis of cancer have been established. The first is that the genetic and epigenetic alterations that underlie colon cancer formation promote the cancer formation process because they

provide a clonal growth advantage to the cells that acquire them. The second tenet is that cancer emerges via a multi-step progression at both the molecular and the morphologic levels [74]. The third is that loss of genomic stability is a key molecular step in cancer formation [75]. The fourth is that hereditary cancer syndromes frequently correspond to germ line forms of key genetic defects whose somatic occurrences drive the emergence of sporadic colon cancers [76].

Genetic alterations:

Much progress has been made in understanding the molecular mechanism of CRC since 1990, when Fearon and Vogelstein proposed their genetic model for CRC tumourigenesis [74]. A progression from normal mucosa to adenoma to carcinoma was supported by the demonstration of accumulating mutations in genes of *APC, K-RAS, P53 and DCC*, all of which are thought to be of significance, but are not able successfully to account for all CRCs. The earliest identifiable lesion in colon-cancer formation is the aberrant crypt focus (ACF). The true neoplastic potential of this lesion is still undetermined, but it does appear that some of these lesions can progress to frank adenocarcinoma and harbor mutations in *K-RAS* or *APC*. In particular, dysplastic aberrant crypt foci frequently carry mutations in *APC* and appear to have the highest potential for progressing to colon cancer. Thus, alterations in *APC*, which result in overactivation of theWingless/Wnt signalling pathway, appear to initiate tumour formation in the colon. Subsequent alterations in other genes then play a role in tumour growth and the eventual acquisition of other malignant characteristics such as tissue invasiveness and the ability to metastasize.

APC:

The *Adenomatous polyposis coli (APC)* gene encodes a protein that possesses multiple functional domains that mediate oligomerization as well as binding to a variety of intracellular proteins including ß-catenin, γ-catenin, glycogen synthase kinase (GSK)-3ß, axin, tubulin, EB1, and hDLG [76]. Germline mutations in *APC* result in FAP or one of its variants, Gardner's syndrome, attenuated FAP, Turcott's syndrome, or the flat adenoma syndrome [77, 78]. In addition; studies have shown that APC is mutated in up to 70% of all sporadic colon

adenocarcinomas, which is a high *APC* mutation frequency unique to colorectal cancers [79, 80]. These mutations are present beginning in the earliest stages of colon-cancer formation and precede the other alterations observed during colon-cancer formation [81, 82]. One of the central tumour-promoting effects of these mutations results in over-activation of theWingless/Wnt signalling pathway, with the subsequent expression of genes that favor cell growth. APC mutations disrupt the association of APC with ß-catenin, resulting in excessive amounts of ß-catenin and overactivation of the Wnt signaling pathway. Consequently, genes that promote tumour formation are transcribed. The over-activation of the Wnt signalling pathway occurs because normally GSK-3ß forms a complex with APC, ß-catenin, and axin and phosphorylates these proteins. The phosphorylation of ß-catenin targets it for ubiquitin-mediated proteasomal degradation. Truncating APC mutations prevent this process from happening and cause an increase in the amount of cytoplasmic ß-catenin, which can then translocate to the nucleus and interact with other transcription factors

K-RAS:

Kirstein rat sarcoma (*K-RAS)* is a member of the *RAS* family of genes and present one of the most prominent proto-oncogenes in colon carcinogenesis. The *RAS* family genes encode highly conserved proteins that are involved in signal transduction. One major function of the RAS protein family is to couple growth factors to the Raf-mitogen-activated protein (MAP) kinase kinase-MAP kinase signal transduction pathway, which leads to the nuclear expression of early response genes [83]. *K-RAS* mutations have been found in 37%–41% of colon carcinomas and appear to occur relatively early in colon-cancer formation [84, 85]. Vogelstein *et al*. [82] found *K-RAS* mutations in 13% of small tubular adenomas, 42% of large adenomas, and 57% of adenomas that contained areas of invasive carcinoma. In fact, 58% of adenomas greater than 1 cm in size had *RAS* mutations, compared to 9% of adenomas less than 1 cm in size [82]. These results have been supported by other investigators who have found an incidence of approximately 40% in colon adenomas [84]. The *K-RAS* mutations appear to follow *APC* mutations and are associated with advanced adenomatous lesions. Evidence for this model comes from the observation that small adenomas with *APC* mutations carry *K-RAS* mutations in approximately 20% of the tumours,

whereas approximately 50% of more advanced adenomas have *K-RAS* mutations [81, 86]. Thus, alterations of *K-RAS* appear to promote colon-cancer formation early in the adenoma-carcinoma sequence by mediating adenoma growth. Of interest, however, they do not appear necessary for the malignant conversion of *adenomas* to adenocarcinomas.

P53:

Tumour protein-53 (*p53*) was initially identified as a protein forming a stable complex with the SV40 large T antigen and was originally suspected to be an oncogene [87]. Subsequent studies demonstrated that *P53* is a transcription factor with tumour suppressor activity, is located at chromosome 17p13.1, and is mutated in 50% of primary human tumours, including tumours of the gastrointestinal tract [88]. *P53* is currently believed to be a transcription factor that is involved in maintaining genomic stability through the control of cell cycle progression and apoptosis in response to genotoxic stress [88]. In colon cancers, *P53* mutations have not been observed in colon adenomas, but rather appear to be late events in the colon adenoma-carcinoma sequence that may mediate the transition from adenoma to carcinoma [82]. Furthermore, mutation of *P53* coupled with loss of heterozygosity (LOH) of the wild-type allele was found to coincide with the appearance of carcinoma in an adenoma, thus providing further evidence of its role in the transition to malignancy [89, 90]. The function of *P53* to recognize DNA damage and induce cell cycle arrest and DNA repair or apoptosis has led to *P53* being called the “guardian of the genome” [91]. Thus, *P53* normally acts as a tumour suppressor gene by inducing genes that can cause cell cycle arrest or apoptosis and also by inhibiting angiogenesis through the induction of *TSP1* [92]. Mutant *P53* can block these functions by forming oligomers with wild-type *TP53*, thereby causing diminished DNA-binding specificity [93].

DCC:

Since it was first discovered in a colorectal cancer study in 1990 [94], *DCC* (*Deleted in colorectal cancer*) has been the focus of a significant amount of research. *DCC* held a controversial place as a tumour suppressor gene for many

years, and is well known as an axon guidance receptor that responds to netrin-1 [95]. More recently DCC has been characterized as a dependence receptor, and theories have been put forward that have revived interest in *DCC*'s candidacy as a tumour suppressor gene, as it may be a ligand-dependent suppressor that is frequently epigenetically silenced. One of the most frequent genetic abnormalities that occur in advanced colorectal cancer is loss of heterozygosity (LOH) of *DCC* in region 18q21. *DCC* elimination is not believed to be a key genetic change in tumour formation, but one of many alterations that can promote existing tumour growth.

Epigenetic alterations:

The finding of aberrant *hMLH1* promoter methylation in sporadic MSI colon cancers dramatically illustrated the role of epigenetic changes as potential pathogenetic alterations in cancer [96-99]. The term DNA methylation refers to the methylation of cytosine residues (5-methylcytosine) at CpG sites found throughout the genome [100]. These epigenetic alterations are characteristically clustered in so-called CpG islands in gene promoter regions, and hypo and hypermethylation of these regions are related to activation and inhibition of transcription, respectively. This type of gene regulation is essential to cell differentiation as well as embryological development [101]. Furthermore, DNA methylation is closely related to the mechanism by which one copy of a gene is preferentially silenced according to parental origin, generally referred to as genomic imprinting [102]. Aberrant methylation of the cancer genome, and associated silencing of the genes whose promoters demonstrated such methylation, has been well described at multiple genetic loci [103, 104]. Reversion of the methylation using demethylating agents such as 5-deoxy-azacytidine frequently restores expression of these genes, demonstrating methylation in fact induces gene silencing. As inactivation of *hMLH1* plays an initiating role in the pathogenesis of MSI colon cancers, the finding of aberrant methylation of *hMLH1* in sporadic MSI colon cancers, and the restoration of hMLH1expression by demethylating the *hMLH1* promoter in cell lines derived from such cancers, strongly suggests that such aberrant methylation could be a cause rather than a consequence of colon carcinogenesis [97-99]. Moreover,

Grady *et al*. [105] provided additional evidence for the primary role of aberrant methylation in gastrointestinal carcinogenesis. They demonstrated that, loss of expression of E-cadherin (*CDH1*) in association with CpG methylation of the wild-type *CDH1* allele in tumours occurs in the setting of the cancer family syndrome Hereditary Diffuse Gastric Cancer. Epigenetic and genetic changes also appear to cooperate to promote cancer formation [106]. Moreover, the aberrant hypermethylation of 50 CpG dinucleotides that has been demonstrated to silence a variety of tumour suppressor genes including *CDH1, CDKN2A/p16, TSP1,* and *GSTP1* may be similarly pathogenic in the tumours in which these changes have been identified [97, 106-109].

1.2.3 Classification of CRC

Colorectal carcinoma (CRC) is generally classified into three categories, based on increasing hereditary influence and cancer risk [3]. Sporadic CRC (60%) comprises patients with no notable family history and, by definition, with no identifiable inherited gene mutation that accelerates cancer development. Familial CRC (30%) refers to patients who have at least one blood relative with CRC or an adenoma, but with no specific germline mutation or clear pattern of inheritance. Hereditary CRC syndromes (10%) which result from germline inheritance of mutations in highly penetrant cancer susceptibility genes. Although the last group is observed with the lowest frequency, however they have been instrumental in the elucidation of molecular mechanisms of carcinogenesis applicable to sporadic CRC.

Sporadic CRC:

Sporadic colorectal cancers arise at a median age of 70-75 years. Seventy percent arise in the left side of the colon and there are differences in the age, sex and regional distribution of both adenomas and carcinomas between both sides of the large bowel. Sporadic cancers caused by the development of a series of genetic abnormalities in tumour suppressor genes and oncogenes that give cells an evolutionary advantage over their neighbours.

Hereditary and familial CRC syndromes:

Hereditary non-polyposis colorectal cancer (Lynch Syndrome):

Hereditary non-polyposis colorectal cancer, also referred to as the Lynch syndrome, is the most common form of hereditary colorectal cancer. It is inherited in an autosomal dominant fashion its clinical consequences develop from germline mutations in mismatch repair (MMR) genes. The lack of functional MMR proteins leads to genomic instability and development of various cancers. Multiple generations are affected with colorectal cancer at an early age (mean, approximately 45 years) with a predominance of right-sided colorectal cancer (approximately 70 percent proximal to the splenic flexure). There is an excess of synchronous colorectal and metachronous colorectal cancer. In addition, there is an excess of extracolonic cancers, namely, carcinoma of the endometrium, ovary, stomach, small bowel, pancreas, hepatobiliary tract, brain, and upper uroepithelial tract [110, 111] As compared with sporadic colorectal cancer, tumours in hereditary non-polyposis colorectal cancer are more often poorly differentiated, with an excess of mucoid and signet-cell features, a Crohn's-like reaction, and the presence of infiltrating lymphocytes within the tumour [112-114].

Familial adenomatous polyposis

Familial adenomatous polyposis (FAP) is characterized by numerous (>100, usually several hundreds in fully developed cases) of adenomatous colorectal polyps. In general, less than 1% of all new CRC arise in FAP patients. FAP is an autosomal dominant hereditary cancer syndrome caused by a germline mutation in the APC gene (adenomatous polyposis coli). Because this syndrome may be associated fewer number of colonic polyps (''attenuated'' FAP), it may first present with extra-intestinal manifestations and because as many as 50% of FAP patients result from new germline mutations in APC gene, pathologists may be the first to suspect this hereditary condition. Gardner syndrome is characterized by epidermoid cysts, osteomas, dental anomalies and desmoid tumours (fibromatoses). Turcot syndrome is an association between colorectal polyposis and primary central nervous system (CNS) tumour (usually medulloblastomas) [115, 116]. Extra-gastrointestinal manifestations may be of importance for practicing pathologists in diagnosis of unsuspected FAP. Desmoid tumour

(fibromatosis) is rare in the general population, but it is commonly seen in FAP and it may be the first manifestation of disease. Patients with FAP typically develop in retroperitoneal tissues or in the abdominal wall following surgical trauma (abdominal desmoids), while fibromatoses unrelated to FAP are more common in extra-abdominal localizations [117]. Papillary carcinoma of the thyroid and its rare cribriform-morular variant may be associated with FAP, and this could lead to detection of unsuspected FAP [118, 119]. The risk of hepatoblastoma in children of patients with FAP is highly increased and new germline mutations can be identified in 10% of cases [120].

Hamartomatous polyposes syndromes:

The hamartomatous polyposis syndromes include Peutz-Jeghers syndrome, juvenile polyposis, Cronkhite-Canada, and Cowden disease/Bannayan-Riley-Ruvalcaba syndrome [121-123]. Hamartomatous polyposes syndromes are distinguished by their characteristic clinicopathologic and radiologic features. All of these syndromes are characterized by hamartomatous polyps and most of them are associated with increased risk of development of gastrointestinal and extraintestinal carcinomas [124].

- *Peutz-Jeghers Syndrome (PJS):*

Peutz-Jeghers syndrome (PJS) is characterized by mucocutanous pigmentation and GI hamartomas, which occur anywhere from stomach to anus. It was first described by Peutz in 1921 [125] and Jeghers in 1944 [126]. It is inherited in an autosomal dominant fashion with no sex predilection [127]. A prototypic PJS polyp is a hamartoma of the muscularis mucosae. Therefore, the core of the polyp consists of smooth muscle covered by lamina propria and mature glandular epithelium [128, 129] which gives rise to a characteristic arborising smooth muscle core of the polyp. Germ-line mutations in the serine/threonine kinase gene (STK11/LKB1) on chromosome 19p13.3 cause Peutz- Jeghers syndrome in about half of the affected families. Additional loss of the wild-type allele in hamartomas and adenocarcinomas suggests that STK11/LKB1 is a tumour suppressor gene [130].

- *Juvenile polyposis coli (JP):*

Juvenile polyposis (JP) coli is inherited in an autosomal dominant fashion at least in 30% of patients. Patients develop numerous hamartomatous colorectal polyps, which are characterized by dilated crypts [131]. The number of polyps is smaller than in FAP and the disease course is less malignant [132]. The diagnosis of juvenile polyposis syndrome is made when multiple (3–10) juvenile polyps are found in the gastrointestinal tract, even though there is still some variation in criteria used in diagnosis. Mutations of SMAD4/MADH4 gene were initially described and explain about 30% of cases [133]. Mutations in BMPR1A can also lead to juvenile polyposis in additional 30% of patients [134, 135].

Table 1.4: Pathologic and genetic features of hereditary CRC [136]

Syndrome	Preinvasive colon pathology	Characteristic CRC morphology	Common extracolonic pathology	Defining mutation
HNPCC	Tubulo-villous adenoma, (right-sided)	Mucinous, medullary, signet-ring and mixed types	Endometrial carcinoma; sebaceous skin tumours	MMR genes
FAP	Tubular adenoma, microadenoma	Adenocarcinoma NOS	Fibromatoses, hepatoblastoma	*APC*
MUTYH-associated polyposis (MAP)	Tubular adenoma	Adenocarcinoma NOS	Duodenal carcinoma	*MUTYH*
PJS	Hamartomatous polyp with smooth muscle core	Adenocarcinoma NOS	Esophagus, Stomach, small intestine and pancreas carcinomas; sex cord tumours	*STK11*
JP	Hamartomatous polyp with dilated crypts	Not specified	Pancreatic, gastric duodenal carcinomas	*SMAD4/BMPR1A*
Cowden/ BRR	Hamartomatous polyp	No increased risk for CRC	Breast, thyroid and uterus carcinomas	*PTEN*
Hereditary Mixed Polyposis Syndrome (HMPS)	Adenomatous, hyperplastic, serrated and mixed polyps	Not specified	No increased risk	Putative *HMPS/ CRAC 1* (15q)
Familial CRC type X	Undefined	Not specified	Undefined	Unknown, probably complex.

1.3 Genetic instability of CRC

Colorectal cancer is a heterogeneous disease that can develop through different genetic pathways. The most common is termed the chromosomal instability pathway and accounts for 70% to 85% of colorectal cancers [82]. These tumours are characterized by mutations in *APC*, *P53*, and *KRAS* and by frequent allelic loss at 18q [82, 137]. Aneuploidy, amplifications, and translocations are also common in these tumours. Familial adenomatous polyposis is the hereditary syndrome associated with these changes [114]. The microsatellite instability

(MSI) pathway, comprising the remaining 15% of colorectal cancers, is characterized by loss of proficiency of the DNA mismatch repair (MMR) system and MSI.

1.3.1 Microsatellite instability (MSI)

Microsatellites are repeated DNA sequences, usually 1 to 10 nucleotides long, present throughout the genome. Instability is mostly characterized by single base-pair insertions or deletions in these repeat loci, causing widespread genomic instability due to the failure of the cell's mismatch repair (MMR) mechanism. MSI occurs as a consequence of inactivation of the mutation mismatch repair system and is recognized by frame shift mutations in microsatellite repeats located throughout the genome. Inactivation of the MMR system due to germ line gene defects accounts for the colon cancer family syndrome, hereditary non-polyposis colon cancer syndrome (HNPCC). Somatic inactivation of the mismatch repair system additionally gives rise to approximately 15% of sporadic colon cancers. In either instance the resulting colon cancers display the phenotype of microsatellite instability. The demonstration of microsatellite unstable cancers is generally performed by assaying for alterations at microsatellite loci that are particularly frequently mutated in the setting of MMR inactivation. Since many colon cancers demonstrate frame shift mutations at a small percentage of microsatellite repeats, the designation of a colon tumour as showing microsatellite instability depends on the detection of at least two unstable loci out of five from a panel of loci that were selected by a National Cancer Institute consensus conference [138].

Genes that possess such microsatellite-like repeats in their coding regions appear to be the targets relevant to carcinogenesis. Indeed, frequently, many genes that possess microsatellite repeats are observed to be mutated in MSI colon cancers. The relationship between the microsatellite mutator pathway and other genetic alterations frequently found in colon cancer is only partially understood. Alteration of the Wnt/Wingless pathway can be observed in tumours irrespective of MSI [139]. Mutations in *APC* and *CTNNB1* can be found in 21% and 43% of MSI tumours, respectively [140, 141]. In addition, the incidence of *K-RAS* mutations appears to be as high as 22–31%, which is similar to the incidence

observed in microsatellite stable (MSS) colon cancers [142]. Mutations in *P53* do appear to be less frequent in MSI cancers than in MSS cancers. The mutation incidence in MSI colon cancers has been demonstrated to range between 0–40%, whereas the incidence in MSS tumours is between 31–67% [140, 142, 143].

The MSI tumour formation process has been termed the microsatellite mutator phenotype and is a pathway to tumour formation that is distinct from that seen in colon cancers that are microsatellite stable [144-146]. The most frequently targeted gene for mutation in this pathway is the TGF-ß receptor type II tumour suppressor gene (TGFBR2). Other less frequently targeted genes include the *ACVR2, BAX, RIZ, CDX2, SEC63, AIM2, MSH3* and *MSH6* [146-149]. *CTNNB1* mutations are also found in 25% of MSI colon cancers but are not found in MSS cancers.

The Lynch syndrome is caused mainly by germ-line mutations in the DNA mismatch repair genes and heterozygosity for a mutation results in susceptibility to the cancer. Lynch syndrome can be identified based on age at onset, previous medical history and the characteristics of family history that fulfil the Amsterdam criteria and Bethesda guidelines for the diagnosis of hereditary non-polyposis colorectal cancer (HNPCC) [150-153]. The early recognition of Lynch syndrome is essential to identify patients at high risk who will require intensive surveillance. Nevertheless; its diagnosis can be difficult to make due to incomplete family history information and lake of characteristic clinical phenotype. Although the Amsterdam criteria and Bethesda guidelines continue to be used widely, several studies have underscored the limitations of their accuracy in predicting the presence of MMR gene mutations [154, 155]. Therefore, new strategies for screening for and diagnosis of Lynch syndrome need to be investigated.

In addition to screening for Lynch syndrome, testing for MSI is important because of its possible prognostic and therapeutic implications. Cancers with high microsatellite instability (H-MSI) were reported to have a more favourable clinical out come than non-MSI tumours and the survival advantage conferred by the MSI phenotype is independent of tumour stage and other clinicopathological

variables [156-158]. Moreover, tumours with H-MSI are thought to be less responsive to 5-fluorouracil and other anticancer agents *in vitro* and *in vivo* [159-161].

1.4 Molecular profiling

1.4.1 Source of biological data

One of the major factors influencing the performance and accuracy of molecular profiling is the source and processing of patient samples. So far, reliable analysis is limited to fresh blood or fresh-frozen tissue samples. However, these samples may be unavailable from subsets of patients.

Routine histology processing uses formalin fixation to preserve the histological architecture of tissue specimens. Archival collections of formalin-fixed tissues, linked to clinical databases, provide a rich resource from which biological insights could be derived far more expeditiously than the prospective collection of frozen samples. In addition, any biomarker developed from formalin-fixed paraffin-embedded (FFPE) samples could be more readily translated into clinical practice. Unfortunately, RNA is degraded in tissues before, during, and after formalin fixation [162, 163] and can continue to deteriorate even during storage, leading to shortened fragments of RNA [164]. Interestingly, miRNAs appear to be better preserved, perhaps because of their intrinsically shorter lengths. Therefore, it should be possible to perform genome-wide screening for miRNAs using FFPE tissues [165-167] Indeed, this approach has been successful for colon and breast cancers; however, the technical robustness of these platforms has not been thoroughly investigated.

1.4.2 Techniques for molecular profiling

1.4.2.1 Real-time polymerase chain reaction (RQ-PCR)

Developed in 1983 by Kary Mullis [168], PCR is now a common and often indispensable technique used in medical and biological research labs for a variety of applications. It is a technique to amplify a single or a few copies of a piece of DNA across several orders of magnitude, generating thousands to millions of copies of a particular DNA sequence. The method relies on thermal cycling,

consisting of cycles of repeated heating and cooling of the reaction for DNA melting and enzymatic replication of the DNA.

1.4.2.2 Microarray analysis

A DNA microarray is a multiplex technology consists of an arrayed series of thousands of microscopic spots of DNA oligonucleotides, called features, each containing picomoles of a specific DNA sequence, known as probes (or reporters). These can be a short section of a gene or other DNA element that are used to hybridize a cDNA or cRNA sample (called target) under high-stringency conditions. Probe-target hybridization is usually detected and quantified by detection of fluorophore-, silver-, or chemiluminescence-labeled targets to determine relative abundance of nucleic acid sequences in the target. Since an array can contain tens of thousands of probes, a microarray experiment can accomplish many genetic tests in parallel. Therefore arrays have dramatically accelerated many types of investigation. In standard microarrays, the probes are attached via surface engineering to a solid surface by a covalent bond to a chemical matrix. The solid surface can be glass, plastic or a silicon chip commonly known as a *genome chip*, *DNA chip* or *gene array*. Other microarray platforms, such as Illumina, use microscopic beads, instead of the large solid support. DNA arrays are different from other types of microarray only in that they either measure DNA or use DNA as part of its detection system. DNA microarrays can be used to measure changes in expression levels, to detect single nucleotide polymorphisms (SNPs), or to genotype or resequence mutant genomes

The core principle behind microarrays is hybridization between two DNA strands, the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs. A high number of complementary base pairs in a nucleotide sequence means tighter non-covalent bonding between the two strands. After washing off of non-specific bonding sequences, only strongly paired strands will remain hybridized. So fluorescently labelled target sequences that bind to a probe sequence generate a signal that depends on the strength of the hybridization determined by the number of paired bases, the hybridization conditions (such as temperature), and washing after hybridization. Total strength of the signal, from a

spot (feature), depends upon the amount of target sample binding to the probes present on that spot. Microarrays use relative quantitation in which the intensity of a feature is compared to the intensity of the same feature under a different condition, and the identity of the feature is known by its position.

The development of high throughput technologies such as microarray profiling has necessitated the development of an entirely new field of bioinformatics. Optimisation of data analysis tools has been a matter of intense debate centering on issues such as normalization and appropriate interpretation of such large datasets. Many studies have employed a clustering approach in the analysis of microarray experiments. This involves grouping genes on the basis of their similarity in expression patterns across samples and tumour samples on the basis of their similarities in gene expression. Hierarchical clustering is one of the most commonly used bioinformatics methods in the analysis of microarrays. This requires two main steps that are repeated in order to find the genes that are most similar. In brief, the pair of genes that are most similar are identified and joined together then, the next most similar pair of genes is identified. This process continues until all of the genes are joined into one giant cluster. A dendrogram tree serves as a graphical representation of the data, with the lengths of branches representing the degree of relationship between genes or subjects

The disadvantage of this technique is that it does not involve input of any prior functional information. Moreover; conventional methods of analysing microarray experiments have limitations in their applicability to such highly dimensional data. The inherent “noise” (e.g. experimental error, sample and chip variability) can significantly interfere with the development of accurate predictive models and their performance has been compromised by their modelling of extraneous portions of the dataspace. For this reason, much attention has been paid to the development of novel data analysis methodologies to cope with data of this non-linear and highly dimensional nature.

1.4.3 Gene expression profiling in CRC

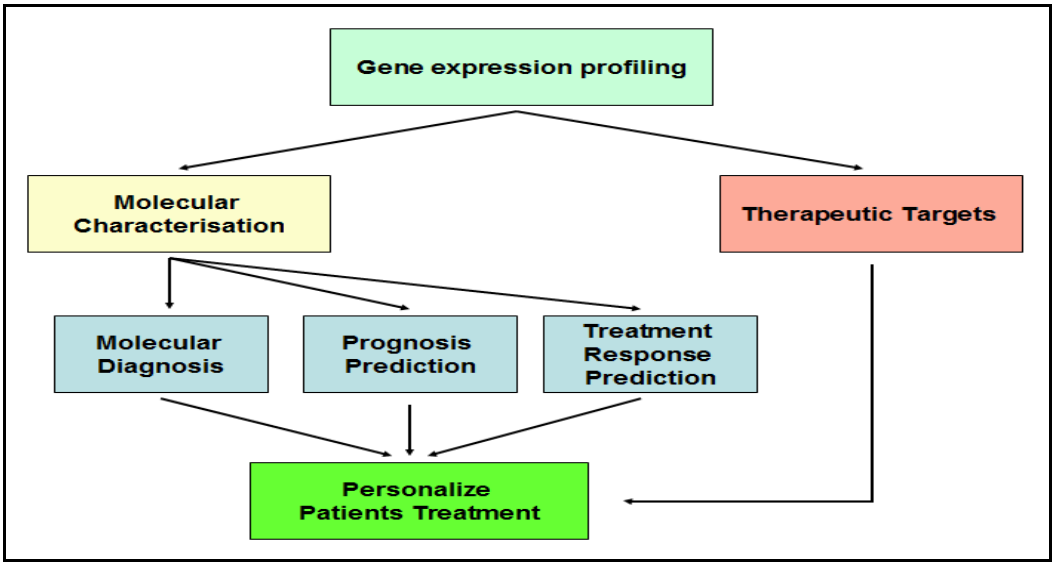
Molecular biology represents one of the most interesting topics in medical oncology, because it provides a global and detailed view on the molecular

changes involved in tumour progression, leading to a better understanding of the carcinogenesis process, to discovering new prognostic markers and novel therapeutic targets. Despite of clinical and pathological parameters are available for the classification and prognostic stratification of cancer, they may be inadequate in everyday practice due to the great biologic and genetic heterogeneity of this multiform disease.

Gene expression of breast cancer is the intensively studied among cancers and represents a model for gene expression profiling experiments of solid tumours. Perou et al. provided a distinctive molecular portrait of 65 breast carcinomas, from 42 different individuals, using complementary microarrays, suggesting that tumours could be classified into subtypes distinguished by pervasive differences in their gene expression patterns [169]. Subsequently they found a correlation between those subtypes and clinical outcome, suggesting that gene expression patterns of tumours have both a taxonomic and prognostic value [170].

CRC represents an interesting field of molecular profiling research for several reasons: CRC is considered a biological model of tumourigenesis, because clinical progression from adenoma to early stage carcinoma until advanced stage carcinoma seems to parallel distinctive molecular alterations [82]. In addition; traditional clinical and pathological parameters are not always sufficient to discriminate high risk from low risk CRC and validated molecular markers with prognostic value are still not available. The studies on molecular profiling in CRC have mainly focused on carcinogenesis process, disease prognosis prediction and therapeutic targets and response prediction.

Figure 1.2: Potential clinical application of gene expression profiling



1.4.3.1 Diagnostic and prognostic biomarkers

The application of gene expression profiling on carcinogenesis studies purposes to identify specific alterations on gene expression according to tumour development and to diagnose and classify tumours on the basis of molecular features. Studies comparing gene expression between normal mucosa, adenoma and carcinoma or between primary tumour and metastases, as well as between left-side and right-side tumours are performed, in order to discover distinctive genetic signatures belonging to each. Furthermore; studies on prognosis prediction aim to identify specific alterations to the gene expression profile that may be useful to discriminate high risk from low risk CRC, to provide a molecular stratification according to the clinical outcome and to predict the metastatic potential of the primary tumour.

Several studies were set to investigate the difference in gene expression levels between tumour and normal colorectal tissues. In 1999, Alon and colleagues reported a clustered data set of 2000 genes able to separate 22 normal and 40 tumour colon tissues with the highest minimal intensity across samples. Subsequently many studies reported other sets of genes that were differentially expressed between cancer and normal tissue and therefore potentially involved in the development of colorectal carcinogenesis [171-175]. In addition, some studies

reported significant differences in gene expression profile between adenoma and normal mucosa, suggesting that different mechanisms of development of these precancerous lesions may exist [176, 177]. Furthermore, and in order to clarify the molecular modifications underlying the development of metastases, some studies compared the gene expression profile of primary tumours with their corresponding metastases [178-182]. Agrawal et al. reported that among all genes associated with disease progress, osteopontin expression seemed to be the leading candidate [178]. Moreover; Yanagawa et al. [180] analysed genome-wide expression profiles of 10 primary CRCs and their corresponding liver metastasis and identified 40 genes whose expression was commonly up-regulated in metastatic lesions, and 7 that were commonly down-regulated. On the other hand; Watanabe et al. studied 89 CRC patients to identify a set of discriminating genes that can be used for characterization and prediction of lymph node metastasis and identified 73 genes in which expression was significantly different between patients with and without lymph node metastasis [183]. Using this gene set, they established a model to predict the presence of lymph node metastasis with an accuracy of 88.4%. In addotion, the controversial data on the benefit of adjuvant chemotherapy in stage II CRC led to the identification of molecular prognostic factors that may identify stage II CRC patients who develop disease recurrence and may benefit by adjuvant treatment. Wang et al. studied the gene expression profile in this set of patients and, using two supervised class prediction approaches of analysis, they identified a 23-genes signature that may predict recurrence in stage II CRC with 78% accuracy [184]

Some studies have also investigated differences in gene expression between CRC of the right side and left side, due to their epidemiological, morphological and pathogenetic diversity and found distinct profiles according to the anatomical stratification. Birkenkamp-Demtroder et al. [185] investigated the difference in gene expression between the caecum vs. sigmoid and rectosigmoid and identified 58 genes to be differentially expressed between the normal mucosa of caecum and the sigmoid and rectosigmoid.

1.4.3.2 Therapeutic targets and treatment response prediction

While gene expression profiling has been widely applied to CRC for diagnosis, classification and prognosis prediction based on molecular patterns of expression, its application to response prediction to medical treatment is still lacking reliable results due to few currently available studies [186-190]. In a panel of 30 colon carcinoma cell lines Mariadason et al. identified 420 genes correlated with response to 5-fluorouracil (5-FU) and involved in two main biological processes, DNA replication and repair and protein processing/targeting [188]. The predictive value of 50 genes best correlated with 5-FU response was subsequently validated using a leave one out cross validation approach and it was higher than the traditional markers, such as thymidylate synthase, thymidine phosphorylase, mismatch repair and p53 status. Furthermore they also found that 149 genes best-correlated with CPT-11-induced apoptosis significantly predicted response of colon cancer cell lines to this agent. In addition; Del Rio et al. analyzed gene expression profile of 21 primary advanced CRC tissues, in order to identify an expression pattern that could predict response to leucovorin, fluorouracil and irinotecan as first-line treatment: 14 genes were found expressed differently between responders and non responders and were able to predict treatment response with 95% accuracy [189]. In the same year Khambata-Ford et al. investigated gene expression pattern of metastatic biopsies of 80 advanced CRC patients treated with cetuximab to identify genes whose expression correlates with best clinical response [190]. They found that, among 629 genes expressed differently between 25 patients with disease control and 55 non responders, the top candidate markers based on lowest p value were epiregulin and amphiregulin, both ligands for epidermal growth factor receptor (EGFr), suggesting that these markers could select patients for cetuximab therapy.

Some studies evaluated the ability of gene expression profiling for predicting response of advanced rectal cancer (RC) to preoperative chemoradiotherapy [186, 191-193]. Ghadimi et al. analyzed gene expression signatures of biopsies from 30 locally advanced RC and found 54 genes differentially expressed between responders and non responders [186]. Kim et al. reported 261 genes differentiated between 20 partial response and 11 complete response patients affected by locally

advanced RC treated with preoperative chemoradiotherapy.In their study 95 genes predicted complete and partial response with an 84% accuracy [191]. Similarly another study identified a gene expression signature of 42 genes that was able to distinguish responder from non responder locally advanced RC patients with a 71% accuracy [192]. Recently; Spitzner et al. were able to identify a gene expression signature for chemoradiosensitivity of colorectal cancer cells [193]. They exposed 12 colorectal cancer cell lines to of 5-fluorouracil and radiation therapy. The differences in treatment sensitivity were then correlated with the pretherapeutic gene expression profiles of these cell lines. Their data have suggested a potential relevance of the insulin and Wnt signalling pathways for treatment response, and they also identified *STAT3, RASSF1, DOK3*, and *ERBB2* as potential therapeutic targets [193].

Although colorectal cancer (CRC) is still one of the leading causes of cancer related death, the introduction of new therapeutic options like oxaliplatin and irinotecan in addition to 5-fluorouracil, the standard therapeutic for CRC has increased the overall survival of affected patients from 10 to 18–24 months. Furthermore, the ‘‘biological” therapeutics cetuximab, an IgG1 chimeric monoclonal antibody against epidermal growth factor receptor (*EGFR*), and bevacizumab, a monoclonal antibody against vascular endothelial growth factor (*VEGF*), have augmented the course of the disease and brought in the new era of targeted therapy against cancer specific molecular pathways [194-197]. Although these biologicals have entered clinical routine due to their encouraging results, their effect has been shown to be limited due to adaptation or previously existing resistance of tumour cells. This has been clearly shown in the case of patients with mutations of *K-RAS*, which lead to resistance against cetuximab. Therefore, several new pathways are currently investigated for therapeutic targeting in CRC. These include WNT-signaling, downstream mediators of EGFR as the mitogen-activated protein kinase (*MAPK*) or the phosphatidylinositol 3-kinase (PI3K)-pathway, the hypoxia response system involving hypoxia inducible factor-1 (*HIF-1*), mechanisms of tumour development following chronic inflammation, and many others [198].

Table 1.5: Genes that are consistently represented in CRC literature.
Of many published studies on colorectal gene expression profiling, little correlation exists between validated candidate genes associating with disease status. Some candidate genes are consistently represented in the literature however, examples are shown below:

Gene	Function	Expression Level	references
FABP1	Lipid transport & metabolism	D	[171, 199-201]
CA2	Zinc metallo-enzyme	D	[171, 181, 202, 203]
IL8	Neutrophils activation and migration	U	[200, 204, 205]
GPX2	GIT protection	U	[171, 175, 206, 207]
ADH1A	Alcohol metabolism	U	[172, 208, 209]
COL1A2	Cell growth & maintenance	D	[172, 206, 210]
ITGA5	Cell adhesion	D	[210-212]
HSP90B1	Protein folding & degradation	U	[172, 174, 213]
PLAU	Haemostasis & cell migration	U	[200, 213, 214]
MMP1	Collagenase	D	[181, 200, 209]
MMP2	Gelatinase	D	[175, 181, 210]
COL5A2	Cell growth & maintenance	D	[206, 211]
COL4A2	Cell growth & maintenance	D	[206, 207, 210]
CDH17	Cell adhesion & transport	D	[181, 184, 211]
CXCL12	Immunity	U	[181, 211]
CDK6	Cell proliferation	U	[174, 211]
CDK8	Cell proliferation	U	[175, 205, 211]
MUC2	Immunity	U	[172, 175]
EPOR	Cell proliferation	U	[181, 212]
ATP6V0E	Cell transport	U	[181, 204, 215]
PDCD4	Tumour suppressor	U	[191, 203]
AXIN2	Signal transduction	U	[175, 200]
IGFBP7	IGF availability & function	D	[199, 210]

U = up-regulated, D = Down-regulated

1.4.4 miRNA expression profiling

MiRNAs have recently emerged as an exciting new class of disease biomarker with further potential as the next generation of cancer therapeutics. Although elucidating their mechanisms of action is still in its infancy, the discovery of miRNAs has uncovered an entirely new and exciting repertoire of molecular

factors upstream of gene expression, with great potential for new developments in current diagnostic and therapeutic strategies in the management of cancer patients. MiRNAs are small 19 to 22 nucleotide sequences of RNA found in both prokaryotes and eukaryotes that are intimately involved in cell differentiation, cell cycle progression, and apoptosis. MiRNAs have been demonstrated recently to potentially play a significant role in tumourigenesis. Additionally, miRNAs may be useful tools for characterizing specific cancers and for determining patient prognosis and response to therapy. The study of miRNA has been extended into many types of cancer, including leukemias, lung, breast, and colon cancer. The first description of miRNA appeared in 1993 by Lee et al., who proved that *lin-4* is involved in controlling the temporal progression of cell differentiation in *C. elegans* [216]. In 2000, another miRNA, *miR-7*, was found to share the control of developmental progression of nematodes [217]. Discoveries of other miRNAs that regulate apoptosis, proliferation, and differentiation in Drosophila, mice, and humans soon followed [218, 219]. Calin et al. [220] published the first study to link miRNAs to cancer in 2002. These authors demonstrated that *miR-15* and *miR-16* are located on chromosome 13 in a position where deletion of a putative tumour suppressor, known to be associated with greater than half of chronic lymphocytic leukemia cases, was identified. Researchers have proposed that specific miRNA expression patterns could help identify human solid tumours, suggest patient prognosis, and even represent a novel molecular target for cancer treatment.

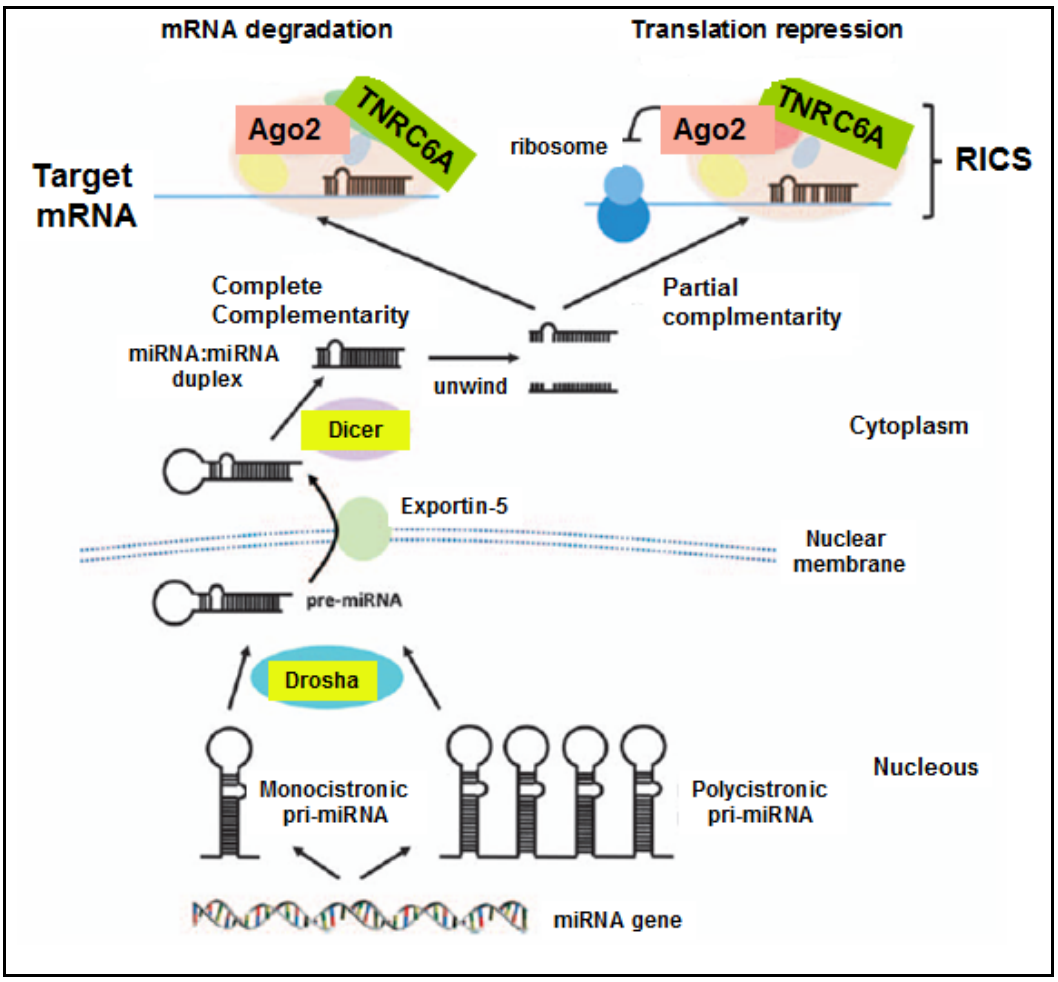
1.4.4.1 miRNA biology and functions

There are three steps in the maturation of miRNA: transcription of pri-miRNA, cleavage in the nucleus to form pre-miRNA, and a final cleavage in the cytoplasm to form mature miRNA [221, 222]. Pri-miRNA is synthesized from DNA by RNA polymerase II and may be up to 1 kb in length, forming hairpin loops. Pri-miRNAs are found as independent transcripts or within the introns of another gene. After the pri-miRNA is transcribed, it is cleaved by the RNase Drosha on the non-loop end to form 60–70 bp length precursor miRNA (pre-miRNA) [223, 224]. Pre-miRNA then moves from the nucleus into the cytoplasm via Ran-GTP-dependent Exportin 5, a transporter on the nuclear membrane [225, 226]. In the

cytoplasm, pre-miRNA is cleaved on the loop end by Dicer to form a miRNA: miRNA duplex that is unwound by a helicase to release two mature miRNAs, of which one or both may be active [227].

MiRNAs exert their functionality via sequence-specific regulation of post-transcriptional gene expression and it is estimated that they regulate up to 30% of all protein-coding genes [228]. The specific region important for mRNA target recognition is located in the 5'-end of the mature miRNA strand, from bases 2 to 8, often referred to as the 'seed-sequence' [229]. Governance of gene expression and protein translation by these noncoding RNA molecules occurs largely through one of two mechanisms, dependent upon the complementarity of the miRNA seed sequence with its target mRNA. Although remarkably small, miRNAs harbor enough sequence content to be relatively specific. Generally, if a miRNA–target duplex contains imperfect complementarity, protein expression is inhibited without target mRNA destruction. However, if the duplex has nearly perfect base-pairing, then the mRNA target is marked for degradation [229, 230]. The Argonaute proteins present in the RNA-induced silencing complex (RISC) appear to dictate the mode of regulation elicited by the miRNA–target duplex. Recruitment of specific Argonaute proteins can catalyze cleavage of mRNA sequences perfectly base-paired to the miRNA, or inhibit translation of mRNAs that form an imperfect duplex with the miRNA [231, 232].

Figure 1.3: **Pathways of miRNA biogenesis and mechanism of action**



1.4.4.2 miRNA in cancer

The recent explosion of miRNA research and discovery further underscores the importance of these regulatory molecules in many key biological processes, such as development, cellular differentiation, cell cycle control and apoptosis [233-235]. There is enough evidence to show that miRNAs are involved in human cancer [236, 237]. It was suggested previously that miRNAs assert their function as oncogenes or tumour suppressor genes via several potential mechanisms. If a particular miRNA targets key tumour suppressor genes, it is supposed to be an oncogene; but, if a miRNA targets an oncogene, it might be viewed as a tumour suppressor gene. However, the matter may be far more complicated than this simple view because one particular miRNA can mediate the expression of up to several hundred mRNAs. We speculate that to a large extent, the function of miRNAs is to fine tune gene expression in response to acute changes in growth

conditions rather than as a traditional tumour suppressor or oncogene by definition. The first evidence that miRNAs may function as tumour suppressor genes came from a recent study by Calin et al. [220] that showed that patients with B-cell chronic lymphocytic leukemia (CLL) have frequent deletions or down regulation of two miRNA genes, *miR-15a* and *miR-16-1* . Cimmino et al. [238] showed that an anti-apoptotic gene *BCL2*, was negatively regulated by *miR-15a* and *miR-16-1* . This suggests that deletion or down regulation of *miR-15a* and *miR-16-1* results in an elevated level of *BCL2* to promote leukaemogenesis and lymphomagenesis in haematopoietic cells. However, Borkhardt et al. [239] reported recently that among 69 B-cell cases with 13q deletion, none of them showed mutations in *miR-15a* and *miR-16-1*.. Fulci et al. [240] also reported that the down regulation of *miR-15a* and *miR-16-1* only occurred in 11% of 56 cases of B-cell CLL. In another report, Linsley et al. [241]demonstrated that *miR-15a* and hsa-*miR-16-1* do not behave as classical tumour suppressor genes and most importantly, they do not regulate *BCL2* expression at both mRNA and protein level. These results suggest that our notion of miRNAs can not be simply classified as traditional tumour suppressor genes or oncogenes and more studies are clearly needed to address this issue.

miRNA in Breast cancer:

Several miRNAs are associated with breast cancer. Iorio et al. [242] identified 29 miRNAs that were differentially expressed in breast cancer tissues compared to normal. Among the miRNAs differentially expressed; *miR 10b, miR-125b* and *mR-145* were downregulated whilst *miR-21* and *miR-155* were consistently over-expressed in breast tumour tissues. Furthermore, Ma et al. [243] showed increased expression of the gene encoding *miR-10b*, which is upregulated by the transcription factor *Twist1* and can promote tumour invasion in vivo. *MiR-21* has also been found to be upregulated in breast cancer, and this upregulation causes downregulation of 2 important targets: programmed cell death 4 (*PDCD4*) and tropomyosin1 (*TPM1*) [244-246]

Differential expression of genes encoding some miRNAs seems to be associated with particular pathologic features of breast cancer. Mattie et al. [247]

subsequently identified unique sets of miRNAs associated with breast tumours defined by their HER2/neu or ER/PR status . Moreover, Lowery et al. [248] has described 3 miRNA signatures predictive of ER, PR and Her2/neu receptor status, which were identified by applying artificial neural network analysis to miRNA microarray expression data. In addition; expression of the gene encoding *miR-30* seems to correlate with estrogen receptor and progesterone receptor status; downregulation of this miRNA is found in estrogen receptor– and progesterone receptor–negative tumours [242]. *MiR-206* has been found to target the 3' UTR of the estrogen receptor α protein, leading to an inverse correlation between miR-206 concentration and estrogen receptor status [249, 250]. Recently, Heneghan et al. [251] identified a systemic miRNA profile diagnostic of breast cancer, based largely on circulating *miR-195* levels. Their results highlight *miR-195* as a potentially ideal breast tumour marker; circulating levels of which reflect tumour *miR- 195* levels and correlate with tumour size and stage of disease. In addition; *miR-195* levels decrease to basal level two weeks post-curative tumour resection, and are not elevated in blood from patients with other malignancies. *MiR-213* and *miR-203* appear to correlate with tumour stage; increased expression of the genes encoding these miRNAs is found in higher-stage tumours [242]. Other miRNAs with prognostic value for breast cancer include *miR-10b, miR-21, miR-145, miR-9-3* and *let-7*; levels of these miRNAs correlate with tumour grade, degree of vascular invasion, lymph node metastases, or metastatic potential [252].

miRNAs in Gastric Cancer:

There is an increasing number of studies showing the overexpression or downregulation of specific miRNA in H. pylori-infected gastric mucosa and gastric cancer [253]. Dysregulated miRNAs include *miR-21, miR-181* family, *miR-191* and *miR-17,* which are upregulated in gastric cancer. The downregulated group of miRNAs includes *miR-200* family, *miR-143, miR-145 miR-31* and *Let-7* family. In addition; it has been demonstrated that the plasma concentrations of various miRNAs, such as *miR-17-5p, miR-21, miR-106a, miR-106b,* are higher whereas *let-7a* is lower in gastric cancer patients [254]. High levels of *miR-17 and miR-106a* in peripheral blood of gastric cancer patients have also been confirmed

in another study [255]. These findings suggest that miRNAs are useful biomarkers for early diagnosis of gastric cancer.
Recent studies suggest that polymorphisms in the miRNA genes may serve as novel risk predictors for gastric cancer. Arisawa et al. [256] in 2007 first reported that a polymorphism of miR-27a genome region is associated with a higher risk for the development of gastric mucosal atrophy in Japanese men. Peng et al. [257] later reported an association of miRNA-196a-2 gene polymorphism with gastric cancer risk in a Chinese population. Moreover; miRNAs have recently been used to predict the outcome of patients with gastric cancer. For example, a seven-miRNA signature (*miR-10b, miR-21, miR-223, miR-338, let-7a, miR-30a-5p* and *miR-126*) is closely associated with relapse-free and overall survival among patients with gastric cancer [258]. High expression levels of *miR-20b* or *miR-150 [259]* or downregulation of *miR-451* [260] or *miR-218* [261] are also associated with poor survival, whereas there is a correlation between *miR-27a* and lymph node metastasis [259]. In addition, Ueda et al. [262] recently reported that *miR-125b, miR-199a* and *miR-100* represents a progression-related signature, whereas low expression of *let-7g* and high expression of *miR-214* are associated with shorter overall survival independent of depth of invasion, lymph-node metastasis and stage [262]. These prognostic miRNAs could be applicable to future decisions concerning treatment

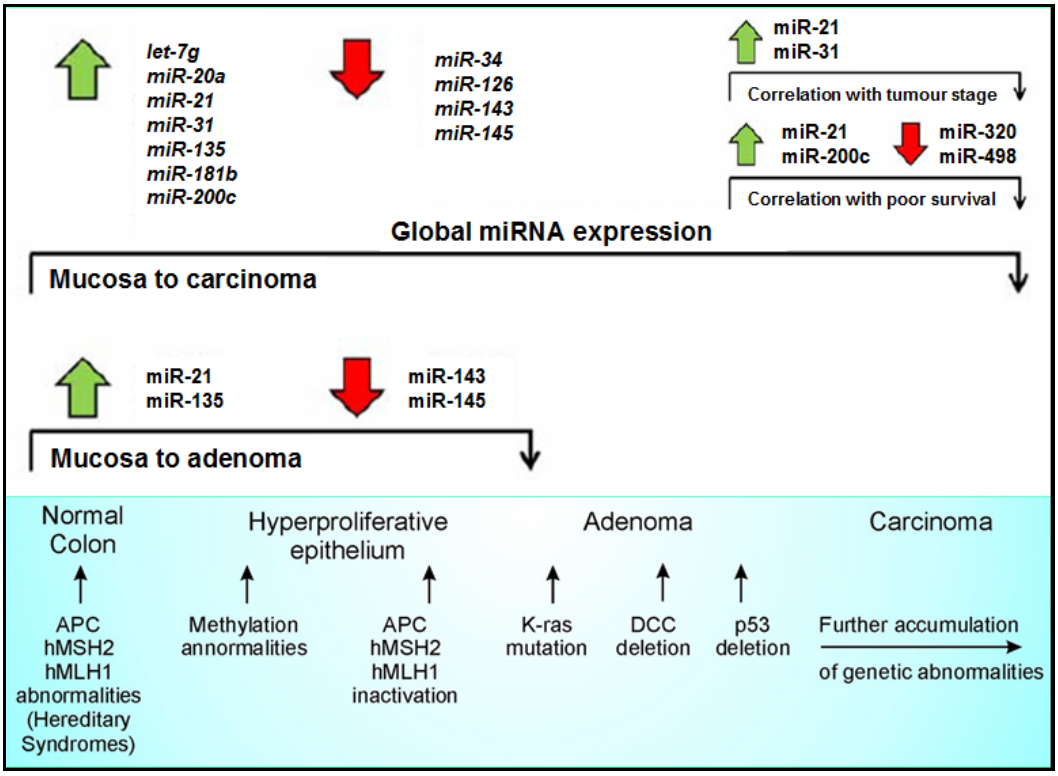
1.4.4.3 miRNA expression and functions in CRC

In 2003 Michael et al. published the first report to profile miRNAs in CRC. Using cloning technology followed by Northern blotting, they reported consistently reduced accumulation of the specific mature *miR-143* and *miR-145* in the adenomatous and carcinoma stages of colorectal tumours [263]. Thereafter, several studies were set to investigate the role of miRNAs in colorectal cancer.

MiRNAs with tumour suppressor properties which are under-expressed in CRC specimens, and thus potentially function as tumour suppressors, include *miR-31, miR-34a, miR-96, miR-143, miR-145,* and *let-7a* [264, 265]. *MiR-34a* is a well described tumour suppressor miRNA which regulates the *p53* pathway and when overexpressed induces apoptosis and acute senescence. Conversely reduction of

miR-34 expression and function attenuates *p53*-mediated cell death and is therefore implicated in tumourigenesis, including initiation of CRC [266, 267]. It is postulated therefore that loss of *miR-34a* expression in colorectal biopsy specimens may be an early biomarker of CRC. Other miRNAs like m*iR-31, miR-96, miR-135b,* and *miR-183* have been found to be upregulated in colorectal neoplasm. The transcription factor *CHES1* which is involved in repressing apoptosis *is a* potential target of *miR-96*. Schetter et al. identified miRNAs which can distinguish cancerous from normal colon tissue, with *miR-21* over-expressed in 87% of colon cancers [264]. Subsequent mechanistic investigations provide evidence for the oncogenic role of *miR-21* in CRC by demonstrating how it suppresses the cell cycle regulator *CDC25A* [268], and can also target and repress the tumour suppressor gene *PDCD4* thus inducing invasion, intravasation and metastatic potential [269]. *MiR-21* may also target *PTEN* and *TPM1*. In addition; m*iR-135a* and *miR-135b* are upregulated, and this upregulation correlates with reduced expression of the *APC* [270]. Moreover; *miR-143* and *miR-145* are both downregulated in colorectal cancer. The genes encoding these miRNAs are both located on 5q23, and these miRNAs possibly originate from the same primary miRNA [263, 265]. *MiR-126* promotes cell proliferation through modulation of phosphatidylinositol 3-kinase signaling [271]. *MiR-133b* is also downregulated, and one of its putative targets is *KRAS [272]*, which is a member of the Ras family of proteins, that regulates signaling pathways involved in cellular proliferation, differentiation, and survival. Moreover; over-expression of the oncogenic *miR-17-92* cluster is also implicated in the etiology of CRC, specifically in adenoma to adenocarcinoma progression.

Figure 1.4: **miRNA expression in colorectal tumours and their correlations**



1.4.4.4 Clinical value of miRNA expression in CRC

Accumulating evidence shows that miRNA expression patterns are unique to certain cancers and may be used clinically as prognostic and diagnostic factors as well as therapeutic targets.

Diagnostic and prognostic value:

To test the function of miRNAs in the pathogenesis of CRC, expression of 156 miRNAs was measured in both tumour and normal tissues from patients with CRC and cell lines [272]. Expression of 13 miRNAs was significantly altered, and the most significantly dysregulated miRNAs were *miR-31*, *miR-96*, *miR-133b*, *miR-135b*, *miR-145* and *miR-183*. In addition, the expression level of *miR-31* was significantly correlated with tumour stage. Xi et al. [273] analysed patients with adenocarcinoma of the colon and rectum and found that tumours expressing high levels of *miR-200c* are correlated with poorer prognosis, regardless of tumour stage: approximately 12 months decreased survival compared with patients whose tumour expresses low levels of *miR-200c*. Furthermore; Arndt et al. [274]

identified 37 miRNAs that were differentially expressed between CRC and normal tissues. They also reported that loss of *miR-133a* and gain of *miR-224* are associated with tumour progression. Overexpression of miR-21 was shown in many reports to be associated with worse prognosis, lymph node and distant metastasis and poor response to chemotherapy in CRC. Moreover, Asangani et al. [269] reported that overexpression of miR-21causes tumour cells to invade and metastasize more aggressively when implanted into mouse models. In addition; a study by Motoyama and colleagues [275] showed that expression of *miR-31, miR-183, miR-17-5p, miR-18a* and *miR-92* were significantly higher in tumour tissues compared to normal, while expression of *miR-143* and *miR-145* in cancer were lower than in normal tissues. They also showed that *miR-18a* expression was associated with poor disease prognosis. Moreover; *miR-31* expression was positively related to advanced TNM stage and tumour invasion suggesting its role in CRC initiation and progression [276]. Of further interest; Lanza et al. [277] identified a molecular signature consisting of 27 differentially expressed genes, inclusive of 8 miRNAs that can correctly distinguish high microsatellite instable (MSI-H) vs. microsatellite stable (MSS) colon cancers of

Therapeutic potential:

The synthesis and functions of miRNAs can be manipulated with various oligonucleotides that encode the sequences complementary to mature miRNAs [278]. Overexpression of miRNAs can be induced either by using synthetic miRNA mimics or chemically modified oligonucleotides [279]. Conversely, miRNAs can be silenced by antisense oligonucleotides and synthetic analogues of miRNAs [280, 281]. Cross-sensitivity with endogenous miRNAs and lack of specificity for target miRNAs can cause non-specific side-effects with miRNA modulation therapy. However, the use of an effective delivery system and less toxic synthetic anti-miRNA oligonucleotides may minimize such side-effects. The role of miRNAs in pathogenesis of cancer makes them important targets for therapeutic intervention. Gene therapies may be designed to treat colorectal cancers and to block the progression of precursor lesions by manipulating the tumour suppressor or promoter miRNAs [282]. Such manipulation may control

the tumour growth rate and have potential as a new therapy for both early and advanced cancers

Studies have revealed that inhibition of *miR-21* and *miR-17-92* activity is associated with reduced tumour growth, invasion, angiogenesis and metastasis [283, 284]. Targeting such miRNAs may help to prevent the recurrence of disease in high-risk tumours and may control the growth of advanced metastatic tumours. Overexpression of *miR-21* is associated with low sensitivity and poor response to chemotherapy [282]. Its inhibition may improve the response to chemotherapy. In addition; some drugs were found to alter the expression of miRNAs. Rossi et al. reported a suggestive pattern of miRNA rearrangement in HT-29 and HCT-116 human colon cancer cell lines after exposure to 5-FU [285]. It leads to down-regulation of *miR-200*, which is a microRNA known to inhibit a tumour-suppressor gene, protein tyrosine phosphatase, non-receptor type 12 (*PTPN12*) [286]. 5-FU treatment also induces up-regulation of *miR-133a*, which is thought to inhibit the proto-oncogene K-Ras. Strangely, treatment with 5-FU also causes up-regulation of microRNA known to be mitogenic. To this, Rossi suggests that the cytotoxic effect of 5-FU induces cells to express anti-apoptotic factors, of which are these abnormally up-regulated, and tumourigenic, miRNAs. Besides these, 5-FU treatment leads to significant elevation in expression of many other miRNAs, and it remains to be seen what genes these miRNAs target. 5-FU induces *p53* protein expression at a posttranscriptional level without correspondingly elevated mRNA level in a pattern that has become a hallmark for microRNA involvement. When wild-type HCT-116 cells are treated with 5-FU, they express high levels of certain miRNAs, and a great majority of these affected microRNA have a binding site for *p53* in the gene. When HCT-116 cells knocked out for *p53* are treated with 5-FU, these miRNAs are not up-regulated. These results suggest that 5-FU acts as a switch to turn on *p53* and, through *p53*, a cascade of miRNAs that may act with or independently of *p53*.

1.5 Study Rationale

The involvement of certain molecules in initiation and progression of human malignancy holds much potential for new developments in current diagnostic and therapeutic strategies in the management of CRC patients. While a number of miRNAs with a functional role in CRC have been identified and functionally characterised, the heterogeneity and molecular complexity of CRC makes it likely that there are many more molecules involved in the pathways that promote CRC progression and response to therapeutics. The identification of novel genes and miRNAs involved in colorectal carcinogenesis and understanding their functional effects, particularly in relation to the current indicators will improve our knowledge of the roles of these novel biomarkers in carcinogenesis and promises to open avenues for potential therapeutic intervention.

The purpose of this study was to investigate the role of mRNA, miRNA and MMR proteins by analysing their expression using the following approaches:

I- mRNA expression profile in CRC:

Analysis of gene expression patterns represents one of the most interesting topics in medical oncology, because it provides a global and detailed view on the molecular changes involved in tumour progression, leading to a better understanding of the carcinogenesis process, to discovering new prognostic markers and novel therapeutic targets. Despite of clinical and pathological parameters are available for the classification and prognostic stratification of cancer, they may be inadequate in everyday practice due to the great biologic and genetic heterogeneity of this multiform disease. Therefore, we selected a panel of candidate genes, based on literature review, to quantitate their expression in colorectal cancer using RQ-PCR in order to:

1. Determine the expression levels of candidate genes in tumour and tumour-associated normal colorectal tissue.
2. Investigate correlation between serum carcinoembryonic antigen (CEA) and tissue *CEACAM5* levels to identify a relationship that could further refine the role of *CEACAM5* as a biomarker in CRC.

3. Correlate candidate genes expression levels and clinicopathological variables

II- Prediction of rectal cancer response to neoadjuvant chemoradiation therapy (CRT):

Recently, post-transcriptional and translational controls of protein coding genes regulated by miRNA have emerged as an interesting field of cancer research. Due to their small size, they are more stable and resistant to environmental, physical and chemical stresses compared to mRNAs. Therefore; their analysis in formalin-fixed paraffin embedded (FFPE) tissue samples is likely to provide more accurate replication of what would be observed in fresh tissues than that of mRNA species. The aims of miRNA expression analysis in this study were to:

1. Optimise miRNA extraction methods from FFPE tissue samples and to systematically investigate the miRNA expression profiles between FFPE samples and fresh-frozen samples using RQ-PCR.
2. Characterise miRNA expression in tumour compared to tumour-associated normal (TAN) FFPE colorectal tissues. Moreover; we aimed to identify predictors of response to neoadjuvant chemoradiation therapy in colorectal cancer using FFPE tissues as source of genetic materials, and microarray analysis as investigation tool.
3. Correlate the expression levels of candidate mRNA to a panel of miRNAs in order to identify miRNA/mRNA duplexes and to investigate the miRNA and target gene expression patterns in colorectal tissue samples.

III- MMR protein expression in CRC:

Information about MMR protein status in colorectal cancer is important because it will identify those most likely to have Lynch syndrome and those most likely to have microsatellite instability in their tumours which has been proven to have better prognosis and may affect their treatment regimens in the future. We undertook this study to:

1. Develop and optimise a protocol for MMR protein immunohistochemistry testing in colorectal cancer.
2. Analyse the proportion of patients with colorectal cancer with loss of immunostaining for MMR proteins (hMLH1, hMPS2, hMSH2 and hMSH6) in order to determine the feasibility of molecular screening for the loss of MMR proteins through the study of unselected patients with colorectal cancer.

Chapter 2:
Materials and Methods

2.1 Colorectal cancer tissues

2.1.1 Department of Surgery Bio-Bank:

The NUI, Galway Department of Surgery Bio-Bank was established at Galway University Hospital in 1992. The original purpose was to archive breast cancer-related tissues and clinical data for research. Since 2007 additional tumour related tissues in particular colorectal specimens, have been added to this repertoire which subsequently expanded again to incorporate the acquisition and storage of prostate, renal and skin cancers. With regard to colorectal tissue, ethical approval from the Galway University Hospitals Research Ethics Committee was sought and granted to BioBank tissue, blood and clinically relevant data from consenting patients (Appendix 1: Consent Form) undergoing investigations for colorectal cancer and to use them for departmental research programmes. The colorectal cancer-related material being collected and stored includes:

- Primary colorectal tumour and tumour associated normal (TAN) tissue retrieved from patients at time of diagnostic procedures or at time of resection.
- Serum, plasma and whole blood samples retrieved from patients pre- and post tumour resection.
- Serum, plasma and whole blood samples retrieved from non-cancer controls (Appendix 2: Specimen Request Form).

In accordance with the guidelines [287] including:

- Study description and approval
- Participants selection and recruitment, capacity, age and informed written consent
- Research procedures
- Access to medical records and data protection and processing
- Human biological materials
- Genetic testing
- Clinical assessment

Tissue samples are routinely collected at the time of surgical resection or diagnostic procedure and immediately transported to the Galway University Hospital Histopathology laboratory for histopathological review by a consultant

histopathologist. Subsequently two pairs of tumour and TAN samples are retrieved: one pair of samples is immediately snap-frozen in liquid nitrogen and then transported to the department of surgery research laboratory for data input and storage at -80°C. The second pair is placed in formalin for fixation and prior to paraffin embedding in the research laboratory, as described in section (2.1.2) of this chapter.

Blood samples are routinely collected in two Vacuette EDTA K3E blood bottles (5 mL) (Grenier Bio-one) and one Vacutainer Serum Separator Tubes II (10 mL) (Becton Dickinson). Of the samples collected in Vacuette EDTA tubes, one is processed for the plasma and the other is stored as whole blood. Serum samples collected in Vacutainer tubes and are incubated at room temperature for 30 minutes to enable clotting. All samples assigned for serum and plasma collection are centrifuged at 2000 rpm at 4°C for 10 minutes in Sorval RT 6000D refrigerated centrifuge (GMI Inc.). Plasma or serum is aspirated using sterile disposal Pasteur pipettes, aliquotted and stored at -20°C in 2mL Sarstedt tubes until required. Whole blood samples are stored at 4°C prior to processing for nucleic acid extraction.

Table 2.1: Author's contribution to colorectal Bio-Banking

Tissues				Blood		
Fresh-frozen		Formalin-fixed		Serum	Whole blood	Negative control
Tumour	Normal	Tumour	Normal			
45	48	36	45	16	40	14

2.1.2 Formalin-fixed paraffin embedded (FFPE) tissues

2.1.2.1 Formalin fixation

Following excision, pairs of tissues (tumour and TAN) were placed in 10% formalin (Lennox) for fixation, prior to paraffin embedding. The 10% formalin solution was prepared as follows:

Sodium phosphate monobasic	4 g
Sodium phosphate dibasic	6.5 g

Formaldehyde (37%)	100 mL
Distilled water	900 mL

Biopsies were fixed and stored at room temperature until embedding for a minimum of 24 hours.

2.1.2.2 Paraffin embedding

After fixation, tissue samples (10mm×5mm×2mm) were removed from formalin and placed in open cassettes. The cassettes were then closed and placed in 250 mL of industrial methylated spirit (VWR) to wash the formalin from the tissue. Next, the cassettes were removed and placed in JFC solution (Milestone) filled JFC beaker and placed in the histoprocessor (MicroMED) for 60 minutes (70°C). Thereafter, the cassettes were transferred to paraffin wax (VWR)-filled beakers and placed in the histoprocessor (MicroMED) for 30 minutes. The cassettes were then removed from the wax beaker and tissue was blocked out carefully. The blocks were left at 4°C until hard and then stored at fridge or room temperature until sectioning.

2.1.2.3 Sectioning

Sectioning of formalin-fixed paraffin-embedded tissues was carried out using a Slee microtome (LIS Ltd). Tissue blocks were inserted into the holder with the label facing downwards. Section thickness was set to 30µM to pare the block down until even sections were being cut and the outer layer of wax was removed. Then the section thickness was adjusted to the required size (5 µM for immunohistochemistry staining and 10 µM for molecular analysis and RNA extraction experiments). For immunohistochemistry the 5 µM sections were placed in a floating out bath to stretch them out, before being placed onto a Superfrost plus (positive charged) slides (VWR). The slides were allowed to air-dry overnight at room temperature and then stored at 4°C until further use. While for molecular studies 3 of the 10 µM sections (after the first 2 sections been discarded) were placed into a 2 mL sterile tube and immediately preceded for RNA extraction process. (Section 2.2.2)

2.1.2.4 Haematoxylin-Eosin staining:

Prior to enrolment in any further analysis each slide is stained in H & E and reviewed by a pathologist to determine the quality of the block and the percentage of tumour tissues in the section (should be >50%). The sections were deparaffinised in two changes of 100% Xylene (Sigma-Aldrich) for 10 minutes each at room temperature and then re-hydrated through two changes of 100% ethanol (Sigma-Aldrich) for 5 minutes each, 95% ethanol for 2 minutes and 75% ethanol for 2 minutes at room temperature. Then the slides were washed in distilled water before stained in Mayer haematoxylin solution for 5 minutes. The slides were washed in warm running tap water for 10 minutes and rinsed in 95% ethanol prior to counter-staining in Eosin-phaloxine B solution for 2 minutes. The slides were then dehydrated through 95% ethanol and two changes of 100% ethanol and cleared in two changes of 100% Xylene, 5 minutes each. A layer of DPX mounting medium (Sigma-Aldrich) was added to each slide , followed by the application of cover glass, taking care to avoid bubble formation. Slides were allowed to dry overnight and then examined.

Table 2.2: **Percentage of tumour per section of the FFPE tissues**

Tissue no	Tumour location	Tumour %
T08-0418	Proximal colon	100%
T07-2238	Rectum	100%
T07-2256	Proximal colon	100%
T08-0605	Rectum	70%
T08-0727	Proximal colon	100%
T08-1055	Proximal colon	100%
T08-1056	Proximal colon	100%
T08-1167	Distal colon	100%
T08-1095	Distal colon	100%
T08-1102	Distal colon	100%
T08-0907	Distal colon	100%
T08-0713	Distal colon	100%
T08-0534	Proximal colon	75%

2.1.3 Clinical data collection:

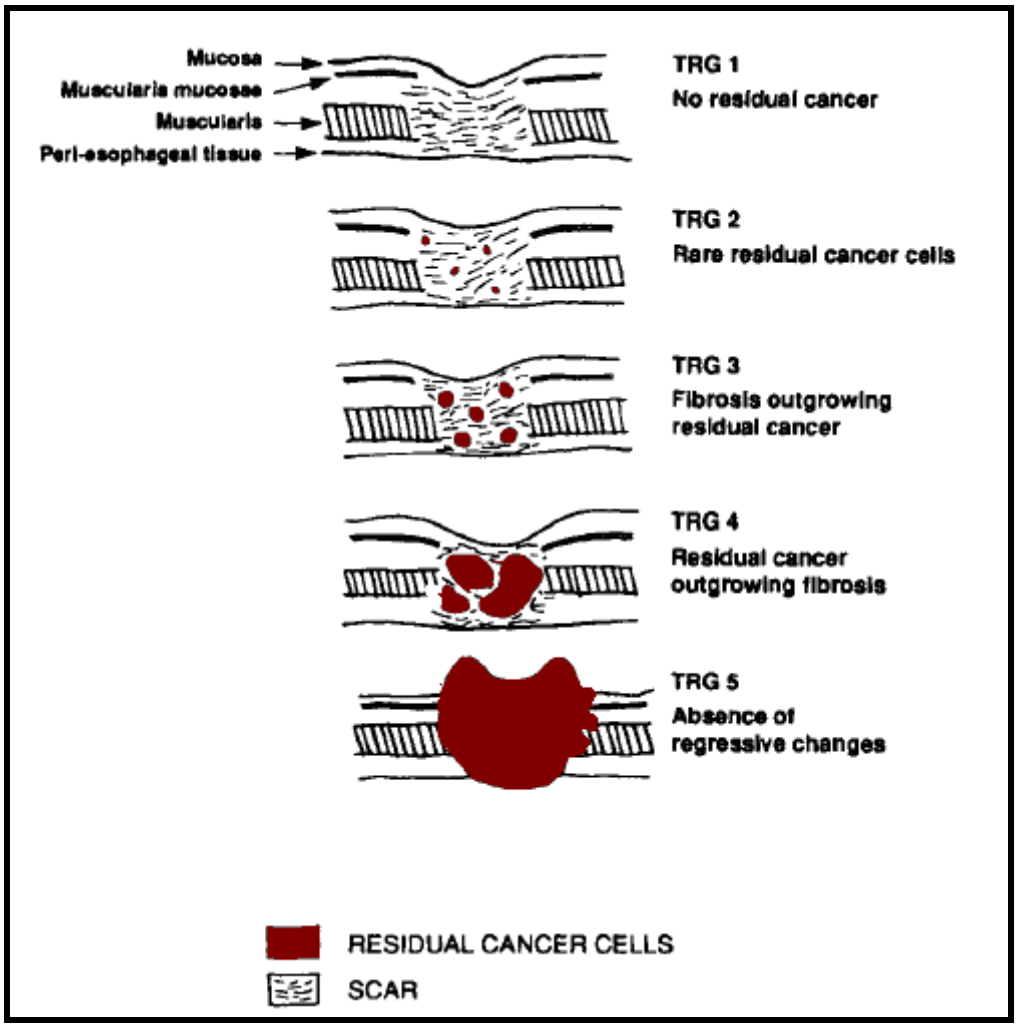
Clinical and pathological data related to patients are obtained through patient interview and review of clinical charts. This information is prospectively updated every 6 months. Relevant clinical data includes:

- Demographic data like sex and age
- Tumour location (colonic or rectal - proximal or distal)
- Tumour thickness (mm), diameter (mm) and the percentage of circumferential bowel wall involvement as measurements of tumour size
- Presence or absence of distant metastasis
- Nodal status (mesenteric nodal involvement)
- Tumour grade which represent the degree of differentiation, as gauged primarily by architectural features and is defined base on TNM classification [288], i.e., grade 1 as well- differentiated, grade 2 is moderately-differentiated, grade 3 as poorly-differentiated and grade 4 as undifferentiated [289].
- Pathological data includes perineural and lymphovascular invasion, and mucin secretion
- Response to neoadjuvant therapy was scored based on tumour regression score by Mandard [44].
- Dukes' [8] and American Joint Committee on Cancer (AJCC) [290] systems were used for disease classification and staging.
- Tumour markers (CEA and CA 19.9) serum levels

Table 2.3: Mandard tumour regression score:

Tumour regression grade 1	No residual cancer
Tumour regression grade 2	Rare residual cancer cells
Tumour regression grade 3	Fibrosis outgrowing residual cancer
Tumour regression grade 4	Residual cancer outgrowing Fibrosis
Tumour regression grade 5	Absence of regressive changes

Figure 2.1: Tumour regression grade.
As described by Mandard in oesophageal carcinoma [44]



2.2 Study groups:

Clinicopathological data on all patients were examined in order to select suitable samples for study groups appropriate to address specific questions. Informed written consent was obtained from each patient prior to enrolment in the study and ethical approval for this study was granted by the Clinical Research Ethics Committee, Galway University Hospitals.

A study group of 64 biopsies of human colon tissue samples was gathered to identify the best endogenous control gene/s to normalise data in colorectal cancer gene expression profiling experiments using RQ-PCR. The cohort consists of 30 colorectal tumour specimens and 34 tumour-associated normal (TAN) tissues. A heterogeneous group of 107 patients with colorectal tumours, all of which had matched TAN samples was selected for gene expression profiling experiment using real-time quantitative PCR (RQ-PCR) (Table2.5)

A group of 9 patients was selected for miRNA extraction from FFPE tissues in order to evaluate miRNA quality in relation to miRNA extracted from fresh-frozen tissues. Each patient in this group has both FFPE and fresh frozen/ tumour and TAN tissues available. Then a group of 12 patients who had pre-operative chemoradiation therapy and had pre-treatment tissue samples available was selected for microarray analysis in order to determine predictors of response to neoadjuvant therapy.

For the microsatellite instability experiment we selected a group of 33 patients with colorectal cancer to test the expression of MSI proteins in their tumour tissues using immunohistochemistry techniques.

A group of 65 patients, in whom the expression levels of a panel of miRNA was carried out before at the surgical research laboratory, were selected for the miRNA:mRNA correlations in order to identify miRNA/target genes duplexes.

Table 2.4: Study groups.
The number of samples in each group (n) and the criteria for selecting samples for each group.

Group Name	Samples (n)		Group Details
	Tumour	TAN	
Endogenous control group	30	34	Heterogeneous group of colorectal tissues
Gene expression profiling group	101 cancer & 8 polyps	107	Heterogeneous group of colorectal cancer patients
FFPE miRNA extraction group	9	9	Each patient in this group has both FFPE and fresh frozen/ tumour and TAN tissues available
Neo-adjuvant therapy response prediction group	12	0	Had neo-adjuvant therapy and had pre-treatment tissue samples available
Microsatellite instability (MSI) group	33	0	Had colorectal cancer and had FFPE tissues available
mRNA:miRNA correlation group	58 cancer & 7 polyps	62	MiRNA expression analysis was carried out on them before in the surgical laboratory.

Table 2.5: **Clinicopathological data of gene expression study group.**

Clinicopathological Variable	Number of Patients N=(107)
Tissue type	
Carcinoma	101
Polyp	8
Gender	
Males	67
Females	40
Mean Age (SD)	69.72 (11.89)
Tumour Location	
Colon	43
Rectum	58
Tumour Location	
Proximal	27
Distal	74
Tumour thickness (mm)	
<10	23
10-15	33
>15	22
Unknown	23
Tumour Diameter (mm)	
<30	29
30-40	26
>40	31
Unknown	15
Distant Metastasis	
M_0	80
M_1	21
Nodal Status	
N_0	22
N_1	11
N_2	9
UICC Stage	
Stage 0	2
Stage I	17
Stage II	28
Stage III	28
Stage IV	21
pCR	5
Tumour Differentiation	
Grade 1: Well differentiated	11
Grade 2: Moderate differentiated	72
Grade 3: Poor differentiated	10
Not applicable	8
Mucin Secretion	
Mucinous	19
Non-mucinous	82

2.3 RNA extraction and analysis

2.3.1 RNA extraction from fresh-frozen tissue

Two methods of RNA extraction were employed in the study, the total RNA extraction (co-purification) and the separate purification of mRNA and miRNA. The co-purification method includes isolation of total RNA with a subsequent mRNA or small RNA purification from the total RNA pool. The second method purifies mRNA and miRNA directly out of solution via poly-A isolation or sequence-specific isolation. The separate purification was used when miRNA analysis was required. To ensure both methods were working properly correlation of RNA extraction were carried out and showed good results (Table 2.6)

Table 2.6: RNA co-purification vs. separate purification extraction methods

No	Total RNA (Co-purification)		miRNA (Separate purification)		mRNA (Separate purification)	
	RNA Conc. (ng/µl)	A_{260}/A_{280}	RNA Conc. (ng/µl)	A_{260}/A_{280}	RNA Conc. (ng/µl).	A_{260}/A_{280}
1	3290.6	2	285.57	1.94	1807.6	2.06
2	1511	2.07	219.66	1.96	1664.3	2.04
3	2199.2	2.03	298.97	1.86	2103.8	2.04
4	1165.3	2.05	46.15	1.7	1005.5	2.06
5	2551.8	2.01	326.13	1.85	2789.1	2.02
6	2617.3	2.04	356.27	1.89	2451.1	2.02
7	139.03	2.07	71.32	1.78	294.8	2.06
8	34.23	1.94	24.29	1.7	55.01	2.03
9	744.31	2.06	17.44	1.57	654.59	2.06
10	1745	2.07	311.46	1.81	1848.4	2.04
11	1140.6	2.07	168.73	1.87	1367.2	2.05
12	1876.2	2.05	229.24	1.9	1968.2	2.04
13	691.12	2.06	123.49	1.81	1083.7	2.04
14	232.61	1.98	103.51	1.79	474.55	2.01
15	727.89	2.03	198.49	1.87	1244.6	2.09
16	749.96	2.07	191.29	1.87	1145.8	2.04
17	2375.2	2.02	278.38	1.93	3052.3	2.02
18	1301.4	2.05	195.52	1.91	980.92	2.1
19	1631.7	2.04	64.58	1.8	1078.6	2.1

2.3.1.1 Total RNA extraction (co-purification)

Approximately 50-100 mg of fresh-frozen colorectal tissue samples were homogenised using a hand-held homogenizer (Polytron PT1600E) in 1-2 mL of QIAzol reagent (Qiagen). To minimise variation in sample processing, tumour and TAN samples were homogenised separately, but on the same day. Total RNA was isolated from homogenised tissues using RNeasy Plus Mini kit (Qiagen) according to the manufacturer's instructions. An Eppendorf Micromax refrigerated centrifuge was used throughout the RNA extraction process. Aliquote of 500 µL of the homogenate was transferred to sterile 1.5 mL tubes and centrifuged at 14000 rpm for 10 minutes at 4°C, before the addition of chloroform to sediment insoluble material and precipitate fats. Three and a half volumes of 100% ethanol were added to the upper aqueous phase and the entire volume was transferred to RNeasy mini kit column. To pass the sample through the column, a centrifugation of 12000×g for 21 seconds at 4°C was applied. The column was washed with 350 µL buffer RW1 by centrifuging at 12000g for 21 seconds at 4°C. Then a DNase I treatment was performed. A volume of 80µL of DNase I mix made using reagents from an RNase-free DNase set (Qiagen) was applied onto the membrane of the column and left at room temperature for 15 minutes. The buffer RW1 wash step was repeated. Two further wash steps, using buffer RPE, were carried out. The second of these steps had an increased centrifugation time of 2 minutes to dry the membrane. Total RNA was eluted from the RNeasy column by applying 60µL RNase –free water to the membrane and centrifuging at 12000×g for I minute at 4°C. A volume of the total RNA was aliquoted for quantitative and qualitative analysis using NanoDrop spectrophotometry (NanoDrop technologies) and the Agilent 2100 Bioanalyzer System (Agilent technologies) respectively. The remaining RNA was stored at -80°C until further use.

2.3.1.2 Large and micro RNA extraction (separate purification)

Using this method large (> 200 nt) and small RNA (< 200 nt) fractions were isolated separately using the RNeasy Plus Mini Kit and RNeasy MinElute cleanup kit (Qiagen) according to the manufacturer's supplementary protocol: purification of miRNA from animal cells. Approximately 100 mg of fresh-frozen tissue was

homogenised) in 1-2 mL of QIAzol reagent (Qiagen) using a hand-held homogenizer, as above. An initial centrifugation step of 12000×g for 10 minutes at 4°C was used, before addition of chloroform, to bring excess fat to the surface and allow for its removal by pipetting. One volume of 80%ethanol was added to the upper aqueous phase before being added to the RNeasy column. A centrifugation of 12000×g for 21 seconds at 4°C was then used. 1.4 volumes of 100% ethanol were added to the flow-through from the RNeasy column and this was mixed thoroughly by vortexing. The RNeasy column was stored at 4°C for subsequent isolation of large RNA. The sample was passed through a MinElute column by centrifuging at 12000×g for 21 seconds at 4°C. The same centrifugation conditions were used for two subsequent wash steps using 500 µL buffer RPE, from the RNeasy Plus Mini Kit (Qiagen), and 500µL 80% ethanol respectively. The small RNA was eluted from the MinElute column in 20µL RNase-free water by centrifuging at 12000×g for 1 minute at 4°C. The RNeasy column was removed from the 4°C and the column was washed with 350 µL buffer RW1 by centrifuging at 12000×g for 21 seconds at 4°C. DNase I treatment was then performed using 80µL of DNase I mix, made using reagents from the RNase-free DNase kit (Qiagen). Enzyme was applied onto the membrane of the column and left at room temperature for 15 minutes. The buffer RW1 wash step was repeated. Two further wash steps, using 500µL of buffer RPE, were carried out. The second of these steps had an increased centrifugation time of 2 minutes to dry the membrane. The large RNA was eluted from the RNeasy column by applying 50µL RNase-free water to the membrane and centrifuging at 12000×g for 1 minute at 4°C. A portion of the purified large and small RNA was aliquoted for quantitative and qualitative analysis using NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies) and the Agilent 2100 Bioanalyzer System (Agilent technologies) respectively. The remaining RNA was stored at -80°C until further use.

2.3.2 RNA extraction from FFPE tissue

2.3.2.1 Qiagen RNeasy FFPE kit

This method was employed using RNeasy FFPE kit (Qiagen) according to the Qiagen supplementary protocol. Paraffin sections (3×10µm) were cut from FFPE

block using a Microtec 4050 cut microtome (Techno-Med Biefield). The first two whole sections were discarded because of exposure to the atmosphere and 3 subsequent sequential sections were placed in a 2 mL microcentrifuge tube. To each sample 1mL of 100% xylene (Sigma Aldrich, Germany) was added, samples were vortexed vigorously for 10 seconds, and centrifuged at full speed (20000×g) for 2 minutes at 20-25°C. The supernatant was removed using a pipette and discarded. 1mL of 100% ethanol (Sigma, Germany) was added to each tube; samples were mixed by vortexing and centrifuged at 20000×g for 2 minutes at 20-25°C. The supernatant was removed and the tube was left open and incubated at room temperature (15-25°C) for 10 minutes to completely evaporate all residual ethanol which may reduce RNA yield. The pellet was the resuspended in 240 µL buffer PKD and 10 µL of Proteinase K (Qiagen) and mixed by vortexing. Sample was then incubated on an orbital heating block at 55°C for 15 minutes and then at 80°C for 15 minutes to partially reverse formaldehyde modification of nucleic acids. Longer incubation times or higher the temperature may have resulted in more fragmented RNA, and so were avoided. Buffer RBC (500 µL) was added to lysate and mixed thoroughly to adjust binding conditions. The entire volume was transferred to a gDNA eliminator spin column placed in a 2mL collection tube. Samples were centrifuged at ≥8000×g for 30 seconds at 20-25°C, after which the column was discarded and flow-through saved. To the flow-through 1.75 mL of 100% ethanol was added and thoroughly mixed by pipetting. Of the sample, 700 µL (including any precipitate) was transferred to an RNeasy MinElute spin column placed in a 2 mL collection tube and centrifuged for 15 seconds at ≥8000×g at 20-25°C. The flow-through was discarded and the column retained. This was repeated until the entire sample drawn through the column. Two wash steps, using 500µL of buffer RPE, were carried out. The second of these steps had an increased centrifugation time of 2 minutes to dry the membrane. The RNeasy MinElute column was placed in a new 2mL collection tube, and the old collection tube with the flow-through was discarded. The lid of the spin column was opened and centrifuged at full speed for 5 minutes. The collection tube with the flow-through was then discarded. The RNeasy MinElute column was the placed in a new 1.5 mL collection tube. The RNA was eluted from the RNeasy column by applying 30µL RNase-free water to the membrane and centrifuging at full speed

for 1 minute at 20-25°C. A portion of the purified RNA was aliquoted for quantitative and qualitative analysis using NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies) and the Agilent 2100 Bioanalyzer System (Agilent technologies) respectively. The remaining RNA was stored at -80°C until further use.

2.3.2.2 Qiazol and chloroform protocol

Paraffin sections (3×10µm) were prepared as previously described and placed in 2 mL microcentrifuge tubes. To each sample 1mL of 100% xylene (Sigma Aldrich) was added, samples were vortexed vigorously for 10 seconds, and centrifuged at full speed (20000×g) for 2 minutes at 20-25°C. The supernatant was removed using a pipette and discarded. Then, 1mL of 100% ethanol (Sigma, Germany) was added to each tube; samples were mixed by vortexing and centrifuged at 20000×g for 2 minutes at 20-25°C. Ethanol wash was repeated twice. The supernatant was removed and the tubes were left open and incubated at room temperature (15-25°C) for 10 minutes to completely evaporate. 1mL of QIAzol reagent (Qiagen) was added to each sample prior to disruption using a needle and syringe before the addition of chloroform. Three and a half volumes of 100% ethanol were added to the upper aqueous phase and the entire volume was transferred to RNeasy mini kit column and the process was continued as for total RNA extraction from fresh-frozen tissues. A portion of the RNA was aliquoted for quantitative and qualitative analysis. The remaining RNA was stored at -80°C until further use.

2.3.2.3 TRI reagent RT-Blood protocol

Paraffin sections (3×10µm) were placed in a 2 mL microcentrifuge tube. Xylene and 100% ethanol (Sigma-Aldrich) wash were carried out as described in the preceeding section. After complete evaporation of the ethanol at room temperature for 10 minutes 1mL of QIAzol reagent (Qiagen) was added to the sample and then homogenized using needle and syringe. To precipitate RNA, 80% of each aqueous phase (about 1 mL) was then transferred to a new 2 mL round tube and mixed with a similar volume of isopropanol (Sigma-Aldarich). The mixture was stored at room temperature for 5 minutes and the centrifuged at 14000 rpm for 8 minutes at 18°C. RNA precipitate formed a gel-like pellet at the

bottom of the tube. Two wash steps using 75%ethanol were then carried out in order to improve the quality of RNA (260/280 ratio). The RNA pellet was air-dried for 5 minutes before dissolving RNA in 30 µL of nuclease-free water. The dissolve was incubated at room temperature for 5 minutes, vortexed and spin down for 10 seconds. A volume of the RNA was aliquoted for quantitative and qualitative analysis. The remaining RNA was stored at -80°C until further use.

2.3.3 RNA extraction from blood

Total RNA was extracted from 1mL of whole blood using the Tri Reagent BD (http://www.mrcgene.com/rna.htm) and a modified protocol from that provided by the manufacturers. In brief, 1-bromo-4-methoxybenzene was used to augment the RNA phase separation and an additional ethanol (75%) wash was performed to improve the purity of RNA isolated as reflected in an improved 260/280 ratio. RNA concentration and integrity were determined by spectrophotometery (NanoDrop technologies) and bioanalyzer (Agilent technologies)

2.3.4 RNA concentration and quality analysis

RNA concentration and purity was assessed in duplicate samples (1 µL) using a NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies) while RNA integrity was evaluated using the RNA 6000 Nano Chip Kit (Series II) and the Agilent 2100 Bioanalyzer System (Agilent technologies).

2.3.4.1 Large and total RNA analysis

***Nanodrop Spectrophotometery*:**

Total and large RNA concentration and purity was assessed using the NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies). Aliquote of 1 µL of RNA was pipetted onto the apparatus pedestal. The sample arm was used to compress the sample and a sample column formed, held in place by surface tension. Spectral measurements were made with a tightly controlled pathlength of 0.1cm. RNA concentration was automatically calculated using the formula:

$$\text{RNA concentration (ng/µL)} = (A_{260} \times e)/b$$

A_{260} = Absorbance at 260 nm, e = extinction coefficient (ng-cm/mL), b = pathlengh (cm)

When analysing total or large RNA sample, RNA-40 was selected as the sample type and an extinction coefficient of 40 was used. RNA with an absorbance ratio at 260 and 280 nm (A_{260}/A_{280}) between 1.8 and 2.2 was deemed indicative of pure RNA. The presence of protein or phenol results in high absorption at 280 nm, producing a lower A_{260}/A_{280} ratio. A ratio at 260 and 230 nm (A_{260}/A_{230}) between 1.8 and 2.2 was considered acceptable. Lower ratio can indicate the presence of contamination.

Agilent Bioanalyzer:

The large-RNA enriched fractions and the total RNA were also analysed using the RNA 6000 Nano LabChip series II Assay and the Agilent 2100 Bioanalyzer (Agilent technologies). RNA samples loaded onto the Agilent chip were separated by capillary electrophoresis according to their molecular weights. The intensity of fluorescence on each sample's electropherogram represented the amount of RNA of a given size. The Total RNA Assay was carried out according to the Agilent RNA 6000 Nano Assay protocol. The electrodes were cleaned with RNase-free ZAP for 1 minute and RNase-free water for 30 seconds prior to use. To prepare the gel, RNA 6000 Nano gel matrix and dye concentrate were allowed to equilibrate at room temperature for 30 minutes; the latter reagent was protected from light throughout by covering the tube in tin-foil. The complete volume of gel was spun at 10000×g for 15 minutes. The dye concentrate was vortexed for 10 seconds and briefly centrifuged. In a new 0.5 mL RNase–free tube, 1 µL of dye concentrate and 65 µL of the filtered gel were mixed thoroughly by careful pipetting. The gel/dye mix was the spun at 13000×g for 10 minutes at room temperature. Samples were diluted to 1 ng/µL within the quantitative and qualitative range of the assay. The RNA samples and RNA ladder were denatured at 70°C for 2 minutes and then placed on ice prior to use. A new Nano Chip was placed on the Chip Priming Station where 9.0 µL of gel/dye mix were pipetted into a marked well before closing the chip priming station for 30 seconds. Another 9.0 µL of gel/dye mix were pipetted into the second marked well before discarding the rest of the mix. A volume of 5 µL of the RNA 6000 Nano marker was pipetted into all the 12 sample wells and the ladder well. Then 1 µL of sample was drawn into each of the 12 sample wells and 1 µL of the ladder was

pipetted into the ladder well. Before the chip was run in the Agilent 2100 bioanalyzer it was vortexed horizontally in the adaptor at 2000rpm for 5 minutes. An RNA integrity number [291] was generated for each sample using the Agilent 2100 Expert Software (Version B.02.03) based on the ratio of ribosomal bands and also the presence or absence of degradation products on the electrophoretic and gel-like images. A threshold value of RIN ≥ 7 was applied; ensuring only RNA of good integrity was used in these experiments.

Figure 2.2: Analysis of RNA quality.
(A) Small RNA and (B) large RNA concentration and quality analysis. RNA concentration and purity was assessed by NanoDrop spectrophotometry while RNA integrity was evaluated using an Agilent 2100 Bioanalyzer

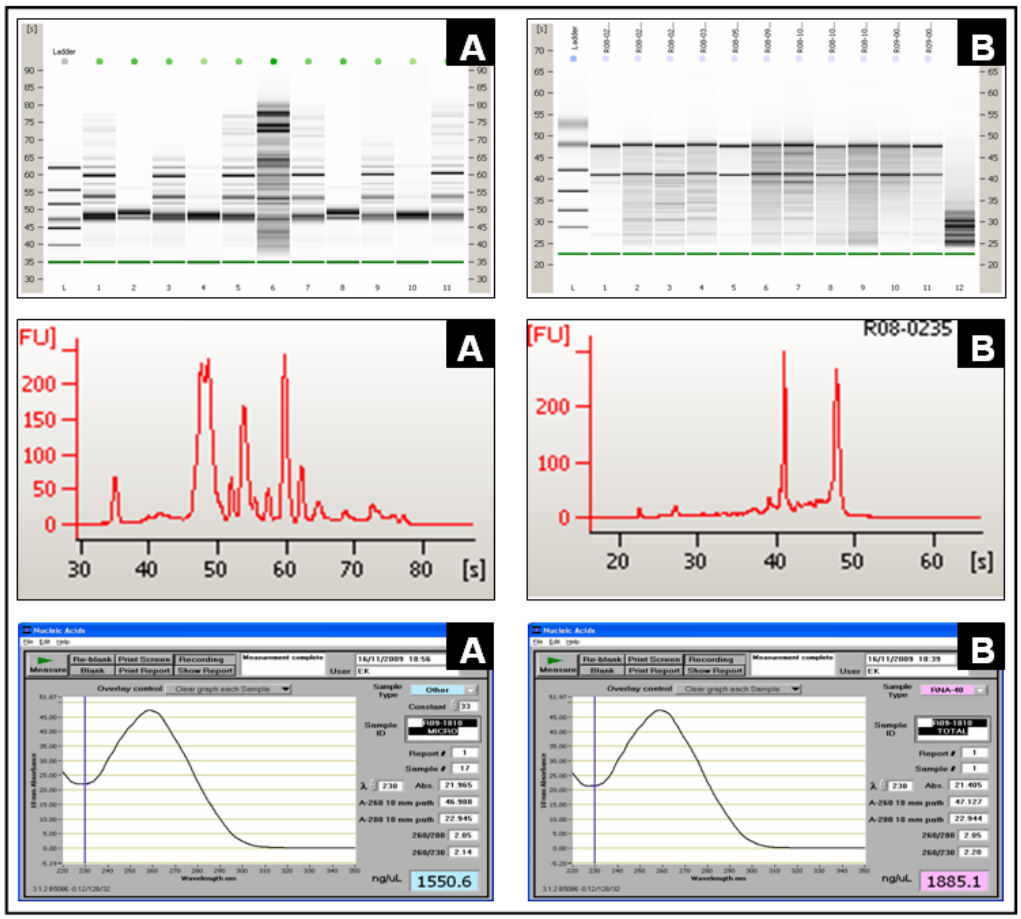


Table 2.7: **RNA concentration and quality**

Sample ID	RNA integrity number (RIN)	A_{260}/A_{230}	A_{260}/A_{280}	RNA Conc. ng/µL
R07-1671	7.7	1.9	2.06	242.2
R07-1929	8.2	2.08	2.07	1142.6
R07-1673	7.4	2.09	2	1807.6
R07-1674	8.3	2.18	2.07	1664.3
R08-0219	8.6	2.13	2.06	1081.6
R08-0222	7.3	2.18	2.04	1978.4
R08-0225	7.8	2.19	2.04	1973.1
R08-0226	7.4	2.19	2.03	2103.8
R08-0228	8.2	2.03	1.94	2398.8
R08-0229	7.1	2.02	2.04	1669.8

2.3.4.2 MiRNA analysis

***Nanodrop Spectrophotometery*:**

The concentration and purity of small RNA were assessed using a NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies). 'Other' was selected as the sample type and the wavelength-dependent extinction coefficient of 33 was used. RNA integrity was assessed using Small RNA Assay with the bioanalyzer (Agilent technologies).

Agilent Bioanalyzer

The Small RNA Assay was chosen for its high resolution in the 6-150 nucleotide range, allowing verification of small RNA retrieval and comparison of the small RNA component between tissue samples. The small RNA assay was carried out according to the Small RNA Assay kit guide. The electrodes were cleaned with RNase-free water prior to use. To prepare the gel, Small RNA gel matrix and dye concentrate were allowed to equilibrate at room temperature for 30 minutes. The complete volume of gel was spun at 10000×g for 15 minutes. The dye concentrate was vortexed for 10 seconds and briefly centrifuged. In a new 0.5 mL RNase–free tube, 2 µL of dye concentrate and 40 µL of the filtered gel were mixed thoroughly by careful pipetting. The gel/dye mix was the spun at 13000×g for 10 minutes at

room temperature. Samples were diluted to 1 ng/µL within the quantitative and qualitative range of the assay. The RNA samples and RNA ladder were denatured at 70°C for 2min and then placed on ice prior to use. To prepare the chip 9.0 µL of gel/dye mix were pipetted into a marked well before closing the chip priming station for 60 seconds. Another 9.0 µL of gel/dye mix were pipetted into the second marked well before discarding the rest of the mix. Small RNA conditioning solution was then drawn into the well marked CS. 5.0 µL of the Small RNA marker was pipetted into all the 11 sample wells and the ladder well. Then 1 µL of sample is drawn into each of the 11 sample wells and 1 µL of the ladder is pipetted into the ladder well. Before the chip was run in the Agilent 2100 bioanalyzer it was vortexed horizontally in the adaptor at 2000rpm for 5 minutes.

2.4 Reverse transcription (cDNA synthesis)

2.4.1 mRNA Reverse transcription

Aliquots of large RNA equivalent to 2 µg were reverse transcribed using Superscript III reverse transcriptase (Invitrogen). RNA (in a final volume of 23.34 µL water) was combined with the following reagents:

RNA (2µg)	23.34 µL	65°C 15 minutes
Random nonomer primers (1.0 µg)	0.66 µL	
dNTP (10mM)	2.0 µL	

After this initial RNA denaturation/primer annealing step, the remaining reagents were added:

RT buffer (5X)	4.0 µL	25°C 30 mins / 50°C 60 mins 70°C 15 mins / 4°C$^{\infty}$
DTT (0.1M)	1.0 µL	
RNaseOUT(40U/µl)	1.0 µL	
SuperScript III RT	1.0µL	

After a gentle mix, the tubes were briefly centrifuged. The mixture, (40µL in total) was incubated as above in an AB9700 GeneAmp thermal cycler (Applied Biosystems). Samples were subsequently diluted to 100 µL in nuclease-free water and stored at -20º C. An RT-negative control was included in each batch of reactions.

2.4.2 miRNA Reverse transcription

Small RNA (5ng or 100ng, depending on tissue type) was reverse transcribed to cDNA using MultiScribe Reverse Transcriptase (Applied Biosystems). Each reaction was primed using a gene-specific stem-loop primer. Where sequences were available, primers were obtained from MWG Biotech. Otherwise, assays containing the RT stem-loop primer and the PCR primers and probes were used (Applied Biosystems). MiRNA was reverse-transcribed as follows:

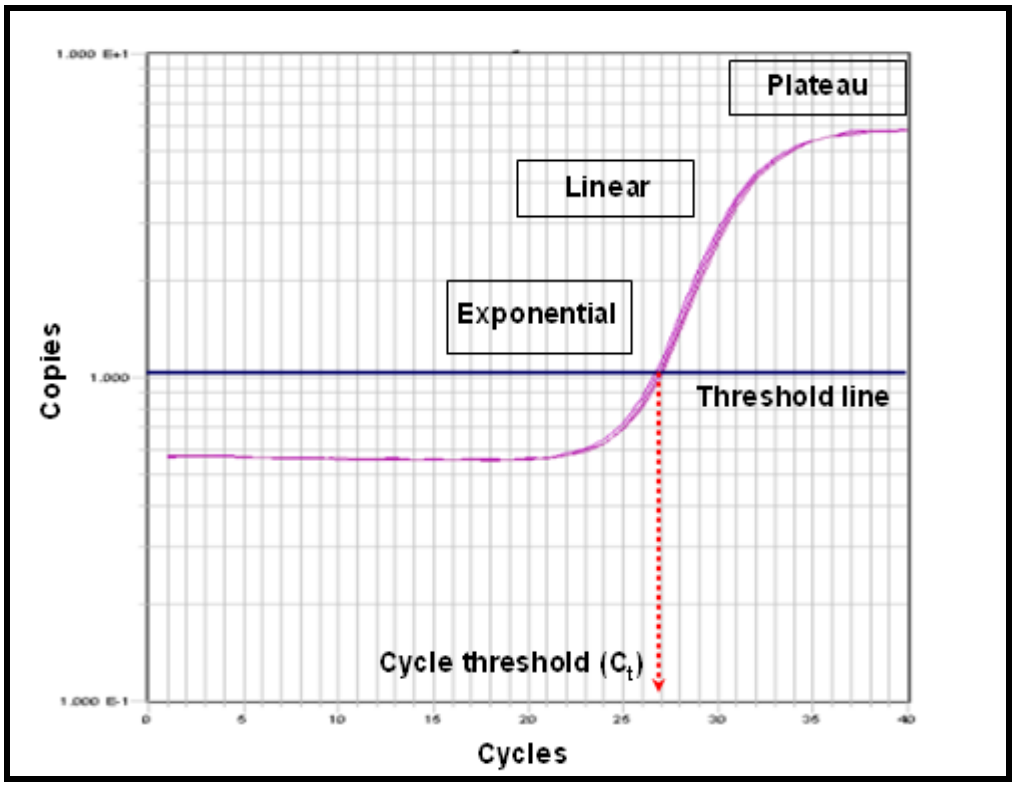
Small RNA (1ng / μL)	5.0 μL
dNTPs (100 mM)	0.17 μL
10X RT Buffer	1.65 μL
Nuclease-free water	4.57 μL
RNase inhibitor (20U / μL)	0.21 μL
Stem-loop primer (50 nM)	3.1 μL
MultiScribe RT (50U / μL)	1.1 μL

Samples were incubated at 16ºC for 30 minutes, 42ºC for 30 minutes and finally 85ºC for 5 minutes to denature the strands. The reaction was performed using a Gene Amp PCR system 9700 thermal cycler (Applied Biosystems). An RT-negative control was included in each batch of reactions.

2.5 Real time quantitative PCR

RQ-PCR allows accumulating amplified DNA/cDNA to be detected and measured as the reaction progresses, *i.e*. in real time. It is possible to detect the amount of amplified product by incorporating a DNA-binding dye or fluorescently-labelled gene-specific probe in the reaction. The RQ-PCR reaction consists of an exponential phase in which the amount of amplified product approximately double during the each cycle of denaturation, primer annealing and template extension, and a non-exponential or plateau phase in which reduced reagents limit the reaction. The point at which enough amplified product has accumulated to produce a detectable fluorescence signal is known as the threshold cycle or C_t and the greater the amount of starting template, the lower the C_t value.

Figure 2.3: **RQ-PCR phases**.
Basic PCR run can be broken up into three phases: **Exponential**: Exact doubling of product is accumulating at every cycle .It occurs because all of the reagents are fresh and available. **Linear**: As the reaction progresses, some of the reagents are being consumed as a result of amplification. **Plateau**: The reaction has stopped, no more products are being made and if left long enough, the PCR products will begin to degrade. The RQ- PCR calculates two values. The Threshold line is the level of detection at which a reaction reaches a fluorescent intensity above background. The PCR cycle at which the sample reaches this level is called the Cycle Threshold, C_t.



2.5.1 Amplification efficiency

In a PCR reaction with optimised primer conditions, reagent concentration *etc.* the amplification efficiency should approach 100% in the exponential phase, *i.e.* a doubling of amplification product for each cycle. To determine the amplification efficiency of the RQ-PCR assay, serial dilution (neat to 10^{-6}) of cDNA template were prepared and amplified using the same conditions used for subsequent gene expression analysis. A dilution curve was constructed by plotting C_t versus the dilution factor of cDNA. Amplification efficiencies (E) were calculated for each RQ-PCR assay using the formula:

$$\text{Amplification efficiencies (E)} = (10\text{-}1/\text{slope} - 1) \times 100$$

Slope = Slope of the dilution curve.

The R^2 Amplification efficiencies value of the dilution curve represents the linearity of the data. R^2 value should be ≥ 0.98 for each dilution curve. A threshold of 10% above and below 100% efficiency was applied to indicate a relatively robust and producible RQ-PCR assay.

2.5.2 mRNA RQ-PCR

RQ-PCR reactions were carried out in final volumes of 20µL using a 7900HT instrument (Fast Real-Time PCR System) (Applied Biosystems), TaqMan gene expression assays and optical 96-well fast plates and sequence detection system (SDS) software (Applied Biosystems). Negative controls were included for each gene target under assay. Reactions consisted of:

First strand cDNA	2.0 µL
TaqMan Fast Master Mix (2X)	10.0 µL
TaqMan Probe (20X)	1.0 µL
Nuclease-free water	1.68 µL

Standard ‘fast’ thermal cycling conditions were applied. This consisted of 40 cycles at 95ºC for 15 seconds and 60ºC for 60 seconds. Complimentary DNA synthesised from commercially available breast cancer cell lines RNA was included on each 96-well plate as an interassay control. All reactions were performed in triplicate. The threshold standard deviation for intra- and inter-assay replicates was 0.3.

2.5.3 miRNA RQ-PCR

RQ-PCR reactions were carried out in final volumes of 20 µL using an AB7900HT. Reactions consisted of:

First strand miRNA-specific cDNA	1.33 µL
TaqMan Fast Master Mix (2X)	10.0 µL
TaqMan Probe (0.2µM)	1.0 µL
Forward primer (1.5µM)	3.0 µL
Reverse primer (0.7 µM)	1.4 µL

Nuclease-free water	3.27 μL

As before, standard fast thermal cycling conditions were used, consisting of 40 cycles at 95ºC for 15 seconds and 60ºC for 60 seconds. On each plate, an interassay control was included to account for any variations between runs.

2.5.4 Endogenous control

Central to the reliable determination of gene expression is the choice of control gene with which to normalise real-time data from target genes. Normalisation can be achieved using endogenous or exogenous controls; however the use of endogenous control (EC) genes is the most widely adopted approach as it excludes variation associated with differences in amounts of template RNA. An ideal EC gene (or genes) should be stably expressed and unaffected by parameters such as disease status and in the case of CRC, should remain unaffected by whether a tissue was derived from normal, adenoma or carcinoma lesions.

B2M and *PPIA* were used as endogenous control (EC) genes to normalise gene expression levels in RQ-PCR reactions measuring gene expression levels [292]. This pair of genes was chosen on the basis that they had been validated as the most stably expressed genes in a large group of colorectal tissues, as will be described in detail in the following chapter. For miRNA expression analysis, the combined expression of *miR-16* and *miR-345* was used to normalise expression data, as previous work in the department of Surgery had validated these miRNAs in colorectal tissue [293].

2.5.5 Relative quantity

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. In order to correct for non-biological variation in gene expression potentially introduced during RQ-PCR process, an endogenous control (EC) gene, which has verified stable expression across samples, is used. QBasePlus was used for

calculation of expression levels of target genes relative to each of the EC genes. It applies $\Delta\Delta C_t$ method where:

$\Delta\Delta Ct$ = (C_t target gene, test sample – C_t endogenous control, test sample) - (C_t target gene, calibrator sample - C_t endogenous control, calibrator sample).

Relative quantities were corrected for efficiency of amplification and fold change in gene expression between groups was calculated as $E^{-\Delta\Delta Ct} \pm$ s.e.m. Where more than one endogenous control are used , fold change estimates were calculated using the geometric mean of EC quantities relative to the calibrator sample which could be the minimum, maximum or a named sample or an average.

2.6 Microarray analysis

Microarray analysis was carried out on total RNA extracted from sections of colorectal FFPE tissues using Megaplex pool A primers (Applied Biosystems) according to the manufacturer's protocol. These TaqMan microfluidic real-time PCR array cards (TLDAs) contained of 384 TaqMan sequence-specific miRNA assays and were prepared in at two-step process as follows:

2.6.1 Megaplex RT reactions:

Total RNA was extracted from paraffin sections as described in section 2.3.2.1. Concentrations of 700 ng (in 3.0 µL volumes of nuclease-free water) were reverse-transcribed using pooled primers and reagents in the Megaplex kit (Applied Biosystems) as follows:

Megaplex RT primers (10X)	0.8 µL
dNTPs with dTTP (100 mM)	0.2 µL
Multiscribe RT (50 U / µL)	1.5 µL
RT buffer (10X)	0.8 µL
$MgCl_2$ (25 mM)	0.9 µL
RNase inhibitor (20 U / µL)	0.1 µL
Nuclease-free water	0.2 µL

Reactions were performed in total volumes of 7.5µL of total RNA and RT reaction mix. Thereafter, samples were incubated for 40 cycles at 16°C for 2

minutes, 42°C for 1 minute and 50°C for 1 second and finally left at 85°C for 5 minutes to denature the strands. The reaction was performed using a Gene Amp PCR system 9700 thermal cycler (Applied Biosystems).

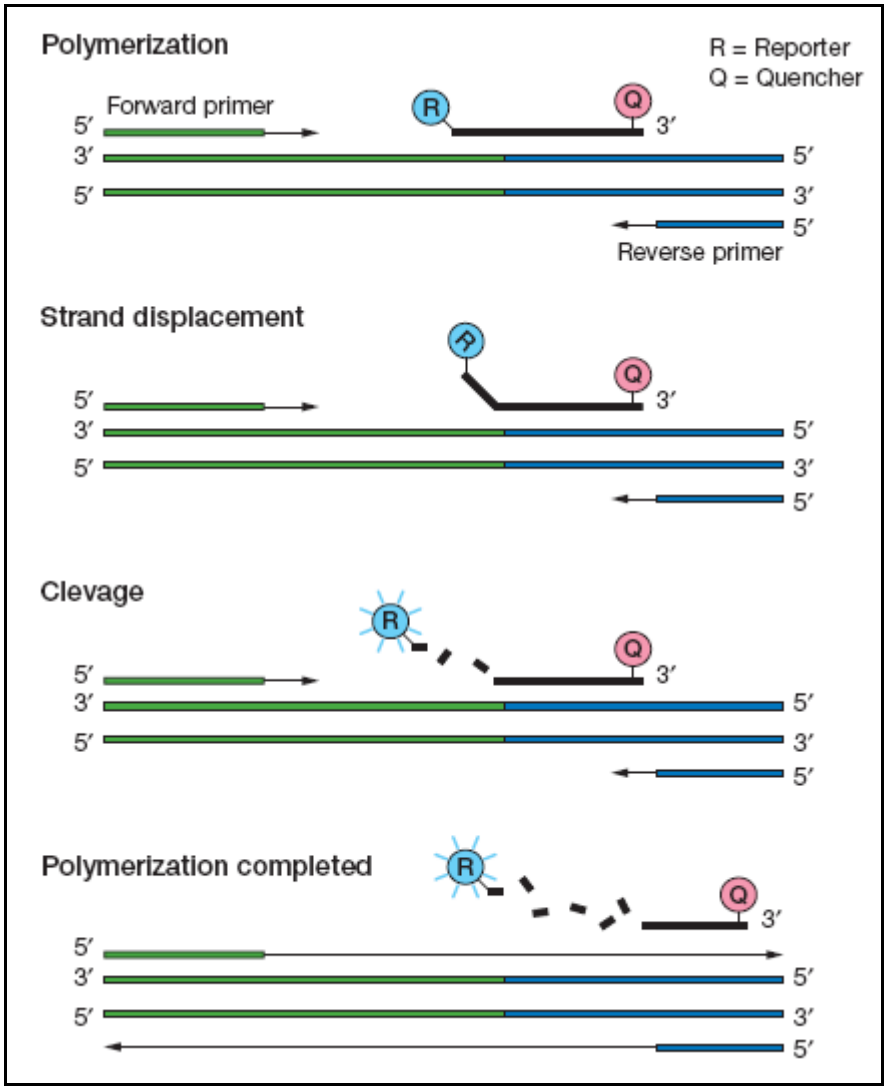
2.6.2 TLDA RQ-PCR reactions:

Reactions mixes (900 µL in total) for samples for TLDA RQ-PCR array profiling using 384-well microfluidic cards were prepared by combining:

TaqMan Master Mix (2X)	450 µL
Megaplex cDNA	6 µL
Nuclease-free water	444 µL

100 µL of the above pre-mix was dispensed into each port of the TLDA card, which was then centrifuged and sealed. Thermal cycling was performed using a 7900 HT instrument (Applied Biosystems) and default thermal cycling conditions of 50°C for 2 minutes and 95°C for 10 minutes then 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. The presence of the targets was detected in real-time through cleavage of TaqMan probes by polymerase 5' – 3' exonuclease activity (figure 2.4).

Figure 2.4: The steps in real-time PCR (the 5´ nuclease assay)
Each TaqMan MGB probe anneals specifically to its complementary sequence between the forward and reverse primer sites. When the hybridized probes are cleaved by AmpliTaq Gold® enzyme, the quencher is separated from the reporter dye, increasing the fluorescence of the reporter dye. Therefore, the fluorescence signal generated by PCR amplification indicates the gene expression level in the sample.



2.6.3 Microarray Data processing

Artificial neural network

Algorithms and architecture:

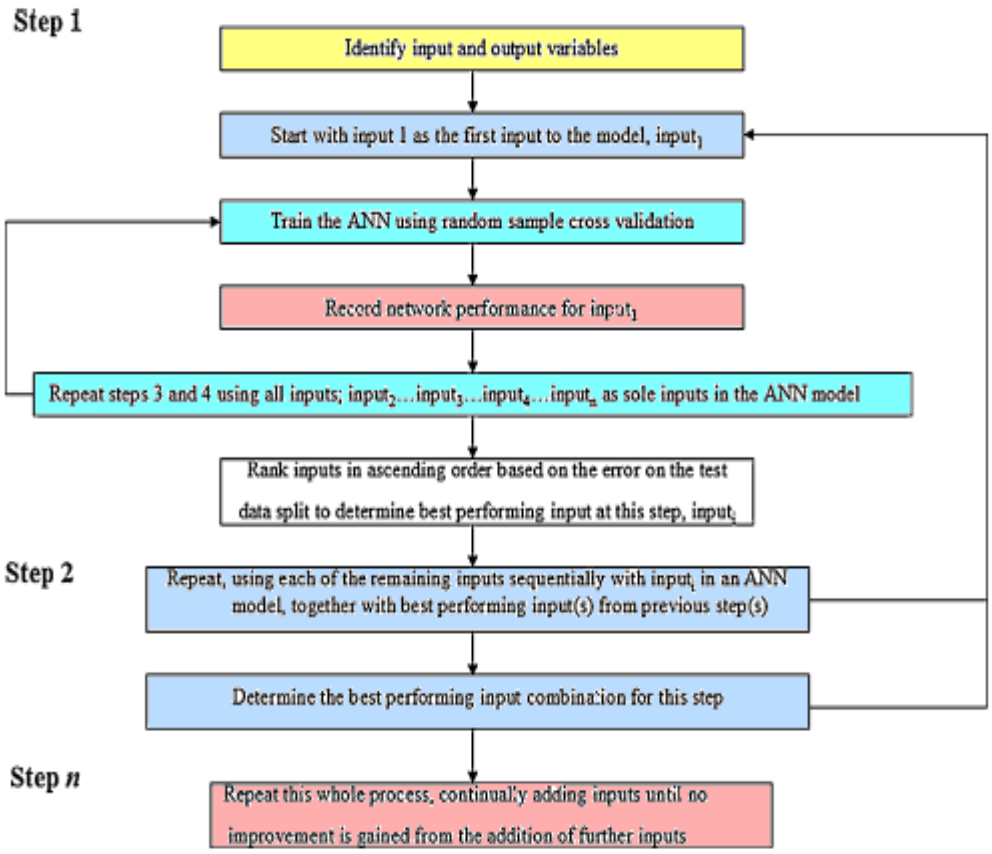
In this study, a three-layer multi-layer Perceptron (MLP) modified with a feed-forward back-propagation algorithm and a sigmoidal transfer function was used. The learning rate and momentum were respectively set at 0.1 and 0.5. An automatic pre-processing normalised the data between 0 and 1 for each variable.

The intensity values for the miRNA for each individual were represented in the input layer, the hidden layer contained 2 hidden nodes, and the class was represented in the output layer coded as 0 for negative and 1 for positive. A randomLy selected subset of the cases developed for training is presented to the network to train it (training data) while it is constantly monitored with a randomLy selected subset of unseen cases (test data). These test data are used to stop the training process once the model has reached predetermined conditions like an optimal error value preventing overtraining. Once training is stopped the efficiency of the model is further assessed by presenting a third, randomLy selected blind subset to the model to determine performance for unseen cases not involved in the training process. This subset selection process was repeated up to 50 times for randomLy selected subsets, a process known as Monte Carlo Cross Validation (MCCV). The suite of 50 models produced was analysed and screened for model optimisation purposes.

Model optimisation:

An additive stepwise approach was employed to identify an optimal set of markers explaining variations in the population for each question explored [294]. The stepwise approach consists of taking each single variable as an input to the ANN, and training 50 sub-models with MCCV. Each single input model subset is then analysed and the median classification performance (based on predictive error for the blind test set) determined. The median performance for all single inputs is then analysed and the inputs ranked accordingly. The best predictor input (with the lowest error) is then selected and a second single variable added, creating a two-input model. This was repeaed for all the variables in the dataset, and the best pair determined again based on classification error. Further inputs are then added in the stepwise fashion (generating 3-input models, 4-input models and so on), until no further improvement is obtained and an obtimal model with the best predictive performance is generated.

Figure 2.5: **Stepwise ANN algorithm modelling process** [294]



2.7 Protein analysis

Immunohistochemistry was used to examine the expression of the DNA mismatch repair (MMR) proteins *hMLH1, hMSH2, hMSH6* and *hPMS2* in colorectal cancer tissues.

2.7.1 Immunohistochemistry (IHC)

IHC is the localisation of antigens in fixed cells by the use of labelled antibodies as specific reagents through antigen-antibody interactions that are visualised by a marker such as an enzyme or a fluorescent label. In most routine IHC methods (*e.g.* DABMap), an unlabelled primary antibody is incubated on the tissue section, binding the antigen of interest. A biotinylated secondary antibody directed against the primary antibody is then applied. Streptavidin-horseradish peroxidise conjugate is used to catalyse the 3, 3'-Diaminobenzidine tetrahydrochloride (DAB)/H_2O_2 reaction. A streptavidin-biotin complex (ABC) that possesses

biotin-binding sites is then added, cross reacts with the biotin molecules on the secondary antibody, amplifying the signal intensity.

UltraMap HRP anti-mouse is biotin-free detection system based on property multimer technology. It consists of robust chemistry that provides clean background in combination with enhanced specificity and sensitivity, which increase the signal-to-noise ratio.

Staining was carried out on 5 µm thick paraffin sections of normal and tumour tissue from each patient, using mouse monoclonal antibodies specific for each of the four human MMR proteins: *hMLH1* (BD PharMingen), *hMSH2* (Calbiochem), *hMSH6* (BD Transduction Laboratories), and *hPMS2* (BD PharMingen) (Table 2.6). HRP-conjugated anti mouse IgG (Dako) was used as the secondary antibodies. Although general tissue processing protocols are similar among laboratories, a single universal protocol is not in place. Closed Loop Assay Development (CLAD) (figure 3.1) was employed to optimize the staining systems. It allows for consistent and reproducible results for both routine and complex projects and empowers the user to optimize development protocols based on crisp morphology, signal intensity and high signal to noise ratio

Table 2.8: Antibodies used for MMR analysis.

Antibody	**Clonality**	**Clone**	**Dilution**	**Detection System**
hMLH1	Mouse Monoclonal	G168-15	1:50	UltraMap
hMSH2	Mouse Monoclonal	FE11	1:300	UltraMap
hMSH6	Mouse Monoclonal	44	1:500	DABMap
hPMS2	Mouse Monoclonal	A16-4	1:20	UltraMap

2.7.1.1 DABMap protocol:

DapMap system was used to detect *hMSH6* protein expression employing the automated Vantana Discovery instrument or manual IHC steps as follows:

1- Deparaffinization: The slides were warmed up to 75°C for 8minutes. The EZPrep and Depar volumes (Ventana) were added and the cover slip was

applied to each slide and incubated for 8 minutes. Then the slides were warmed to 37°C for 2 minutes and rinsed.

2- Cell conditioning: The cells were conditioned by two applications of long cell conditioner (LCC) followed by slides warming up to 95°C for 8 minutes. Then medium cell conditioner (MCC) was applied and the slides were warmed up to 100°C for 4 minutes. This was followed by two applications of cell conditioner (CC), one application of MCC, one application of CC for 4 minutes each. The last cycle of medium and cell conditioner was repeated again before disabling slide heater and incubated for 8 minutes.
3- The slides were rinsed and warmed up to 37°C for 2 minutes. One drop of Inhibitor D (Ventana) was added to each slide and slides were then incubated for 4 minutes, this was followed by disabling the heater and rinsing the slides.
4- A 100 µL aliquot of primary antibody was added to each experimental slide and incubated for 1 hour at room temperature. Experimental slides and positive control slides were kept separate to avoid crossover of antibodies.
5- This was followed by wash and blotting of slides.
6- The HRP-conjugate secondary antibodies (Dako) was then added to each slide and incubated for 30 minutes.
7- Slides were washed, one drop of Blocker D (Ventana) was applied to each slide and incubated for 2 minutes
8- Then one drop of SA-HRP D (Ventana) was applied to each slide and incubated for 16 minutes
9- The slides were then washed 4 times before applying one drop of DAB and one drop of DAB H_2O_2 D (Ventana) to each slide and incubated for 8 minutes. Then Copper D (Ventana) was applied and slides were incubated for 4 minutes.
10- Then the slides were washed twice for 2 minutes before disabling slide heater.

11- Counter-stain: Slides were warmed up to 3°C for 2minutes, washed and then a drop of haematoxylin was applied to each slide and incubated for 2 minutes.
12- Post Counter-stain: Slides were then washed twice and Bluing reagent (Ventana) was applied to each slide and incubated for 2 minutes and again washed and blotted dry.
13- Once Ventana staining was complete, sections were washed in warm soapy water and then dehydrated in serial alchol immersions as follow:
    - Dipped many times in distilled water
    - Immersed for 3 minutes in 70% ethanol.
    - Immersed for 3 minutes in 95% ethanol.
    - Immersed for 3 minutes in 100% ethanol.
    - Immersed for 3 minutes in 100% Xylene
    - Immersed for 3 minutes in 100% Xylene.
14- A layer of DPX mounting medium (Sigma-Aldrich) was added to each slide , followed by the application of cover glass, taking care to avoid bubble formation. Slides were allowed to dry overnight and then examined.

2.7.1.2 UltraMap protocol:

UltraMap anti-MS HRP detection system was used to determine the expression of *hMLH1*, *hMSH2* and *hPMS2* in colorectal cancer tissues. The steps were much similar to DABMap system with differences in reagents.

***Extended UltraMap*:**

It was used for detection of *hMLH1* and *PMS2*.

1- Deparaffinization
2- Extended cell conditioning: the medium cell conditioner (MCC) and cell conditioner (CC) application cycle (1×MCC, 2×CC, 1×MCC, 1×CC) was repeated three times.
3- The slides were rinsed and warmed up to 37°C for 2 minutes. One drop of Inhibitor CM (Ventana) was added to each slide and slides were then incubated for 4 minutes

4- After washing the slides, a 100 µL aliquot of primary antibody was added to each experimental slide and incubated for 12 hours at 37°C.
5- No secondary antibody was added.
6- Thereafter UMap anti-Ms HRP (Multimer HRP) (Ventana) was added to each slide and incubated for 32 minutes. Then slides were washed 4 times and DAB CM and H_2O_2 CM (Ventana) were added and incubated for 8 minutes. This was again followed by two slides washes before applying Copper CM (Ventana) for 4 minutes.
7- Counter-stain, post counter-stain, dehydration and DPX mounting was carried out as for DABMap protocol.

Standard UltraMap:

It was used to detect *hMSH2* protein expression. The steps were similar to these of extended protocol with the exception of cell conditioning which was carried out in two cycles of (1×MCC, 2×CC, 1×MCC, 1×CC) compared to the three cycles in the extended UltraMap system.

2.7.2 IHC analysis

Changes in protein expression following transfection of colorectal tissues were observed in stained cells using Olympus BX60 microscope and image analySIS software. Adjacent normal tissue served as an internal control for positive staining. As a negative control staining was carried out without the primary antibody. MMR protein staining was considered negative when all of the tumour cell nuclei failed to react with the antibody.

2.8 mRNA target prediction

It is thought that functional characterisation of miRNAs will depend heavily on identification of their specific target mRNAs. However, experimental studies have touched on only a handful of the possible ranges of function of miRNAs, and numerous bioinformatics methods have been developed to allow high-throughput prediction of miRNA target genes. Results derived using these computational algorithms have been validated biologically and feedback from validation results

have greatly improved performance of *in- silico* miRNA target prediction algorithms.

2.8.1 Computational target predictions

There are several computational target prediction programmes available (table 2.6), all of which place emphasis on the seed region of the miRNA and the 3' UTR of mRNA sequence. However, they differ in their exact scoring system. Computational prediction of miRNA target sites consists of four main steps:

- Extraction of rules related to formation of miRNA–mRNA Duplexes
- Incorporation of those rules in computational algorithms;
- Prediction of novel miRNA target sites using those algorithms
- Validation of the results, and thus the algorithm itself, using computational and experimental approaches.

For the purpose of this study predicted targets of specific miRNAs were determined by searching the miRBase, miRDB, PicTar and TargetScans for putative mRNA with a known role in colorectal cancer or other cancer-associated signal cascade.

MiRBase [295-297], is a programme which predicts mRNA targets in vertebrates through a fully automated pipeline (figure 2.4), using the miRanda algorithm to identify potential binding sites for a given miRNA. The miRNA sequence is scanned against the 3'-untranslated region (3'UTRs) of all available species in Ensembl [291]. It scores complementary sites between 0 and 100, where 0 represents no complementary and 100 represents complete complementary. For example, matched pairs get positive scores (e.g., +5 for G:C and A:U pairs and +2 for G:U pairs) mismatched pairs get a negative scores (e.g., -3), and there is a gap-opening and a gap-elongation penalty of -8 and -2 respectively. The scoring system is weighted for complementarity at the 5'end of the miRNA. An alignment score (S score) is calculated based on all of these factors. Next the free energy score (ΔG score) of the resulting duplex is computed using the RNA lib package [298]. Cut-offs for S and ΔG score must be met before conservation of the 3'UTR

target site is examined a cross species. For a site to be conserved it must be detected at the same position in a cross-species orthologous UTR alignment by a miRNA of the same family. The position of the target site can be shifted slightly (e.g., ±10 residues), and sequence identity does not have to be perfect (e.g., 90% identity may be required). Each target must be conserved in at least two species for inclusion in the database.

In determining putative mRNA targets for a miRNA, TargetScan [299-301] requires target site conservation in the human, mouse, rat, dog and chicken genome. The requirement of a 7-nucleotide match of the seed region of the miRNA (nucleotides 2-8) has been relaxed to requirement of a 6-nucleotide match comprising nucleotides 2-7. A mRNA is declared to be a target of miRNA if there is a conserved seed match and a conserved anchoring adenosine nucleotide on the 3;UTR downstream of the seed region or conserved m8-t8match, i.e. an A:U or G:C match between the 8 nucleotide of miRNA and the corresponding position on the 3'UTR, or both.

The PicTar [302] algorithm identifies seed matches, which are seven nucleotide segments in the 3'UTR of the mRNA which have perfect Watson-Crick complementarity to the miRNA of interest. The seed match region must start at nucleotide 1 or 2 of the miRNA. Conservation of this region is examined across species. The free energy of the miRNA:mRNA duplex is calculated and compared to a cut-off. A score is calculated for each a alignment and the average of the scores a cross all species is reported in the PicTar predictions.

MiRDB [303, 304] is a free online database for miRNA target prediction and functional annotations. All the targets were predicted by a bioinformatics tool called MirTarget2, which was developed by analyzing thousands of genes downregulated by miRNAs with an SVM learning machine

Table 2.9: Computational algorithms for miRNA target prediction

Software	**URL or availability**	**Supported organisms *(ref)***
TargetScan,	http://genes.mit.edu/targetscan/	Vertebrates [299, 300]
miRanda	http://www.miRNA.org/	Flies, vertebrates [305, 306]
DIANA-microT	http://diana.pcbi.upenn.edu/DIANA-microT/	Vertebrates [307]
RNAhybrid	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/	Flies [308]
PicTar	http://pictar.bio.nyu.edu/	Nematodes, flies, Vertebrates [309, 310]
FastCompare	http://tavazoielab.princeton.edu/mirnas/	Nematodes, flies [311]
rna22	http://cbcsrv.watson.ibm.com/rna22.htmL	Nematodes, flies, Vertebrates [228]

Figure 2.6: Pipeline for miRNA target prediction.
The main steps in identifying miRNA target genes are shown. When miRNA and mRNA (3'UTR) sequences are provided as input data sets, similar data sets from related species are constructed using data on putative orthologs. After preparation of the data sets, miRNA binding sites are identified by determining the base pairing pattern of miRNAs and mRNAs according to the complementarity within specific regions (Step1); determining the strength of the resulting miRNA-mRNA duplex by calculation of the free energy (Step 2); comparative sequence analysis (Step 3); and checking for the presence of multiple target sites per transcript (Step 4). [312]

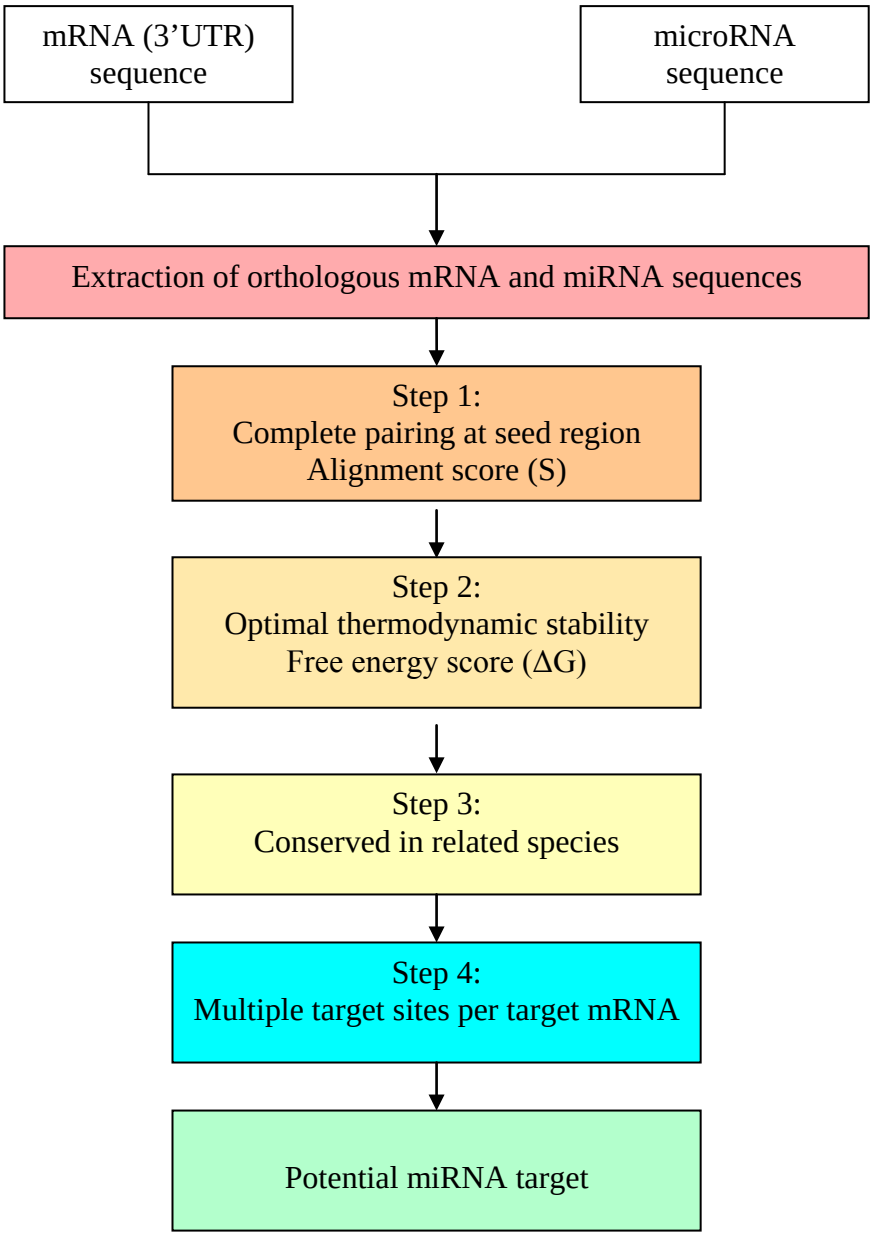
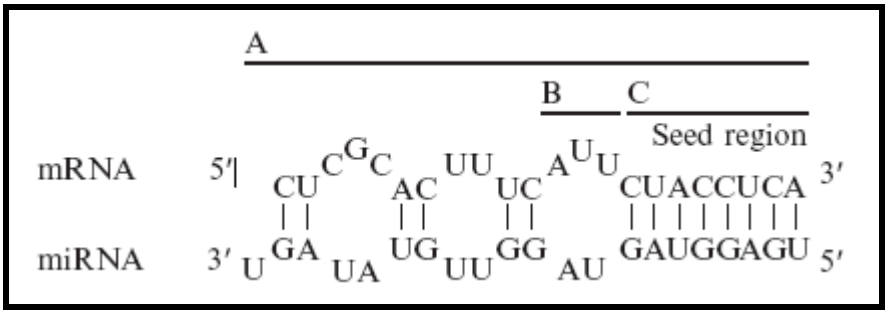


Figure 2.7: **Typical pattern of base pairing between miRNAs and target mRNAs**.
Typically, the miRNA binds to a specific site or sites within the 30UTR region of the mRNA sequence. According to thermodynamic analysis, some degree of complex formation occurs along the entire miRNA-mRNA duplexed region (A). Base pairing is particularly weak in the central region (B) and particularly strong at the 50 end (seed region) of the miRNA (C).These aspects are commonly used to identify putative novel binding sites. Base pairing between let-7 miRNA and hbl-1 mRNA in C. elegans is shown as an example [313]



2.8.2 Correlation of mRNA-miRNA expression levels

The expression levels of the examined mRNA was quantitated by RQ-PCR from colorectal tissues and correlated with miRNA expression levels quantitated by stem-loop RQ-PCR from the same tumour samples. The correlated genes were then checked against the miRNA target databases to see if any of the correlated genes were designated targets worthy of further investigation.

2.9 Statistical Analysis

Statistical analysis was carried out with Minitab 15 (Minitab Ltd) and IBM SPSS Statistics 17.0 (SPSS Inc.). Data was tested for normal distribution graphically using histograms and also using the Kolmogorov–Smirnov, Shapiro-Wilk and Anderson-Darling tests. Parametric tests were used where appropriate. One-way ANOVA and independent t-test were used to determine association and comparisons between independent groups. Correlation analysis used Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively. The correlation data interpretation was carried out following Cohen's guidelines [314] (table 2.7). Univariate analysis and paired-T test were used to assess related samples. The statistical significance of differences in

survival between groups was determined by log rank which compares differences along all points of the curve and multivariate analysis was done using Cox regression. P values <0.05 were considered statistically significant.

Table 2.10: Cohen's guidelines for interpretation of correlation data.

Correlation coefficient (r)	Strength of relationship
0.10 - 0.29	Small
0.30-0.49	Medium
0.50-1.00	Large

Chapter 3:

Normalisation of RQ-PCR data

Chapter 3

3.1 Introduction:

The majority of colorectal tumours originate from adenomatous precursor lesions and develop along a well-defined adenoma-carcinoma sequence. According to this model the culmination of mutational events including activation of oncogenes and loss of function of tumour suppressor genes results in the emergence of carcinomas [315]. Molecular profiling across the spectrum of normal-adenoma-tumour tissue types has yielded many candidate genes in the search for novel molecular diagnostic and prognostic markers and treatment strategies [316-318]. In latter years real-time quantitative (RQ-) PCR has become established as the gold standard for accurate, sensitive and rapid quantification of gene expression [319, 320]. In comparison to alternative methods such as Northern blotting and Ribonuclease Protection Assays (RPA), RQ-PCR has been universally adopted as the transcriptomic method of choice due to its superiority with regard to speed, sensitivity, reproducibility and the wide range of instrumentation and reagents commercially available.

To accurately quantify an mRNA target by RQ-PCR, samples are assayed during the exponential phase of the PCR reaction during which the amount of target is assumed to double with each cycle of PCR without bias due to limiting reagents. Analysis of cycle threshold (C_t), the cycle number at which signals are detected above background, can be used to estimate gene expression levels by relating C_t values either to a standard curve (absolute quantification) or to a control gene (relative quantification). The latter method requires the generation of standard curves of known copy number for each target and so is limited due to logistical issues associated with the generation of standards in studies of multiple gene targets. Relative quantification is the most widely adopted approach and as the name suggests, quantification of gene expression is based on the analysis of a target gene whose expression is normalised relative to the expression of a control gene. Central to the reliable determination of gene expression is the choice of control gene with which to normalise real-time data from target genes. Normalisation can be achieved using endogenous or exogenous controls; however the use of endogenous control (EC) genes is the most widely adopted approach as it excludes variation associated with differences in amounts of template RNA.

Vandestompele *et al* 2002 described a normalisation method whereby geometrical averaging of multiple EC genes improved accuracy [321]. This approach has been adopted to reliably measure levels of gene expression in many studies in different tissue types including breast [322-324], lung [325], kidney [191], brain [326] and liver [327].

An ideal EC gene (or genes) should be stably expressed and unaffected by parameters such as disease status and in the case of CRC, should remain unaffected by whether a tissue was derived from normal, adenoma or carcinoma lesions. Traditionally *GAPDH* (glyceraldehyde phosphate dehydrogenase) has been widely used to normalise RQ-PCR data. A common feature of earlier studies was that the stability of reference gene expression between different sample types was assumed with little consideration paid to validation of these EC genes as suitable normalisers. More recent studies have brought into question the stability of commonly used EC genes such as *GAPDH* on the basis that gene expression levels have been found to vary in response to treatment or as a result of physiological, pathological or experimental changes. For example, alteration in oxygen tension and hypoxia were found to be associated with wide variation in *GAPDH, B-ACTIN* and *CYCLOPHILIN* expression [328]. In addition, *GAPDH* expression was found to be strongly unregulated in diabetic patients and down-regulated in response to the administration of bisphosphonate compounds in the treatment of metastatic breast cancer [329]. Other evidence indicates that neoplastic growth can affect EC expression levels [330]. Goidin *et al* [331] found differences in the expression of *GAPDH* and *B-ACTIN* in two sub-populations of melanoma cells derived from a tumour in a single patient. Treatment agents such as dexamethasone, deprenyl and isatin also affect EC gene expression [332, 333]. Schmittgen *et al* [334] reported increased expression of *GAPDH, B2M, 18S rRNA* and *β-ACTIN* in fibroblasts after the addition of serum: evidence of the effect of experimental conditions on EC expression. These findings were further supported by Wu *et al* [335] in their investigation of the effect of different skin irritants on *GAPDH* and PolyA+ RNA expression. *GAPDH* was found to be involved in age-induced apoptosis in mature cerebellar cells [336] and also as a tRNA binding protein present in the nuclei of HeLa cells [337].

As the use of unreliable ECs can result in inaccurate results, the identification of the most reliable gene or set of genes at the outset of an investigation is critical. Thus far, a pervasive stably expressed gene (or genes) has yet to be identified across all tissue types [338, 339]. This would indicate that the identification of robust ECs at the outset of transcriptomic analysis would yield more reliable and meaningful RQ-PCR data.

3.2 Aims

The aim of this study was to evaluate a panel of thirteen candidate EC genes from which to identify the most stably expressed gene (or genes) to normalise RQ-PCR data derived from primary colorectal tumour and tumour associated normal (TAN) tissue. Six of the candidate EC genes were selected from the literature and represent the most frequently studied reference genes in cancer including, but not limited to, colorectal cancer. Each gene was previously reported as being constitutively expressed in various tissues. These EC genes included *B2M (beta-2-microglogulin) [318], HPRT (hypoxanthine guanine phosphoribosyl transferase 1) [316, 340], GAPDH [341], ACTB (beta-actin) [342], PPIA (peptidyl-prolyl isomerise A) [322]* and *MRPL19 (mitochondrial ribosomal protein L19) [322]*. The remaining seven genes included *HCRT, SLC25A23, DTX3, APOC4, RTDR1, KRTAP12-3, and CHRNB4.* The latter candidates were selected from an unpublished whole genome microarray dataset of 20 human tumour specimens and represented the most stably expressed probes with a fold-change of 1.0-1.2, ($p< 0.05$). Expression of *CXCL12* [343], *FABP1* [344], *MUC2* [345] and *PDCD4* genes were chosen as targets against which to measure the effects of candidate EC expression on the basis of their previously identified roles in tumourigenesis. In addition to its tumour suppressor properties, *PDCD4* [346] also has diagnostic and prognostic utility and represents a promising target for anti-cancer therapy.

3.3 Materials and methods

3.3.1 Study group

A study group of 64 biopsies of human colon tissue samples was gathered from consenting patients at the time of primary curative surgical resection at Galway University Hospital, Ireland. The cohort comprised of 30 colorectal tumour specimens and 34 and tumour-associated normal (TAN) tissues. Following excision, all samples were subject to histopathological review prior immediate snap-freezing in liquid nitrogen and archival at -80ºC until further use. Concomitant clinicopathological data on patients and specimens was obtained from the Department of Surgery Biobank, NUI Galway as detailed in Table 4. Ethical approval for this study was granted by the Clinical Research Ethics Committee, Galway University Hospitals.

Table 3.1: Clinicopathological data of study group

Clinicopathological Variable	Number of Patients N=42
Gender	
Males	29
Females	13
Mean Age (SD)	66.5 (12.84)
Tumour Location	
Colon	12
Rectum	30
Tumour Diameter (mm)	
<10	11
10-15	15
>15	17
Tumour Thickness (mm)	
<30	15
30-40	12
>40	15
Distant Metastasis	
M_0	36
M_1	6
Nodal Status	
N_0	22
N_1	11
N_2	9
UICC Stage	
Stage 0	6
Stage I	10
Stage II	10
Stage III	11
Stage IV	5
Tumour Differentiation	
Well	12
Moderate	24
Poor	6
Mucin Secretion	
Mucinous	8
Non-mucinous	34

3.3.2 RNA Extraction and Analysis

Tissue samples (50-100 mg) were homogenised using a hand-held homogenizer (Polytron PT1600E) in 1-2 mL of QIAzol reagent (Qiagen, Crawley, UK). To minimise variation in sample processing, tumour and TAN samples were homogenised separately, but on the same day. RNA was extracted as previously described (Davoren *et al*) using the RNeasy Plus Mini Kit and RNeasy MinElute cleanup kit (Qiagen) according to the manufacturer's instructions. Briefly, large (> 200 nt) and small RNA (< 200 nt) fractions were isolated separately. For this study, only large RNA was utilised for further analysis. RNA was eluted in 60µl volumes and stored at -80^{o}C.

RNA concentration and purity was assessed in duplicate samples using a using a NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies). RNA integrity was evaluated using the RNA 6000 Nano Chip Kit (Series II) and the Agilent 2100 Bioanalyzer System (Agilent technologies). An RNA integrity number (RIN) was generated for each sample using the Agilent 2100 Expert Software (Version B.02.03) based on the ratio of ribosomal bands and also the presence or absence of degradation products on the electrophoretic and gel-like images. A threshold value of RIN $\geq$ 7 was applied and RNA purity was verified by an average A260/A280 ratio of 1.98 (range 1.97-2.01) and A260/A230 ration of 1.7 (range 1.5-1.83).

3.3.3 Candidate Endogenous Control Genes

Based on literature search six commonly used candidate endogenous control genes were selected for analysis: *ACTB, GAPDH, HPRT, B2M, PPIA* and *MRPL19*. An additional panel of seven genes: *HCRT, SLC25A23, DTX3, APOC4, RTDR1, KRTAP12-3 and CHRNB4*, was also selected for analysis (table 3.2). To our knowledge all genes have independent cellular functions and were assumed not to be co-regulated.

Table 3.2: Candidate EC genes and their PCR amplification efficiencies (E)

EC	Function	Location	Assay Identifier*	E (%)
B2M	Defence and immunity	15q21-22.2	Hs00187842_m1	101.8
GAPDH	Oxidoreductase, dehydrogenase	12p13	Hs99999905_m1	99.8
PPIA	Isomerase	7p13	Hs99999904_m1	96.6
HPRT	Glycosyl transferase	Xq26.1	Hs99999909_m1	97.9
MRPL19	Protein biosynthesis	2q11.1-11.2	Hs00608519_m1	102.2
ACTB	Cytoskeletal structure	7p15-12	Hs99999903_m1	95.2
DTX3	Signals transduction	12q13.3	Hs00400987_m1	99.1
SLC25A23	Mitochondrial carrier	19p13.3	Hs00225469_m1	97.8
CHRNB4	Nicotinic receptor	15q24	Hs00609523_m1	103.6
RTDR1	Aminopeptidase transport	22q11.2	Hs00205353_m1	UD
HCRT	Homeostatic regulator	12q21	Hs00533664_m1	UD
APOC4	Apo-lipoprotein	19q.2	Hs00155791_m1	UD
KRTAP12-3	Acetylgalactoa-minyltransferase	3q25	Hs01651247_s1	UD

* Applied Biosystems TaqMan® gene expression assay ID.

UD: undetermined

3.3.4 cDNA Synthesis and RQ-PCR

First strand cDNA was synthesised using Superscript III reverse transcriptase (Invitrogen) and random primers (N9; 1µg, MWG Biotech, AG, Ebersberg). Negative control samples were included in each set of reactions. Reactions were incubated at 25º C for 5 minutes followed by 50º C for 1 hour and final denaturation at 72º C for 15 minutes. Samples were subsequently diluted to 50 µL in nuclease-free water and stored at -20º C. The expression of each EC gene was analysed by RQ-PCR using TaqMan gene expression assays using a 7900HT

instrument (Applied Biosystems). All reactions were performed in 20 µL reactions, in triplicate within the same PCR run. Negative controls were included for each gene target under assay. On each plate, an interassay control was included to account for any variations between runs. For each well 2µl of cDNA from each sample was added to 18µl of PCR reaction mix which consisted of 10x TaqMan universal master mix, No AmpErase UNG, 7X nuclease free water and 1X gene expression assay primer-probe mix (Applied Biosystems). The PCR reactions were initiated with a 10 minute incubation at 95º C followed by 40 cycles of 95º C for 15 seconds and 60º C for 60 seconds, in accordance with the manufacturer's recommendations.

3.3.5 PCR Amplification Efficiency

Amplification efficiencies for each EC gene assay were calculated applying the formula E= (10-1/slope - 1) × 100, using the slope of the plot of Ct versus log input of cDNA (10-fold dilution series). A threshold of 10% above and below 100% efficiency was applied. PCR amplification efficiency for each candidate EC gene is shown in table 3.1

3.3.6 Data Analysis

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. QBasePlus was used for calculation of *PDCD4* expression relative to each of the EC genes. It applies $\Delta\Delta$ C_t method was used where $\Delta\Delta Ct$ = (C_t target gene, test sample – C_t endogenous control, test sample) - (C_t target gene, calibrator sample - C_t endogenous control, calibrator sample). Relative quantities were corrected for efficiency of amplification and fold change in gene expression between groups was calculated as E-$\Delta\Delta Ct$ ± s.e.m. Where more than one endogenous control are used , fold change estimates were calculated using the geometric mean of EC quantities relative to the calibrator sample which could be the minimum, maximum or a named sample or an average.

Stability of the EC genes expression was evaluated with two freely available statistical models, geNorm and NormFinder. It is further validated with qBasePlus. Statistical analysis was carried out with Minitab 15 (Minitab Ltd). Anderson-Darling normality test was applied and parametric tests were used where appropriate. The equivalence test was used to assess the equivalently of expression of the candidate genes between tumour and normal tissues. One-way ANOVA, two-sample t-test, Levene's test and Spearman and Pearson correlations were used to determine association and comparisons between groups. P values <0.05 were considered statistically significant.

3.4 Results

3.4.1 Range of Expression of Candidate EC Genes

A range of C_t values was observed across the candidate EC genes in tumour and TAN tissue from CRC patients as indicated in table 1. Only samples with a standard deviation < 0.3 from the mean C_t of the triplicates were included for further analysis. The expression of *RTDR1, HCRT, APOC4* and *KRTAP12-3* could not be determined in all 64 tissue samples, resultantly these candidates were excluded from further analysis.

Mean C_t values for the remaining genes ranged from 19.48 (± 0.14 s.e.m) for *B2M* to 32.30 (± 0.19 s.e.m) for *CHRNB4*. *B2M* displayed the narrowest range of C_t values between 17.5 and 21.5 (mean 19.5 ± 0.14 s.e.m, range of 4.04) followed by *PPIA* and *MRPL19*, while *ACTB* had the widest range of C_t values between 33.8 and 21.1. The genes broadly fell into three categories, those least abundant genes with mean C_t values of 27-32 (*SLC25A23, MRPL19, DTX3* and *CHRNB4*), moderately abundant genes with mean C_t values of 22-26 (*HPRT* and *ACTB*) and the most abundant highly expressed genes with mean C_t values of 19-21 (*B2M, PPIA* and *GAPDH*) (table 3.2).

Table 3.3: Ct values of candidate EC genes and target genes
B2M and *PPIA* were the most abundantly expressed genes, having the lowest mean Ct values while *MRPL19* was the least abundantly expressed gene with average Ct values > 26. Both *B2M* and *PPIA* had the lowest range in their Ct values.

EC Gene	C_t Range	C_t Min	C_t Max	Mean C_t ± s.e.m	Standard deviation (SD)
B2M	4.03	17.47	21.51	19.48 ± 0.14	1.04
PPIA	4.13	17.78	21.91	19.90 ± 0.14	1.06
GAPDH	5.80	18.51	24.32	21.00 ± 0.17	1.29
ACTB	12.74	21.08	33.32	25.14 ± 0.34	2.61
HPRT	8.54	22.74	31.28	26.68 ± 0.25	1.89
DTX3	6.6	24.95	31.56	28.62 ±0.17	1.37
SLC25A23	7.26	24.48	31.74	27.36 ±0.19	1.54
CHRNB4	9.40	27.99	37.38	32.30 ±0.19	2.15
RTDR1	-	30.59	UD	35.82 ±0.36	2.15
HCRT	-	33.96	UD	38.46 ±0.29	1.67
APOC4	-	UD	UD	-	-
KRTAP12-3	-	33.16	UD	36.95 ±0.19	1.46
MRPL19	4.10	26.70	30.80	28.62 ± 0.13	0.98
CXCL12	13.54	21.85	35.39	25.77 ± 0.32	2.61
FABP1	15.57	16.61	32.19	20.83 ± 0.40	3.24
MUC2	17.71	17.62	35.33	22.43 ± 0.53	4.16
PDCD4	11.92	21.35	33.27	24.56 ± 0.32	2.59

UD: undetermined

3.4.2 Identification of Optimal EC genes

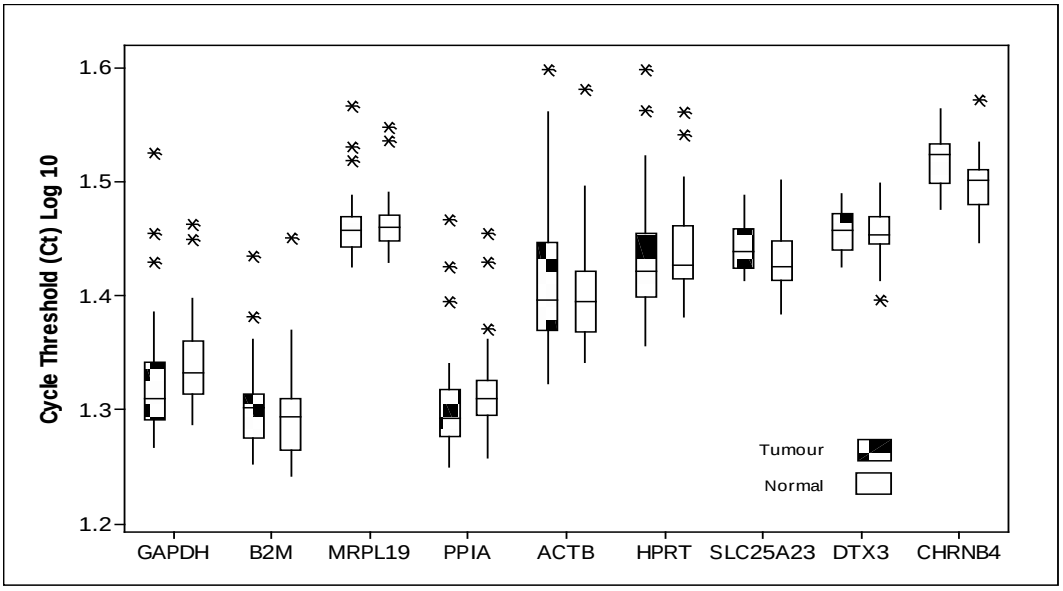
Scaled expression levels across the remaining nine candidate ECs analysed (figure 3.1) indicated within-gene differences in expression between tumour and normal tissue groups in both *SLC25A23* (p= 0.040) and *CHRNB4* (P=0.002) but not in the remaining genes (p>0.05), (figure 3.1A). Therefore, *SLC25A23* and *CHRNB4* genes were excluded from further analysis. Significant differences in variance of EC expression were identified using Levene's test (p<0.001, figure 2B). These findings necessitated further evaluation of each candidate EC gene prior to their

possible use to accurately quantitate gene expression levels of the target genes *CXCL12, FABP1, MUC2* and *PDCD4*.

The stability of candidate EC genes was analysed using geNorm [321] and NormFinder [347] programmes. Stability was further evaluated using qBasePlus [321, 348], a commercially available RQ-PCR data mining package. These programmes were used to calculate amplification efficiency-corrected relative quantities from raw fluorescence data. The ranking of candidate EC genes as determined by each of these programmes is illustrated in Table 3. In the case of GeNorm the variable V indicating the pairwise variation (Vn/Vn+1) between two sequential normalisation factors (NFn/NFn+1) indicated that three EC genes was the optimal number of genes for accurate normalisation (figure 3.2), however, target genes expression did not differ significantly if two rather than three EC genes were used (figure 3.3). Use of all three programmes confirmed that *B2M* and *PPIA* was the best combination of genes for normalising RQ-PCR data in CRC tissues (table 3.3). The Equivalence test [349] was used to examine the expression of candidate ECs. All genes were equivalently expressed between the normal and tumour colorectal tissues using a fold cut-off of 2 (figure 3.4).

Figure 3.1: Scaled expression levels and variation of each candidate EC gene. (A) Log 10 of cycle threshold of candidate EC genes *ACTB*, *B2M*, *GAPDH*, *HPRT*, *MRPL19*, *SLC25A23*, *DTX3*, *CHRNB4* and *PPIA* in tumour and normal colorectal tissues. Boxplot shows interquartile range box, median, range whiskers and outliers (*). Within gene, differences were found in expression between tissue groups in both *SLC25A23* (p=0.040) and *CHRNB4* (p=0.002) but not the other genes (p>0.05) (ANOVA). (B) Variation associated with EC gene expression. There was a significant difference in variation associated with gene expression (p<0.001) with ACTB, *GAPDH* and *HPRT* showing greater variation than *B2M*, *MRPL19* or *PPIA*. *DTX3*, *CHRNB4* and *SLC25A23* showed the least variations (Levene's test).

(A)



(B)

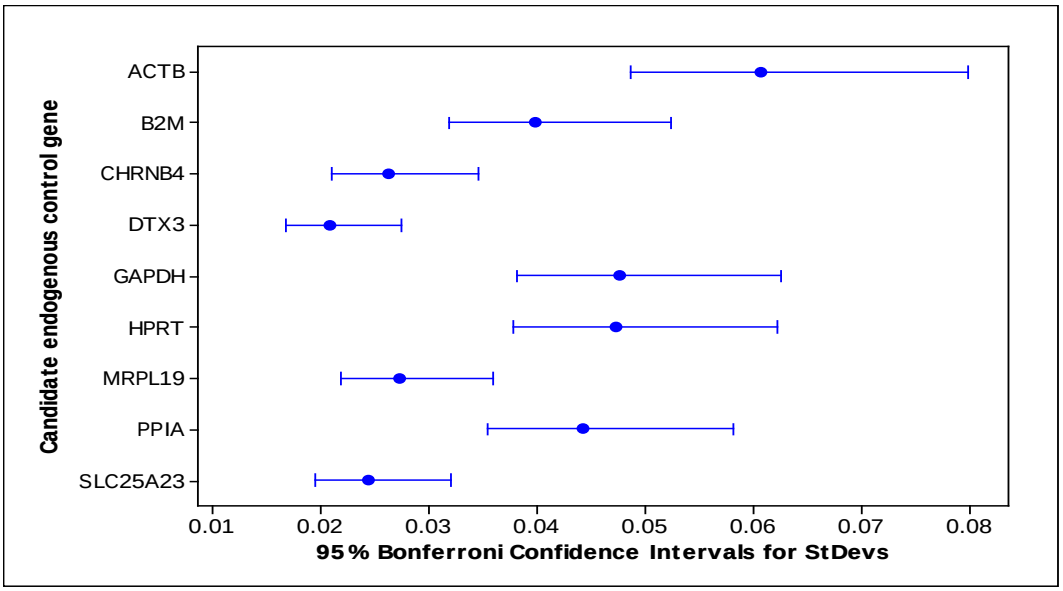
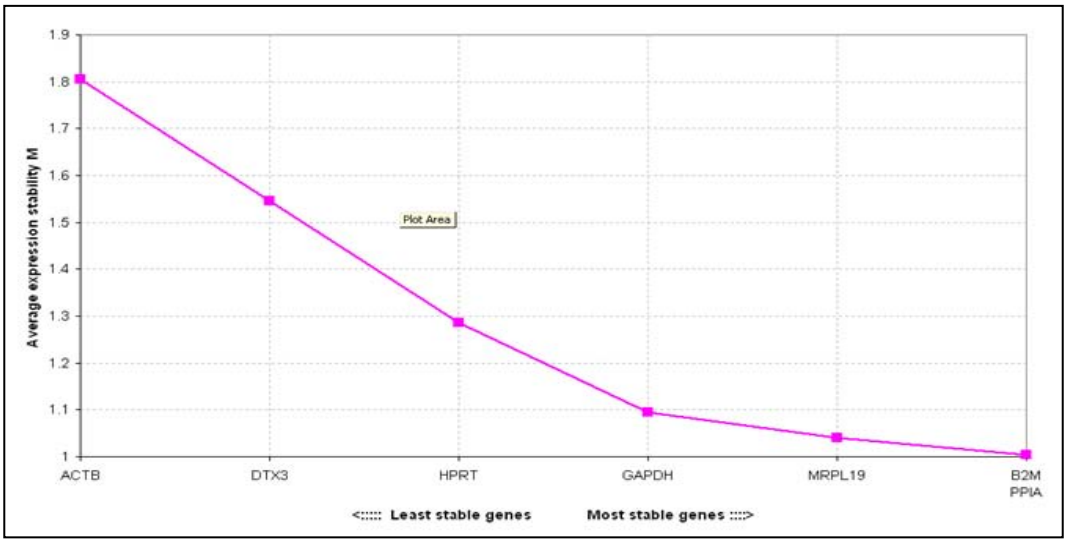


Figure 3.2: Analysis of candidate EC genes using geNorm.
(A): Average expression stability values of eligible EC genes. Expression stability of the control genes as calculated by geNorm. Stability value M is based on the average pair-wise variation between all genes. The least stable gene with highest M value was excluded and M value recalculated till end up with the most stable pair. (B): Determination of optimal number of control genes for normalisation. The GeNorm programme calculates a normalisation factor (NF) which is used to determine the optimal number of EC genes required for accurate normalisation. This factor is calculated using the variable V as the pairwise variation (Vn/Vn + 1) between two sequential NFs (NFn and NFn + 1). To meet the recommended cut off V-value which is the point at which it is unnecessary to include additional genes in a normalisation strategy. The recommended limit for V value is 0.15 but it is not always achievable. In this instance, the GeNorm output file indicated that the optimal number of genes required for normalisation was three.

(A)



(B)

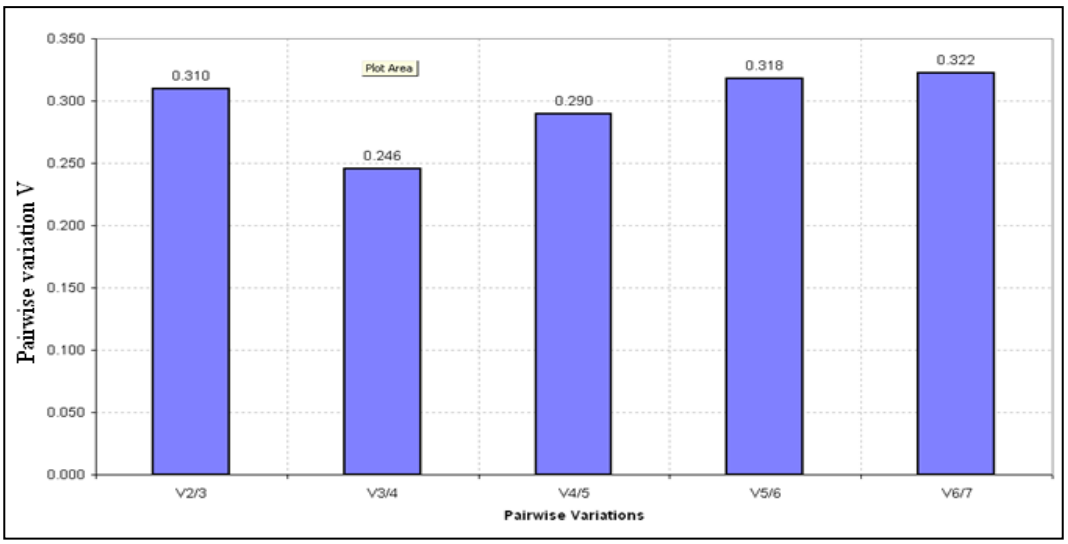


Figure 3.3: RQ of *CXCL12, FABP1, MUC2* and *PDCD4* in CR tissue.
Error bars indicate 95% confidence intervals. No significant differences in the relative quantities of target genes were found using a combination o*f PPIA* and *B2M* (PB) genes in comparison to the use of combination of *PPIA, B2M* and *MRPL19* (PBM) EC genes (ANOVA). (A) Target gene expression in tumour vs. normal colorectal tissues using *PPIA* and *B2M* (PB) two gene combination compared to the *PPIA, B2M, MRPL19* (PBM) three gene combination. (B) Target gene expression levels in all tissues using the *PPIA, B2M* (PB) two gene combination compared to the *PPIA, B2M, MRPL19* (PBM) three gene combination. (C) Estimation of error of target gene expression when normalised to the *PPIA, B2M* (PB) combination compared to the *PPIA, B2M, MRPL19* (PBM) three gene combination.

(A)

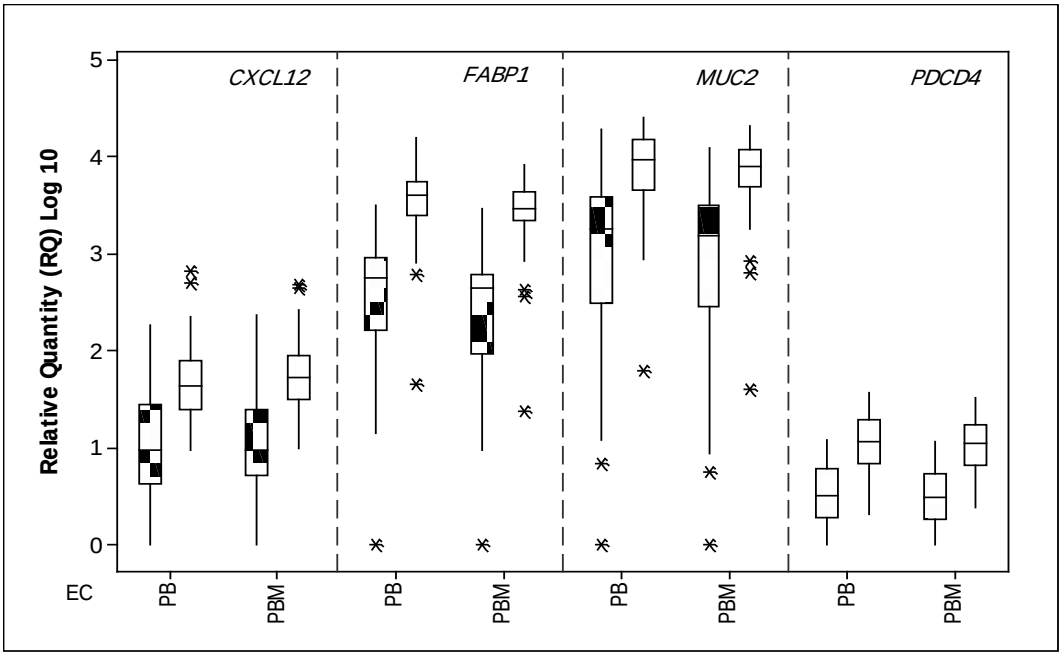
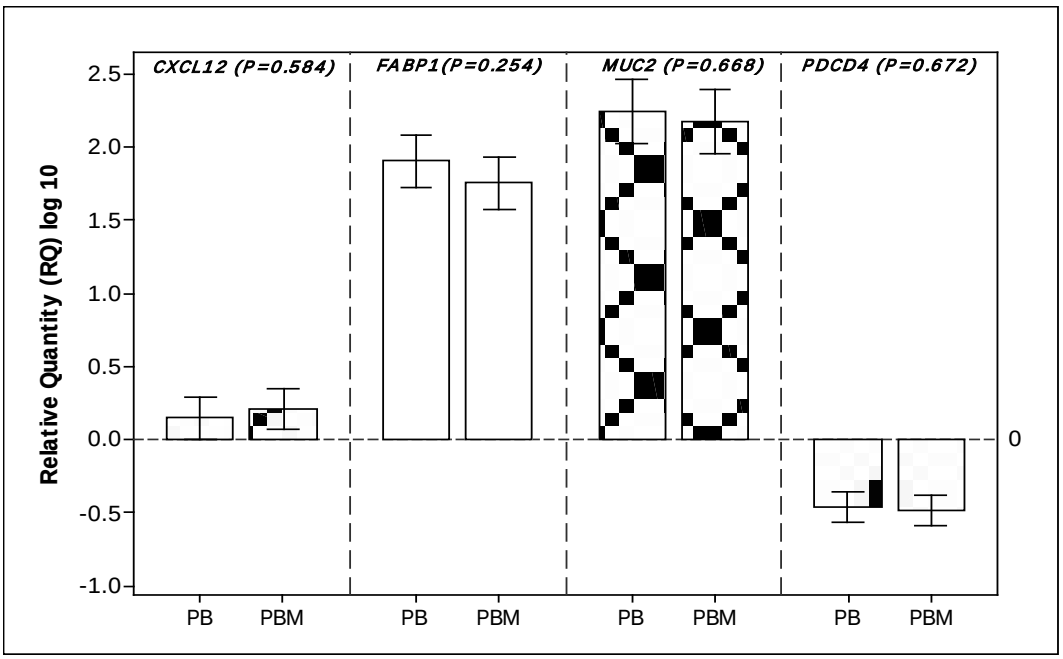


Figure 3.3: continued.

(B)



(C)

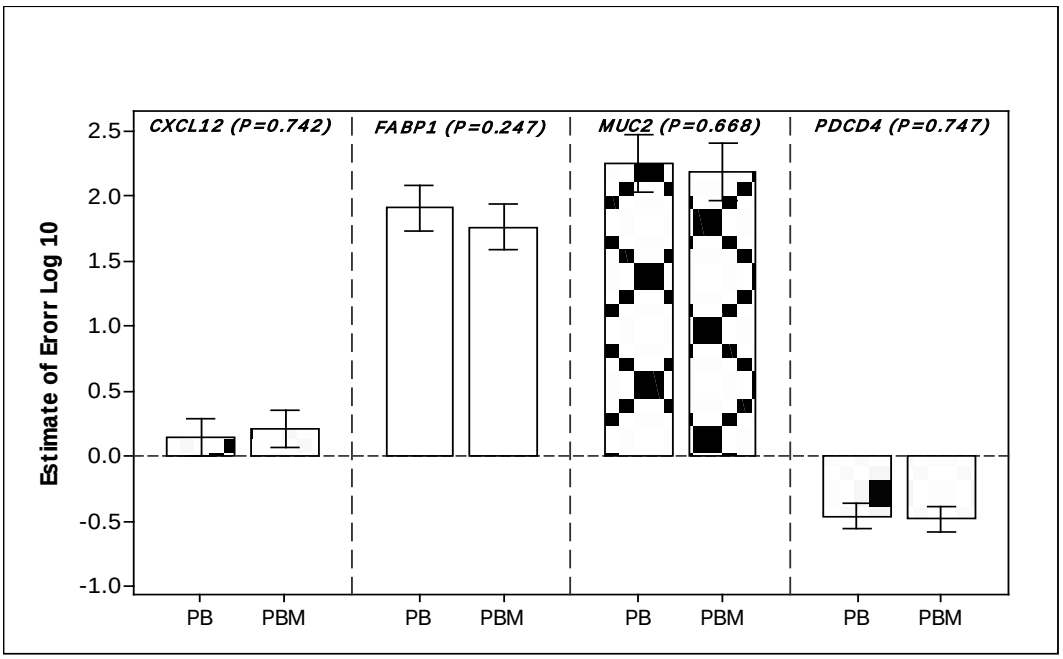


Figure 3.4: Equivalence test for candidate control genes in colorectal tissue. Differences in logarithmic expression levels between tumour and normal tissues (●) are indicated. The upper and lower bars of each line indicate the upper and lower limits of the symmetrical confidence intervals, respectively. The deviation area (-1, 1) for a fold change of 2 or less is plotted as a continuous line while the deviation area of (-1.58, 1.58) for a fold change of 3 is plotted as a dotted line.

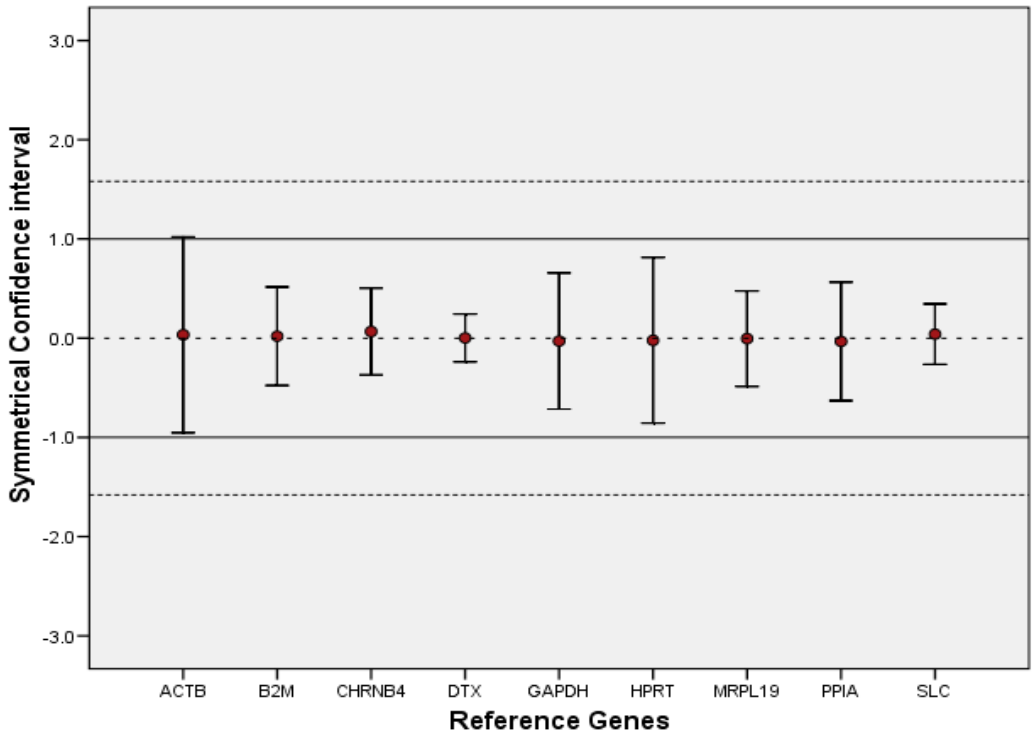


Table 3.4: Ranking and best combination of EC genes
As determined by geNorm, NormFinder and qBasePlus**.** For GeNorm, lower stability values (M) indicate greater stability. In the case of NormFinder, stability is calculated from inter- and intra-group variation. By grouping the tissues into tumour and normal the best combination of genes was identified. For geNorm stability was based on the estimation of pair-wise variation. QBasePlus through its components, geNorm and qBase, identified coefficient of variation (CV) and stability (M) values and thereby the best combination of genes for normalisation only when more than one gene is used.

Rank	**GeNorm**		**NormFinder**		**qBaseplus**	
	Gene	**Stability (M)**	**Gene**	**Stability (M)**	**Gene**	**CV value**
1	*GAPDH*	1.477	*MRPL19*	0.008	*GAPDH*	0.555
2	*MRPL19*	1.467	*B2M*	0.015	*PPIA*	0.659
3	*PPIA*	1.535	*HPRT*	0.016	*HPRT*	0.775
4	*B2M*	1.636	*PPIA*	0.017	*MRPL19*	0.914
5	*HPRT*	1.813	*GAPDH*	0.018	*B2M*	0.923
6	*DTX3*	2.251	*DTX3*	0.020	*ACTB*	0.957
7	*ACTB*	2.454	*ACTB*	0.026	*DTX3*	5.829
Best Combination	***B2M / PPIA***	**1.005**	***B2M / PPIA***	**0.007**	***B2M / PPIA***	**0.460**

3.4.3 Association between EC genes and target genes

There was a significant effect of the expression of the candidate EC genes on relative expression of *CXCL12* ($p<0.001$), *FABP1* ($p<0.001$), *MUC2* ($p<0.001$) and *PDCD4* ($p<0.001$) (figure 5A and 5B). Moreover, a significant effect of the choice of EC with regard to the estimation of error (figure 3.5) was also detected. These findings were further confirmed for each EC gene compared to each other by ANOVA Tukey post hoc tests (table 3.5). The combined use of *B2M* and *PPIA* significantly reduced the magnitude of error in comparison to the use of either gene individually for both *CXCL12* and *PDCD4* expression. The addition of a third EC gene (*MRPL19*) to the *B2M* / *PPIA* combination did not result in any further improvement of the estimation of error (figure 3.3).

Figure 3.5: RQ of target gene expression in CR tissues
Relative quantity of target gene expression in colorectal tissues relative to each EC gene and to the geometric mean of the combined use of *PPIA* and *B2M* (PB). (A) Target gene expression in tumour versus normal using either individual candidate EC genes or the PB combination. (B) Significant differences in relative gene expression values as determined using ANOVA to compare mean expression levels across all tissues using either individual EC genes or PB in combination. (C) One way ANOVA indicating a reduction in the magnitude of error when the PB combination was used to normalise expression of *CXCL12* ($p<0.001$) and *PDCD4* ($p<0.001$) in comparison to the use of individual EC genes. See Table 1 Additional files for Post Hoc tests. Error bars indicate 95% confidence intervals.

(A)

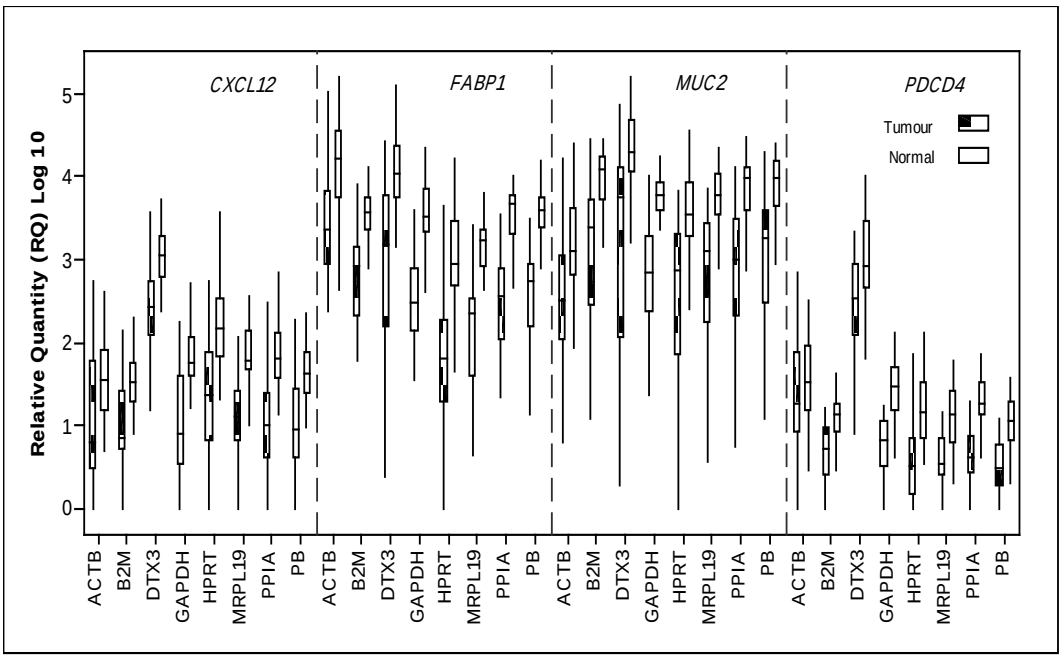
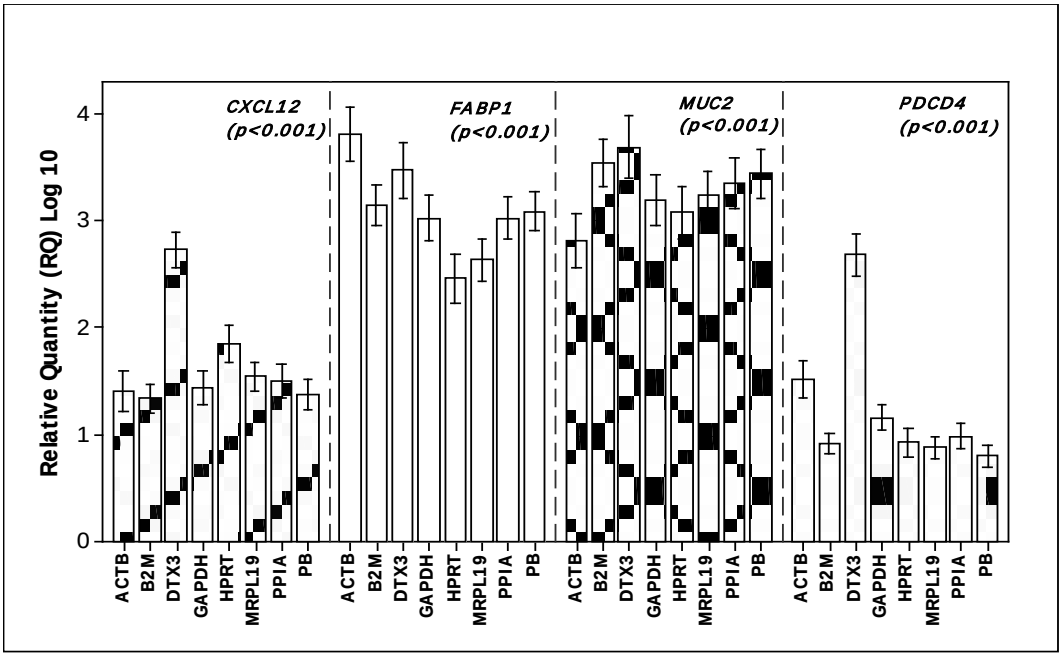
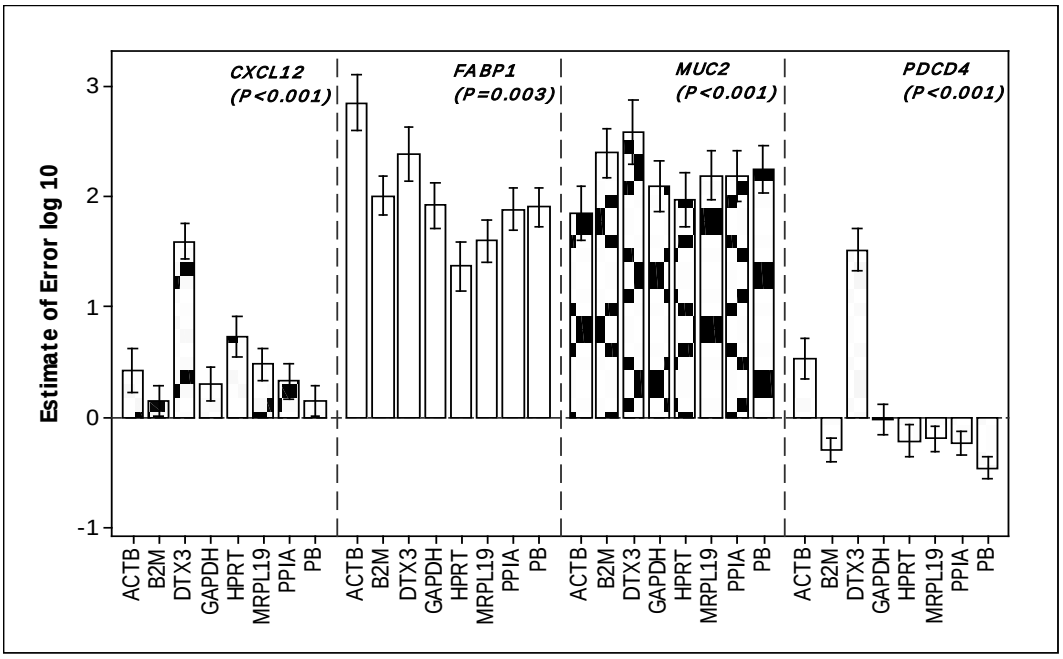


Figure 5.3: continued.

(B)



(C)



3.4.4 Non-normalised expression levels of target genes

To assess whether normalisation was necessary in a large cohort such as this in which the biological effect of the target genes is already established, we compared

the expression of the four target genes in tumour *vs*. normal tissues using non-normalised cycle threshold (C_t) values in the entire sample set (n=64) and in a sub-set of randomLy selected 10 normal and 10 tumour tissues (n=20). This analysis showed down-regulated target gene (*CXCL12, FABP1, MUC2* and *PDCD4*) expression in tumour compared to normal tissues (figure 3.6), in keeping with their documented tumour suppressor functions, when using the larger set of samples. The unchanged target gene expression levels in the large cohort could be explained by the fact that in larger sample sizes the biological milieu may diminish subtle variations in individual samples. In contrast, when the smaller sample size was used, no significant differences in target gene expression were observed. Furthermore the expression levels of *PDCD4* appeared slightly higher in tumours than in normal tissues. When the same subset of 20 samples were normalised with *PPIA/B2M*, significant differences in target gene expression were observed.

Figure 3.6: Non-normalised C_t of *target genes* in CRC
Using this approach, the expression of *CXCL12*, *FABP1*, *MUC2* and *PDCD4* appeared to be down-regulated in tumours compared to normal tissues in the large cohort of patients (30 tumour and 34 normal tissue specimens), similar to previous published reports of reduced expression in colorectal tumours. No significant differences were noted in expression levels of target genes when using the small cohort of patients (10 tumour and 10 normal tissue specimens) (2-sample t-test). This confirms the effect of sample size on findings when using non-normalised C_t values and therefore the importance of normalisation especially in such type of studies

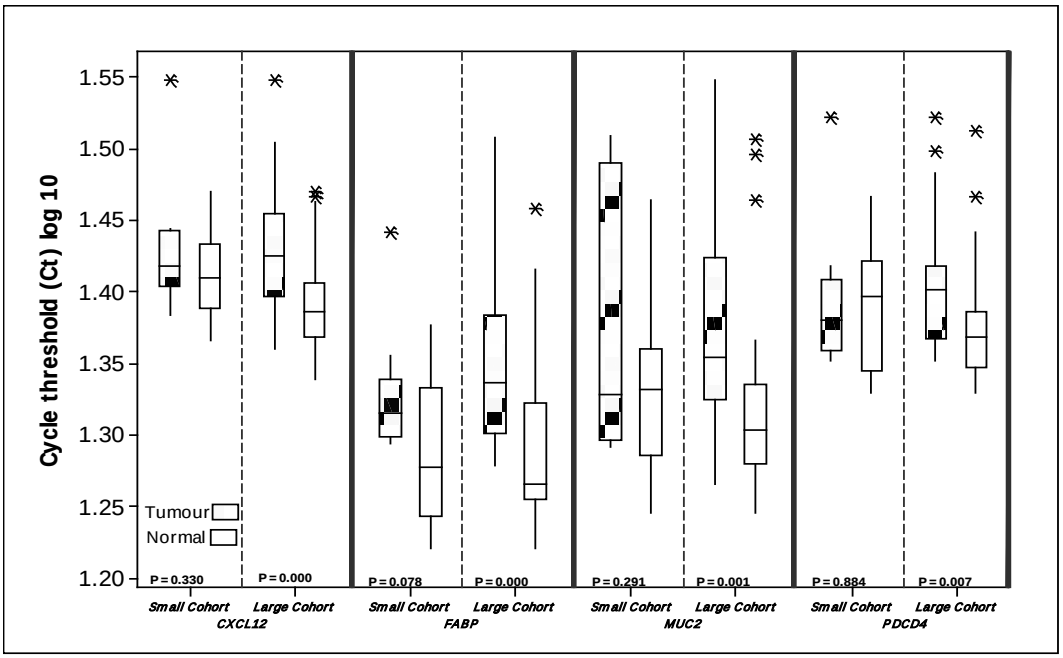


Table 3.5: Target genes relative quantities and estimation of error.
Differences in relative quantities and estimation of error using each of the EC genes and the combination of PPIA and B2M. *P value* of significant differences in expression (A) and estimation of error (B) of CXCL12, FABP1, MUC2 and PDCD4 using each of the candidate EC gene compared to the others. (ANOVA Tukey post hoc tests)

(A)

Targets	CE genes	ACTB	B2M	DTX3	GAPDH	HPRT	MRPL19	PPIA	PPIA & B2M
CXCL12	***DTX3***	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001
	HPRT	0.003	<0.001	<0.001	0.008				0.001
FABP1	***ACTB***		<0.001		<0.001	<0.001	<0.001	<0.001	<0.001
	HPRT	<0.001	<0.001	<0.001	0.005			0.005	0.001
	MRPL19		0.017	<0.001					
MUC2	***ACTB***		0.001	<0.001				0.046	0.007
	B2M								
	DTX3					0.010			
PDCD4	***ACTB***		<0.001	<0.001	0.003	<0.001	<0.001	<0.001	<0.001
	DTX3	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001
	GAPDH								0.004

(B)

Targets	CE genes	ACTB	B2M	DTX3	GAPDH	HPRT	MRPL19	PPIA	PPIA & B2M
CXCL12	***DTX3***	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001
	HPRT		<0.001		0.004			0.010	<0.001
FABP1	***ACTB***		0.010	0.022	0.011	0.009	0.009	0.010	0.009
MUC2	***DTX3***	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001
PDCD4	***DTX3***	<0.001	<0.001		<0.001	<0.001	<0.001	<0.001	<0.001

3.5 Discussion

Since its introduction in 1996 [350] many methods have been developed for the analysis real-time quantitative PCR data. Relative quantification has come to the fore as the method of choice due to its superior flexibility and reduction in inherent variation associated with sample preparation. Prior to the availability of high-throughput realtime PCR instrumentation, a handful of genes were commonly used to normalise real-time data. Major technological advances enabling high throughput analysis of both samples and target genes have enabled investigation and validation of putative EC genes prior to their use to normalise target gene expression. It is now accepted that the use of more than one gene to normalise RQ data improves experimental accuracy compared to the use of a single EC gene [321, 347, 350]

In their study of EC gene expression in breast and colon cancer tissues Tricarico *et al* [351] illustrated significant variation in the expression levels of 10 commonly used housekeeping genes including 18S rRNA, both between individuals and between biopsies taken from the same patient. They concluded that normalisation to a single EC gene was inappropriate for human tissue samples. Moreover, Vandesompele *et al* identified errors of up to 6.5 fold when a single EC gene was used in comparison to the use of multiple genes for data normalisation [321] thereby clearly indicating the potential for superior accuracy when due consideration is paid to the choice of EC genes.

Many analytical programmes for relative quantification have been developed, certain of which enable the identification of EC genes from a study population [349, 352, 353]. In the present study the stability of expression of candidate EC genes was determined using a pair-wise comparison model: geNorm [321] and an MS Excel ANOVA based model, NormFinder [347]. No effect of disease status EC gene expression was identified in colorectal tissue. Since both geNorm and NormFinder are based on the assumption that candidate genes are not differentially expressed between samples, this was an important first step prior to their continued use [322, 323].

In this study GeNorm was used to identify the most stably expressed EC genes from our panel of candidates and also provided a measure of the optimal number of EC genes. *B2M* and *PPIA* were identified as the most stable pairing. In order to achieve a pair-wise variation value (V) below the cut-off of 0.15 additional genes should theoretically be used; however this cut-off point is not absolute [326]and may not always be achievable [354]. No significant difference in target gene expression was observed when the top three most stable EC genes identified by geNorm were used confirming that using of a pair of genes may be more practicable given cost, work load and sample availability considerations.

NormFinder was designed to identify EC genes with the lowest stability values; these values are calculated based on intra- and inter-group variation. In this study NormFinder was used to define the best combination of genes using tumour and normal as group identifiers in the calculations. *MRPL19* was selected as the most stable gene using these criteria; however *B2M* and *PPIA* were highlighted as the best combination of genes with even lower stability value compared to *MRPL19* alone. QBasePlus real-time PCR data manager programme was developed based on geNorm and qBase [348] algorithms. QBasePlus was used to confirm our selection of the *B2M* and *PPIA* pairing as the best combination of ECs in colorectal tissue.

Equivalence testing was developed in biostatistics to address the situation where the aim is not to show the difference between groups, but rather to establish that two methods are equal to one another. In equivalence testing, the null hypothesis is that the two groups are not equivalent to one another, and hence rejection of the null indicates that the two groups are equivalent. Therefore, as stated by Haller *et al*, there is a risk of accepting non-differentially expressed genes as suitable controls although they are not equivalently expressed [355]. Equivalence of expression between tumour and normal colorectal tissue was confirmed for all candidate EC genes using the equivalence test and a fold cut-off of 2. *DTX3, B2M*, *MRPL19* and PPIA showed the minimum of variability in the confidence interval hence can be used for normalisation.

In their study to identify EC genes to monitor enterocyte differentiation and to compare normal and adenocarcinoma of the colon from microarray data, Dydensborg *et al* [318] recommended *RPLP0* for normalising gene quantification in human intestinal epithelial cells and *B2M* for studying gene expression in human colon cancer. In addition, Blanquicett [356] analysed the extent of variability in gene expression between tumour and normal colorectal and liver tissues using two-tailed T tests. They showed that *18S, S9* and *GUS* were the least variable genes in normal and metastatic liver specimens and were also appropriate for normal and tumour colorectal tissues. In the present study, we confirmed that more than one EC gene is required for optimal normalisation in colorectal tissue. We used clinico-pathologically diverse tissues to systematically evaluate normalisation of gene expression data in colorectal tissues. We also conducted equivalence testing to confirm the equality of expression of each EC gene. Thereby, the risks of incorrect rejection (type 1 error) and of false negativity (type 2 error) were minimised.

As stated above significant differences in target gene expression were noticed when using each of the EC genes and the combination of *PPIA* and *B2M*. Moreover, significant effect of EC on the magnitude of error associated with estimation of target gene expression was also determined in this study (figure 3.6). Our results were further confirmed by post hoc testing of individual levels of EC gene expression (table 3.5). Reduction in the magnitude of error achieved using the combination of *PPIA* and *B2M* in comparison to using individual EC genes alone, further indicates that using two EC genes to normalise real-time data achieves greater accuracy in the determination of gene expression levels.

The findings reported in this study confirm that use of two EC genes to normalise RQ-PCR data resulted in superior accuracy in the quantification of gene expression in colorectal tissue. The combined use of *B2M* and *PPIA* was validated as the optimal pair of EC genes with which to estimate the expression of all four target genes in colorectal cancer tissue. Although these ECs may not be ideal in other tissue types, the approach described herein could serve as a template to identify valid ECs in other tissue types.

Chapter 4:

Gene expression profiling

4.1 Introduction

Despite several advances in diagnosis and treatment, CRC remains a threat to life for a large number of people and approximately 20% of patients present with metastatic disease, and 30% of colorectal cancer recur [2]. In general, Colorectal carcinoma (CRC) is generally classified into three categories, based on increasing hereditary influence and cancer risk [3] sporadic CRC (60%) and comprises patients with no notable family history and, by definition, with no identifiable inherited gene mutation that accelerates cancer development, familial CRC (30%) and refers to patients who have at least one blood relative with CRC or an adenoma, but with no specific germLine mutation or clear pattern of inheritance), and hereditary CRC syndromes (10%) which result from inheritance of a single gene mutation in highly penetrant cancer susceptibility genes. Although the last group has the lowest frequency, it has elucidated molecular mechanisms of carcinogenesis applicable to sporadic CRC.

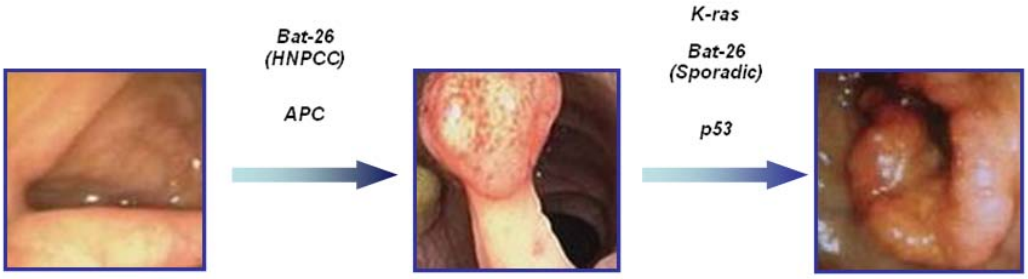
The increasing use of adjuvant and neoadjuvant therapy has led to improved outcomes in the management of colorectal cancer. Post-operative adjuvant chemotherapy has been shown to improve the outcome in patients with Dukes' C tumours and is generally accepted as standard care[357] , however, only selected patients of Dukes' B group would benefit from this treatment.. Moreover, neoadjuvant chemoradiation is becoming the standard of care in the treatment of locally advanced rectal cancer. It is associated with significant improvements in down staging of the disease which correlates with improved rates of sphincter sparing surgery, decreased regional recurrence, and improved overall survival as confirmed by the prospective randomized trial of the DCCG and the German Rectal Cancer Study Group [36, 358]. The response to neoadjuvant therapy is quantified by tumour regression grade which was originally described for tumours of the oesophagus [44]. Currently the ability to predict response to RCT does not exist. Moreover, the selection criteria for patients who would benefit from adjuvant chemotherapy have not yet been defined.

At the molecular level, activation of oncogenes and inactivation of tumour suppressor genes [359] are processes known to be involved in colorectal

carcinogenesis. Additionally, abrogation of mismatch repair systems [360] contributes to some colorectal cancers. Nevertheless, exactly how those genetic alterations bring about the development and progression of colorectal carcinomas remains to be resolved. To complicate the picture, accumulations of mutant genes in neoplasms tend to be accompanied by other genetic and epigenetic changes including loss of heterozygosity, inactivation of important genes by methylation or loss of imprinting [361] or gene amplifications, all of which can alter gene expression profiles. Therefore, genome wide monitoring of gene expression is of great importance if we are to disclose the numerous and diverse events associated with carcinogenesis. Molecular profiling, a tool of genome monitoring, is an attempt to identify the different combinations of genetic events or alternative pathways that may be represented by cancers of a similar type.

In 1990 Fearon and Vogelstein [74] developed the principle of an adenoma-carcinoma sequence, postulating that the transition from adenoma to carcinoma was associated with an accumulation of genetic events in key regulatory oncogenes and tumour suppressor genes that confer a growth advantage to a clonal population of cells (Figure 4.1). Since then, molecular detection methods based on gene mutation determination for *APC, p53* and *K-ras,* have been carried out. The APC pathway is documented to be altered in approximately 95% of colorectal tumours [81]. The same number holds for the fraction of tumours with mutations in the p53 pathway [89] and somatic mutations that lead to alteration of the Ras/Raf pathway have so far been found in about 70%of tumours. There are additional genes with somatic mutations in colorectal cancers, but their respective roles and pathways are less well understood and examined.Despite the usefulness of these molecular markers, the applications remain limited for CRC patients; therefore, new molecular markers are needed for Identification of signature gene lists - potential for prediction of clinical outcome and tumour subclassification like the Oncotype DX of breast cancer [362]

Figure 4.1: **The genetic model of colorectal tumorigenesis** [74]
The model described by Fearon and Voglestein includes several genetic changes that are required for cancer initiation and progression



Advances in molecular technology such as microarray analysis and PCR allow comparisons of expression for thousands of genes within individual tumours and thus serve as tools for prognostication and novel tumour treatment strategies.

Microarrays are microscope glass slides containing thousands of addressable genes that are used as probes to quantify the relative amount of RNA transcripts extracted from tissues on the basis of florescent signal produced by the labelled cDNA bound to the microarrays. Two different colours are used to visualize the difference between up- and down-regulation of gene expression. The most currently used microarray plat-forms are spotted cDNA microarrays and high-density oligonucleotide microarrays (e.g. Affymetrix GeneChip system). Gene expression profiling in colorectal cancer using microarray analysis was used to investigate carcinogenesis process, prognosis prediction and treatment response prediction [363]

On the other hand, real-time quantitative polymerase chain reaction (TR-PCR) is a combination of the reverse transcriptase (RT)-dependent conversion of RNA into cDNA, the amplification of the cDNA using the PCR and the detection and quantification of amplification products in real time [364]. It addresses the evident requirement for quantitative data analysis in molecular medicine and has become the gold standard method for the quantification of mRNA and therefore for validation of microarray data.

4.1.1 Candidate genes

In order to identify a list of genes associated with deregulated expression in colorectal cancer and thereby might have a role in colorectal cancer tumourogenesis, we carried out a detailed analysis of published colorectal cancer microarray data and identify the most prominent genes. Furthermore, a literature review was performed to identify mRNA highly associated with cancer to identify their role in colorectal cancer pathogenecity and progression [185, 363, 365]. Table 5.1 showed the list of candidate genes selected for analysis in this study

4.1.1.1 Cadherin 17 (*CDH17)*

CDH17 (liver-intestinal cadherin) encodes a membrane associated glycoprotein, a structurally and functionally unique member of cadherin superfamilly. It was originally found to be expressed in the rate intestine and liver [366], although being exclusively expressed in the intestinal but not in liver epithelial cells in man [367]. It functions as peptide transporter and cell to cell adhesion molecule. Several studies on clinical implication of cadherin17 in human cancer have been performed. This intestines specific cell adhesion molecule is found to be associated with intestinal metaplasia, carcinoma of the stomach[368-370], hepatocellularcarcinoma [371] and adenocarcinoma of the pancreas [372]. In colorectal cancer *CDH17* studied mainly at proteins level using IHC. Its reduced expression was found to be associated with advanced stage, tumour dedifferentiation and poor survival [373, 374]

4.1.1.2 Carcinoembryonic antigen related cell adhesion molecule 5 *(CEA, CEACAM5)*

CEACAM5 (CEA), a member of *CEA* genes family, encodes the tumour marker carcinoemberyonic antigen (CEA). Since 1965 CEA serum level is used clinically to monitor patients with colorectal and other cancers [16] but a group of these patients have extraordinarily high CEA levels that cannot be attributed solely to tumour load. Moreover, elevated levels of serum CEA are found in patients with benign liver disease including cirrhosis, biliary obstruction, and hepatitis. Both factors significantly affected the specificity of the test in clinical practice. *CEA* mRNA was detected in the peripheral blood colorectal cancer patients and was

Table 4.1: Candidate genes

Gene name	Gene symbol	Location	Assay ID	bp	References
Cadherin 17	***CDH17***	8q22.1	Hs00184865_m1	72	[184, 211]
Carcinoembryonic antigen related cell adhesion molecule 5	***CEACAM5***	19q13.2	Hs00944023_m1	71	[375, 376]
Chemokine ligand 12	***CXCL12***	10q11.1	Hs00171022_m1	77	[181, 211]
Chemokine, cxc motif, receptor4	***CXCR4***	2q21	Hs00237052_m1	78	[377, 378]
Chemokine, cxc motif, receptor7	***CXCR7***	2q37	Hs00171022_m1	129	[377, 379]
Fatty acid binding protein 1, liver	***FABP1***	2p11	Hs00155026_m1	71	[171, 200, 201, 365, 380]
Interleukin-8	***IL-8***	4q13-q21	Hs99999034_m1	81	[185, 200, 204, 205]
Mucin2	***MUC2***	11p15.5	Hs03005094_m1	64	[172, 175, 185]
Programmed cell death 4	***PDCD4***	10q24	Hs00205438_m1	94	[191, 203]
Transforming growth factor beta1	***TGFB1***	19q13.1	Hs00998133_m1	57	[185, 365, 381-383]
Transforming growth factor-beta receptor type 1	***TGFBR1***	9q22	Hs00610320_m1	73	[384, 385]
Transforming growth factor-beta receptor type 2	***TGFBR2***	3p22	Hs00234253_m1	70	[381, 386]

used as an indicator of circulating tumour cells load [387, 388]. Quantification oftumour tissues *CEACAM5* expression levels may establish an alternative approach with higher sensitivity and specificity for diagnosis and follow up of colorectal cancer patients.

4.1.1.3 *CXCL12* and its receptors (*CXCR4* and *CXCR7*)

Chemokines are small, chemotactic cytokines that direct migration of leukocytes, activate inflammatory responses and participate in many other functions, including regulation of tumour growth. Chemokines contribute to cancer biology by three important mechanisms: regulation of tumour-associated angiogenesis, activation of a host tumour-specific immunological response, and direct stimulation of tumour cell proliferation, migration and survival [389-391]. All of these mechanisms are promising points of cancer intervention, and preclinical experiments suggest that chemokine antagonists and agonists could become important in the development of new anticancer therapies.

The chemokine *CXCL12*, also known as stromal cell-derived factor-1 (*SDF-1*), has been shown to play a significant role in tumourgenesis, promoting angiogenesis and tumour invasion and migration to metastatic sites [378, 392-394]. These observed effects of *CXCL12* have been previously thought to be mediated entirely through its receptor *CXCR4*. However, the recently described receptor for CXCL12, CXCR7 [395, 396] may required a re-examination of much of the previous work that presumed an exclusive effect of CXCL12/CXCR4 axis.

4.1.1.4 Fatty acid-binding protein 1, liver *(FABP1)*

Liver fatty acid-binding protein (*L-FABP, FABP1*) is a member of intracellular proteins family that mediate transportation and utilization of lipids. It is specifically expressed in the hepatocytes and enterocytes and could serve as a sensitive marker of enterocytes differentiation [397, 398]. It increases solubility of fatty acids in cell cytoplasm and facilitate their up take and processing [399-401]. *FABP1* plays an active part in several physiological functions including signal transduction, modulation of cell division, cell growth and differentiation and regulation of gene expression [402, 403]. All these functions are deregulated in

tumourogenesis and tumour progression, supporting the possible role of this molecule in colorectal cancer development and progression.

4.1.1.5 Interleukin 8 (*IL-8, CXCL8)*

IL8 is a major mediator of the inflammatory response. It is secreted by leukocytes and tumour cells and functions as a chemoattractant, and a potent angiogenic factor [404, 405]. Increased IL-8 expression has been found in most of metastatic and solid tumours of the breast, melanoma and ovaries [406-408]. In colorectal cancer, studies have found that IL-8 serum levels correlate with poorer prognosis, tumour progression and metastasis [409].

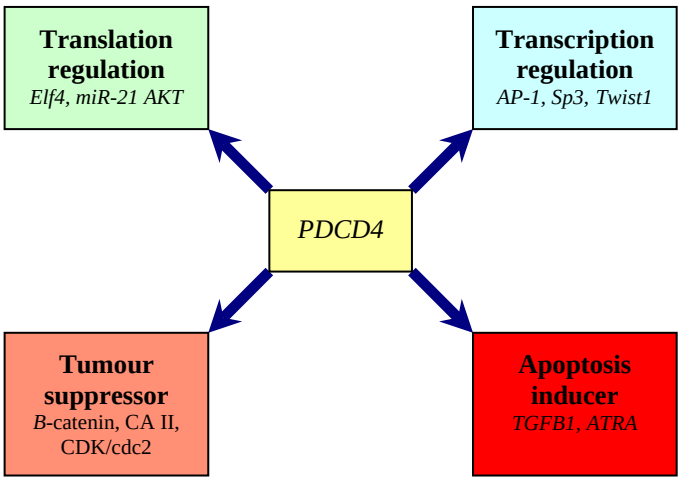
4.1.1.6 Mucin 2 (*MUC2)*

MUC2, intestinal-type gel-forming secretary mucin, is produced and secreted by globlet cells and is a major constituents of mucus, which acts to lubricate and protect intestinal epithelial tissues [410]. In animals, inactivation of MUC2 caused intestinal tumour formation which was accompanied by increase proliferation, decrease apoptosis and increase migration of the cells [411]. These alterations might primarily relate to MUC 2 absence or could be secondary to the inadequate protection of intestinal mucosa. Reduced *MUC2* expression was reported to be associated with development and progression of colorectal cancer [412, 413], however, in tumours like gastric and bladder cancer overexpression of MUC2 was noticed[414, 415]

4.1.1.7 Programmed cell death 4 (*PDCD4)*

PDCD4 is a novel tumour suppressor that inhibits tumour promotion and progression in both cell lines and animal models. The main functions of the gene are to inhibit translation, suppress proliferation and cell cycle progression and induce apoptosis [416, 417]. It achieves these functions through interaction with many other molecules and pathways. *PDCD4* is down-regulated in several human cancers including lung, ovary, and brain [418-420]. In colorectal cell lines, down-regulation of *PDCD4* found to promote invasion and metastasis.

Figure 4.2: Functions and molecular interactions of *PDCD4* [416]



4.1.1.8 *TGFB1* and its receptors (*TGFBR1* and *TGFBR2*)

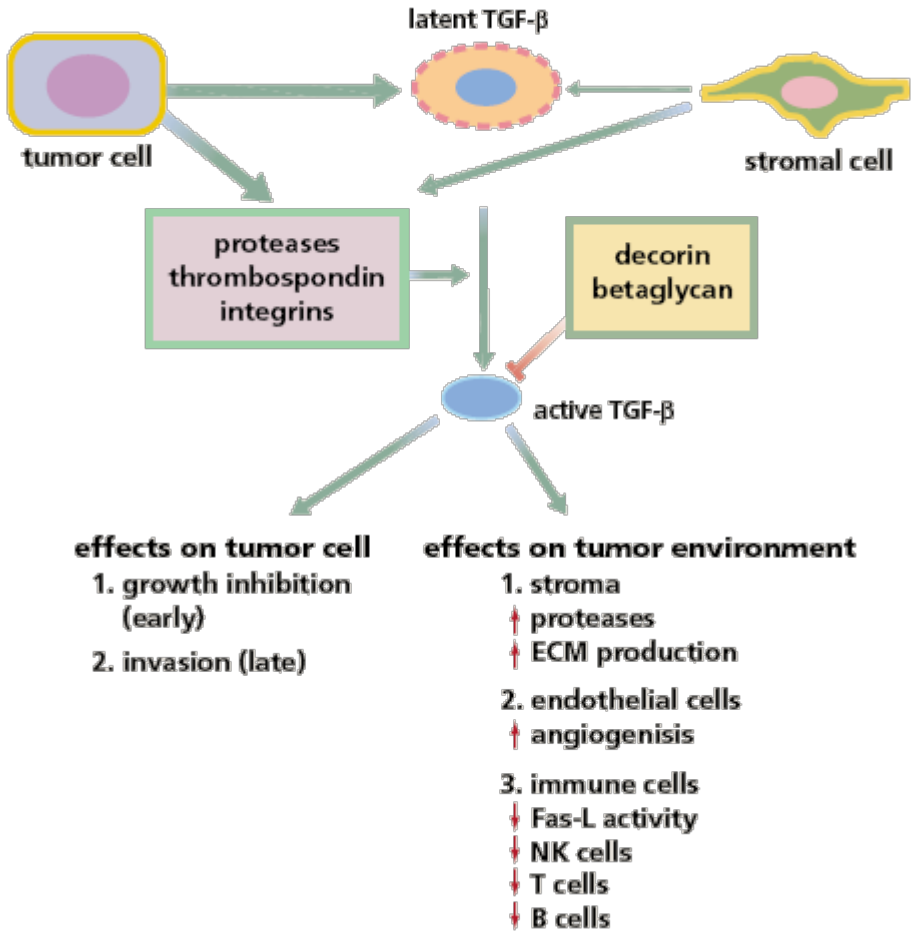
Transforming growth factor-β1 (*TGFB1*), a multifunctional cytokine, mediates its effect on cells through a heteromeric receptor complex that consist of type I and type II components. The pathway signalling is initiated by binding of *TGFB* to type II receptor (*TGFBR2*) which consequently recruits and phosphorylates the type I (*TGFBR1*) receptor. This will lead to stimulation of *TGFBR1* protein kinase activity. Activated *TGFBR1* then phosphorylates two downstream transcription factors, *SMAD2* and SMAD*3*, allowing them to form a complex with *SMAD4*. The complexes then translocate into the nucleolus and interact with other transcription factors to regulate the transcription of *TGFB1* responsive genes.

The *TGFB1* pathway, the most commonly altered cellular pathway in human cancer, is involved in several physiological functions including cell proliferation, differentiation, migration and apoptosis [421]. It also stimulates angiogenesis, directly through induction of expression of *VEGF* or indirectly through attracting monocytes which release angiogenic cytokines [422]. In addition, *TGFB1* is involved in regulation of extracellular matrix production, cell adhesion and immune surveillance. These functions are integral part of tissue homeostasis and represent logical targets for dysregulation in carcinogenesis.

Although *TGFBR1* and *TGFBR2* function as tumour suppressors in the development of carcinoma [423-425], *TGFB1* acts as both a tumour suppressor and as a significant stimulator of tumour progression. At early stage of tumourogenesis, it acts directly on cancer cell to suppress tumour growth. As tumour progress genetics and biochemical changes allow *TGFB1* to stimulate tumour progression by its activity on both cancer cells and stromal cells of the tumour [426]. Various factors can contribute to increase tumour progression by *TGFB* pathway. Of these, early genetic loss of signalling components like *TGFBR2*(in more than 30% of colorectal cancer) leads to rapid growth through increase cell division and inhibition of apoptosis; hence increase the probability of further mutations and cytogenic changes that ultimately drive tumour progression [422, 426, 427].

In colorectal cancer high serum levels of *TGFB1* protein was found to be associated with advanced Dukes' stage, depth of tumour invasion and metastasis [428]. However, at an early stage of the disease *it* was found to suppress nonmetastatic tumour growth [429]. Regarding TGFB1 receptors, *TGFBR1*6A* polymorphism was linked to hereditary colorectal cancer [430, 431] while *TGFBR2* inactivation was identified in more than 90% of tumours with microsatellite instability [147]

Figure 4.3: *TGFB* in tumour inviroment
TGFB secreted by tumour cells and non-malignant stromal cells acts on both the tumour cells and its environment [422].



4.2 Aims

The aim of the study was to quantitative candidate genes expression in colorectal cancer tissues using RT-PCR in order to:

- Determine the expression levels of candidate genes in tumour and tumour-associated normal colorectal tissue
- Investigate correlation between serum carcinoembryonic antigen (CEA) and tissue CEACAM5 levels
- Correlate candidate genes expression levels and clinicopathological variables.

4.3 Materials and methods

4.3.1 Study group

Clinicopathological data on all patients were examined in order to select suitable samples for study groups appropriate to address specific questions. A heterogeneous group of 107 patients with colorectal tumours, all of which had matched TAN samples was selected for gene expression profiling experiment using real-time quantitative PCR (RQ-PCR) (Table 4.3). Tissue samples were gathered from consenting patients at the time of diagnostic procedure or at primary curative surgical resection at Galway University Hospital, Ireland. The cohort comprised of 101 colorectal tumour specimens, 8 polyps and 107 tumour-associated normal (TAN) tissues. Following retrieval, all samples were subject to histopathological review prior immediate snap-freezing in liquid nitrogen and archival at -80ºC until further use. Concomitant clinicopathological data on patients and specimens was obtained through patient interview and review of clinical notes. Ethical approval for this study was granted by the Clinical Research Ethics Committee, Galway University Hospitals.

Whole blood from 4 colorectal cancer patients and 4 negative controls was used to determine gene expression in blood. Informed written consent was obtained from each participant before enrolment in the study.

The group used to investigate the effect of neoadjuvant therapy in gene expression of colorectal cancer consisted of 58 patients with rectal cancer. Of them, 25 had neoadjuvant therapy before surgical resection. Only 1 patient had no response to treatment (Mandard TRG5) while 6 patients were TRG4, 4 patients were TRG3, 9 patients were TRG2 and 5 patients were TRG1 (complete response).

Table 4.2: Concentration of RNA extracted from 4 colorectal cancer patients

Sample ID	ng/µl	A260	260/230	260/280	constant
R08-0948 1	3629.1	90.729	0.95	1.68	40
R08-0949 1	3990.5	99.764	1.01	1.63	40
R08-0950 1	28.07.2	70.181	0.72	1.70	40
R08-0951 1	3858.8	96.472	0.98	1.62	40

Table 4.3: **Clinicopathological data for patients used for gene expression analysis**

Clinicopathological Variable	Number of Patients N=(107)
Tissue type	
Carcinoma	101
Polyp	8
Gender	
Males	67
Females	40
Mean Age (SD)	69.72 (11.89)
Tumour Location	
Colon	43
Rectum	58
Tumour Location	
Proximal	27
Distal	74
Tumour thickness (mm)	
<10	23
10-15	33
>15	22
Unknown	23
Tumour Diameter (mm)	
<30	29
30-40	26
>40	31
Unknown	15
Distant Metastasis	
M_0	80
M_1	21
Nodal Status	
N_0	22
N_1	11
N_2	9
UICC Stage	
Stage 0	2
Stage I	17
Stage II	28
Stage III	28
Stage IV	21
pCR	5
Tumour Differentiation	
Grade 1: Well differentiated	11
Grade 2: Moderate differentiated	72
Grade 3: Poor differentiated	10
Not applicable	8
Mucin Secretion	
Mucinous	19
Non-mucinous	82

4.3.2 RNA extraction and analysis

Tissue samples (50-100 mg) were homogenised using a hand-held homogenizer (Polytron PT1600E) in 1-2 mL of QIAzol reagent (Qiagen). Tumour and TAN samples were homogenised separately but on the same day. Two methods of RNA extractions were employed in the study, the total RNA extraction (co-purification) and the separate purification of mRNA and miRNA.To ensure both methods were working properly correlation of RNA concentration and quality were carried out and showed good results (Table 2.5).

RNA was extracted using the RNeasy Plus Mini Kit and RNeasy MinElute cleanup kit (Qiagen) according to the manufacturer's instructions. For this study, only large RNA was utilised for further analysis. RNA was eluted in 60µl volumes and stored at -80^{o}C.

Total RNA was extracted from 1mL of whole blood using the Tri Reagent BD (Molecular Research Centre) and a slightly modified protocol from that provided by manufacturers. 1-bromo-4-methoxybenzene was used to augment the RNA phase separation and an additional ethanol (75%) wash was performed to improve the purity of RNA isolated as reflected in an improved 260/280 ratio.

RNA concentration and purity was assessed in duplicate samples using a using a NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies). RNA integrity was evaluated using the RNA 6000 Nano Chip Kit (Series II) and the Agilent 2100 Bioanalyzer System (Agilent technologies). An RNA integrity number (RIN) was generated for each sample using the Agilent 2100 Expert Software (Version B.02.03) based on the ratio of ribosomal bands and also the presence or absence of degradation products on the electrophoretic and gel-like images. A threshold value of RIN $\geq$ 7 was applied and RNA purity was verified by an average A260/A280 ratio of 1.98 (range 1.97-2.01) and A260/A230 ration of 1.7 (range 1.5-1.83).

4.3.3 Reverse transcription

First strand cDNA was synthesised using Superscript III reverse transcriptase (Invitrogen) and random primers (N9; 1µg, MWG Biotech). Negative control

samples were included in each set of reactions. Reactions were incubated at 25º C for 5 minutes followed by 50º C for 1 hour and final denaturation at 72º C for 15 minutes. Samples were subsequently diluted to 100 µL in nuclease-free water and stored at -20º C.

4.3.4 Real-time quantitative PCR

4.3.4.1 Amplification efficiency

In determining gene expression using RQ-PCR and relative quantification, it is important to consider the amplification efficiency for the assay in use. The PCR efficiency impacts greatly on the accuracy of the calculated expression result and is influence by PCR reaction component. For 100% efficiency there will be doubling of the amount of DNA at each cycle, while for 80% and 70% the amount of DNA will increase from 1 to 1.8 and 1.7, respectively. Therefore, a small difference in efficiency makes a large difference in the amount of the final product.

Amplification efficiencies for each EC gene assay were calculated applying the formula E= (10-1/slope - 1) × 100, using the slope of the plot of Ct versus log input of cDNA (10-fold dilution series). A threshold of 10% above and below 100% efficiency was applied (Table 4.4)

4.3.4.2 Endogenous control

Relative quantification is the most widely adopted approach whereby quantification of gene expression is normalised relative to an endogenously expressed control (EC) gene. Central to the reliable determination of gene expression is the choice of control gene. To identify the most stably expressed gene(s) to normalise RQ-PCR data derived from primary colorectal cancer tissue the expression of thirteen candidate EC genes: *B2M, HPRT, GAPDH, ACTB, PPIA, HCRT, SLC25A23, DTX3, APOC4, RTDR1, KRTAP12-3, CHRNB4* and *MRPL19* were analysed in a cohort of 64 colorectal tumours and tumour associated normal specimens. Data was analysed using geNorm, NormFinder and qBasePlus. We determined that two genes were required for optimal normalisation and identified *B2M* and *PPIA* as the most stably expressed and reliable EC genes [292].

Figure 4.4: **Dilution curves of *IL-8* and *PDCD4*** as examples of assay efficiency calculation

(A) *IL-8* dilution curve (B) *PDCD4* dilution curve

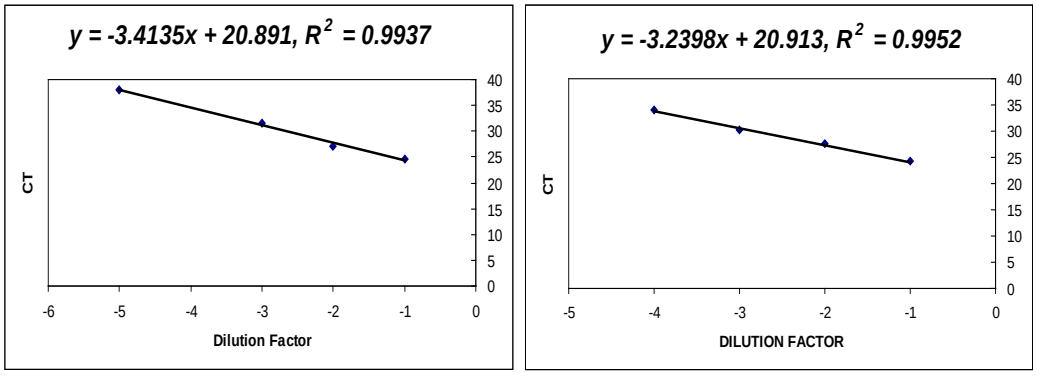


Table 4.4: Amplification efficiencies of candidate genes
Amplification efficiencies (E) and standard error of the mean (SE) of the candidate genes

Candidate gene	E%	SE
CDH17	95 %	0.007
CEACAM5	92.8 %	0.004
CXCL12	91.8 %	0.008
CXCR4	100.6 %	0.018
CXCR7	102.5 %	0.022
FABP1	99 %	0.034
IL8	96.1 %	0.013
MUC2	94.5 %	0.007
PDCD4	102.9 %	0.011
TGFB1	99.4 %	0.018
TGFBR1	100.1 %	0.011
TGFBR2	100.1 %	0.003

4.3.4.3 RQ-PCR of mRNA

The expression of each EC gene was analysed by RQ-PCR using TaqMan gene expression assays using a 7900HT instrument (Applied Biosystems). All reactions were performed in 20 μL reactions, in triplicate within the same PCR run.

Negative controls were included for each gene target under assay. On each plate, an interassay control was included to account for any variations between runs. For each well 2µl of cDNA from each sample was added to 18µl of PCR reaction mix which consisted of 10x TaqMan universal master mix, No AmpErase UNG, 7X nuclease free water and 1X gene expression assay primer-probe mix (Applied Biosystems). The PCR reactions were initiated with a 10 minute incubation at 95º C followed by 40 cycles of 95º C for 15 seconds and 60º C for 60 seconds, in accordance with the manufacturer's recommendations.

4.3.5 Relative quantification

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. QBasePlus was used for calculation of candidate expression relative to the endogenous control genes. It applies ΔΔ C_t method was used where ΔΔCt = (C_t target gene, test sample – C_t endogenous control, test sample) - (C_t target gene, calibrator sample - C_t endogenous control, calibrator sample). Relative quantities were corrected for efficiency of amplification and fold change in gene expression between groups was calculated as E-ΔΔCt ± s.e.m. The lowest expressed sample was used as a calibrator.

4.3.6 Statistical analysis

Statistical analysis was carried out with IBM SPSS Statistics 17.0 (SPSS Inc.). Data was tested for normal distribution graphically using histograms and also using the Kolmogorov–Smirnov and Shapiro-Wilk tests. Parametric tests were used where appropriate. One-way ANOVA and independent t-test were used to determine association and comparisons between independent groups. Correlation analysis used Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively. Univariate analysis and paired-T test were used to assess related samples. The statistical significance of differences in survival between groups was determined by log rank which compares differences along all points of the curve and multivariate analysis was done using Cox regression. P values <0.05 were considered statistically significant.

4.4 Results

4.4.1 Gene expression profiling and clinicpathological correlations.

4.4.1.1 In blood

The expression levels of all the candidate genes and the endogenous control genes were undetermined in blood using RQ-PCR despite the high concentration and the good quality of RNA extracted.

4.4.1.2 In tissues

Cadherin 17 (CDH17)

The expression of CDH17 was significantly lower in colorectal cancer compared to TAN tissues (*p<0.001*, t-test, figure 4.6). Regarding the clinicopathological variables, the CDH17 expression significantly increased with increased tumour diameter (*p=0.043*) and tumour thickness (*p=0.035*), however, its expression reduced with increased bowel wall involvement (*p=0.002*). This finding could be explained by CDH17 adhesion function (figure 4.7). Its expression was also reduced in poorly differentiated tumours (*p=0.045*) and in patients with increased CA 19.9 serum level (p=0.014) (Kruskal-Wallis and Mann-Whitney tests, table 4.5).

Figure 4.5*: **CDH17*** **expressions in CRC**. *CDH17* expression was significantly lower in colorectal tumours compared to tumour-associated normal tissues (*p<0.001*)

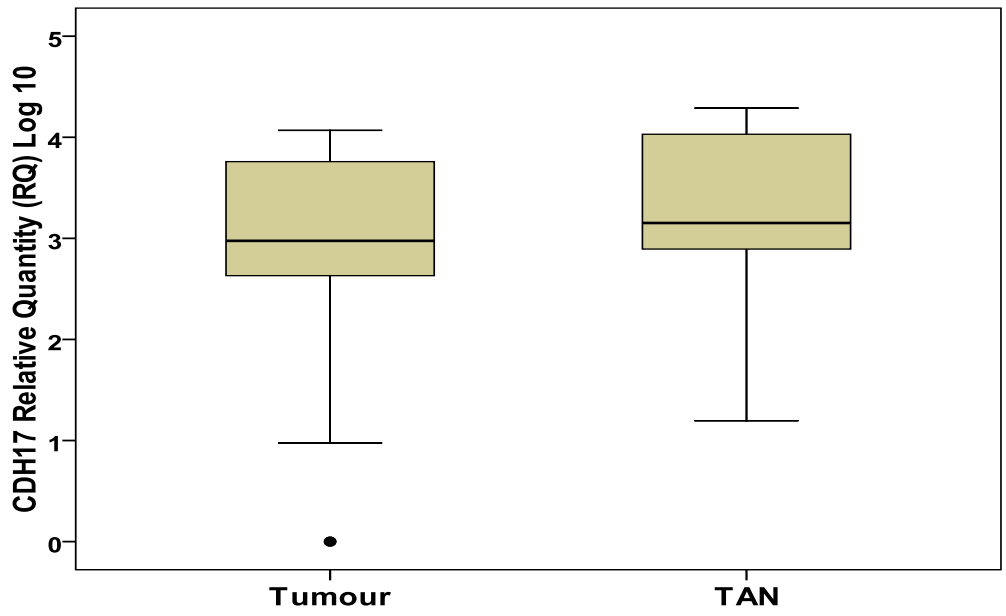
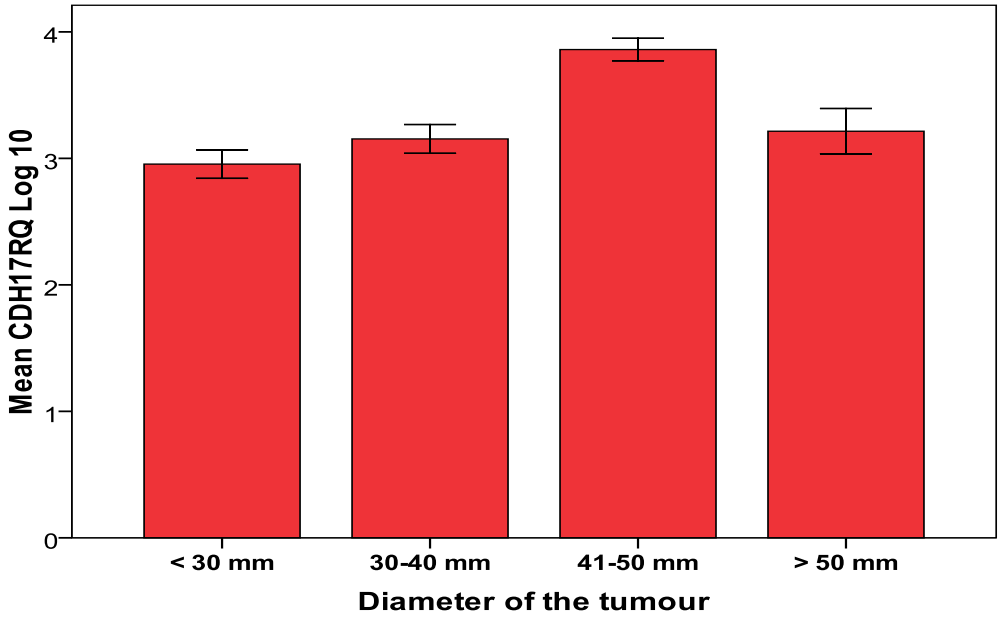
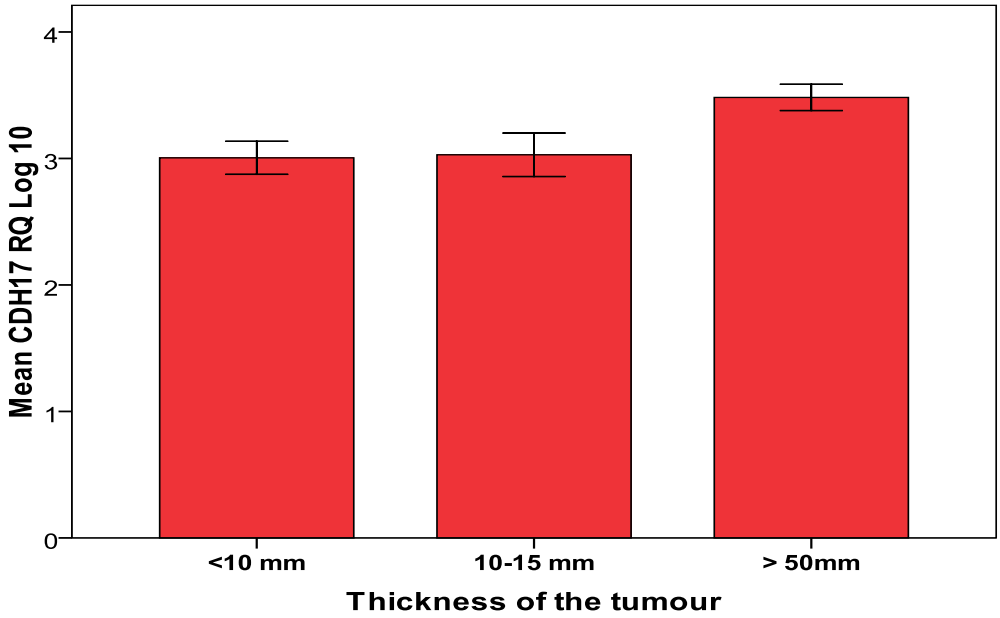


Figure 4.6: **Correlations of *CDH17* expression**. The CDH17 expression significantly increased with increased tumour diameter (A, *p=0.043*) and tumour thickness (B, *p=0.035*) while reduced with increased bowel wall involvement (C, *p=0.002*)

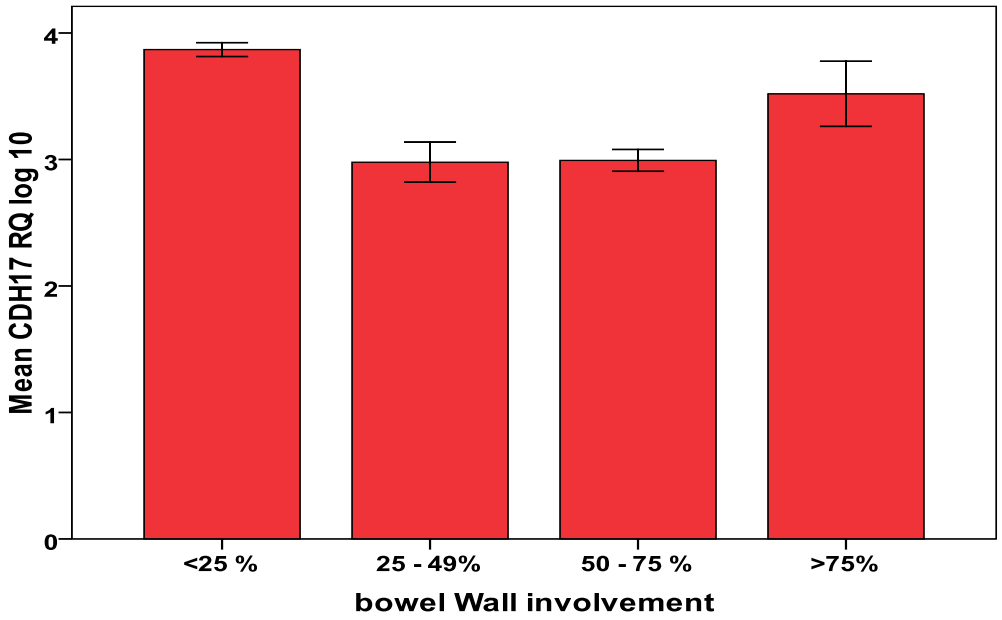
(A)



(B)



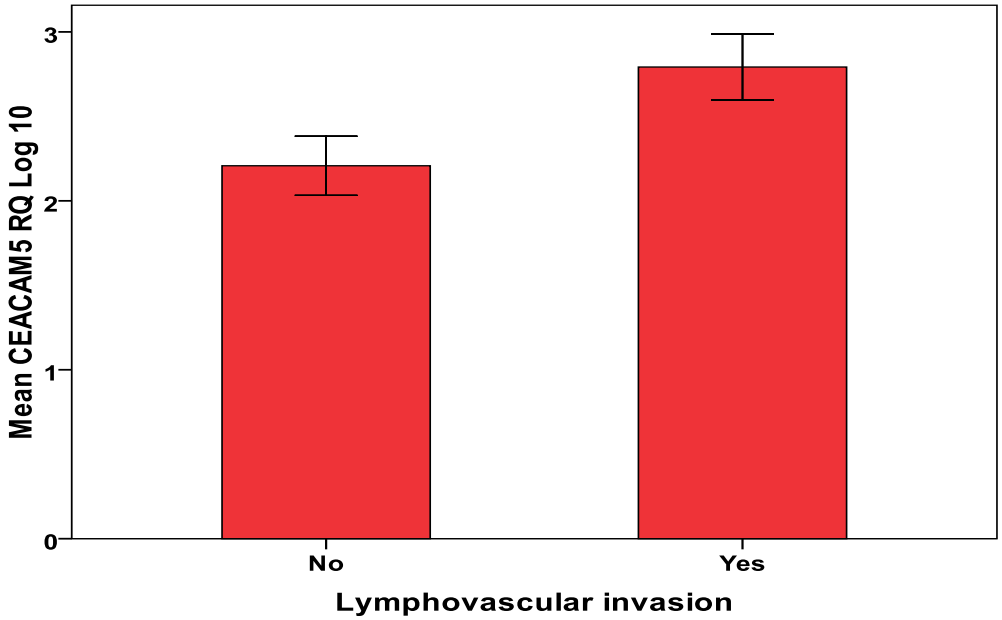
(C)


Mean CDH17 RQ log 10
4
3
2
1
0
<25 %
25 - 49%
50 - 75 %
>75%
bowel Wall involvement

Carcinoembryonic antigen related cell adhesion molecule 5 (CEACAM5)

No significant differences were identified in *CEACAM 5* expression levels in tumour compared to TAN colorectal tissues (*p=0.981*, t-test). In addition, no significant correlations were found between *CEACAM5* expression and the CEA serum level (*r= -134, n=79, p=0.240)*. Higher expression of CEACAM 5 was associated with moderately differentiated tumours (*p=0.016*) and local (*p=0.002*) and lymphovascular invasion (*p=0.019*) (Kruskal-Wallis and Mann-Whitney tests, table 4.5, figure 4.7).

Figure 4.7: **Correlations of *CEACAM5* expression**
Higher expression of the gene was associated with lymphovascular invasion (*p=0.019*)



CXCL12 and its receptors (CXCR4 and CXCR7)

Paired t test was used to investigate the difference in gene expression between tumour and normal colorectal cancer patients in 101 paired tissues. *CXCR4* expression levels were higher in tumour tissues in contrast to the expression of *CXCL12*. However, these differences were only significant in relation to CXCL12 (*p<0.001,*). No differences in *CXCR7* expression were noted between tumour and TAN tissues. Although significant differences were observed in *CXCL12* expression in tumour and polyps compared to TAN tissues (*p<0.001* and *0.003*,

respectively), no differences were found between tumours and polyps (*p=0.907*) (ANOVA, table 4.6, figure 4.8)

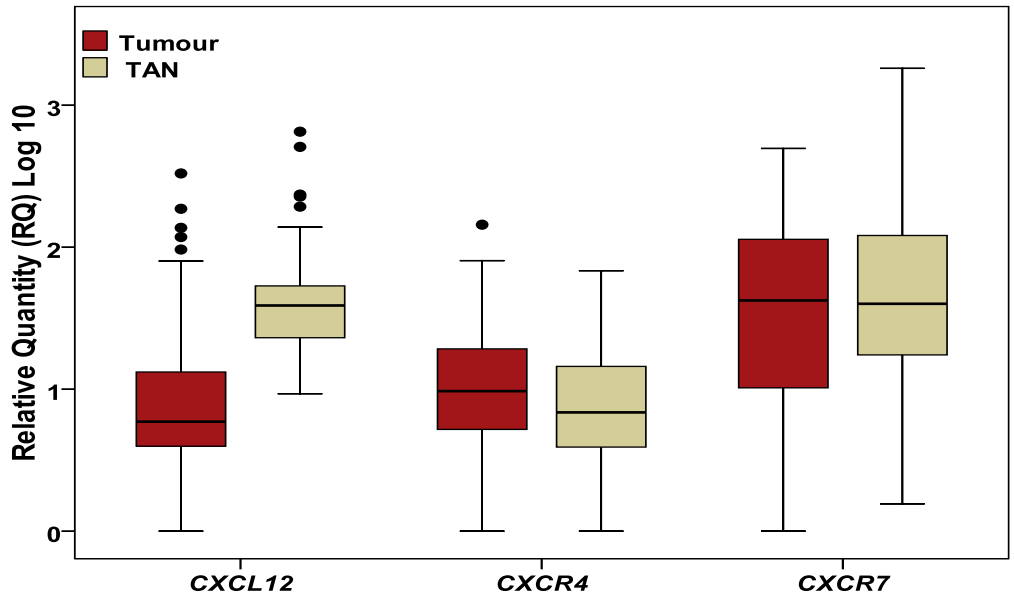
The relationship between *CXCL12*, CXCR7 and CXCR4 was further investigated using Pearson correlation. Preliminary analysis was performed to ensure no violation of the assumption of normality, linearity and homogenecity. There was strong positive correlation between all the variables in both tumour and TAN colorectal tissues with high expression level of the ligand associated with high expression of the receptors (Figure 4.9).

A one- ways ANOVA and t-test were conducted to explore the relation of chemokine expression level and the clinico-pathological parameters. Both *CXCL12* and *CXCR7* were significantly under-expressed in proximal colon. Down-regulation of *CXCL12* and its receptors was significantly associated with survival, advanced stage, poor differentiation and tumour size, invasion and metastesis (figures 4.10, 4.11, 4.12 and 4.13 and table 4.5)

Comparison of patients with higher (above median) tumour *CXCR7* expression to those with lower *CXCR7* expression (below median) yielded a significant difference in overall survival with p value of 0.010 (log rank test, figure 4.11). With median follow up of 15 months, *CXCR7* under-expressors (below median) had a high mortality from colorectal cancer with mean survival of 27 months compared to 46 months in over-expressors (*CXCR7* above median). A multivariate Cox regression analysis was used to determine the prognostic factors for overall survival. After simultaneous adjustment of all these variables there continue to be a significant difference in survival between both groups (p=0.044).

Figure 4.8: **Chemokine expressions in colorectal tumours and TAN**
(A) *CXCL4 (p=0.159)* was up-regulated in tumours, in contrast to *CXCR12 (p<0.001*) which was down-regulated in tumours. No difference was noted in case of *CXCR7*. (B) Although significant differences were observed in *CXCL12* expression in tumour and polyps compared to TAN tissues (*p<0.001* and *0.003*, respectively), no differences were found between tumours and polyps (*p=0.907*)

(A)



(B)

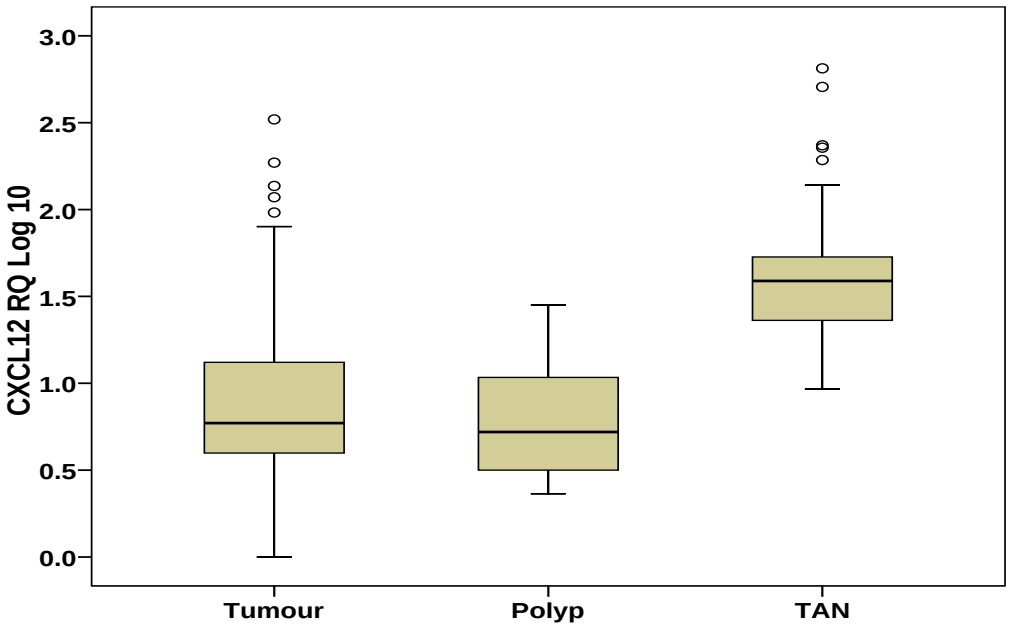
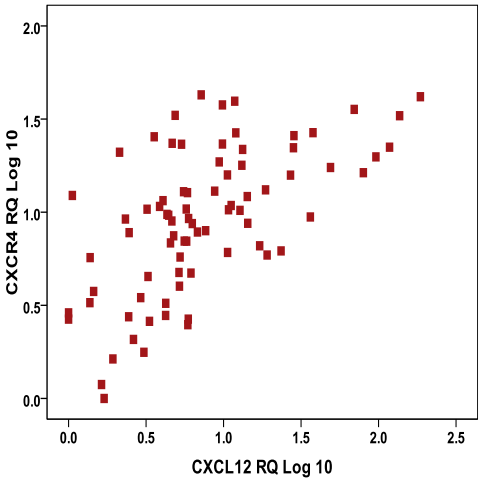
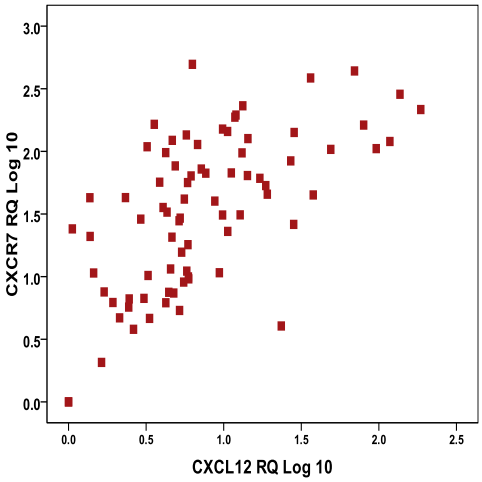


Figure 4.9: Correlations of expression levels of *CXCL12* and its receptors
Pearson's correlations of expression levels of *CXCL12*, *CXCR4* and *CXCR7* in tumour (A, B, C) and tumour-associated normal (D, E, F) colorectal tissues.

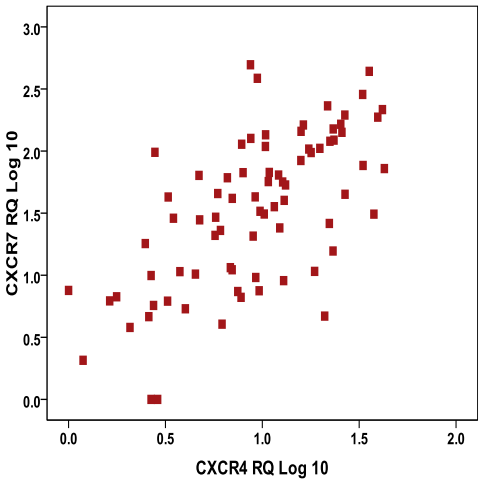
(A) *(r= 0.61, p<0.001)*



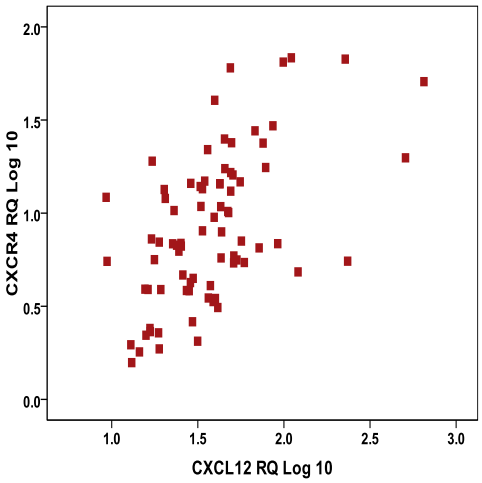
(B) *(r= 0.58, p<0.001)*



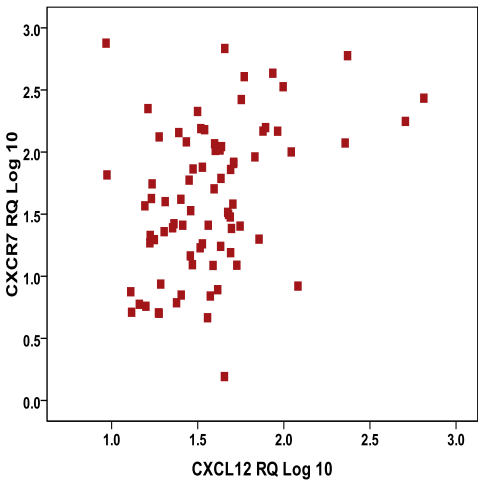
(C) *(r= 0.55, p<0.001)*



(D) *(r= 0.55, p<0.001)*



(E) *(r= 0.54, p<0.001)*



(F) *(r= 0.55, p<0.001*

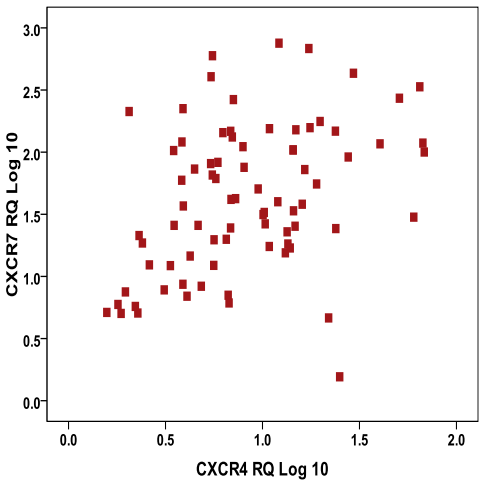
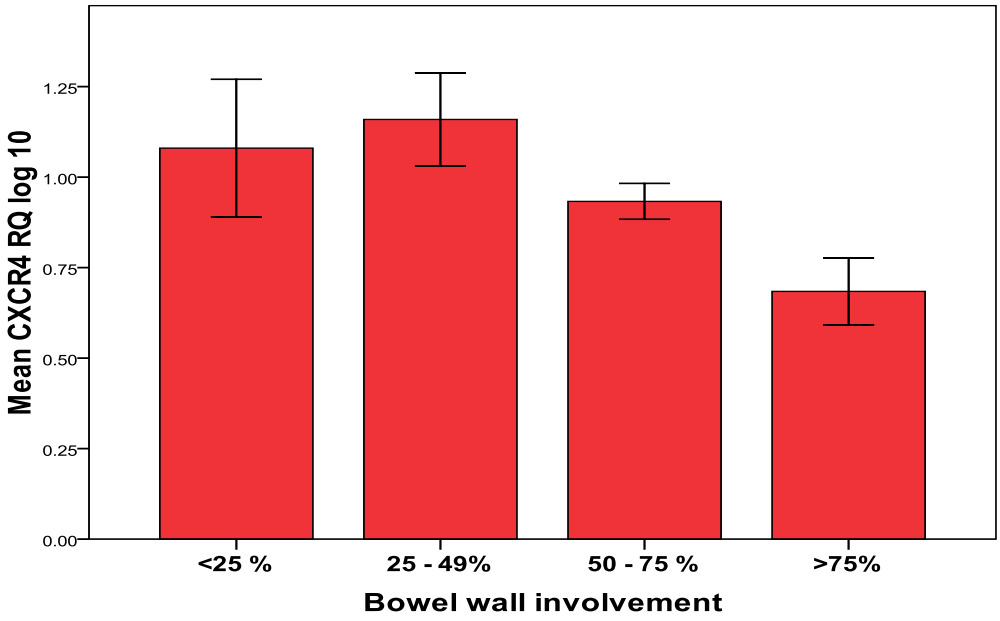


Figure 4.10: Correlations of *CXCR4* expression levels
Down-regulation of *CXCR4* was associated with local invasion (A, *p=0.005*), poor differentiation (B, *p=0.043*) and metastasis (C, *p=0.044*)

(A)



(B)

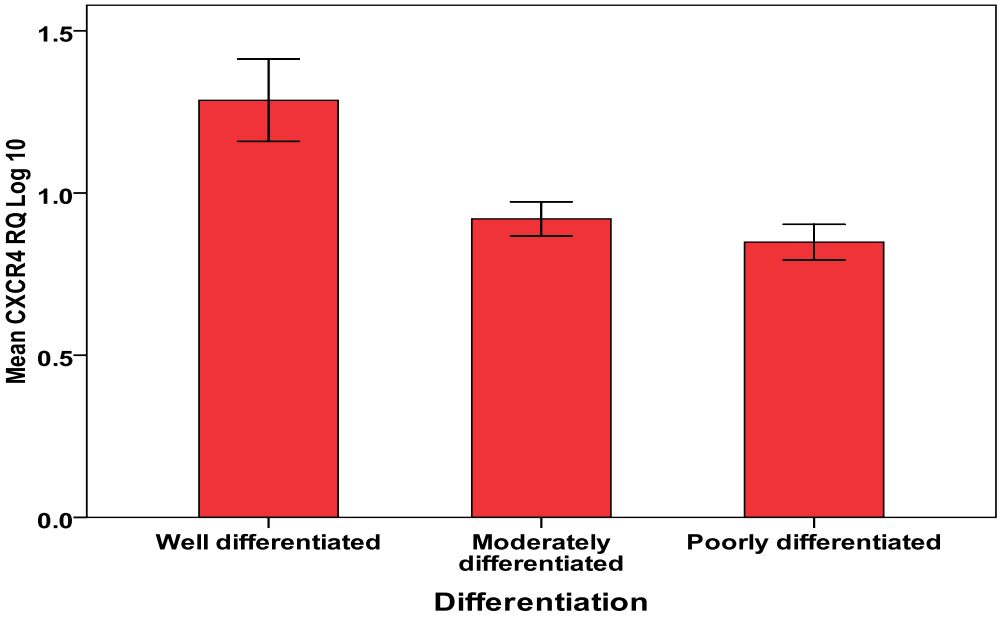


Figure 4:10: Continued

(C)

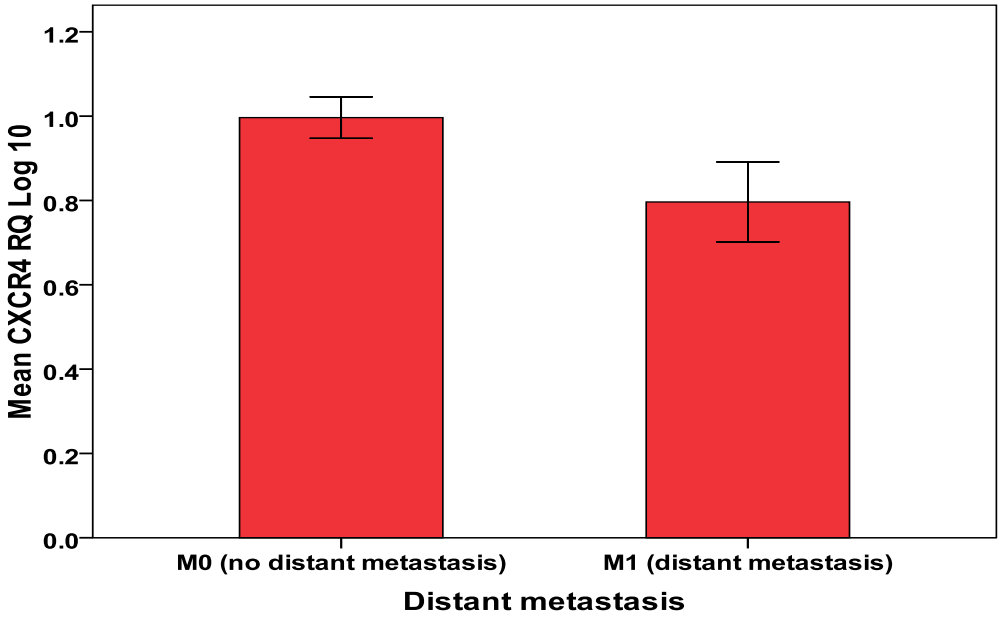


Figure 4.11: ***CXCR7* expression levels and over-all survival of CRC patients**
Patients with tumour tissues expressing *CXCR7 (*expression levels above median) had significantly better over-all survival compared to CXCR7 under-expressers. Significant difference in over-all survival between both groups was determined by Kaplan-Meire survival analysis (X^2*=6.5, p=0.010*) and Cox-regression (*p=0.044*)

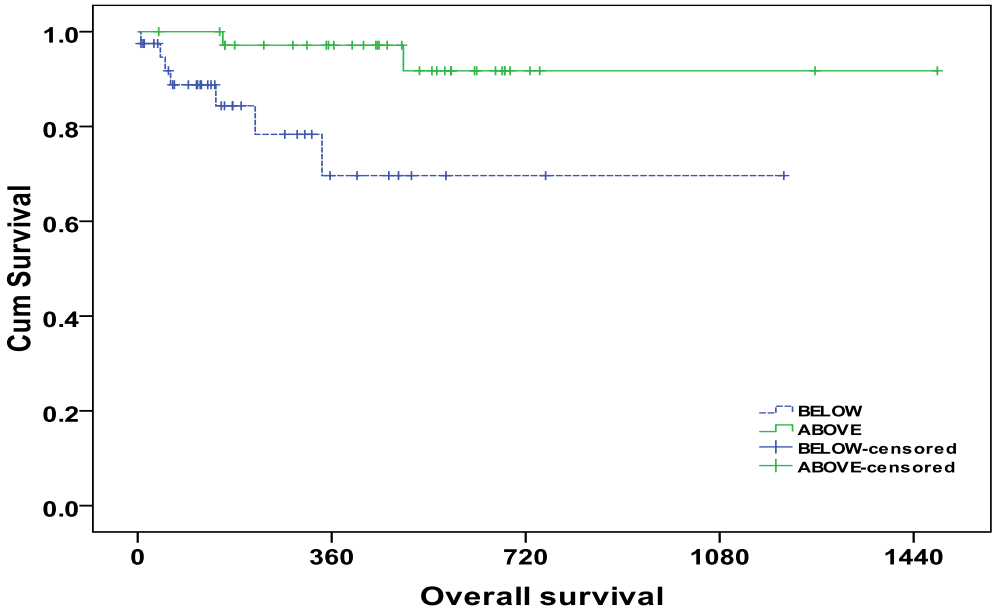
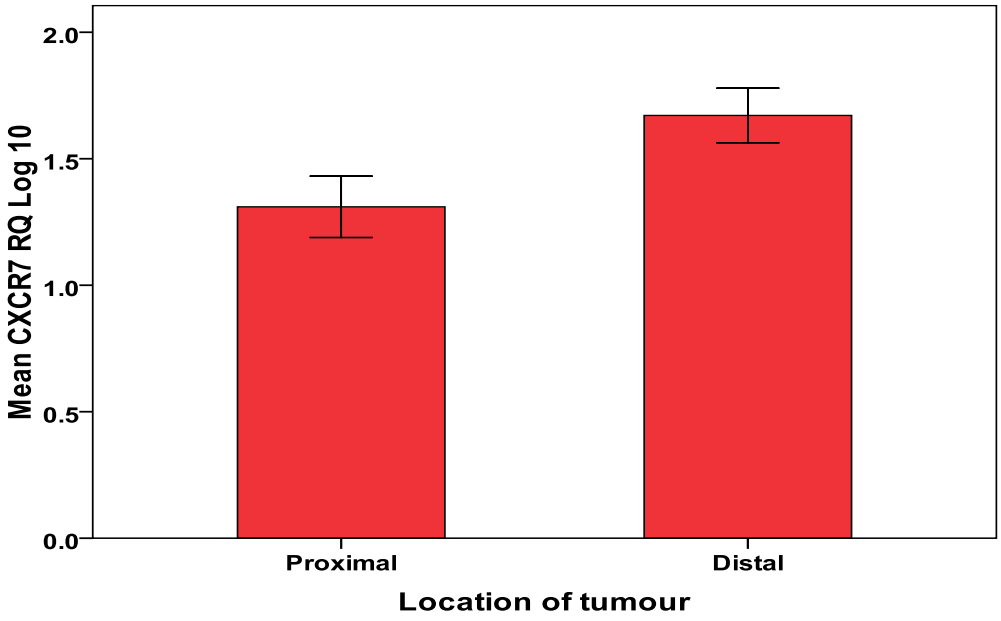


Figure 4.12: **Correlations of *CXCR7* expression levels**
CXCR7 was down-regulated in proximal colon (A, *p=0.021*) and its reduced expression was associated with lymphovascular invasion *(B, p=0.020)*

(A)



(B)

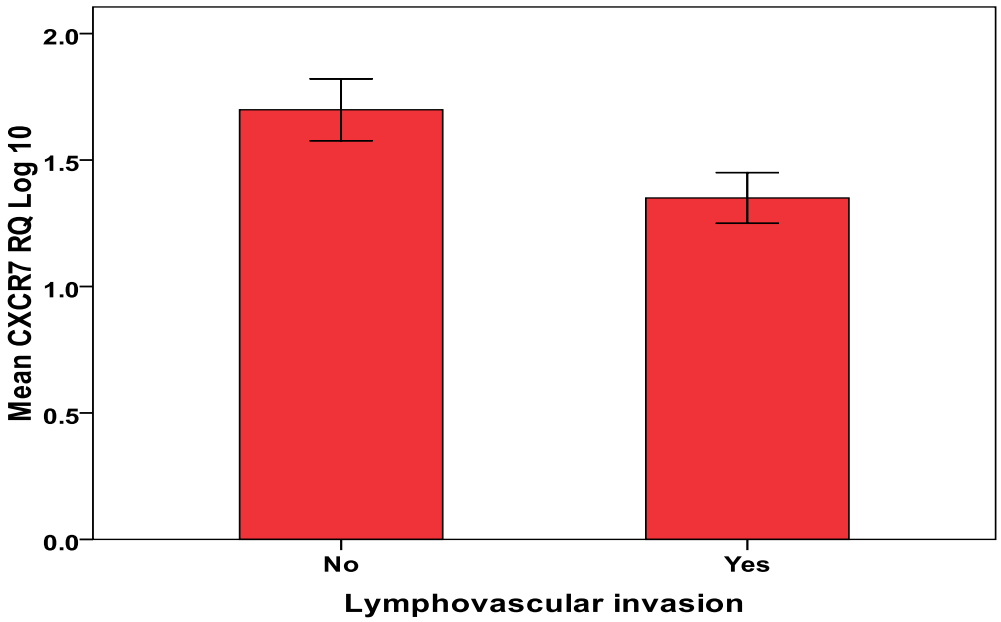
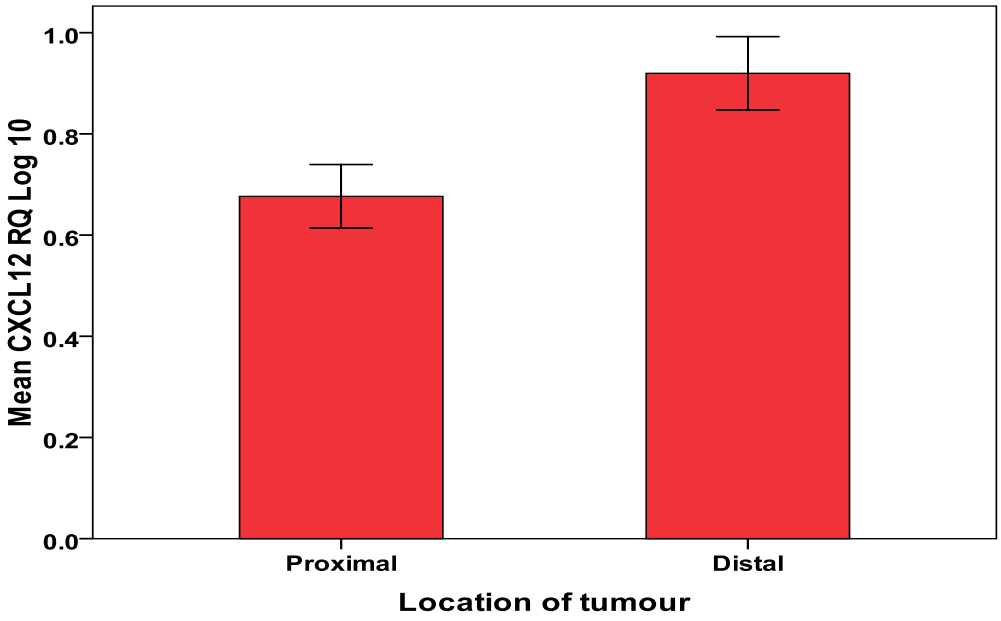


Figure 4.13: **Correlations of *CXCL12* expression levels**
The expression levels of *CXCL12* were down-regulated in colon compared to rectum (A, *p=0.015*) and in proximal compared to distal colon (B, *p=0.020).* Reduced expression of CXCL12 was associated with poorly differentiated tumour (C, *p=0.043*), local *(D, p=0.019*) and lymphovascular invasion (E, *p=0.033*) and advanced lymph nodes stage (F, *p=0.040*)

(A)



(B)

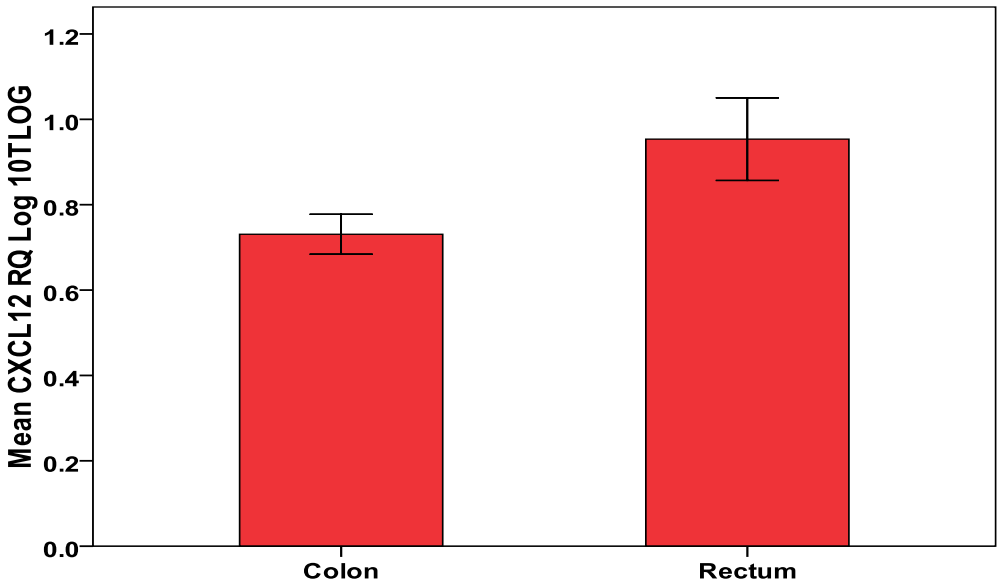
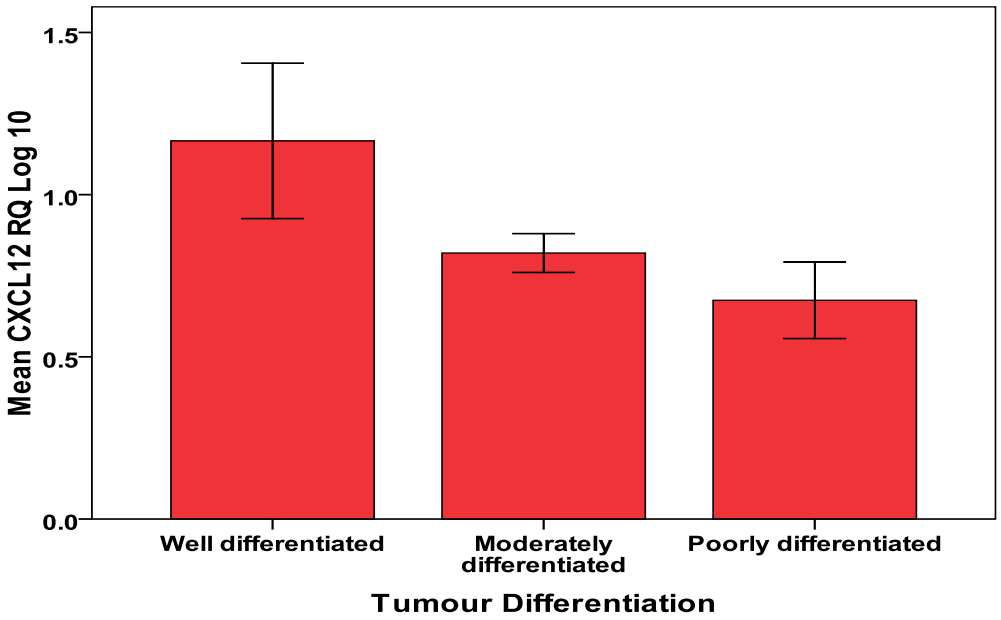


Figure 4.13: Continued

(C)



(D)

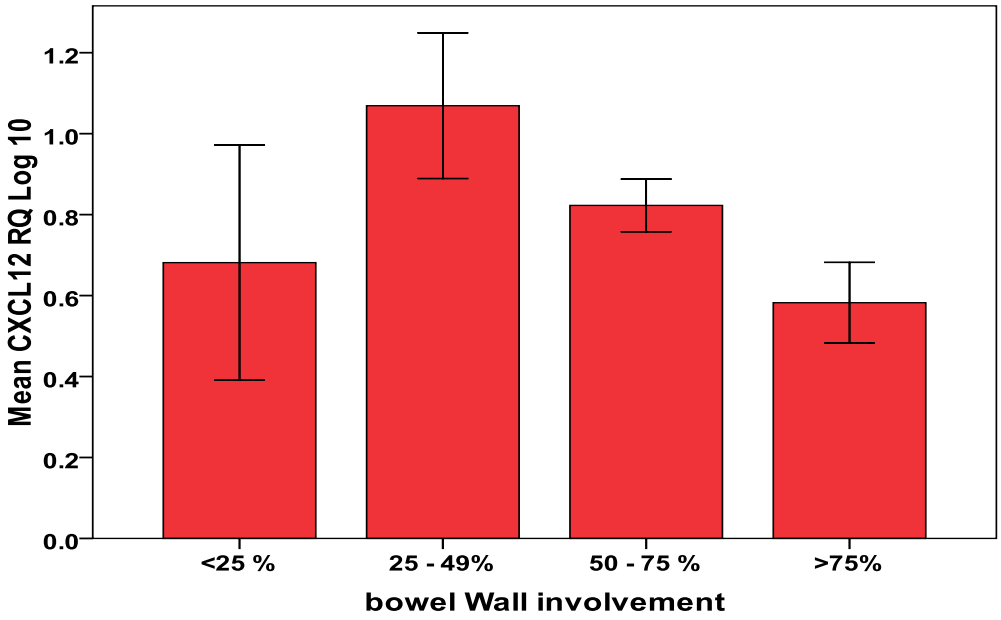
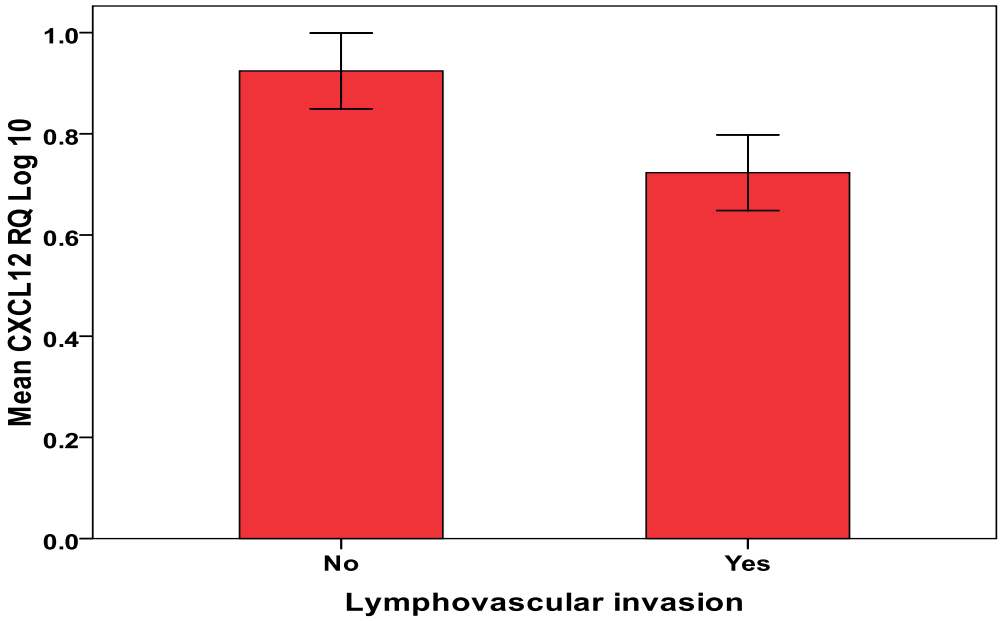
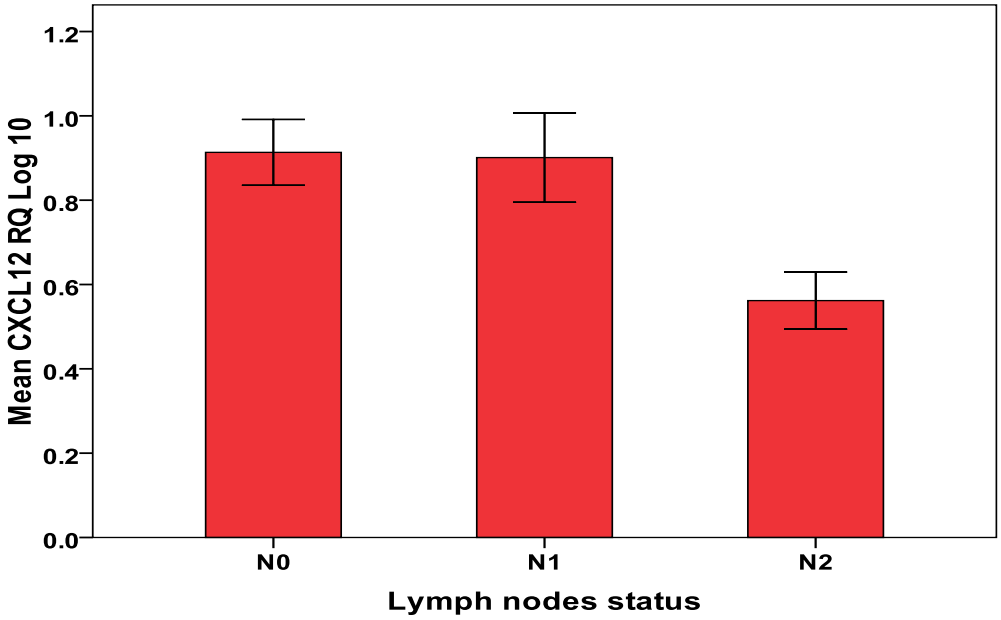


Figure 4.13: Continued

(E)



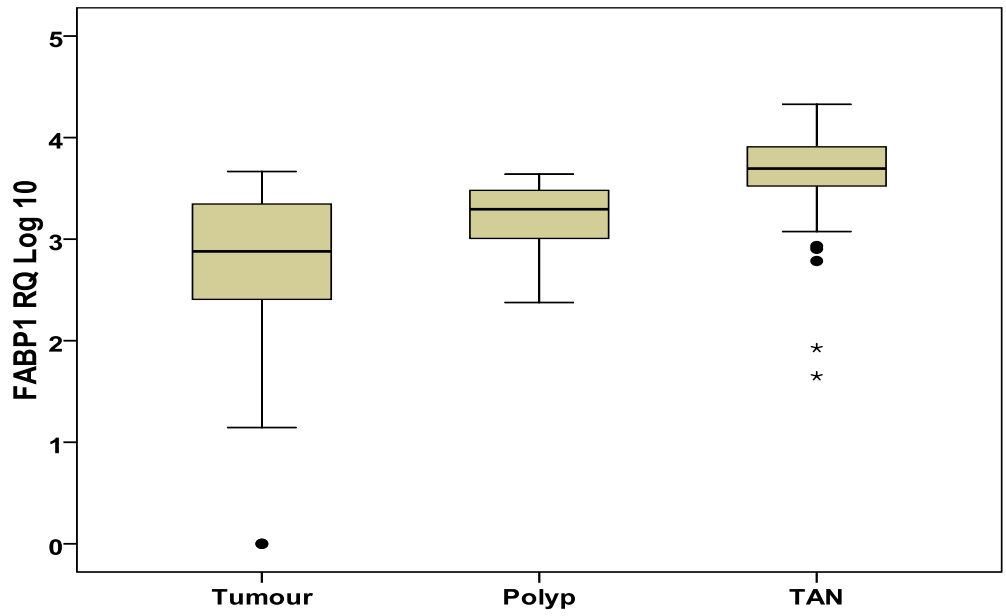
(F)



Fatty-acid binding protein 1, liver (FABP1)

Reduced expression of *FABP1* was observed in a progressive manner from TAN, to polyp, to tumour (*p<0.001*, Kruskal-Wallis t-test, table 4.6). Between groups analysis revealed significant differences in *FABP1* expression levels between tumour and TAN ($p<0.001$) and between polyps and TAN ($p=0.001$), but not between tumours and polyp ($p=0.055$). There was no significant association of FABP1 with other clinicopathological variables of the colorectal tumours (figure 4.14, table 4.5)

Figure 4.14: ***FABP1* expressions in colorectal tumours and TAN**
Reduced expression of *FABP1* was observed in a progressive manner from TAN, to polyp, to tumour



Interleukin 8 (IL-8)

Expression levels of IL-8 increased progressively from tumour-associated normal, to polyps, to tumours (*p<0.001*, ANOVA). Post-Hoc Tukey analysis revealed significant differences in *IL-8* expression levels between tumour and TAN ($p<0.001$) and between polyps and TAN ($p=0.025$), but not between tumours and polyp ($p=0.068$) (table 4.6, figure 4.15).

Although the expression of *IL-8* increased in tumours compared to normal colorectal tissues, its reduced expression was significantly associated with poor differentiation (*p=0.008*), advanced nodal stage (p=*0.015*) and disease recurrence (*p=0.036*) (ANOVA, figure 4.16 A and B). A non-significant trend of reduced IL-8 expression was also associated with perineural (*p=0.670*) and lymphovascular invasion (*p=0.687*), advanced Dukes' stage (*p=0.425*) and distal metastasis (*p=0.062*) (ANOVA, table 4.5)

Figure 4.15: ***IL8* expression in colorectal tumours and TAN**
IL-8 expression increased progressively from tumour-associated normal, to polyps, to tumours (A, *p<0.001*).

(A)

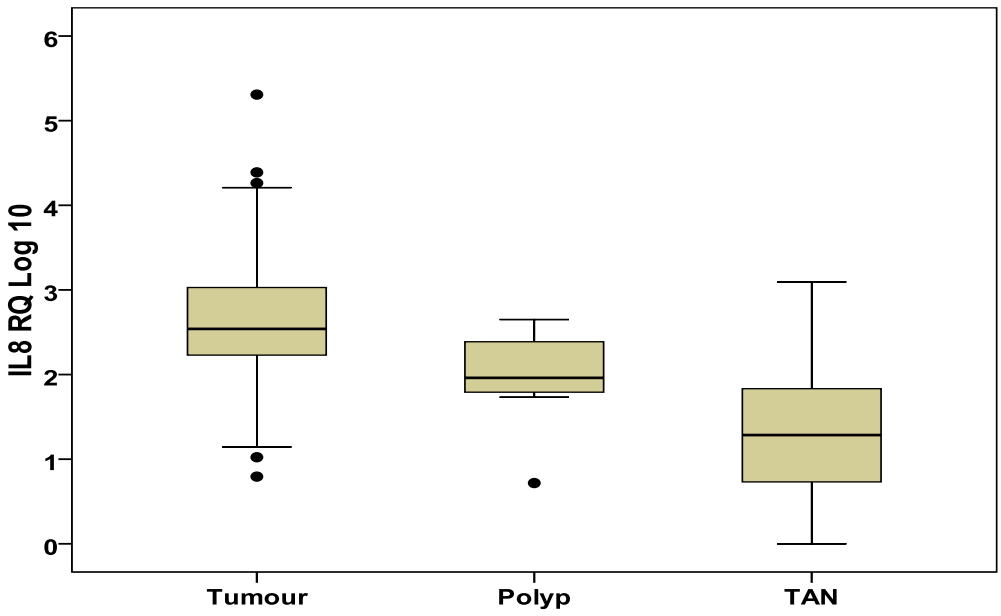
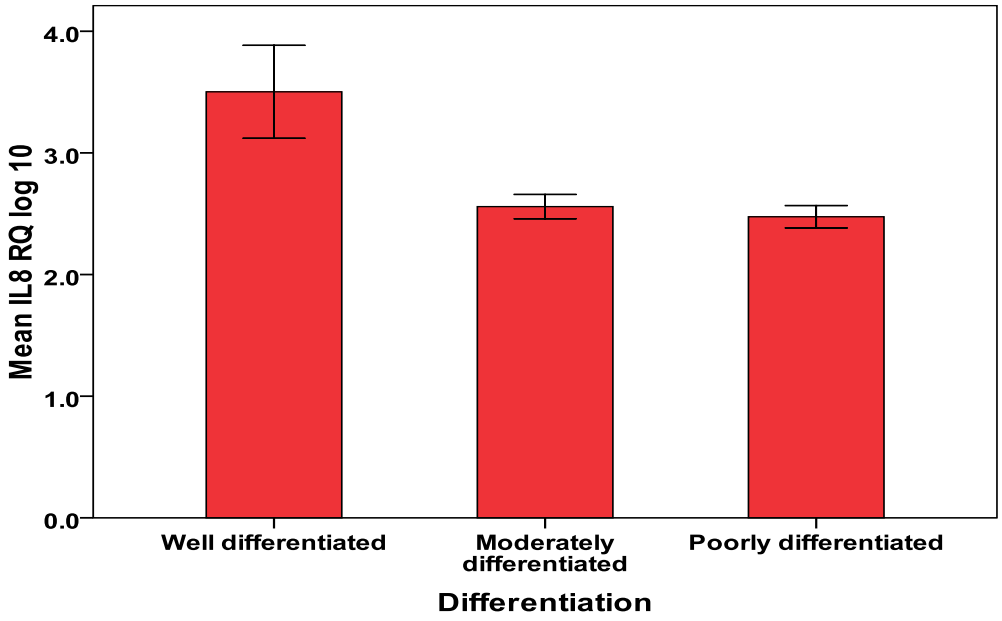
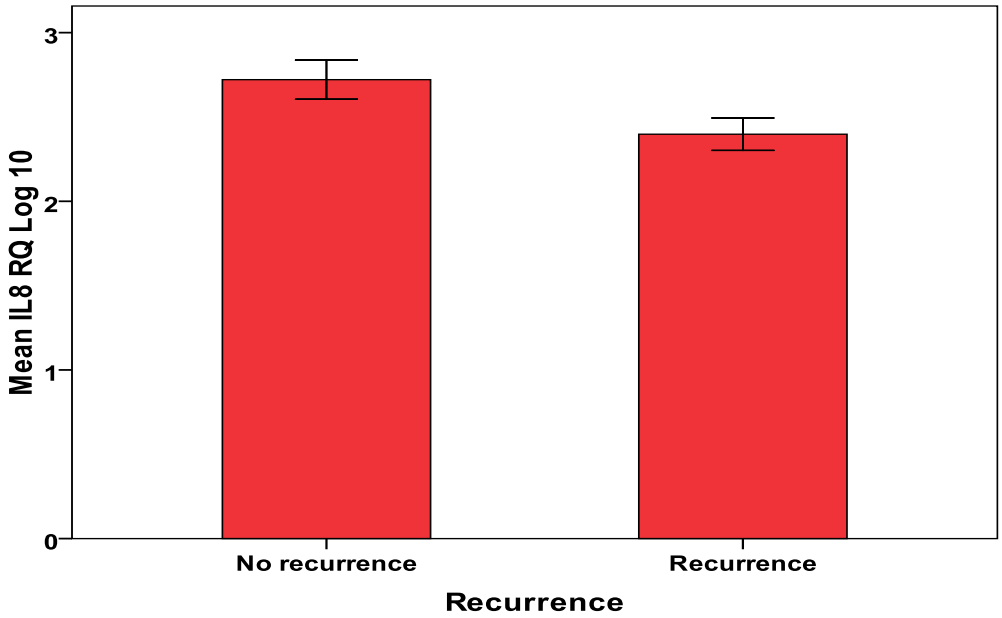


Figure 4.16: **Correlations of *IL8* expression levels**
Reduced IL-8 expression was associated with poor differentiation (B, *p=0.008*) and disease recurrence (C, *p=0.036*)

(A)



(B)



Mucin 2 (MUC2)

Again a progressive manner of expression from tumour, to polyp, to tumour associated normal was observed in *MUC2* (*p<0.001*, Kruskal-Wallis t-test, Figure 4.16A)). Further analysis confirmed a significant differences in *MUC2* expression levels between tumour and TAN ($p<0.001$) but not between polyps and TAN (*p=0.081*), and between tumours and polyp (*p=0.218*) (table 4.6). MUC2 expression was higher in mucinous tumours compared to non-mucinous (*p=0.013*, Mann-Whitney test); however, it was reduced in patients with high CA19.9 serum level (*p=0.037*) (Mann-Whitney test, figure 4.17B).

Figure 4.17: ***MUC2* expression in colorectal tumours and TAN**
(A) A progressive manner of expression from tumour, to polyp, to tumour associated normal was observed in *MUC2 (p<0.001*). (B) In regards to clinicopathological parameters, higher expression of *MUC2* was noted in mucinous tumours (*p=0.013*)

(A)

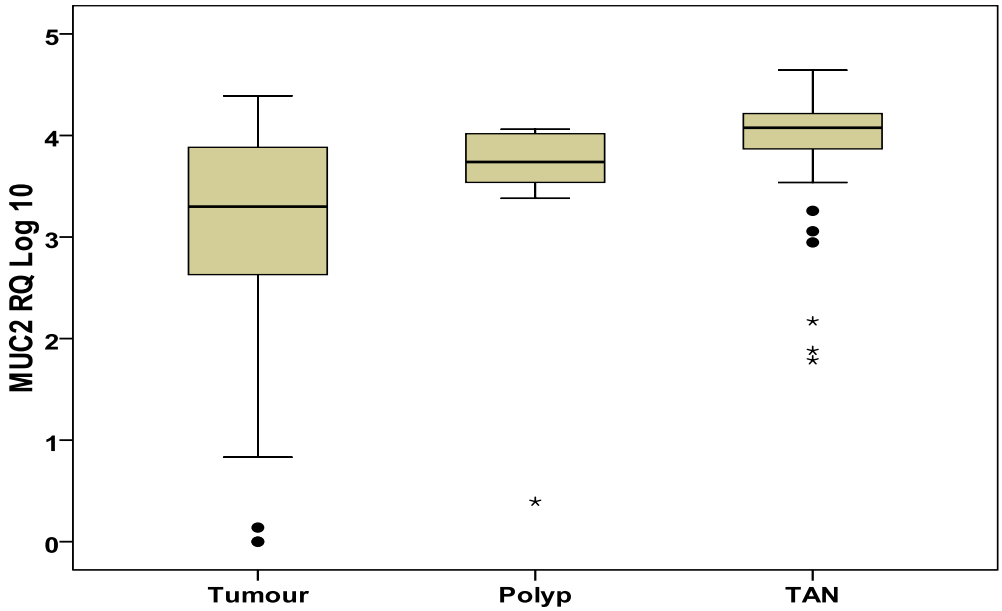
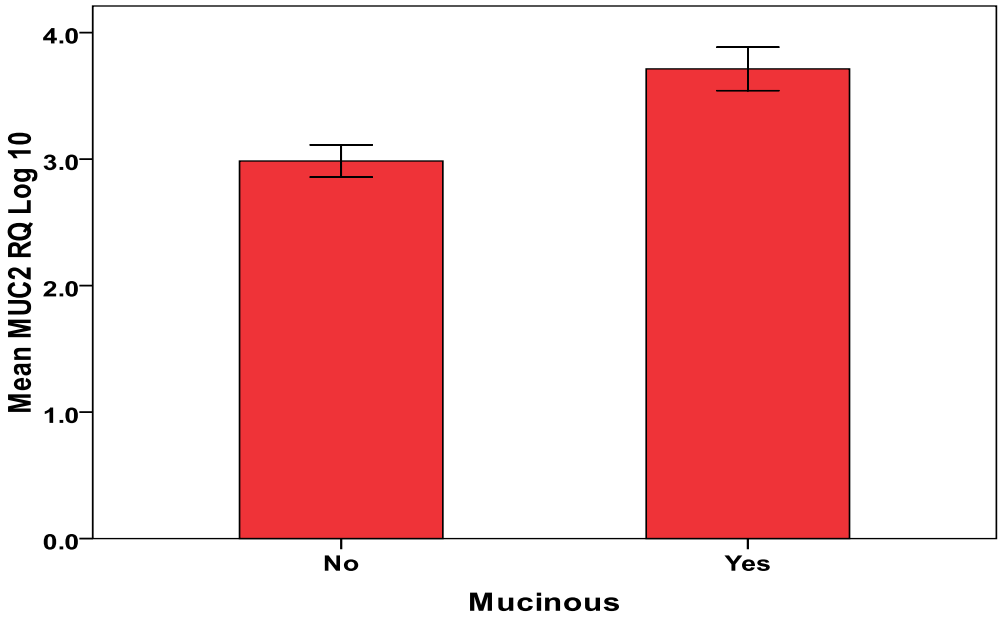


Figure 4.17: Continued

(B)

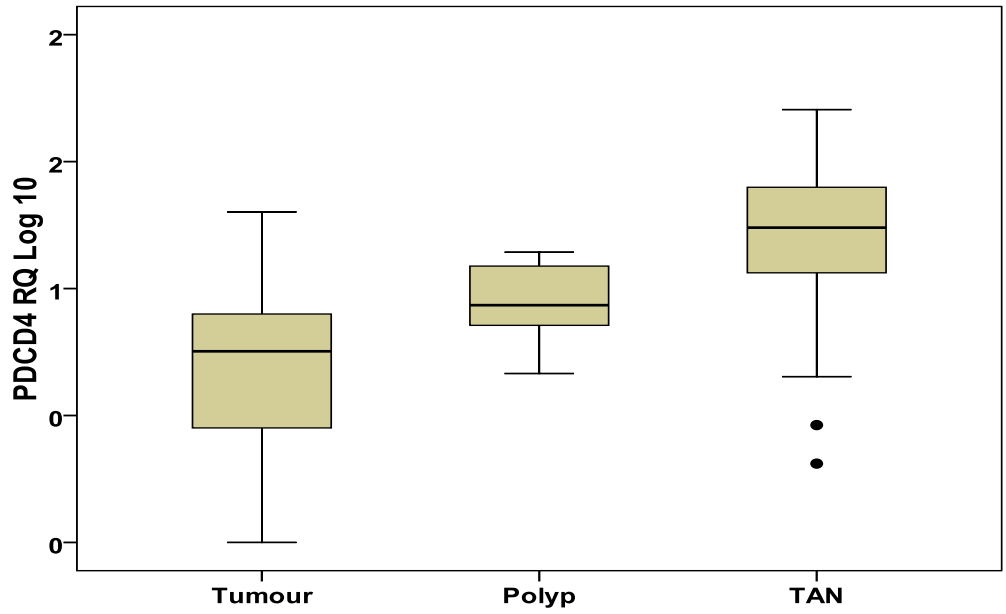


Programmed cell death 4 (PDCD4)

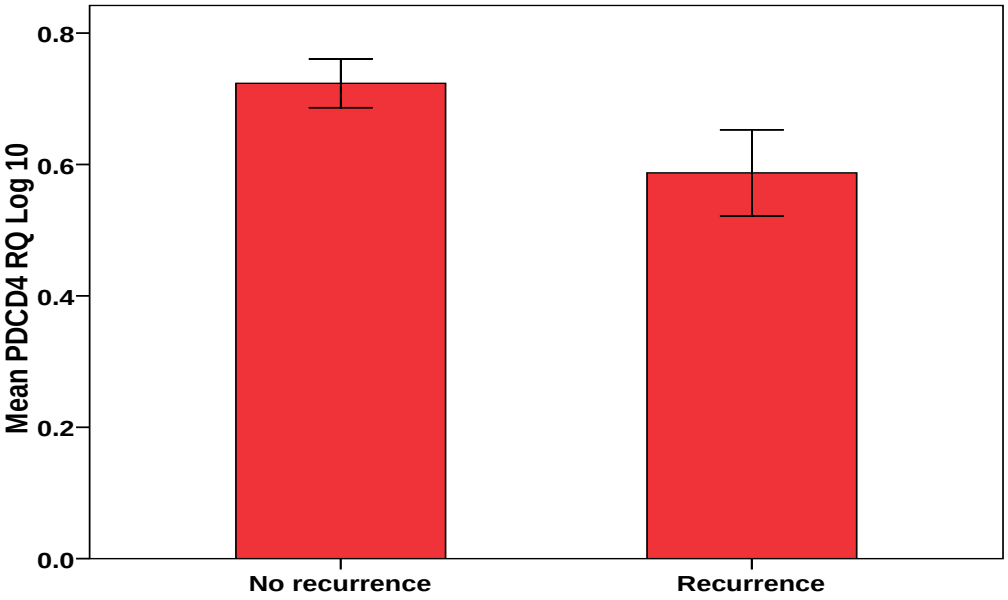
PDCD4 showed step-wise increase in expression from tumours, to polyps, to tumour associated normal tissues (*p<0.001*, ANOVA, figure 4.18A). Further between groups analysis (Post-Hoc Tukey test) identified significant differences in expression between tumour and TAN (*p<0.001*) and between polyp and TAN (*p=0.002*) but not between tumour and polyp (*p=0.065*). Additionally, down-regulation of PDCD4 was significantly associated with proximal colon tumours *(p=0.007*), tumour recurrence (*p=0.023*) and raised CA19.9 serum level (p=*0.003*) (t-test, figure 4.18B, table 4.5)

Figure 4.18: ***PDCD4* expression in colorectal tumours and TAN**
(A) Step-wise increase in expression from tumour, to polyp, to TAN *(p<0.001*). (B) Low expression of *PDCD4* was associated with tumours recurrence (*p=0.023*)

(A)



(B)



TGFB1 and its receptors (TGFBR1 and TGFBR2)

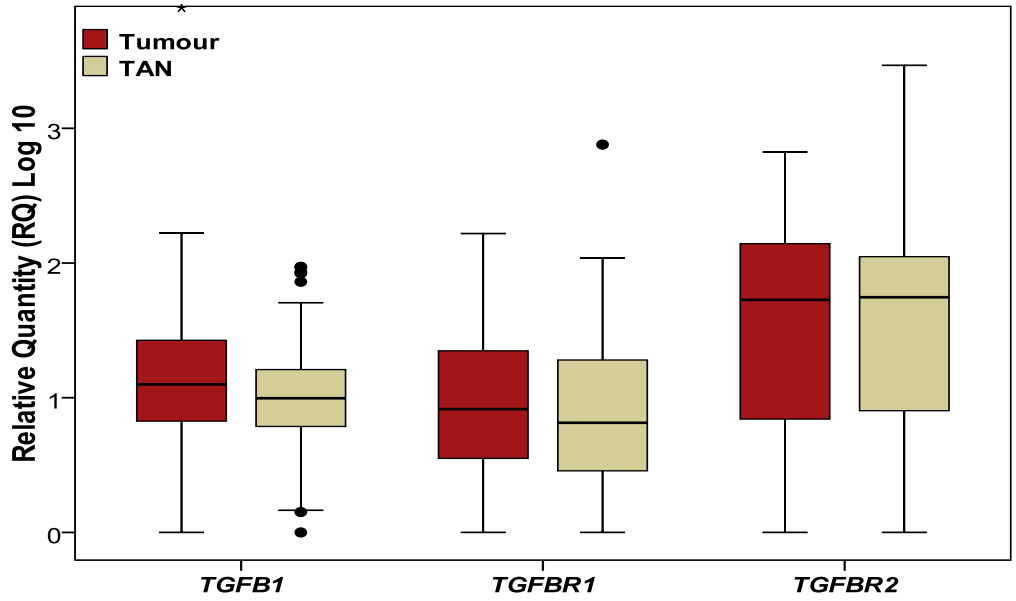
TGFB1 expression levels were higher in tumour compared to TAN tissues (*p=0.109,* paired t-test, figure 4.19A) in contrast to the expression of its receptors *TGFBR1* and *TGFBR2* whish showed low expression trends in tumour compared to TAN (*p=0.044* and *0.460* respectively, paired t-test, table 4.6, figure 4.18). Interestingly, TGFB1 expression showed step-wise increase from polyp, to normal, to tumour (*p=0.016*, ANOVA). Further analysis (Post-Hoc Tukey test) pointed out significant differences in expression between tumours and polyps (*p=0.029*), but not between tumours and TAN (*p=0.345*) and between polyps and TAN (*p=0.914*) (figure 4.19 B).

The relationship between *TGFB1*, *TGFBR1* and *TGFBR2* was further investigated using Pearson correlation. No violation of the assumption of normality, linearity and homogenecity was ensured before conducting further analysis. There was positive correlation between all the variables in both tumour and TAN colorectal tissues with high expression level of the ligand associated with high expression of the receptors (Figure 4.20).

The relation of *TGFB1* and its receptors expression levels and the clinico-pathological parameters were examined using ANOVA and t-test (figure 4.21). Although high level of *TGFB1* was documented in tumours compared to normal colorectal tissues, we noticed an association of TGFB1 down-regulation and lymphovascular invasion (*p=0.035*). Both *TGFBR1* and *TGFBR2* were under-expressed in proximal colon, however, the difference was only significant for *TGFBR2 (p=0.003)*. *TGFBR1* showed reduced expression in association with advanced disease clinicopathological parameters like tumour size, poor differentiation, advanced nodal stage, advanced Dukes' stage and tumour invasion and metastasis (table 4.4), However, these associations were only significant in relation to bowel wall involvement (*p<0.001*), and raised CEA serum level (*p=0.045*). Down-regulation of *TGFBR2* was significantly associated with increased bowel wall involvement (*p=0.006*), in colon cancer compared to rectal cancer (*p=0.031*) and in association with perineural (*p=0.030*) and lymphovascular invasion (*p=0.012*).

Figure 4.19: ***TGFB1, TGFBR1* and *TGFBR2* expression in tumours and TAN**
(A) High expression of *TGFB1* in tumours compared to TAN (*p=0.109*), in contrast to *TGFBR1* and *TGFBR2* which were down-regulated in tumours (*p=0.044* and *0.460*, respectively). (B) Reduced expression in early colorectal tumourogenesis and increased during late tumour progression (*p=0.016*, ANOVA).

(A)



(B)

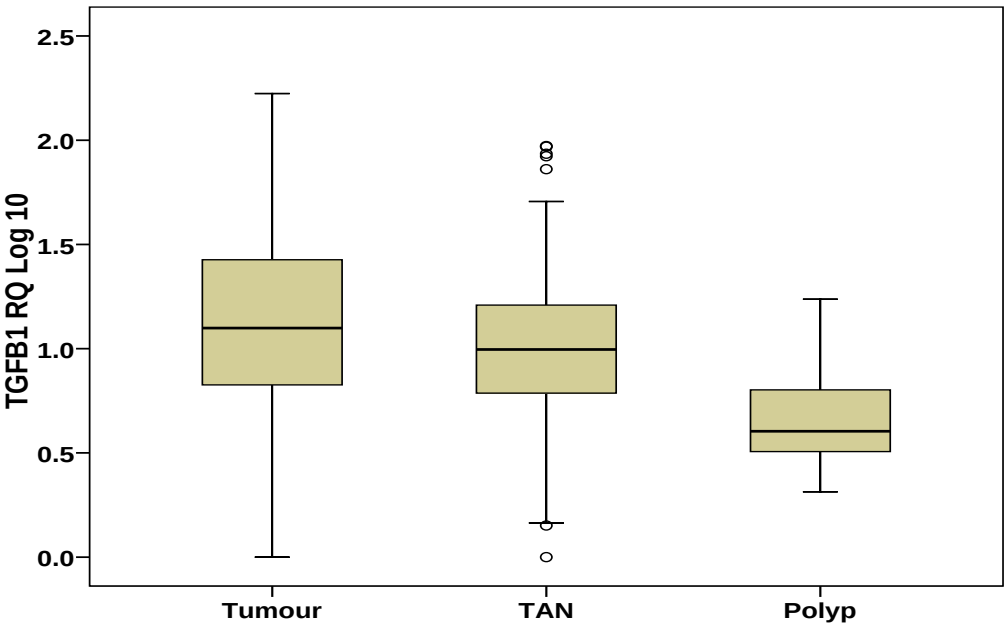
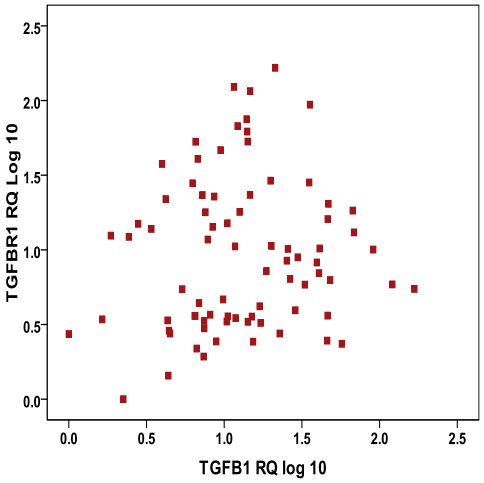
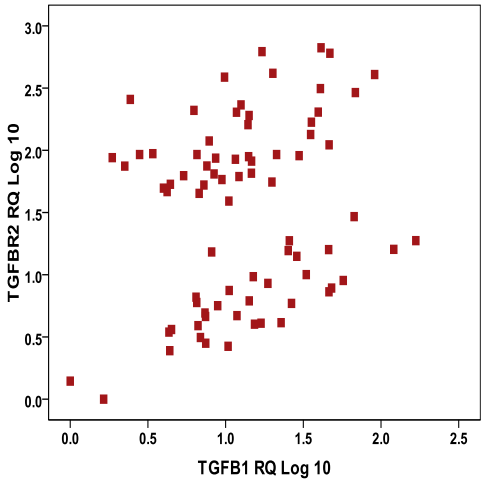


Figure 4.20: Correlations of expression levels of *TGFB1* and its receptors
Pearson's correlations of expression levels of *TGFB1, TGFBR1 and TGFBR2* in tumour (A, B, C) and tumour-associated normal (D, E, F) colorectal tissues.

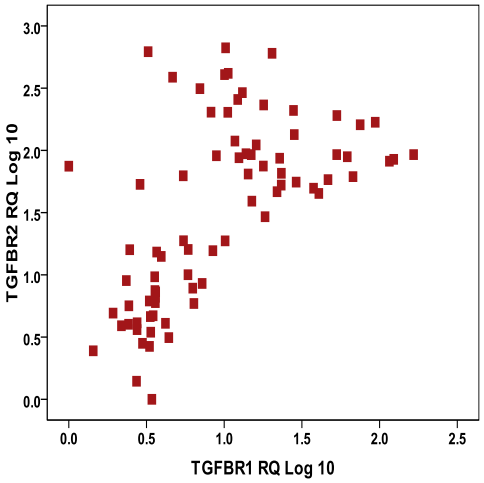
(A) *(r= 0.40, p<0.001)*



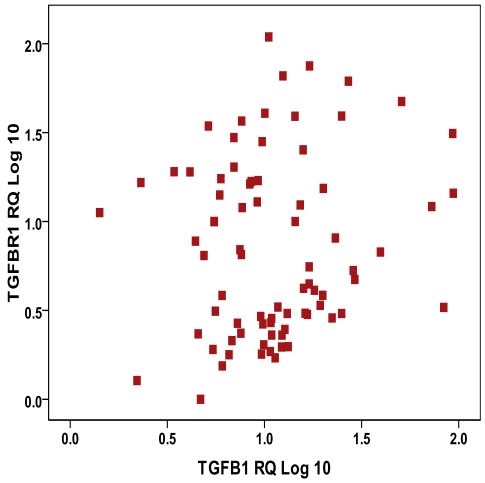
(B) *(r= 0.42, p<0.001)*



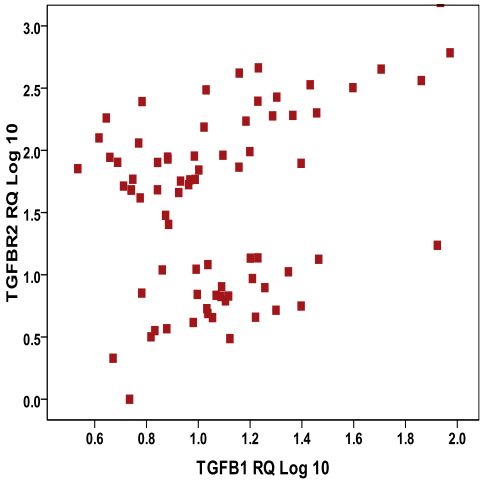
(C) *(r= 0.70, p<0.001)*



(D) *(r= 0.23, p=0.043)*



(E) *(r= 0.30, p=0.007)*



(F) *(r= 0.24, p=0.037)*

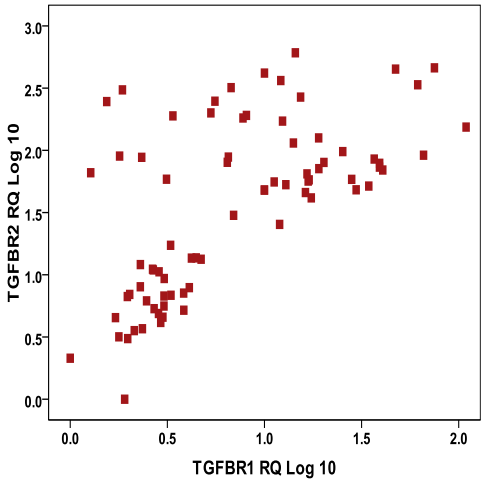
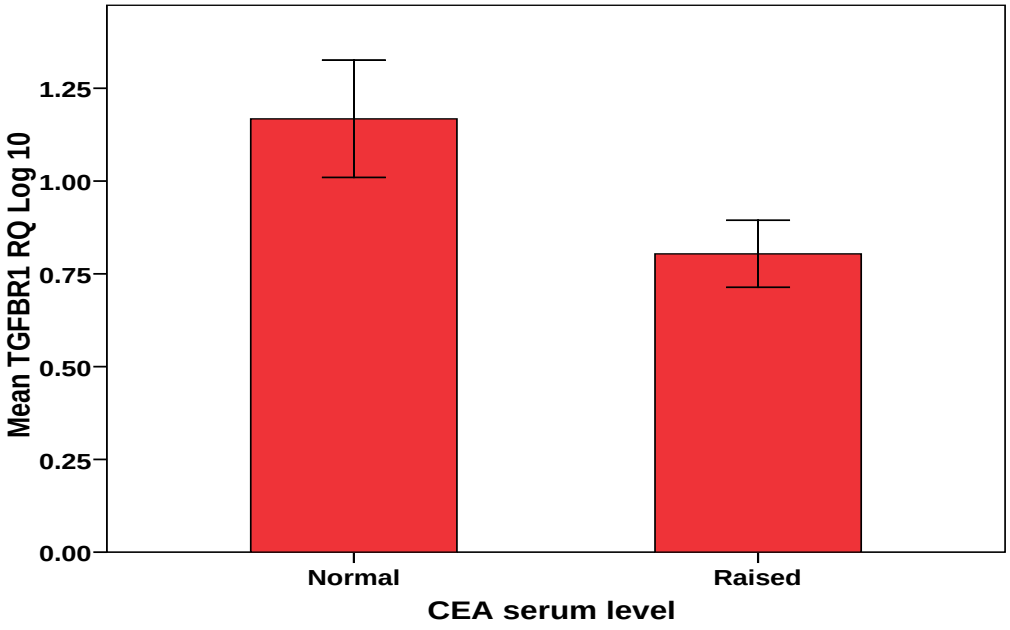
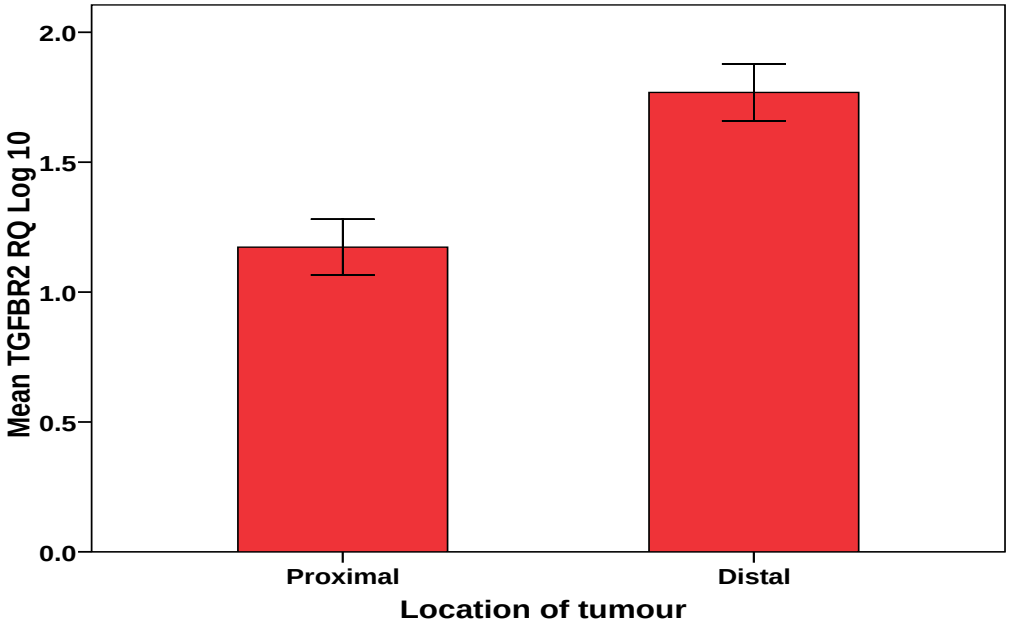


Figure 4.21: **Correlations of *TGFB1, TGFBR1* and *TGFBR2* expression levels** (A) Reduced expression of *TGFBR1* in patients with increased serum CEA level (*p=0.045*). (B, C and D) Expression of *TGFBR2* was lower in proximal compared to distal colon (*p=0.003*) and in association with perineural (*p=0.030*) and lymphovascular invasion (*p=0.012*)

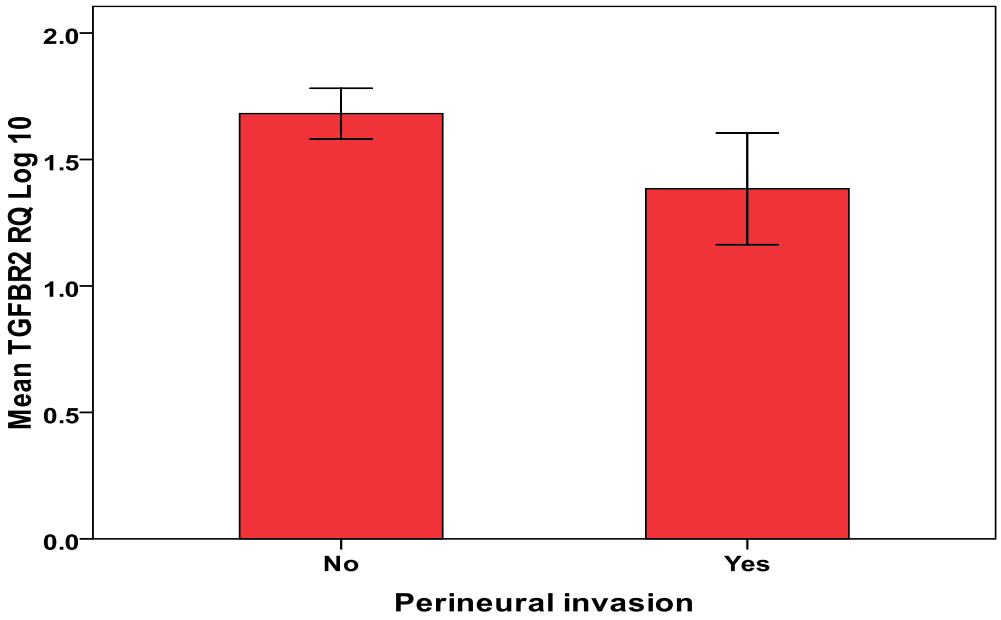
(A)



(B)



(C)


Mean TGFBR2 RQ Log 10
2.0
1.5
1.0
0.5
0.0
No
Yes
Perineural invasion

(D)

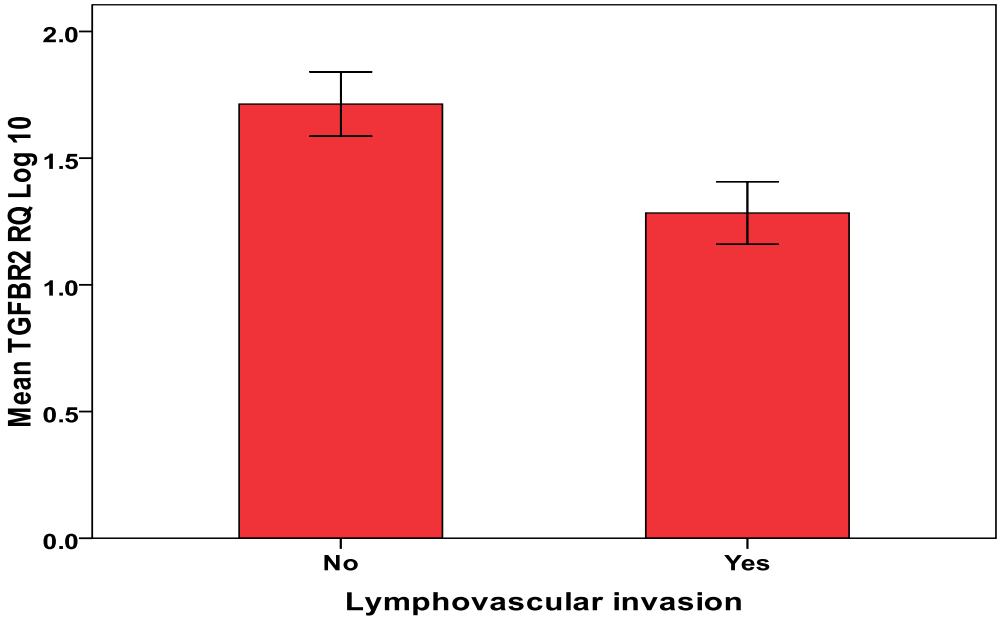

Mean TGFBR2 RQ Log 10
2.0
1.5
1.0
0.5
0.0
No
Yes
Lymphovascular invasion

Table 4.5: Gene expression and clinicopathological correlations

Variable	*CDH17*		*CXCL12*		*CXCR4*		*CXCR7*		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Tumour diameter		0.043*		0.481		0.860		0.035*	ANOVA/ Kruskal-Wallis H
<30 mm	2.937		0.917		0.977		1.490		
30-40 mm	3.154		0.858		0.935		1.848		
41-50 mm	3.860		0.573		0.970		0.760		
>50 mm	3.215		0.743		0.881		1.356		
Tumour thickness		0.035*		0.094		0.242		0.036*	ANOVA/ Kruskal-Wallis H
<10 mm	2.982		0.975		0.903		1.527		
10-15 mm	3.029		0.868		1.036		1.763		
>15 mm	3.842		0.650		0.847		1.278		
bowel Wall involvement		0.002*		0.019*		0.005*		0.002*	ANOVA/ Kruskal-Wallis H
<25%	3.689		0.681		1.080		0.927		
25-49%	2.945		1.133		1.163		1.714		
50-75%	2.993		0.823		0.933		1.719		
>75%	3.519		0.582		0.684		0.929		
Tumour location		0.473		0.020*		0.381		0.021*	t-test/ Mann-Whitney U
Proximal colon	3.027		0.676		0.903		1.310		
Distal colon	3.138		0.959		0.983		1.689		
Tumour location		0.706		0.015*		0.559		0.181	t-test/ Mann-Whitney U
Colon	3.081		0.731		0.935		1.456		
rectum	3.124		1.005		0.981		1.679		
Tumour differentiation		0.045*		0.043*		0.043*		0.596	ANOVA/ Kruskal-Wallis H
Well	2.775		1.166		1.286		1.827		
Moderate	3.227		0.820		0.920		1.535		
Poor	2.722		0.674		0.849		1.486		
Mucin secretion		0.407		0.342		0.272		0.679	t-test/ Mann-Whitney U
No	3.125		0.843		0.934		1.584		
Yes	2.977		1.019		1.095		1.487		
Depth of invasion		0.587		0.001*		0.093		0.485	ANOVA/ Kruskal-Wallis H
Tx (can't be assessed)	2.414		1.987		1.281		2.145		
Tis (carcinoma in-situ)	3.120		1.451		1.346		1.417		
T1 (Submicosal)	3.079		0.536		0.879		1.557		
T2 (Muscularis propria)	2.873		1.192		1.228		1.903		
T3 (subserosa)	3.111		0.867		0.865		1.573		
T4 (adjacent structures)	3.223		0.711		0.947		1.416		

Variable	*CDH17*		*CXCL12*		*CXCR4*		*CXCR7*		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Lymph nodes status		0.175		0.040*		0.059		0.287	ANOVA/ Kruskal-Wallis H
N0 (No metastasis)	3.079		0.913		0.985		1.576		
N1 (1-3 nodes)	3.016		0.901		1.023		1.673		
N2 (>3 nodes)	3.404		0.562		0.739		1.295		
Distant metastasis		0.434		0.163		0.044*		0.138	t-test/ Mann-Whitney U
No	3.097		0.913		1.006		1.626		
Yes	3.121		0.723		0.796		1.376		
Perineural invasion		0.180		0.389		0.126		0.904	t-test/ Mann-Whitney U
Yes	3.284		0.812		0.855		1.648		
No	2.989		0.923		1.026		1.618		
Lymphovascular invasion		0.016*		0.033*		0.132		0.020*	t-test/ Mann-Whitney U
Yes	3.334		0.723		0.877		1.350		
No	2.945		0.970		1.013		1.718		
Tumour stage		0.949		0.016*		0.253		0.749	ANOVA/ Kruskal-Wallis H
Stage 0	3.120		1.451		1.346		1.417		
Stage I	3.001		0.886		1.054		1.655		
Stage II	3.164		0.925		0.942		1.624		
Stage III	3.085		0.800		1.012		1.582		
Stage VI	3.121		0.723		0.796		1.376		
pCR	2.414		1.987		1.281		2.145		
Dukes' stage		0.877		0.492		0.201		0.711	ANOVA/ Kruskal-Wallis H
A	3.103		0.938		1.081		1.633		
B	3.164		0.925		0.942		1.624		
C	3.085		0.800		1.012		1.582		
D	3.121		0.723		0.796		1.376		
Recurrence		0.891		0.169		0.058		0.124	t-test/ Mann-Whitney U
Yes	3.022		0.754		0.821		1.380		
No	3.134		0.915		1.011		1.642		
CEA Group		0.045*		0.305		0.399		0.183	t-test/ Mann-Whitney U
Normal	3.023		0.952		1.044		1.361		
Raised	3.401		0.791		0.940		1.706		
CA 19.9 Group		0.014*		0.230		0.310		0.971	t-test/ Mann-Whitney U
Normal	3.103		1.054		1.076		1.592		
Raised	2.739		0.797		0.924		1.582		

Table 4.5: Continued

Variable	*CEACAM5*		*FABP1*		*IL8*		*MUC2*		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Tumour diameter		0.213		0.449		0.285		0.271	ANOVA/ Kruskal-Wallis H
<30 mm	2.200		2.739		2.678		3.157		
30-40 mm	2.481		2.812		2.658		3.201		
41-50 mm	3.952		2.925		1.755		3.851		
>50 mm	2.639		2.884		2.588		2.883		
Tumour thickness		0.282		0.049		0.616		0.654	ANOVA/ Kruskal-Wallis H
<10 mm	2.344		2.818		2.620		3.298		
10-15 mm	2.438		2.554		2.712		2.943		
>15 mm	2.998		3.069		2.487		3.257		
bowel Wall involvement		0.002*		0.949		0.055		0.400	ANOVA/ Kruskal-Wallis H
<25%	3.698		2.951		1.849		2.984		
25-49%	2.194		2.714		3.001		3.462		
50-75%	2.132		2.853		2.602		2.996		
>75%	3.567		2.708		2.308		3.187		
Tumour location		0.858		0.466		0.285		0.711	t-test/ Mann-Whitney U
Proximal colon	2.436		2.669		2.501		2.966		
Distal colon	2.416		2.870		2.683		3.186		
Tumour location		0.741		0.541		0.914		0.935	t-test/ Mann-Whitney U
Colon	2.396		2.822		2.635		3.016		
rectum	2.447		2.792		2.616		3.214		
Tumour differentiation		0.016*		0.109		0.008*		0.910	ANOVA/ Kruskal-Wallis H
Well	1.673		2.435		3.502		2.945		
Moderate	2.646		2.883		2.558		3.112		
Poor	1.893		2.671		2.475		3.041		
Mucin secretion		0.782		0.398		0.155		0.013*	t-test/ Mann-Whitney U
No	2.444		2.832		2.566		3.005		
Yes	2.302		2.667		2.957		3.734		
Depth of invasion		0.442		0.389		0.319		0.645	ANOVA/ Kruskal-Wallis H
Tx (can't be assessed)	1.326		2.456		2.336		4.161		
Tis (carcinoma in-situ)	1.942		2.951		2.528		3.540		
T1 (Submicosal)	2.481		2.648		2.386		3.381		
T2 (Muscularis propria)	1.811		2.653		3.156		3.292		
T3 (subserosa)	2.618		2.872		2.522		2.929		
T4 (adjacent structures)	2.494		2.827		2.617		3.146		

Variable	*CEACAM5*		*FABP1*		*IL8*		*MUC2*		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Lymph nodes status		0.071		0.976		0.015*		0.716	ANOVA/ Kruskal-Wallis H
N0 (No metastasis)	2.531		2.778		2.753		2.939		
N1 (1-3 nodes)	2.049		2.850		2.648		3.191		
N2 (>3 nodes)	2.953		2.848		2.318		3.271		
Distant metastasis		0.547		0.373		0.062		0.938	t-test/ Mann-Whitney U
No	2.369		2.862		2.693		3.127		
Yes	2.602		2.618		2.398		3.080		
Perineural invasion		0.154		0.057		0.670		0.443	t-test/ Mann-Whitney U
Yes	2.806		3.008		2.575		3.293		
No	2.189		2.696		2.683		3.025		
Lymphovascular invasion		0.019*		0.515		0.687		0.131	t-test/ Mann-Whitney U
Yes	2.792		2.893		2.583		3.342		
No	2.170		2.748		2.654		2.962		
Tumour stage		0.681		0.704		0.676		0.911	ANOVA/ Kruskal-Wallis H
Stage 0	1.942		2.951		2.528		3.540		
Stage I	2.222		2.760		2.724		3.241		
Stage II	2.728		2.849		2.801		2.985		
Stage III	2.176		2.948		2.608		3.118		
Stage VI	2.602		2.618		2.398		3.080		
pCR	1.326		2.456		2.336		4.161		
Dukes' stage		0.514		0.576		0.425		0.818	ANOVA/ Kruskal-Wallis H
A	2.196		2.778		2.706		3.268		
B	2.728		2.849		2.801		2.985		
C	2.176		2.948		2.608		3.118		
D	2.602		2.618		2.398		3.080		
Recurrence		0.801		0.170		0.036*		0.685	t-test/ Mann-Whitney U
Yes	2.391		2.591		2.400		3.102		
No	2.434		2.890		2.713		3.122		
CEA Group		0.081	2.746	0.220		0.198		0.120	t-test/ Mann-Whitney U
Normal	2.167		2.918		2.779		3.095		
Raised	2.972				2.520		3.415		
CA 19.9 Group		0.207		0.559		0.756		0.037*	t-test/ Mann-Whitney U
Normal	2.178		2.878		2.610		3.454		
Raised	1.672		2.814		2.677		2.703		

Table 4.5: Continued

Variable	PDCD4		TGFB1		TGFBR1		TGFBR2		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Tumour diameter		0.674		0.766		0.189		0.155	ANOVA/ Kruskal-Wallis H
<30 mm	0.639		1.155		1.035		1.571		
30-40 mm	0.725		1.247		1.112		1.698		
41-50 mm	0.821		1.091		0.422		0.864		
>50 mm	0.659		1.071		0.816		1.295		
Tumour thickness		0.562		0.348		0.317		0.234	ANOVA/ Kruskal-Wallis H
<10 mm	0.683		1.087		0.993		1.524		
10-15 mm	0.640		1.303		1.066		1.686		
>15 mm	0.739		1.098		0.779		1.262		
bowel Wall involvement		0.886		0.201		0.000*		0.006*	ANOVA/ Kruskal-Wallis H
<25%	0.622		1.165		0.535		0.799		
25-49%	0.704		1.236		1.017		1.696		
50-75%	0.683		1.165		1.152		1.755		
>75%	0.618		1.067		0.520		0.681		
Tumour location		0.007*		0.347		0.621		0.003*	t-test/ Mann-Whitney U
Proximal colon	0.593		1.266		0.957		1.209		
Distal colon	0.738		1.112		1.032		1.707		
Tumour location		0.976		0.741		0.873		0.031*	t-test/ Mann-Whitney U
Colon	0.693		1.127		1.019		1.352		
rectum	0.691		1.168		0.997		1.742		
Tumour differentiation		0.209		0.335		0.443		0.560	ANOVA/ Kruskal-Wallis H
Well	0.590		1.375		1.245		1.838		
Moderate	0.714		1.160		0.974		1.510		
Poor	0.550		0.972		0.864		1.427		
Mucin secretion		0.217		0.585		0.225		0.870	t-test/ Mann-Whitney U
No	0.674		1.135		0.967		1.555		
Yes	0.793		1.222		1.238		1.518		
Depth of invasion		0.487		0.828		0.217		0.116	ANOVA/ Kruskal-Wallis H
Tx (can't be assessed)	0.927		1.043		1.541		2.109		
Tis (carcinoma in-situ)	1.077		0.816		1.724		1.966		
T1 (Submicosal)	0.494		1.017		0.451		2.046		
T2 (Muscularis propria)	0.708		1.356		1.188		2.071		
T3 (subserosa)	0.696		1.135		0.974		1.407		
T4 (adjacent structures)	0.680		1.128		1.000		1.410		

Variable	PDCD4		TGFB1		TGFBR1		TGFBR2		Statistical test (parametric & non-parametric)
	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	Expression level	*p-value*	
Lymph nodes status		0.934		0.238		0.688		0.049*	ANOVA/ Kruskal-Wallis H
N0 (No metastasis)	0.675		1.260		0.995		1.572		
N1 (1-3 nodes)	0.687		1.084		1.058		1.735		
N2 (>3 nodes)	0.710		1.003		0.883		1.098		
Distant metastasis		0.443		0.861		0.161		0.235	t-test/ Mann-Whitney U
No	0.707		1.154		1.054		1.605		
Yes	0.641		1.128		0.852		1.359		
Perineural invasion		0.969		0.792		985		0.030*	t-test/ Mann-Whitney U
Yes	0.694		1.200		1.033		1.384		
No	0.691		1.152		1.038		1.698		
Lymphovascular invasion		0.600		0.035*		0.208		0.012*	t-test/ Mann-Whitney U
Yes	0.714		1.001		0.906		1.284		
No	0.677		1.284		1.077		1.730		
Tumour stage		0.649		0.490		0.534		0.371	ANOVA/ Kruskal-Wallis H
Stage 0	1.077		0.816		1.724		1.966		
Stage I	0.683		1.228		0.958		1.950		
Stage II	0.712		1.303		1.042		1.411		
Stage III	0.679		0.977		1.038		1.598		
Stage VI	0.641		1.128		0.852		1.359		
pCR	0.927		1.043		1.541		2.109		
Dukes' stage		0.880		0.287		0.749		0.222	ANOVA/ Kruskal-Wallis H
A	0.719		1.190		1.028		1.952		
B	0.712		1.303		1.042		1.411		
C	0.679		0.997		1.038		1.599		
D	0.641		1.128		0.852		1.359		
Recurrence		0.023*		0.476		0.126		0.213	t-test/ Mann-Whitney U
Yes	0.608		1.080		0.859		1.372		
No	0.725		1.174		1.066		1.618		
CEA Group		0.860		0.467		0.045*		0.062	t-test/ Mann-Whitney U
Normal	0.704		1.142		1.201		1.800		
Raised	0.687		1.279		0.804		1.287		
CA 19.9 Group		0.003*		0.820		0.995		0.550	t-test/ Mann-Whitney U
Normal	0.853		1.123		1.159		1.638		
Raised	0.491		1.170		1.158		1.837		

Table 4.6: Gene expression in tumours, polyps and TAN colorectal tissues

Gene	*P-Value*	Between groups differences			Gene	*P-Value*	Between groups differences		
CDH17	*0.005*	Tumour	TAN	0.002	***CEACAM5***	*0.981*	Tumour	TAN	0.851
			Polyp	0.837				Polyp	0.863
		Polyp	TAN	0.491			Polyp	TAN	0.662
CXCL12	*<0.001*	Tumour	TAN	<0.001	***CXCR4***	*0.437*	Tumour	TAN	0.479
			Polyp	0.907				Polyp	0.704
		Polyp	TAN	0.003			Polyp	TAN	0.984
CXCR7	*0.602*	Tumour	TAN	0.889	***FABP1***	*<0.001*	Tumour	TAN	<0.001
			Polyp	0.933				Polyp	0.055
		Polyp	TAN	0.994			Polyp	TAN	0.001
IL8	*<0.001*	Tumour	TAN	<0.001	***MUC2***	*<0.001*	Tumour	TAN	<0.001
			Polyp	0.068				Polyp	0.218
		Polyp	TAN	0.025			Polyp	TAN	0.081
PDCD4	*<0.001*	Tumour	TAN	<0.001	***TGFB1***	*0.016*	Tumour	TAN	0.354
			Polyp	0.064				Polyp	0.029
		Polyp	TAN	0.002			Polyp	TAN	0.914
TGFBR1	*0.044*	Tumour	TAN	0.640	***TGFBR2***	*0.460*	Tumour	TAN	0.640
			Polyp	0.679				Polyp	0.679
		Polyp	TAN	0.756			Polyp	TAN	0.756

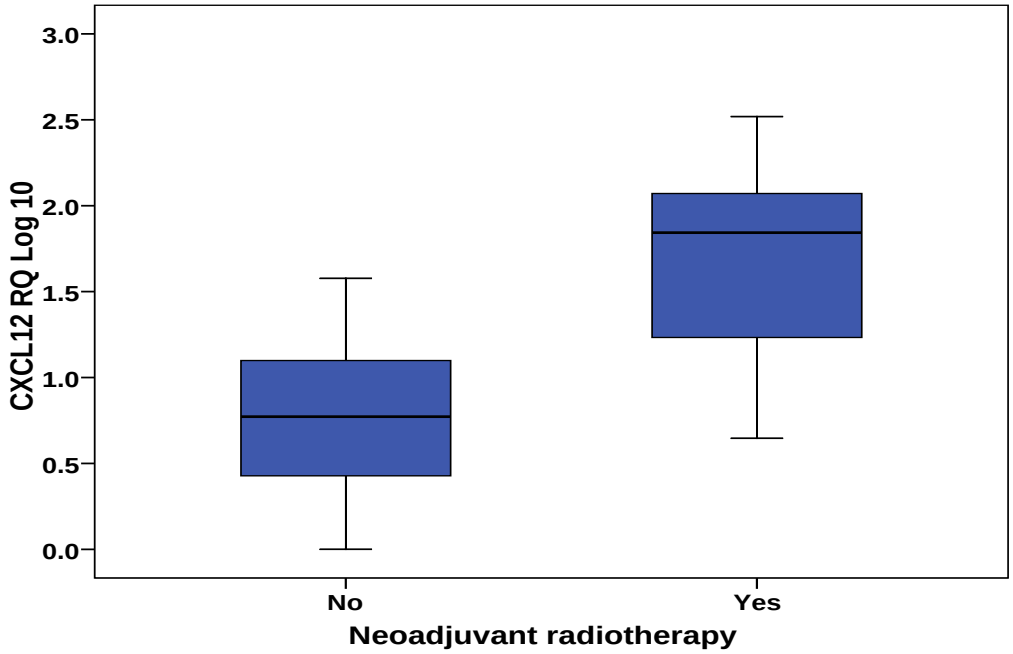
4.4.2 Neoadjuvant therapy and colorectal cancer genes expression

In the cohort of rectal cancer patients (n=58) we analysed the differences in gene expression in patients who had neoadjuvant chemoradiation (n=25) compared to those who did not (n=33) using t-test. Univariate analysis of variance was further conducted to test for interaction effect and to control for confounding factors. We demonstrated decrease expression of *CDH17* (p=0.020) and *CEACAM5 (p=0.032)* and increase expression of *CXCL12 (p<0.001), CXCR4 (p=0.004)* and *MUC2 (p=0.041)* in response to neoadjuvant therapy. However, the differences only persisted for *CXCL12 (p=0.035)* and CXCR4 *(p=0.001)* after univariate analysis (figure 4.21).

Interestingly, expression levels of *CDH17* (p=0.003), *CEACAM5* (p=0.036), *CXCL12* (p=<0.001) and *CXCR4 (*p=0.003*)* significantly correlated with Mandard tumour regression grade (TRG). Higher expression of *CXCL12* and *CXCR4* was noticed in good responders (TRG1, TRG2 and TRG3) compared to poor responders (TRG4 and TRG5) in contrast to the expression levels of *CDH17* and *CEACAM5* which were lower in good responders (ANOVA test, figure 4.22)

Figure 4.22: Dysregulation of gene expression in response to neoadjuvant CRT
Neoadjuvant chemoradiation associated with significant up-regulation of *CXCL12* (A, univariate analysis, *p=0.035*) and *CXCR4* (B, univariate analysis, *p=0.001*) expression.

(A)



(B)

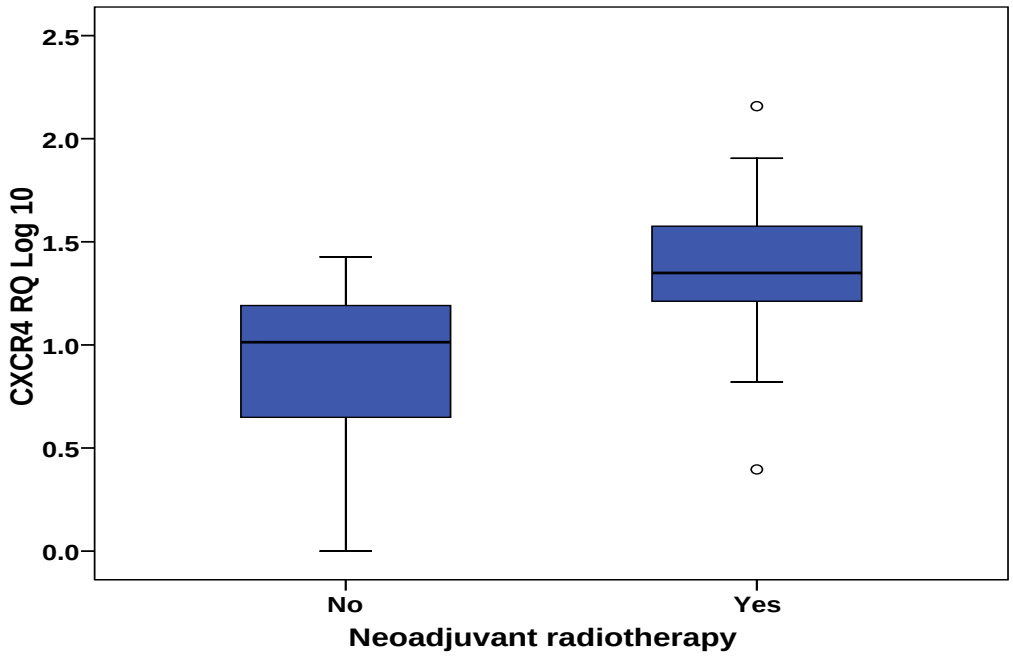
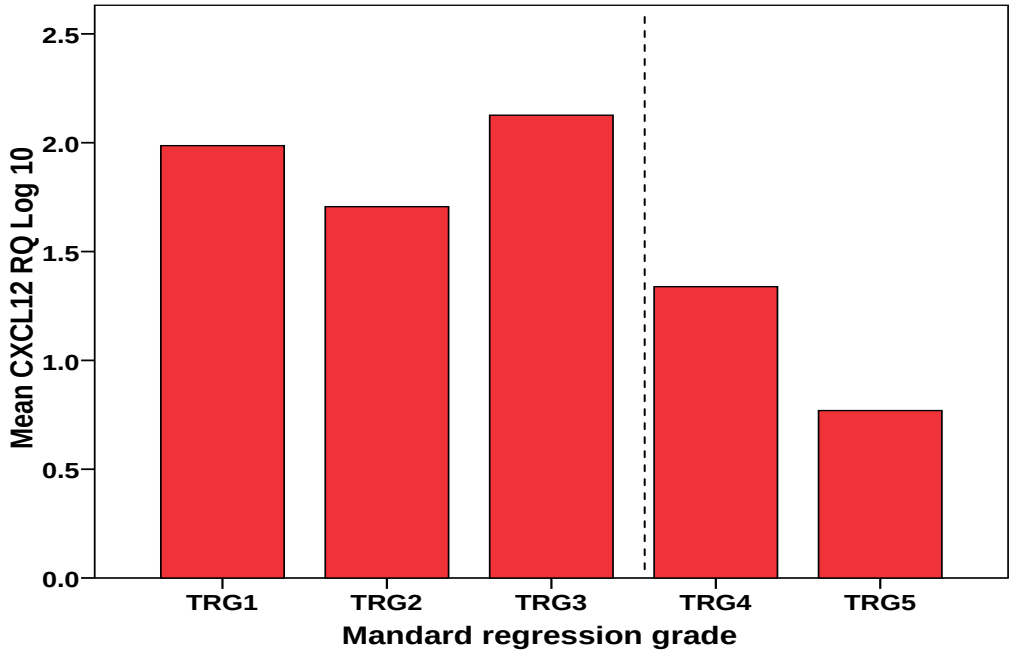


Figure 4.23: Correlation of gen*e* expression with TRG
Increased expression of *CXCL12 (A, p<0.001)* and *CXCR4* (B, p=0.003) was associated with lower TRG (good response) in contrast to *CDH17* (C, p=0.003) and *CEACAM5 (D, p=0.036).*

(A)



(B)

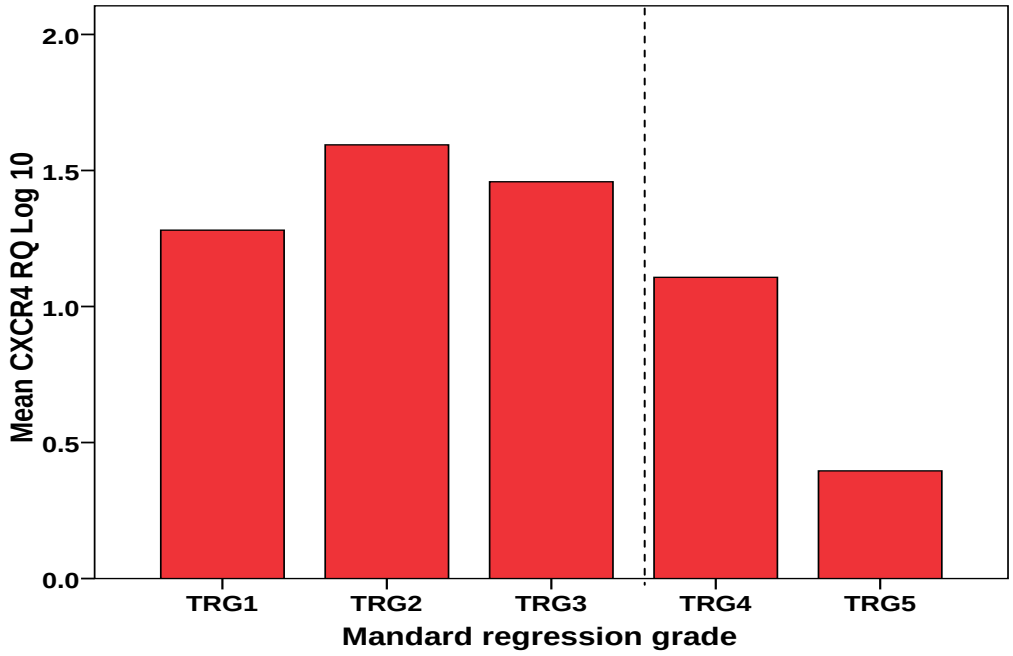
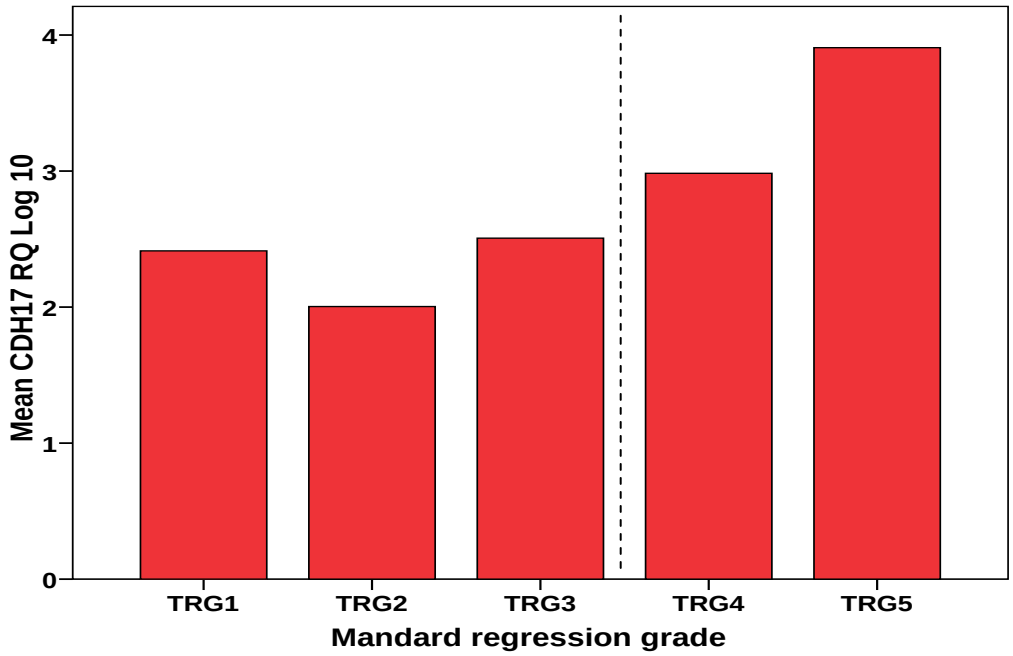
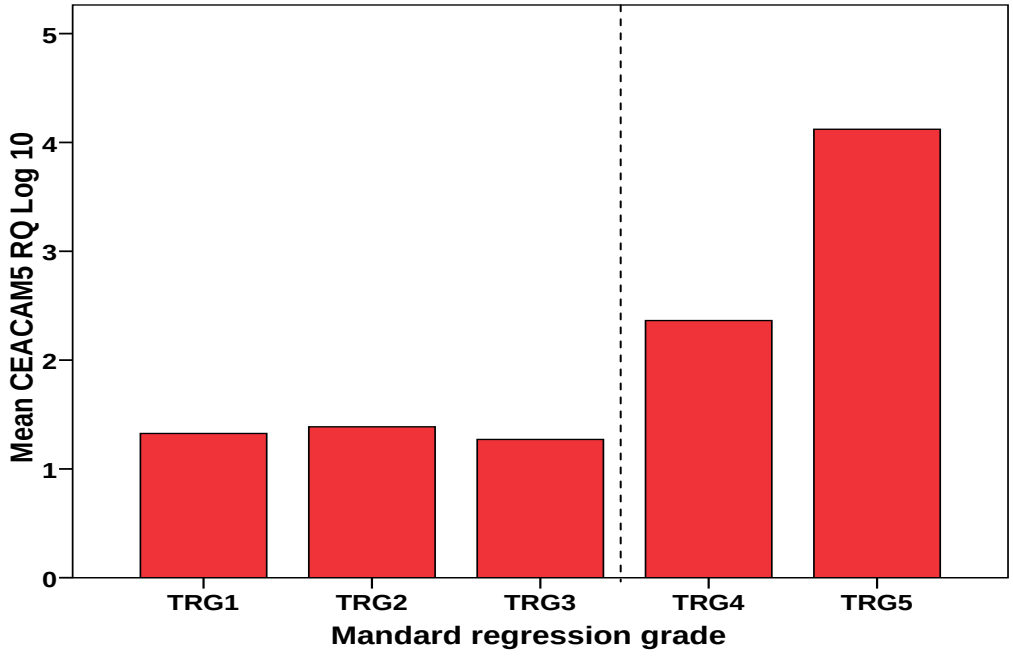


Figure 4.23: Continued

(C)



(D)



4.5 Discussion:

The clinical and pathological parameters of colorectal cancer are still the basis of treatment options, classification and prognostic stratification; however, they may be inadequate in everyday practice due to the great biologic and genetic heterogeneity of the disease. Furthermore, the molecular factors involved in prognosis and response to therapy in CRC is poorly understood

Detection of disseminated tumour cells in peripheral blood of colorectal cancer patients has been achieved primarily using immunocytological or flow cytometry based techniques [432, 433]. The clinical usefulness and high sensitivity of PCR in detecting cancer markers in circulating tumour cells was confirmed before [434, 435]. In colorectal cancer, RQ-PCR was used to determine the expression levels of CEA, CK20 and CK19 mRNA in peripheral blood and indicated a valuable tool for cancer staging and disease monitoring [387, 436, 437]. The undetermined expression of mRNA in peripheral blood in this study may be explained by the low concentration of these molecules in the blood. This problem could be overcome by optimisation of the extraction methods and applying RNA concentration techniques.

In this study we characterised the expression of a group of genes in colorectal cancer. Although their expression levels were undetermined in blood, we identified a comprehensive list of genes with highly differential expression patterns in colorectal cancer tissues that could serve as molecular markers to complement existing histopathological factors in diagnosis, follow up and therapeutic strategies for individualised care of patients.

CDH17 is a member of the cadherin superfamily of genes which encode calcium-dependent membrane-associated glycoproteins that mediate cell-cell adhesion in the intestinal epithelium. The protein is a component of the gastrointestinal tract and pancreatic ducts where it functions as an intestinal proton-dependent peptide transporter. The mechanism of the adhesive function of *CDH17* is unclear but it could be complementary to the classical cadherins like E-cadherins [438]. Cell to cell adhesion by *CDH17* is apparently independent of any interaction with cytoskeleton component because its cytoplasmic dominant is very short [439].

High expression levels of *CDH17* were noted in hepatocellular carcinoma and gastric cancer and found to be associated with intestinal type gastric carcinoma, poor survival, tumour invasion and lymph nodes metastasis, while in pancreatic and colorectal cancer the reverse is true [370, 372-374, 440-443]. Those observations in diverse types of tumours highlight tumour-specific expression patterns and presumably reflect tissue specific regulatory mechanism for *CDH17*.

In colorectal cancer, *CDH17* expression was only investigated at protein level using immunohistochemistry and immunoblotting. Hinoi et al. examined the protein expression in human colorectal cancer cell lines. In their study, *CDH17* was not detected in cell lines showing dedifferentiated phenotypes [444]. This was further confirmed by Takamura et al. who examined the *CDH17* expression in four cell lines and 45 human primary colorectal carcinoma using monoclonal antibodies. In cell lines the protein was expressed in differentiated but not the dedifferentiated phenptypes while in tissues reduced *CDH17* expression was associated with high tumour grade, advanced stage and lymphatic invasion and metastasis [373]. Moreover, Kwak et al. found reduced expression in 51% of the 207 colorectal cancers he studied using immunohistochemistry and he significantly correlated down-expression of *CDH17* with poor survival and lymph nodes metastasis [374].

To our knowledge, this is the first study to investigate *CDH17* mRNA in colorectal cancer using RQ-PCR. Our findings support the above reports and confirm that down-regulation of *CDH17* in colorectal cancer is associated with poor differentiation, raised CA19.9 tumour marker serum level and local tumour invasion indicated by increase bowel wall involvement. Interestingly, *CDH17* expression correlated with increased tumour diameter and tumour thickness (indices of intraluminal tumour growth) and decreased with increased bowel wall involvement (index of local tumour invasion). Those findings could be explained by the adhesion function of the protein. Generally, for the tumour to grow in diameter and thickness it needs to retain adhesion molecules expression, while loss or inactivation of those adhesion molecules correlate with inhibition of cell aggregation and promotion of tumour invasiveness. Therefore, *CDH17* expression

levels could be used as marker to guide total mesenteric excision (TME) planning in rectal cancer patient.

Carcinoembryonic antigen (*CEA, CEACAM5*), a member of immunoglobulin superfamily, has been assigned numerous physiological functions including immunological defence, cell adhesion and cell signalling. Its expression started during early fetal life and continues thereafter at lower levels mainly in the epithelial cells of the gastrointestinal tract, cervix, sweat glands and pancreas [445]. Overexpression of *CEACAM5* antigen was identified in majority of carcinomas involving the gastrointestinal, respiratory and genitourinary tracts and in breast cancer [445, 446]. Moreover, it has been reported to promote the metastatic potential in some experimental tumours [447]. The antigenic characteristics of *CEACAM5*, in addition to its role in tumour biology and metastasis, make it a favourable target for immunotherapy [448-451].

We identified non-significant overexpression of *CEACAM5* in tumour compared to tumour-associated normal colorectal tissues. Overexpression of *CEACAM5* was significantly associated with moderately differentiated tumours and tumour invasion. In relation to CEA protein, although the expression of the gene was high in patients with raised CEA protein serum level, we failed to identify any correlation between the tumour *CEACAM5* expression and the CEA serum level. Previous reports have shown that increase CEA protein level does not involve gene amplification or rearrangement but may be due to hypomethylation of upstream regions and factor changes leading to altered rate transcription [452, 453].

Chemokines are low molecular weight proteins that share a high degree of structural homology and the ability to attract specific cell types like leukocytes and tumour cells. They exert their biological effect by coupling to G protein-linked transmembrane receptors called chemokine receptor. Binding of chemokines to their receptors trigger activation of many signalling pathways including activation of calcium fluxes and protein kinases [454]. Dysregulated chemokine expression has fundamental roles in tumour initiation and progression and recent studies have shown that chemokines and chemokine receptors

contribute to cancer biology in different ways. They are able to modulate the local inflammatory reaction harbouring pro- or anti-tumorigenic activity and may promote angiogenesis and tumour cell proliferation, migration and survival [389, 455, 456].

CXCL12 (SDF-1), binds and signals through the chemokine receptors *CXCR4* and *CXCR7*, regulates many essential biological processes including angiogenesis, apoptosis, cell motility, migration and adhesion and cardiac and neural development [457, 458]. In mice genetically deleted of *CXCL12* early stage embryos exhibit profound defects in the vascular and brain development, hematopoisis and cardiogenesis which lead ultimately to embryonic lethality [459, 460]. The recent evidence that *CXCL12* also binds *CXCR7* receptor raised many questions on the potential contribution of the *CXCL12/CXCR7* axis to these processes that were previously attributed solely to *CXCL12/CXCR4* interaction. While *CXCR4* mRNA and protein were reported to be expressed on several cells as immune cells, epithelial cells and various types of cancer cells, *CXCR7* protein rarely expressed on the surface of normal nontransformed adult tissues [395, 461-463]. Interestingly, nontransformed tissues that lake surface *CXCR7* expression expressed *CXCR7* mRNA, suggesting that *CXCR7* could be regulated in a posttranslational manner [395]. Furthermore, in contrast to *CXCR4* which only binds to *CXCL12, CXCR7* is able to interact with two chemokines, *CXCL11* (*I-TAC*) and *CXCL12 [395, 464]*. Hence, *CXCL12* mediated response could be potentially modulated by *I-TAC*. These facts must always be considered when the biological effect of *CXCL12* via these receptors is evaluated.

A considerable number of previous reports have demonstrated that overexpression of *CXCR4* on tumours cells is associated with increased tumour growth and metastasis [465, 466]. *CXCR4* signalling was found to play a crucial role in metastasis homing of breast [467], ovarian [457] and prostatic [468] cancer cells by inducing chemotactic and invasive responses. Muller et al reported that primary breast cancer cells expressed *CXCR4*, whereas *CXCL12* was found in elevated levels in the metastatic sites of breast cancer like bone marrow, lung, lymph nodes and liver. Neutralization of CXCL12/CXCR4 interactions lead to inhibition of breast cancer lymphatic and lung metastasis [467]. In addition,

marked expression of *CXCR7* was determined in variety of tumour cell lines and in primary human tumours with correlation with tumour aggressiveness, angiogenesis, metastasis and promotion of tumour growth [395, 461-463].

Intestinal epithelium produces chemokines to regulate the trafficking of leukocytes into and out of the lamina properia [469]. Both *CXCL12* and *CXCR4* are normally expressed in these cells, however, in colorectal cancer cells *CXCR4* is over expressed while *CXCL12* seems to be partially or irregularly expressed as shown by Immunohistochemistry and RQ-PCR [392, 470]. In colorectal cancer, the overexpression of *CXCR4* was significantly associated with advanced tumours and metastasis [393]. Furthermore, *CXCR4* was found to induce stimulation of colon growth, VEGF release and ICAM-1 upregulation [471]. Similarly, *CXCL12* was reported to increase VEGF expression and to induce cell proliferation, metastasis and migration in colorectal cancer cells [392]. On the other hand, *CXCR7* in colorectal cancer was only investigated in colorectal cancer cell lines and animal colon cancer models and was determined to regulate angiogenesis and induce proliferation and chemotactic of cancer cells [377, 379]. Although the role of *CXCL12* and its receptors in colorectal cancer was investigated before, there is only sparse information on their function in carcinogenesis *in vivo.* Moreover, the data provided was controversy and generated on small sample size. In this study we used RQ-PCR, a gold standard method for gene analysis, to determine the clinicopathological correlation of *CXCL12, CXCR4* and CXCR7 mRNA in 107 tumour and tumour associated normal tissues, which is the largest sample size to date.

We demonstrated a significant down-regulation of *CXCL12* in tumour compared to tumour-associated normal colorectal tissues in contrast to *CXCR4*, which showed non-significant up-regulated expression levels in tumour tissues. The reduced expression of *CXCL12* was noticed in both polyps and cancer. This could be explained by the role of *CXCL12* in tumour immunology; however, it may highlight a possible tumour suppressor function of this gene. Investigation of *CXCL12, CXCR4* and *CXCR7* relationship may provides some help in understanding their functions and the role of each gene in regulating the expression of the others. Despite the reciprocal pattern of expression we identified

a strong positive correlation of *CXCL12/CXCR4* and *CXCL12/CXCR7* in both tumour and tumour-associated normal colorectal tissues. Moreover, *CXCR4* and *CXCR7* showed the same manner of correlation. Saigusa et al (2010) also reported a significant positive correlation between expression levels of *CXCL12* and *CXCR4* in 53 patients with rectal cancer underwent preoperative chemoradiotherapy. Moreover, the expression of *CXCR7* in *CXCR4* positive cells appears to enhance the responsiveness to *CXCL12* as reported by Sierro[472]. These findings suggest a possible receptors interaction in tumour and normal colorectal tissues.

Regarding clinicopathological correlations, in this study the expression levels of *CXCL12* and *CXCR7* were noted to be low in proximal colon. This may indicate a possible role of this axis in microsatellite instability (MSI) as tumours associated with MSI arise mainly in the proximal colon. Down-regulation of *CXCL12* and its receptors was also found to be associated with increase tumour size, local invasion, poor differentiation, advanced nodal stage, advance tumour stage and lymphovascular invasion. Of further interest, we identified for the first time the prognostic significant of *CXCR7* mRNA in colorectal cancer. We found that people with high expression of CXCR7 in their tumour cells live longer than their counterparts with low *CXCR7* gene expression. This was further confirmed by multivariate analysis.

In summary, in our cohort of CRC patients we found a reciprocal pattern of *CXC12* and *CXCR4* expression with increase expression of *CXCR4* is associated with decrease expression of its ligand. Further more we demonstrated significant correlation of expression of *CXCL12, CXCR4* and *CXCR7* in both tumour and normal colorectal tissues. We also addressed for the first time the significant association of clinicopathological variables like tumour size, location, grade, invasion and lymph node status and the expression of target genes. To our knowledge, this is the first study to report the prognostic significant of *CXCR7* expression in cancer patients.

Liver fatty-acid binding protein (*FABP1*) is specific marker of hepatocytes and enterocytes. In gastrointestinal tract it exists only in the epithelial absorptive cells

of small intestine and the colon, but not in oesophagus and stomach [473]. The expression of *FABP1* was investigated before in numerous cancers. High expression was identified in hepatocellular carcinoma, lung cancer and gastric cancer [474-476].

Evidence of dysregulated *FABP1* gene expression has been reported in colorectal gene expression array datasets [365, 477], however, little is known of its expression profile with regard to clinical data. Lawrie et al. used the proteomic and immunohistochemistry to determine the changes of *FABP1* in 20 colorectal tumours. They identified consistent loss of *FABP1* in tumour compared to normal colon. They also noted the association of decreased protein expression and poorly differentiated tumours and large adenomas [344]. Moreover, *FABP1* expression was significantly associated with good prognosis after liver resection of colorectal cancer metastasis in the study by Yamazaki et al. who investigated 68 liver metastasis and 10 primary colorectal cancers using immunohistochemistry [478].

Although no statistically significant correlation between *FABP1* expression and clinicopathological parameters was identified in this study, we observed that *FABP1* is differentially expressed in normal-adenoma-carcinoma sequence and its loss occurred early in colorectal cancer tumourogenesis. This indicates tumour suppressor function of *FABP1* in colorectal cancer. The loss of *FABP1* in colorectal cancer contrast with the findings in other tumours types which might be explained by the organ-specific distribution and the different role of *FABP1* through distinct intracellular interacting molecules.

In addition to their role in inflammation *IL8* may promotes tumour progression, invasion and metastasis by stimulating neoangiogenesis and activation of matrix proteases [479, 480]. *IL8* was shown to directly modulate endothelial cells proliferation and migration, hence promoting angiogenesis [481, 482]. It also exerts its effect on endothelial cells indirectly via increase secretion of vascular growth factors such as vascular endothelial growth factor and basic fibroblast growth factor [407]. Furthermore, inflammatory cells recruited by *IL8* to the cancer site may contribute to tumour progression through release of growth and angiogenic factors and promote invasion and distal metastasis [483]. Although

high expression of IL8 in colorectal cancer was noted before [484, 485], no *in vivo* correlations with clinicopatological variables were identified.

In keeping with the previous reports, we noted overexpression of *IL8* in tumour compared to normal colorectal tissue. In addition, we identified a progressive manner of increase gene expression from normal, to polyps, to tumour. The early dysregulation of *IL8* in colorectal cancer suggest that the gene may play a role in carcinogenesis in addition to its confirmed role in tumour progression.
Correlations with clinicopathological parameters revealed significant association of reduced *IL8* expression and poor tumour differentiation, advanced nodal stage and disease recurrence. Although the significant of these findings is unclear, it should be considered when planning *IL8* targeting therapy.

Mucus is viscoelastic secretion that is secreted by specialised epithelial cells called goblet cells which are abundant in the epithelial surface of tubular organs and the secretary surfaces of organs like liver, pancreas, kidney and salivary glands [486]. Mucus secretions serve as a protective barrier against harmful substances and act as a lubricant between lumen and cell surface. Mucins are the most abundant components in mucus and are responsible for its properties and functions. In addition to their physical protection, mucins play an important role in diverse biological roles such as differentiation, adhesion, immune response and cell signalling.
The human mucin family consist of at least 21 members designated *MUC1* to *MUC21* and have been classified into secreted gel-forming and membrane-bound (transmembrane) forms [487, 488].

The intestinal mucosa is covered by mucus, partly consist of secreted mucins, which provide a physical barrier and limit damage to the epithelium by luminal contents including bile, enzymes, ingested toxins and normal flora. The mucus consists of a less dense outer layer and a highly enzyme resistant packed inner layer. The inner layer is comprised of uncleaved *MUC2* and is free of bacterial colonisation. Therefore, loss of *MUC2* expression could contributes to numerous colonic pathologies including, but not limited to, ulcerative colitis and carcinoma [488, 489]. However, the role of MUC2 as a tumour suppressor may seems

confusing as *MUC2* was reported to be expressed in increased levels in certain malignancies including gastrointestinal tract [490-492]. This might reflect the origin of these tumours from cells normally expressed *MUC2*, rather than a role in carcinogenesis.

Inactivation of *MUC2* in mice caused tumour formation in small intestine and colon. This was accompanied by increase proliferation, decreased apoptosis and increased migration of epithelial cells. It is unclear if these changes are a primary response to loss of MUC2 or secondary to the inadequate protection of the intestinal epithelium [411]. In human, loss of *MUC2* expression was identified in non-mucinous colorectal cancer and showed to correlate with tumour progression [487-489, 493]. Moreover, down regulation of MUC2 was noted to be associated with progression along the adenoma-carcinoma sequence pathway [412, 494]. On the other hand, overexpression of MUC2 was noticed in mucinous type colorectal cancer and found to be associated with poor prognosis and depth of invasion [495]. This might be due to the barrier formed by mucins secreted by tumour cells, which protect against recognition by anti-tumour immune effectors. In this study, keeping with previous reports, we confirmed *MUC2* mRNA down-regulation in non-mucinous and over-regulation in mucinous colorectal cancer. We also showed decreased expression of *MUC2* in a progressive manner from tumour-associated normal, to polyps, to tumours. No significant association of *MUC2* and clinicopathological variables other than CA19.9 serum levels has been determined in this study.

Programmed cell death 4 (*PDCD4*) is a tumour suppressor gene, its overexpression was found to inhibit chemicals induced neoplastic transformation *in vitro* [496, 497]. It was also shown to suppress tumour promotion and progression in animal models [498]. In the JB6 mouse epidermal clonal genetic variant cell system *PDCD4* was found to be highly expressed in JB6 transformation-resistant but not in transformation-susceptible cells. Moreover, reduction of *PDCD4* expression in transforming-resistant cells was accompanied by acquisition of a transforming-susceptible phenotype, while its overexpression in transforming-susceptible cells render them resistant to tetradecanoyl phorbol

acetate-induced transformation and inhibit the expression of tumour phenotype [497].

PDCD4 has been identified as a suppressor of transformation, tumourogenesis and progression [496, 497, 499]. It also inhibit apoptosis and tumour growth, invasion and intravasation [416, 418, 500, 501]. The mechanisms by which *PDCD4* exerts this list of functions is unclear, however, interaction with other molecules could be involved. It has been demonstrated that *PDCD4* regulates molecules like *p21* [502] , *Cdk4* , carbonic anhydrase-II [503], *elF4A/elF4G* [504, 505] and urokinase receptor [506], while other molecules like protein kinase B [507], *Myb* [508], *Cox-2* inhibitors [501] and miR-21 [509, 510] regulate *PDCD4* expression.

Although the role of *PDCD4* in suppressing different phenomena associated with cancer has been extensively explored, few studies have investigated the potential use of PDCD4 as a prognostic or diagnostic biomarker. Furthermore, the role of *PDCD4* in tumour progression has mainly been based on studies that used cell lines. Mudduluru et al carried out the only prognostic study of *PDCD4* in colorectal cancer [511]. They analysed *PDCD4* expression in 71 colorectal cancer patients and 42 adenomas using immunohistochemistry and western blot and noticed a significant reduction in *PDCD4* expression comparing tumour and polyps to tumour associated tissues. They also identified loss of *PDCD4* expression as an independent predictor of disease free survival.

To our knowledge, we carried out the first study to characterised *PDCD4* expression in colorectal cancer tissues using RQ-PCR. *PDCD4* mRNA was significantly lower in tumour and polyp compared to tumour-associated tissue in keeping with the protein expression levels described before [477, 502, 511]. Furthermore, we identified the novel association of reduced *PDCD4* expression with disease recurrence and raised CA19.9 serum level. These findings suggest that *PDCD4* involves in both tumour promotion and tumour progression and represent a potential biomarker for evaluating the transition of normal colorectal tissue to adenoma and carcinoma. Reduced expression of PDCD4 in proximal

compared to distal colon may indicate a potential role in microsatellite instability (MSI) and Lynch syndrome.

TGFB1 serves as tumour suppressor in the normal intestinal epithelial cells as they move out of intestinal crypts to the tips of villous by inhibiting proliferation and inducing apoptosis [429, 512]. However, during the late stage of carcinogenesis it acts as a tumour promoter and is usually highly expressed [422, 426, 513]. Experimentally, prolong exposure of intestinal epithelium to *TGFB1* promotes neoplastic transformation and it stimulates proliferation and invasion of poorly differentiated and metastatic colon cancer cells [428, 514, 515]. The molecular changes that result in redirection of *TGFB1* growth inhibition signals during tumourogenesis are essentially unknown. A subset of colorectal cancer has been shown to have mutations or down regulation of type I and type II receptors [516, 517] *SMAD2 [518]* and *SMAD4 [519]*, hence increase the production of several mitogenic growth factors including *TGFα, FGF* and *EGF [520]*. The role of *TGFB1* signalling pathway is best illustrated by presence of inactivating mutation in genes encoding *TGFB* receptors and *SMAD*s in human cancer and by studies of tumour development in mouse models.

Silencing of *TGFB* receptors has been observed to promote establishment and progression of cancer [516]. Type I receptor inactivating mutations were described in ovarian and pancreatic cancers, T-cells lymphoma and metastatic breast cancer [422], while type II receptor gene inactivation have been identified in colon, head and neck cancers [425]. Transgenic mice that lack a copy of *tgfbr2* have an increase susceptibility to develop cancer and restoration of functional receptors reverse the malignant behaviour of several human cancer cell lines that lack functional *TGFB* receptors. These observations suggest that *TGFB* receptors might work as tumour suppressor in development of cancer.

Many previous studies have examined the relations between *TGFB* pathway and the disease progression in colorectal cancer. Nevertheless, this is the first study to explore the relation of *TGFB1* and its receptors mRNA in colorectal cancer using RT-PCR. Moreover, the large cohort of patients in this study gives it further advantage compared to the other studies.

Although no significant differences were identified in *TGFB* receptors expression in colorectal tumours compared to normal, *TGFB1* expression levels were significantly lower in polyp and higher in cancer compared to tumour-associated normal tissues. This is in keeping with previous reports. Matsushita et al (1999) found that *TGFB* receptors mRNA expressed mainly by normal and adenoma colorectal tissues whereas *TGFB1* expressed by cancer [516]. Moreover, Daniel et al (2007), using immunohistochemistry, identified higher *TGFB1* protein expression in colorectal cancer, than in high-grade dysplastic polyp, than in low-grade dysplastic polyp [521]. The significant positive correlation between *TGFB1* and its receptors expression levels in both tumour and normal colorectal tissues confirm that their role in colorectal cancer is more complex than a simple legend-receptor feedback.

Interestingly, we identified for the first time the relationship of *TGFB* pathway and some established prognostic clinicopathological parameters. Low expression of *TGFBR1* was found to be associated with raised CEA serum level and local tumour invasion. In addition, *TGFBR2* down-regulation was associated with local, perineural and lymphovascular invasion and advanced nodal stage. These findings will further confirm the role of *TGFB* receptors as tumour suppressor. The down-regulation of *TGFBR2* in proximal compared to distal tumours was described before and highlights the role of this gene in microsatellite instable tumours.

Tumours of proximal and distal parts of the colon may form different but related groups of tumours because of their different embryological origin, different exposure to bowel contents and differences in clinical presentation, progression and possible genetic and environmental epidemiology [522]. The concept that proximal and distal colon and rectum represent different entities is supported by evidence that two different genetic mechanisms, microsatellite instability (MSI) and chromosomal instability (CIN), contributes unevenly to the carcinogenesis in the different parts of gastrointestinal tract [523]. The incidence of CIN is similar in the distal colon and the rectum and associated with tumours located distal to the splenic flexure [524, 525]. In sporadic colorectal cancer, MSI results from inactivation of DNA-mismatch repair (MMR) genes and secondary mutations of genes with coding microsatellite. Reports indicate that MSI tumours are located

proximally to splenic flexure [524] and also suggested that if MSI is present in rectal cancer, this would strongly suggest a hereditary predisposition for the cancer as MSI is rare in the lower part of the colon and rectum.

Using oligonucleotide microarrays, Birkenkamp-Demtroder and his colleagues investigated the differences in gene expression in colon cancer of the caecum versus the sigmoid and rectosigmoid [185]. They identified 58 genes to be differentially expressed between the normal mucosa and 16 genes differentially expressed comparing Dukes' B and C carcinoma of the caecum with those of the sigmoid or rectosigmoid. Moreover, Kapiteijin et al. reported increased beta-catenin and P53 in rectal cancer compared to proximal [526], whereas Fric et al. noted increased expression of *EGFR, c-erb B2, PCNA* and *DPP IV* in right sporiadic compared to left colon cancer [527].

In this study we found decreased expression levels of *CXCL12, CXCR7, PDCD4* and *TGFBR2* in proximal compared to distal tumours. Furthermore, we determined decreased expression of *CXCL12* and *TGFBR2* in colon compared to rectal cancer. These findings may indicate to the role of these genes in tumourogenesis of microsatellite instable tumours. In addition, *CXCL12* and *TGFBR2* may serve as markers to differentiate colon and rectal cancer.

The aim of neoadjuvant therapy is to sterilize the resection margins and destroy the microscopic collections of cancer cells, a potential focus of recurrence and source of distal metastasis. The increasing use of neoadjuvant chemoradiation therapy and improvement in the quality of rectal surgery, have led to reduce local recurrence and improved survival in the management of rectal cancer. Nevertheless, these benefits only noticed in limited group of patients (10-30%). In addition, such therapy is expensive, associated with increase risk of second cancer in adjacent organs and increases the risk of postoperative mortality and morbidity. Therefore, accurate selection of patients who are suitable candidates for neoadjuvant therapy will significantly improve the outcomes. Many molecular markers, including *P53, p21, BCL2, BAX, EGFR, COX2, PTMA* and *ELF5a1,* were identified before as predictors of response to this modality of treatment; however, their clinical application is still under evaluation [67, 528]. Although no

pre-treatment biopsies were used, the list of genes identified to be dysregulated in response to neoadjuvant therapy in this study is consistent with previous reports. Ambrosini-Spaltro et al. investigated 32 pre-treatment biopsies by immunohistochemistry and determined *MUC2* as a predictor of poor response [529]. Moreover, stromal *CXCL12* and *CXCR4* expression was found to be associated with recurrence and poor survival after neoadjuvant therapy in the 53 patients analysed using RQ-PCR and immunohistochemistry [528]. However; comparison of expression levels of these genes in pre- and post-treatment biopsies is required to further validate their use as predictors of response.

Measurement and quantifying of tumour response to neoadjuvant chemoradiation therapy is an important parameter in order to elucidate factors that may allow for response prediction and planning of next step of treatment in rectal cancer patients. Clinical response (cCR), pathological response (pCR) and tumour downstaging are the commonly used methods to measure response. Both clinical response and tumour downstaging compared the tumour characteristics before and after treatment clinically and using radiological tools like magnetic resonance imaging (MRI) and trans-rectal ultrasound (TRUS). Whereas pathological response (regression grade) stratifies response base on biological effect of radiation on tumours. Mandard tumour regression grade, originally described for oesophageal cancer, is the most commonly used [44]. It consists of five different grades based on ratio of fibrosis to tumours. We identified, for the first time, a group of genes that can be used as markers to quantify tumour response following neoadjuvant therapy in rectal cancer patients.

In summary, Expression levels of *CXCL12*, *CDH17*, *MUC2*, *L-FABP* and *PDCD4* were down-regulated and *IL8* was up-regulated in tumours compared to normal colorectal tissues. No significant differences were noted in expression of *CEACAM5, CXCR4, CXCR7, TGFB1, TGFBR1* and *TGFBR2*. We also identified significant associations of gene expression levels and clinicopathological variables such as tumour location, size, grade and invasion, overall survival and lymph node status. Moreover, we determined the significant relationship of *CDH17, CEACAM5, CXCL12* and *CXCR4* expression levels and tumour regression grade in response to neoadjuvant chemoradiation therapy.

Chapter 5:
Prediction of rectal cancer response to neoadjuvant chemoradiation therapy

5.1 Introduction

The increasing use of adjuvant and neoadjuvant therapy has led to improved outcomes in the management of colorectal cancer [530]. Post-operative adjuvant chemotherapy has been shown to improve the outcome in patients with Dukes' C tumours and is generally accepted as standard care [357] , however, only selected patients of Dukes' B group would benefit from this treatment. Moreover, neoadjuvant chemoradiation is becoming the standard of care in the treatment of locally advanced rectal cancer. It is associated with significant improvements in down staging of the disease which correlates with improved rates of sphincter sparing surgery, decreased regional recurrence, and improved overall survival as confirmed by the prospective randomized trial of the DCCG and the German Rectal Cancer Study Group [36, 358]. The response to neoadjuvant therapy is quantified by tumour regression grade which was originally described for tumours of the oesophagus [44]. Tumour regression grading (TRG) is pathological grading system based on the histological degree of tumour regression and fibrosis present in the specimen after preoperative treatment [531]. It has proven to be of prognostic significant when assessed in multi-centric preoperative therapy trials [41]. Preoperative treatment revealed increase risk of second cancer and associated with considerable perioperative morbidity. Moreover; there are emerging opinions for non-surgical management of patients with complete response to neoadjuvant therapy [532-534]. These facts make the ability to predict response to neoadjuvant therapy of great importance in clinical settings.

Recently, post-transcriptional and translational controls of protein coding genes regulated by miRNA have emerged as an interesting field of cancer research. Control mediated by miRNA provides the cell with a more precise, energy-efficient way of controlling the expression of proteins and greater flexibility in responding to numerous cytotoxic stresses. The exact function of miRNAs is just emerging, however, their ability to regulate cell proliferation and cell death has been previously shown [535]. Due to their small size, miRNA are more stable and resistant to environmental, physical and chemical stresses compared to mRNAs. Therefore; their analysis in formalin-fixed paraffin embedded (FFPE) tissue samples is likely to provide more accurate replication of what would be observed in fresh tissues than that of mRNA species [167, 536-538].

Relationships between radiosensitivity and functions of several genes have been reported, including *P53, BCL-2, BAX and P27* [67, 539], however, little is known about the clinical significant of these genes or their usefulness for estimating radiotherapy effectiveness. The potential role of miRNAs as good biomarkers for cancer diagnosis and prognosis has been confirmed before [540-543]. However, no previous study has investigated their role as predictors of response to neoadjuvant therapy in colorectal cancer.

FFPE tissue offers a widely available and rich archive of well characterised tissue specimens and patient data for comparative molecular and clinical retrospective studies [544]. New extraction methods have made it possible to retrieve total RNA from preserved tissue specimens to a level that could be quantified by RQ-PCR. However, the application of these methods to FFPE tissue is limited by extensive RNA fragmentation and modifications. Until recently, FFPE samples have not considered reliable source of mRNA for gene profiling experiments due to difficulty in obtaining intact mRNA from these samples. Therefore, optimising of the extraction methods and RNA quality from FFPE tissues is of particular interest to many research groups.

While formalin fixation helps in preserving the cellular proteins and conserves the tissue architecture, it significantly affects the quality of RNA. The process of preserving tissues involves fixation of tissues in formalin for more than 24 hours followed by dehydration with graded ethanol solution to enhance paraffin embedding. This procedure considerably modifies the RNA structure and questions its quality for standard molecular analysis. The most important modifications involve:

- Extensive cross-linking of RNA with proteins during fixation
- Enzymes degradation
- Chemical degradation
- Formation of mono-methylol adducts with bases of nucleic acids, especially with adenine.

These modifications reduced the efficiency of reverse transcription and negatively affect the performance of RNA in downstream reactions. Many research groups have attempted with success to overcome these limitations. In addition, the

advances in high-throughput molecular genetic techniques like RQ-PCR and microarrays; there has been growing interest in using FFPE archive collections as a source of biological data

Artificial Neural Networks (ANNs) are a form of artificial intelligence that can learn to predict, through modelling, answers to particular questions in complex data. The original inspiration for the term *Artificial Neural Network* came from examination of central nervous systems and their neurons, axons, dendrites and synapses which constitute the processing elements of biological neural networks investigated by neuroscience. In an artificial neural network simple artificial nodes, called variously "neurons", "neurodes", "processing elements" (PEs) or "units", are connected together to form a network of nodes mimicking the biological neural networks — hence the term "artificial neural network".

ANN model development is achieved by a training process involving the adjustment of the weighted interconnections between nodes within the neural network over a defined number of epochs. This adjustment occurs by the iterative propagation of the predictive error back through the entire network with a learning algorithm. The models produced by ANNs have been shown to have the ability to predict well for unseen data and have the ability to cope with complexity and non-linearity within the dataset [545, 546]. Thus ANNs have the potential to identify and model patterns in this type of data to address a particular question. These patterns can combine into a fingerprint that can accurately predict subgroups. For this reason, they have been widely applied to a range of domains including character and face recognition [547], stock market predictions, or survival prognosis for trauma victims [548]. Consequently, ANNs have the ablility to determine patterns or features (*e.g.* in genes or proteins) within a dataset that can discriminate between subgroups of a clinical population (*e.g.* disease and control, or disease grades [549]. ANNs have already been successfully applied in a number of contexts where markers of biological relevance have been identified including polycystic ovarian syndrome [549], melanoma [549], prostate cancer [550, 551] and breast cancer [294, 552]. In colorectal cancer they were used to characterize the disease and to predict survival and outcomes of CRC patients [553-558]. Their application to the analysis of colorectal cancer microarray data

was reported by Selaru *et al* [556] who evaluated the ability of artificial neural networks (ANNs) based on complementary DNA (cDNA) microarray data to discriminate between sporadic colorectal adenomas and cancers (SAC) and inflammatory bowel disease (IBD)-associated dysplasias and cancers. Signatures were identified and validated with 100% accuracy. Notably, significantly fewer genes were included in the signatures compared to signatures generated by other analysis methodologies. This study highlighted the potential application of ANN to microarray analysis and illustrated how this method should be exploited to provide a further understanding of CRC biology.

5.2 Aims

The objectives of this study were to optimise miRNA extraction methods from FFPE tissue samples and to systematically investigate the miRNA expression profiles between FFPE samples and fresh-frozen samples using RQ-PCR. Also to characterise miRNA expression in tumour compared to tumour-associated normal (TAN) FFPE colorectal tissues. Moreover; we aimed to identify predictors of response to neoadjuvant chemoradiation therapy in colorectal cancer using FFPE tissues as source of genetic materials, and microarray analysis as investigation tool.

5.3 Materials and methods

5.3.1 Study group

A group of 9 patients was selected for optimization of RNA extraction methods from FFPE tissues and to evaluate RNA quality in relation to RNA extracted from fresh-frozen tissues. Each patient in this group has both FFPE and fresh frozen/ tumour and TAN tissues available. Then a group of 12 rectal cancer patients who had neoadjuvant therapy and had pre-treatment biopsies available was selected to examine the expression of miRNA by microarray analysis in order to identify predictors of response to neoadjuvant chemoradiation therapy. Response to neoadjuvant therapy was quantified using Mansard tumour regression grade.

Table 5.1: Characteristics of patients used in optimisation of RNA extraction.

Tissue sample	Tissue type	FFPE RNA Conc.	FF RNA Conc.	Tumour % in FFPE block	Location	Dukes Stage
T08-0732	Tumour	119.4	597.9	100%	Rectal	C
T08-0733	Normal	60.68	778.64			
T08-0655	Tumour	61.50	1360	100%	Rectal	C
T08-0656	Normal	78.99	840.32			
T08-0605	Tumour	63.9	1747.7	70%	Colon	B
T08-0607	Normal	43.20	704.78			
T08-0727	Tumour	55.8	1190.4	100%	Colon	C
T08-0728	Normal	45.79	253.86			
T08-0907	Tumour	61.4	1286.9	100%	Colon	B
T08-0908	Normal	21.90	17.03			
T08-0685	Tumour	66.4	1865.9	100%	Rectal	D
T08-0686	Normal	34.50	740.94			
T08-0713	Tumour	28.0	1049.5	100%	Rectal	B
T08-0714	Normal	21.02	1336.6			
T08-0594	Tumour	70.9	1570.8	100%	Colon	A
T08-0595	Normal	19.37	476.91			
T08-0700	Tumour	65.9	1834.5	100%	Rectal	C
T08-0701	Normal	41.48	753.93			

FFPE = Formalin-fixed paraffin embedded, FF = Fresh-frozen

Table 5.2: Microarray analysis cohort

No	Biopsy No	Response to CRT	TRG	RNA Conc. (ng/µL)
1	11628/05	No residual cancer	1	95.88
2	19664/06	Rare residual cancer cells	2	284.07
3	11275/07	No residual cancer	1	219.36
4	11992/07	No residual cancer	1	1094
5	16232/07	Rare residual cancer cells	2	201.96
6	13370/07	Fibrosis outgrowing residual cancer	3	335.58
7	18383/07	Rare residual cancer cells	2	1177
8	19808/07	Fibrosis outgrowing residual cancer	3	124.03
9	22376/07	Fibrosis outgrowing residual cancer	3	140.46
10	22075/07	Rare residual cancer cells	2	196.12
11	4201/08	Residual cancer outgrowing Fibrosis	4	103.66
12	10077/08	Rare residual cancer cells	2	92.6
13	4223/08	No residual cancer	1	89.72
14	5060/08	Residual cancer outgrowing Fibrosis	4	54.31
15	9556/08	Fibrosis outgrowing residual cancer	3	160.84
16	15736/08	Fibrosis outgrowing residual cancer	3	223.76
17	358/09	Absence of regressive changes	5	113.31

5.3.2 Formalin-fixed paraffin embedded (FFPE) tissues

A pair of tissues (tumour and TAN) was placed in 10% formalin (Lennox) for fixation and prior to paraffin embedding. The 10% formalin mixed was made of 4 gm of Sodium phosphate monobasic, 6.5 gm of Sodium phosphate dibasic, 100 mL of 37% formaldehyde and 900 mL of distilled water. Biopsies were fixed and stored at room temperature until embedding for a minimum of 24 hours. Tissue was then removed from the formalin and placed on an open cassette. The cassette was closed and placed in 250 mL of Industrial Methylated Spirit (VWR) to wash the formalin from the tissue. Then, the cassette was removed and placed in JFC solution (Milestone) filed JFC beaker and placed in the histoprocessor (MicroMED) for 60 minutes (70°C). Thereafter, the cassette was transferred to the paraffin wax (VWR) filled wax beaker and placed in the histoprocessor (MicroMED) for 30 minutes. The cassette was removed from the wax beaker and tissue was blocked out carefully. The blocks were left at 4°C until hard and then stored at fridge or room temperature until sectioning. Sectioning of formalin-fixed paraffin-embedded tissues was carried out using Slee microtome (LIS Ltd). The section thickness was adjusted to 5 µM for immunohistochemistry staining and 10 µM for molecular analysis and RNA extraction experiments. For immunohistochemistry, the sections were placed onto a Superfrost plus (positive charged) slides (VWR). While for molecular studies 3 of the 10 µM sections (after the first 2 sections been discarded) were placed into a 2 mL sterile tube and immediately preceded for RNA extraction process. Prior to enrolment in any further analysis each slide is stained in H & E and reviewed by a pathologist to determine the quality of the block and the percentage of tumour tissues in the section (should be >50%).

5.3.3 RNA extraction and analysis

For the initial evaluation of extraction techniques, RNA was isolated from colorectal tissues using three previously described methods for RNA extraction: Qiagen RNeasy FFPE method (Qiagen), Qiazol and chloroform protocol (Qiagen) and the TRI reagent RT-Blood protocol (Qiagen) according to the manufacturer's instruction. Thereafter Qiagen RNeasy FFPE method was employed in the proceeding experiments. RNA concentration and purity was assessed in duplicate samples (1 µL) using a NanoDrop ND-1000 Spectrophotometer (NanoDrop

technologies) while RNA integrity was evaluated using the RNA 6000 Nano Chip Kit (Series II) and the Agilent 2100 Bioanalyzer System (Agilent technologies).

5.3.4 Reverse transcription

Small RNA (5ng or 100ng) was reverse transcribed to cDNA using MultiScribe Reverse Transcriptase (Applied Biosystems). Each reaction was primed using a gene-specific stem-loop primer. Where sequences were available, primers were obtained from MWG Biotech. Otherwise, assays containing the RT stem-loop primer and the PCR primers and probes were used (Applied Biosystems). The 5 or 100ng of small RNA (in a final volume of 5 µL) was combined with the cDNA synthesis mix consisting of 0.17 µL dNTP (100mM), 1.65 µL of 10×RT buffer, 4.57 µL nuclease-free water,0.21 µL RNase inhibitor (20U/ µL), 3.1 µL stem-loop primer (50nM) and 1.1 µL MultiScribe RT (50U/µL). Samples were incubated at 16°C for 30 minutes, 42°C for 30 minutes and finally 85°C for 5 minutes to denature the strands. An RT-negative control was included in each batch of reactions.

5.3.5 Real-time quantitative PCR

The PCR reactions were carried out using a 7900 HT Fast Real-Time PCR System (Applied Biosystems). All reactions were performed in 20 µL reactions, in triplicate within the same PCR run. On each plate, an interassay control was included to account for any variations between runs. Reactions consist of 1.33 µL cDNA, 10 µL 2× TaqMan universal PCR master mix (Applied Biosystems), 1µL 0.2µM TaqMan probe (Applied Biosystems), 3 µL 1.5 µM forward primer, 1.4 µL 0.7µM reverse primer and 3.27 µL nuclease-free water. As before, The PCR reactions were initiated with a 10 minute incubation at 95ºC followed by 40 cycles of 95ºC for 15 seconds and 60º C for 60 seconds.

5.3.6 Relative quantification

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. QBasePlus was used for calculation of candidate expression relative to the endogenous control

miRNA (miR-26). It applies ΔΔ C_t method was used where ΔΔCt = (C_t target gene, test sample – C_t endogenous control, test sample) - (C_t target gene, calibrator sample - C_t endogenous control, calibrator sample). Relative quantities were corrected for efficiency of amplification and fold change in miRNA expression between groups was calculated as E-ΔΔCt ± s.e.m. The lowest expressed sample was used as a calibrator.

5.3.7 Microarray analysis

Microarray analysis was carried out on total RNA extracted from FFPE tissues using Megaplex pool A (Applied Biosystems) according to the manufacturer's protocol. It consists of matching primer pool and TaqMan arrays.

5.3.7.1 Megaplex RT reactions:

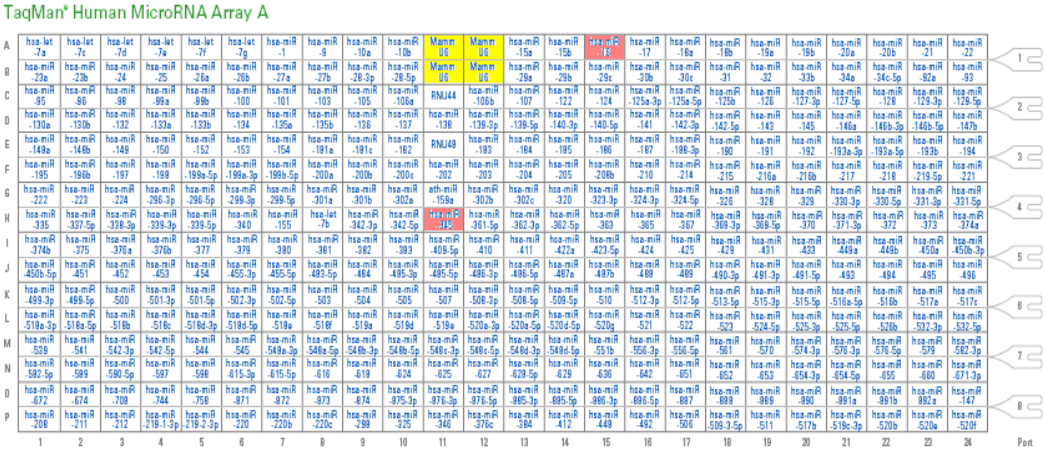
We used TaqMan MiRNA Reverse Transcription Kit and the Megaplex RT primers (Applied Biosystems) to synthesise single-stranded cDNA from total RNA samples. The reaction was performed in a total volume of 7.5µL of total RNA and RT reaction mix. RT reaction mix for one sample consists of 0.8µL of Megaplex RT primers (10X), 0.2µL of dNTPs with dTTP (100mM), 1.5µL of MultiScribe Reverse Transcriptase (50U/µL), 0.8µL of 10X RT buffer, 0.9 µL of $MgCl_2$ (25mM), 0.1µL of RNase inhibitor (20U/µL) and 0.2 of µL nuclease-free water (Applied Biosystems). Thereafter, samples were incubated for 40 thermal cycles at 16°C for 2 minutes, 42°C for 1 minute and 50°C for 1 second and finally left at 85°C for 5 minutes to denature the strands. The reaction was performed using a Gene Amp PCR system 9700 thermal cycler (Applied Biosystems).

5.3.7.2 TLDA RQ-PCR reactions:

The DNA polymerase from TaqMan Universal PCR Master Mix amplifies the target cDNA using sequence specific primers and probe on TaqMan MiRNA Array (Applied Biosystems). The reaction was carried out using a 7900 HT instrument (Applied Biosystems). The presence of the target is detected in real time through cleavage of the TaqMan probe by the polymerase 5' – 3' exonuclease activity (figure 2.4). For each array, a volume of 900 µL was prepared by combining 450µL of TaqMan Universal PCR Master Mix No

AmpErase UNG, 2X, 6 µL of Megaplex RT product and 444 µL of nuclease-free water. Of the PCR reaction mix, 100 µL was dispensed into each port of the TaqMan MiRNA Assay, centrifuged and sealed. Then, the array was loaded and run in using the 384 well TaqMan Low density Array default thermal cycling conditions of 50°C for 2 minutes and 95°C for 10 minutes then 40 cycles of 95°C for 15 seconds and 60°C for 1minute.

Figure 5.1: TaqMan Human MiRNA Array A
Eacch card contains 377 target miRNA and 4 endogenous controls (yellow). For the purpose of this study *miR-16* and *miR-345* (red) were used as endogenous control.



5.3.7.3 Data processing

Artificial neural network

Data was analysed using an Artificial Neural Network algorithm. A three layer Multi-Layer Perceptron (MLP) modified with a feed forward back-propagation algorithm and a sigmidal transfer function [559] was employed for development of the model using randomLy selected training and testing data sets. An additive stepwise approach [546] was employed to identify an optimal set of markers explaining variation in the population of each of questions explored.

5.3.8 Statistical analysis

Statistical analysis was carried out with IBM SPSS Statistics 17.0 (SPSS Inc.). Data was tested for normal distribution. Parametric tests were used where appropriate. One-way ANOVA and independent t-test were used to determine

association and comparisons between independent groups. Correlation analysis used Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively. Univariate analysis and paired-T test were used to assess related samples. P values <0.05 were considered statistically significant.

5.4 Results

5.4.1 RNA extraction from FFPE tissues

Extraction Methods

As initial step towards identifying the most optimal technique for RNA isolation fro FFPE tissues we isolated total RNA in duplicate from 4 tissue samples using three different extraction protocols: Qiagen RNeasy FFPE method (Qiagen), Qiazol and chloroform protocol (Qiagen) and the TRI reagent RT-Blood protocol (Qiagen). The comparison was based mainly on yield and purity of RNA extracted. Qiagen RNeasy FFPE method (Qiagen) demonstrated the best performance (table 6.1). Thereafter, to evaluate the quality and efficiency of miRNA extracted from FFPE tissues we integrated the expression levels of let7a, *miR-10b, miR-21* and *miR-16* in the same 4 samples using TaqMan primers and probes and equal input of total RNA for each sample. The expression levels of all miRNAs examined were undetermined in all samples when Qiazol and chloroform and the TRI reagent RT-Blood methods were used for RNA extraction. When Qiagen RNeasy method was used all miRNAs were detected at average Ct values of 26.9, 29.6, 24.1 and 23.7 for *let7a, miR-10b, miR-21* and *miR-16*, respectively.

Table 5.3: Yield and purity of RNA extracted using three different methods. As measured by NanoDrop spectrophotometery (NanoDrop technologies). RNA with an absorbance ratio at 260 and 280 nm (A_{260}/A_{280}) between 1.8 and 2.2 was deemed indicative of pure RNA

Sample	Qiagen RNeasy		Qiazol & chloroform		TRI reagent	
	Conc. (ng/µL)	A_{260}/A_{280}	Conc. (ng/µL)	A_{260}/A_{280}	Conc. (ng/µL)	A_{260}/A_{280}
1	32.04	2.05	13.51	1.49	3.83	1.72
2	66.39	1.82	5.82	1.40	2.11	2.61
3	65.90	1.89	11.12	1.49	1.72	20.5
4	85.92	1.88	11.26	1.28	3.61	2.77

Number of Slices:

To determine the number of slices required for optimal analysis of miRNA expression in FFPE we analysed the RNA yield and 2 miRNAs expression levels in two colorectal samples using 1, 2, 3 or 4 × 10 micron slices and under the same experimental conditions. The RNA showed stepwise increase in concentration depending on the number of slides, however, this increase was not statistically significant ($p>0.05$) (figure 6.1). Moreover, we examined the expression levels of *miR-145* and *miR-143* in two colorectal tissues using the same number of slices (1, 2, 3 or 4 × 10 micron). No significant difference in any of the miRNA expression level was noted based on the number of slices ($p>0.05$) (figure 6.2).

Figure 5.2: RNA yield in colorectal samples
RNA yield in two colorectal samples (C & D) using 1, 2, 3 or 4 × 10 micron slices. No significant difference in RNA yield was noted depending on the number of slices ($p>0.05$).

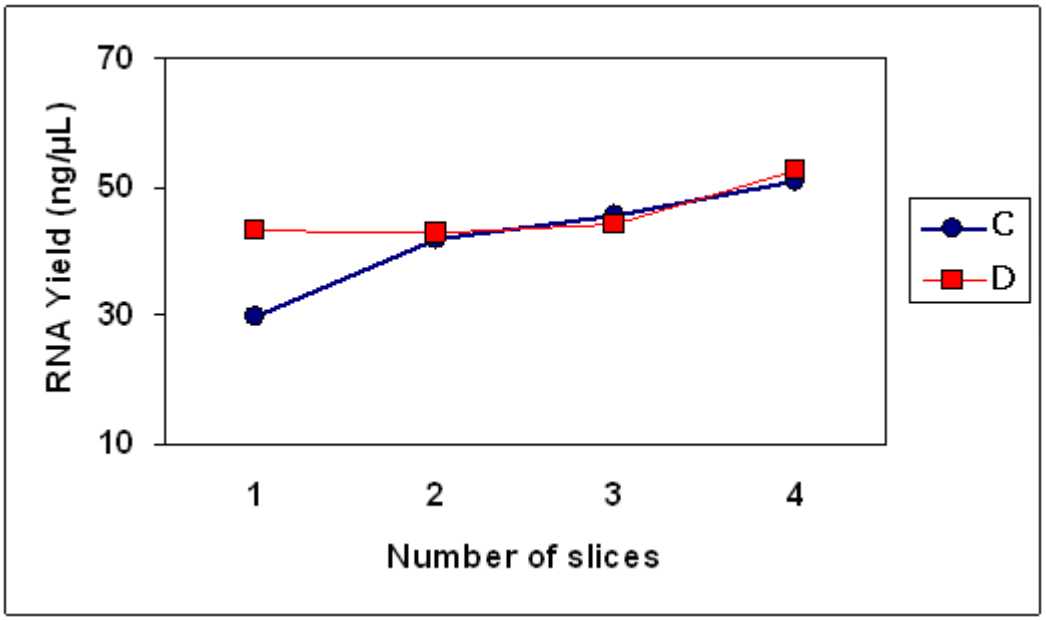
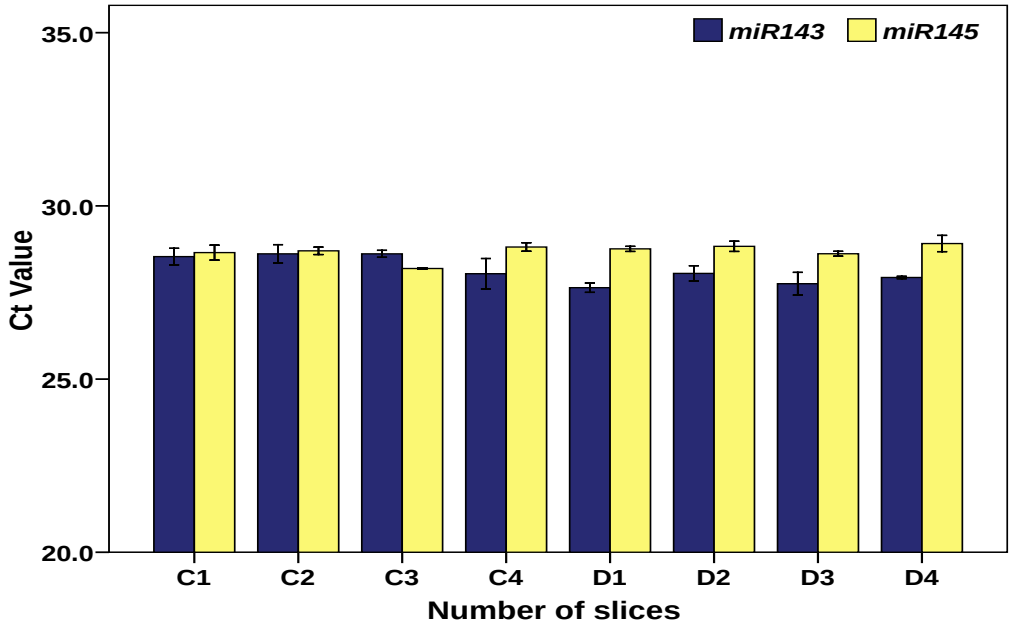


Figure 5.3: Expression levels of miRNAs in FFPE CRT
Expression levels of *miR145* and *miR-143* in two colorectal tissues (C & D) using 1, 2, 3 or 4 × 10 micron slices. No significant difference in miRNA expression levels was noted depending on the number of slices (p>0.05).

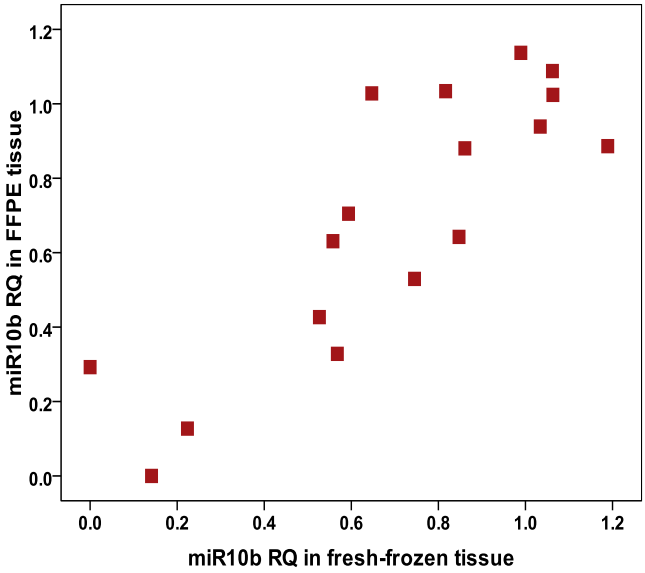


5.4.2 MiRNA expression in FFPE vs. fresh-frozen colorectal tissues

To further evaluate the efficiency of the RNA extraction method we compared the expression levels of miR-*10b, miR-143, miR-145* and *miR-21* in eight matched FFPE and Fresh-frozen colorectal tissues. In general our results showed a good correlation of miRNA expression profiles of FFPE samples compared to fresh-frozen samples, with correlation coefficients ranging from 0.85 to 0.61. This result was further strengthened by analysing the miRNA expression profiles in tumour compared to tumour-associated normal FFPE and matched fresh-frozen colorectal tissues. The expression levels of *miR-10b, miR-143* and *miR-145* were down-regulated while that of *miR-21* was up-regulated in tumour compared to normal in both FFPE and fresh-frozen tissues (figures 6.3, 6.4, 6.5 and 6.7)

Figure 5.4: **miR-10b expression in FFPE and FF CR tissues** (A) Significant correlation was noted in comparing the expression profiles of *miR-10b* in FFPE with fresh-frozen tissues (*r= 0.85, p<0.001*). (B) The expression levels of *miR-10b* were down-regulated in tumour compared to tumour-associated normal colorectal tissues in both FFPE (*p=0.001*) and fresh-frozen (*p=0.001*) colorectal tissues.

(A)



(B)

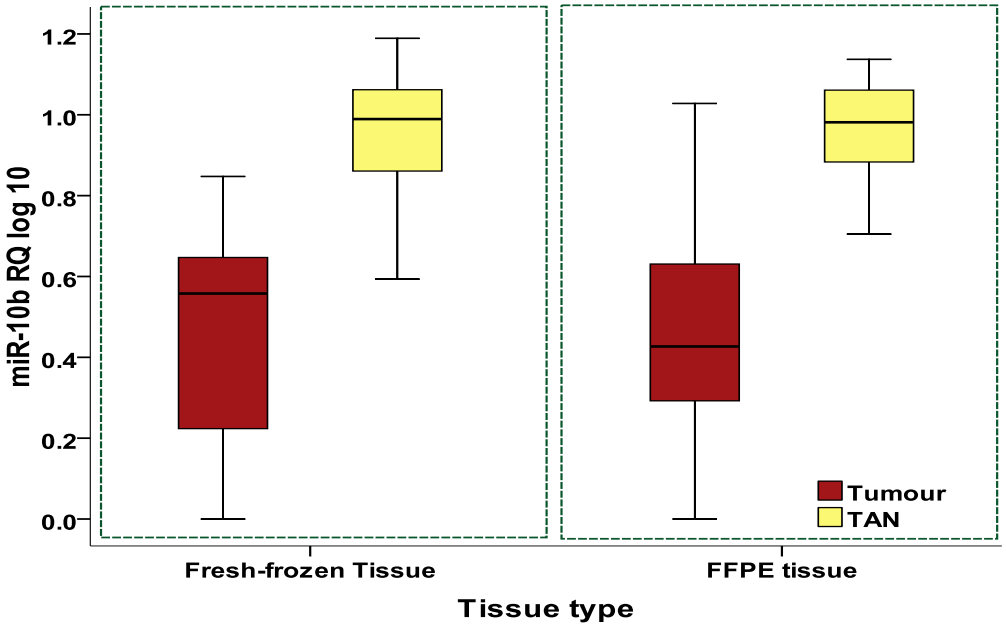
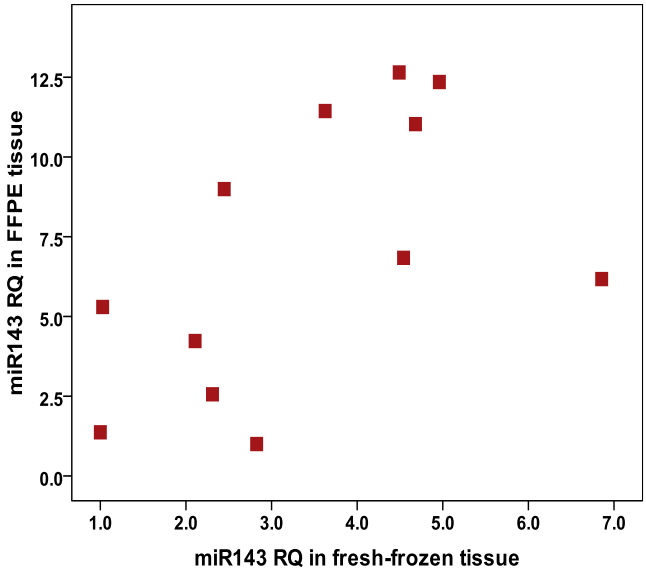


Figure 5.5: **miR-143 expression in FFPE and FF CR tissues**
(A) Significant correlation was noted in comparing the expression profiles of *miR-143* in FFPE with fresh-frozen tissues (*r= 0.61, p=0.036*). (B) The expression levels of *miR-143*were down-regulated in tumour compared to tumour-associated normal colorectal tissues in both FFPE (*p=0.016*) and fresh-frozen (*p<0.001*) colorectal tissues.

(A)



(B)

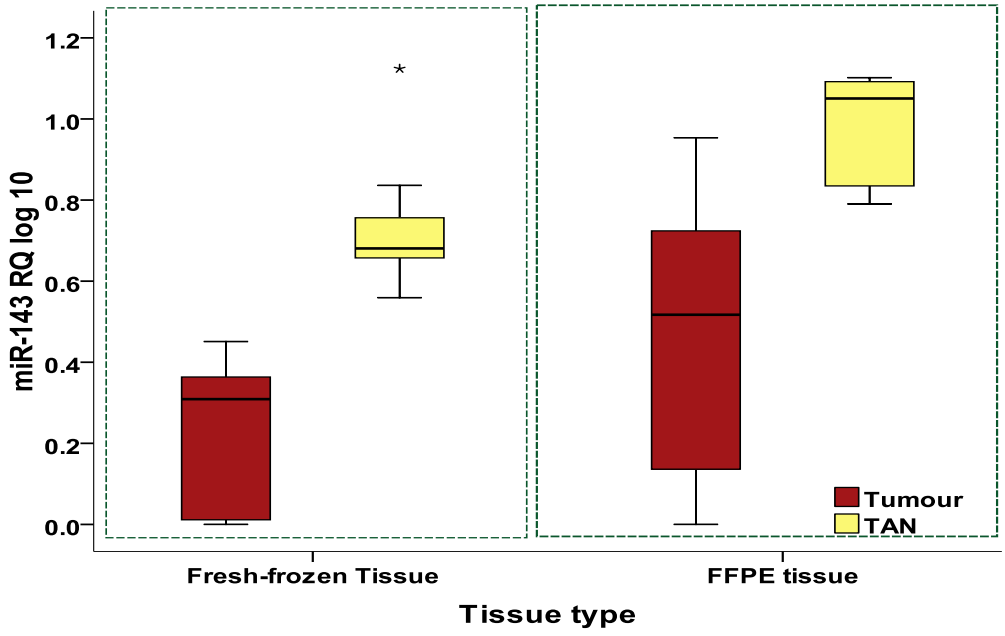
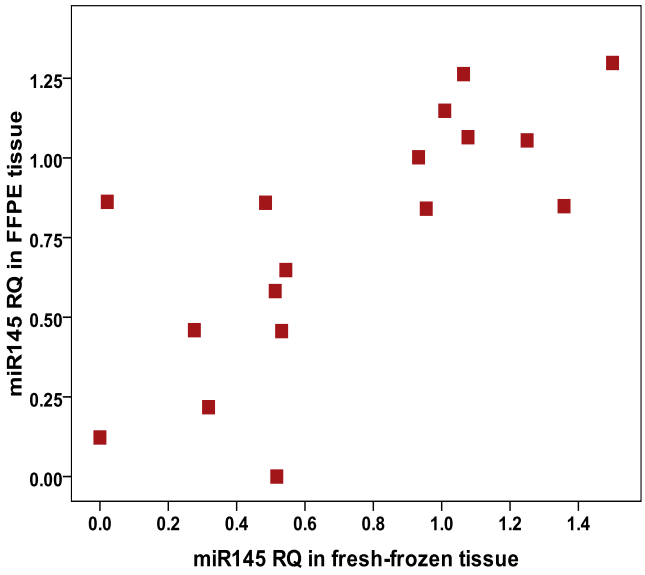


Figure 5.6: **miR-145 expression in FFPE and FF CR tissues**
(A) Significant correlation was noted in comparing the expression profiles of *miR-145* in FFPE with fresh-frozen tissues (*r=0.73, p=0.001*). (B) The expression levels of *miR-145* were down-regulated in tumour compared to tumour-associated normal colorectal tissues in both FFPE (*p=0.003*) and fresh-frozen (*p<0.001*) colorectal tissues.

(A)



(B)

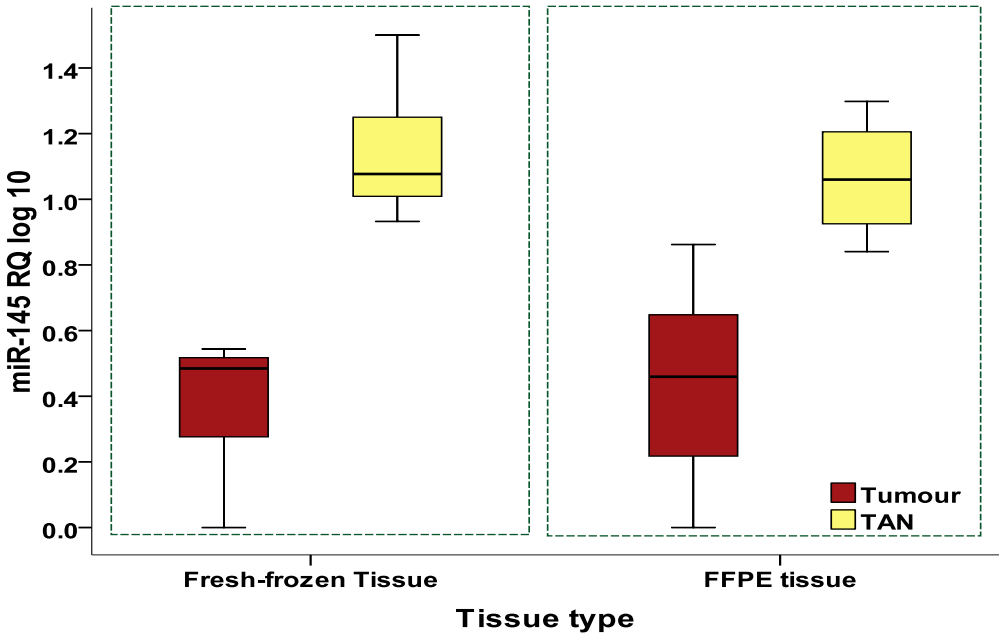
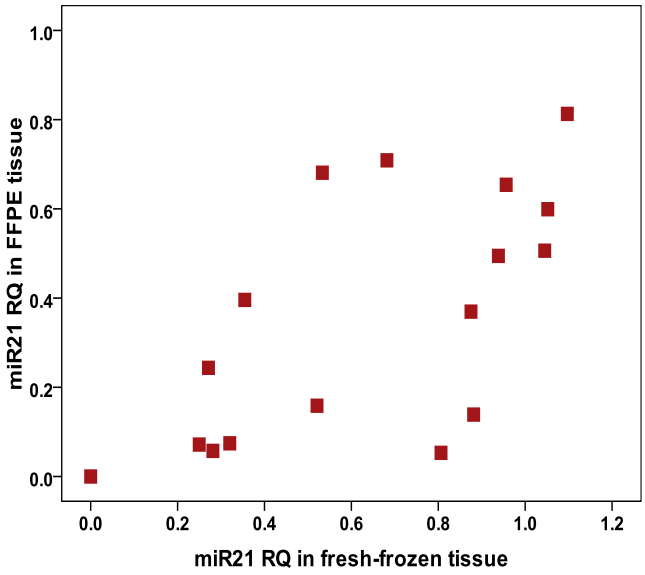
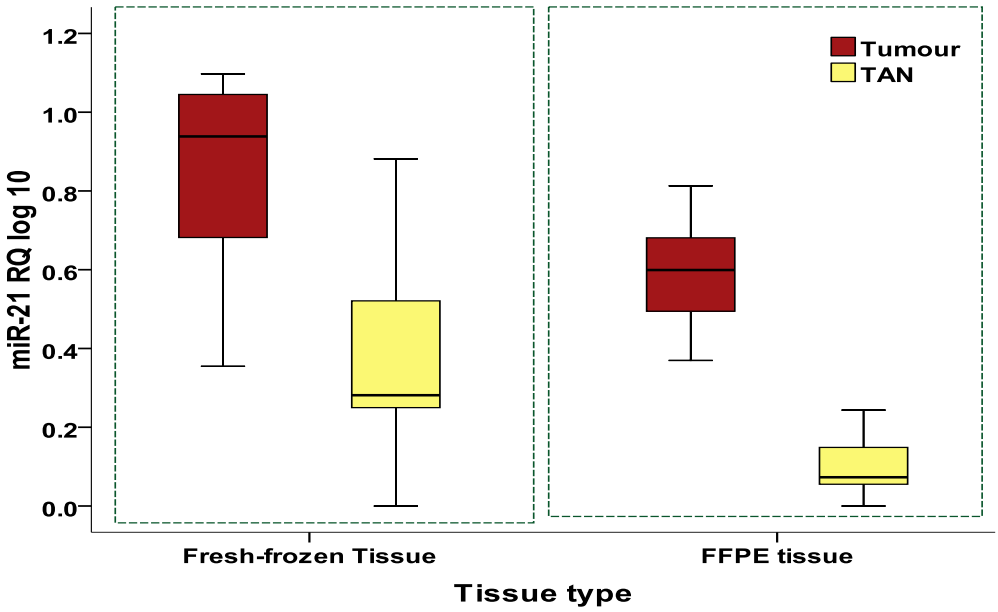


Figure 5.7: **miR-21 expression in FFPE and FF CR tissues**.(A) Significant correlation was noted in comparing the expression profiles of *miR-21* in FFPE with fresh-frozen tissues (*r= 0.67, p=0.003*). (B) The expression levels of *miR-21* were down-regulated in tumour compared to tumour-associated normal colorectal tissues in both FFPE (*p<0.001*) and fresh-frozen (*p=0.003*) colorectal tissues.

(A)



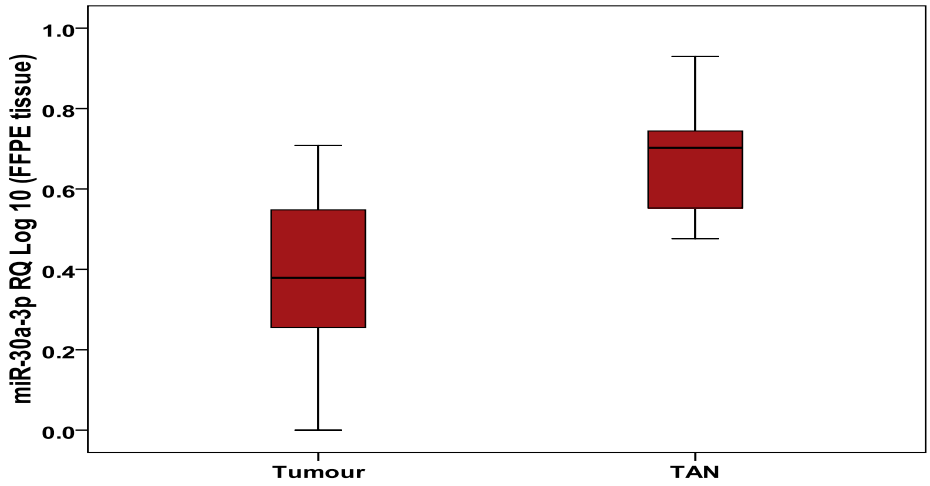
(B)



5.4.3 miR-30a-3p in colorectal

Expression profiles of *miR-30a-3p* were determined in 8 pairs of tumour and tumour-associated normal (TAN) colorectal tissues. We observed down-regulation of *miR-30a-3p* in tumours compared to TAN tissues (*p=0.029*, figure 6.8)

Figure 5.8: Expression profile of miR-30a-3p in CRC
MiR-30a-3p was down-regulated in tumour compared to TAN colorectal tissues (*p=0.029*)



5.4.3 MiRNAs as predictors of neoadjuvant chemoradiation response in RC

Using ANN to analyse the miRNA profiling data, we identified a distinct miRNA expression signature predictive of response to neoadjuvant CRT in 12 FFPE pre-treatment rectal cancer tissue samples. These signatures consisted of three miRNA transcripts (*miR-16, miR-590-5p* and *miR-153)* to predict complete vs. incomplete response and two miRNAtranscript *(miR-519c-3p and miR-561)* to predict good vs. poor response with a median accuracy of 100%. Details of this miRNAs in the signatures are presented in Table 3

Table 5.4: miRNAs predictor of complete response to neoadjuvant CRT
Complete Vs Incomplete response 10 steps model. The optimal model is in red (after the third, the predictive error increases)

Rank	miRNA	Average Valid. Error
1	**hsa-miR-16**	**0.145692**
2	**hsa-miR-153**	**0.138909**
3	**hsa-miR-590-5p**	**0.109172**
4	hsa-miR-487b	0.181193
5	hsa-miR-32	0.134736
6	hsa-miR-101	0.134252
7	hsa-miR-219-2-3p	0.138436
8	hsa-miR-488	0.152497
9	hsa-let-7f	0.157066
10	hsa-miR-494	0.183324

Table 5.5: Complete *vs.* Incomplete response to neoadjuvant CRT
The top 12 single predictors of complete vs. incomplete response to neoadjuvant CRT ranked by predictive error

Rank	miRNA	Median Train Perf	Median Valid. Perf	Average Valid. Error
1	hsa-miR-16	0.857143	0.8833334	0.1686571
2	hsa-miR-153	0.857143	0.9166667	0.1741115
3	hsa-miR-598	0.857143	0.7000003	0.207054
4	hsa-miR-375	0.842857	0.666667	0.2324498
5	hsa-miR-429	0.7928572	0.666667	0.2326782
6	hsa-miR-200a	0.7714287	0.666667	0.2241302
7	hsa-miR-339-3p	0.857143	0.6833336	0.2091335
8	hsa-miR-196b	0.7857145	0.666667	0.216506
9	hsa-let-7c	0.807143	0.666667	0.2370371
10	hsa-miR-198	0.857143	0.6833336	0.204131
11	hsa-miR-210	0.8142859	0.666667	0.2369278
12	hsa-miR-376a	0.7285717	0.666667	0.2404988

Table 5.6: miRNAs predictor of good response to neoadjuvant CRT
Good VS Poor response 10 steps model. The optimal model is in red (after the second, the predictive error increases)

Rank	miRNA	Average Valid. Error
1	**hsa-miR-519c-3p**	**0.074871**
2	**hsa-miR-561**	**0.042763**
3	hsa-miR-410	0.118914
4	hsa-miR-146b-3p	0.099968
5	hsa-miR-367	0.123437
6	hsa-miR-642	0.070075
7	hsa-miR-329	0.117644
8	hsa-miR-616	0.092262
9	hsa-miR-376b	0.133452
10	hsa-miR-551b	0.112602

Table 5.7: Good *vs.* Poor response to neoadjuvant CRT
The top 12 single predictors of good vs. poor response to neoadjuvant CRT ranked by predictive error

Rank	miRNA	Median Train Perf	Median Valid. Perf	Average Valid. Error
1	hsa-miR-519c-3p	1	1	0.0734598
2	hsa-miR-517b	0.857143	0.7500002	0.1663234
3	hsa-miR-582-3p	0.714286	0.666667	0.2330682
4	hsa-miR-149	0.714286	0.666667	0.2397969
5	hsa-miR-520g	0.714286	0.666667	0.2525955
6	hsa-miR-127-3p	0.714286	0.666667	0.2759746
7	hsa-miR-205	0.7357145	0.666667	0.2358426
8	hsa-miR-488	0.7571431	0.666667	0.2422655
9	hsa-miR-224	0.714286	0.666667	0.2616239
10	hsa-miR-184	0.714286	0.666667	0.2572127
11	hsa-miR-137	0.714286	0.666667	0.2587016
12	hsa-miR-96	0.714286	0.6333336	0.2481213

5.5 Discussion

Various limitations are associated with the retrieval of pre-treatment fresh biopsies from patients undergoing neoadjuvant treatment. Therefore, an alternative source of genetic material should be investigated. A valuable well characterised archival collection of FFPE tissue, linked to clinical databases was created worldwide for over a century. This archive provides a rich resource from which biological insight could be derived beyond the prospective collection of fresh-frozen sample. Moreover, the biomarkers developed from FFPE samples could be more readily translated into clinical practice. There have been extensive evaluations conducted on the quality of mRNA and miRNA isolated from FFPE tissues. miRNAs may be less prone to degradation and modification compared to mRNA, and good quality miRNAs were extracted from up to 12 years preserved tissues [538]. Direct comparison of mRNA profiling of FFPE versus fresh-frozen tissues has showed a correlation coefficient of only 0.28 compared to above 0.9 coefficient in case of miRNA expression analysis [167, 537]. Hence; the only mRNA experiments that could be conducted using FFPE samples is to measure a previously determined transcript, which will not allow for the identification of novel biomarkers. These results provide a solid foundation for using miRNAs as biomarkers when using FFPE samples in targets discovery studies.

In 1993 the first miRNA, *Lin-4*, was discovered in *C.elegans* [216, 560]. Michael et al. in 2003 have published the first report of miRNA in colorectal cancer [263]. They found reduced accumulation of specific miRNA in colorectal neoplasia and identified 28 different miRNA sequences between colonic cancer and normal mucosa. They also identified the human homologues of murine *miR-143* and *miR-145*. Since then numerous reports have demonstrated the role of miRNAs in colorectal carcinogenesis and highlight their potential use as diagnostic and prognostic biomarkers [274, 275, 561-563]. Moreover, increasing evidence support the use of miRNA profiling to characterise human tumours and distinct predictive signatures have been reported for hepatocellular carcinoma, oesophageal cancer, colon and lung cancer [264, 564-567]. Nevertheless, no study has investigated the possible use of miRNA in predicting response to neoadjuvant chemoradiation in rectal cancer.

For the purpose of this study, we compared the performance of three RNA extraction methods, and identified Qiagen RNeasy FFPE kit as a preferred methodology. The main reasons why RNA extracted from FFPE tissues is of poor quality are RNA fragmentation and cross-linked with other molecules including proteins [568]. The problem of fragmentation is solved by choosing small fragments for detection by PCR-based methods [568, 569]. Qiagen RNeasy FFPE kit uses Proteinase k at 55°C to break the cross-linked RNA formed with proteins. Incubation at 80°C in buffer PKD is an important step in RNA isolation process using this method. It partially reverses formaldehyde modification of nucleic acids; thereby improves the quality of RNA harvested (chapter 2). To ensure that the recovery of miRNA was adequately assessed it was crucial to select appropriate miRNA targets for integration by RQ-PCR. *miR-10b*, *miR-143*, *miR-145*, *miR-21* and *miR-30a-3p* were chosen because they were intensively investigated in colorectal cancer before [263, 264, 269, 570, 571]. Using FFPE and fresh-frozen tissue samples we were able to demonstrate the previously confirmed down-regulation of *miR-10b*, *miR-143*, miR-*145* and *miR-30a-3p*, and the over-expression of *miR-21* in colorectal cancer compared to tumour-associated normal tissues. The RQ-PCR amplification results reported here demonstrate that miRNA targets are detected at levels nicely matched expression levels from reference fresh-frozen tissues.

When comparing miRNAs expression levels between fresh-frozen and matched FFPE tissue samples, we observed some variations with correlation coefficients of 0.85 - 0.61. These variations could be attributed to the technical variations from one replicate to another. Also the amount of miRNA used in each reaction would have an impact on the miRNA expression level, as 5ng reaction was carried out when examining fresh-frozen tissues while 100ng were used in FFPE reactions. In addition some of these variations could be explained by the sample heterogeneity.

Furthermore, to enable extraction of miRNA from FFPE tissue blocks with different cross-sectional areas in quantities adequate for multiple analyses of the purified miRNA, we determined the number of slices required for optimal RNA yield. The purified RNA yield increased stepwise when we used 1, 2, 3, or 4 slices; however, the changes in concentration were not statistically significant.

Doleshal et al. performed RNA isolation in duplicate using 4, 8, 12 or 16 slices of FFPE tissues from two prostatic cancer locks that differ in their tissue cross-sectional area [538]. For the tissue blocks with smaller cross-sectional area they observed a linear increase in RNA recovery, while for the blocks with larger area not all the tissue was digested in tubes containing more than 4 slices resulting in yields that were lower than expected. To further evaluate the RNA recovered we selected *miR-143* and *miR-145* isolated from 1, 2, 3 or 4 slices for integration by RQ-PCR. The reactions were carried in triplicate for each slice number. Regardless of the number of slices used for miRNA extraction, the mean expression level of miRNAs was stable with standard deviation less than 0.3. This will confirm the suitability of this method for RNA isolation from tissue as small as a colonic biopsy retrieved during endoscopy procedure.

Microarray studies are frequently used to identify differential biosignature that distinguish two or more groups. Routinely processed FFPE samples represent an extensive and valuable resource for large-scaled, microarray-based molecular analysis. Major improvements were achieved in improving RNA extraction techniques and further RNA processing for microarray hybridisation, and recent reports provided evidence of the validity and utility of conducting mRNA and miRNA microarray profiling using FFPE tissues [536, 537, 572]. Roberts et al. obtained mRNA expression data for colon and lung tumour and normal FFPE samples and matched frozen samples and found significant agreement between the biosignature identified by each samples group using microarray technology [572]. Their microarray results were further validated and confirmed using RQ-PCR. Moreover, Hui et al. employed TaqMan low density arrays (TLDA) to assess the expression levels of miRNAs in FFPE breast cancer and normal tissues. They identified a high technical reducibility with intra-sample correlations above 0.9 and 92.8% accuracy in differential expression comparisons, indicating that such profiling studies are technically and biologically robust [537].

Several neoadjuvant treatment regimens have been described and established, including short-term radiotherapy, long-term radiotherapy alone or in combination with chemotherapy. The benefits of these therapeutic regimens have been examined and confirmed by numerous prospective trials [28, 31, 37, 38, 573-575].

Therefore, neoadjuvant therapy has become the preferred treatment modality for locally advanced rectal adenocarcinoma with a complete pathological response observed in up to 30% of patients [59, 576]. The ability to predict response to pre-treatment chemoradiation may spare poorly responding patients from undergoing aggressive and severely toxic treatment [577, 578] from which they would derive no benefit. At present there is no reliable technique to predict clinical or pathological complete tumour regression after treatment and limited data exist for each potential modality in this regards. Hence; molecular markers have been assessed for their predictive values. Studies on *P53* has shown *P53* gene mutations are significantly associated with radioresistance [46, 47]; however, immunohistochemistry and conformational polymorphism analyses have showed no potential of its use in this particular area [67]. Bengala et al., in their study looking at the predictive value of *K-ras* mutations and *EGFR* gene copy number for pathological response to pre-operative Cetuximab, 5-FU and radiation therapy in locally advancer rectal cancer, reported that 52.4% of patients with high gene copy number had a tumour regression grade 3 and 4 compared to 5.6% with low number [579]. Other molecules including *P21*[63], *EGFR* [64], *COX2*[580], *MUC2*[529] and growth hormone receptor [581] were also examined as potential markers; nevertheless, it seems unlikely that they will prove to be clinically useful response predictors

While expression profiling with microarray technologies has been broadly used to colorectal cancer for diagnosis, classification and prognostication based on pattern of expression, its application to response prediction to treatment is still unclear due to few currently available studies [186, 191, 192, 206]. The first report using microarray for prediction of response to pre-treatment radiotherapy was published in 2005. Ghadimi et al. used microarray for 30 patients in pre-operative group of the German Rectal Cancer Trial. Analysis of a 54 gene set allowed prediction of tumour response with 78% sensitivity and 86% specificity [186].Watanabe et al. achieved an accuracy of 82.4% in predicting radiotherapy response by using the 33 identified genes of which the expression differed significantly between responders and non-responders [206]. In addition, Kim and colleagues were able to predict response to chemoradiation neoadjuvant therapy in 84% of their training samples and 87% of the validation samples when they used the 95 top-

ranked of the 261 genes identified as differing between responders and non-responders [187]. Moreover, in the most recent report, Rimkus et al. identified a signature of 42 genes in 43 pre-treatment biopsies. Using the identified 42-gene set they were able to predict pre-therapeutic response with 71% sensitivity and 86% specificity. Applying different methods, Ojima et al. compared gene expression profiles in parent and radioresistant colorectal cancer cell lines using microarray analysis [528]. Then, they validated the identified differentially expressed genes using RQ-PCR in 30 irradiated rectal cancer patients. Their results suggest the potential of PTMA as a marker for predicting the response to radiotherapy.

Change in miRNA expression profiles during treatment of cancer could potentially provide a tool to predict and estimate the success of certain therapies. By enabling screening of tissue samples for several miRNAs simultaneously, microarrays revealed convincing evidence that a large number of miRNAs are deregulated in therapy resistance or sensitive cancer cells. The extent of changes in miRNA expression were reported following anticancer treatment with various chemotherapeutic drugs in different cancer cell lines and patient samples [582]. Kovalchuck et al. found 63 up-regulated and 75 down-regulated miRNA species when comparing doxorubicine-resistant and -sensitive breast cancer cell lines [583]. In gastric cancer cells, 10 miRNAs were reported to be down-regulated and 2 miRNAs were up-regulated in multidrug-resistant cell lines [584]. Two colorectal cancer cell lines, C22.20 and HC.21 cells, treated with 5-flurouracil were used to examine the expression of 153 miRNAs [285]. A group of 22 miRNA were found to be differentially expressed more than 2-fold after chemotherapy in C22.20 cells. Based on colorectal cancer biopsies analysis *miR-21, miR-181b* and *Let-7g* were reported to associate with response to 5-flurouracil-based anticancer therapy [264, 582]. Moreover, experiments with cancer cell lines in therapeutic radiotherapy revealed that irradiation causes significant variations in miRNA expression during therapy. The effect of radiotherapy on miRNA expression patterns was examined in both lung and prostatic cancers and in animal models and miRNAs like *Let-7* family were reported to induce radiosensitivity *in vivo* and *in vitro* [582, 585, 586].

To our knowledge, this is the first study to investigate the role of miRNA as predictors of response to neoadjuvant CRT therapy in rectal cancer. Using ANN to analyse the miRNA profiling data, we identified a distinct miRNA expression signature predictive of response to neoadjuvant CRT in 12 FFPE pre-treatment rectal cancer tissue samples. These signatures consisted of three miRNA transcripts (*miR-16, miR-590-5p* and *miR-153)* to predict complete vs. incomplete response and two miRNAtranscript *(miR-519c-3p and miR-561)* to predict good vs. poor response with a median accuracy of 100%.

Although *miR-16* was described as being stably expressed in both colorectal and breast tissues and has been highlighted as a good endogenous control for miRNA profiling in cancer research using RQ-PCR [293, 323], several studies confirmed its dysregulation in many cancers including CRC [587-590]. Moreover; Schaefer et al examined the expression of four putative reference genes including *miR-16* with regard to their use as normalizer in prostatic cancer and they found that normalization to *miR-16* can lead to biased results [591].

Since 2002, when Calin et al. have identified *miR-16* as a potential cancer gene in the pathogenesis of chronic lymphocytic leukemia (CLL) [220], neumours studies were designed to investigate the role of *miR-16* in tumourogenesis. Cimmino et al. suggested that *miR-16* induces apoptosis by targetting *BCL2* [238]. In addition, Liu et al. found that *miR-16* induces cell cycle arrest by regulating multiple cell cycle genes including *CDK6* and Cyclins D1, D3 and E1 [592]. The antiproliferative function of *miR-16* was also confirmed by Kaddar et al who demonistrated that *miR-16* can negatively regulate *HMGA1* and caprin-1 which are involved in cell proliferation [593]. Moreover; a recent report confirmed that *miR-16* is antiproliferative in enterocyte and exhibits diurnal rhythmicity in intestinal crypts [594]. Of further interest; Zhang et al. investigated the temporal and functional regulation of Wip1, which is a critical inhibitor in the ATM/ATR-p53 DNA damage signaling pathway [595]. Wip1 is transcriptionally induced by p53 at the early stage of the DNA damage response. Zhang and his colleaques identified *miR-16* that specifically targets the mRNA of Wip1 and thus negatively regulates the expression level of Wip1 [595]. They further examined *miR-16* expression in mammary tumour stem cells and found that overexpression of *miR-*

16 or inhibition of Wip 1 suppresses the self-renewal and growth of mammary stem cells and sensitizees MCF-7 human breast cancer to the chemotherapeutic agent doxorubicin [595].

Role of *miR-16* was also explored in relation to cancer therapy. Xia et al. investigated the possible role of miRNAs in the development of multidrug resistance in gastric cancer cells and suggest that *miR-15b* and *miR-16* could play a role in the development of multidrug resistance in gastric cancer cells at least in part by modulation of apoptosis via targeting *BCL2 [584]*. In their study, in vitro drug sensitivity assay demonstrated that overexpression of *miR-15b* or *miR-16* sensitized GC7901/VCR cells to anticancer drugs whereas inhibition of them conferred SGC7901 cells MDR. The downregulation of *miR-15b* and *miR-16* in SGC7901/VCR cells was concurrent with the upregulation of Bcl-2 protein. Enforced *miR-15b* or *miR-16* expression reduced Bcl-2 protein level and the luciferase activity of a *BCL2* 3' UTR-based reporter construct in SGC7901/VCR cells, suggesting that *BCL2* is a direct target of *miR-15b* and *miR-16.* Moreover; Takeshita et al. reported that systemic delivery of synthetic *miR-16* inhibits the growth of metastaitc prostate tumours via down-regulation of multiple cell-cycle genes [596]. Expression level of *miR-16* was also found to affect breast cancer cell lines response to tamoxifen [597]. Cittelly et al. showed that ectopic expression of a clinically important oncogenic isoform of *HER2*, *HER2Δ16*, promotes tamoxifen resistance and estrogen independence of MCF-7 xenografts. MCF-7/HER2Δ16 cells evade tamoxifen through upregulation of BCL2, whereas mediated suppression of *BCL2* expression or treatment of MCF-7/HER2Δ16 cells with the *BCL2* family inhibitor ABT-737 restores tamoxifen sensitivity. Tamoxifen-resistant MCF-7/HER2Δ16 cells upregulate BCL2 protein levels in response to suppressed ERα signaling mediated by estrogen withdrawal, tamoxifen treatment or fulvestrant treatment. In addition, HER2Δ16 expression results in suppression of BCL2-targeting miRNAs *miR-15a* and *miR-16*. Reintroduction of *miR-15a/16* reduced tamoxifen-induced *BCL2* expression and sensitized MCF-7/HER2Δ16 to tamoxifen. Conversely, inhibition of *miR-15a/16* in tamoxifen-sensitive cells activated *BCL2* expression and promoted tamoxifen resistance. Expression of miRNAs after ionizing radiation in human endothelial cells was investigated by Wanger-Ecker et al [598]. They reported that radiation

up-regulate *miR-16* expression levels. Their data also suggested that the miRNAs which are differentially expressed after radiation modulate the intrinsic radiosensitivity of endothelial cells in subsequent irradiations. This indicates that miRNAs are part of the innate response mechanism of the endothelium to radiation [598].

Altough no report has determined the significance of *miR-153* and *miR-590* in CRC, their role in carcinogenesis ws highlighted before [599]. Shan et al. [599] investigated the role of miRNAs on the expression and regulation of transforming growth factor-beta1 (*TGFB1*), TGF-beta receptor type II (*TGFBRII*), and collagen production *in vivo* and *in vitro*. They found that nicotine produced significant upregulation of expression of *TGFB1* and *TGFBRII* at the protein level, and a decrease in the levels of miRNAs *miR-133* and *miR-590*. This downregulation of *miR-133* and *miR-590* partly accounts for the upregulation of *TGFB1* and *TGFBRII*. Transfection of *miR-133* or miR-590 into cultured atrial fibroblasts decreased *TGFB1* and *TGFBRII* levels and collagen content. These effects were abolished by the antisense oligonucleotides against *miR-133* or *miR-590*. Therefore, their data stablished *TGFB1* and *TGFBRII* as targets for *miR-133* and *miR-590* repression. On the other hand, *miR-153* associations were confirmed in relation to glioplastoma and ovarian and endometrial cancers [600-602]. Xu et al. reported that down regulation of *BCL2* and *MCL1* by *miR-153* induce apoptosis in glioplastoma cell lines [602]. In addition; Myatt et al. reported *miR-153* as one of the miRNAs that repress expression of *FOXO1*, a tumour supressor gene, in endometrial cancer [601].

As mentioned above we identified two miRNAs signature to predict good vs. poor response to neoadjuvant CRT in rectal cancer. None of the miRNAs identified in our study was reported to be associated with CRC.Howevere; the role of *miR-519* in cancer was documented before [603-605]. *miR-519* was reported as a tumour suppressor and was found to reduce cell proliferation by lowering RNA-binding protein HuR levels [603]. It decreases HuR translation without influencing HuR mRNA abundance [603, 604]. Abdelmoshen et al examined the level of miR-519 and HuR in pairs of cancer and adjacent normal tissues from ovary, lung and kidney and reported significant high levels of HuR, unchanged HuR mRNA

concentration and reduced *miR-519* levels in cancer specimens compared to normal tissues [604]. They also found that tumour cells overexpressing *miR-519* fpormed significantly smaller tumours while those expressing reduded *miR-519* gave rise to substancilally larger tumours.

Taken together, therefore, using microarray analysis of pretreatment FFPE rectal cancer tissues we identified for the first time a group of miRNAs predictors of response to neoadjuvant CRT. This, indeed, can lead to a significant improvement in patient selection criteria and personalized rectal cancer management. However; before clinically applying this data, a validation study using a large cohort of patients needs to be designed.

Chapter 6:

miRNA: mRNA correlations

6.1 Introduction

MiRNAs are small noncoding RNA of about 22-nucleotide (nt) long, which are cleaved from 70-100 nt endogenous double stranded precursors. Although their biological role is not fully understood, miRNAs are found in diverse organisms and epigenetically functions as negative regulators of gene expression. They are complementary to genomic regions and one of their modes of action is to bind 3-UTR and inhibit gene translation [606, 607]. Moreover; they can also function by cleaving and degrading a target mRNA, in which case the miRNA may target sequences out side the 3-UTR [608]. MiRNAs are crucial in eukaryotes gene regulation, especially in development and differentiation [609, 610], and their expression in cancers has indicated that they may have a tumour suppressor or oncogenic function [611]. Functional characterisation of miRNAs will depends heavily on identification of their specific gene targets. In addition; a number of studies have shown that more than one miRNA can potentially bind to a single targeted gene; hence multiple miRNAs may cooperatively control the expression of target genes [305, 306]. Numerous bioinformatics methods were developed to high-throughput prediction of miRNA target genes [295, 300-302, 304], although it is understood that the presumed targets have to be validated experimentally.

Computational approaches have been developed based on an understanding of the relationship between the miRNA seed region and the 3-UTR of the target gene. To develop computational algorithms, empirical evidence is examined carefully and principles of miRNA target recognition are extracted. After preparation of the data set, miRNA binding sites are identified by determining the base pairing pattern of miRNAs and mRNAs according to the complementarity within specific region, thermodynamic analysis of the miRNA/mRNA duplexes via calculation of the free energy, and comparative sequence analysis. Then, the number of target sites of miRNA is counted [312, 612-614]. However; most of the available computational prediction algorithms have a group of drawbacks, includes [308]:

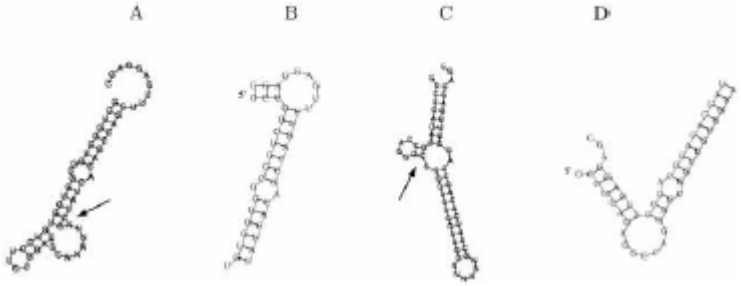
1- The sequences have to fit together with a short linker sequence that can lead to artefact in the prediction.
2- The hybridizations of the target or the miRNA with itself or with the linker may form an internal hairpin in either sequence; hence the energy calculated

would include the energy of the hairpin and be a misleading indicator of combination strength (figure 5.1).

3- For prediction of multiple binding sites in one target, the appropriate potential binding sites have to be cut out and folded separately.

Therefore, these algorithms can result in prediction of false-positives. Moreover; some targets may pass undetected. The false positive rates were estimated at 22%, 24% and 30% for TargetScan, miRanda and PicTar, respectively [312]. The PicTar and EMBL algorithms have a reported sensitivity of 70-80% [615] indicating 20-30% of targets may go undetected.

Figure 6.1: **Artefacts of target/miRNA combinations**
The structure exhibits hybridization between target and linker (A) or self-hybridization of the target (C) (arrow), and Corresponding prediction from that shows no artefact (*B, D*). [308]



It is necessary to confirm the bioinformatically presumed miRNA target experimentally. Verification of the biological function or target gene of a miRNA can be performed using gain-of-function and loss-of-function approaches *in-vitro* and *in-vivo*. As miRNAs modulate gene expression by both mRNA degradation and translation inhibition, the effect of the miRNA and its mimic and inhibitors on a target gene should be verified. That could be achieved using RQ-PCR and western blot at mRNA level and ELISA and immunohistochemistry at protein level.

The experimental verification is usually based on demonstrating that [616]:

1- The target protein is down-regulated by the predicted miRNA.
2- A reporter gene expressing the 3-UTR or the miRNA-binding sites of the targeted mRNA is also down-regulated
3- The targeted protein is not down-regulated when the 3-UTR is missing or blocked.
4- The miRNA has a biological function predicted by the biological function of the targeted protein.

6.2 Aims:

The aims of this study were to correlate the expression levels of candidate mRNA to a panel of miRNAs in order to identify miRNA/mRNA duplexes and to investigate the miRNA and target gene expression patterns in colorectal tissue samples using RQ-PCR.

6.3 Materials and methods

6.3.1 Study group

A group of 58 consecutive patients undergoing surgical resection for CRC, and in whom the expression levels of a panel of miRNA was carried out before in the surgical research laboratory, were selected for the miRNA: mRNA correlations study in order to determine miRNA targeting a panel of genes (table 5.1).

Table 6.1: **Clinicopathological data of study group**

Clinicopathological Variable	Number of Patients N=(58)
Mean Age (SD)	68.5 (12.0)
Gender	
Males	39
Females	19
Tumour Location	
Colon	19
Rectum	39
Distant Metastasis	
M_0	48
M_1	10
Nodal Status	
N_0	26
N_1	22
N_2	6
UICC Stage	
Stage 0	1
Stage I	11
Stage II	14
Stage III	17
Stage IV	10
pCR	5
Tumour Differentiation	
Grade 1: Well differentiated	8
Grade 2: Moderate differentiated	35
Grade 3: Poor differentiated	7
Not applicable	8
Mucin Secretion	
Mucinous	13
Non-mucinous	45

6.3.2 RNA extraction and analysis

Tissue samples (50-100 mg) were homogenised using a hand-held homogenizer (Polytron PT1600E) in 1-2 mL of QIAzol reagent (Qiagen). Two methods of RNA extractions were employed in the study, the total RNA extraction (co-purification) and the separate purification of mRNA and miRNA. RNA was extracted using the RNeasy Plus Mini Kit and RNeasy MinElute cleanup kit (Qiagen) according to the manufacturer's instructions. RNA concentration, purity and integrity were assessed in duplicate samples using a NanoDrop ND-1000 Spectrophotometer (NanoDrop technologies) and the Agilent 2100 Bioanalyzer System (Agilent technologies).

6.3.3 Reverse transcription

6.3.3.1 mRNA Reverse transcription

First strand cDNA was synthesised using Superscript III reverse transcriptase (Invitrogen) and random primers (N9; 1µg, MWG Biotech). Negative control samples were included in each set of reactions. Reactions were incubated at 25º C for 5 minutes followed by 50º C for 1 hour and final denaturation at 72º C for 15 minutes. Samples were subsequently diluted to 100 µL in nuclease-free water and stored at -20º C.

6.3.3.2 miRNA Reverse transcription

First strand cDNA was synthesised using gene specific stem-loop primers. Primers were obtained from MWG Biotech (Germany) if sequences were available. Otherwise, assays containing stem-loop primer were purchase from Applied Biosystems. All reagents were included in High-capacity cDNA reverse transcription kit (Applied Biosystems). The reactions were performed using a GeneAmp PCR system 9700 thermal cycle (Applied Biosystems) with sample incubated at 16 º C for 30 minutes, 42 º C for 30 minutes and 85 º C for 5 minutes. An RT-negative control was included in each batch of reactions.

6.2.4 Real-time quantitative PCR

The expression of each gene was analysed by RQ-PCR using TaqMan gene expression assays using a 7900HT instrument (Applied Biosystems). All reactions were performed in 20 µL reactions, in triplicate within the same PCR run.

Negative controls were included for each gene target under assay. On each plate, an interassay control was included to account for any variations between runs. For mRNA, each well 2µl of cDNA from each sample was added to 18µl of PCR reaction mix which consisted of 10 x TaqMan universal master mixes, No AmpErase UNG, 7x nuclease free water and 1x gene expression assay primer-probe mix (Applied Biosystems). For miRNA, reaction mix consist of 10 µL 2 x TaqMan universal master mix, No AmpErase UNG, 1 µL 0.2 µM TaqMan probe, 3 µL 1.5 µM of forward primer, 1.4 µL 0.7 µM reverse primer (Applied Biosystems), and 1.33 µL of cDNA. The PCR reactions were initiated with 10 minute incubation at 95º C followed by 40 cycles of 95º C for 15 seconds and 60º C for 60 seconds, in accordance with the manufacturer's recommendations. The threshold standard deviation for intra- and inter-assay replicates was 0.3.

Percent PCR amplification efficiencies (E) for each assay were calculated as E = (10-1/slope - 1) × 100, using the slope of the semi-log regression plot of C_t versus log input of cDNA (10-fold dilution series of five points) A threshold of 10% above or below 100% efficiency was applied.

6.3.5 Relative quantification

Cycle threshold (C_t) is defined as the PCR cycle number at which the fluorescence generated from amplification of the target gene within a sample increases to a threshold value of 10 times the standard deviation of the base line emission and is inversely proportionate to the starting amount of the target cDNA. QBasePlus was used for calculation of candidate expression relative to the endogenous control genes. It applies ΔΔ C_t method was used where ΔΔCt = (C_t target gene, test sample – C_t endogenous control, test sample) - (C_t target gene, calibrator sample - C_t endogenous control, calibrator sample). Relative quantities were corrected for efficiency of amplification and fold change in gene expression between groups was calculated as E-ΔΔCt ± s.e.m. The lowest expressed sample was used as a calibrator.

6.3.6 Statistical analysis

Statistical analysis was carried out with IBM SPSS Statistics 17.0 (SPSS Inc.). Data was tested for normal distribution. Correlation analysis used Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively. The correlation data interpretation was carried out following Cohen's guidelines [314] (table 2.10). Differences between groups were calculated using Analysis of Variance (ANOVA) with post-hoc Tukey and Student's t-tests. P values of less than 0.05 were considered statistically significant for all tests

6.4 Results

6.4.1 Computationally predicted miRNA/mRNA duplexes:

We performed bioinformatics analyses to search for putative miRNA/target gene duplexes from our panels of miRNAs and mRNAs. TargetScan, PicTar, miRDB and miRBase were each used to independently predict miRNA targets. A list of miRNA/mRNA duplexes was determined for further validation analysis (table 5.2). These candidate miRNA/targets were validated based on an inverse relationship between miRNA and their targets, their expression patterns in colorectal tissues, and their relation to clinicopathological parameters

Table 6.2: Computationally predicted miRNA/mRNA duplexes using miRBase, miRDB, PicTar and TargetScan

Database / mRNA	miRBase	miRDB	PicTar	TargetScan
CDH17	*miR-143*	-	-	*miR-143*
CEACAM5	-	-	-	*miR-143* *miR-21* *miR-17*
CXCL12	*miR-10*	-	*miR-31*	
CXCR4	*miR-139*	-	*miR-139*	*miR-139*
CXCR7	*miR-10b*	-	-	*miR-10b*
FABP1	-	-	-	*miR-362*
IL8	*miR-145*	*miR-17*	-	-
MUC2	-	-	-	-
PDCD4	*miR-17* *miR-21*	*miR-145* *miR-21*	-	-
TGFB1	*miR-139*	-	-	*miR-139*
TGFBR1	-	*miR-145*	-	-
TGFBR2	-	*miR-17* *miR-145*	-	-

6.4.2 Correlation of miRNA/mRNA expression:

Gene and miRNA expression was quantitated in 103 colorectal tissue specimens by RQ-PCR and normalized expression data was correlated using Spearman's Rho and Pearson's correlations coefficient for nonparametric and parametric data respectively.

6.4.2.1 miRNA/ putative mRNA correlations:

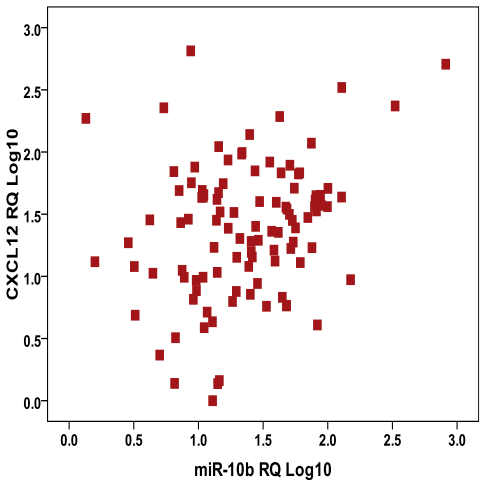
Expected positive correlations were noted between miR-145 and *PDCD4* (r = +0.57, $p<0.001$), mir-145 and *TGFBR1*(r = +0.48, $p<0.001$), and miR-10b and *CXCL12* (r = +0.28, $p=0.005$). An inverse significant correlations were observed between miR-31 and *CXCL12* (r = -0.41, $p<0.001$), miR-145 and *IL8* (r = -0.28, $p=0.004$), and miR-21 and *PDCD4* (r = -0.60, $p<0.001$) (figure 5.2). Other correlations did not reach statistical significance.

6.4.2.2 miRNA/ non-putative mRNA correlations:

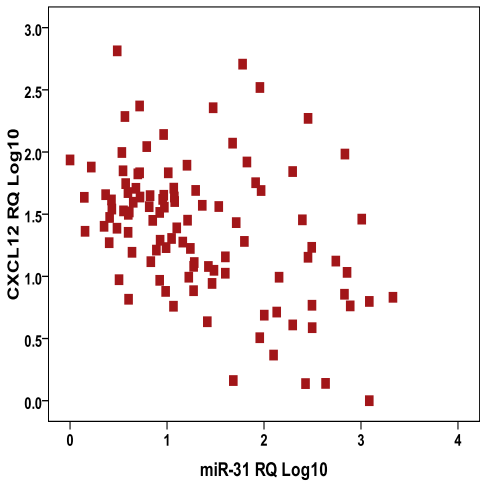
As described before, some of the miRNA targets might not be detected by computational algorithms. Therefore we correlated the expression levels of the candidate genes to the whole panel of miRNAs. The significant correlations are shown in figure 5.3. Inverse correlations were identified between miR-10b and *IL8* (r = -0.54, $p<0.001$), miR-21 and *FABP1* (r = -0.51, $p<0.001$), miR-17 and *CXCL12* (r = -0.48, $p<0.001$) and miR-31 and *PDCD4* (r = -0.52, $p<0.001$) and *FABP1*(r = -0.55, $p<0.001$) (figure 5.3). Although none of these pairs were determined by bioinformatics analysis, the might still be a valid miRNA/target gene duplexes.

Figure 6.2: miRNA and their putative targets with significant correlations

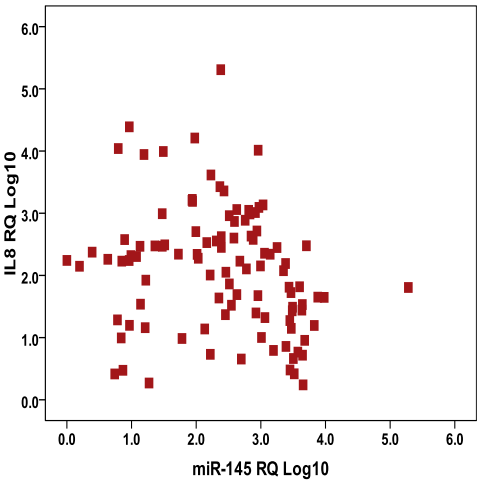
(A) (r = +0.28, *p = 0.005*)



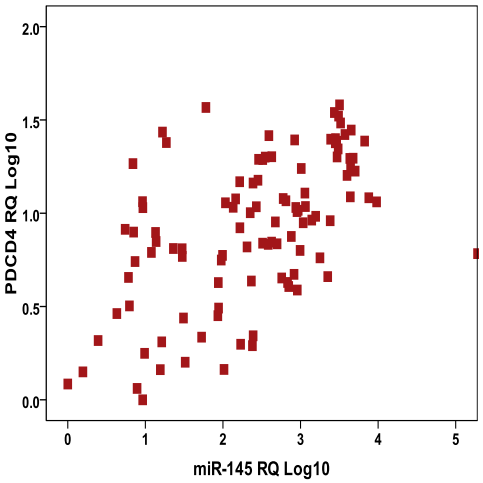
(B) (r = -0.41, *p <0.000*)



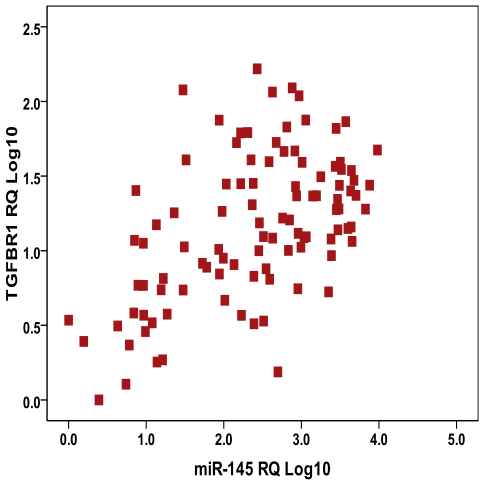
(C) (r = -0.28, *p =0.004*)



(D) (r = +0.57, *p <0.000*)



(E) (r = +0.48, *p <0.000*)



(F) (r = -0.60, *p <0.000*)

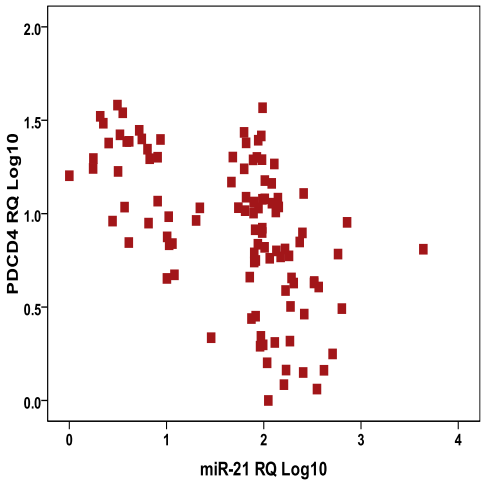
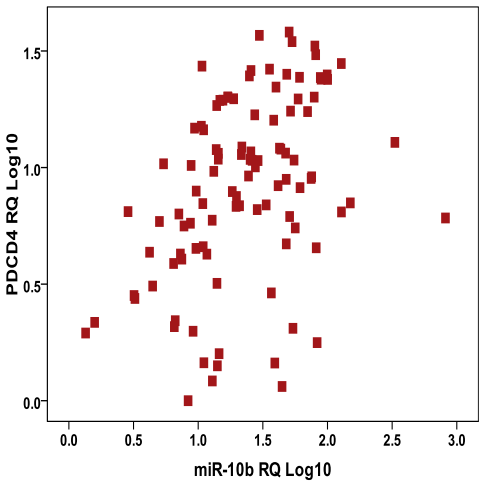
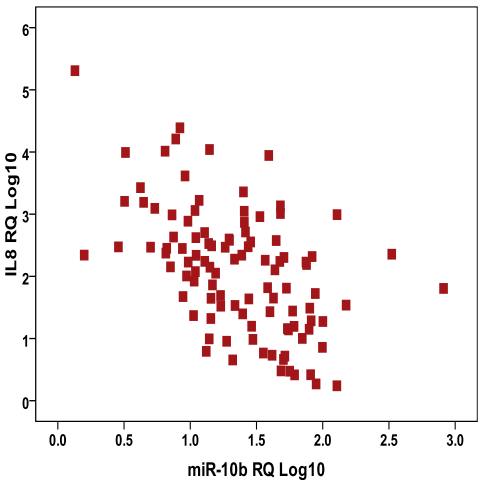


Figure 6.3: miRNA/mRNA duplexes with significant correlations

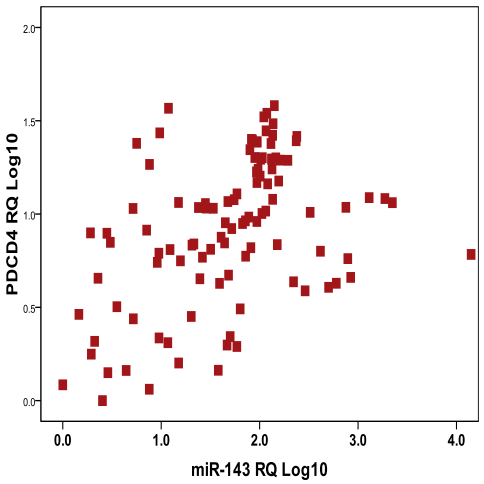
(A) (r = 0.41, *p <0.000*)



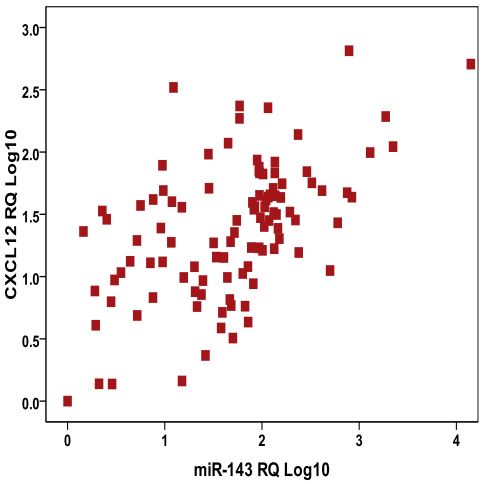
(B) (r = - 0.54, *p <0.000*)



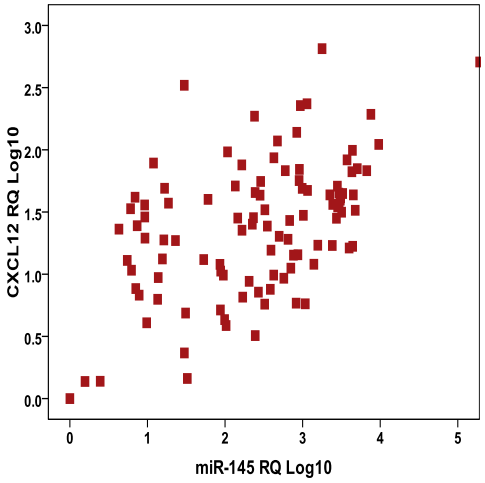
(C) (r = 0.42, *p <0.000*)



(D) (r = 0.55, *p <0.000*)



(E) (r = 0.47, *p <0.000*)



(F) (r = - 0.51, *p <0.000*)

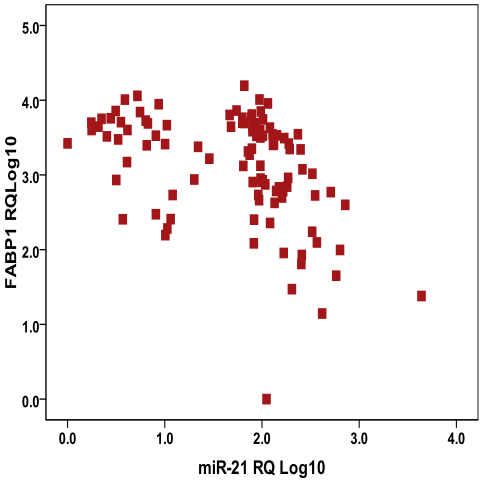
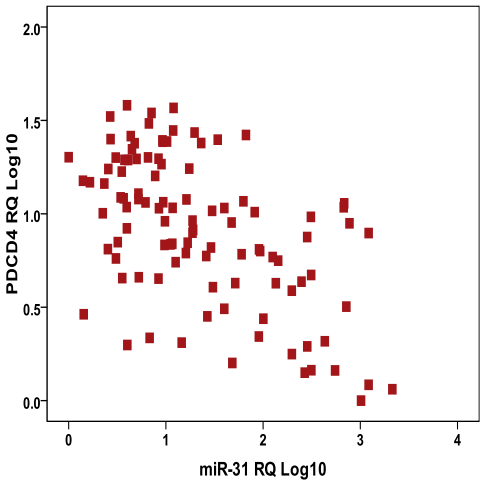
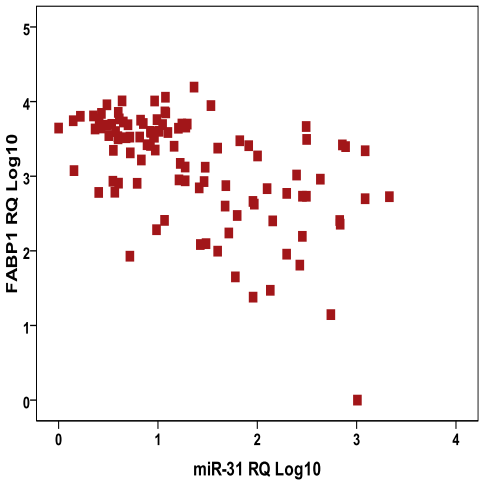


Figure 6.3: Continued

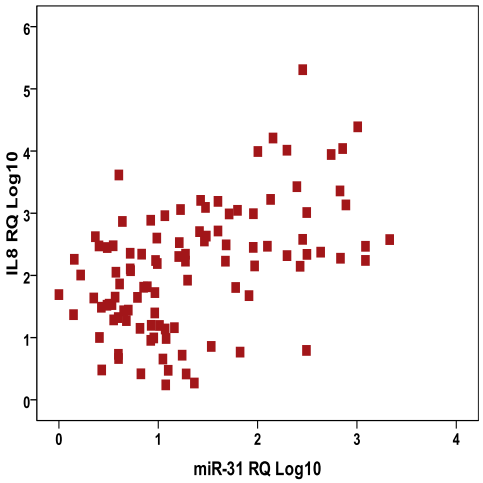
(G) (r = - 0.52, *p <0.000*)



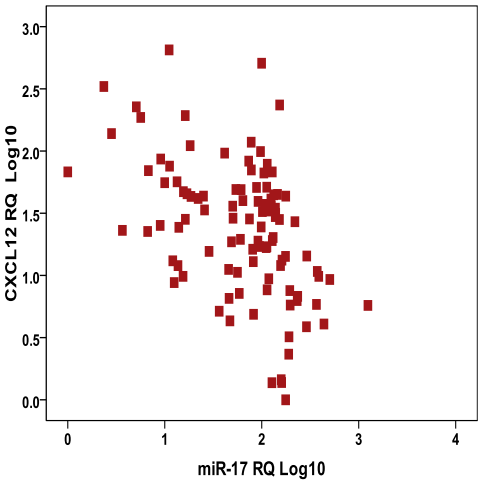
(H) (r = - 0.55, *p <0.000*)



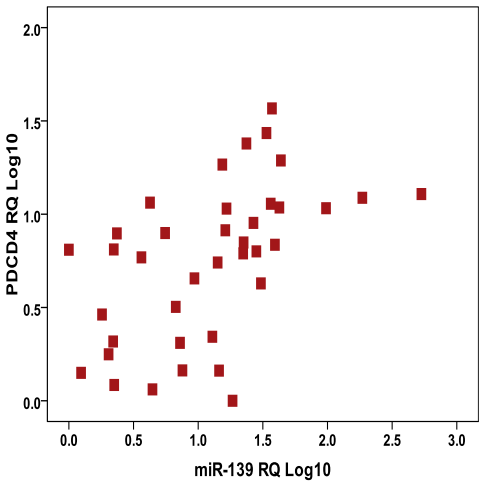
(I) (r = 0.47, *p <0.000*)



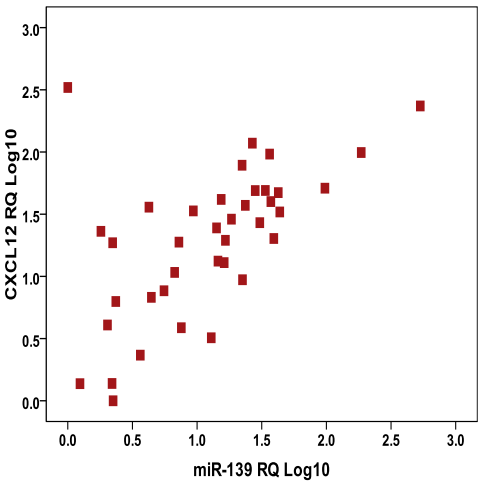
(J) (r = - 0.48, *p <0.000*)



(K) (r = 0.53, *p = 0.001*)



(L) (r = 0.59, *p = 0.001*)



6.4.3 Expression levels of miRNA/mRNA duplexes

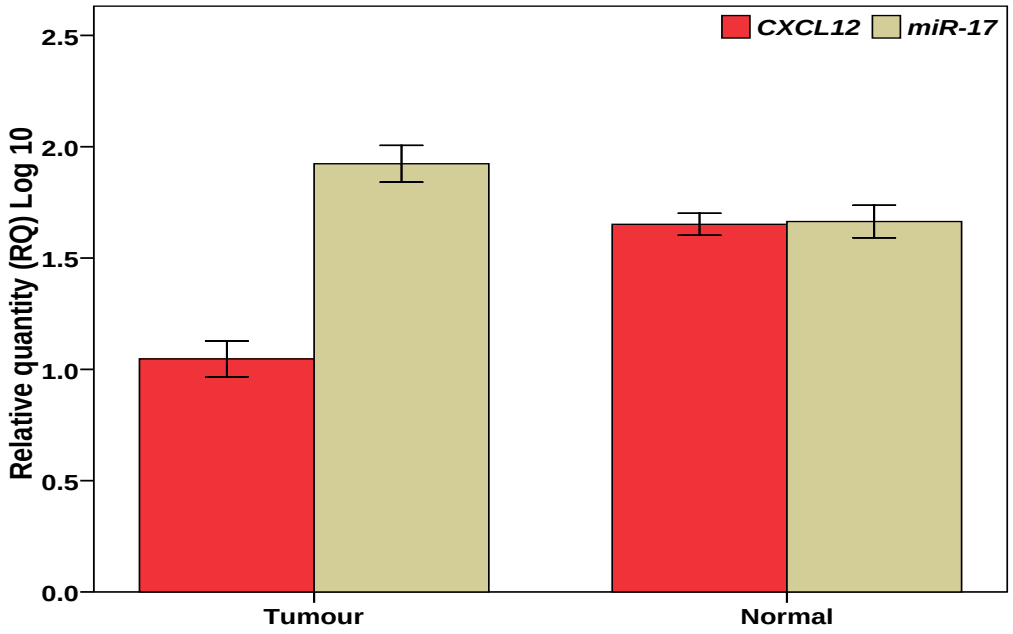
The expression patterns of the inversely correlated mRNA and miRNA were analysed in 46 tumour and 57 TAN colorectal tissues. Reciprocal patterns of expression were noted for all miRNA/ target mRNA duplexes in tumour compared to TAN tissues. The expression levels of *IL8 (p<0.001)*, miR-17*(p=0.021)*, miR-21*(p=0.003)* and miR-31*(p<0.001)* were significantly higher in tumours versus TAN tissues, in contrast to the expression of *CXCL12 (p<0.001), FABP1 (p<0.001), PDCD4 (p<0.001),* miR-10b *(p<0.001)* and miR-145 *(p=0.001)* which were down-regulated in tumour (figures 5.4, 5.5, 5.6 and 5.7).

Table 6.3: **mRNA/miRNA Expression patterns in colorectal tissues**

	Tissue type	**N**	**Mean Expression**	**Std. Deviation**	***P Value***
miR-10b	Tumour	46	1.127	.446	<0.001
	Normal	57	1.554	.413	
miR-145	Tumour	46	2.021	.864	0.001
	Normal	57	2.692	1.063	
miR-21	Tumour	46	1.930	.673	0.003
	Normal	57	1.515	.731	
miR-31	Tumour	46	1.977	.730	<0.001
	Normal	57	.832	.413	
miR-17	Tumour	46	1.923	.560	0.021
	Normal	57	1.664	.552	
PDCD4	Tumour	46	.640	.331	<0.001
	Normal	57	1.138	.287	
CXCL12	Tumour	46	1.047	.544	<0.001
	Normal	57	1.652	.372	
IL8	Tumour	46	2.854	.808	<0.001
	Normal	57	1.511	.697	
FABP1	Tumour	46	2.615	.697	<0.001
	Normal	57	3.543	.435	

Figure 6.4: Expression of *CXCL12 vs.miR-17* and *miR-31* in CR tissue
CXCL12 was down-regulated in tumours compared to normal tissues (p<0.001), in contrast to the expression levels of miR-17 (A, p=0.021) and miR-31 (B, p<0.001)

(A)



(B)

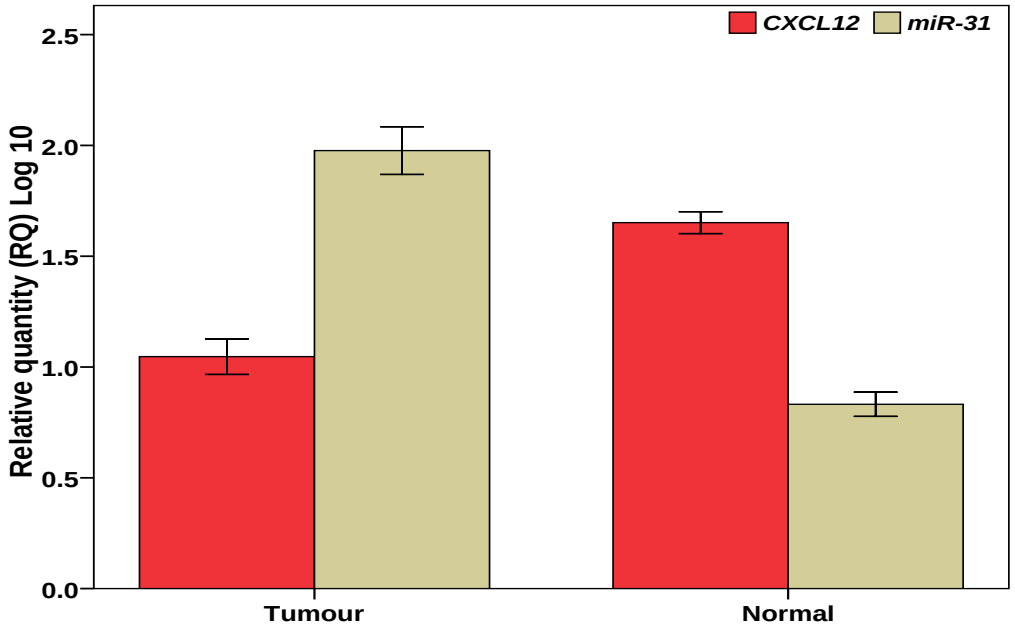
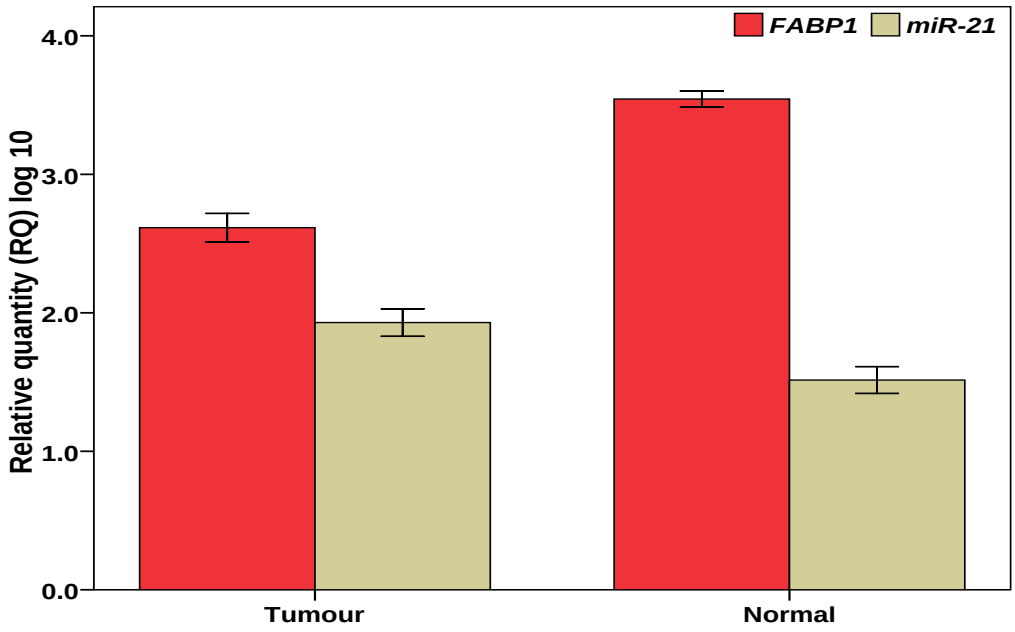


Figure 6.5: Expression of *FABP1 vs. miR-21* and *miR-31* in CR tissue
FABP1 was down-regulated in tumours compared to normal tissues (p<0.001), in contrast to the expression levels of miR-21 (A, p=0.003) and miR-31 (B, p<0.001)

(A)



(B)

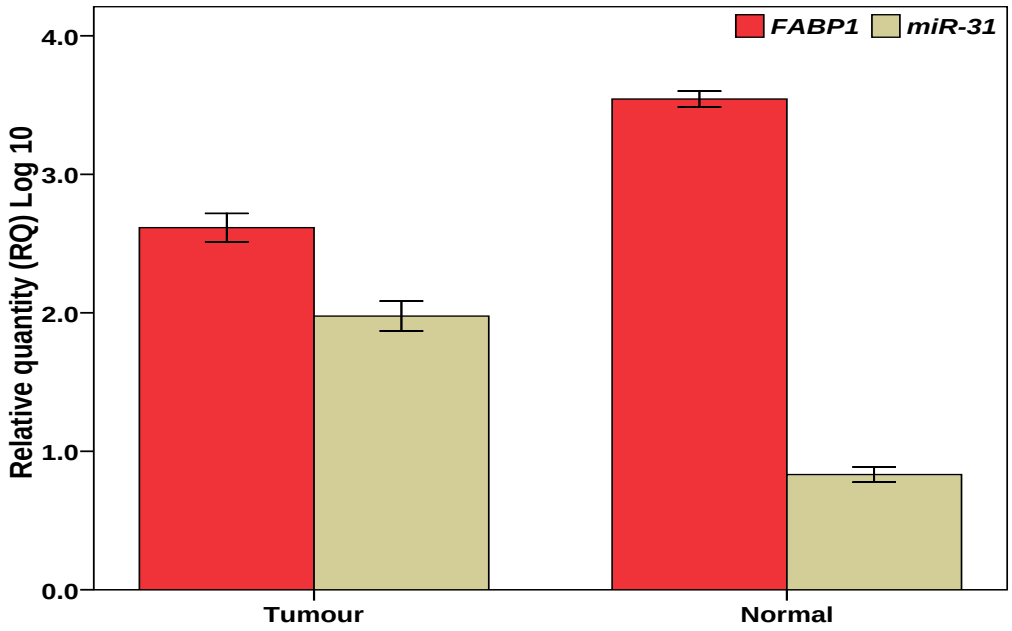
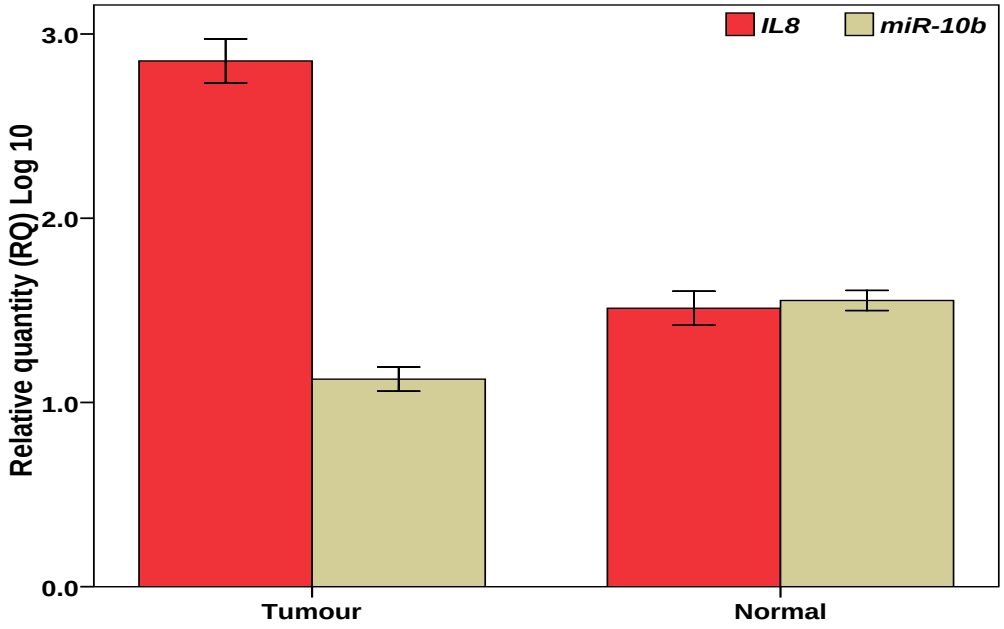


Figure 6.6: Expression of *IL8 vs. miR-10b* and *miR-145* in CR tissues
IL8 was up-regulated in tumours compared to normal tissues (p<0.001), in contrast to the expression levels of miR-10b (A, p=0.003) and miR-145 (B, p<0.001)

(A)



(B)

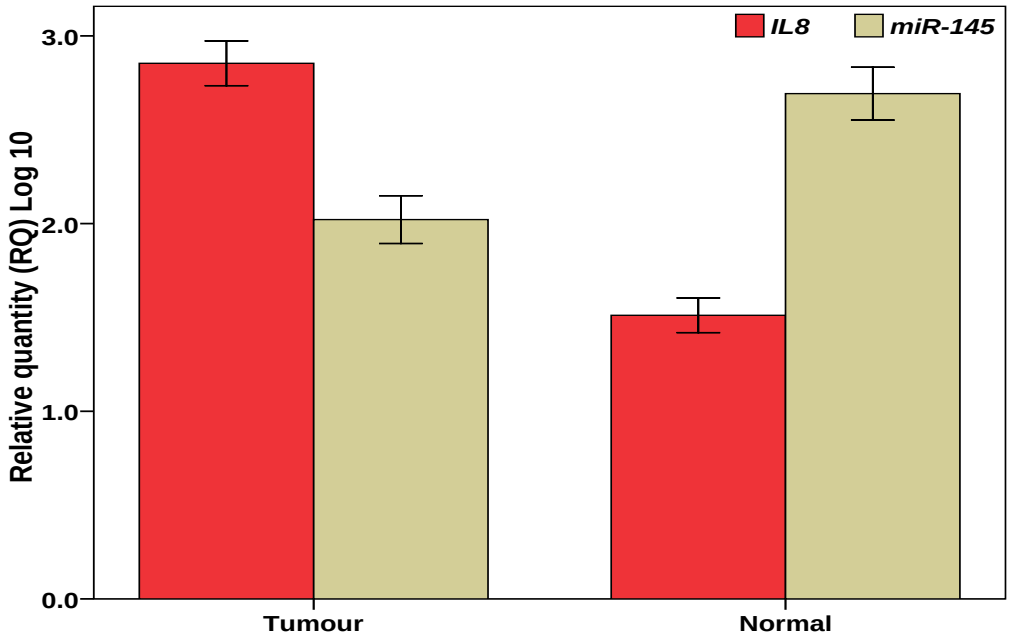
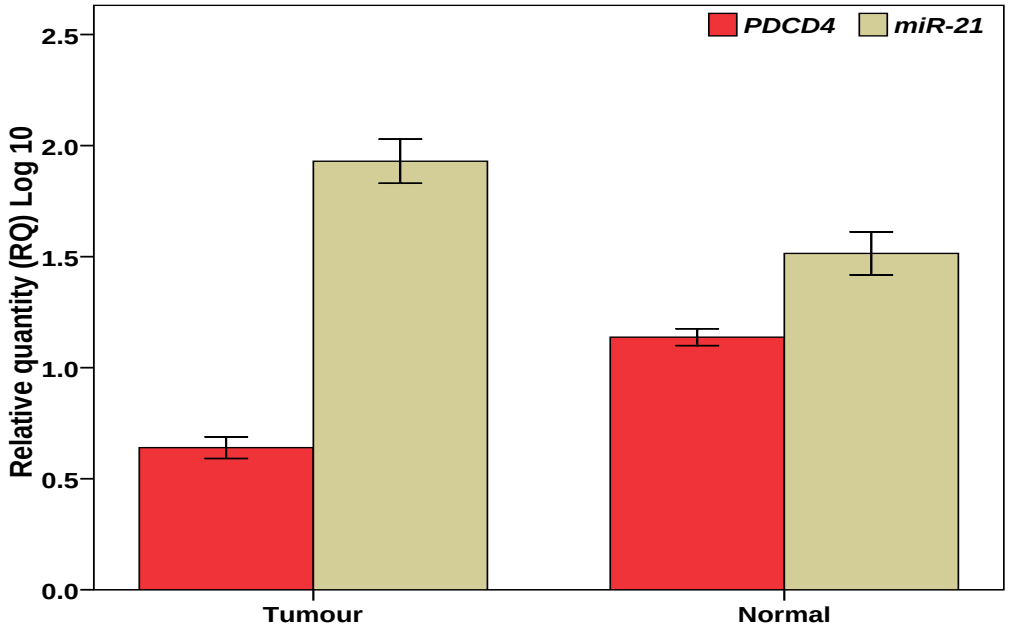
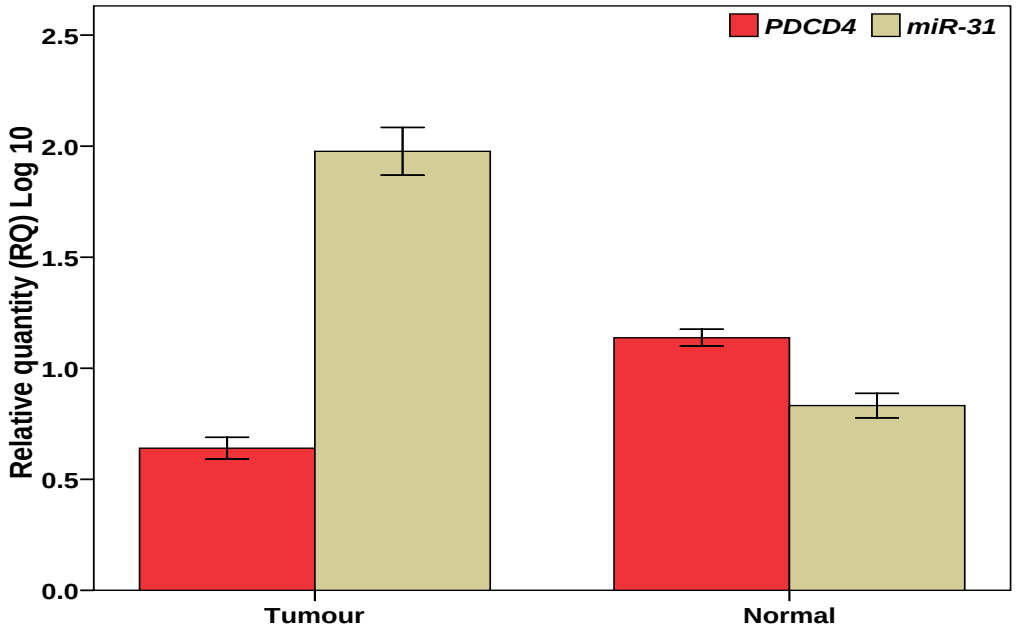


Figure 6.7: Expression of *PDCD4 vs. miR-21* and *miR-31* in CRC tissue
PDCD4 was down-regulated in tumours compared to normal tissues (p<0.001), in contrast to the expression levels of miR-21 (A, p=0.003) and miR-31 (B, p<0.001)

(A)



(B)



6.4.4 Clinicopathological associations of miRNA/mRNA duplexes

To further validate the miRNA/mRNA relationship, we examined the association of the duplexes in colorectal cancer. *CXCL12* and miR-17 showed significant association with tumour differentiation. The expression levels of *miR-17* progressively increased from well differentiated, to moderately differentiated, to poorly differentiated tumours, in contrast to *CXCL12* expression which decreased in the same manner (figure 5.8, table 5.4). Moreover, reciprocal pattern of expression of miR-17 and *CXCL12* was also noted in relation to tumour location as *CXCL12* expression was higher in proximal compared to distal tumours, while miR-17 expression was higher in distal tumours (figure 5.9).

In addition, *CXCL12* and miR-31 duplex was also investigated. Both miR-31 and *CXCL12* showed significant association with tumour location. The expression of *CXCL12* was higher in rectal tumours, while miR-31 was higher in colonic tumours (figure 5.10).

Figure 6.8: *miR-17* and *CXCL12* in association with tumour grade
Progressive up-regulation of *CXCL12* (p=0.019) and down-regulation of miR-17 (p=0.002) in poor, moderate and well differentiated CRC.

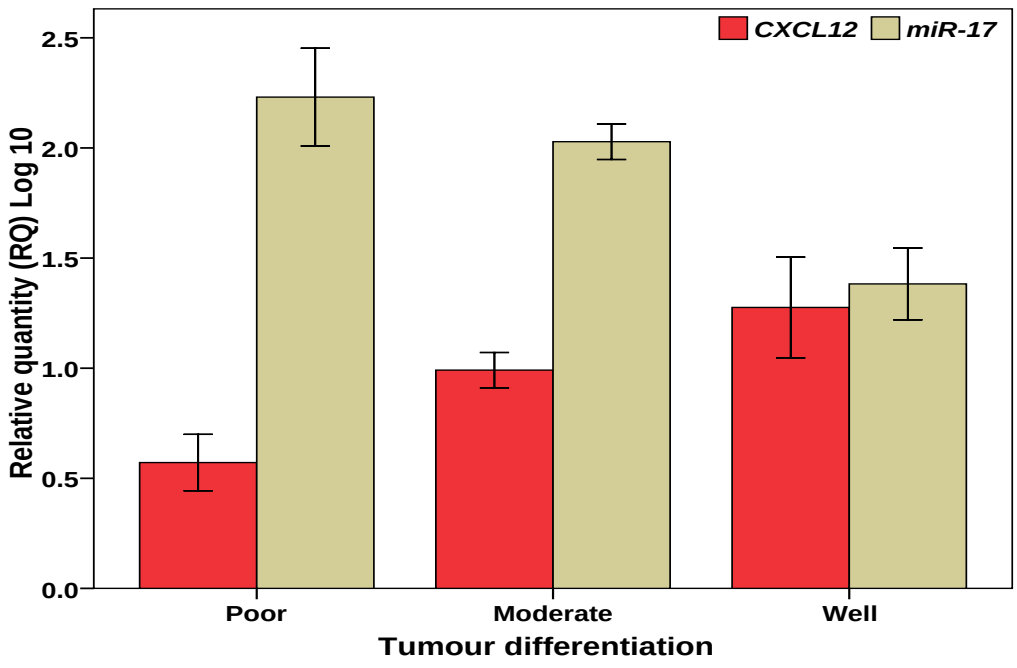


Table 6.4: ***CXCL12* and *miR-17* expression and tumour differentiation**

Variables		***CXCL12* (p=0.019)**	***miR-17* (p=0.002)**
Poor	Well	0.015	0.004
	Moderate	0.080	0.554
Moderate	Well	0.300	0.005

Figure 6.9: *miR-17* and *CXCL12* in association with tumour location
Up-regulation of *CXCL12* (p=0.009) and down-regulation of miR-17 (p<0.001) in rectal compared to colonic cancer.

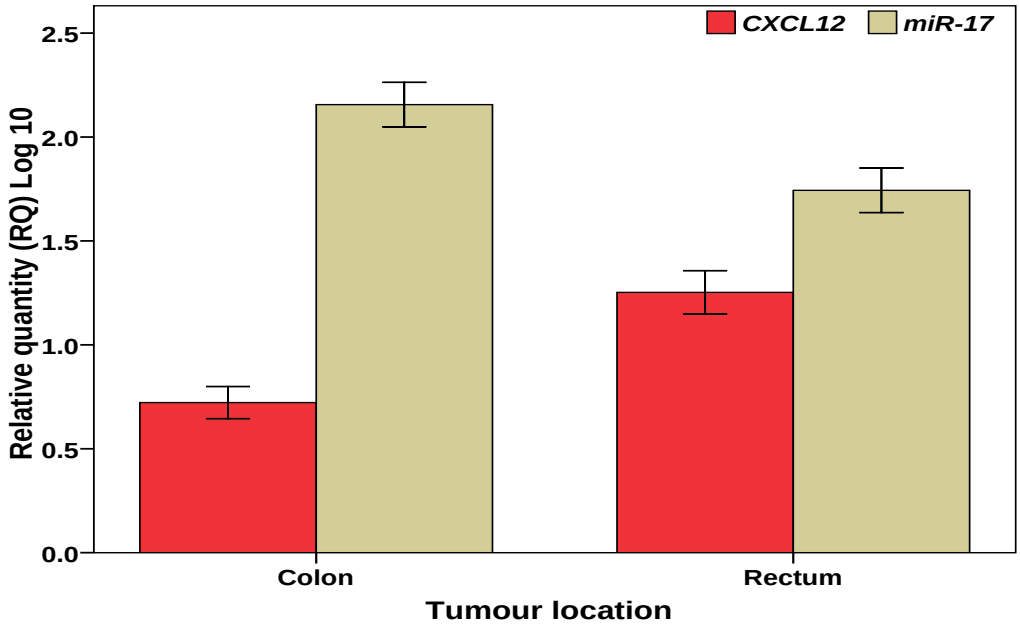
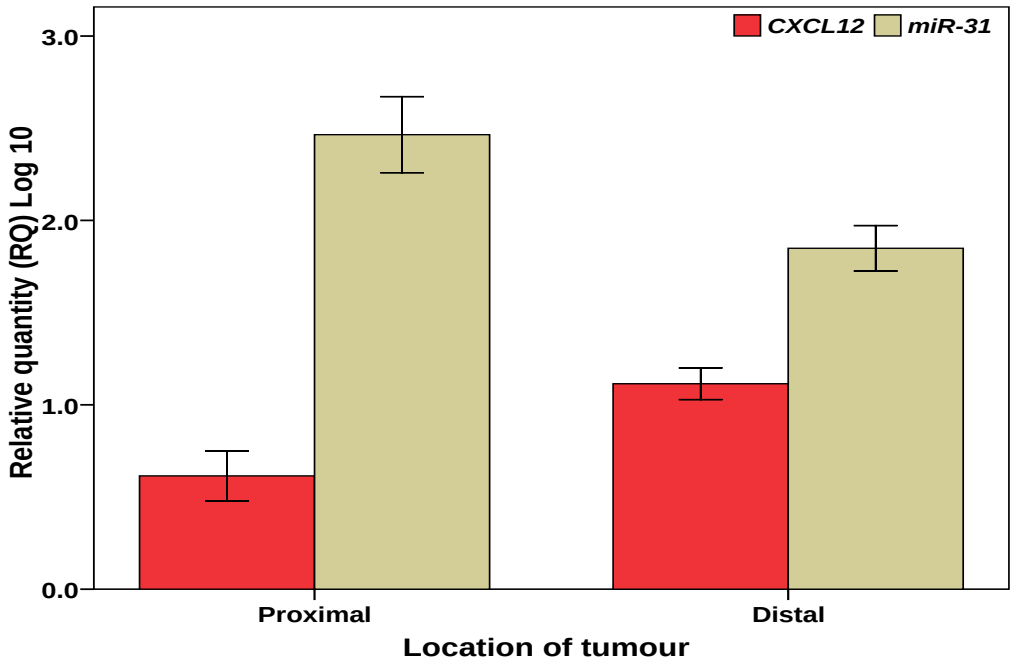


Figure 6.10: *miR-31* and *CXCL12* in association with tumour location
Up-regulation of *CXCL12* (p=0.004) and down-regulation of miR-31 (p=0.018) in distal compared to proximal CRC.



6.5 Discussion

It is known that miRNA are key regulators of gene expression and that these are aberrantly expressed in diverse cancers, including colorectal cancer [248, 272, 540, 561, 617, 618]. It is becoming apparent that miRNAs act as both tumour suppressors and oncogenes in the gene regulatory network and markedly contribute to tumourogenesis. Hence; identification of the miRNA functions may help in understanding cancer pathogenesis, prognosis and response to treatment. Prediction and recognition of miRNA target genes is the first step towards understanding the biology of miRNAs.

Gatt et al. recently indicated that the miRNAs may have more targets than anticipated by conventional prediction methods [619]. In this study, we confirmed the bioinformatics predicted relationship of miR-21/*PDCD4*, miR-31/*CXCL12* and miR-145/*IL8* duplexes. Moreover; and although not computationally predicted, we identified the novel miR-10b/*IL8*, miR-17/*CXCL12*, miR-21/*FABP1*, miR-31/*FABP1* and miR-31/*PDCD4* combinations. These combinations could still represent valid miRNA/target gene duplexes especially if further validated. The interaction of the reported functions of miRNAs and mRNAs, in addition to their reciprocal patterns of expression in tumours and tumour-associated normal tissues and in association to clinicopathological parameters might support the relationship of miRNA/mRNA pairs highlighted by the computational algorithms and correlation analysis.

CXCL12 has been found to play a critical role in tumourogenesis, angiogenesis and tumour cells migration through binding to its *CXCR4* and *CXCR7* receptors [378, 455, 456]. In colorectal cancer, *CXCL12* was reported to be down-regulated and increased with tumour differentiation [392]. Its possible association with MSI and adenoma-carcinoma sequence was discussed in the previous chapters. Moreover; *CXCL12* expression was found to correlate with tumour stage, lymph node status and survival [393, 620]. Although a considerable number of previous studies have investigated the *CXCL12* and its receptors in cancer, the mechanisms by which it exerts its effects are not fully understood. Brand et al. postulated that *CXCL12* activates *ERK-1/2*, *SAPK/JNK* kinases, *AKT* and matrix metalloproteinase-9 which mediate reorganization of actin cytoskeleton resulting

in increase cell migration and invasion [392]. In addition, Yang et al. have demonstrated that stimulation of glioblastoma cells with *CXCL12* contributes to the production of *VEGF* in vitro and thereby synergistically induce tumour angiogenesis [621, 622]. The miRNAs identified to target *CXCL12* in this study will help in further understanding of *CXCL12* role in carcinogenesis.

FABP1 was documented to be involved in several physiological functions including intracellular signalling and cell division and proliferation. Loss of *FABP1* mRNA in colorectal cancer contrasts with the finding in other tumours like gastric cancer, prostatic cancer and hepatocellular carcinoma in which over-expression of the gene was reported [344, 475, 476, 623]. These differences in expression might highlight tumour-specific expression patterns. Expression levels of *FABP1* were found to correlate with survival and grade of colorectal cancer [344, 478]. Lee et al. analysed the expression profiles in the sequence of normal colon crypts, adenoma and early stage carcinoma using cDNA microarray analysis. They identified a group of genes, including *PDCD4* and *FABP1*, to be down regulated in the sequence [477].

PDCD4 is a tumour suppressor gene that inhibits neoplastic transformation, tumour promotion and progression and induces apoptosis in response to different oncogenic factors [497, 499, 500]. It exerts its functions interaction with different other molecules including *eIF4A, eIF4G* and *p21* [416]. Down-regulation of *PDCD4* was showed to lead to increase colon cancer cells invasion[624, 625]. Moreover; its expression levels were found to correlate with poor survival and disease progression in colon and lung cancer [416, 626]. In addition; Mudduluru et al. in their study of 71 cancer and 42 adenoma patients, observed significant reduction in *PDCD4* expression between normal mucosa and adenoma and between adenoma and cancer. The negative regulation of *PDCD4* by *miR-21* was described before by numerous reports. Asangani et al. found that transfection of Colo206f-cells with *miR-21* significantly suppressed a luciferase reporter containing the Pcdc4 3’-UTR, whereas transfection of RKO cells with anti-miR-21 increased its activity. This effect was abolished when a construct with a mutation in the target miR-21-binding site was used instead. Also the anti-miRNA-21 transfected cells showed increase levels of *PDCD4* proteins [269].

miR-21 is one of the most prominent miRNAs implicated in the promotion and progression of human malignancies. It is over-expressed in different tumour types and has been implicated in promotion of growth, proliferation and inhibition of apoptosis [242, 246, 286, 535, 571, 627]. *miR-21* expression has been associated with advanced lymph node and disease stage and tumour invasion and metastasis [571, 628]. Moreover, high expression levels of miR-21 were reported to be associated with disease recurrence, prognosis and therapeutic outcome in colorectal cancer [264, 629]. m*iR-21* is shown to target and down-regulate the expression of tropomyosin 1, *PTEN, SPRY2* and *PDCD4* [245, 269, 286, 630]. Both *miR-17, miR-31*and *miR-21* have been found to be up-regulated in tumours, including colorectal cancer [275, 276, 588, 631-634]. *miR-17,* a member of *miR-17-92* cluster, was reported to be overexpressed during colorectal adenoma to carcinoma progression and induced proliferation of lung cancer cells [635]. Furthermore, *miR-17* activity involves cell anon-autonomous functions that include induction of angiogenesis. Dews et al. demonstrated that the angiogenic activity of *c-MYC* is due, at least in part, to downstream activation of *miR17-92* cluster [284]. On the other hand, *miR-31* overexpression was noted to be associated with advanced tumour stage and local invasion [276]. The role of *miR-31* in cell proliferation was investigated by Liu et al. who found that knockdown of the miRNA repress proliferation of both murine and human lung cancer cell lines [636]. Moreover; *miR-31* was reported to be over-expressed in right-sided colon tumours and associated with microsatellite instability [637].

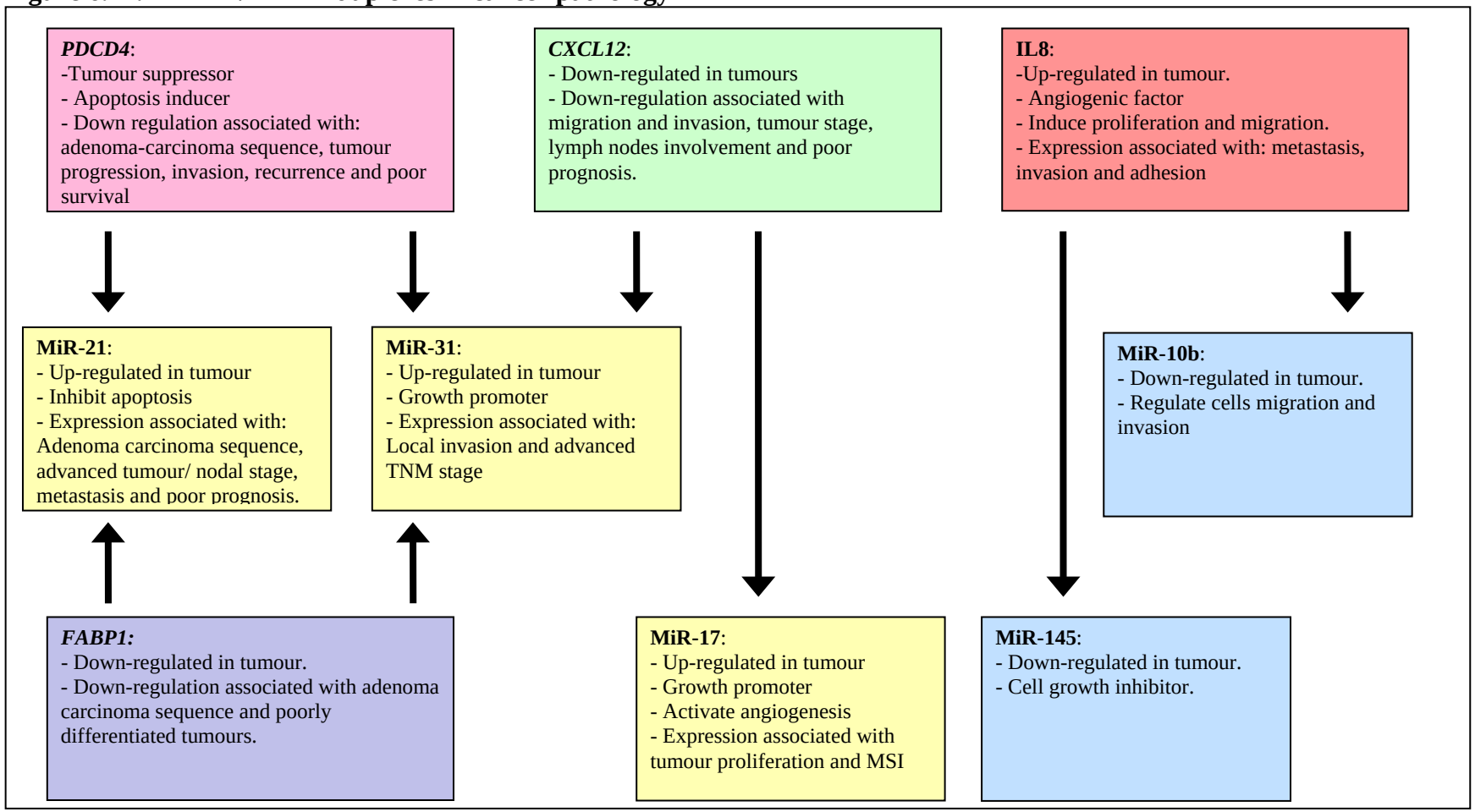
We identified reciprocal pattern of expression of *CXCL12* and *miR-17* and *miR-31* in tumour compared to normal colorectal tissues. Furthermore; the inverse relationship of *CXCL12* and miRNAs was also seen is association with tumour differentiation and tumour location. Dysregulation of *CXCL12* and *miR-31* expression in proximal compared to distal colonic cancer may support their role in MSI tumours. Moreover; inversely related expression levels were noted when comparing the down-regulated miRNAs *miR-21* and *miR-31* to their overexpressed putative targets *PDCD4* and *FABP1*.

Interleukin 8 (*IL8*) has been reported to be overexpressed in cancer and modulate proliferation and migration of tumour cells [406, 638, 639]. Evidences exist that

IL8 is a critical angiogenic factor in a multitude of human cancer. Blocking of the angiogenic activity of *IL8* have proven effective to inhibit angiogenesis, metastasis and tumour progression in murine models [640-642]. The two miRNAs identified in this study to target *IL8*, *miR-10b* and *miR-145*, are reported to be down-regulated in human cancer [242, 263, 272, 274-276, 543]. *miR-145* is a tumour suppressor that inhibits the growth of tumour cells, although the targets are not fully identified. Shi et al. have confirmed targeting of the *IRS-I* 3'-UTR by *miR-145* using a reporter gene expressing the *miR-145* binding sites of the *IRS-1*. They concluded that *miR-145* down-regulate *IRS-I* protein and inhibits the growth of human cancer cells [616]. Moreover, type 1 insulin-like growth factor receptor (*IGF-IR*) is also confirmed as *miR-145* target gene [616, 643]. No significant correlations of *miR-145* with clinicopathological variables were previously identified. Regarding *miR-10b*, the available data is confusing. Although down-regulation of miR-10b was identified in relation to many cancers like colorectal, breast and head and neck aquamous cell carcinoma [242, 263, 543, 644, 645], some other reports described over-expression of miR-10b in tumours and correlate its expression to poor prognostic features like invasion and metastasis [243, 646, 647]. The significance of this apparent paradox is unclear but might highlight tumour-specific expression patterns of *miR-10b*. Our results support the down-regulation opinion of *miR-10b* in colorectal cancer as shown in this and previous chapters. Both *miR-10b* and *mi-145* might target *IL8* and cause its up-regulation and thereby potentiate its angiogenic effect.

Our results and the previous reports, in addition to the negative correlations between miRNA and mRNAs, might support our hypothesis that miRNA/mRNA duplexes identified above represent miRNA/target gene pair. The identified miRNA/ mRNA combinations will not only help in understanding of molecular pathology of colorectal cancer, but may have a potential therapeutic capacity for the disease.

Figure 6.11: miRNA/mRNA duplexes in cancer pathology



Chapter 7:
Mismatch-repair (MMR) protein expression

7.1 Introduction

There are likely to be important clinical indications for determining the molecular subtypes of colorectal cancer. One parameter by which colorectal cancers can be classified involves the expression patterns of Mismatch repair (MMR) proteins. MMR proteins are nuclear enzymes, which participate in repair of base-base mismatch that occur during DNA replication in proliferating cells. The proteins form complexes (heterodimers) that bind to areas of abnormal DNA and initiates its removal. Loss of MMR proteins leads to an accumulation of DNA replication errors, particularly in areas of the genome with short repetitive nucleotide sequences, a phenomenon known as microsatellite instability (MSI) [360, 648, 649]. MSI can be identified in more than 90% of colorectal cancers that arise in patients with Lynch syndrome, while in sporadic colorectal cancer it occurs in 15% of cases [650].

Mechanisms for MSI

Alterations in at least six of the genes that encode proteins involved in the MMR system have been identified in either HNPCC or sporadic colon cancer. These genes include *MSH2, MSH3, MSH6, MLH1, PMS1,* and *PMS2*. Study of the biochemistry of the MMR proteins has revealed that recognition of the base-base mismatches and insertion/deletion loops is performed by a heterodimer of either *MSH2* and *MSH6* or *MSH2* and *MSH3*. Of interest, the *MSH2-MSH3* heterodimer preferentially recognizes insertion/ deletion loops and thus cannot compensate for loss of *MSH6*. Consequently, cancers arising with a loss of *MSH6* function display microsatellite instability only in mononucleotide repeats [651]. The *MLH1*, *PMS2*, and *PMS1* proteins appear to operate primarily in performing the repair of the base-base mismatches and insertion/deletion loops. A heterodimer of *MLH1-PMS2* operates as a molecular matchmaker and is involved in executing the repair of the mismatches in conjunction with other molecules [651, 652].

HNPCC related colon cancers account for 3–6% of all colon cancers, and germline mutations in *MSH2* and *MLH1* have been found in 45–70% of families that meet the Amsterdam criteria for HNPCC [653, 654]. Since inactivation of both alleles of *MSH2* or *MLH1* is required to generate MSI, the cancers that arise in HNPCC kindred frequently show loss of heterozygosity at the loci of these

genes, or alternatively show somatic mutation of the sole wild-type MMR allele. The germline mutations that occur in *MSH2* and *MLH1* are widely distributed throughout either gene and are missense, deletion, or insertion mutations. These mutations result in frame shifts (60% of h*MSH2* mutations and 40% of MLH1 mutations), premature truncations (23% of *MSH2* mutations), or missense mutations (31% of *MLH1* mutations) [655]. The lack of a mutation hotspot has hampered the development of an inexpensive clinical assay to detect germline mutations in the genes known to cause HNPCC. Furthermore, because one wild-type allele is sufficient to maintain MMR activity, functional assays to detect MMR gene mutation carriers have not been developed for clinical use to date. However, proof-of-principle studies have demonstrated that it may be possible to develop such an assay by forcing a cell to a haploid state in which case a mutant MMR allele could be detected [656, 657]. Studies of the 15% of sporadic colon cancers that display MSI demonstrated these arose due to somatic inactivation of MMR genes and not due to germline MMR gene mutations with low penetrance. While occasional somatic mutations of *MSH2* and *MLH1* were detected , the predominant mechanism for inactivating MMR unexpectedly proved to be the epigenetic silencing of the *MLH1* promoter due to aberrant promoter methylation [98, 99].

Clinical implications of MSI

The CRC microsatellite profile provides useful prognostic information [138, 658], showing the patients with microsatellite unstable neoplasms have a better overall survival rate and a modified response to conventional chemotherapy [161, 659-663]. MSI also helps in predicting the treatment response of CRC [161, 661, 664], and could modify the chemotherapy protocols offered to the patients in the future [161], but these results should be applied with caution before this predictive tool is verified.

Molecular markers as predictive factors in treatment decisions have been developed in the last few years. The initial studies in sporadic CRC showed that the retention of heterozygosity at one or more 17p or 18q alleles in microsatellite-stable CRCs and mutation of the gene for the type II receptor for TGF-β1 in CRCs with high levels of microsatellite instability correlated with a favorable outcome

after adjuvant chemotherapy with fluorouracil based regimens, especially for stage III CRC [661, 664]. However, most recent studies have revealed that fluorouracil-based adjuvant chemotherapy benefited patients with stage II or stage III CRC with MSS tumors or tumors exhibiting low frequency MSI but not those with CRCs exhibiting high frequency MSI [161]. The reasons for these responses must be related to the distinctive cell kinetics associated with MMR down-regulation (significantly increased apoptosis and decreased proliferation), which can certainly contribute to tumor cell resistance to conventional chemotherapy.

Testing for MSI and MMR defects:

Clinical Criteria*:*

The recognition that certain types of cancers cluster in families with HNPCC and that cancer develops at relatively early ages compared with the general population provided the rationale for development of criteria that could be used to aid in the diagnosis. Two sets of criteria (the Amsterdam criteria and Bethesda guidelines) developed by a consensus of experts, have been most widely accepted and best studied.

The Amsterdam criteria (table 1.5) were designed to establish the diagnosis of HNPCC based upon familial clustering of HNPCC-related tumors. On the other hand, Bethesda guidelines (table 1.6) were designed to help predict which patients with colorectal cancer are likely to have a mismatch-repair mutation and should thus undergo further testing. However, both the Amsterdam criteria and Bethesda guidelines have been studied for predicting the presence of mismatch repair mutations. Although the Bethesda guidelines and Amsterdam criteria continue to be used widely, several studies evaluating them (both the original and revised) have underscored the limitations of their accuracy in predicting the presence of mismatch repair mutations [665-668], and review of the literature reported that the sensitivity of the original Amsterdam criteria ranged from 54 to 91% [669]. Such a wide range of estimates leaves substantial uncertainty as to the role of the Amsterdam criteria as a screening test for mismatch repair mutations. In addition to the limitations regarding their predictive accuracy, there are practical problems with policies based on the implementation of these clinical criteria. Patients' report of the family history may not be accurate, particularly for cancers other

than colorectal that are potentially related to HNPCC.[670]. Issues of uncertain paternity may also be relevant in some families while some families may be too small or have insufficient contact among family members to obtain a clinically meaningful family history.

Table 7.1: **Amsterdam criteria**

Original (Amsterdam I) [151]	**Revised (Amsterdam II)** [150]
- At least 3 relatives with colorectal cancer, one of whom must be a first degree relative of the other two - Involvement of 2 or more generations - At least 1 case diagnosed before age 50 - Familial adenomatous polyposis has been excluded	- At least 3 relatives with HNPCC-associated cancer - One should be 1st degree relative of other two - At least 2 successive generations affected - At least 1 diagnosed before age 50 - Familial adenomatous polyposis excluded - Tumors should be verified by pathologic examination

Table 7.2: **Bethesda guidelines** [152]

Original	Revised
- Individuals with cancer in families that meet the Amsterdam criteria - Patients with two HNPCC-related cancers, including synchronous and metachronous colorectal cancer or associated extracolonic cancers - Patients with colorectal cancer and a first-degree relative with colorectal cancer and/or HNPCC-related extracolonic cancer and/or a colorectal adenoma with one of the cancers diagnosed before age 45 years, and the adenoma diagnosed before age 40 years. - Patients with right-sided colorectal cancer having an undifferentiated pattern on histopathologic diagnosis before age 45 years. - Patients with signet-ring cell type colorectal cancer diagnosed before age 45. - Patients with adenomas diagnosed before age 40.	- Colorectal cancer (CRC) diagnosed in a patient <50 - Presence of synchronous, metachronous colorectal or other HNPCC-associated tumors regardless of age - CRC with the MSI-H-like histology diagnosed in a patient less than 60 - CRC diagnosed in a patient with one or more 1st degree relatives with an HNPCC related tumor, with one of the cancers being diagnosed under age 50 - CRC in a patient with two or more 1st or 2nd degree relatives with HNPCC-related tumors, regardless of age

Clinical testing for MSI and MMR:

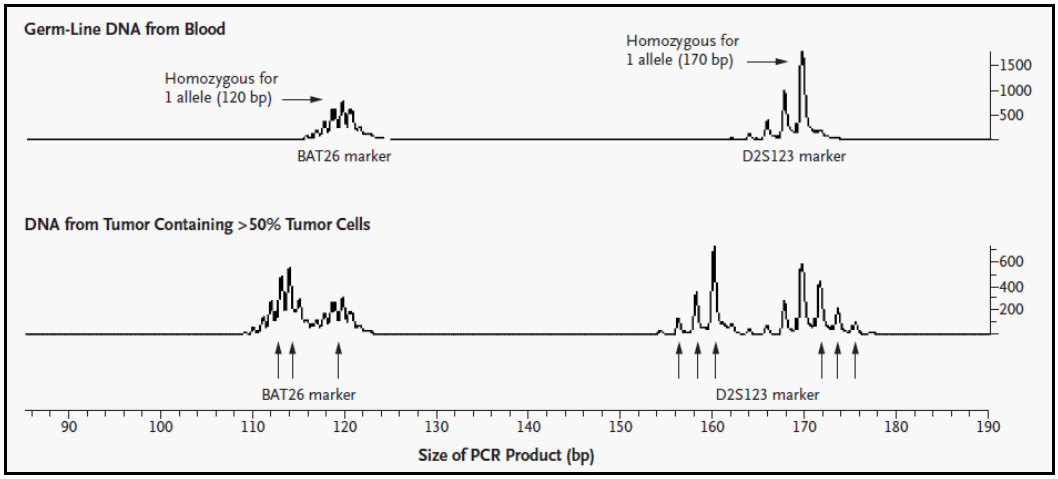
Because of the limitations of relying on clinical criteria to guide testing, some authorities have proposed that tumors from patients with colorectal cancer be evaluated for markers of HNPCC regardless of the family history [671, 672]. One of the largest studies evaluating this approach included 1066 patients with colorectal cancer whose tumors were tested for MSI [671]. Patients with suggestive MSI results were tested for germ-line mutations in the mismatch repair genes (*MSH2*, *MLH1*, *MSH6*, and *PMS2*) by IHC, genomic sequencing, and deletion studies. A mutation causing HNPCC was detected in 23 patients (2.2 percent) of whom ten were older than 50 and five did not meet the Amsterdam criteria or Bethesda guidelines. These data suggest that the Amsterdam or Bethesda criteria alone may miss as many as 22 percent of patients with HNPCC. However, only five additional individuals from the cohort of 1066 subjects (0.5%)

would have been identified by routine molecular analysis of all colon cancers fulfilling the Bethesda criteria, making such an approach impractically expensive for routine clinical use. Therefore; most expert guidelines on HNPCC suggest a combination of sequential laboratory testing in patients who fulfill the Amsterdam criteria or Bethesda guidelines to minimize costs and maximize test accuracy [673, 674]. Approaches based on such a strategy have been considered to be cost-effective [675]. However, the exact methods and order of testing are unsettled. Proposed strategies include initial testing of tumors for MSI with or without IHC for loss or expression of mismatch repair proteins, with germline gene sequencing reserved for patients with suggestive results.

Microsatellite instability (MSI) testing:

MSI testing involves amplification of a standardized panel of DNA markers; five markers were agreed upon by a consensus panel convened by the National Institutes of Health in 1997 [138]. The reference panel included two mononucleotide markers (BAT25 and BAT26) and three dinucleotide microsatellites (D5S346, D2S123 and D17S250), previously tested by Fishel [676], plus a list of several alternative loci. Three categories of MSI have been recognized based upon these panels: MSI-high (instability of two or more markers), MSI-low (instability of one marker), and MS-stable (no instability). More recently, some laboratories have begun using ten or more markers. In such cases MSI is defined as stable when fewer than 10% of markers are unstable, low when 10 to 30% of markers are unstable and high when greater than 30-40% of markers are unstable. There are several pitfalls of MSI testing. First, it is labor intensive, relatively costly, and requires expert pathologic services. In addition, tissue to be amplified should ideally be microdissected to avoid amplifying DNA from normal colonic mucosa.

Figure7.1: **Detection of Microsatellite Instability using Fluorescent Labeling PCR.** Two markers are analyzed in the same track: the mononucleotide repeat marker BAT26 is shown on the left, and the dinucleotide marker D2S123 is shown on the right. The upper tracing is from germ-line DNA from blood. The lower tracing is from DNA extracted from a histologic section of a tumor containing more than 50 percent tumor cells. For marker BAT26, germ-line DNA shows a single peak, indicating that the patient is homozygous for this marker (arrow). Tumor DNA shows, in addition to the normal allele (single arrow), a new allele (double arrows) that has lost approximately five nucleotides. This constitutes microsatellite instability. For marker D2S123, germ-line DNA is homozygous, whereas tumor DNA shows two new alleles (triple arrows), one with a loss of approximately 10 nucleotides (left) and one with a gain of 2 nucleotides (right). Thus, the tumor shows microsatellite instability with both markers



From: Lynch et al. Hereditary colorectal cancer. *NEJM* [155]

Immunohistochemistry (IHC) testing:

Pathogenic mutations in MMR proteins usually lead to the absence of a detectable gene product providing the rational for immunohistochemistry testing to determine loss of expression. Tumours from patients suspected to have MSI can be stained for MMR proteins and the surrounding normal tissues can be used as a positive control. IHC has an advantage over MSI analysis as it is much easier to perform and less expensive. Moreover, it provides gene specific information to direct further genetic analysis. However; the technique is vulnerable to the quality of tissue preparation, staining and interpretation.

The understanding of how the MMR proteins interact during DNA repair can help in the interpretation of the results of such testing. MSH2 forms a heterodimer with

MSH6, while MLH1 binds to PMS2 and complexes MSH2/MSH6 heterodimer. Therefore, when MSH6 is not detected in a tumour MSH6 may also not detected. The situation is more complex with lack of MLH1 expression. Hypermethylation of h*MLH1* gene, which is common in sporadic colorectal cancer, may lead to loss of protein expression.

IHC has a role in detecting MMR defects, with data suggesting that the effectiveness of IHC screening of the MMR proteins would be similar to that of the more complex strategy of microsatellite genotyping [671, 677] . This technique can guide which gene to sequence and can help differentiating sporadic from hereditary mutations: MSH2 loss is likely to be HNPCC, whereas MLH1 loss could be HNPCC or sporadic CRC (*MLH1* promoter methylation). MMR proteins heterodimerize to function; the MSH2 loss almost always accompanies MSH6 loss and when MLH1 is lost, generally so is hPMS2 [678, 679]. In addition, IHC can miss functional loss; i.e. presence of the protein with antigen positivity in the absence of function.

MMR IHC studies are based on a complete absence of at least one MMR protein [113, 677, 680-682]. But these studies do not consider the immunostaining topographic heterogeneity. Since the MMR proteins function as heterodimers, it could be advocated to validate the IHC results of MSH2/MSH6 and MLH1/PMS2. More studies are required to clarify the influence of this predictable tumor heterogeneity to select the appropriate sample for immunohistochemical and/or MSI analyses

*Genetic Testing***:**

Multiple methods have been used for genetic testing in HNPCC. The methods used should ideally be able to detect the many potential genotypes associated with HNPCC like nonsense, missense, and frame shift mutations, genomic deletions, duplications, and rearrangements. The commonly used tests includes: high output screening techniques, DNA sequencing, conversion analysis and methods to detect large structural DNA abnormalities like Southern blot and Multiplex ligation-dependent probe amplification.

7.2 Aims

Information about MMR protein status in colorectal cancer is important because it will identify those most likely to have Lynch syndrome and those most likely to have microsatellite instability in their tumours which has been proven to have better prognosis and may affect their treatment regimens in the future. We undertook this study to develop and optimise a protocol for MMR protein immunohistochemistry testing in colorectal cancer. We also aimed to analyse the proportion of patients with colorectal cancer with loss of immunostaining for MMR proteins (hMLH1, hMPS2, hMSH2 and hMSH6) in order to determine the feasibility of molecular screening for the loss of MMR proteins through the study of unselected patients with colorectal cancer.

7.3 Materials and methods

7.3.1 Study group

A group of 33 patients with colorectal cancer was randomLy selected from the department of surgery bio-bank to determine the expression of MMR proteins in their FFPE tumour tissues using immunohistochemistry techniques. The age of the patients at diagnosis of their cancers and their family history were collected by reviewing the medical charts.

7.3.2 FFPE tissues

Tumour tissues collected at time of surgery were collected and placed in 10% formalin (Lennox) for fixation at room temperature until embedding for a minimum of 24 hours. Tissue was then removed from the formalin and placed on an open cassette. The cassette was closed and placed in 250 mL of Industrial Methylated Spirit (VWR) to wash the formalin from the tissue. Then, the cassette was removed and placed in JFC solution (Milestone) filed JFC beaker and placed in the histoprocessor (MicroMED) for 60 minutes (70°C). Thereafter, the cassette was transferred to the paraffin wax (VWR) filled wax beaker and placed in the histoprocessor (MicroMED) for 30 minutes. The cassette was removed from the wax beaker and tissue was blocked out carefully. The blocks were left at 4°C until hard and then stored at fridge or room temperature until sectioning. Sectioning of formalin-fixed paraffin-embedded tissues was carried out using Slee microtome (LIS Ltd). With section thickness set to 30µM the block was pared down until even sections were being cut and the outer layer of wax was removed. Then the section thickness was adjusted to 5 µM. The sections were then placed in a floating out bath to stretch it out, before being placed onto a Superfrost plus (positive charged) slides (VWR). The slides were allowed to air-dry overnight at room temperature and then stored at 4°C until further use. Before enrolment in any further experiments each slide is stained in H & E and reviewed by a pathologist to determine the quality of the block and the percentage of tumour tissues in the section (should be >50%)

7.3.3 Immunohistochemistry

Immunostaining was carried out on 5 µm thick paraffin sections of tumour tissue from each patient, using mouse monoclonal antibodies specific for each of the

four human MMR proteins and employing automated DABMap system (Ventana) for hMSH6 detection and UltraMap system (Ventana) to detect *hMLH1*, *hMSH2*, and hPMS2 proteins.

DABMap protocol:

It was consist of deparaffinization and cell conditioning, followed by addition of primary antibody and incubation at room temperature for I hour. Then the secondary antibody was added before counterstaining with haematoxylin and slides dehydration.

UltraMap protocol:

The standard UltraMap was used to detect hMSH2. It was again consist of deparrafinization and cell conditioning followed by primary antibody titration. The tissue section was incubated with primary antibody for 12 hours at 37°C. No secondary antibody was added. This was followed by counterstaining and dehydration in serial ethanol alcohol dilution and Xylene (Sigma).

The extended UltraMap protocol was used to determine the expression of hMLH1 and hPMS2.It was different from the standard one in that the cell conditioning was extended to three cycles of medium cell conditioner and cell conditioner compared to two cycles in case standard protocol.

7.3.4 IHC analysis

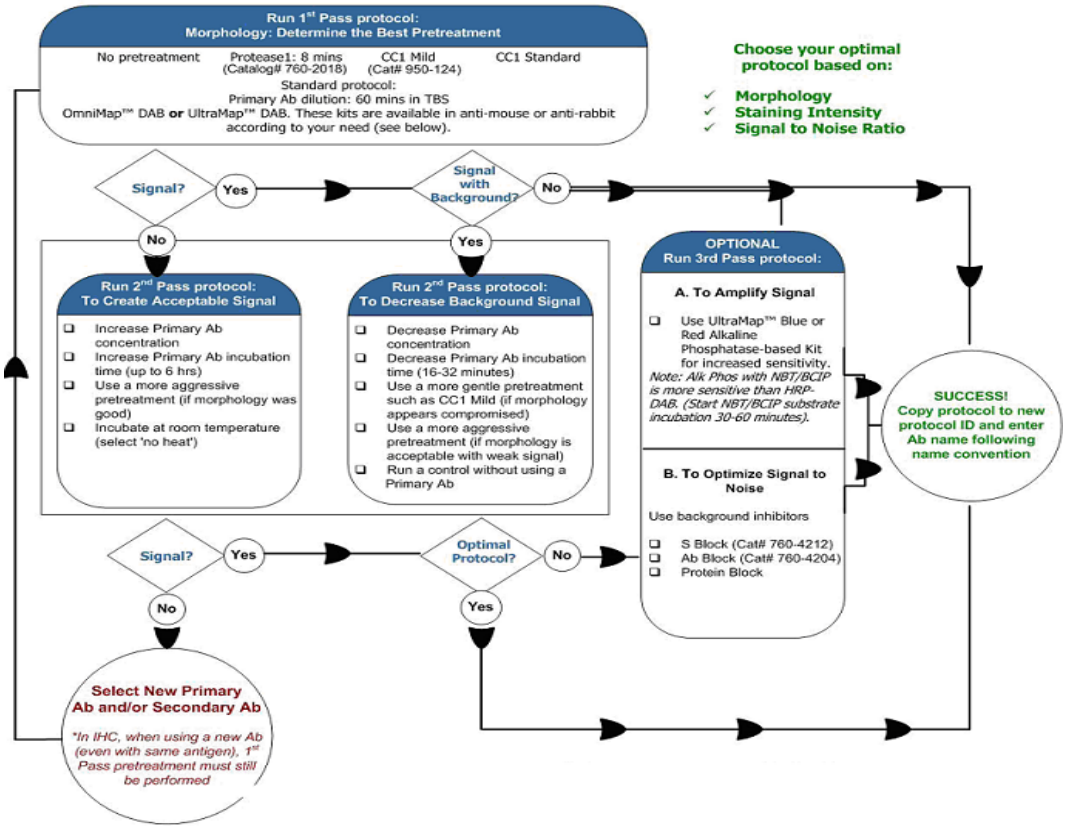
Changes in protein expression following transfection of colorectal tissues were observed in stained cells using Olympus BX60 microscope and image analySIS software. Adjacent normal tissue served as an internal control for positive staining and a negative control staining was carried out without the primary antibody. MMR protein staining was considered negative when all of the tumour cell nuclei failed to react with the antibody.

7.4 Results

7.4.1 Optimization of MMR protein staining protocol

Tissue processing has the greatest single impact on the end result of IHC and different tissue types often require slightly different pre-treatments for optimum results. To optimized staining protocols we employed the Closed Loop Assay Development (CLAD) for IHC (figure 3.1).

Figure 7.2: Closed Loop Assay Development (CLAD)



Optimal staining was achieved for hMSH6 using DABMap system, however; acceptable stating for hMLH1, hMSH2 and hMPS2 was only achievable using UltarMap system.

7.4.2 MMR protein expression

IHC staining was performed on 33 colorectal cancer tissue specimens. Loss of MMR protein is defined as complete absence of nuclear staining within the tumour. While MMR proteins expression is defined as the presence of nuclear

staining within the tumour regardless its intensity or the number of positive nuclei (figures 3.2, 3.3, 3.4, 3.5)

Of the tissue specimens in which acceptable immunostaining was achieved, three samples showed loss of one or more of the MMR proteins (table 3.1). Both hMLH1 and hPMS2 proteins were not expressed in a 36 years old woman (case 3) with cancer of the caecum (Proximal to the splenic flexure). She had history of breast cancer on her mother and colorectal cancer on one of her grandfathers (undocumented weather on paternal or maternal side).The expression of hMSH6 protein was undetermined in tumour tissues retrieved from a 61 years old man (case 13) with cancer of the proximal colon (proximal to the splenic flexure). He had no documented family history of cancer. The third case was a 77 years old man (case 27), again with no documented family history of cancer, who had carcinoma of the rectum. He showed loss of *hMLH1* expression in the tumour tissues.

Figure 7.3: **Pedigree of case 3**
The index case was 38 years old when diagnosed with caecal cancer. One of her grandfathers was diagnosed with colorectal cancer (weather paternal or maternal side, site of tumour and age at diagnosis were not documented). Her mother died of breast cancer (age was not documented). One of her paternal cousin was diagnosed with breast cancer; also age at onset was no documented.

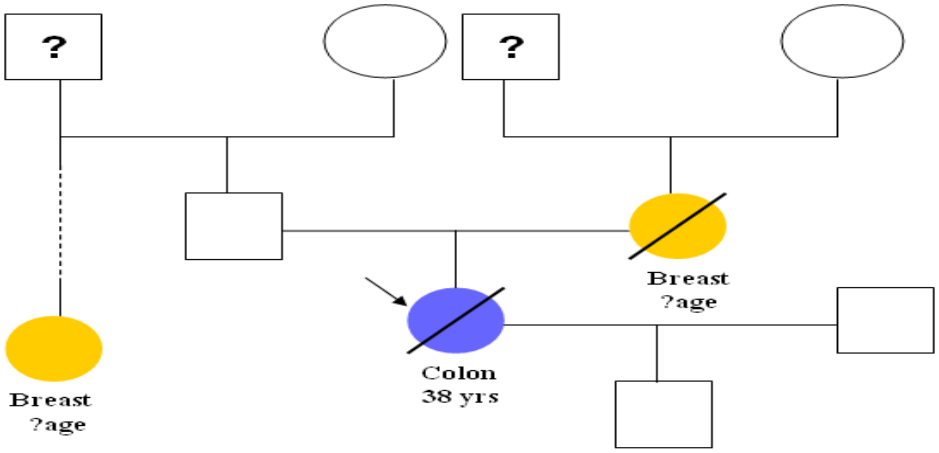


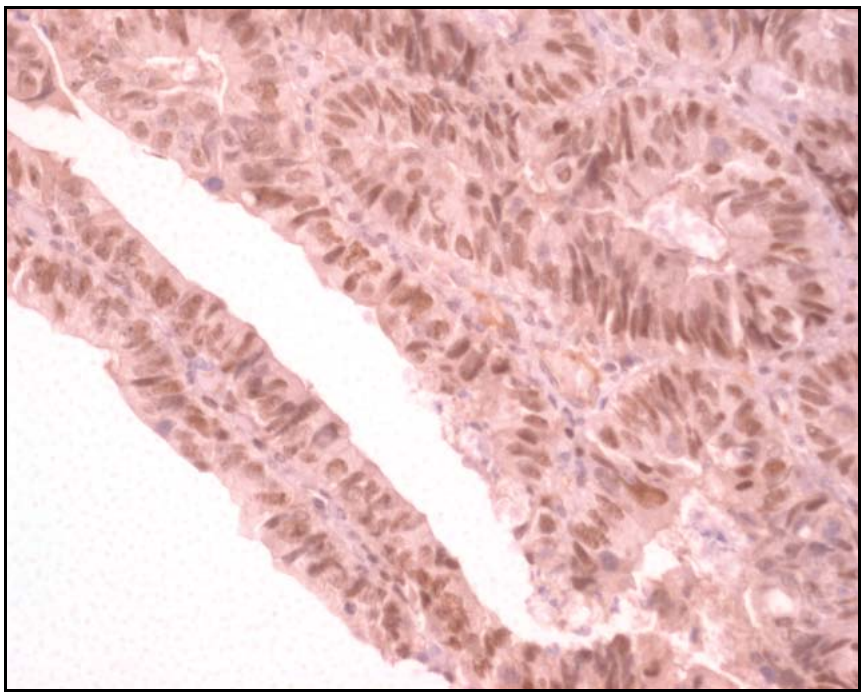
Table 7.3: **Characteristics and MMR protein status of the study cohort**

No	Specimen No	Age	Tumour location	Family history of CRC	Extra-colonic tumour	*Lost Protein*
1	T08-1102	60	D colon	No		
2	T07-2256	82	P colon	Brother		
3	T07-2240	36	P colon	Grandfather	Breast, mother & cousin	hMLH1& hPMS2
4	T07-2244	76	D colon	No		
5	T08-0655	82	D colon	No		
6	T08-0907	70	D colon	No	Lung, brother	
7	T08-1055	90	P colon	No		
8	T08-1167	71	D colon	Unknown		
9	T08-0505	48	P colon	No		
10	T08-0143	60	D colon	Brother		
11	T08-0418	79	P colon	No		
12	T08-0096	75	P colon	Cousin		
13	T08-0727	61	P colon	No		hMSH6
14	T08-0713	78	D colon	No		
15	T08-0534	81	P colon	No		
16	T08-0144	63	D colon	No		
17	T08-0594	77	D colon	5 siblings		
18	T08-700	77	D colon	No		
19	T08-1095	76	D colon	No		
20	T08-1056	90	P colon	No		
21	T08-0413	46	D colon	No		
22	T08-0732	86	Rectal	No	Lung, patient & brother	
23	T08-0615	77	Rectal	No		
24	T09-0060	73	Rectal	No		
25	T07-2238	50	Rectal	No		
26	T08-0605	88	Rectal	No		
27	T08-0285	77	Rectal	No		hMLH1
28	T08-0146	74	Rectal	Father & brother		
29	T08-0138	68	Rectal	No		
30	T08-0142	78	Rectal	No	Prostate, patient	
31	T08-0273	78	Rectal	Mother		
32	T08-0279	81	Rectal	No		
33	T08-0299	94	Rectal	No		

P colon = Proximal to splenic flexure, D colon = Distal to splenic flexure

Figure 7.4: hMLH1 expression
Immunohistochemical staining of tumours expressing hMLH1 (A) or lacking the expression of hMLH1 (B). The nuclei stained brown in hMLH1 positive tumours, while taking the blue colour of haematoxylin in hMLH1 negative tumours

(A)



(B)

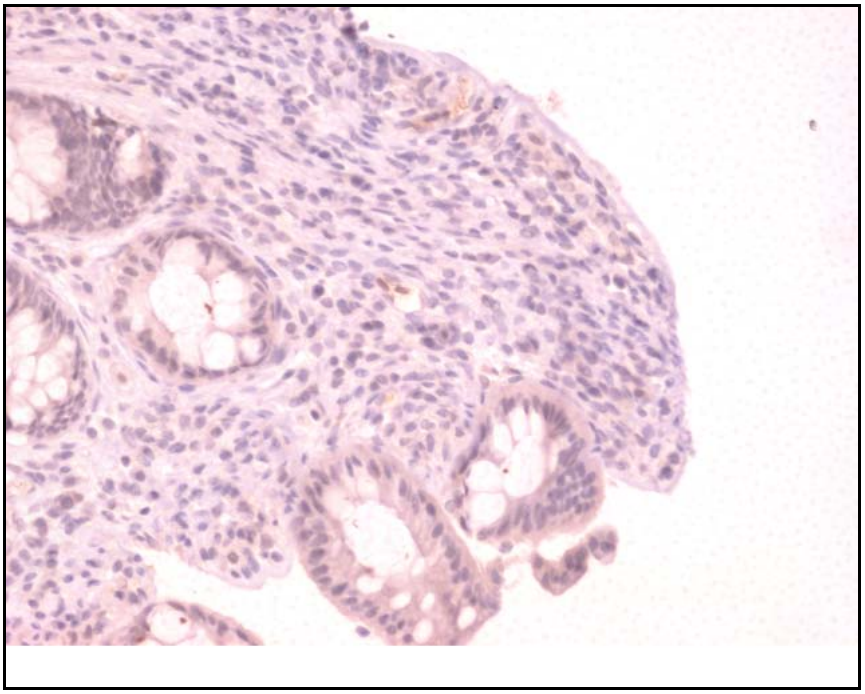
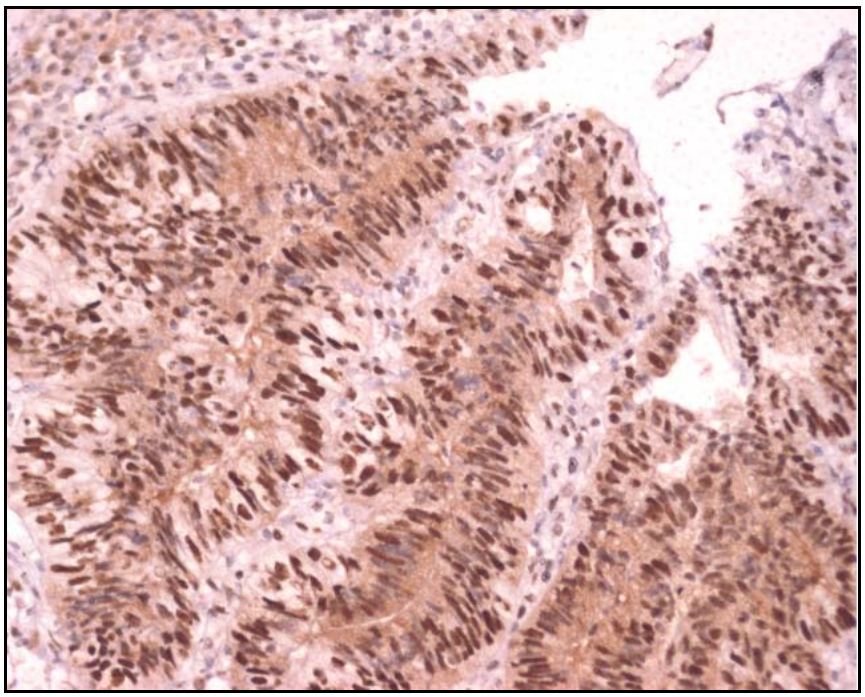
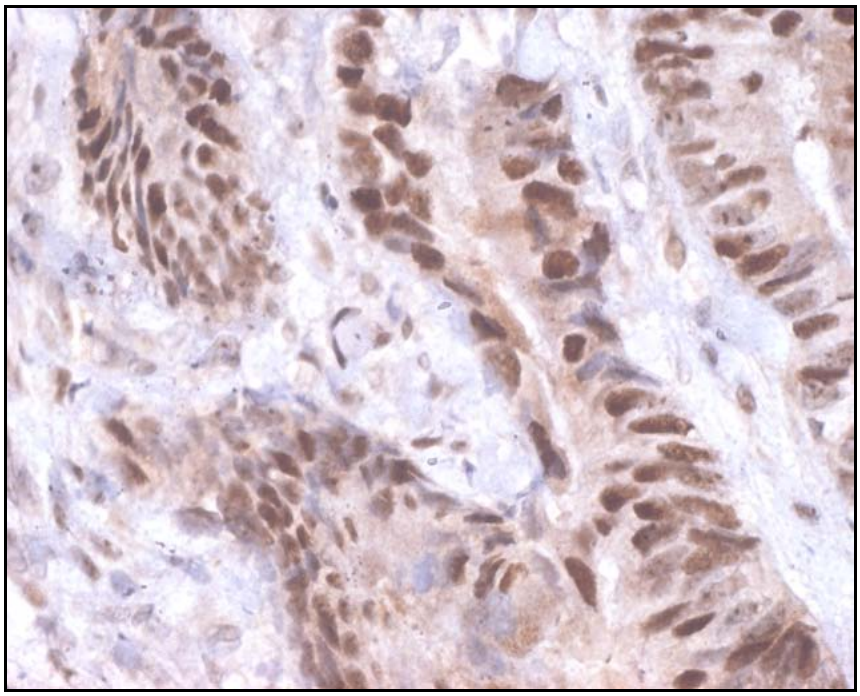


Figure 7.5: hMSH6 expression
Immunohistochemical staining of tumours expressing hMSH6 **(**(A) ×20 and (B) ×40**)** or lacking the expression of hMSH6 (C). The nuclei stained brown in hMSH6 positive tumours, while taking the blue colour of haematoxylin in hMSH6 negative tumours

(A)



(B)



(C)

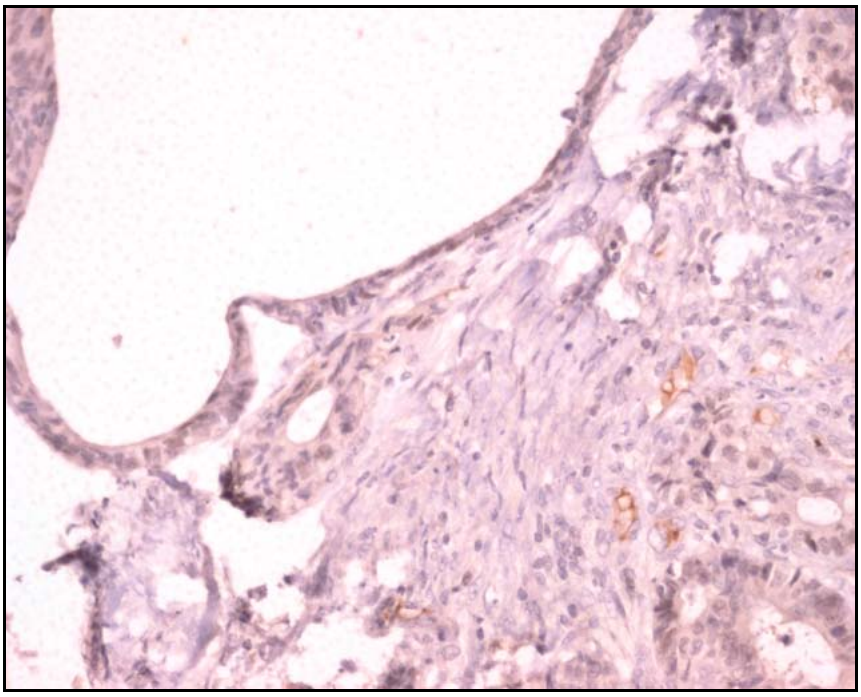


Figure 7.6: hMSH2 expression
Immunohistochemical staining of tumours expressing hMSH2. The nuclei stained brown in hMSH2 positive tumours, while taking the blue colour of haematoxylin in hMSH2 negative tumours.

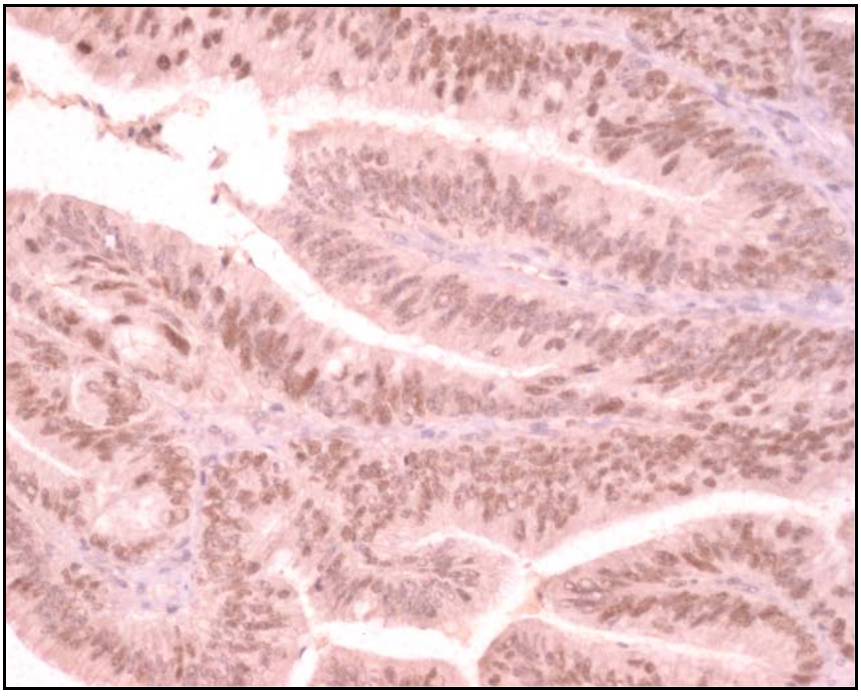
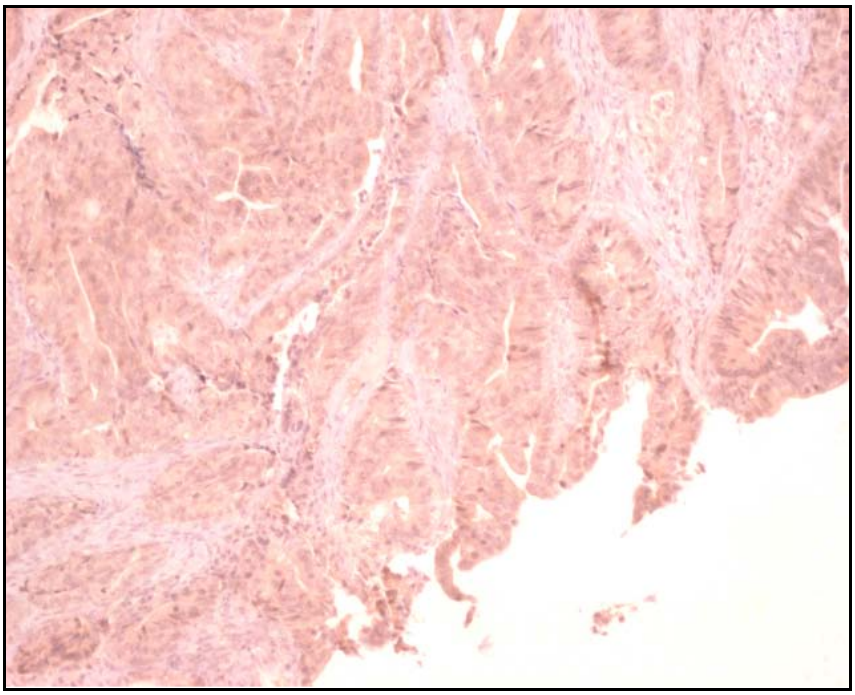
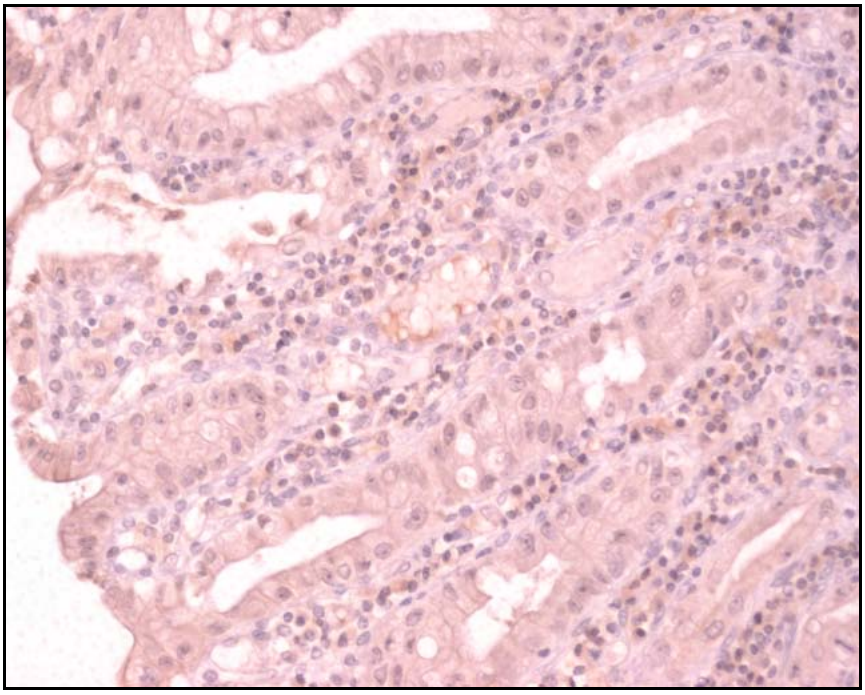


Figure 7.7: hPMS2 expression
Immunohistochemical staining of tumours expressing hPMS2 (A) or lacking the expression of hPMS2 (B). The nuclei stained brown in hPMS2 positive tumours, while taking the blue colour of haematoxylin in hPMS2 negative tumours

(A)



(B)



7.5 Discussion

The identification of HNPCC can be lifesaving as it can lead to early detection of cancer. Jarvinen et al. in a controlled clinical trial extending over 15 year period concluded that screening for colorectal cancer in HNPCC families more than halves the risk of colorectal cancer, prevents deaths from colorectal cancer and decreases the overall mortality rate by about 65% [683]. Furthermore; the cost-effectiveness of screening was quantified by Ramsey et al. as $7,556 per year of life gained [675].When clinical and pedigree criteria such as Amsterdam criteria are used to determine what proportion of all colorectal cancers are due to HNPCC, estimate range from 1-6% [155]. However; molecular screening has suggested that more 3% of all such patients have HNPCC. Moreover, the mean age at presentation with HNPCC diagnosed by molecular screening was 54 years old in a study included several patients over 60 years of age [154, 684].

In addition, experiments have recently shown the differences in the response of MSI-H tumours to chemotherapeutic agents. DNA mismatched repair-deficient cells are resistant to the alkylating agents (e.g. melphalan and busulphan), methylating agents (e.g. temozolomide), the platinum-containing agents (e.g. cisplatin and carboplatin), antimetabolites (e.g. fluorouracil and thioguanine) and topoisomerase inhibitors (e.g. doxorubicin) [685, 686]. The clinical significance of these observations remained unclear till recently. A meta-analysis of 32 studies with 7642 cases found the hazard ratio (HR) for overall survival in patients whose tumours have high microsatellite instability (MSI-H) is 0.65 (95% CI= 0.59-0.71). Two studies, in this review, have assessed the benefit of 5-fluorouracil (5-FU) in stage II and III colorectal cancer patients by MSI status. The analysed data indicates that patients without MSI benefited significantly from 5-FU (HR=0.72, 95% CI= 0.61-0.84), while patients with MSI did not benefit from 5-FU (HR=1.24, 95% CI=0.72-2.14) [157].

Because of the limitations of relying on clinical criteria to guide testing for Lynch syndrome and the prognostic information that could be provided by MSI status, molecular screening of all patients with colorectal cancer for MMR protein expression is now both feasible and desirable. In most Lynch syndrome colorectal tumours, MSI has been shown to result from defects in DNA

mismatch repair mechanism [687]. Mutations in *hMLH1* or *hMSH2* genes are the most common defects in these families making up to 94% of the germ line mutations detected. In addition, a few families have been found to have *hMSH6* or *hPMS2* mutations [76, 655]. On the other hand, about 10-15% of sporadic colorectal cancer also exhibit MSI, and loss of one or more of the MMR proteins has been found in these tumours [658, 688]. Lack of expression of hMLH1 as the result of promoter methylation occurs in most of sporadic MSI-positive tumours [97]. Loss of the other MMR proteins is rare in sporadic tumours and in one study loss of either hMSH2 or hPMS2 was found in only 2% of tumours [689].

The major laboratory tests used in the evaluation of patients suspected to have Lynch syndrome include testing of tumour tissues using immunohistochemistry (IHC), MSI testing or germ line testing for mismatch defects. IHC has the advantage over the other methods, as the primary screening method, since it is less demanding to perform and is available as part of routine services in general pathology laboratories. In addition, IHC will determine which protein is affected and provides gene specific information; thereby direct the genetic analysis rather than performing exhausting, time and material consuming unnecessary tests. Nevertheless, while most of mutations will results in total loss of the protein expression , in some cases mutations only result in loss of function rather than the expression of the protein which will still be detectable by IHC.

Many studies have provided information about the sensitivity and specificity of IHC for predicting MMR mutations [665, 677, 690-696]. A recent meta-analysis determined the sensitivity to range from 27%-100% and specificity from 43%-100%, however, analysis of good quality studies only had a summary sensitivity of 74% (955 CI: 54-87) and specificity of 77% (955 CI: 61-88) [697]. In one study of unselected 131 colorectal cancer patients diagnosed younger than 45 years of age the sensitivity of IHC testing for the main 4 MMR proteins was reported as 100% and its specificity was 69% [696]. Lindoe et al. have assessed 1144 patients with colorectal cancer for MMR deficiency by MSI testing and IHC detection for hMLH1 and hMSH2. They determined 92% specificity and 100% specificity of IHC for screening for MMR defects [698].

In evaluating the expression of MMR proteins using IHC, any tumour cell nuclear expression is considered positive due to the heterogeneity of expression and difficulties in test standardisation [136]. The intensity of staining in normal mucosa decreased towards the surface. Moreover, the normal enterocytes can serve as positive internal controls and should always be observed to determine the quality of staining [699]. In sporadic tumours due to hypermethylation of the promoter of *hMLH1* there is consistent loss of the protein expression [700]. Therefore, this feature alone can not differentiate sporadic MSI-H tumours from Lynch syndrome due to germLine mutation in *hMLH1* (approximately half of the cases) and methylation analysis would more help in the determination of the nature of mutation.

In this study, we looked the MMR protein expression without considering the family history or the result of previous tumour testing for microsatellite status in a prospective of newly diagnosed colorectal cancer patients. We identified three patients with loss of one or more MMR protein. The first patient (case 3) was less than 40 years old when diagnosed with caecal cancer. Although her family history was not fully documented (figure 3.2), she showed history of colorectal and breast cancer in some members of her family. Her tumour loss the expression of hMLH1 and hPMS2, making her more likely to have Lynch syndrome. The other two cases were more than 60 years of age when diagnosed with colorectal cancer which is not a typical age for tumour onset in Lynch syndrome patients. However; case 13 who loss the expression of hMSH6 in his proximal colon tumour can still have Lynch syndrome. Case 27 was 77 years old when developed a rectal cancer. The loss of hMLH1 expression in his tumour in addition to the lack of family history of cancer makes him more likely to have microsatellite instable sporadic cancer. Our results are in keeping with previous report by Hamplel et al. [671]. They examined 1066 patients with newly diagnosed colorectal adenocarcinoma for MSI. Among patients whose screening results were positive for MSI, they looked for germLine mutations in the 4 main MMR genes using IHC, genomic sequencing and deletion studies. MSI was detected in 19.5% of their study population and 2.2% were confirmed to have Lynch syndrome. Of the patients who were found to have Lynch syndrome 10 were more than 50 years and 5 did not meet the clinical criteria for diagnosis of

HNPCC. Their data suggested the similar efficiency of IHC and the more complex genetic analysis for MSI testing.

Our findings and the previous reports pointed out the importance of molecular screening of patients with colorectal cancer for MSI using immunohistochemistry. This strategy managed to identify mutations in patients otherwise would not have been detected. Therefore, we recommend it as a policy for all newly diagnosed colorectal cancer patients due to its important prognostic implications.

Chapter 8:

Discussion

Colorectal cancer is the fourth most common cancer in men and the third most common cancer in women worldwide [701]. In the USA, colorectal cancer is the second most common cause of cancer death among men aged 40 to 79 years and accounts for 9% of all cancer related deaths [702]. In Ireland, the National Cancer Registry predicts that the incidence of colorectal cancer will increase from 2111 cases in 2005 to 5537 in 2035 [703], indicating a more than 100% increase over the next 30 years. In this setting of increasing disease burden, translational research is of vital importance to clinical advancement. At the molecular level, activation of oncogenes and inactivation of tumour suppressor genes [359] are processes known to be involved in colorectal carcinogenesis. Additionally, abrogation of mismatch repair systems [360] contributes to some colorectal cancers. Nevertheless, exactly how these genetic alterations bring about the development and progression of colorectal carcinomas remains to be resolved. To complicate the picture, accumulation of mutant genes in neoplasms tends to be accompanied by other genetic and epigenetic changes including loss of heterozygosity, inactivation of important genes by methylation or loss of imprinting [361] or gene amplifications, all of which can alter gene expression profiles. Therefore, genome wide monitoring of gene expression is of great importance if we are to disclose the numerous and diverse events associated with carcinogenesis. Molecular profiling, a tool of genome monitoring, is an attempt to identify the different combinations of genetic events or alternative pathways that may be represented by cancers of a similar type.

The principle of an adenoma-carcinoma sequence, described in 1990, postulates that the transition from adenoma to carcinoma is associated with an accumulation of genetic events in key regulatory genes that confer a growth advantage to a clonal population of cells [74]. Since then, although molecular detection methods based on gene mutation determination have been carried out for several years, the clinical utility of the many molecular markers and their clinical applications remain limited for colorectal cancer patients. Therefore, there is real need for new molecular markers to to improve umour subclassification and prediction of clinical outcome.

Microarray technology and gene expression profiling studies in colorectal cancer stimulated an interest in potential results that could be directly used in the routine clinical setting. Gene expression signatures predictive of disease outcome and response to adjuvant therapy have been generated and are being evaluated in the clinical setting. Such molecular diagnostics and their promise of tailored therapy generated much excitement among researchers however they have yet to be fully incorporated into today's standard of care as they are limited by difficulties in reproducibility, standardisation and lack proof of significance beyond traditional prognostic tools.

Gene expression in colorectal cancer:

One of the primary aims of this study was to characterise the expression profiles of candidate genes in colorectal tissue. Rigourous evaluation of appropriate genes with which to normalise real-time quantitative PCR data identified *PPIA* and *B2M* as the most stably expressed genes in colorectal tissue samples. This enabled the development of a robust experimental approach which ensured that subsequent profiling of gene expression levels would be measured accurately and reproducibly in colorectal tissue. As a result, a comprehensive list of genes with highly differential expression patterns was derived.

CXCL12 and its receptors CXCR4 and CXCR7:

The first candidates to be examined were the chemokine *CXCL12* and its receptors *CXCR4* and *CXCR7*, whose gene expression levels were determined in 107 tumour and tumour associated normal colorectal tissues, the largest patient cohort reported to date. Significant down-regulation of *CXCL12* in tumour compared to normal colorectal tissue was found, in contrast to *CXCR4*, which showed non-significant up-regulated expression levels in tumour tissues. The reduced expression of *CXCL12* was noticed in both polyps and tumours. This could be explained by the role of *CXCL12* in tumour immunology; however, it may highlight a possible tumour suppressor function of this gene. Investigation of the interaction between *CXCL12, CXCR4* and *CXCR7* may provide some understanding of their functions and the role of each gene in regulating the expression of the others. Despite the reciprocal patterns of expression, strong positive correlation of *CXCL12/CXCR4* and *CXCL12/CXCR7* in both tumour and

normal colorectal tissue was found. Moreover, *CXCR4* and *CXCR7* expression patterns correlated in the same manner. Saigusa *et al*. also reported significant positive correlation between expression levels of *CXCL12* and *CXCR4* in patients with rectal cancer who underwent preoperative CRT. Moreover, the expression of *CXCR7* in *CXCR4* positive cells appears to enhance the responsiveness to *CXCL12* as reported by Sierro [472]. These findings suggest a possible receptor interaction in tumour and normal colorectal tissues.

Correlation of gene expression levels with clinicopathological data indicated that levels of *CXCL12* and *CXCR7* were lower in the proximal colon. This may indicate a possible role of this axis in microsatellite instability (MSI), as tumours associated with MSI arise mainly in the proximal colon. Down-regulation of *CXCL12* and its receptors was also found to be associated with increased tumour size, local invasion, poor differentiation, advanced nodal stage, advance tumour stage and lymphovascular invasion. Of further interest, we identified for the first time the prognostic significance of *CXCR7* mRNA in colorectal cancer. We found that patients with high expression of *CXCR7* in their tumour cells lived longer than their counterparts with lower *CXCR7* gene expression. This was further confirmed by multivariate analysis.

TGFB1 and its receptors TGFBR1 and TGFBR2:
Although no significant differences were identified in gene expression levels of the chemokine receptor molecules *TGFBR1* and *TGFBR2* in tumour versus normal tissue, the expression of their ligand *TGFB1* was found to be significantly lower in polyps and higher in tumours compared to normal tissue. These findings confirm previous work by Daniel *et al* (2007), investigating TGFB1 protein expression by IHC in colorectal cancer. The authors demonstrated than in high-grade dysplastic polyps, than in low-grade dysplastic polyp [521]. Matsushita *et al* (1999) found that *TGFB* receptor mRNA was expressed mainly by normal and adenoma colorectal tissues whereas *TGFB1* expressed by cancer [516].
Moreover, the significant positive correlation between *TGFB1* and the expression levels of its receptors in both tumour and normal tissue confirms that their role in colorectal cancer is more complex than a simple legend-receptor feedback.

Interestingly, we identified for the first time the relationship of *TGFB* pathway and some established prognostic clinicopathological parameters. Low expression of *TGFBR1* was found to be associated with raised CEA serum level and local tumour invasion. In addition, *TGFBR2* down-regulation was associated with local, perineural and lymphovascular invasion and advanced nodal stage. These findings will further confirm the role of *TGFB* receptors as tumour suppressor. The down-regulation of *TGFBR2* in proximal compared to distal tumours was described before and highlights the role of this gene in microsatellite instable tumours.

Tumours of proximal and distal parts of the colon may form different but related groups of tumours because of their different embryological origin, different exposure to bowel contents and differences in clinical presentation, progression and possible genetic and environmental epidemiology [522].

Many previous studies have examined the relationship between *TGFB* pathway and the disease progression in colorectal cancer. Nevertheless, this is the first study to explore the relation of *TGFB1* and its receptors mRNA in colorectal cancer using RT-PCR. Moreover, the large cohort of patients in this study gives it further advantage compared to the other studies.

Other genes shown to be potential biomarkers in this study included *CDH17, FABP1, IL8, MUC2* and *PDCD4.* In colorectal cancer, *CDH17* expression was only investigated at protein level using IHC and immunoblotting. Hinoi et al. examined the protein expression in human colorectal cancer cell lines. In their study, *CDH17* was not detected in cell lines showing dedifferentiated phenotypes [444]. This was further confirmed by Takamura et al. who examined the *CDH17* expression in four cell lines and 45 human primary colorectal carcinoma using monoclonal antibodies. In cell lines the protein was expressed in differentiated but not the dedifferentiated phenptypes while in tissues reduced *CDH17* expression was associated with high tumour grade, advanced stage and lymphatic invasion and metastasis [373]. Moreover, Kwak et al. found reduced expression in 51% of the 207 colorectal cancers he studied using immunohistochemistry and he significantly correlated down-expression of *CDH17* with poor survival and

lymph nodes metastasis [374]. To our knowledge, this is the first study to investigate *CDH17* mRNA in colorectal cancer using RQ-PCR. Our findings support the above reports and confirm that down-regulation of *CDH17* in colorectal cancer is associated with poor differentiation, raised CA19.9 tumour marker serum level and local tumour invasion indicated by increase bowel wall involvement. Interestingly, *CDH17* expression correlated with increased tumour diameter and tumour thickness (indices of intraluminal tumour growth) and decreased with increased bowel wall involvement (index of local tumour invasion). Those findings could be explained by the adhesion function of the protein. Generally, for the tumour to grow in diameter and thickness it needs to retain adhesion molecules expression, while loss or inactivation of those adhesion molecules correlate with inhibition of cell aggregation and promotion of tumour invasiveness. This finding may highlight the potential role of *CDH17* as a marker for rectal cancer surgical management planning. In other wards, decrease level of *CDH17* may indicat local invasion of tumour and therefore total mesorectal excission (TME) will be indicated.

Evidence of dysregulated *FABP1* gene expression has been reported in colorectal gene expression array datasets [365, 477], however, little is known of its expression profile with regard to clinical data. Lawrie et al. identified consistent loss of *FABP1* in tumour compared to normal colon and also noted the association of decreased protein expression and poorly differentiated tumours and large adenomas [344]. Moreover, *FABP1* expression was found to be associated with good prognosis after liver resection of colorectal cancer metastasis [478]. Although no statistically significant correlation between *FABP1* expression and clinicopathological parameters was identified in this study, we observed that *FABP1* is differentially expressed in normal-adenoma-carcinoma sequence and its loss occurred early in colorectal cancer tumourogenesis. This indicates tumour suppressor function of *FABP1* in colorectal cancer. The loss of *FABP1* in colorectal cancer contrast with the findings in other tumours types which might be explained by the organ-specific distribution and the different role of *FABP1* through distinct intracellular interacting molecules.

In keeping with the previous reports, we noted overexpression of *IL8* in tumour compared to normal colorectal tissue. In addition, we identified a progressive manner of increase gene expression from normal, to polyps, to tumour. The early dysregulation of *IL8* in colorectal cancer suggest that the gene may play a role in carcinogenesis in addition to its confirmed role in tumour progression. Correlations with clinicopathological parameters revealed significant association of reduced *IL8* expression and poor tumour differentiation, advanced nodal stage and disease recurrence. Although the significant of these findings is unclear, it should be considered when planning *IL8* targeting therapy.

Furthermore, we confirmed *MUC2* mRNA down-regulation in non-mucinous and over-regulation in mucinous colorectal cancer. We also showed decreased expression of *MUC2* in a progressive manner from tumour-associated normal, to polyps, to tumours. No significant association of *MUC2* and clinicopathological variables other than CA19.9 serum levels has been determined in this study. Regarding *PDCD4* mRNA, its expression was significantly lower in tumour and polyp compared to tumour-associated tissue in keeping with the protein expression levels described before [477, 502, 511]. Furthermore, we identified the novel association of reduced *PDCD4* expression with disease recurrence and raised CA19.9 serum level. These findings suggest that *PDCD4* involves in both tumour promotion and tumour progression and represent a potential biomarker for evaluating the transition of normal colorectal tissue to adenoma and carcinoma. Reduced expression of PDCD4 in proximal compared to distal colon may indicate a potential role in microsatellite instability (MSI) and Lynch syndrome.

Measurement and quantifying of tumour response to neoadjuvant CRT is an important parameter in order to elucidate factors that may allow for response prediction and planning of next step of treatment in rectal cancer patients. Clinical response (cCR), pathological response (pCR) and tumour downstaging are the commonly used methods to measure response. Both clinical response and tumour downstaging compared the tumour characteristics before and after treatment clinically and using radiological tools like magnetic resonance imaging (MRI) and trans-rectal ultrasound (TRUS). Whereas pathological response

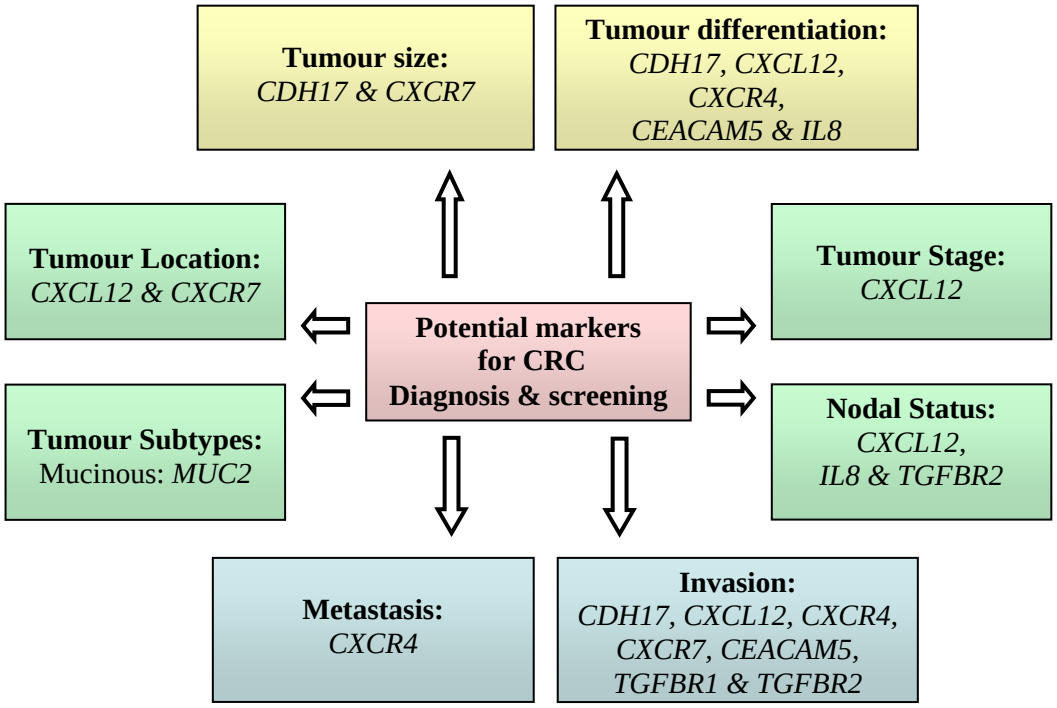
(regression grade) stratifies response base on biological effect of radiation on tumours. Mandard tumour regression grade, originally described for oesophageal cancer, is the most commonly used [44]. It consists of five different grades based on ratio of fibrosis to tumours. We identified, for the first time, a group of genes that can be used as markers to quantify tumour response following neoadjuvant therapy in rectal cancer patients.

Clinical applications:

the list of the genes identified in this study could serve as molecular markers to complement existing histopathological factors in screening, diagnosis, follow up and therapeutic strategies for individualised care of patients (Figures 8.1 and 8.2)

Figure 8.1: Potential biomarkers for CRC.
Genes identified in the study as potential biomarkers for CRC screening, diagnosis and disease progression.

- Biomarkers for diagnosis and screening:



- Biomarkers of disease progression:

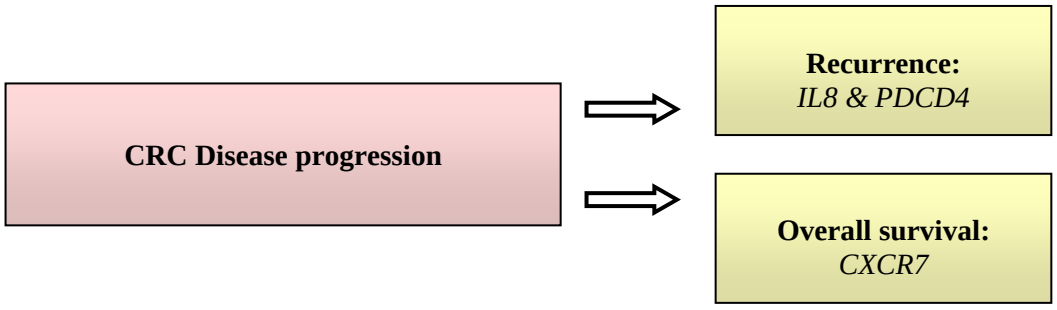
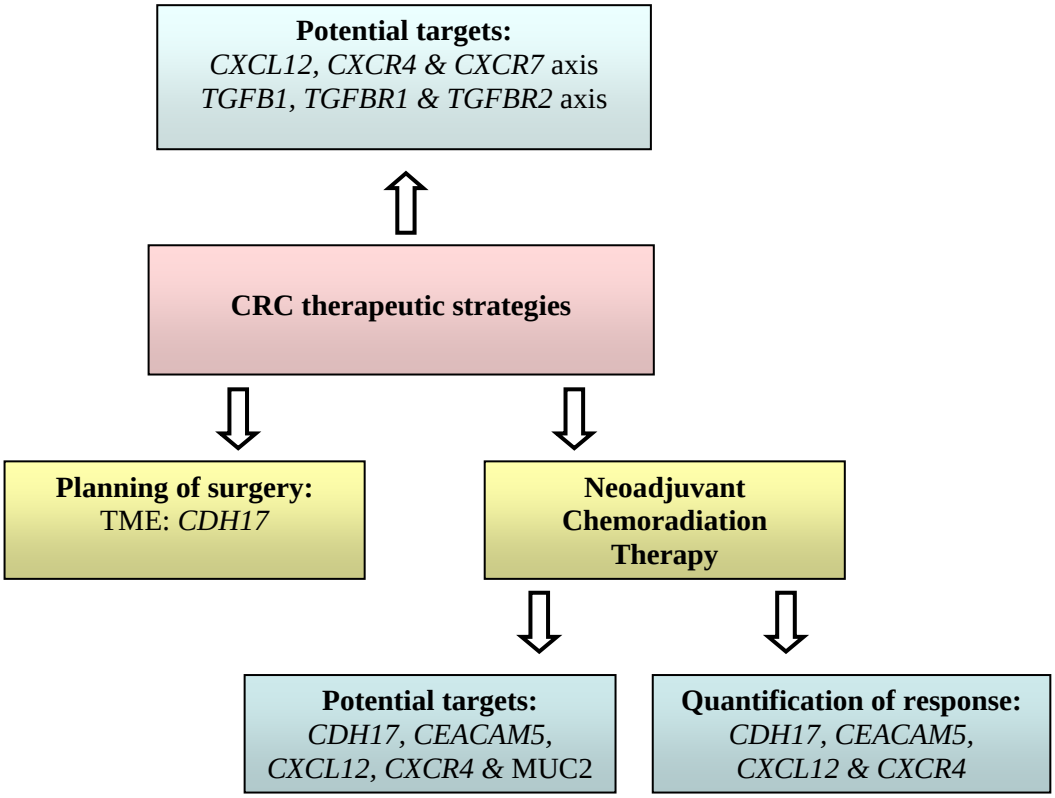


Figure 8.2: Gene expression and CRC management stratigies.



miRNA expression and prediction of rectal cancer response to neoadjuvant chemoradiation therapy (CRT):

Michael et al. in 2003 have published the first report of miRNA in colorectal cancer [263]. They found reduced accumulation of specific miRNA in colorectal neoplasia and identified 28 different miRNA sequences between colonic cancer and normal mucosa. They also identified the human homologues of murine *miR-143* and *miR-145*. Since then numerous reports have demonstrated the role of miRNAs in colorectal carcinogenesis and highlight their potential use as diagnostic and prognostic biomarkers [274, 275, 561-563]. Moreover, increasing evidence support the use of miRNA profiling to characterise human tumours and distinct predictive signatures have been reported for hepatocellular carcinoma, oesophageal cancer, colon and lung cancer [264, 564-567]. Nevertheless, no study has investigated the possible use of miRNA in predicting response to neoadjuvant chemoradiation in rectal cancer.

On the other hand; FFPE tissue offers a widely available and rich archive of well characterised tissue specimens and patient data for comparative molecular and clinical retrospective studies [544]. New extraction methods have made it possible to retrieve total RNA from preserved tissue specimens to a level that could be quantified by RQ-PCR. However, the application of these methods to FFPE tissue is limited by extensive RNA fragmentation and modifications. Until recently, FFPE samples have not considered reliable source of mRNA for gene profiling experiments due to difficulty in obtaining intact mRNA from these samples. Therefore, optimising of the extraction methods and RNA quality from FFPE tissues is of particular interest to many research groups.

For the purpose of this study, we compared the performance of three RNA extraction methods, and identified Qiagen RNeasy FFPE kit as a preferred methodology. The main reasons why RNA extracted from FFPE tissues is of poor quality are RNA fragmentation and cross-linked with other molecules including proteins [568]. The problem of fragmentation is solved by choosing small fragments for detection by PCR-based methods [568, 569]. Qiagen RNeasy FFPE kit uses Proteinase k at 55°C to break the cross-linked RNA formed with proteins. Incubation at 80°C in buffer PKD is an important step in RNA isolation process using this method. It partially reverses formaldehyde modification of nucleic acids; thereby improves the quality of RNA harvested. To ensure that the recovery of miRNA was adequately assessed it was crucial to select appropriate miRNA targets for integration by RQ-PCR. *miR-10b*, *miR-143*, *miR-145*, *miR-21* and *miR-30a-3p* were chosen because they were intensively investigated in colorectal cancer before [263, 264, 269, 570, 571]. Using FFPE and fresh-frozen tissue samples we were able to demonstrate the previously confirmed down-regulation of *miR-10b*, *miR-143*, miR-*145* and *miR-30a-3p*, and the over-expression of *miR-21* in colorectal cancer compared to tumour-associated normal tissues. The RQ-PCR amplification results reported here demonstrate that miRNA targets are detected at levels nicely matched expression levels from reference fresh-frozen tissues.

When comparing miRNA expression levels between fresh-frozen and matched FFPE tissue samples, we observed some variations with correlation coefficients

of 0.85 - 0.61. These variations could be attributed to the technical variations from one replicate to another. Also the amount of miRNA used in each reaction would have an impact on the miRNA expression level, as 5ng reaction was carried out when examining fresh-frozen tissues while 100ng were used in FFPE reactions. In addition some of these variations could be explained by the sample heterogeneity.

Furthermore, to enable extraction of miRNA from FFPE tissue blocks with different cross-sectional areas in quantities adequate for multiple analyses of the purified miRNA, we determined the number of slices required for optimal RNA yield. The purified RNA yield increased stepwise when we used 1, 2, 3, or 4 slices; however, the changes in concentration were not statistically significant. Doleshal et al. performed RNA isolation in duplicate using 4, 8, 12 or 16 slices of FFPE tissues from two prostatic cancer locks that differ in their tissue cross-sectional area [538]. For the tissue blocks with smaller cross-sectional area they observed a linear increase in RNA recovery, while for the blocks with larger area not all the tissue was digested in tubes containing more than 4 slices resulting in yields that were lower than expected. To further evaluate the RNA recovered we selected *miR-143* and *miR-145* isolated from 1, 2, 3 or 4 slices for integration by RQ-PCR. The reactions were carried in triplicate for each slice number. Regardless of the number of slices used for miRNA extraction, the mean expression level of miRNAs was stable with standard deviation less than 0.3. This will confirm the suitability of this method for RNA isolation from tissue as small as a colonic biopsy retrieved during endoscopy procedure.

Neoadjuvant CRT has become the preferred treatment modality for locally advanced rectal adenocarcinoma with a complete pathological response observed in up to 30% of patients [59, 576]. The ability to predict response to pre-treatment chemoradiation may spare poorly responding patients from undergoing aggressive and severely toxic treatment [577, 578] from which they would derive no benefit. At present there is no reliable technique to predict clinical or pathological complete tumour regression after treatment and limited data exist for each potential modality in this regard. Hence; many molecular markers have

been assessed for their predictive values. Nevertheless, it seems unlikely that they will prove to be clinically useful response predictors.

Change in miRNA expression profiles during treatment of cancer could potentially provide a tool to predict and estimate the success of certain therapies. By enabling screening of tissue samples for multiple miRNAs simultaneously, microarrays revealed convincing evidence that a large number of miRNAs are deregulated in therapy resistance or sensitive cancer cells. The extent of changes in miRNA expression were reported following anticancer treatment with various chemotherapeutic drugs in different cancer cell lines and patient samples [582]. To the author's knowledge, this is the first study to investigate the role of miRNA as predictors of response to neoadjuvant CRT therapy in rectal cancer. Using ANN to analyse the miRNA profiling data, a distinct miRNA expression signature predictive of response to neoadjuvant CRT in 12 FFPE pre-treatment rectal cancer tissue samples was identified. These signatures consisted of three miRNA transcripts (*miR-16, miR-590-5p* and *miR-153)* to predict complete vs. incomplete response and two miRNAtranscript *(miR-519c-3p and miR-561)* to predict good versus poor response with a median accuracy of 100%.

Although *miR-16* was described as being stably expressed in both colorectal and breast tissues and has been highlighted as a good endogenous control for miRNA profiling in cancer research using RQ-PCR [293, 323], several studies have confirmed its dysregulation in many cancers including CRC [587-590]. Moreover; Schaefer *et al* examined the expression of four putative reference genes including *miR-16* with regard to their use as normalizer in prostatic cancer and they found that normalization to *miR-16* can lead to biased results [591]. Although no report has determined the significance of *miR-153* and *miR-590* in CRC, their role in carcinogenesis ws highlighted before [599]. Shan *et al*. [599] investigated the role of miRNAs on the expression and regulation of transforming growth factor-beta1 (*TGFB1*), TGF-beta receptor type II (*TGFBRII*), and collagen production *in vivo* and *in vitro*. They found that nicotine produced significant upregulation of expression of *TGFB1* and *TGFBRII* at the protein level, and a decrease in the levels of miRNAs *miR-133* and *miR-590*. The role of *miR-519* in cancer was documented before [603-605]. *miR-519*

was reported as a tumour suppressor and was found to reduce cell proliferation by lowering RNA-binding protein HuR levels [603]. It decreases HuR translation without influencing HuR mRNA abundance [603, 604]. Abdelmoshen *et al* examined the level of *miR-519* and HuR in pairs of cancer and adjacent normal tissues from ovary, lung and kidney and reported significant high levels of HuR, unchanged HuR mRNA concentration and reduced *miR-519* levels in cancer specimens compared to normal tissues [604]. They also found that tumour cells overexpressing *miR-519* fpormed significantly smaller tumours while those expressing reduded *miR-519* gave rise to substancilally larger tumours.

Taken together, therefore, using microarray analysis of pretreatment FFPE rectal cancer tissues, for the first time a group of miRNAs predictors of response to neoadjuvant CRT was identified. This, indeed, can lead to a significant improvement in patient selection criteria and personalized rectal cancer management. However; before clinically applying this data, a validation study using a large cohort of patients needs to be performed.

miRNA: mRNA correlations in colorectal cancer:

MiRNAs are crucial in eukaryotes gene regulation, especially in development and differentiation [609, 610], and their expression in cancers has indicated that they may have a tumour suppressor or oncogenic function [611]. Functional characterisation of miRNAs will depends heavily on identification of their specific gene targets. In addition; a number of studies have shown that more than one miRNA can potentially bind to a single targeted gene; hence multiple miRNAs may cooperatively control the expression of target genes [305, 306]. Numerous bioinformatic methods have been developed to high-throughput prediction of miRNA target genes [295, 300-302, 304], although it is understood that the presumed targets have to be validated experimentally.

Computational approaches have been developed based on an understanding of the relationship between the miRNA seed region and the 3-UTR of the target gene. To develop computational algorithms, empirical evidence is examined carefully and principles of miRNA target recognition are extracted. After preparation of the data set, miRNA binding sites are identified by determining

the base pairing pattern of miRNAs and mRNAs according to the complementarity within specific region, thermodynamic analysis of the miRNA/mRNA duplexes via calculation of the free energy, and comparative sequence analysis. Then, the number of target sites of miRNA is counted [312, 612-614]. However, a recent report by Gatt *et al*. indicated that the miRNAs may have many more targets than anticipated by convensional prediction methods [619]. In addition; these algorithms can result in prediction of false-positives or some targets may pass undetected. The false positive rates were estimated at 22%, 24% and 30% for TargetScan, miRanda and PicTar, respectively [312]. The PicTar and EMBL algorithms have a reported sensitivity of 70-80% [615] indicating 20-30% of targets may go undetected.

To further understand the factors control gene expression, and therefore the protein biosynthesis, we performed bioinformatics analysis to search for putative miRNA/target genes duplexes from the panel of genes and miRNA previously investigated by our research group in the Department of Surgery, NUI Galway. In addition, correlation analysis was performed between miRNA and mRNA which identified novel pairs of miRNA:mRNA duplexes not previously identified by any of the computational approaches mentioned above. In this study, the *in sillico* predicted relationship of miR-21/*PDCD4*, miR-31/*CXCL12* and miR-145/*IL8* duplexes was confirmed by real-time PCR expression analysis. Moreover, novel combinations of: miR-10b/*IL8*, miR-17/*CXCL12*, miR-21/*FABP1*, miR-31/*FABP1* and miR-31/*PDCD4* were also identified. These combinations could represent valid miRNA/target gene duplexes. The interaction of the reported functions of miRNAs and mRNAs, in addition to their reciprocal patterns of expression in tumours and tumour-associated normal tissues and in association to clinicopathological parameters might support the relationship of miRNA/mRNA pairs highlighted by the computational algorithms and correlation analysis.

Mismatch-repair (MMR) protein expression in colorectal cancer:
MMR proteins are nuclear enzymes, which participate in repair of base-base mismatch that occur during DNA replication in proliferating cells. The proteins form heterodimers that bind to areas of abnormal DNA and initiates its removal.

Loss of MMR proteins leads to an accumulation of DNA replication errors, particularly in areas of the genome with short repetitive nucleotide sequences, a phenomenon known as microsatellite instability (MSI) [360, 648, 649]. In addition to screening for Lynch syndrome, testing for MSI is important because of its possible prognostic and therapeutic implications. Cancers with high microsatellite instability (H-MSI) were reported to have a more favourable clinical out come than non-MSI tumours and the survival advantage conferred by the MSI phenotype is independent of tumour stage and other clinicopathological variables [156-158]. Moreover, tumours with H-MSI are thought to be less responsive to 5-fluorouracil and other anticancer agents *in vitro* and *in vivo* [159-161].

The major laboratory tests used in the evaluation of patients suspected to have Lynch syndrome include testing of tumour tissues using immunohistochemistry (IHC), MSI testing or germ line testing for mismatch defects. IHC has the advantage over the other methods, as the primary screening method, since it is less demanding to perform and is available as part of routine services in general pathology laboratories. In addition, IHC will determine which protein is affected and provides gene specific information; thereby direct the genetic analysis rather than performing exhausting, time and material consuming unnecessary tests. Nevertheless, while most of mutations will results in total loss of the protein expression , in some cases mutations only result in loss of function rather than the expression of the protein which will still be detectable by IHC.
In this study, MMR protein expression was tested without considering the family history in a prospective of newly diagnosed colorectal cancer patients. This analysis identified three patients with loss of one or more MMR protein. The first patient (case 3) was less than 40 years old when diagnosed with caecal cancer. Although her family history was not fully documented (figure 3.2), she showed history of colorectal and breast cancer in some members of her family. Her tumour loss the expression of hMLH1 and hPMS2, making her more likely to have Lynch syndrome. The other two cases were more than 60 years of age when diagnosed with colorectal cancer which is not a typical age for tumour onset in Lynch syndrome patients. However; case 13 who loss the expression of hMSH6 in his proximal colon tumour can still have Lynch syndrome. Case 27 was 77

years old when developed a rectal cancer. The loss of hMLH1 expression in his tumour in addition to the lack of family history of cancer makes him more likely to have microsatellite instable sporadic cancer. Our results are in keeping with previous report by Hamplel *et al*. [671]. They examined 1066 patients with newly diagnosed colorectal adenocarcinoma for MSI. Among patients whose screening results were positive for MSI, they looked for germLine mutations in the 4 main MMR genes using IHC, genomic sequencing and deletion studies. MSI was detected in 19.5% of their study population and 2.2% were confirmed to have Lynch syndrome. Of the patients who were found to have Lynch syndrome 10 were more than 50 years and 5 did not meet the clinical criteria for diagnosis of HNPCC. Their data suggested the similar efficiency of IHC and the more complex genetic analysis for MSI testing.

Our findings and the previous reports pointed out the importance of molecular screening of patients with colorectal cancer for MSI using immunohistochemistry. This strategy managed to identify mutations in patients otherwise would not have been detected. Therefore, we recommend it as a policy for all newly diagnosed colorectal cancer patients due to its important prognostic implications.

Future Work

The study of gene expression in colorectal cancer has yielded interesting results and opened new avenues of exploration. The challenge we now face is the translation of new scientific knowledge into clinically applicable diagnostic, prognostic and therapeutic tools for use in the management of colorectal cancer

The microarray analysis of pretreatment FFPE rectal cancer tissues has identified a group of miRNAs predictors of response to neoadjuvant CRT. This, indeed, can lead to a significant improvement in patient selection criteria and personalized rectal cancer management. However; before clinically applying this data, a validation study using a large cohort of patients needs to be designed.

Reciprocal expression observed between several genes and their miRNAs partners, suggestive of novel mechanisms which could become uncoupled in colorectal carcinogenesis. These findings support the hypothesis that these miRNAs:mRNA duplexes may hold potential as therapeutic agents/targets in colorectal cancer. *In-vivo* functional analysis is warranted to further investigate this potential. One possible direction would be the development of a model of colorectal cancer in which the effect of specific miRNA up- or down-regulation on gene expression, and therefore tumour behavior, could be assessed. In this manner, the potential of these duplexes as therapeutic targets could be explored.

MMR protein analysis has pointed out the importance of molecular screening of patients with colorectal cancer for MSI using immunohistochemistry. Expansion of this analysis to a wider scale via microarray promises to identify novel biomarkers that could be used for prognostication and personalized patient treatment in colorectal cancer.

References:

1. Parkin M, et al., *Global Cancer Statistics, 2002.* CA Cancer J Clin 2005. 55: p. 74-108.
2. Comber H and Walsh PM, *Patterns of care and survival of cancer patients in Ireland 1994 to 2004. Summary report. National Cancer Registry*. 2008: Cork, Ireland.
3. Ivanovich JL, et al., *A practical approach to familial and hereditary colorectal cancer.* Am J Med, 1999. 107(1): p. 68-77.
4. Hamilton S and Aaltonen LA, *Pathology and genetics of tumours of digestive system*. 2000, World Health Organisation: New York.
5. Compton, C.C., *Colorectal carcinoma: Diagnostic, prognostic, and molecular features.* Modern Pathology, 2003. 16(4): p. 376-388.
6. Jessurun, J., M. Romero-Guadarrama, and J.C. Manivel, *Medullary adenocarcinoma of the colon: Clinicopathologic study of 11 cases.* Human Pathology, 1999. 30(7): p. 843-848.
7. Rankin FW and Broders AC, *Factors influencing prognosis in carcinoma of the rectum.* Surg Gynecol Obstet, 1928. 46: p. 660–667.
8. Dukes C, *The classification of cancer of the rectum.* J Pathol Bacteriol, 1932. 35: p. 323–332.
9. Dukes C, *Histological grading of rectal cancer.* Proc R Soc Med, 1937. 30: p. 371–376.
10. Compton, C., et al., *American Joint Committee on Cancer prognostic factors consensus conference - Colorectal working group.* Cancer, 2000. 88(7): p. 1739-1757.
11. Chung, C.K., R.J. Zaino, and J.A. Stryker, *Colorectal-Carcinoma - Evaluation of Histologic Grade and Factors Influencing Prognosis.* Journal of Surgical Oncology, 1982. 21(3): p. 143-148.
12. Riboli, E.B., et al., *Colorectal-Cancer - Relationship of Histologic Grading to Disease Prognosis.* Tumori, 1983. 69(6): p. 581-584.
13. Sobin LH and Wittekind C, *TNM classification of malignant tumours*. 6th ed. 2002, New York: J. Wiley.
14. Compton, C.C., et al., *Prognostic factors in colorectal cancer - College of American Pathologists Consensus Statement 1999.* Archives of Pathology & Laboratory Medicine, 2000. 124(7): p. 979-994.
15. Hammarström S, *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues.* Seminars in Cancer Biology, 1999. 9(2): p. 67-81.
16. Gold P and Freedman SO, *Demonstration of Tumor-Specific Antigens in Human Colonic Carcinomata by Immunological Tolerance and Absorption Techniques.* Journal of Experimental Medicine, 1965. 121(3): p. 439-&.
17. Astler, V.B. and F.A. Coller, *The Prognostic Significance of Direct Extension of Carcinoma of the Colon and Rectum.* Annals of Surgery, 1954. 139(6): p. 846-852.
18. Gunderso.Ll and H. Sosin, *Areas of Failure Found at Reoperation (Second or Symptomatic Look) Following Curative Surgery for Adenocarcinoma of Rectum - Clinicopathologic Correlation and Implications for Adjuvant Therapy.* Cancer, 1974. 34(4): p. 1278-1292.

19. Enker, W.E. and R.J. Heald, *Potency, Cure, and Local-Control in the Operative Treatment of Rectal-Cancer.* Archives of Surgery, 1992. 127(12): p. 1396-1402.
20. Heald, R.J., *Rectal-Cancer - the Surgical Options.* European Journal of Cancer, 1995. 31A(7-8): p. 1189-1192.
21. Mendenhall, W.M., R.R. Million, and W.W. Pfaff, *Patterns of Recurrence in Adenocarcinoma of the Rectum and Rectosigmoid Treated with Surgery Alone - Implications in Treatment Planning with Adjuvant Radiation-Therapy.* International Journal of Radiation Oncology Biology Physics, 1983. 9(7): p. 977-985.
22. Pilipshen, S.J., et al., *Patterns of Pelvic Recurrence Following Definitive Resections of Rectal-Cancer.* Cancer, 1984. 53(6): p. 1354-1362.
23. Gastrointestinal Tumor Study Group, *Prolongation of the Disease-Free Interval in Surgically Treated Rectal-Carcinoma.* New England Journal of Medicine, 1985. 312(23): p. 1465-1472.
24. Douglass, H.O., et al., *Survival after Postoperative Combination Treatment of Rectal-Cancer.* New England Journal of Medicine, 1986. 315(20): p. 1294-1295.
25. Oconnell, M., et al., *Improving Adjuvant Therapy for Rectal-Cancer by Combining Protracted-Infusion Fluorouracil with Radiation-Therapy after Curative Surgery.* New England Journal of Medicine, 1994. 331(8): p. 502-507.
26. Tepper, J.E., et al., *Adjuvant postoperative fluorouracil-modulated chemotherapy combined with pelvic radiation therapy for rectal cancer: Initial results of intergroup 0114.* Journal of Clinical Oncology, 1997. 15(5): p. 2030-2039.
27. Tepper, J.E., et al., *Adjuvant therapy in rectal cancer: Analysis of stage, sex, and local control - Final report of intergroup 0114.* Journal of Clinical Oncology, 2002. 20(7): p. 1744-1750.
28. Wagman, R., et al., *Sphincter preservation in rectal cancer with preoperative radiation therapy and coloanal anastomosis: Long term follow-up.* International Journal of Radiation Oncology Biology Physics, 1998. 42(1): p. 51-57.
29. Frykholm, G.J., B. Glimelius, and L. Pahlman, *Preoperative or Postoperative Irradiation in Adenocarcinoma of the Rectum - Final Treatment Results of a Randomized Trial and an Evaluation of Late Secondary Effects.* Diseases of the Colon & Rectum, 1993. 36(6): p. 564-572.
30. Sauer, R., et al., *Preoperative versus postoperative chemoradiotherapy for rectal cancer.* New England Journal of Medicine, 2004. 351(17): p. 1731-1740.
31. Bosset, J.F., et al., *Chemotherapy with preoperative radiotherapy in rectal cancer.* New England Journal of Medicine, 2006. 355(11): p. 1114-1123.
32. Pahlman, L., et al., *Improved survival with preoperative radiotherapy in resectable rectal cancer.* New England Journal of Medicine, 1997. 336(14): p. 980-987.
33. Kapiteijn, E., et al., *Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer.* New England Journal of Medicine, 2001. 345(9): p. 638-646.

34. Francois, Y., et al., *Influence of the interval between preoperative radiation therapy and surgery on downstaging and on the rate of sphincter-sparing surgery for rectal cancer: The Lyon R90-01 randomized trial.* Journal of Clinical Oncology, 1999. 17(8): p. 2396-2402.
35. Crane, C.H., et al., *Response to preoperative chemoradiation increases the use of sphincter-preserving surgery in patients with locally advanced low rectal carcinoma.* Cancer, 2003. 97(2): p. 517-524.
36. Sauer, R., et al., *Preoperative versus postoperative chemoradiotherapy for rectal cancer.* New England Journal of Medicine, 2004. 351(17): p. 1731-1740.
37. Bosset, J.F., et al., *Enhanced tumorocidal effect of chemotherapy with preoperative radiotherapy for rectal cancer: Preliminary results - EORTC 22921.* Journal of Clinical Oncology, 2005. 23(24): p. 5620-5627.
38. Bosset, J.F., et al., *Preoperative radiation (Preop RT) in rectal cancer: Effect and timing of additional chemotherapy (CT) 5-year results of the EORTC 22921 trial.* Journal of Clinical Oncology, 2005. 23(16): p. 247S-247S.
39. Collette, L. and J.F. Bosset, *Adjuvant chemotherapy following neoadjuvant therapy of rectal cancer: The type of neoadjuvant therapy (chemoradiotherapy or radiotherapy) may be important for selection of patients - Reply.* Journal of Clinical Oncology, 2008. 26(3): p. 508-509.
40. Sebag-Montefiore, D., et al., *An international phase II study of Capecitabine, Oxaliplatin, Radiotherapy and Excision (CORE) in patients with MRI-defined locally advanced rectal adenocarcinoma. Interim results.* Ejc Supplements, 2005. 3(2): p. 170-170.
41. Rodel, C., et al., *Prognostic significance of tumor regression after preoperative chemoradiotherapy for rectal cancer.* Journal of Clinical Oncology, 2005. 23(34): p. 8688-8696.
42. Wheeler, J.M.D., et al., *Quantification of histologic regression of rectal cancer after irradiation - A proposal for a modified staging system.* Diseases of the Colon & Rectum, 2002. 45(8): p. 1051-1056.
43. Ryan, R., et al., *Pathological response following long-course neoadjuvant chemoradiotherapy for locally advanced rectal cancer.* Histopathology, 2005. 47(2): p. 141-146.
44. Mandard AM, et al., *Pathologic assessment of tumor regression after preoperative chemoradiotherapy of esophageal carcinoma. Clinicopathologic correlations.* Cancer, 1994. 73(11): p. 2680-2688.
45. Bouzourene, H., et al., *Importance of tumor regression assessment in predicting the outcome in patients with locally advanced rectal carcinoma who are treated with preoperative radiotherapy.* Cancer, 2002. 94(4): p. 1121-1130.
46. Rebischung, C., et al., *Prognostic value of p53 mutations in rectal carcinoma.* International Journal of Cancer, 2002. 100(2): p. 131-135.
47. Kandioler, D., et al., *TP53 genotype but not p53 immunohistochemical result predicts response to Preoperative short-term radiotherapy in rectal cancer.* Annals of Surgery, 2002. 235(4): p. 493-498.
48. Vanagunas, A., D.E. Lin, and S.J. Stryker, *Accuracy of endoscopic ultrasound for restaging rectal cancer following neoadjuvant*

chemoradiation therapy. American Journal of Gastroenterology, 2004. 99(1): p. 109-112.
49. Rau, B., et al., *Accuracy of endorectal ultrasound after preoperative radiochemotherapy in locally advanced rectal cancer.* Surgical Endoscopy and Other Interventional Techniques, 1999. 13(10): p. 980-984.
50. Kwok, H., I.P. Bissett, and G.L. Hill, *Preoperative staging of rectal cancer.* International Journal of Colorectal Disease, 2000. 15(1): p. 9-20.
51. Kuo, L.J., et al., *Interpretation of magnetic resonance imaging for locally advanced rectal carcinoma after preoperative chemoradiation therapy.* Diseases of the Colon & Rectum, 2005. 48(1): p. 23-28.
52. Chen, C.C., et al., *How accurate is magnetic resonance imaging in restaging rectal cancer in patients receiving preoperative combined chemoradiotherapy?* Diseases of the Colon & Rectum, 2005. 48(4): p. 722-728.
53. Denecke, T., et al., *Comparison of CT, MRI and FDG-PET in response prediction of patients with locally advanced rectal cancer after multimodal preoperative therapy: Is there a benefit in using functional imaging?* European Radiology, 2005. 15(8): p. 1658-1666.
54. Guillem, J.G., et al., *Sequential preoperative fluorodeoxyglucose-positron emission tomography assessment of response to preoperative chemoradiation: A means for determining longterm outcomes of rectal cancer.* Journal of the American College of Surgeons, 2004. 199(1): p. 1-7.
55. Calvo, F.A., et al., *F-18-FDG positron emission tomography staging and restaging in rectal cancer treated with preoperative chemoradiation.* International Journal of Radiation Oncology Biology Physics, 2004. 58(2): p. 528-535.
56. Amthauer, H., et al., *Response prediction by FDG-PET after neoadjuvant radiochemotherapy and combined regional hyperthermia of rectal cancer: correlation with endorectal ultrasound and histopathology.* European Journal of Nuclear Medicine and Molecular Imaging, 2004. 31(6): p. 811-819.
57. Rodel, C., et al., *Apoptosis as a cellular predictor for histopathologic response to neoadjuvant radiochemotherapy in patients with rectal cancer.* International Journal of Radiation Oncology Biology Physics, 2002. 52(2): p. 294-303.
58. Myerson, R.J., et al., *Pretreatment clinical findings predict outcome for patients receiving preoperative radiation for rectal cancer.* International Journal of Radiation Oncology Biology Physics, 2001. 50(3): p. 665-674.
59. Reerink, O., et al., *Molecular prognostic factors in locally irresectable rectal cancer treated preoperatively by chemo-radiotherapy.* Anticancer Research, 2004. 24(2C): p. 1217-1221.
60. Kim, N.K., et al., *p53, BCL-2, and Ki-67 expression according to tumor response after concurrent chemoradiotherapy for advanced rectal cancer.* Annals of Surgical Oncology, 2001. 8(5): p. 418-424.
61. Qiu, H.M., et al., *Molecular prognostic factors in rectal cancer treated by radiation and surgery.* Diseases of the Colon & Rectum, 2000. 43(4): p. 451-459.

62. Fu, C.G., et al., *Role of p53 and p21/WAF1 detection in patient selection for preoperative radiotherapy in rectal cancer patients.* Diseases of the Colon & Rectum, 1998. 41(1): p. 68-74.
63. Rau, B., et al., *Dynamic expression profile of p21 (WAF1/CIP1) and Ki-67 predicts survival in rectal carcinoma treated with preoperative radiochemotherapy.* Journal of Clinical Oncology, 2003. 21(18): p. 3391-3401.
64. Giralt, J., et al., *Epidermal growth factor receptor is a predictor of tumor response in locally advanced rectal cancer patients treated with preoperative radiotherapy.* International Journal of Radiation Oncology Biology Physics, 2002. 54(5): p. 1460-1465.
65. Rodel, F., et al., *High survivin expression is associated with reduced apoptosis in rectal cancer and may predict disease-free survival after preoperative radiochemotherapy and surgical resection.* Strahlentherapie Und Onkologie, 2002. 178(8): p. 426-435.
66. Smith, F.M., et al., *Cox-2 overexpression in pretreatment biopsies predicts response of rectal cancers to neoadjuvant radiochemotherapy.* International Journal of Radiation Oncology Biology Physics, 2006. 64(2): p. 466-472.
67. Smith, F.M., et al., *Pathological and molecular predictors of the response of rectal cancer to neoadjuvant radiochemotherapy.* Ejso, 2006. 32(1): p. 55-64.
68. George, M.L., et al., *Non-invasive methods of assessing angiogenesis and their value in predicting response to treatment in colorectal cancer.* British Journal of Surgery, 2001. 88(12): p. 1628-1636.
69. Charara, M., et al., *Microsatellite status and cell cycle associated markers in rectal cancer patients undergoing a combined regimen of 5-FU and CPT-11 chemotherapy and radiotherapy.* Anticancer Research, 2004. 24(5B): p. 3161-3167.
70. Friedberg, E.C., *DNA damage and repair.* Nature, 2003. 421(6921): p. 436-440.
71. Nowell, P.C., *Tumor progression: a brief historical perspective.* Seminars in Cancer Biology, 2002. 12(4): p. 261-266.
72. Maley, C.C., et al., *Selectively advantageous mutations and hitchhikers in neoplasms: p16 lesions are selected in Barrett's esophagus.* Cancer Research, 2004. 64(10): p. 3414-3427.
73. Knudson, A.G., *Cancer genetics.* American Journal of Medical Genetics, 2002. 111(1): p. 96-102.
74. Fearon ER and Vogelstein B, *A genetic model for colorectal tumorigenesis.* Cell, 1990. 61(5): p. 759-767.
75. Lengauer, C., K.W. Kinzler, and B. Vogelstein, *Genetic instabilities in human cancers.* Nature, 1998. 396(6712): p. 643-649.
76. Kinzler, K.W. and B. Vogelstein, *Lessons from hereditary colorectal cancer.* Cell, 1996. 87(2): p. 159-170.
77. Foulkes, W.D., *A tale of four syndromes: Familial adenomatous polyposis, Gardner syndrome, attenuated APC and Turcot syndrome.* Qjm-Monthly Journal of the Association of Physicians, 1995. 88(12): p. 853-863.

78. Soravia, C., et al., *Genotype-phenotype correlations in attenuated adenomatous polyposis coli.* American Journal of Human Genetics, 1998. 62(6): p. 1290-1301.
79. Miyaki, M., et al., *Characteristics of Somatic Mutation of the Adenomatous Polyposis-Coli Gene in Colorectal Tumors.* Cancer Research, 1994. 54(11): p. 3011-3020.
80. Chung, D.C., *The genetic basis of colorectal cancer: Insights into critical pathways of tumorigenesis.* Gastroenterology, 2000. 119(3): p. 854-865.
81. Powell SM, et al., *APC mutations occur early during colorectal tumorigenesis.* Nature, 1992. 359(6392): p. 235-237.
82. Vogelstein, B., et al., *Genetic Alterations During Colorectal-Tumor Development.* New England Journal of Medicine, 1988. 319(9): p. 525-532.
83. Bokoch, G.M. and C.J. Der, *Emerging Concepts in the Ras Superfamily of Gtp-Binding Proteins.* Faseb Journal, 1993. 7(9): p. 750-759.
84. Arber, N., et al., *Activation of c-K-ras mutations in human gastrointestinal tumors.* Gastroenterology, 2000. 118(6): p. 1045-1050.
85. Bos, J.L., et al., *Prevalence of Ras Gene-Mutations in Human Colorectal Cancers.* Nature, 1987. 327(6120): p. 293-297.
86. Tsao, J.I. and D. Shibata, *Further Evidence That One of the Earliest Alterations in Colorectal Carcinogenesis Involves Apc.* American Journal of Pathology, 1994. 145(3): p. 531-534.
87. Ochiai A and Hirohashi S, *Multiple genetic alterations in gastric cancer.*, ed. T Sugimura and M Sasako. 1997, New York: Oxford Univ. Press.
88. Somasundaram, K. and W.S. El-Deiry, *Tumor suppressor P53: Regulation and function.* Frontiers in Bioscience, 2000. 5: p. D424-D437.
89. Ohue M, et al., *A frequent alteration of p53 gene in carcinoma in adenoma of colon.* Cancer Res, 1994. 54(17): p. 4798-4804.
90. Kikuchiyanoshita, R., et al., *Genetic Changes of Both P53 Alleles Associated with the Conversion from Colorectal Adenoma to Early Carcinoma in Familial Adenomatous Polyposis and Nonfamilial Adenomatous Polyposis Patients.* Cancer Research, 1992. 52(14): p. 3965-3971.
91. Lane, D.P., *Cancer - a Death in the Life of P53.* Nature, 1993. 362(6423): p. 786-787.
92. Levine, A.J., *p53, the cellular gatekeeper for growth and division.* Cell, 1997. 88(3): p. 323-331.
93. Howe, J.R. and J.G. Guillem, *The genetics of colorectal cancer.* Surgical Clinics of North America, 1997. 77(1): p. 175-&.
94. Fearon, E.R., et al., *Identification of a Chromosome-18q Gene That Is Altered in Colorectal Cancers.* Science, 1990. 247(4938): p. 49-56.
95. Bernet A and Mehlen P, *Dependence receptors: when apoptosis controls tumor progression.* Bull Cancer, 2007. 98(4): p. E12-7.
96. Breivik, J. and G. Gaudernack, *Genomic instability, DNA methylation, and natural selection in colorectal carcinogenesis.* Seminars in Cancer Biology, 1999. 9(4): p. 245-254.
97. Herman, J.G., et al., *Incidence and functional consequences of hMLH1 promoter hypermethylation in colorectal carcinoma.* Proceedings of the National Academy of Sciences of the United States of America, 1998. 95(12): p. 6870-6875.

98. Kane, M.F., et al., *Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines.* Cancer Research, 1997. 57(5): p. 808-811.
99. Veigl, M.L., et al., *Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers.* Proceedings of the National Academy of Sciences of the United States of America, 1998. 95(15): p. 8698-8702.
100. Ramsahoye, B.H., C.S. Davies, and K.I. Mills, *DNA methylation: Biology and significance.* Blood Reviews, 1996. 10(4): p. 249-261.
101. Monk, M., *Epigenetic Programming of Differential Gene-Expression in Development and Evolution.* Developmental Genetics, 1995. 17(3): p. 188-197.
102. Feil, R. and G. Kelsey, *Genomic imprinting: A chromatin connection.* American Journal of Human Genetics, 1997. 61(6): p. 1213-1219.
103. Barbacid, M., *Ras Genes.* Annual Review of Biochemistry, 1987. 56: p. 779-827.
104. Jones, P.A. and P.W. Laird, *Cancer epigenetics comes of age.* Nature Genetics, 1999. 21(2): p. 163-167.
105. Grady, W.M., et al., *Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer.* Nature Genetics, 2000. 26(1): p. 16-17.
106. Baylin, S.B. and J.G. Herman, *DNA hypermethylation in tumorigenesis - epigenetics joins genetics.* Trends in Genetics, 2000. 16(4): p. 168-174.
107. Herman, J.G., et al., *Inactivation of the Cdkn2/P16/Mts1 Gene Is Frequently Associated with Aberrant DNA Methylation in All Common Human Cancers.* Cancer Research, 1995. 55(20): p. 4525-4530.
108. Toyota, M., et al., *CpG island methylator phenotype in colorectal cancer.* Proceedings of the National Academy of Sciences of the United States of America, 1999. 96(15): p. 8681-8686.
109. Toyota, M., et al., *Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification.* Cancer Research, 1999. 59(10): p. 2307-2312.
110. Watson, P. and H.T. Lynch, *The Tumor Spectrum in Hnpcc.* Anticancer Research, 1994. 14(4B): p. 1635-1639.
111. Aarnio, M., et al., *Cancer risk in mutation carriers of DNA-mismatch-repair genes.* International Journal of Cancer, 1999. 81(2): p. 214-218.
112. Smyrk, T.C., et al., *Tumor-infiltrating lymphocytes are a marker for microsatellite instability in colorectal carcinoma.* Cancer, 2001. 91(12): p. 2417-2422.
113. Alexander, J., et al., *Histopathological identification of colon cancer with microsatellite instability.* American Journal of Pathology, 2001. 158(2): p. 527-535.
114. Jass, J.R., J. Young, and B.A. Leggett, *Evolution of colorectal cancer: Change of pace and change of direction.* Journal of Gastroenterology and Hepatology, 2002. 17(1): p. 17-26.
115. Galiatsatos, P. and W.D. Foulkes, *Familial adenomatous polyposis.* American Journal of Gastroenterology, 2006. 101(2): p. 385-398.
116. Talbot IC, et al., *Familial adenomatous polyposis*, in *WHO classification of tumours. Pathology and genetics tumours of the digestive system*,

Hamilton SR and Aaltonen, Editors. 2000, IARC Press: Lyon. p. 120–125.
117. Fallen, T., et al., *Desmoid tumors - a characterization of patients seen at Mayo Clinic 1976-1999.* Familial Cancer, 2006. 5(2): p. 191-194.
118. Tomoda, C., et al., *Cribriform-morular variant of papillary thyroid carcinoma: Clue to early detection of familial adenomatous polyposis-associated colon cancer.* World Journal of Surgery, 2004. 28(9): p. 886-889.
119. Cetta, F., et al., *Germline mutations of the APC gene in patients with familial adenomatous polyposis-associated thyroid carcinoma: Results from a European cooperative study.* Journal of Clinical Endocrinology & Metabolism, 2000. 85(1): p. 286-292.
120. Aretz, S., et al., *Should children at risk for familial adenomatous polyposis be screened for hepatoblastoma and children with apparently sporadic hepatoblastoma be screened for APC germline mutations?* Pediatric Blood & Cancer, 2006. 47(6): p. 811-818.
121. Munitz MI, *Polyposis syndromes*, in *Polypoid diseases of the colon*, Ott DJ and Wu WC, Editors. 1986, Urban & Schwarzenberg: Baltimore. p. .
122. Dachman, A.H., et al., *Cronkhite-Canada Syndrome - Radiologic Features.* Gastrointestinal Radiology, 1989. 14(4): p. 285-290.
123. Gorlin, R.J., et al., *Bannayan-Riley-Ruvalcaba Syndrome.* American Journal of Medical Genetics, 1992. 44(3): p. 307-314.
124. Salem, O.S. and W.D. Steck, *Cowdens Disease (Multiple Hamartoma and Neoplasia Syndrome) - a Case-Report and Review of the English-Literature.* Journal of the American Academy of Dermatology, 1983. 8(5): p. 686-696.
125. Peutz JLA, *Over een zeer merkwaardige, gecombineerde familiaire polyposis van de slijmliezen van den tractus intestinalis met die van de neuskeelholte en gepaard met eigenaardige pigmentaties van huid-en slijmvliezen.* Ned Maandschr v Gen, 1921. 10: p. 134–146.
126. Jeghers, H., V.A. McKusick, and K.H. Katz, *Generalized Intestinal Polyposis and Melanin Spots of the Oral Mucosa, Lips and Digits - a Syndrome of Diagnostic Significance.* New England Journal of Medicine, 1949. 241(25): p. 993-1005.
127. Tomlinson, I.P.M. and R.S. Houlston, *Peutz-Jeghers syndrome.* Journal of Medical Genetics, 1997. 34(12): p. 1007-1011.
128. Buck, J.L., et al., *Peutz-Jeghers Syndrome.* Radiographics, 1992. 12(2): p. 365-378.
129. Spigelman AD and Phillips RKS, *Peutz-Jeghers syndrome*, in *Familialmadenomatous polyposis and other polyposis syndromes*, Phillips RKS, Spigelman AD, and Thomson JPS, Editors. 1994, Edward Arnold: London. p. 188–202.
130. Wang, Z.J., et al., *Allelic imbalance at the LKB1 (STK11) locus in tumours from patients with Peutz-Jeghers' syndrome provides evidence for a hamartoma-(adenoma)-carcinoma sequence.* Journal of Pathology, 1999. 188(1): p. 9-13.
131. Aaltonen, L.A., *Hereditary intestinal cancer.* Seminars in Cancer Biology, 2000. 10(4): p. 289-298.
132. McColl, I., et al., *Juvenile Polyposis Coli.* Proceedings of the Royal Society of Medicine-London, 1964. 57(10): p. 896-897.

133. Houlston, R., et al., *Mutations in DPC4 (SMAD4) cause juvenile polyposis syndrome, but only account for a minority of cases.* Human Molecular Genetics, 1998. 7(12): p. 1907-1912.
134. Howe, J.R., et al., *Germline mutations of the gene encoding bone morphogenetic protein receptor 1A in juvenile polyposis.* Nature Genetics, 2001. 28(2): p. 184-187.
135. Sayed, M.G., et al., *Germline SMAD4 or BMPR1A mutations and phenotype of juvenile polyposis.* Annals of Surgical Oncology, 2002. 9(9): p. 901-906.
136. Gatalica, Z. and E. Torlakovic, *Pathology of the hereditary colorectal carcinoma.* Familial Cancer, 2008. 7(1): p. 15-26.
137. Popat, S. and R.S. Houlston, *A systematic review and meta-analysis of the relationship between chromosome 18q genotype, DCC status and colorectal cancer prognosis.* European Journal of Cancer, 2005. 41(14): p. 2060-2070.
138. Boland, C.R., et al., *A National Cancer Institute Workshop on Microsatellite Instability for Cancer Detection and Familial Predisposition: Development of International Criteria for the Determination of Microsatellite Instability in Colorectal Cancer.* Cancer Research, 1998. 58(22): p. 5248-5257.
139. Huang, J., et al., *APC mutations in colorectal tumors with mismatch repair deficiency.* Proceedings of the National Academy of Sciences of the United States of America, 1996. 93(17): p. 9049-9054.
140. Konishi, M., et al., *Molecular nature of colon tumors in hereditary nonpolyposis colon cancer, familial polyposis, and sporadic colon cancer.* Gastroenterology, 1996. 111(2): p. 307-317.
141. Miyaki, M., et al., *Frequent mutation of ss-catenin and APC genes in primary colorectal tumors from patients with hereditary nonpolyposis colorectal cancer.* Cancer Research, 1999. 59(18): p. 4506-4509.
142. Fujiwara, T., et al., *Accumulated clonal genetic alterations in familial and sporadic colorectal carcinomas with widespread instability in microsatellite sequences.* American Journal of Pathology, 1998. 153(4): p. 1063-1078.
143. Eshleman, J.R., et al., *Chromosome number and structure both are markedly stable in RER colorectal cancers and are not destabilized by mutation of p53.* Oncogene, 1998. 17(6): p. 719-725.
144. Yamamoto, H., H. Sawai, and M. Perucho, *Frameshift somatic mutations in gastrointestinal cancer of the microsatellite mutator phenotype.* Cancer Research, 1997. 57(19): p. 4420-4426.
145. Perucho, M., *Cancer of the microsatellite mutator phenotype.* Biological Chemistry, 1996. 377(11): p. 675-684.
146. Yamamoto, H., et al., *Somatic frameshift mutations in DNA mismatch repair and proapoptosis genes in hereditary nonpolyposis colorectal cancer.* Cancer Research, 1998. 58(5): p. 997-1003.
147. Markowitz, S., et al., *Inactivation of the Type-Ii Tgf-Beta Receptor in Colon-Cancer Cells with Microsatellite Instability.* Science, 1995. 268(5215): p. 1336-1338.
148. Piao, Z., et al., *Frequent frameshift mutations of RIZ in sporadic gastrointestinal and endometrial carcinomas with microsatellite instability.* Cancer Research, 2000. 60(17): p. 4701-4704.

149. Mori, Y., et al., *Instabilotyping: Comprehensive identification of frameshift mutations caused by coding region microsatellite instability.* Cancer Research, 2001. 61(16): p. 6046-6049.
150. Vasen, H.F.A., et al., *New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch Syndrome) proposed by the International Collaborative Group n HNPCC.* Gastroenterology, 1999. 116(6): p. 1453-1456.
151. Vasen, H.F.A., et al., *The International-Collaborative-Group-on-Hereditary-Non-Polyposis-Colorectal -Cancer (Icg-Hnpcc).* Diseases of the Colon & Rectum, 1991. 34(5): p. 424-425.
152. RodriguezBigas, M.A., et al., *A National Cancer Institute workshop on hereditary nonpolyposis colorectal cancer syndrome: Meeting highlights and Bethesda guidelines.* Journal of the National Cancer Institute, 1997. 89(23): p. 1758-1762.
153. Umar, A., et al., *Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability.* Journal of the National Cancer Institute, 2004. 96(4): p. 261-268.
154. Aaltonen, L.A., et al., *Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease.* New England Journal of Medicine, 1998. 338(21): p. 1481-1487.
155. Lynch, H.T. and A. de la Chapelle, *Genomic medicine - Hereditary colorectal cancer.* New England Journal of Medicine, 2003. 348(10): p. 919-932.
156. Gryfe, R., et al., *Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer.* New England Journal of Medicine, 2000. 342(2): p. 69-77.
157. Popat, S., R. Hubner, and R.S. Houlston, *Systematic review of microsatellite instability and colorectal cancer prognosis.* Journal of Clinical Oncology, 2005. 23(3): p. 609-618.
158. Lanza, G., et al., *Immunohistochemical test for MLH1 and MSH2 expression predicts clinical outcome in stage II and III colorectal cancer patients.* Journal of Clinical Oncology, 2006. 24(15): p. 2359-2367.
159. Carethers, J.M., et al., *Use of 5-fluorouracil and survival in patients with microsatellite-unstable colorectal cancer.* Gastroenterology, 2004. 126(2): p. 394-401.
160. Arnold, C.N., A. Goel, and C.R. Boland, *Role of hMLH1 promoter hypermethylation in drug resistance to 5-fluorouracil in colorectal cancer cell lines.* International Journal of Cancer, 2003. 106(1): p. 66-73.
161. Ribic, C.M., et al., *Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer.* New England Journal of Medicine, 2003. 349(3): p. 247-257.
162. Macabeo-Ong M, et al., *Effect of duration of fixation on quantitative reverse transcription polymerase chain reaction analyses.* Mod Pathol, 2002. 15(9): p. 979-87.
163. Bresters D, et al., *The duration of fixation influences the yield of HCV cDNA-PCR products from formalin-fixed, paraffin-embedded liver tissue.* J Virol Methods, 1994. 48(2-3): p. 267-272.
164. Cronin M, et al., *Measurement of gene expression in archival paraffin-embedded tissues: development and performance of a 92-gene reverse*

transcriptase-polymerase chain reaction assay. Am J Pathol, 2004. 164(1): p. 35-42.
165. Blenkiron, C., et al., *MicroRNA expression profiling of human breast cancer identifies new markers of tumor subtype.* Genome Biology, 2007. 8(10).
166. Jiang, J.M., et al., *Real-time expression profiling of microRNA precursors in human cancer cell lines.* Nucleic Acids Research, 2005. 33(17): p. 5394-5403.
167. Xi, Y.G., et al., *Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples.* Rna-a Publication of the Rna Society, 2007. 13: p. 1668-1674.
168. Bartlett J and Stirling D, *A Short History of the Polymerase Chain Reaction.* Methods Mol Biol, 2003. 226: p. 3-6.
169. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. 406(6797): p. 747-752.
170. Sorlie, T., et al., *Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications.* Proceedings of the National Academy of Sciences of the United States of America, 2001. 98(19): p. 10869-10874.
171. Kitahara O, et al., *Alterations of gene expression during colorectal carcinogenesis revealed by cDNA microarrays after laser-capture microdissection of tumor tissues and normal epithelia.* Cancer Res, 2001. 61(9): p. 3544-3549.
172. Lin YM, et al., *Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas.* Oncogene, 2002. 21(26): p. 4120-4128.
173. Zou, T.T., et al., *Application of cDNA microarrays to generate a molecular taxonomy capable of distinguishing between colon cancer and normal colon.* Oncogene, 2002. 21(31): p. 4855-4862.
174. Kwon, H.C., et al., *Gene expression profiling in lymph node-positive and lymph node-negative colorectal cancer.* Diseases of the Colon & Rectum, 2004. 47(2): p. 141-152.
175. Friederichs J, et al., *Gene expression profiles of different clinical stages of colorectal carcinoma: toward a molecular genetic understanding of tumor progression.* Int J Colorectal Dis, 2005. 20(5): p. 391-402.
176. Kim, K., et al., *Gene profiling of colonic serrated adenomas by using oligonucleotide microarray.* International Journal of Colorectal Disease, 2008. 23(6): p. 569-580.
177. Kita, H., et al., *Differential gene expression between flat adenoma and normal mucosa in the colon in a microarray analysis.* Journal of Gastroenterology, 2006. 41(11): p. 1053-1063.
178. Agrawal, D., et al., *Osteopontin identified as lead marker of colon cancer progression, using pooled sample expression profiling.* Journal of the National Cancer Institute, 2002. 94(7): p. 513-521.
179. Koehler, A., et al., *Gene expression profiling of colorectal cancer and metastases divides tumours according to their clinicopathological stage.* Journal of Pathology, 2004. 204(1): p. 65-74.
180. Yanagawa, R., et al., *Genome-wide screening of genes showing altered expression in liver metastases of human colorectal cancers by cDNA microarray.* Neoplasia, 2001. 3(5): p. 395-401.

181. Ki DH, et al., *Whole genome analysis for liver metastasis gene signatures in colorectal cancer.* Int J Cancer, 2007. 121(9): p. 2005-2012.
182. Lin, H.M., et al., *Genome wide expression profiling identifies genes associated with colorectal liver metastasis.* Oncology Reports, 2007. 17(6): p. 1541-1549.
183. Watanabe, T., et al., *Gene Expression Signature and the Prediction of Lymph Node Metastasis in Colorectal Cancer by DNA Microarray.* Diseases of the Colon & Rectum, 2009. 52(12): p. 1941-1948.
184. Wang Y, et al., *Gene expression profiles and molecular markers to predict recurrence of Dukes' B colon cancer.* J C Oncol, 2004. 22(9): p. 1564-1571.
185. Birkenkamp-Demtroder, K., et al., *Differential gene expression in colon cancer of the caecum versus the sigmoid and rectosigmoid.* Gut, 2005. 54(3): p. 374-384.
186. Ghadimi, B.M., et al., *Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy.* Journal of Clinical Oncology, 2005. 23(9): p. 1826-1838.
187. Kim, I.J., et al., *Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer.* Diseases of the Colon & Rectum, 2007. 50(9): p. 1342-1353.
188. Mariadason, J.M., et al., *Gene expression profiling-based prediction of response of colon carcinoma cells to 5-fluorouracil and camptothecin.* Cancer Research, 2003. 63(24): p. 8791-8812.
189. Del Rio, M., et al., *Gene expression signature in advanced colorectal cancer patients select drugs and response for the use of leucovorin, fluorouracil, and irinotecan.* Journal of Clinical Oncology, 2007. 25(7): p. 773-780.
190. Khambata-Ford, S., et al., *Expression of epiregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab.* Journal of Clinical Oncology, 2007. 25(22): p. 3230-3237.
191. Kim IJ, et al., *Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer.* Dis Colon Rectum., 2007. 50(9): p. 1342-1353.
192. Rimkus, C., et al., *Microarray-based prediction of tumor response to neoadjuvant radiochernotherapy of patients with locally advanced rectal cancer.* Clinical Gastroenterology and Hepatology, 2008. 6(1): p. 53-61.
193. Spitzner M, et al., *A gene expression signature for chemoradiosensitivity of colorectal cancer cells.* Int J Radiat Oncol Biol Phys, 2010. 78(4): p. 1184-92.
194. Cunningham, D., et al., *Cetuximab monotherapy and cetuximab plus irinotecan in irinotecan-refractory metastatic colorectal cancer.* New England Journal of Medicine, 2004. 351(4): p. 337-345.
195. Hurwitz, H., et al., *Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer.* New England Journal of Medicine, 2004. 350(23): p. 2335-2342.

196. Mayer, R.J., *Targeted Therapy for Advanced Colorectal Cancer-More Is Not Always Better.* New England Journal of Medicine, 2009. 360(6): p. 623-625.
197. Saltz, L.B., et al., *Phase II trial of cetuximab in patients with refractory colorectal cancer that expresses the epidermal growth factor receptor.* Journal of Clinical Oncology, 2004. 22(7): p. 1201-1208.
198. Waldner MJ and Neurath MF, *The molecular therapy of colorectal cancer.* Mol Aspects Med, 2010. 31(2): p. 171-8.
199. Birkenkamp-Demtroder, K., et al., *Gene expression in colorectal cancer.* Cancer Research, 2002. 62(15): p. 4352-4363.
200. Grade M, et al., *Gene expression profiling reveals a massive, aneuploidy-dependent transcriptional deregulation and distinct differences between lymph node-negative and lymph node-positive colon carcinomas.* Cancer Res, 2007. 67(1): p. 41-56.
201. Pei H, et al., *Proteome analysis and tissue microarray for profiling protein markers associated with lymph node metastasis in colorectal cancer.* J Proteome Res., 2007. 6(7): p. 2495-2501.
202. Lin, Y.H., et al., *Multiple gene expression classifiers from different array platforms predict poor prognosis of colorectal cancer.* Clinical Cancer Research, 2007. 13(2): p. 498-507.
203. Bianchini M, et al., *Comparative study of gene expression by cDNA microarray in human colorectal cancer tissues and normal mucosa.* Int J Oncol, 2006. 29(1): p. 83-94.
204. Barrier A, et al., *Colon cancer prognosis prediction by gene expression profiling.* Oncogene, 2005. 24(40): p. 6155-6164.
205. Colliver DW, et al., *Molecular profiling of ulcerative colitis-associated neoplastic progression.* Exp Mol Pathol, 2006. 80(1): p. 1-10.
206. Watanabe, T., et al., *Prediction of sensitivity of rectal cancer cells in response to preoperative radiotherapy by DNA microarray analysis of gene expression profiles.* Cancer Research, 2006. 66(7): p. 3370-3374.
207. Chiu, C.T., et al., *Clinicopathological analysis of colorectal cancer liver metastasis and intrahepatic cholangiocarcinoma: Are they just apples and oranges?* Digestive and Liver Disease, 2008. 40(9): p. 749-754.
208. Ieta, K., et al., *Biological and genetic characteristics of tumor-initiating cells in colon cancer.* Annals of Surgical Oncology, 2008. 15(2): p. 638-648.
209. Mazzanti, R., et al., *Differential expression proteomics of human colon cancer.* American Journal of Physiology-Gastrointestinal and Liver Physiology, 2006. 290(6): p. G1329-G1338.
210. Wiese, A.H., et al., *Identification of gene signatures for invasive colorectal tumor cells.* Cancer Detection and Prevention, 2007. 31(4): p. 282-295.
211. Varghese S, et al., *Site-specific gene expression profiles and novel molecular prognostic factors in patients with lower gastrointestinal adenocarcinoma diffusely metastatic to liver or peritoneum.* Ann Surg Oncol., 2007. 14(12): p. 3460-3471.
212. Gao, X.Q., et al., *Identification of the differential expressive tumor associated genes in rectal cancers by cDNA microarray.* World Journal of Gastroenterology, 2007. 13(3): p. 341-348.

213. Maglietta, R., et al., *Selection of relevant genes in cancer diagnosis based on their prediction accuracy.* Artificial Intelligence in Medicine, 2007. 40(1): p. 29-44.
214. Galamb, O., et al., *mRNA expression, functional profiling and multivariate classification of colon biopsy specimen by cDNA overall glass microarray.* World Journal of Gastroenterology, 2006. 12(43): p. 6998-7006.
215. Croner, R.S., et al., *Molecular signature for lymphatic metastasis in colorectal carcinomas.* Annals of Surgery, 2008. 247(5): p. 803-810.
216. Lee, R.C., R.L. Feinbaum, and V. Ambros, *The C-Elegans Heterochronic Gene Lin-4 Encodes Small Rnas with Antisense Complementarity to Lin-14.* Cell, 1993. 75(5): p. 843-854.
217. Reinhart, B.J., et al., *The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans.* Nature, 2000. 403(6772): p. 901-906.
218. Brennecke, J., et al., *bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila.* Cell, 2003. 113(1): p. 25-36.
219. Lagos-Quintana, M., et al., *Identification of tissue-specific microRNAs from mouse.* Current Biology, 2002. 12(9): p. 735-739.
220. Calin, G.A., et al., *Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia.* Proceedings of the National Academy of Sciences of the United States of America, 2002. 99(24): p. 15524-15529.
221. Hannon, G.J., *RNA interference.* Nature, 2002. 418(6894): p. 244-251.
222. Ambros, V., *The functions of animal microRNAs.* Nature, 2004. 431(7006): p. 350-355.
223. Lee, Y., et al., *MicroRNA maturation: stepwise processing and subcellular localization.* Embo Journal, 2002. 21(17): p. 4663-4670.
224. Lee, Y., et al., *The nuclear RNase III Drosha initiates microRNA processing.* Nature, 2003. 425(6956): p. 415-419.
225. Yi, R., et al., *Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs.* Genes & Development, 2003. 17(24): p. 3011-3016.
226. Bohnsack, M.T., K. Czaplinski, and D. Gorlich, *Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs.* Rna-a Publication of the Rna Society, 2004. 10(2): p. 185-191.
227. Hutvagner, G., et al., *A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA.* Science, 2001. 293(5531): p. 834-838.
228. Miranda KC, et al., *A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes.* Cell, 2006. 126(6): p. 1203-1217.
229. Bartel, D.P., *MicroRNAs: Genomics, biogenesis, mechanism, and function.* Cell, 2004. 116(2): p. 281-297.
230. Dugas, D.V. and B. Bartel, *MicroRNA regulation of gene expression in plants.* Current Opinion in Plant Biology, 2004. 7(5): p. 512-520.
231. Liu, J.D., et al., *Argonaute2 is the catalytic engine of mammalian RNAi.* Science, 2004. 305(5689): p. 1437-1441.

232. Meister, G., et al., *Human Argonaute2 mediates RNA cleavage targeted by miRNAs and siRNAs.* Molecular Cell, 2004. 15(2): p. 185-197.
233. Lee, Y.S., et al., *Depletion of human micro-RNA miR-125b reveals that it is critical for the proliferation of differentiated cells but not for the downregulation of putative targets during differentiation.* Journal of Biological Chemistry, 2005. 280(17): p. 16635-16641.
234. Carmell, M.A., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis.* Genes & Development, 2002. 16(21): p. 2733-2742.
235. Karube, Y., et al., *Reduced expression of Dicer associated with poor prognosis in lung cancer patients.* Cancer Science, 2005. 96(2): p. 111-115.
236. Saito, Y., et al., *Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells.* Cancer Cell, 2006. 9(6): p. 435-443.
237. Costinean, S., et al., *Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E mu-miR155 transgenic mice.* Proceedings of the National Academy of Sciences of the United States of America, 2006. 103(18): p. 7024-7029.
238. Cimmino, A., et al., *miR-15 and miR-16 induce apoptosis by targeting BCL2.* Proceedings of the National Academy of Sciences of the United States of America, 2005. 102(39): p. 13944-13949.
239. Borkhardt, A., U. Fuchs, and T. Tuschl, *MicroRNA in chronic lymphocytic leukemia.* New England Journal of Medicine, 2006. 354(5): p. 524-524.
240. Fulci, V., et al., *Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia.* Blood, 2007. 109(11): p. 4944-4951.
241. Linsley, P.S., et al., *Transcripts targeted by the MicroRNA-16 family cooperatively regulate cell cycle progression.* Molecular and Cellular Biology, 2007. 27(6): p. 2240-2252.
242. Iorio, M.V., et al., *MicroRNA gene expression deregulation in human breast cancer.* Cancer Research, 2005. 65(16): p. 7065-7070.
243. Ma, L., J. Teruya-Feldstein, and R.A. Weinberg, *Tumour invasion and metastasis initiated by microRNA 10b in breast cancer.* Nature, 2007. 449(7163): p. 682-U2.
244. Frankel, L.B., et al., *Programmed cell death 4 (PDCD4) is an important functional target of the microRNA miR-21 in breast cancer cells.* Journal of Biological Chemistry, 2008. 283(2): p. 1026-1033.
245. Zhu, S.M., et al., *MicroRNA-21 targets the tumor suppressor gene tropomyosin 1 (TPM1).* Journal of Biological Chemistry, 2007. 282(19): p. 14328-14336.
246. Si, M.L., et al., *miR-21-mediated tumor growth.* Oncogene, 2007. 26(19): p. 2799-2803.
247. Mattie, M.D., et al., *Optimized high-throughput microRNA expression profiling provides novel biomarker assessment of clinical prostate and breast cancer biopsies.* Molecular Cancer, 2006. 5.
248. Lowery, A.J., et al., *MicroRNA signatures predict oestrogen receptor, progesterone receptor and HER2/neu receptor status in breast cancer.* Breast Cancer Research, 2009. 11(3).

249. Adams, B.D., H. Furneaux, and B.A. White, *The micro-ribonucleic acid (miRNA) miR-206 targets the human estrogen receptor-alpha (ER alpha) and represses ER alpha messenger RNA and protein expression in breast cancer cell lines.* Molecular Endocrinology, 2007. 21(5): p. 1132-1147.
250. Kondo, N., et al., *MicroRNA-206 expression is down-regulated in estrogen receptor a-positive human breast cancer.* Cancer Research, 2009. 69(2): p. 217S-217S.
251. Heneghan, H.M., et al., *Systemic miRNA-195 Differentiates Breast Cancer from Other Malignancies and Is a Potential Biomarker for Detecting Noninvasive and Early Stage Disease.* Oncologist. 15(7): p. 673-682.
252. Iorio, M.V., et al., *MicroRNA profiling as a tool to understand prognosis, therapy response and resistance in breast cancer.* European Journal of Cancer, 2008. 44(18): p. 2753-2759.
253. Wu, W.K.K., et al., *MicroRNA dysregulation in gastric cancer: a new player enters the game.* Oncogene. 29(43): p. 5761-5771.
254. Tsujiura, M., et al., *Circulating microRNAs in plasma of patients with gastric cancers.* British Journal of Cancer. 102(7): p. 1174-1179.
255. Zhou, H., et al., *Detection of circulating tumor cells in peripheral blood from patients with gastric cancer using microRNA as a marker.* Journal of Molecular Medicine-Jmm, 2010. 88(7): p. 709-717.
256. Arisawa, T., et al., *A polymorphism of microRNA 27a genome region is associated with the development of gastric mucosal atrophy in Japanese male subjects.* Digestive Diseases and Sciences, 2007. 52(7): p. 1691-1697.
257. Peng, S., et al., *Association of MicroRNA-196a-2 Gene Polymorphism with Gastric Cancer Risk in a Chinese Population.* Digestive Diseases and Sciences, 2010. 55(8): p. 2288-2293.
258. Li, X.H., et al., *Survival prediction of gastric cancer by a seven-microRNA signature.* Gut. 59(5): p. 579-585.
259. Katada, T., et al., *microRNA expression profile in undifferentiated gastric cancer.* International Journal of Oncology, 2009. 34(2): p. 537-542.
260. Bandres, E., et al., *microRNA-451 Regulates Macrophage Migration Inhibitory Factor Production and Proliferation of Gastrointestinal Cancer Cells.* Clinical Cancer Research, 2009. 15(7): p. 2281-2290.
261. Tie, J., et al., *MiR-218 Inhibits Invasion and Metastasis of Gastric Cancer by Targeting the Robo1 Receptor.* Plos Genetics. 6(3).
262. Ueda, T., et al., *Relation between microRNA expression and progression and prognosis of gastric cancer: a microRNA expression analysis.* Lancet Oncology. 11(2): p. 136-146.
263. Michael, M.Z., et al., *Reduced accumulation of specific microRNAs in colorectal neoplasia.* Molecular Cancer Research, 2003. 1(12): p. 882-891.
264. Schetter, A.J., et al., *MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma.* Jama-Journal of the American Medical Association, 2008. 299(4): p. 425-436.
265. Akao, Y., Y. Nakagawa, and T. Naoe, *MicroRNA-143 and-145 in colon cancer.* DNA and Cell Biology, 2007. 26(5): p. 311-320.

266. Braun, C.J., et al., *p53-Responsive MicroRNAs 192 and 215 Are Capable of Inducing Cell Cycle Arrest.* Cancer Research, 2008. 68(24): p. 10094-10104.
267. He, X.Y., L. He, and G.J. Hannon, *The guardian's little helper: MicroRNAs in the p53 tumor suppressor network.* Cancer Research, 2007. 67(23): p. 11099-11101.
268. Wang, P., et al., *microRNA-21 Negatively Regulates Cdc25A and Cell Cycle Progression in Colon Cancer Cells.* Cancer Research, 2009. 69(20): p. 8157-8165.
269. Asangani, I.A., et al., *MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer.* Oncogene, 2008. 27(15): p. 2128-2136.
270. Nagel, R., et al., *Regulation of the adenomatous polyposis coli gene by the miR-135 family in colorectal cancer.* Cancer Research, 2008. 68(14): p. 5795-5802.
271. Guo, C.G., et al., *The noncoding RNA, miR-126, suppresses the growth of neoplastic cells by targeting phosphatidylinositol 3-kinase signaling and is frequently lost in colon cancers.* Genes Chromosomes & Cancer, 2008. 47(11): p. 939-946.
272. Bandres, E., et al., *Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues.* Molecular Cancer, 2006. 5.
273. Xi Y, et al., *Prognostic Values of microRNAs in Colorectal Cancer.* Biomark Insights, 2006. 2: p. 113-121.
274. Arndt, G.M., et al., *Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer.* Bmc Cancer, 2009. 9.
275. Motoyama, K., et al., *Over- and under-expressed microRNAs in human colorectal cancer.* International Journal of Oncology, 2009. 34(4): p. 1069-1075.
276. Wang, C.J., et al., *Clinicopathological significance of microRNA-31,-143 and-145 expression in colorectal cancer.* Disease Markers, 2009. 26(1): p. 27-34.
277. Lanza, G., et al., *mRNA/microRNA gene expression profile in microsatellite unstable colorectal cancer.* Molecular Cancer, 2007. 6.
278. Kloosterman, W.P., et al., *Targeted inhibition of miRNA maturation with morpholinos reveals a role for miR-375 in pancreatic islet development.* Plos Biology, 2007. 5(8): p. 1738-1749.
279. Hutvagner, G. and P.D. Zamore, *A microRNA in a multiple-turnover RNAi enzyme complex.* Science, 2002. 297(5589): p. 2056-2060.
280. Elmen, J., et al., *Antagonism of microRNA-122 in mice by systemically administered LNA-antimiR leads to up-regulation of a large set of predicted target mRNAs in the liver.* Nucleic Acids Research, 2008. 36(4): p. 1153-1162.
281. Krutzfeldt, J., et al., *Silencing of microRNAs in vivo with 'antagomirs'.* Nature, 2005. 438(7068): p. 685-689.
282. Brosens, L.A.A., et al., *Gastrointestinal polyposis syndromes.* Current Molecular Medicine, 2007. 7(1): p. 29-46.

283. Lu, Z., et al., *MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene.* Oncogene, 2008. 27(31): p. 4373-4379.
284. Dews, M., et al., *Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster.* Nature Genetics, 2006. 38(9): p. 1060-1065.
285. Rossi L, Bonmassar E, and Faraoni I, *Modification of miR gene expression pattern in human colon cancer cells following exposure to 5-fluorouracil in vitro.* Pharmacological Research, 2007. 56(3): p. 248-253.
286. Meng, F.Y., et al., *Involvement of human micro-RNA in growth and response to chemotherapy in human cholangiocarcinoma cell lines.* Gastroenterology, 2006. 130(7): p. 2113-2129.
287. MMI. *INSTRUCTIONS: Fifth and Sixth Drafts Application Form12 for the Ethical Review of Health Related Research Studies which are not interventional clinical trials of medicinal products for human use as defined in Statutory Instrument 190/2004 IT.* http://www.moleculармedicineireland.ie/page/g/s/82 2010 [cited.
288. Sobin LH, Wittekind C, and (eds). *TNM: Classification of Malignant Tumours*. 6 ed. 2002, New York, NY: Wiley-Liss.
289. Compton CC and Greene FL, *The Staging of Colorectal Cancer: 2004 and Beyond.* CA Cancer J Clin, 2004. 54(6): p. 295-308.
290. Greene FL, et al., *AJCC Cancer Staging Manual*. 6 ed. 2002, New York, NY: Springer.
291. Birney E, et al., *Ensembl 2006.* Nucleic Acids Res, 2006. 34(Database issue): p. D556-561.
292. Kheirelseid EA, et al., *Identification of endogenous control genes for normalisation of real-time quantitative PCR data in colorectal cancer.* BMC Mol Biol, 2010. 11(1): p. 12.
293. Chang KH, et al., *MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer.* BMC Cancer, 2010. 10: p. 173.
294. Lancashire, L.J., R.C. Rees, and G.R. Ball, *Identification of gene transcript signatures predictive for estrogen receptor and lymph node status using a stepwise forward selection artificial neural network modelling approach.* Artificial Intelligence in Medicine, 2008. 43(2): p. 99-111.
295. Griffiths-Jones S, et al., *miRBase: tools for microRNA genomics.* Nucleic Acids Res, 2008. 36(Database issue): p. D154-8.
296. Griffiths-Jones S, *The microRNA Registry.* Nucleic Acids Res, 2004. 32(Database issue): p. D109-111.
297. Griffiths-Jones S, et al., *miRBase: microRNA sequences, targets and gene nomenclature.* Nucleic Acids Res, 2006. 34(Database issue): p. D140-144.
298. Wuchty S, et al., *Complete suboptimal folding of RNA and the stability of secondary structures.* Biopolymers, 1999. 49(2): p. 145-165.
299. Lewis BP, Burge CB, and Bartel DP, *Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets.* Cell, 2005. 120(1): p. 15-20.
300. Lewis BP, et al., *Prediction of mammalian microRNA targets.* Cell, 2003. 115(7): p. 787-798.
301. Grimson A, et al., *MicroRNA targeting specificity in mammals: determinants beyond seed pairing.* Mol Cell, 2007. 27(1): p. 91-105.

302. Krek A, et al., *Combinatorial microRNA target predictions.* Nat Genet, 2005. 37(5): p. 495-500.
303. Wang X and El Naqa IM, *Prediction of both conserved and nonconserved microRNA targets in animals.* Bioinformatics 2008. 24(3): p. 325-332.
304. Wang X, *miRDB: a microRNA target prediction and functional annotation database with a wiki interface.* RNA, 2008. 14(6): p. 1012-1017
305. Enright AJ, et al., *MicroRNA targets in Drosophila.* Genome Biol, 2003. 5(1): p. R1.
306. John B, et al., *Human MicroRNA targets.* PLoS Biol, 2004. 2(11): p. e363.
307. Kiriakidou M, et al., *A combined computational-experimental approach predicts human microRNA targets.* Genes Dev, 2004. 18(10): p. 1165-1178.
308. Rehmsmeier M, et al., *Fast and effective prediction of microRNA/target duplexes.* RNA, 2004. 10(10): p. 1507-1517.
309. Grün D, et al., *microRNA target predictions across seven Drosophila species and comparison to mammalian targets.* PLoS Comput Biol., 2005. 1(1): p. e13.
310. Lall S, et al., *A genome-wide map of conserved microRNA targets in C. elegans.* Curr Biol. , 2006. 16(5): p. 460-471.
311. Chan CS, Elemento O, and Tavazoie S, *Revealing posttranscriptional regulatory elements through network-level conservation.* PLoS Comput Biol., 2005. 1(17): p. e69.
312. Watanabe Y, Tomita M, and Kanai A, *Computational methods for microRNA target prediction.* Methods Enzymol, 2007. 427: p. 65-86.
313. Lin SY, et al., *The C. elegans hunchback Homolog, hbl-1, Controls Temporal Patterning and Is a Probable MicroRNA Target.* Developmental Cell, 2003. 4(5): p. 639–650.
314. Cohen J, *Statistical power analysis for the behavioral sciences* 2nd ed. 1988, Hillsdale, NJ: Lawrence Erlbaum Associates.
315. Fearon, E.R. and B. Vogelstein, *A genetic model for colorectal tumorigenesis.* Cell, 1990. 61(5): p. 759-67.
316. de Kok, J.B., et al., *Normalization of gene expression measurements in tumor tissues: comparison of 13 endogenous control genes.* Lab Invest. , 2005. 85(1): p. 154-159.
317. Rubie, C., et al., *Housekeeping gene variability in normal and cancerous colorectal, pancreatic, esophageal, gastric and hepatic tissues.* Mol Cell Probes. , 2005. 19(2): p. 101-109.
318. Dydensborg, A.B., et al., *Normalizing genes for quantitative RT-PCR in differentiating human intestinal epithelial cells and adenocarcinomas of the colon.* Am J Physiol Gastrointest Liver Physiol. , 2006. 290(5): p. 1067-1074.
319. Arya, M., et al., *Basic principles of real-time quantitative PCR.* BioTechniques 2005. 5(2): p. 209-219.
320. Nolan, T., R.E. Hands, and S.A. Bustin, *Quantification of mRNA using real-time RT-PCR.* Nat. Protoc., 2006. 1: p. 1559–1582.
321. Vandesompele, J., et al., *Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes.* Genome Biol, 2002. 3(7): p. RESEARCH0034.

322. McNeill, R.E., N. Miller, and M.J. Kerin, *Evaluation and validation of candidate endogenous control genes for real-time quantitative PCR studies of breast cancer.* BMC Mol Biol 2007, 8:107, 2007. 8: p. 107.
323. Davoren, P.A., et al., *Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer.* BMC Molecular Biology, 2008. 9: p. 76.
324. Lyng, M.B., et al., *Identification of genes for normalization of real-time RT-PCR data in breast carcinomas.* BMC Cancer, 2008. 8: p. 20.
325. Saviozzi, S., et al., *Selection of suitable reference genes for accurate normalization of gene expression profile studies in non-small cell lung cancer.* BMC Cancer, 2006. 6: p. 200.
326. Coulson, D.T., et al., *Identification of valid reference genes for the normalization of RT qPCR gene expression data in human brain tissue.* BMC Mol Biol 2007, 8:107, 2008. 9: p. 46.
327. Cicinnati, V.R., et al., *Validation of putative reference genes for gene expression studies in human hepatocellular carcinoma using real-time quantitative RT-PCR.* BMC Cancer, 2008. 8: p. 350.
328. Zhong, H. and J. Simons, *Direct comparison of GAPDH, beta-actin, cyclophilin, and 28S rRNA as internal standards for quantifying RNA levels under hypoxia.* Biochem Biophys Res Commun, 1999. 259(3): p. 523–526.
329. Valenti, M.T., et al., *The effect of bisphosphonates on gene expression: GAPDH as a housekeeping or a new target gene?* BMC Cancer, 2006. 6: p. 49.
330. Kunth, K., H. Höfler, and M.J. Atkinson, *Quantification of messenger RNA expression in tumors: which standard should be used for best RNA normalization?* Verh Dtsch Ges Pathol. , 1994. 78: p. 226-230.
331. Goidin, D., et al., *Ribosomal 18S RNA prevails over glyceraldehyde-3-phosphate dehydrogenase and beta-actin genes as internal standard for quantitative comparison of mRNA levels in invasive and noninvasive human melanoma cell subpopulations.* Anal Biochem. , 2001. 295(1): p. 17-21.
332. Oikarinen, A., et al., *Comparison on collagen gene expression in the developing chick emberyotendon and heart. Tissue and development tissue-dependent action of dexamethasone.* Biochim Biophys Acta, 1991. 1089(1): p. 40-46.
333. Fedchenko, V., et al., *The effect of short-term administration of (-)-deprenyl and isatin on the expressions of some genes in the mouse brain cortex.* Med Sci Monit, 2008. 14(12): p. BR269-73.
334. Schmittgen, T.D. and B.A. Zakrajsek, *Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR.* J Biochem Biophys Methods. , 2000. 46(1-2): p. 69-81.
335. Wu, Y.Y. and J.L. Rees, *Variation in epidermal housekeeping gene expression in different pathological states.* Acta Derm Venereol. , 2000. 80(1): p. 2-3.
336. Ishitani , R., et al., *Evidence that glyceraldehyde-3-phosphate dehydrogenase is involved in age-induced apoptosis in mature cerebellar neurons in culture.* Journal of Neurochemistry, 1996. 66(3): p. 928-935.

337. Singh, R. and M.R. Green, *Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase.* Science, 1993. 259(5093): p. 365-368.
338. Bustin SA, *Quantification of mRNA using real-time reverse transcription PCR (RT–PCR): trends and problems.* J. Mol. Endocrinol, 2002. 29(1): p. 23-39.
339. Kubista M, et al., *The real-time polymerase chain reaction.* Mol Aspects Med, 2006. 27(2-3): p. 95-125.
340. Foss DL, Baarsch MJ, and Murtaugh MP, *Regulation of hypoxanthine phosphoribosyltransferase, glyceraldehyde-3-phosphate dehydrogenase and beta-actin mRNA expression in porcine immune cells and tissues.* Animal Biotechnology, 1998. 9(1): p. 67-78.
341. Suzuki T, Higgins PJ, and Crawford DR, *Control selection for RNA quantitation.* Biotechniques. , 2000. 29(2): p. 332-337.
342. Wei, Q., et al., *Expression of five selected human mismatch repair genes simultaneously detected in normal and cancer cell lines by a nonradioactive multiplex reverse transcription-polymerase chain reaction.* Pathobiology, 1997. 65(6): p. 293-300.
343. Akishima-Fukasawa Y, et al., *Prognostic significance of CXCL12 expression in patients with colorectal carcinoma.* Am J Clin Pathol, 2009. 132(2): p. 202-10.
344. Lawrie LC, et al., *Liver fatty acid binding protein expression in colorectal neoplasia.* Br J Cancer, 2004. 90(10): p. 1955-60.
345. Yang K, et al., *Interaction of Muc2 and Apc on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation.* Cancer Res., 2008. 68(18): p. 7313-22.
346. Wang, Q., Z. Sun, and H.S. Yang, *Downregulation of tumor suppressor Pdcd4 promotes invasion and activates both beta-catenin/Tcf and AP-1-dependent transcription in colon carcinoma cells.* Oncogene, 2008. 27(11): p. 1527-35.
347. Andersen, C.L., J.L. Jensen, and T.F. Orntoft, *Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets.* Cancer Res., 2004. 64(15): p. 5245-50.
348. Hellemans, J., et al., *qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data.* Genome Biology, 2007. 8(2): p. R19.
349. Haller, F., et al., *Equivalence test in quantitative reverse transcription polymerase chain reaction: confirmation of reference genes suitable for normalization.* Anal Biochem., 2004. 335(1): p. 1-9.
350. Huggett , J., et al., *Real-time RT-PCR normalisation; strategies and considerations.* Genes and immunity, 2005. 6(4): p. 279-284.
351. Tricarico, C., et al., *Quantitative real-time reverse transcription polymerase chain reaction: normalization to rRNA or single housekeeping genes is inappropriate for human tissue biopsies. .* Anal Biochem., 2002. 309(2): p. 293-300.
352. Szabo, A., et al., *Statistical modeling for selecting housekeeper genes.* Genome Biol, 2004. 5(8): p. R59.

353. Pfaffl, M.W., et al., *Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations.* . Biotechnol Lett, 2004. 26(6): p. 509-515.
354. Goossens, K., et al., *Selection of reference genes for quantitative real-time PCR in bovine preimplantation embryos.* BMC Dev Biol, 2005. 5: p. 27.
355. Wellek, S., *Testing statistical hypothesis of equivalence*. 2003, London: Chapman & Hall.
356. Blanquicett, C., et al., *Housekeeping gene variability in normal and carcinomatous colorectal and liver tissues: applications in pharmacogenomic gene expression studies.* Anal Biochem., 2002. 303(2): p. 209–214.
357. Galanis E, Alberts SR, and O'Connell MJ, *New adjuvant therapy for colon cancer: justified hope or commercial hype.* Surg Oncol Clin N Am, 2000. 9(4): p. 813-823.
358. Kapiteijn, E., et al., *Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer.* New England Journal of Medicine, 2001. 345(9): p. 638-646.
359. Forrester K, et al., *Detection of high incidence of K-ras oncogenes during human colon tumorigenesis.* Nature, 1987. 327(6120): p. 298-303.
360. Ionov Y, et al., *Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis.* Nature, 1993. 363(2429): p. 558-561.
361. Cui H, et al., *Loss of imprinting in normal tissue of colorectal cancer patients with microsatellite instability.* Nat Med, 1998. 4(11): p. 1276-1280.
362. Paik S, et al., *A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer.* N Engl J Med, 2004. 351(27): p. 2817-2826.
363. Nannini, M., et al., *Gene expression profiling in colorectal cancer using microarray technologies: Results and perspectives.* Cancer Treatment Reviews, 2009. 35(3): p. 201-209.
364. Gibson UE, Heid CA, and Williams PM, *A novel method for real time quantitative RT-PCR.* Genome Res, 1996. 6(10): p. 995-1001.
365. Chan, S.K., et al., *Meta-analysis of colorectal cancer gene expression profiling studies identifies consistently reported candidate biomarkers.* Cancer Epidemiology Biomarkers & Prevention, 2008. 17(3): p. 543-552.
366. Berndorff, D., et al., *Liver-Intestine Cadherin - Molecular-Cloning and Characterization of a Novel Ca2+-Dependent Cell-Adhesion Molecule Expressed in Liver and Intestine.* Journal of Cell Biology, 1994. 125(6): p. 1353-1369.
367. Gessner, R. and R. Tauber, *Intestinal cell adhesion molecules - Liver-intestine cadherin.* Epithelial Transport and Barrier Function, 2000. 915: p. 136-143.
368. Grötzinger C, et al., *LI-cadherin: a marker of gastric metaplasia and neoplasia.* Gut, 2001. 49(1): p. 73-81.
369. Park SS, et al., *Expression of liver-intestine cadherin and its correlation with lymph node metastasis in gastric cancer: can it predict N stage preoperatively?*

. Ann Surg Oncol., 2007. 14(1): p. 94-99.
370. Ko, S., et al., *CDX2 co-localizes with liver-intestine cadherin in intestinal metaplasia and adenocarcinoma of the stomach.* Journal of Pathology, 2005. 205(5): p. 615-622.
371. Wang, X.Q., et al., *Liver intestine-cadherin (CDH17) haplotype is associated with increased risk of hepatocellular carcinoma.* Clinical Cancer Research, 2006. 12(17): p. 5248-5252.
372. Takamura M, et al., *Expression of liver-intestine cadherin and its possible interaction with galectin-3 in ductal adenocarcinoma of the pancreas.* Cancer Science, 2003. 94(5): p. 425-430.
373. Takamura M, et al., *Reduced expression of liver-intestine cadherin is associated with progression and lymph node metastasis of human colorectal carcinoma.* Cancer Lett., 2004. 212(2): p. 253-259.
374. Kwak JM, et al., *The prognostic significance of E-cadherin and liver intestine-cadherin expression in colorectal cancer.* Dis Colon Rectum., 2007. 50(11): p. 1873-1880.
375. Xu D, et al., *Quantitative real-time RT-PCR detection for CEA, CK20 and CK19 mRNA in peripheral blood of colorectal cancer patients*.* J Zhejiang Univ Sci B, 2006. 7(6): p. 445-451.
376. Iinuma H, et al., *Usefulness and clinical significance of quantitative real-time RT-PCR to detect isolated tumor cells in the peripheral blood and tumor drainage blood of patients with colorectal cancer.* Int J Oncol, 2006. 28(2): p. 297-306.
377. Kollmar O, et al., *CXCR4 and CXCR7 regulate angiogenesis and CT26.WT tumor growth independent from SDF-1.* Int J Cancer, 2010. 126(6): p. 1302-1315.
378. Ingold B, et al., *The role of vascular CXCR4 expression in colorectal carcinoma.* Histopathology., 2009. 55(5): p. 576-586.
379. Meijer J, Ogink J, and Roos E, *Effect of the chemokine receptor CXCR7 on proliferation of carcinoma cells in vitro and in vivo.* Br J Cancer, 2008. 99(9): p. 1493-1501.
380. Birkenkamp-Demtroder K, et al., *Gene expression in colorectal cancer.* Cancer Res, 2002. 62(15): p. 4352-4363.
381. Xu WQ, et al., *Expression of TGF-β1, TβRII and Smad4 in colorectal carcinoma.* Exp Mol Pathol, 2007. 82(3): p. 284-291.
382. Xu Y and Pasche B, *TGF-beta signaling alterations and susceptibility to colorectal cancer.* Hum Mol Genet, 2007. 16(Spec No 1): p. R14-20.
383. Saltzman BS, et al., *Association of Genetic Variation in the Transforming Growth Factor B-1 Gene with Serum Levels and Risk of Colorectal Neoplasia.* Cancer Res, 2008. 68(4): p. 1236-1244.
384. Skoglund Lundin J, et al., *TGFBR1 variants TGFBR1(*)6A and Int7G24A are not associated with an increased familial colorectal cancer risk.* Br J Cancer, 2009. 100(10): p. 1674-1679.
385. Zeng Q, et al., *Tgfbr1 Haploinsufficiency Is a Potent Modifier of Colorectal Cancer Development.* Cancer Res, 2009. 69(2): p. 678-686.
386. Biswas S, et al., *Mutational inactivation of TGFBR2 in microsatellite unstable colon cancer arises from the cooperation of genomic instability and the clonal outgrowth of transforming growth factor beta resistant cells.* Genes Chromosomes Cancer, 2008. 47(2): p. 95-106.

387. Wharton, R.Q., et al., *Increased detection of circulating tumor cells in the blood of colorectal carcinoma patients using two reverse transcription-PCR assays and multiple blood samples.* Clinical Cancer Research, 1999. 5(12): p. 4158-4163.
388. Xu D, et al., *Quantitative real-time RT-PCR detection for CEA, CK20 and CK19 mRNA in peripheral blood of colorectal cancer patients.* J Zhejiang Univ Sci B, 2006. 7(6): p. 445-451.
389. Seruga, B., et al., *Cytokines and their relationship to the symptoms and outcome of cancer.* Nature Reviews Cancer, 2008. 8(11): p. 887-899.
390. Satake, M., et al., *The Comparative Study of Angiogenetic and Angiostatic CXC Chemokines in Human Pancreatic Cancer.* Pancreas, 2009. 38(8): p. 1043-1043.
391. Gerber, P.A., et al., *Chemokines in tumor-associated angiogenesis.* Biological Chemistry, 2009. 390(12): p. 1213-1223.
392. Brand, S., et al., *CXCR4 and CXCL12 are inversely expressed in colorectal cancer cells and modulate cancer cell migration, invasion and MMP-9 activation.* Experimental Cell Research, 2005. 310(1): p. 117-130.
393. Yoshitake, N., et al., *Expression of SDF-1 alpha and nuclear CXCR4 predicts lymph node metastasis in colorectal cancer.* British Journal of Cancer, 2008. 98(10): p. 1682-1689.
394. Kollmar, O., et al., *Stromal cell-derived factor-1 promotes cell migration and tumor growth of colorectal metastasis.* Neoplasia, 2007. 9: p. 862-870.
395. Burns, J.M., et al., *A novel chemokine receptor for SDF-1 and I-TAC involved in cell survival, cell adhesion, and tumor development.* Journal of Experimental Medicine, 2006. 203(9): p. 2201-2213.
396. Balabanian, K., et al., *The chemokine SDF-1/CXCL12 binds to and signals through the orphan receptor RDC1 in T lymphocytes.* Journal of Biological Chemistry, 2005. 280(42): p. 35760-35766.
397. Gordon, J.I., et al., *Tissue Specific Expression and Developmental Regulation of 2 Genes-Coding for Rat Fatty-Acid Binding-Proteins.* Journal of Biological Chemistry, 1985. 260(4): p. 1995-1998.
398. Carroll, S.L., K.A. Roth, and J.I. Gordon, *Liver Fatty-Acid Binding-Protein - a Marker for Studying Cellular-Differentiation in Gut Epithelial Neoplasms.* Gastroenterology, 1990. 99(6): p. 1727-1735.
399. Veerkamp, J.H., et al., *Structural and functional studies on different human FABP types.* Molecular and Cellular Biochemistry, 1999. 192(1-2): p. 137-142.
400. Storch, J. and A.E.A. Thumser, *The fatty acid transport function of fatty acid-binding proteins.* Biochimica Et Biophysica Acta-Molecular and Cell Biology of Lipids, 2000. 1486(1): p. 28-44.
401. Murphy, E.J., et al., *Liver fatty acid-binding protein expression in transfected fibroblasts stimulates fatty acid uptake and metabolism.* Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism, 1996. 1301(3): p. 191-198.
402. Sorof, S., *Modulation of Mitogenesis by Liver Fatty-Acid-Binding Protein.* Cancer and Metastasis Reviews, 1994. 13(3-4): p. 317-336.

403. Schroeder, F., et al., *Expression of liver fatty acid binding protein alters growth and differentiation of embryonic stem cells.* Molecular and Cellular Biochemistry, 2001. 219(1-2): p. 127-138.
404. Modi, W.S., et al., *Monocyte-Derived Neutrophil Chemotactic Factor (Mdncf/Il-8) Resides in a Gene-Cluster Along with Several Other Members of the Platelet Factor-Iv Gene Superfamily.* Human Genetics, 1990. 84(2): p. 185-187.
405. Huang, S.Y., et al., *Fully humanized neutralizing antibodies to interleukin-8 (ABX-IL8) inhibit angiogenesis, tumor growth, and metastasis of human melanoma.* American Journal of Pathology, 2002. 161(1): p. 125-134.
406. Brat, D.J., A.C. Bellail, and E.G. Van Meir, *The role of interleukin-8 and its receptors in gliomagenesis and tumoral angiogenesis.* Neuro-Oncology, 2005. 7(2): p. 122-133.
407. Fasciani, A., et al., *High concentrations of the vascular endothelial growth factor and interleukin-8 in ovarian endometriomata.* Molecular Human Reproduction, 2000. 6(1): p. 50-54.
408. Youngs, S.J., et al., *Chemokines induce migrational responses in human breast carcinoma cell lines.* International Journal of Cancer, 1997. 71(2): p. 257-266.
409. Berghella, A.M., et al., *Peripheral blood immunological parameters for use as markers of pre-invasive to invasive colorectal cancer.* Cancer Biotherapy and Radiopharmaceuticals, 2002. 17(1): p. 43-50.
410. Chang, S.K., et al., *Localization of Mucin (Muc2 and Muc3) Messenger-Rna and Peptide Expression in Human Normal Intestine and Colon-Cancer.* Gastroenterology, 1994. 107(1): p. 28-36.
411. Velcich, A., et al., *Colorectal cancer in mice genetically deficient in the mucin Muc2.* Science, 2002. 295(5560): p. 1726-1729.
412. Mizoshita, T., et al., *Loss of MUC2 expression correlates with progression along the adenoma-carcinoma sequence pathway as well as de novo carcinogenesis in the colon.* Histology and Histopathology, 2007. 22(3): p. 251-260.
413. Iwase, T., et al., *Overexpression of CD10 and reduced MUC2 expression correlate with the development and progression of colorectal neoplasms.* Pathology Research and Practice, 2005. 201(2): p. 83-91.
414. Wang, R.Q. and D.C. Fang, *Alterations of MUC1 and MUC3 expression in gastric carcinoma: relevance to patient clinicopathological features.* Journal of Clinical Pathology, 2003. 56(5): p. 378-384.
415. Walsh, M.D., et al., *Mucin Expression by Transitional-Cell Carcinomas of the Bladder.* British Journal of Urology, 1994. 73(3): p. 256-262.
416. Allgayer H, *Pdcd4, a colon cancer prognostic that is regulated by a microRNA.* Crit Rev Oncol Hematol, 2010. 73(3): p. 185-191.
417. Zakowicz, H., et al., *Mutational analysis of the DEAD-box RNA helicase eIF4AII characterizes its interaction with transformation suppressor Pdcd4 and eIF4GI.* Rna-a Publication of the Rna Society, 2005. 11(3): p. 261-274.
418. Chen, Y., et al., *Loss of PDCD4 expression in human lung cancer correlates with tumour progression and prognosis.* Journal of Pathology, 2003. 200(5): p. 640-646.

419. Bonome, T., et al., *Expression profiling of serous low malignant potential, low-grade, and high-grade tumors of the ovary.* Cancer Research, 2005. 65(22): p. 10602-10612.
420. Gao, F., et al., *Frequent loss of PDCD4 expression in human glioma: Possible role in the tumorigenesis of glioma.* Oncology Reports, 2007. 17(1): p. 123-128.
421. Massague, J., S.W. Blain, and R.S. Lo, *TGF beta signaling in growth control, cancer, and heritable disorders.* Cell, 2000. 103(2): p. 295-309.
422. Derynck, R., R.J. Akhurst, and A. Balmain, *TGF-beta signaling in tumor suppression and cancer progression.* Nature Genetics, 2001. 29(2): p. 117-129.
423. Wang D, et al., *Analysis of specific gene mutations in the transforming growth factor-beta signal transduction pathway in human ovarian cancer.* Cancer Research, 2000. 60(16): p. 4507-7512.
424. Turco A, et al., *Overexpression of transforming growth factor beta-type II receptor reduces tumorigenicity and metastastic potential of K-ras-transformed thyroid cells.* Int J Cancer, 1999. 80(1): p. 85-91.
425. Biswas, S., et al., *Transforming growth factor beta receptor type II inactivation promotes the establishment and progression of colon cancer.* Cancer Research, 2004. 64(14): p. 4687-4692.
426. Akhurst, R.J. and R. Derynck, *TGF-beta signaling in cancer - a double-edged sword.* Trends in Cell Biology, 2001. 11(11): p. S44-S51.
427. Rojas A, et al., *TGF-beta receptor levels regulate the specificity of signaling pathway activation and biological effects of TGF-beta.* Biochimica et Biophysica Acta, 2009. 1793(7): p. 1165-1173.
428. Xiong, B., et al., *Transforming growth factor-beta 1 in invasion and metastasis in colorectal cancer.* World Journal of Gastroenterology, 2002. 8(4): p. 674-678.
429. Engle, S.J., et al., *Transforming growth factor beta 1 suppresses nonmetastatic colon cancer at an early stage of tumorigenesis.* Cancer Research, 1999. 59(14): p. 3379-3386.
430. Kaklamani, V.G., et al., *TGFBR1*6A and cancer risk: A meta-analysis of seven case-control studies.* Journal of Clinical Oncology, 2003. 21(17): p. 3236-3243.
431. Bian, Y., et al., *TGFBR1*6A may contribute to hereditary colorectal cancer.* Journal of Clinical Oncology, 2005. 23(13): p. 3074-3078.
432. Leather, A.J.M., et al., *Detection and Enumeration of Circulating Tumor-Cells in Colorectal-Cancer.* British Journal of Surgery, 1993. 80(6): p. 777-780.
433. Chen, Y.H., et al., *Detection of bone marrow micrometastasis.* Hybridoma, 1999. 18(5): p. 465-466.
434. Rosenberg, R., et al., *Impact of cytokeratin-20 and carcinoembryonic antigen mRNA detection by RT-PCR in regional lymph nodes of patients with colorectal cancer.* British Journal of Cancer, 2000. 83(10): p. 1323-1329.
435. Giribaldi, G., et al., *Specific detection of cytokeratin 20-positive cells in blood of colorectal and breast cancer patients by a high sensitivity real-time reverse transcriptase-polymerase chain reaction method.* Journal of Molecular Diagnostics, 2006. 8(1): p. 105-112.

436. Wong, I.H.N., et al., *Quantitative relationship of the circulating tumor burden assessed by reverse transcription-polymerase chain reaction for cytokeratin 19 mRNA in peripheral blood of colorectal cancer patients with Dukes' stage, serum carcinoembryonic antigen level and tumor progression.* Cancer Letters, 2001. 162(1): p. 65-73.
437. Schuster, R., et al., *Quantitative real-time RT-PCR for detection of disseminated tumor cells in peripheral blood of patients with colorectal cancer using different mRNA markers.* International Journal of Cancer, 2004. 108(2): p. 219-227.
438. Gessner, R. and R. Tauber, *Intestinal cell adhesion molecules - Liver-intestine cadherin,* in *Epithelial Transport and Barrier Function - Pathomechanisms in Gastrointestinal Disorders,* J.D. Schulzke, et al., Editors. 2000. p. 136-143.
439. Kreft, B., et al., *LI-cadherin-mediated cell-cell adhesion does not require cytoplasmic interactions.* Journal of Cell Biology, 1997. 136(5): p. 1109-1121.
440. Ito, R., et al., *Clinicopathological significant and prognostic influence of cadherin-17 expression in gastric cancer.* Virchows Archiv, 2005. 447(4): p. 717-722.
441. Oue, N., et al., *Gene expression profile of gastric carcinoma: Identification of genes and tags potentially involved in invasion, metastasis, and carcinogenesis by serial analysis of gene expression.* Cancer Research, 2004. 64(7): p. 2397-2405.
442. Ko, S., et al., *Overexpression of LI-cadherin in gastric cancer is associated with lymph node metastasis.* Biochemical and Biophysical Research Communications, 2004. 319(2): p. 562-568.
443. Wang, X.Q., et al., *Alternative mRNA splicing of liver intestine-cadherin in hepatocellular carcinoma.* Clinical Cancer Research, 2005. 11(2): p. 483-489.
444. Hinoi, T., et al., *CDX2 regulates liver intestine-cadherin expression in normal and malignant colon epithelium and intestinal metaplasia.* Gastroenterology, 2002. 123(5): p. 1565-1577.
445. Hammarstrom, S., *The carcinoembryonic antigen (CEA) family: structures, suggested functions and expression in normal and malignant tissues.* Seminars in Cancer Biology, 1999. 9(2): p. 67-81.
446. Thompson, J.A., F. Grunert, and W. Zimmermann, *Carcinoembryonic Antigen Gene Family - Molecular-Biology and Clinical Perspectives.* Journal of Clinical Laboratory Analysis, 1991. 5(5): p. 344-366.
447. Minami, S., J. Furui, and T. Kanematsu, *Role of carcinoembryonic antigen in the progression of colon cancer cells that express carbohydrate antigen.* Cancer Research, 2001. 61(6): p. 2732-2735.
448. Blumenthal, R.D., H.J. Hansen, and D.M. Goldenberg, *Inhibition of adhesion, invasion, and metastasis by antibodies targeting CEACAM6 (NCA-90) and CEACAM5 (Carcinoembryonic antigen).* Cancer Research, 2005. 65(19): p. 8809-8817.
449. Govindan, S.V., et al., *Therapy of human colonic and lung cancer xenografts with SN-38 conjugates of anti-CEACAM5 and anti-EGP-1 humanized monoclonal antibodies.* Molecular Cancer Therapeutics, 2007. 6(12): p. 3617S-3618S.

450. Govindan, S.V., et al., *CEACAM5-Targeted Therapy of Human Colonic and Pancreatic Cancer Xenografts with Potent Labetuzumab-SN-38 Immunoconjugates.* Clinical Cancer Research, 2009. 15(19): p. 6052-6061.
451. Greiner, J.W., et al., *Vaccine-based therapy directed against carcinoembryonic antigen demonstrates antitumor activity on spontaneous intestinal tumors in the absence of autoimmunity.* Cancer Research, 2002. 62(23): p. 6944-6951.
452. Hauck, W. and C.P. Stanners, *Transcriptional Regulation of the Carcinoembryonic Antigen Gene - Identification of Regulatory Elements and Multiple Nuclear Factors.* Journal of Biological Chemistry, 1995. 270(8): p. 3602-3610.
453. Boucher, D., et al., *Studies on the Control of Gene-Expression of the Carcinoembryonic Antigen Family in Human-Tissue.* Cancer Research, 1989. 49(4): p. 847-852.
454. Tilton, B., et al., *Signal transduction by CXC chemokine receptor 4: Stromal cell-derived factor 1 stimulates prolonged protein kinase B and extracellular signal-regulated kinase 2 activation in T lymphocytes.* Journal of Experimental Medicine, 2000. 192(3): p. 313-324.
455. Vandercappellen, J., J. Van Damme, and S. Struyf, *The role of CXC chemokines and their receptors in cancer.* Cancer Letters, 2008. 267(2): p. 226-244.
456. Raman, D., et al., *Role of chemokines in tumor growth.* Cancer Letters, 2007. 256(2): p. 137-165.
457. Scotton, C.J., et al., *Multiple actions of the chemokine CXCL12 on epithelial tumor cells in human ovarian cancer.* Cancer Research, 2002. 62(20): p. 5930-5938.
458. Scotton, C.J., et al., *Epithelial cancer cell migration: A role for chemokine receptors?* Cancer Research, 2001. 61(13): p. 4961-4965.
459. Tachibana, K., et al., *The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract.* Nature, 1998. 393(6685): p. 591-594.
460. Nagasawa, T., et al., *Defects of B-cell lymphopoiesis and bone-marrow myelopoiesis in mice lacking the CXC chemokine PBSF/SDF-1.* Nature, 1996. 382(6592): p. 635-638.
461. Miao, Z., et al., *CXCR7 (RDC1) promotes breast and lung tumor growth in vivo and is expressed on tumor-associated vasculature.* Proceedings of the National Academy of Sciences of the United States of America, 2007. 104(40): p. 15735-15740.
462. Goldmann, T., et al., *CXCR7 transcription in human non-small cell lung cancer and tumor-free lung tissues; possible regulation upon chemotherapy.* Virchows Archiv, 2008. 452(3): p. 347-348.
463. Wang, J.H., et al., *The role of CXCR7/RDC1 as a chemokine receptor for CXCL12/SDF-1 in prostate cancer.* Journal of Biological Chemistry, 2008. 283(7): p. 4283-4294.
464. Maksym, R.B., et al., *The role of stromal-derived factor-1-CXCR7 axis in development and cancer.* European Journal of Pharmacology, 2009. 625(1-3): p. 31-40.

465. Darash-Yahana, M., et al., *Role of high expression levels of CXCR4 in tumor growth, vascularization, and metastasis.* Faseb Journal, 2004. 18(9): p. 1240-+.
466. Ottaiano, A., et al., *Inhibitory effects of anti-CXCR4 antibodies on human colon cancer cells.* Cancer Immunology Immunotherapy, 2005. 54(8): p. 781-791.
467. Muller, A., et al., *Involvement of chemokine receptors in breast cancer metastasis.* Nature, 2001. 410(6824): p. 50-56.
468. Taichman, R.S., et al., *Use of the stromal cell-derived factor-1/CXCR4 pathway in prostate cancer metastasis to bone.* Cancer Research, 2002. 62(6): p. 1832-1837.
469. Dwinell, M.B., P.A. Johanesen, and J.M. Smith, *Immunobiology of epithelial chemokines in the intestinal mucosa.* Surgery, 2003. 133(6): p. 601-607.
470. Zeelenberg, I.S., L. Ruuls-Van Stalle, and E. Roos, *The chemokine receptor CXCR4 is required for outgrowth of colon carcinoma micrometastases.* Cancer Research, 2003. 63(13): p. 3833-3839.
471. Ottaiano, A., et al., *Overexpression of both CXC chemokine receptor 4 and vascular endothelial growth factor proteins predicts early distant relapse in stage II-III colorectal cancer patients.* Clinical Cancer Research, 2006. 12(9): p. 2795-2803.
472. Sierro, F., et al., *Disrupted cardiac development but normal hematopoiesis in mice deficient in the second CXCL12/SDF-1 receptor, CXCR7.* Proceedings of the National Academy of Sciences of the United States of America, 2007. 104(37): p. 14759-14764.
473. Sakai Y, *Quantitative measurement of liver fatty acid binding protein in human gastrointestinal tract.* Nippon Shokakibyo Gakkai Zasshi, 1990. 87(12): p. 2594-2604.
474. Kawamura, T., et al., *Expression of liver-type fatty-acid-binding protein, fatty acid synthase and vascular endothelial growth factor in human lung carcinoma.* Pathobiology, 2005. 72(5): p. 233-240.
475. Hashimoto, T., et al., *Liver-type fatty acid-binding protein is highly expressed in intestinal metaplasia and in a subset of carcinomas of the stomach without association with the fatty acid synthase status in the carcinoma.* Pathobiology, 2004. 71(3): p. 115-122.
476. Suzuki, T., K. Watanabe, and T. Ono, *Immunohistochemical Demonstration of Liver Fatty Acid-Binding Protein in Human Hepatocellular Malignancies.* Journal of Pathology, 1990. 161(1): p. 79-83.
477. Lee, S., et al., *Differential expression in normal-adenoma-carcinoma sequence suggests complex molecular carcinogenesis in colon.* Oncology Reports, 2006. 16(4): p. 747-754.
478. Yamazaki, T., et al., *Liver fatty acid-binding protein is a new prognostic factor for hepatic resection of colorectal cancer metastases.* Journal of Surgical Oncology, 1999. 72(2): p. 83-87.
479. Strieter, R.M., et al., *Cancer CXC chemokine networks and tumour angiogenesis.* European Journal of Cancer, 2006. 42(6): p. 768-778.
480. Murphy, P.M., *Chemokines and the molecular basis of cancer metastasis.* New England Journal of Medicine, 2001. 345(11): p. 833-835.

481. Hu, D.E., Y. Hori, and T.P.D. Fan, *Interleukin-8 Stimulates Angiogenesis in Rats.* Inflammation, 1993. 17(2): p. 135-143.
482. Strieter, R.M., et al., *Interleukin-8 - a Corneal Factor That Induces Neovascularization.* American Journal of Pathology, 1992. 141(6): p. 1279-1284.
483. Lewis, C.E. and J.W. Pollard, *Distinct role of macrophages in different tumor microenvironments.* Cancer Research, 2006. 66(2): p. 605-612.
484. Erreni, M., et al., *Expression of Chemokines and Chemokine Receptors in Human Colon Cancer.* Chemokines, Pt A, 2009. 460: p. 105-121.
485. Baier, P.K., et al., *Chemokines in human colorectal carcinoma.* Anticancer Research, 2005. 25(5): p. 3581-3584.
486. Forstner, J.F., *Intestinal Mucins in Health and Disease.* Digestion, 1978. 17(3): p. 234-263.
487. Rachagani, S., et al., *Current status of mucins in the diagnosis and therapy of cancer.* Biofactors, 2009. 35(6): p. 509-527.
488. Byrd, J.C. and R.S. Bresalier, *Mucins and mucin binding proteins in colorectal cancer.* Cancer and Metastasis Reviews, 2004. 23(1): p. 77-99.
489. Kufe, D.W., *Mucins in cancer: function, prognosis and therapy.* Nature Reviews Cancer, 2009. 9(12): p. 874-885.
490. Leteurtre, E., et al., *Relationships between mucinous gastric carcinoma, MUC2 expression and survival.* World Journal of Gastroenterology, 2006. 12(21): p. 3324-3331.
491. Takikita, M., et al., *Associations between Selected Biomarkers and Prognosis in a Population-Based Pancreatic Cancer Tissue Microarray.* Cancer Research, 2009. 69(7): p. 2950-2955.
492. Yonezawa, S., et al., *Expression profiles of MUC1, MUC2, and MUC4 mucins in human neoplasms and their relationship with biological behavior.* Proteomics, 2008. 8(16): p. 3329-3341.
493. Albarracin, C.T., et al., *Differential expression of MUC2 and MUC5AC mucin genes in primary ovarian and metastatic colonic carcinoma.* Human Pathology, 2000. 31(6): p. 672-677.
494. Li, A.H., et al., *Expression of MUC1 and MUC2 mucins and relationship with cell proliferative activity in human colorectal neoplasia.* Pathology International, 2001. 51(11): p. 853-860.
495. Perez, R.O., et al., *Mucinous colorectal adenocarcinoma: influence of mucin expression (Muc1, 2 and 5) on clinico-pathological features and prognosis.* International Journal of Colorectal Disease, 2008. 23(8): p. 757-765.
496. Yang, H.S., et al., *A novel transformation suppressor, Pdcd4, inhibits AP-1 transactivation but not NF-kappa B or ODC transactivation.* Oncogene, 2001. 20(6): p. 669-676.
497. Cmarik, J.L., et al., *Differentially expressed protein Pdcd4 inhibits tumor promoter-induced neoplastic transformation.* Proceedings of the National Academy of Sciences of the United States of America, 1999. 96(24): p. 14037-14042.
498. Jansen, A.P., et al., *Characterization of programmed cell death 4 in multiple human cancers reveals a novel enhancer of drug sensitivity.* Molecular Cancer Therapeutics, 2004. 3(2): p. 103-110.

499. Jansen, A.P., C.E. Camalier, and N.H. Colburn, *Epidermal expression of the translation inhibitor programmed cell death 4 suppresses tumorigenesis.* Cancer Research, 2005. 65(14): p. 6034-6041.
500. Shibahara, K., et al., *Isolation of a novel mouse gene MA-3 that is induced upon programmed cell death.* Gene, 1995. 166(2): p. 297-301.
501. Zhang, Z.H. and R.N. DuBois, *Detection of differentially expressed genes in human colon carcinoma cells treated with a selective COX-2 inhibitor.* Oncogene, 2001. 20(33): p. 4450-4456.
502. Goke, R., et al., *Programmed cell death protein 4 suppresses CDK1/cdc2 via induction of p21(Waf1/Cip1).* American Journal of Physiology-Cell Physiology, 2004. 287(6): p. C1541-C1546.
503. Lankat-Buttgereit, B., et al., *Pdcd4 inhibits growth of tumor cells by suppression of carbonic anhydrase type II.* Molecular and Cellular Endocrinology, 2004. 214(1-2): p. 149-153.
504. Yang, H.S., et al., *A novel function of the MA-3 domains in transformation and translation suppressor Pdcd4 is essential for its binding to eukaryotic translation initiation factor 4A.* Molecular and Cellular Biology, 2004. 24(9): p. 3894-3906.
505. Yang, H.S., et al., *The transformation suppressor Pdcd4 is a novel eukaryotic translation initiation factor 4A binding protein that inhibits translation.* Molecular and Cellular Biology, 2003. 23(1): p. 26-37.
506. Schewe, D.M., et al., *Tumor-specific transcription factor binding to an activator protein-2/Sp1 element of the urokinase-type plasminogen activator receptor promoter in a first large series of resected gastrointestinal cancers.* Clinical Cancer Research, 2003. 9(6): p. 2267-2276.
507. Palamarchuk, A., et al., *Akt phosphorylates and regulates Pdcd4 tumor suppressor protein.* Cancer Research, 2005. 65(24): p. 11282-11286.
508. Schlichter, U., et al., *Identification of the myb-inducible promoter of the chicken Pdcd4 gene.* Biochimica Et Biophysica Acta-Gene Structure and Expression, 2001. 1520(1): p. 99-104.
509. Baffa, R., et al., *MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets.* Journal of Pathology, 2009. 219(2): p. 214-221.
510. Hiyoshi, Y., et al., *MicroRNA-21 Regulates the Proliferation and Invasion in Esophageal Squamous Cell Carcinoma.* Clinical Cancer Research, 2009. 15(6): p. 1915-1922.
511. Mudduluru, G., et al., *Loss of programmed cell death 4 expression marks adenoma-carcinoma transition, correlates inversely with phosphorylated protein kinase B, and is an independent prognostic factor in resected colorectal cancer.* Cancer, 2007. 110: p. 1697-1707.
512. Mackay, S.L.D., et al., *Colon-Cancer Cells That Are Not Growth Inhibited by Tgf-Beta Lack Functional Type-I and Type-Ii Tgf-Beta Receptors.* Annals of Surgery, 1995. 221(6): p. 767-777.
513. Langenskiod, M., et al., *Increased TGF-beta1 protein expression in patients with advanced colorectal cancer.* Journal of Surgical Oncology, 2008. 97(5): p. 409-415.
514. Sheng, H.M., et al., *Transformation of intestinal epithelial cells by chronic TGF-beta 1 treatment results in downregulation of the type II*

TGF-beta receptor and induction of cyclooxygenase-2. Oncogene, 1999. 18(4): p. 855-867.
515. Schroy, P., et al., *Role of Transforming Growth-Factor Beta-1 in Induction of Colon-Carcinoma Differentiation by Hexamethylene Bisacetamide.* Cancer Research, 1990. 50(2): p. 261-265.
516. Matsushita, M., et al., *Down regulation of TGF-beta receptors in human colorectal cancer: implications for cancer development.* British Journal of Cancer, 1999. 80(1-2): p. 194-205.
517. Wang, J., et al., *Reduced expression of transforming growth factor beta type I receptor contributes to the malignancy of human colon carcinoma cells.* Journal of Biological Chemistry, 1996. 271(29): p. 17366-17371.
518. Riggins, G.J., et al., *Mad-related genes in the human.* Nature Genetics, 1996. 13(3): p. 347-349.
519. MacGrogan, D., et al., *Comparative mutational analysis of DPC4 (Smad4) in prostatic and colorectal carcinomas.* Oncogene, 1997. 15(9): p. 1111-1114.
520. Elliott, R.L. and G.C. Blobe, *Role of transforming growth factor beta in human cancer.* Journal of Clinical Oncology, 2005. 23(9): p. 2078-2093.
521. Daniel, P., et al., *Transforming growth factor beta 1 and metalloproteinase-9 overexpression in colorectal cancer (CC) and adenoma.* International Journal of Colorectal Disease, 2007. 22(10): p. 1165-1172.
522. Bufill, J.A., *Colorectal-Cancer - Evidence for Distinct Genetic Categories Based on Proximal or Distal Tumor Location.* Annals of Internal Medicine, 1990. 113(10): p. 779-788.
523. Gervaz, P. and P. Morel, *Two colons - two cancers paradigm shift and clinical implications.* British Journal of Surgery, 2004. 91(7): p. 907-908.
524. Gervaz, P., et al., *Comparison of microsatellite instability and chromosomal instability in predicting survival of patients with T3N0 colorectal cancer.* Surgery, 2002. 131(2): p. 190-197.
525. Jernvall, P., et al., *Conserved region mutations of the p53 gene are concentrated in distal colorectal cancers.* International Journal of Cancer, 1997. 74(1): p. 97-101.
526. Kapiteijn, E., et al., *Mechanisms of oncogenesis in colon versus rectal cancer.* Journal of Pathology, 2001. 195(2): p. 171-178.
527. Fric, P., et al., *Different expression of some molecular markers in sporadic cancer of the left and right colon.* European Journal of Cancer Prevention, 2000. 9(4): p. 265-268.
528. Ojima, E., et al., *Effectiveness of gene expression profiling for response prediction of rectal cancer to preoperative radiotherapy.* Journal of Gastroenterology, 2007. 42(9): p. 730-736.
529. Ambrosini-Spaltro, A., et al., *High MUC2 lmmunohistochemical expression is a predictor of poor response to preoperative radiochemotherapy (RCT) in rectal adenocarcinoma.* Applied Immunohistochemistry & Molecular Morphology, 2008. 16(3): p. 227-231.
530. Julien, L.A. and A.G. Thorson, *Current Neoadjuvant Strategies in Rectal Cancer.* Journal of Surgical Oncology, 2010. 101(4): p. 321-326.

531. Dworak, O., L. Keilholz, and A. Hoffmann, *Pathological features of rectal cancer after preoperative radiochemotherapy.* International Journal of Colorectal Disease, 1997. 12(1): p. 19-23.
532. Habr-Gama, A., et al., *Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy - Long-term results.* Annals of Surgery, 2004. 240(4): p. 711-717.
533. Habr-Gama, A., et al., *Long-term results of preoperative chemoradiation for distal rectal cancer correlation between final stage and survival.* Journal of Gastrointestinal Surgery, 2005. 9(1): p. 90-99.
534. Habr-Gama, A., *Assessment and management of the complete clinical response of rectal cancer to chemoradiotherapy.* Colorectal Disease, 2006. 8: p. 21-24.
535. Chan, J.A., A.M. Krichevsky, and K.S. Kosik, *MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells.* Cancer Research, 2005. 65(14): p. 6029-6033.
536. Lassmann, S., et al., *A novel approach for reliable microarray analysis of microdissected tumor cells from formalin-fixed and paraffin-embedded colorectal cancer resection specimens.* Journal of Molecular Medicine (Berlin), 2009. 87(2).
537. Hui, A.B.Y., et al., *Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues.* Laboratory Investigation, 2009. 89(5): p. 597-606.
538. Doleshal, M., et al., *Evaluation and validation of total RNA extraction methods for MicroRNA expression analyses in formalin-fixed, paraffin-embedded tissues.* Journal of Molecular Diagnostics, 2008. 10(3): p. 203-211.
539. O'Neill, B.D.P., et al., *Non-operative treatment after neoadjuvant chemoradiotherapy for rectal cancer.* Lancet Oncology, 2007. 8(7): p. 625-633.
540. Cummins, J.M., et al., *The colorectal microRNAome.* Proceedings of the National Academy of Sciences of the United States of America, 2006. 103(10): p. 3687-3692.
541. Negrini, M. and G.A. Calin, *Breast cancer metastasis: a microRNA story.* Breast Cancer Research, 2008. 10(2).
542. Ma, L. and R.A. Weinberg, *MicroRNAs in malignant progression.* Cell Cycle, 2008. 7(5): p. 570-572.
543. Hiroki, E., et al., *Changes in microRNA expression levels correlate with clinicopathological features and prognoses in endometrial serous adenocarcinomas.* Cancer Science, 2010. 101(1): p. 241-249.
544. Lewis, F., et al., *Unlocking the archive - gene expression in paraffin-embedded tissue.* Journal of Pathology, 2001. 195(1): p. 66-71.
545. Khan, J., et al., *Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks.* Nature Medicine, 2001. 7(6): p. 673-679.
546. Lancashire, L., et al., *Classification of bacterial species from proteomic data using combinatorial approaches incorporating artificial neural networks, cluster analysis and principal components analysis.* Bioinformatics, 2005. 21(10): p. 2191-2199.

547. Phung, S.L. and A. Bouzerdoum, *A pyramidal neural network for visual pattern recognition.* Ieee Transactions on Neural Networks, 2007. 18(2): p. 329-343.
548. DiRusso, S.M., et al., *Development of a model for prediction of survival in pediatric trauma patients: Comparison of artificial neural networks and logistic regression.* Journal of Pediatric Surgery, 2002. 37(7): p. 1098-1103.
549. Matharoo-Ball, B., et al., *Characterization of biomarkers in polycystic ovary syndrome (PCOS) using multiple distinct proteomic platforms.* Journal of Proteome Research, 2007. 6(8): p. 3321-3328.
550. Matharoo-Ball, B., et al., *Diagnostic biomarkers differentiating metastatic melanoma patients from healthy controls identified by an integrated MALDI-TOF mass spectrometry/bioinformatic approach.* Proteomics Clinical Applications, 2007. 1(6): p. 605-620.
551. Stephan, C., et al., *An artificial neural network for five different assay systems of prostate-specific antigen in prostate cancer diagnostics.* Bju International, 2008. 102(7): p. 799-805.
552. Lancashire, L.J., et al., *A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks.* Breast Cancer Research and Treatment. 120(1): p. 83-93.
553. Dolgobrodov, S.G., et al., *Artificial neural network: Predicted vs. observed survival in patients with colonic cancer.* Diseases of the Colon & Rectum, 2007. 50(2): p. 184-191.
554. Ahmed, F.E., *Artificial neural networks for diagnosis and survival prediction in colon cancer.* Molecular Cancer, 2005. 4.
555. Chen, Y.D., et al., *Artificial neural networks analysis of surface-enhanced laser desorption/ionization mass spectra of serum protein pattern distinguishes colorectal cancer from healthy population.* Clinical Cancer Research, 2004. 10(24): p. 8380-8385.
556. Selaru, F.M., et al., *Artificial neural networks distinguish among subtypes of neoplastic colorectal lesions.* Gastroenterology, 2002. 122(3): p. 606-613.
557. Bottaci, L., et al., *Artificial neural networks applied to outcome prediction for colorectal cancer patients in separate institutions.* Lancet, 1997. 350(9076): p. 469-472.
558. Burke, H.B., et al., *Artificial neural networks improve the accuracy of cancer survival prediction.* Cancer, 1997. 79(4): p. 857-862.
559. Bishop C, *Neural Networks for Pattern Recognition*. 1995, Oxford: Oxford Universty Press.
560. Wightman, B., I. Ha, and G. Ruvkun, *Posttranscriptional Regulation of the Heterochronic Gene Lin-14 by Lin-4 Mediates Temporal Pattern-Formation in C-Elegans.* Cell, 1993. 75(5): p. 855-862.
561. Yang, L.D., N. Belaguli, and D.H. Berger, *MicroRNA and Colorectal Cancer.* World Journal of Surgery, 2009. 33(4): p. 638-646.
562. Slaby, O., et al., *Significance of microRNA-21 in colorectal cancer pathogenesis.* FEBS Journal, 2008. 275: p. 465-465.
563. Ng, E.K.O., et al., *Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening.* Gut, 2009. 58(10): p. 1375-1381.

564. Yanaihara, N., et al., *Unique microRNA molecular profiles in lung cancer diagnosis and prognosis.* Cancer Cell, 2006. 9(3): p. 189-198.
565. Yu, S.L., et al., *MicroRNA signature predicts survival and relapse in lung cancer.* Cancer Cell, 2008. 13(1): p. 48-57.
566. Budhu, A., et al., *Identification of metastasis-related microRNAs in hepatocellular carcinoma.* Hepatology, 2008. 47(3): p. 897-907.
567. Guo, Y., et al., *Distinctive MicroRNA profiles relating to patient survival in esophageal squamous cell carcinoma.* Cancer Research, 2008. 68(1): p. 26-33.
568. Farragher, S.M., et al., *RNA expression analysis from formalin fixed paraffin embedded tissues.* Histochemistry and Cell Biology, 2008. 130(3): p. 435-445.
569. Vincek, V., et al., *A tissue fixative that protects macromolecules (DNA, RNA, and protein) and histomorphology in clinical samples.* Laboratory Investigation, 2003. 83(10): p. 1427-1435.
570. Chen, X., et al., *Role of miR-143 targeting KRAS in colorectal tumorigenesis.* Oncogene, 2009. 28(10): p. 1385-1392.
571. Slaby, O., et al., *Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer.* Oncology, 2007. 72(5-6): p. 397-402.
572. Roberts L, et al., *Identification of methods for use of formalin-fixed, paraffin-embedded tissue samples in RNA expression profiling.* Genomics, 2009. 94(5): p. 341-8.
573. Graf, W., et al., *Short-term preoperative radiotherapy results in down-staging of rectal cancer: A study of 1316 patients.* Radiotherapy and Oncology, 1997. 43(2): p. 133-137.
574. Marijnen, C.A.M., et al., *No downstaging after short-term preoperative radiotherapy in rectal cancer patients.* Journal of Clinical Oncology, 2001. 19(7): p. 1976-1984.
575. Leibold, T. and J.G. Guillem, *The Role of Neoadjuvant Therapy in Sphincter-Saving Surgery for Mid and Distal Rectal Cancer.* Cancer Investigation, 2010. 28(3): p. 259-267.
576. Valentini, V., et al., *Does downstaging predict improved outcome after preoperative chemoradiation for extraperitoneal locally advanced rectal cancer? A long-term analysis of 165 patients.* International Journal of Radiation Oncology Biology Physics, 2002. 53(3): p. 664-674.
577. Heriot, A.G., et al., *Adjuvant radiotherapy is associated with increased sexual dysfunction in male patients undergoing resection for rectal cancer - A predictive model.* Annals of Surgery, 2005. 242(4): p. 502-511.
578. Birgisson, H., et al., *Adverse effects of preoperative radiation therapy for rectal cancer: Long-term follow-up of the Swedish rectal cancer trial.* Journal of Clinical Oncology, 2005. 23(34): p. 8697-8705.
579. Bengala, C., et al., *Epidermal growth factor receptor gene copy number, K-ras mutation and pathological response to preoperative cetuximab, 5-FU and radiation therapy in locally advanced rectal cancer.* Annals of Oncology, 2009. 20(3): p. 469-474.
580. Kobayashi, H., et al., *Absence of cyclooxygenase-2 protein expression is a predictor of tumor regression in rectal cancer treated with*

Preoperative short-term chemoradiotherapy. Diseases of the Colon & Rectum, 2007. 50(9): p. 1354-1362.
581. Wu, X.Y., et al., *Growth hormone receptor overexpression predicts response of rectal cancers to pre-operative radiotherapy.* European Journal of Cancer, 2006. 42(7): p. 888-894.
582. Hummel, R., D.J. Hussey, and J. Haier, *MicroRNAs: Predictors and modifiers of chemo- and radiotherapy in different tumour types.* European Journal of Cancer, 2010. 46(2): p. 298-311.
583. Kovalchuk, O., et al., *Involvement of microRNA-451 in resistance of the MCF-7 breast cancer cells to chemotherapeutic drug doxorubicin.* Molecular Cancer Therapeutics, 2008. 7(7): p. 2152-2159.
584. Xia, L., et al., *miR-15b and miR-16 modulate multidrug resistance by targeting BCL2 in human gastric cancer cells.* International Journal of Cancer, 2008. 123(2): p. 372-379.
585. Weidhaas, J.B., et al., *MicroRNAs as potential agents to alter resistance to cytotoxic anticancer therapy.* Cancer Research, 2007. 67(23): p. 11111-11116.
586. Josson, S., et al., *Radiation Modulation of MicroRNA in Prostate Cancer Cell Lines.* Prostate, 2008. 68(15): p. 1599-1606.
587. Earle, J.S.L., et al., *Association of MicroRNA Expression with Microsatellite Instability Status in Colorectal Adenocarcinoma.* Journal of Molecular Diagnostics. 12(4): p. 433-440.
588. Schaefer, A., et al., *Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma.* International Journal of Cancer. 126(5): p. 1166-1176.
589. Nam, E.J., et al., *MicroRNA expression profiles in serous ovarian carcinoma.* Clinical Cancer Research, 2008. 14(9): p. 2690-2695.
590. Aqeilan, R.I., G.A. Calin, and C.M. Croce, *miR-15a and miR-16-1 in cancer: discovery, function and future perspectives.* Cell Death and Differentiation. 17(2): p. 215-220.
591. Schaefer A, et al., *Suitable reference genes for relative quantification of miRNA expression in prostate cancer*. Experimental and Molecular Medicine, 2010. 42(11): p. 749-758.
592. Liu, Q., et al., *miR-16 family induces cell cycle arrest by regulating multiple cell cycle genes.* Nucleic Acids Research, 2008. 36(16): p. 5391-5404.
593. Kaddar, T., et al., *Two new miR-16 targets: caprin-1 and HMGA1, proteins implicated in cell proliferation.* Biology of the Cell, 2009. 101(9): p. 511-524.
594. Balakrishnan, A., et al., *MicroRNA mir-16 is anti-proliferative in enterocytes and exhibits diurnal rhythmicity in intestinal crypts.* Experimental Cell Research, 2010. 316(20): p. 3512-3521.
595. Zhang X, et al., *Oncogenic Wip1 phosphatase is inhibited by miR-16 in the DNA damage signaling pathway.* Cancer Research, 2010. 70(18): p. 7176-7186.
596. Takeshita, F., et al., *Systemic Delivery of Synthetic MicroRNA-16 Inhibits the Growth of Metastatic Prostate Tumors via Downregulation of Multiple Cell-cycle Genes.* Molecular Therapy. 18(1): p. 181-187.

597. Cittelly DM, et al., *Oncogenic HER2{Delta}16 suppresses miR-15a/16 and deregulates BCL-2 to promote endocrine resistance of breast tumors.* Carcinogenesis, 2010. 31(12): p. 2049-2057.
598. Wagner-Ecker M, et al., *MicroRNA expression after ionizing radiation in human endothelial cells.* Radiation Oncology, 2010. 5: p. 25.
599. Shan, H.L., et al., *Downregulation of miR-133 and miR-590 contributes to nicotine-induced atrial remodelling in canines.* Cardiovascular Research, 2009. 83(3): p. 465-472.
600. Kim, T.H., et al., *Deregulation of miR-519a, 153, and 485-5p and its clinicopathological relevance in ovarian epithelial tumours.* Histopathology. 57(5): p. 734-743.
601. Myatt, S.S., et al., *Definition of microRNAs That Repress Expression of the Tumor Suppressor Gene FOXO1 in Endometrial Cancer.* Cancer Research. 70(1): p. 367-377.
602. Xu, J.Z., X.M. Liao, and C.W. Wong, *Downregulations of B-cell lymphoma 2 and myeloid cell leukemia sequence 1 by microRNA 153 induce apoptosis in a glioblastoma cell line DBTRG-05MG.* International Journal of Cancer. 126(4): p. 1029-1035.
603. Abdelmohsen, K., et al., *miR-519 reduces cell proliferation by lowering RNA-binding protein HuR levels.* Proceedings of the National Academy of Sciences of the United States of America, 2008. 105(51): p. 20297-20302.
604. Abdelmohsen, K., et al., *miR-519 suppresses tumor growth by reducing HuR levels.* Cell Cycle, 2010. 9(7): p. 1354-1359.
605. Ristimaki, A., *Tumor suppressor effect of the microRNA miR-519 is mediated via the mRNA-binding protein HuR.* Cell Cycle, 2010. 9(7): p. 1234-1234.
606. Lagos-Quintana, M., et al., *Identification of novel genes coding for small expressed RNAs.* Science, 2001. 294(5543): p. 853-858.
607. Huang, J.C., et al., *Using expression profiling data to identify human microRNA targets.* Nature Methods, 2007. 4(12): p. 1045-1049.
608. Lai, E.C., *microRNAs: Runts of the genome assert themselves.* Current Biology, 2003. 13(23): p. R925-R936.
609. Chen, C.Z., et al., *MicroRNAs modulate hematopoietic lineage differentiation.* Science, 2004. 303(5654): p. 83-86.
610. Shi, B., et al., *Expression profiles of micro RNA in proliferating and differentiating 32D murine myeloid cells.* Journal of Cellular Physiology, 2006. 207(3): p. 706-710.
611. Wu, W., et al., *MicroRNA and cancer: Current status and prospective.* International Journal of Cancer, 2007. 120(5): p. 953-960.
612. Bentwich, I., *Prediction and validation of microRNAs and their targets.* Febs Letters, 2005. 579(26): p. 5904-5910.
613. Brown, J.R. and P. Sanseau, *A computational view of microRNAs and their targets.* Drug Discovery Today, 2005. 10(8): p. 595-601.
614. Rajewsky, N., *microRNA target predictions in animals.* Nature Genetics, 2006. 38: p. S8-S13.
615. Stark, A., et al., *Animal microRNAs confer robustness to gene expression and have a significant impact on 3 ' UTR evolution.* Cell, 2005. 123(6): p. 1133-1146.

616. Shi, B., et al., *Micro RNA 145 targets the insulin receptor substrate-1 and inhibits the growth of colon cancer cells.* Journal of Biological Chemistry, 2007. 282: p. 32582-32590.
617. Murakami, Y., et al., *Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues.* Oncogene, 2006. 25(17): p. 2537-2545.
618. Voorhoeve, P.M., et al., *A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors.* Cell, 2006. 124(6): p. 1169-1181.
619. Gatt ME, et al., *MicroRNAs 15a/16-1 function as tumor suppressor genes in multiple myeloma.* Blood, 2010. Epub ahead of print.
620. Matsusue, R., et al., *Hepatic Stellate Cells Promote Liver Metastasis of Colon Cancer Cells by the Action of SDF-1/CXCR4 Axis.* Annals of Surgical Oncology, 2009. 16(9): p. 2645-2653.
621. Yang, S.X., et al., *Activation of chemokine receptor CXCR4 in malignant glioma cells promotes the production of vascular endothelial growth factor (vol 335, pg 523, 2005).* Biochemical and Biophysical Research Communications, 2005. 337(1): p. 411-411.
622. Kryczek, I., et al., *CXCL12 and vascular endothelial growth factor synergistically induce neoanglogenesis in human ovarian cancers.* Cancer Research, 2005. 65(2): p. 465-472.
623. Das, R., et al., *Expression pattern of fatty acid-binding proteins in human normal and cancer prostate cells and tissues.* Clinical Cancer Research, 2001. 7(6): p. 1706-1715.
624. Yang, H.S., et al., *Tumorigenesis suppressor Pdcd4 down-regulates mitogen-activated protein kinase kinase kinase kinase 1 expression to suppress colon carcinoma cell invasion.* Mol Cell Biol, 2006. 26(4): p. 1297-306.
625. Wang, Q., Z. Sun, and H.S. Yang, *Downregulation of tumor suppressor Pdcd4 promotes invasion and activates both beta-catenin/Tcf and AP-1-dependent transcription in colon carcinoma cells.* Oncogene, 2008. 27(11): p. 1527-1535.
626. Mudduluru, G., et al., *Loss of programmed cell death 4 expression marks adenoma-carcinoma transition, correlates inversely with phosphorylated protein kinase B, and is an independent prognostic factor in resected colorectal cancer.* Cancer Res., 2007. 110(8): p. 1697-707.
627. Roldo, C., et al., *MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior.* Journal of Clinical Oncology, 2006. 24(29): p. 4677-4684.
628. Yan, L.X., et al., *MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis.* Rna-a Publication of the Rna Society, 2008. 14(11): p. 2348-2360.
629. Qian, B.Y., et al., *High miR-21 expression in breast cancer associated with poor disease-free survival in early stage disease and high TGF-beta 1.* Breast Cancer Research and Treatment, 2009. 117(1): p. 131-140.
630. Sayed, D., et al., *MicroRNA-21 targets Sprouty2 and promotes cellular outgrowths.* Molecular Biology of the Cell, 2008. 19(8): p. 3272-3282.

631. Liu, X., et al., *MicroRNA-31 functions as an oncogenic microRNA in mouse and human lung cancer cells by repressing specific tumor suppressors.* Journal of Clinical Investigation. 120(4): p. 1298-1309.
632. Liu, X.Q., et al., *MicroRNA Profiling and Head and Neck Cancer.* Comparative and Functional Genomics, 2009.
633. Mi, S.L., et al., *Aberrant overexpression and function of the miR-17-92 cluster in MLL-rearranged acute leukemia.* Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(8): p. 3710-3715.
634. Diosdado, B., et al., *MiR-17-92 cluster is associated with 13q gain and c-myc expression during colorectal adenoma to adenocarcinoma progression.* British Journal of Cancer, 2009. 101(4): p. 707-714.
635. Hayashita, Y., et al., *A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation.* Cancer Research, 2005. 65(21): p. 9628-9632.
636. Liu, X., et al., *MicroRNA-31 functions as an oncogenic microRNA in mouse and human lung cancer cells by repressing specific tumor suppressors.* Journal of Clinical Investigation, 2010. 120(4): p. 1298-1309.
637. Sarver, A.L., et al., *Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states.* Bmc Cancer, 2009. 9.
638. Fujimoto, J., et al., *Clinical implications of expression of interleukin 8 related to angiogenesis in uterine cervical cancers.* Cancer Research, 2000. 60(10): p. 2632-2635.
639. Sparmann, A. and D. Bar-Sagi, *Ras-induced interleukin-8 expression plays a critical role in tumor growth and angiogenesis.* Cancer Cell, 2004. 6(5): p. 447-458.
640. Inoue, K., et al., *Interleukin 8 expression regulates tumorigenicity and metastases in androgen-independent prostate cancer.* Clinical Cancer Research, 2000. 6(5): p. 2104-2119.
641. Galffy, G., et al., *Inhibition of interleukin-8 reduces human malignant pleural mesothelioma propagation in nude mouse model.* Oncology Research, 1999. 11(4): p. 187-194.
642. Kim, S.J., et al., *Expression of interleukin-8 correlates with angiogenesis, tumorigenicity, and metastasis of human prostate cancer cells implanted orthotopically in nude mice.* Neoplasia, 2001. 3(1): p. 33-42.
643. La Rocca, G., et al., *Mechanism of Growth Inhibition by MicroRNA 145: The Role of the IGF-I Receptor Signaling Pathway.* Journal of Cellular Physiology, 2009. 220(2): p. 485-491.
644. Gee HE, et al., *hsa-miR-210 is a marker of tumor hypoxia and a prognostic factor in head and neck cancer.* Cancer, 2010. 16(9): p. 2148-58.
645. Gee HE, et al., *MicroRNA-10b and breast cancer metastasis.* Nature, 2008. 455(7216): p. E8-9.
646. Sasayama, T., et al., *MicroRNA-10b is overexpressed in malignant glioma and associated with tumor invasive factors, uPAR and RhoC.* International Journal of Cancer, 2009. 125(6): p. 1407-1413.

647. Tian, Y.Y., et al., *MicroRNA-10b Promotes Migration and Invasion through KLF4 in Human Esophageal Cancer Cell Lines.* Journal of Biological Chemistry, 2010. 285(11): p. 7986-7994.
648. Aaltonen, L.A., et al., *Clues to the Pathogenesis of Familial Colorectal-Cancer.* Science, 1993. 260(5109): p. 812-816.
649. Thibodeau, S.N., G. Bren, and D. Schaid, *Microsatellite Instability in Cancer of the Proximal Colon.* Science, 1993. 260(5109): p. 816-819.
650. Aaltonen, L.A., et al., *Replication Errors in Benign and Malignant-Tumors from Hereditary Nonpolyposis Colorectal-Cancer Patients.* Cancer Research, 1994. 54(7): p. 1645-1648.
651. Jiricny, J., *Replication errors: cha(lle)nging the genome.* Embo Journal, 1998. 17(22): p. 6427-6436.
652. Kolodner, R.D. and G.T. Marsischky, *Eukaryotic DNA mismatch repair.* Current Opinion in Genetics & Development, 1999. 9(1): p. 89-96.
653. Marra, G. and C.R. Boland, *Hereditary Nonpolyposis Colorectal-Cancer - the Syndrome, the Genes, and Historical Perspectives.* Journal of the National Cancer Institute, 1995. 87(15): p. 1114-1125.
654. Liu, B., et al., *Analysis of mismatch repair genes in hereditary non-polyposis colorectal cancer patients.* Nature Medicine, 1996. 2(2): p. 169-174.
655. Peltomaki, P., et al., *Mutations predisposing to hereditary nonpolyposis colorectal cancer: Database and results of a collaborative study.* Gastroenterology, 1997. 113(4): p. 1146-1158.
656. Yan, H., et al., *Conversion of diploidy to haploidy - Individuals susceptible to multigene disorders may now be spotted more easily.* Nature, 2000. 403(6771): p. 723-724.
657. Langer, S., et al., *Facilitating haplotype analysis by fully automated analysis of all chromosomes in human-mouse hybrid cell lines.* Cytogenetics and Cell Genetics, 2001. 93(1-2): p. 11-15.
658. Lawes, D.A., S. SenGupta, and P.B. Boulos, *The clinical importance and prognostic implications of microsatellite instability in sporadic cancer.* European Journal of Surgical Oncology, 2003. 29(3): p. 201-212.
659. Elsaleh, H., et al., *p53 Gene mutation, microsatellite instability and adjuvant chemotherapy: Impact on survival of 388 patients with Dukes' C colon carcinoma.* Oncology, 2000. 58(1): p. 52-59.
660. Barratt, P.L., et al., *The role of the DNA mismatch repair protein, hMLH1 in sensitivity of colon colon cancer cells to chemotherapeutic agents used clinically.* British Journal of Cancer, 1998. 78: p. P9.
661. Barratt, P.L., et al., *DNA markers predicting benefit from adjuvant fluorouracil in patients with colon cancer: a molecular study.* Lancet, 2002. 360(9343): p. 1381-1391.
662. Elsaleh, H., et al., *p53 alteration and microsatellite instability have predictive value for survival benefit from chemotherapy in stage III colorectal carcinoma.* Clinical Cancer Research, 2001. 7(5): p. 1343-1349.
663. Chen, X.X., M.D. Lai, and Q.O. Huang, *Increased sensitivity of colorectal cancer cell lines with microsatellite instability to 5-fluorouracil in vitro.* Chinese Medical Journal, 2002. 115(7): p. 1048-1052.

664. Watanabe, T., et al., *Molecular predictors of survival after adjuvant chemotherapy for colon cancer.* New England Journal of Medicine, 2001. 344(16): p. 1196-1206.
665. Barnetson, R.A., et al., *Identification and survival of carriers of mutations in DNA mismatch-repair genes in colon cancer.* New England Journal of Medicine, 2006. 354(26): p. 2751-2763.
666. Pinol, V., et al., *Accuracy of revised Bethesda guidelines, microsatellite instability, and immunohistochemistry for the identification of patients with hereditary nonpolyposis colorectal cancer.* Jama-Journal of the American Medical Association, 2005. 293(16): p. 1986-1994.
667. Wullenweber, H.P., et al., *Evaluation of Bethesda guidelines in relation to microsatellite instability.* Diseases of the Colon & Rectum, 2001. 44(9): p. 1281-1289.
668. Raedle, J., et al., *Bethesda guidelines: Relation to microsatellite instability and MLH1 promoter methylation in patients with colorectal cancer.* Annals of Internal Medicine, 2001. 135(8): p. 566-576.
669. Gruber, S.B., *New developments in Lynch syndrome (hereditary nonpolyposis colorectal cancer) and mismatch repair gene testing.* Gastroenterology, 2006. 130(2): p. 577-587.
670. Murff, H.J., D.R. Spigel, and S. Syngal, *Does this patient have a family history of cancer? An evidence-based analysis of the accuracy of family cancer history.* Jama-Journal of the American Medical Association, 2004. 292(12): p. 1480-1489.
671. Hampel, H., et al., *Screening for the Lynch syndrome (Hereditary nonpolyposis colorectal cancer).* New England Journal of Medicine, 2005. 352(18): p. 1851-1860.
672. Ward, R.L., et al., *Routine testing for mismatch repair deficiency in sporadic colorectal cancer is justified.* Journal of Pathology, 2005. 207(4): p. 377-384.
673. Giardiello, F.M., J.D. Brensinger, and G.M. Petersen, *AGA technical review on hereditary colorectal cancer and genetic testing.* Gastroenterology, 2001. 121(1): p. 198-213.
674. Chung, D.C. and A.K. Rustgi, *The hereditary nonpolyposis colorectal cancer syndrome: Genetics and clinical implications.* Annals of Internal Medicine, 2003. 138(7): p. 560-570.
675. Ramsey, S.D., et al., *Cost-effectiveness of microsatellite instability screening as a method for detecting hereditary nonpolyposis colorectal cancer.* Annals of Internal Medicine, 2001. 135(8): p. 577-588.
676. Dietmaier, W., et al., *Diagnostic microsatellite instability: Definition and correlation with mismatch repair protein expression.* Cancer Research, 1997. 57(21): p. 4749-4756.
677. Wahlberg, S.S., et al., *Evaluation of microsatellite instability and inummohistochemistry for the prediction of germ-line MSH2 and MLH1 mutations in hereditary nonpolyposis colon cancer families.* Cancer Research, 2002. 62(12): p. 3485-3492.
678. Drummond, J.T., et al., *Isolation of an Hmsh2-P160 Heterodimer That Restores DNA Mismatch Repair to Tumor-Cells.* Science, 1995. 268(5219): p. 1909-1912.

679. Prolla, T.A., et al., *Mlh1, Pms1, and Msh2 Interactions During the Initiation of DNA Mismatch Repair in Yeast.* Science, 1994. 265(5175): p. 1091-1093.
680. Hawkins, N.J. and R.L. Ward, *Sporadic colorectal cancers with microsatellite instability and their possible origin in hyperplastic polyps and serrated adenomas.* Journal of the National Cancer Institute, 2001. 93(17): p. 1307-1313.
681. Parc, Y.R., et al., *hMSH6 alterations in patients with microsatellite instability-low colorectal cancer.* Cancer Research, 2000. 60(8): p. 2225-2231.
682. Thibodeau, S.N., et al., *Microsatellite instability in colorectal cancer: Different mutator phenotypes and the principal involvement of hMLH1.* Cancer Research, 1998. 58(8): p. 1713-1718.
683. Jarvinen, H.J., et al., *Controlled 15-year trial on screening for colorectal cancer in families with hereditary nonpolyposis colorectal cancer.* Gastroenterology, 2000. 118(5): p. 829-834.
684. Salovaara, R., et al., *Population-based molecular detection of hereditary nonpolyposis colorectal cancer.* Journal of Clinical Oncology, 2000. 18(11): p. 2193-2200.
685. Fink, D., S. Aebi, and S.B. Howell, *The role of DNA mismatch repair in drug resistance.* Clinical Cancer Research, 1998. 4(1): p. 1-6.
686. Anthoney, D.A., et al., *Microsatellite instability, apoptosis, and loss of p53 function in drug-resistant tumor cells.* Cancer Research, 1996. 56(6): p. 1374-1381.
687. Peltomaki, P., *Deficient DNA mismatch repair: a common etiologic factor for colon cancer.* Human Molecular Genetics, 2001. 10(7): p. 735-740.
688. Chaves, P., et al., *Immunohistochemical detection of mismatch repair gene proteins as a useful tool for the identification of colorectal carcinoma with the mutator phenotype.* Journal of Pathology, 2000. 191(4): p. 355-360.
689. Plaschke, J., et al., *Involvement of hMSH6 in the development of hereditary and sporadic colorectal cancer revealed by immunostaining is based on germline mutations, but rarely on somatic inactivation.* International Journal of Cancer, 2002. 97(5): p. 643-648.
690. Curia, M.C., et al., *Unbalanced germ-line expression of hMLH1 and hMSH2 alleles in hereditary nonpolyposis colorectal cancer.* Cancer Research, 1999. 59(15): p. 3570-3575.
691. Debniak, T., et al., *Value of pedigree/clinical data, immunohistochemistry and microsatellite instability analyses in reducing the cost of determining hMLH1 and hMSH2 gene mutations in patients with colorectal cancer.* European Journal of Cancer, 2000. 36(1): p. 49-54.
692. Dieumegard, B., et al., *Extensive molecular screening for hereditary non-polyposis colorectal cancer.* British Journal of Cancer, 2000. 82(4): p. 871-880.
693. Syngal, S., et al., *Sensitivity and specificity of clinical criteria for hereditary non-polyposis colorectal cancer associated mutations in MSH2 and MLH1.* Journal of Medical Genetics, 2000. 37(9): p. 641-645.
694. Terdiman, J.P., et al., *Efficient detection of hereditary nonpolyposis colorectal cancer gene carriers by screening for tumor microsatellite*

instability before germline genetic testing. Gastroenterology, 2001. 120(1): p. 21-30.
695. Durno, C., et al., *Family history and molecular features of children, adolescents, and young adults with colorectal carcinoma.* Gut, 2005. 54(8): p. 1146-1150.
696. Southey, M.C., et al., *Use of molecular tumor characteristics to prioritize mismatch repair gene testing in early-onset colorectal cancer.* Journal of Clinical Oncology, 2005. 23(27): p. 6524-6532.
697. Bonis PA, et al., *Hereditary Nonpolyposis Colorectal Cancer: Diagnostic Strategies and Their Implications*, in *Evidence Report/Technology*. 2007, Tufts-New England Medical Center Evidence-based Practice Center: Rockville, MD:Agency for Healthcare Research and Quality.
698. Lindor, N.M., et al., *Immunohistochemistry versus microsatellite instability testing in phenotyping colorectal tumors.* Journal of Clinical Oncology, 2002. 20(4): p. 1043-1048.
699. Müller W, et al., *The reliability of immunohistochemistry as a prescreening method for the diagnosis of hereditary nonpolyposis colorectal cancer (HNPCC)--results of an international collaborative study.* Fam Cancer. 1(2): p. 87-92.
700. Young, J., et al., *Features of colorectal cancers with high-level microsatellite instability occurring in familial and sporadic settings - Parallel pathways of tumorigenesis.* American Journal of Pathology, 2001. 159(6): p. 2107-2116.
701. Parkin DM, et al., *Global cancer statistics, 2002.* CA Cancer J Clin, 2005. 55(2): p. 74-108.
702. Jemal A, et al., *Cancer statistics, 2010.* CA Cancer J Clin, 2010. 60(5): p. 277-300.
703. *Cancer Projections 2005-2035*. National Cancer Registry C. 2008, National Cancer Registry

Appendix 1: Department of Surgery Biobank Forms

1.1 Study Ethical Approval

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

Merlin Park University Hospital
Ospidéal na h-Ollscoile, Páirc Mheirlinne
GALWAY UNIVERSITY HOSPITALS

Clinical Research Ethics Committee
Unit 4
Merlin Park Hospital
Galway.

24th September, 2007.

Dr. Kah Hoong Chang
Research Fellow
Department of Surgery
University College Hospital
Galway.

Ref: CA 102 – "Investigation of novel factors in colorectal carcinoma

Dear Dr. Hoong Chang,

I have considered your project, and I am happy to grant Chairman's approval.

Yours sincerely,

Dr. Shaun T. O'Keeffe
Chairman Clinical Research Ethics Committee.

Merlin Park University Hospital, Ospidéal na h-Ollscoile, Páirc Mheirlinne,
Galway, Ireland. Tel: 00 353 (0)91 757631

1.2 Consent form Ethical Approval

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

Merlin Park University Hospital
Ospidéal na h-Ollscoile, Páirc Mheirlinne
GALWAY UNIVERSITY HOSPITALS

Clinical Research Ethics Committee
Unit 4
Merlin Park Hospital
Galway.

6th October, 2008.

Ms. Nicola Miller
Department of Surgery
Clinical Science Institute
University College Hospital
Galway.

Ref: C.A. 151 – Galway University Hospitals – BIOBANK INFORMED CONSENT

Dear Ms. Miller,

The Chairman's decision to approve the above project, was ratified at the last Clinical Research Ethics Committee meeting of Wednesday 17th September, 2008.

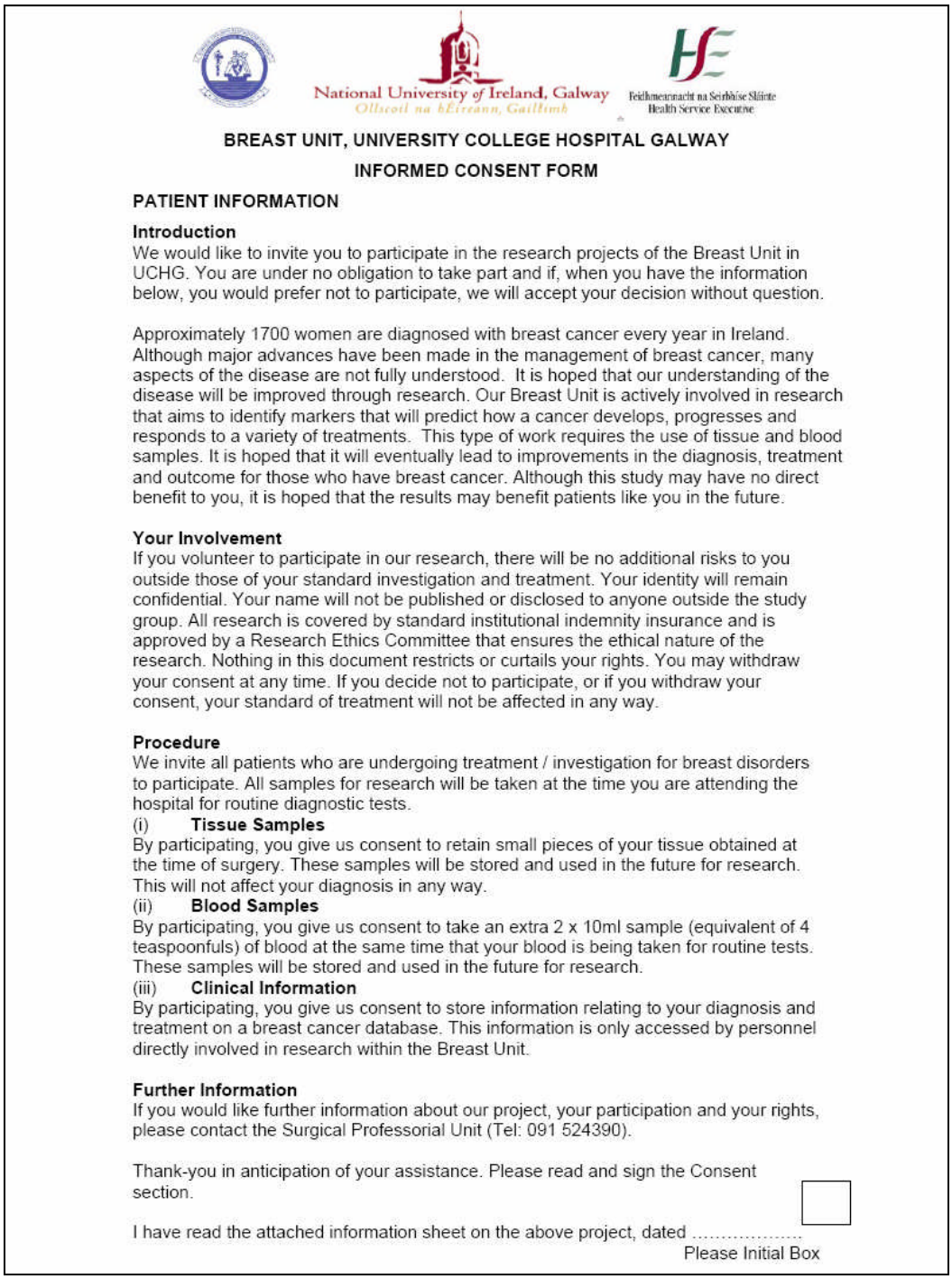
Yours sincerely,

p.p. Colette Collins

Dr. Shaun T. O'Keeffe
Chairman Clinical Research Ethics Committee.

Merlin Park University Hospital, Ospidéal na h-Ollscoile, Páirc Mheirlinne,
Galway, Ireland. Tel: 00 353 (0)91 757631

1.3 Patient Consent Form



National University of Ireland, Galway
Ollscoil na hÉireann, Gaillimh

Feidhmeannacht na Seirbhíse Sláinte
Health Service Executive

BREAST UNIT, UNIVERSITY COLLEGE HOSPITAL GALWAY

INFORMED CONSENT FORM

PATIENT INFORMATION

Introduction
We would like to invite you to participate in the research projects of the Breast Unit in UCHG. You are under no obligation to take part and if, when you have the information below, you would prefer not to participate, we will accept your decision without question.

Approximately 1700 women are diagnosed with breast cancer every year in Ireland. Although major advances have been made in the management of breast cancer, many aspects of the disease are not fully understood. It is hoped that our understanding of the disease will be improved through research. Our Breast Unit is actively involved in research that aims to identify markers that will predict how a cancer develops, progresses and responds to a variety of treatments. This type of work requires the use of tissue and blood samples. It is hoped that it will eventually lead to improvements in the diagnosis, treatment and outcome for those who have breast cancer. Although this study may have no direct benefit to you, it is hoped that the results may benefit patients like you in the future.

Your Involvement
If you volunteer to participate in our research, there will be no additional risks to you outside those of your standard investigation and treatment. Your identity will remain confidential. Your name will not be published or disclosed to anyone outside the study group. All research is covered by standard institutional indemnity insurance and is approved by a Research Ethics Committee that ensures the ethical nature of the research. Nothing in this document restricts or curtails your rights. You may withdraw your consent at any time. If you decide not to participate, or if you withdraw your consent, your standard of treatment will not be affected in any way.

Procedure
We invite all patients who are undergoing treatment / investigation for breast disorders to participate. All samples for research will be taken at the time you are attending the hospital for routine diagnostic tests.

(i) **Tissue Samples**
By participating, you give us consent to retain small pieces of your tissue obtained at the time of surgery. These samples will be stored and used in the future for research. This will not affect your diagnosis in any way.

(ii) **Blood Samples**
By participating, you give us consent to take an extra 2 x 10ml sample (equivalent of 4 teaspoonfuls) of blood at the same time that your blood is being taken for routine tests. These samples will be stored and used in the future for research.

(iii) **Clinical Information**
By participating, you give us consent to store information relating to your diagnosis and treatment on a breast cancer database. This information is only accessed by personnel directly involved in research within the Breast Unit.

Further Information
If you would like further information about our project, your participation and your rights, please contact the Surgical Professorial Unit (Tel: 091 524390).

Thank-you in anticipation of your assistance. Please read and sign the Consent section.

I have read the attached information sheet on the above project, dated .....

Please Initial Box

1.4 contd- Patient Consent Form

PARTICIPANT DECLARATION

I have read, or had read to me, this consent form. I have had the opportunity to ask questions and all my questions have been answered to my satisfaction. I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights. I have received a copy of this agreement and I understand that, if there is a sponsoring company, a signed copy will be sent to that sponsor. I understand I may withdraw from the study at any time.

(Name of sponsor :) .....

PARTICIPANT'S NAME: .....

CONTACT DETAILS: .....

PARTICIPANT'S SIGNATURE: .....

DATE: .....

Where the participant is incapable of comprehending the nature, significance and scope of the consent required, the form must be signed by a person competent to give consent to his or her participation in the research study (other than a person who applied to undertake or conduct the study). If the participant is a minor (under 18 years old) the signature of parent or guardian must be obtained:

NAME OF CONSENTER, PARENT, OR GUARDIAN: .....

SIGNATURE: .....

RELATION TO PARTICIPANT: .....

DECLARATION OF INVESTIGATOR'S RESPONSIBILITY

I have explained the nature and purpose of this research study, the procedures to be undertaken and any risks that may be involved. I have offered to answer any questions and fully answered such questions. I believe that the participant understands my explanation and has freely given informed consent.

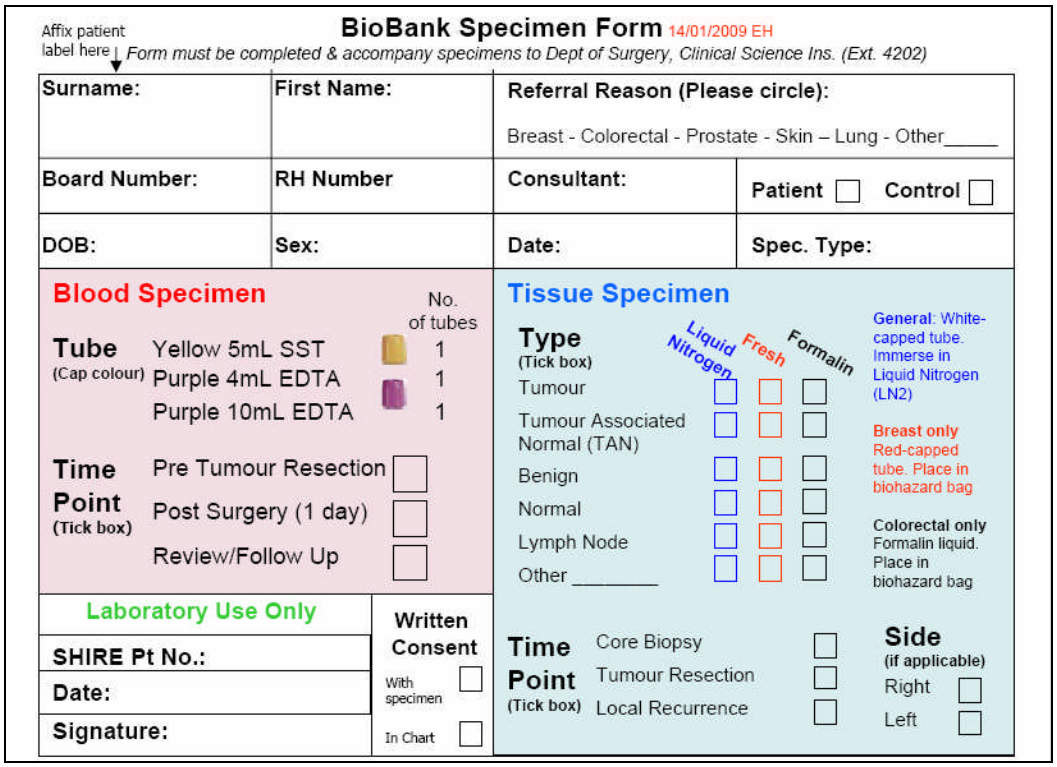
NAME: .....

SIGNATURE: .....

DATE: .....

(Keep the original of this form in the investigator's file, give one copy to the participant, and send one copy to the sponsor (if there is a sponsor).

1.5 Specimen Request Form



Affix patient label here

BioBank Specimen Form 14/01/2009 EH

Form must be completed & accompany specimens to Dept of Surgery, Clinical Science Ins. (Ext. 4202)

Surname:	**First Name:**	**Referral Reason (Please circle):** Breast - Colorectal - Prostate - Skin – Lung - Other_____	
Board Number:	**RH Number**	**Consultant:**	**Patient** ☐ **Control** ☐
DOB:	**Sex:**	**Date:**	**Spec. Type:**

Blood Specimen

		No. of tubes
Tube (Cap colour)	Yellow 5mL SST	1
	Purple 4mL EDTA	1
	Purple 10mL EDTA	1

Time Point (Tick box)
- Pre Tumour Resection ☐
- Post Surgery (1 day) ☐
- Review/Follow Up ☐

Tissue Specimen

Type (Tick box)	Liquid Nitrogen	Fresh	Formalin
Tumour	☐	☐	☐
Tumour Associated Normal (TAN)	☐	☐	☐
Benign	☐	☐	☐
Normal	☐	☐	☐
Lymph Node	☐	☐	☐
Other _____	☐	☐	☐

General: White-capped tube. Immerse in Liquid Nitrogen (LN2)

Breast only Red-capped tube. Place in biohazard bag

Colorectal only Formalin liquid. Place in biohazard bag

Time Point (Tick box)
- Core Biopsy ☐
- Tumour Resection ☐
- Local Recurrence ☐

Side (if applicable)
- Right ☐
- Left ☐

Laboratory Use Only

SHIRE Pt No.:

Date:

Signature:

Written Consent
- With specimen ☐
- In Chart ☐

Appendix 2: Suppliers of Reagents & Equipment used in experiments described in Chapter 2 (Materials & Methods)

Agilent Technologies, Waldbronn, Germany and Palo Alto, CA, USA

- Agilent 2100 Bioanalyzer Sysem
- Agilent 2100 Expert software (version B.02.03)
- RNA 6000 Nano LabChip Series II Assay
- Small RNA Assay (cat. no. 5067-1548)
- Agilent Small RNA kit guide (manual part no. G2938-90093)

Applied Biosystems, Foster City, CA, USA

- MultiScribe™ Reverse Transcriptase (cat. no. 4311235)
- TaqMan® Micro
- RNA reverse transcription kit
- TaqMan® Gene Expression Assays
- TaqMan® Universal PCR Master mix (No AmpErase® UNG, P/N 4324018)
- TaqMan miRNA Arrays
- GeneAmp PCR system 9700 thermal cycler
- ABI Prism 7000 and 7900 Sequence Detection System
- ABI Prism 7900HT instrument with TLDA upgrade
- Nuclease Free water
- Megaplex™ RT Primers

BD Pharmingen™

- Purified Mouse Anti-Human MLH-1 (# 550838)
- Purified Mouse Anti-PMS2 (# 556415)

BD Transduction Laboratories™

- Purified Mouse Anti-MSH6 (# 610919)

Becton Dickinson, New Jersey USA

- Vacutainer Serum Separator Tubes II

Calbiochem®, Germany

- Anti-MSH2 (Ab-2) Mouse mAb (FE11)(#NA27)

Eppendorf UK Ltd, Cambridge CB24 9ZR, United Kingdom

- Eppendorf 5417C Micro Centrifuge

GMI Inc. Ramsey, Minnesota, 55303 USA

- Refregerated centrifuge

Grenier Bio-one, St. Gallen, Switzerland

- Vacuette EDTA K3E blood bottles

Invitrogen, Calsbad, CA, USA

- SuperScript™ III RT (200U/µL, cat. no. 18080-093)
- RNaseOUT™ (cat. no. 10777019)
- ddH2O (cat. no.10977035)

KINEMATICA AG, CH-6014 Littau/Lucerne , Switzerland

- Homogenizer (Polytron® PT1600E)

Lennox Laboratory Supplies Ltd, Dublin, Ireland

- Formalin (Cat.no. CE110036)

Milestone, s.r.l, Sorisole, Italy

- JFC Solution (Code No. MW51408)
- microMed, Histoprocessor

Nanodrop Technologies, Wilmington, DE, USA

- NanoDrop™ 1000 spectrophotometer

Qiagen, Crawley, West Sussex, RH10 9NQ, United Kingdom

- QIAzol lysis reagent (cat no: 79306)
- QIAzol (cat no: 79306)
- RNeasy® Tissue Mini Kit (cat. no 74804)
- RNeasy® Plus Mini Kit (cat. no. 74106)
- RNeasy MinElute® Cleanup Kit (cat no. 74202)
- RNeasy® FFPE Kit(cat no. 74404)
- RNase-free DNase kit (cat. No. 79254)

Sigma-Aldrich® Co. Germany

- Hematoxylin (cat. no. H9627)
- Ammonium Hydroxide (cat. no. 320145)
- DPX Mountant (cat. no. 44581)
- Ethanol
- Xylene (cat. no. 534056)
- Isopropanol

Slee MAINZ, Germany

- Manual microtome (Cat.no.10065000)

Appendix

Ventana Medical Systems Inc. Tucson, AZ 85755 USA

- Ventana Discovery™ System
- UltraMap™ anti-Rb HRP (cat.no. 760-4315)
- UltraMap™ anti-Ms HRP (cat.no. 760-4313)
- DAB Map™ Detection Kit (cat.no. 760-124)

VWR International, LLC

- Industerial methylated Spirit
- Paraffin wax
- Superfrost® Plus slides(#631-0108)
- Acetone (#20066.321)
- S35 Feather blades (# 404011720)

Appendix 3: Details of miRNAs in TaqMan Human MiRNA Array A (384 miRNAs)

Well	Assay ID	Assay Name	Well	Assay ID	Assay Name
A1	000268	dme-miR-7	I1	002176	hsa-miR-933
A2	002909	hsa-miR-548I	I2	002177	hsa-miR-934
A3	000416	hsa-miR-30a-3p	I3	002178	hsa-miR-935
A4	000417	hsa-miR-30a-5p	I4	002179	hsa-miR-936
A5	000420	hsa-miR-30d	I5	002180	hsa-miR-937
A6	000422	hsa-miR-30e-3p	I6	002181	hsa-miR-938
A7	000427	hsa-miR-34b	I7	002182	hsa-miR-939
A8	000451	hsa-miR-126#	I8	002183	hsa-miR-941
A9	000478	hsa-miR-154#	I9	002185	hsa-miR-335#
A10	000483	hsa-miR-182#	I10	002187	hsa-miR-942
A11	001973	U6 snRNA	I11	002188	hsa-miR-943
A12	001973	U6 snRNA	I12	002189	hsa-miR-944
A13	000510	hsa-miR-206	I13	002196	hsa-miR-99b#
A14	000516	hsa-miR-213	I14	002197	hsa-miR-124#
A15	000534	hsa-miR-302c#	I15	002200	hsa-miR-541#
A16	000535	hsa-miR-302d	I16	002203	hsa-miR-875-5p
A17	000567	hsa-miR-378	I17	002213	hsa-miR-888#
A18	000570	hsa-miR-380-5p	I18	002214	hsa-miR-892b
A19	002910	hsa-miR-1257	I19	002231	hsa-miR-9#
A20	001011	hsa-miR-200a#	I20	002238	hsa-miR-411#
A21	001026	hsa-miR-432	I21	002243	hsa-miR-378
A22	001027	hsa-miR-432#	I22	002254	hsa-miR-151-3p
A23	001043	hsa-miR-497	I23	002259	hsa-miR-340#
A24	001046	hsa-miR-500	I24	002263	hsa-miR-190b
B1	002927	hsa-miR-1238	J1	002266	hsa-miR-545#
B2	001106	hsa-miR-488	J2	002270	hsa-miR-183#
B3	001113	hsa-miR-517#	J3	002272	hsa-miR-192#
B4	001149	hsa-miR-516-3p	J4	002274	hsa-miR-200b#
B5	001158	hsa-miR-518c#	J5	002286	hsa-miR-200c#
B6	001166	hsa-miR-519e#	J6	002287	hsa-miR-155#
B7	001170	hsa-miR-520h	J7	002288	hsa-miR-10a#
B8	001173	hsa-miR-524	J8	002293	hsa-miR-214#
B9	001178	mmu-let-7d#	J9	002294	hsa-miR-218-2#
B10	001283	hsa-miR-363#	J10	002298	hsa-miR-129#
B11	001973	U6 snRNA	J11	002301	hsa-miR-22#
B12	001973	U6 snRNA	J12	002302	hsa-miR-425#
B13	001338	rno-miR-7#	J13	002305	hsa-miR-30d#
B14	001510	hsa-miR-656	J14	002307	hsa-let-7a#
B15	001511	hsa-miR-549	J15	002309	hsa-miR-424#
B16	001512	hsa-miR-657	J16	002310	hsa-miR-18b#
B17	001513	hsa-miR-658	J17	002311	hsa-miR-20b#
B18	001514	hsa-miR-659	J18	002312	hsa-miR-431#

Appendix

Well	Assay ID	Assay Name	Well	Assay ID	Assay Name
B19	001519	hsa-miR-551a	J19	002314	hsa-miR-7-2#
B20	001520	hsa-miR-552	J20	002315	hsa-miR-10b#
B21	001521	hsa-miR-553	J21	002316	hsa-miR-34a#
B22	001522	hsa-miR-554	J22	002317	hsa-miR-181a-2#
B23	001523	hsa-miR-555	J23	002325	hsa-miR-744#
B24	001525	hsa-miR-557	J24	002330	hsa-miR-452#
C1	001526	hsa-miR-558	K1	002332	hsa-miR-409-3p
C2	001527	hsa-miR-559	K2	002333	hsa-miR-181c#
C3	001529	hsa-miR-562	K3	002336	hsa-miR-196a#
C4	001530	hsa-miR-563	K4	002339	hsa-miR-483-3p
C5	001531	hsa-miR-564	K5	002342	hsa-miR-708#
C6	001533	hsa-miR-566	K6	002343	hsa-miR-92b#
C7	001534	hsa-miR-567	K7	002346	hsa-miR-551b#
C8	001536	hsa-miR-569	K8	002362	hsa-miR-202#
C9	001539	hsa-miR-586	K9	002366	hsa-miR-193b#
C10	001540	hsa-miR-587	K10	002368	hsa-miR-497#
C11	001094	RNU44	K11	002371	hsa-miR-518e#
C12	001542	hsa-miR-588	K12	002376	hsa-miR-543
C13	001543	hsa-miR-589	K13	002378	hsa-miR-125b-1#
C14	001544	hsa-miR-550	K14	002379	hsa-miR-194#
C15	001545	hsa-miR-591	K15	002380	hsa-miR-106b#
C16	001546	hsa-miR-592	K16	002381	hsa-miR-302a#
C17	001547	hsa-miR-593	K17	002384	hsa-miR-519b-3p
C18	001550	hsa-miR-596	K18	002387	hsa-miR-518f#
C19	001553	hsa-miR-622	K19	002391	hsa-miR-374b#
C20	001554	hsa-miR-599	K20	002400	hsa-miR-520c-3p
C21	001555	hsa-miR-623	K21	002404	hsa-let-7b#
C22	001556	hsa-miR-600	K22	002405	hsa-let-7c#
C23	001557	hsa-miR-624	K23	002407	hsa-let-7e#
C24	001558	hsa-miR-601	K24	002410	hsa-miR-550
D1	001559	hsa-miR-626	L1	002411	hsa-miR-593
D2	001562	hsa-miR-629	L2	002417	hsa-let-7f-1#
D3	001563	hsa-miR-630	L3	002418	hsa-let-7f-2#
D4	001564	hsa-miR-631	L4	002419	hsa-miR-15a#
D5	001566	hsa-miR-603	L5	002420	hsa-miR-16-1#
D6	001567	hsa-miR-604	L6	002421	hsa-miR-17#
D7	001568	hsa-miR-605	L7	002423	hsa-miR-18a#
D8	001569	hsa-miR-606	L8	002424	hsa-miR-19a#
D9	001570	hsa-miR-607	L9	002425	hsa-miR-19b-1#
D10	001571	hsa-miR-608	L10	002432	hsa-miR-625#
D11	001573	hsa-miR-609	L11	002434	hsa-miR-628-3p
D12	001574	hsa-miR-633	L12	002437	hsa-miR-20a#
D13	001576	hsa-miR-634	L13	002438	hsa-miR-21#
D14	001578	hsa-miR-635	L14	002439	hsa-miR-23a#
D15	001581	hsa-miR-637	L15	002440	hsa-miR-24-1#
D16	001582	hsa-miR-638	L16	002441	hsa-miR-24-2#

Well	Assay ID	Assay Name	Well	Assay ID	Assay Name
D17	001583	hsa-miR-639	L17	002442	hsa-miR-25#
D18	001584	hsa-miR-640	L18	002443	hsa-miR-26a-1#
D19	001585	hsa-miR-641	L19	002444	hsa-miR-26b#
D20	001586	hsa-miR-613	L20	002445	hsa-miR-27a#
D21	001587	hsa-miR-614	L21	002447	hsa-miR-29a#
D22	001589	hsa-miR-616	L22	002642	hsa-miR-151-5P
D23	001591	hsa-miR-617	L23	002643	hsa-miR-765
D24	001594	hsa-miR-643	L24	002658	hsa-miR-338-5P
E1	001596	hsa-miR-644	M1	002672	hsa-miR-620
E2	001597	hsa-miR-645	M2	002675	hsa-miR-577
E3	001598	hsa-miR-621	M3	002676	hsa-miR-144
E4	001599	hsa-miR-646	M4	002677	hsa-miR-590-3P
E5	001600	hsa-miR-647	M5	002678	hsa-miR-191#
E6	001601	hsa-miR-648	M6	002681	hsa-miR-665
E7	001602	hsa-miR-649	M7	002743	hsa-miR-520D-3P
E8	001603	hsa-miR-650	M8	002752	hsa-miR-1224-3P
E9	001606	hsa-miR-661	M9	002867	hsa-miR-1305
E10	001607	hsa-miR-662	M10	002756	hsa-miR-513C
E11	001006	RNU48	M11	002757	hsa-miR-513B
E12	001613	hsa-miR-571	M12	002758	hsa-miR-1226#
E13	001614	hsa-miR-572	M13	002761	hsa-miR-1236
E14	001615	hsa-miR-573	M14	002763	hsa-miR-1228#
E15	001617	hsa-miR-575	M15	002766	hsa-miR-1225-3P
E16	001619	hsa-miR-578	M16	002768	hsa-miR-1233
E17	001621	hsa-miR-580	M17	002769	hsa-miR-1227
E18	001622	hsa-miR-581	M18	002773	hsa-miR-1286
E19	001623	hsa-miR-583	M19	002775	hsa-miR-548M
E20	001624	hsa-miR-584	M20	002776	hsa-miR-1179
E21	001625	hsa-miR-585	M21	002777	hsa-miR-1178
E22	001818	rno-miR-29c#	M22	002778	hsa-miR-1205
E23	001986	hsa-miR-766	M23	002779	hsa-miR-1271
E24	001987	hsa-miR-595	M24	002781	hsa-miR-1201
F1	001992	hsa-miR-668	N1	002783	hsa-miR-548J
F2	001993	hsa-miR-767-5p	N2	002784	hsa-miR-1263
F3	001995	hsa-miR-767-3p	N3	002785	hsa-miR-1294
F4	001996	hsa-miR-454#	N4	002789	hsa-miR-1269
F5	001998	hsa-miR-769-5p	N5	002790	hsa-miR-1265
F6	002002	hsa-miR-770-5p	N6	002791	hsa-miR-1244
F7	002003	hsa-miR-769-3p	N7	002792	hsa-miR-1303
F8	002004	hsa-miR-802	N8	002796	hsa-miR-1259
F9	002005	hsa-miR-675	N9	002798	hsa-miR-548P
F10	002087	hsa-miR-505#	N10	002799	hsa-miR-1264
F11	002094	hsa-miR-218-1#	N11	002801	hsa-miR-1255B
F12	002096	hsa-miR-221#	N12	002803	hsa-miR-1282
F13	002097	hsa-miR-222#	N13	002805	hsa-miR-1255A
F14	002098	hsa-miR-223#	N14	002807	hsa-miR-1270

Well	Assay ID	Assay Name	Well	Assay ID	Assay Name
F15	002100	hsa-miR-136#	N15	002810	hsa-miR-1197
F16	002102	hsa-miR-34b	N16	002815	hsa-miR-1324
F17	002104	hsa-miR-185#	N17	002816	hsa-miR-548H
F18	002105	hsa-miR-186#	N18	002818	hsa-miR-1254
F19	002107	hsa-miR-195#	N19	002819	hsa-miR-548K
F20	002108	hsa-miR-30c-1#	N20	002820	hsa-miR-1251
F21	002110	hsa-miR-30c-2#	N21	002822	hsa-miR-1285
F22	002111	hsa-miR-32#	N22	002823	hsa-miR-1245
F23	002113	hsa-miR-31#	N23	002824	hsa-miR-1292
F24	002114	hsa-miR-130b#	N24	002827	hsa-miR-1301
G1	002115	hsa-miR-26a-2#	O1	002829	hsa-miR-1200
G2	002116	hsa-miR-361-3p	O2	002830	hsa-miR-1182
G3	002118	hsa-let-7g#	O3	002832	hsa-miR-1288
G4	002119	hsa-miR-302b#	O4	002838	hsa-miR-1291
G5	002120	hsa-miR-302d#	O5	002840	hsa-miR-1275
G6	002121	hsa-miR-367#	O6	002841	hsa-miR-1183
G7	002125	hsa-miR-374a#	O7	002842	hsa-miR-1184
G8	002126	hsa-miR-23b#	O8	002843	hsa-miR-1276
G9	002127	hsa-miR-376a#	O9	002844	hsa-miR-320B
G10	002128	hsa-miR-377#	O10	002845	hsa-miR-1272
G11	000338	ath-miR159a	O11	002847	hsa-miR-1180
G12	002129	hsa-miR-30b#	O12	002850	hsa-miR-1256
G13	002130	hsa-miR-122#	O13	002851	hsa-miR-1278
G14	002131	hsa-miR-130a#	O14	002852	hsa-miR-1262
G15	002132	hsa-miR-132#	O15	002854	hsa-miR-1243
G16	002134	hsa-miR-148a#	O16	002857	hsa-miR-663B
G17	002135	hsa-miR-33a	O17	002860	hsa-miR-1252
G18	002136	hsa-miR-33a#	O18	002861	hsa-miR-1298
G19	002137	hsa-miR-92a-1#	O19	002863	hsa-miR-1290
G20	002138	hsa-miR-92a-2#	O20	002868	hsa-miR-1249
G21	002139	hsa-miR-93#	O21	002870	hsa-miR-1248
G22	002140	hsa-miR-96#	O22	002871	hsa-miR-1289
G23	002141	hsa-miR-99a#	O23	002872	hsa-miR-1204
G24	002142	hsa-miR-100#	O24	002873	hsa-miR-1826
H1	002143	hsa-miR-101#	P1	002874	hsa-miR-1304
H2	002144	hsa-miR-138-2#	P2	002877	hsa-miR-1203
H3	002145	hsa-miR-141#	P3	002878	hsa-miR-1206
H4	002146	hsa-miR-143#	P4	002879	hsa-miR-548G
H5	002148	hsa-miR-144#	P5	002880	hsa-miR-1208
H6	002149	hsa-miR-145#	P6	002881	hsa-miR-548E
H7	002150	hsa-miR-920	P7	002883	hsa-miR-1274A
H8	002151	hsa-miR-921	P8	002884	hsa-miR-1274B
H9	002152	hsa-miR-922	P9	002885	hsa-miR-1267
H10	002154	hsa-miR-924	P10	002887	hsa-miR-1250
H11	002157	hsa-miR-337-3p	P11	002888	hsa-miR-548N
H12	002158	hsa-miR-125b-2#	P12	002890	hsa-miR-1283

Well	Assay ID	Assay Name	Well	Assay ID	Assay Name
H13	002159	hsa-miR-135b#	P13	002893	hsa-miR-1247
H14	002160	hsa-miR-148b#	P14	002894	hsa-miR-1253
H15	002163	hsa-miR-146a#	P15	002895	hsa-miR-720
H16	002164	hsa-miR-149#	P16	002896	hsa-miR-1260
H17	002165	hsa-miR-29b-1#	P17	002897	hsa-miR-664
H18	002166	hsa-miR-29b-2#	P18	002901	hsa-miR-1302
H19	002168	hsa-miR-105#	P19	002902	hsa-miR-1300
H20	002170	hsa-miR-106a#	P20	002903	hsa-miR-1284
H21	002171	hsa-miR-16-2#	P21	002904	hsa-miR-548L
H22	002172	hsa-let-7i#	P22	002905	hsa-miR-1293
H23	002173	hsa-miR-15b#	P23	002907	hsa-miR-1825
H24	002174	hsa-miR-27b#	P24	002908	hsa-miR-1296