DEVELOPMENT OF EPIGENETIC-BASED PROGNOSTIC BIOMARKERS TO STRATIFY RISK OF RECURRENCE IN EARLY RECTAL CANCER

by

KAI JUEN LEONG

A thesis submitted to the University of Birmingham for the degree of DOCTOR OF PHILOSOPHY

School of Cancer Sciences

College of Medical and Dental Sciences

University of Birmingham

November 2013

Abstract

Organ-preservation surgery for early rectal cancer is an alternative treatment to conventional radical surgery that offers comparable oncological outcome but has lower morbidity and mortality. The efficacy of this strategy relies on accurate preoperative staging of mesorectal nodal metastasis which current staging modalities are unable to deliver. DNA methylation has a prognostic role in colorectal carcinogenesis and may select for tumours suitable for local excision, although this has not been explored in rectal cancer.

A panel of staged rectal cancers was analysed for differential DNA methylation patterns. This identified a unique signature consisting of concomitant hypermethylation of three or more of *APC*, *RARB, GSTP1, TIMP3, CASP8, DAPK1* and *CXCL12* that was associated with histopathologically localised disease. Correlation of protein expression with clinicopathological features found UNC5C expression to be associated with nodal status. Genome-wide screening found >7000 differentially methylated genes between node-negative and node-positive rectal cancer. Validation was performed on a subset of these genes and this confirmed hypermethylation of *SNAP25, SOX7* and *TIAM1* to be associated with favourable histological indices.

This study has provided insight into the methylation patterns of rectal cancer and has identified novel prognostic biomarkers. Further work will determine their clinical usefulness in rectal cancer risk stratification.

To my wife, Deborah and our beautiful daughter, Ashleigh

Acknowledgements

The completion of this thesis would not have been possible without the incredible help and support from my two supervisors, Glenn Matthews and Simon Bach. My heartfelt thanks also go to Professor Dion Morton for his advice and relentless support. I am indebted to Germaine Caldwell, Lesley Tannahill, Carolyn Jones and Tessa Webb for their patience, support and 'sisterly/motherly' advice. Jonathan James has been an incredible technician and I am grateful for his help.

I would also like to thank Jennifer Anderton and Sarah Leonard for their expertise on pyrosequencing. I thoroughly enjoyed their company, particularly when troubleshooting the pyrosequencing machine. I am also grateful for the help and advice received from Kai Sheng and Merlin. I would like to thank Wenbin Wei for his statistical support and patience in explaining his bioinformatic analysis to me.

My gratitude goes to Philippe Taniere, Rahul Hejmadi, John Gregory and Brendan O'Sullivan for the tremendous support and kindness I received while working at the Department of Histopathology Queen Elizabeth Hospital.

Finally, I am grateful to the Medical Research Council for the award of the prestigious Clinical Research Training Fellowship.

Table of contents

List of Figures

Chapter 7 Genome-wide promoter and CpG island methylation alterations in multistage colorectal tumourigenesis

List of Tables

List of Abbreviations

ADP	Adenosine diphosphate
bp	Base pair
cm	centimetre
DNA	Deoxyribonucleic acid
EGF	Epidermal growth factor
HER2	Heregulin 2
KRAS	Kirsten rat sarcoma viral oncogene homolog
miRNA	Micro RNA
mRNA	Messenger RNA
PCR	Polymerase chain reaction
RNA	Ribonucleic acid
UTR	Untranslated region

Chapter 1

Introduction

1.1 Epidemiology of colorectal cancer

Colorectal cancer (CRC) poses a significant morbidity, mortality and economic burden to the world. Globally, its estimated 1.2 million new cases per year makes CRC the third most common cancer in men (after lung and prostate) and second (after breast) in women[1]. Geographical variations in incidence rates are observed, with almost 60% of cases occurring in developed countries. It is currently the forth most common cause of cancer deaths and accounts for over half a million deaths per annum worldwide[1].

In the United Kingdom, 41,142 new cases were diagnosed in 2009, of which approximately a third were rectal cancer[2]. The incidence is age-dependent, with 84% of cases occurring in adults over the age of 60. The lifetime risk of developing this disease in the UK is 1 in 16 for men and 1 in 20 for women. Despite the fact that overall mortality has been falling since the early 1990s, it is still the second most common cause of cancer death in the UK after lung cancer[2]. Survival is inversely related to the stage of the disease at the time of diagnosis. The relative 5-year survival for those with Dukes' A (TNM Stage 1) is 93.2%. In contrast, those with Dukes' 'D' have a 5-year relative survival of only 6.6% (Table 1.1)[3]. More discouragingly, the clinical stage of symptomatic patients at the time of diagnosis is usually late (Dukes B and C, see Table 1.1). This not only highlights the importance of detecting this disease early but also targeting the population at the presymptomatic level.

Duke's stage at diagnosis	Percentage of cases	Five-year relative survival
A	8.7%	93.2%
B	24.2%	77.0%
C	23.6%	47.7%
D	9.2%	6.6%
Unknown	34.3%	35.4%

Table 1.1. Percentage of cases and five year relative survival (%) by Dukes' stage at diagnosis, colorectal cancer patients diagnosed 1996-2002, England. Taken from http://www.cancerresearchuk.org/cancer-info/cancerstats/types/bowel/survival/, accessed 31 October 2012.

1.2 Screening and early detection of colorectal cancer in the United Kingdom

CRC screening is aimed at populations with average risk of developing CRC. Those at a higher risk of CRC have either inherited forms of CRC (Hereditary Non-polyposis Colorectal Cancer/Lynch syndrome and Familial Adenomatous Polyposis being the two most common) or a strong family history of CRC, usually with an autosomal pattern of inheritance (familial clustering). The surveillance modality of choice is usually colonoscopy and in general, the frequency of surveillance is tailored according to individual risks. A detailed discussion on the surveillance for high-risk population is beyond the scope of this thesis.

The NHS National Bowel Cancer Screening Programme was introduced in 2006 and has now achieved nationwide coverage. It targets the average risk population aged between 60-69 years old, as over 80% of CRCs have been shown to occur in people in this age group and beyond[2]. Since 2008, screening centres have started to extend the screening age group to 75 years old, in accordance with the government's strategy to improve cancer outcomes[4]. The current screening tool utilises the guaiac-based faecal occult blood test (G-FOBt), which detects the pseudoperoxidase activity

of haem in faeces. It is, however, not specific to human blood and can be affected by the consumption of certain food (e.g. red meat, grapefruit, horseradish and figs leading to false positive). Individuals with positive results are invited for screening colonoscopies.

G-FOBt's relatively low sensitivity (7.2-12.9%)[5, 6] in detecting colorectal neoplasia has prompted other screening modalities such as immunochemical-based faecal occult blood test and stool-extracted DNA assays to be developed to improve the performance of G-FOBt but they lack the Level I evidence demonstrated by G-FOBt in terms of reduction in colorectal cancer mortality. Screening with G-FOBt has been shown to reduce colorectal cancer mortality in four large population-based randomised control trials after more than 10 years of follow up (Table 1.2)[7]. The results from the UK and Denmark were directly comparable. The apparent improved results in the US trial were ascribed to a far higher colonoscopy rate, which in turn was due to more frequent FOB testing and a lower positive test threshold.

Study	Country	Screening frequency	Age range (yr)	Follow up period (yr)	RR	95% CI
Funen	Denmark	Annual	45-75	17	0.84	0.73-0.96
		Biennial			0.89	0.78-1.01
Minnesota	U.S.	Annual	50-80	18	0.67	0.51-0.83
		Biennial			0.79	0.62-0.97
Nottingham	U.K.	Biennial	45-74	11.7	0.87	0.78-0.97
Goteberg	Sweden	Biennial	60-64	19	0.84	0.71-0.99

Table 1.2. Characteristics, Relative Risk (RR) and Confidence Interval (CI) colorectal cancer mortality of 4 randomised controlled trials using FOBT[7]. Yr=Year.

More recently, a randomised controlled trial in the UK has shown that a one-off screening with flexible sigmoidoscopy results in reduction in CRC incidence and CRC-related mortality by 23% and 31% respectively[8]. This was evident after a median follow-up of 11 years. The results are supported by two other large multicentre randomised trials; one in Italy (SCORE trial)[9] and the other in the US (PLCO trial)[10]. A similar trial in Norway trial did not demonstrate any reduction in CRC incidence and mortality but this could be due to an inadequate length of follow up of only 7 years[11]. Results from the UK trial have prompted the NHS to consider offering flexible sigmoidoscopy screening to asymptomatic adults between the age of 55 to 60, after which they would be offered screening using FOBT. Further details of its nationwide introduction are awaited.

The screening programme is changing how CRC presents. Population-based screening studies in the UK report that approximately 50% of screen detected tumours are 'early' (pT1-2N0M0; Stage I)[12, 13]. This is in marked contrast with the pre-screening era whereby, <10% of CRCs diagnosed were Dukes' A and the majority of cancers were either Dukes' B or C (Table 1.1). Such a shift, favouring the detection of early stage cancers, introduces an interesting management dilemma. The strategy in CRC management has evolved from treating advanced cancer, quite often compromising patient quality of life and function in exchange for satisfactory oncological outcome. Whether such a debilitating strategy is appropriate for early rectal cancers is debatable and is explored in the following section.

1.3 Management strategy for early rectal cancer

The management strategy for rectal cancer in the last 20 years has evolved from a surgeon-managed condition to a multidisciplinary approach, encompassing radiologists, pathologists, oncologists, specialist nurses and surgeons. For discussion purposes, the treatment modality for rectal cancer can be broadly divided into radiotherapy, chemotherapy and surgery but in practice, a multimodal approach is frequently tailored according to individual patients.

1.3.1 Preoperative staging of rectal cancer

Accurate preoperative staging is paramount in the clinical decision making for rectal cancer. The main purpose of preoperative staging is to identify prognostic features that that will stratify patients into different treatment groups. Preoperative staging is performed in accordance to the TNM classification and is divided into the assessment of local invasion (mural wall invasion, circumferential resection margin and mesorectal nodal involvement) and distant metastasis. In addition, clinical assessment of the tumour by the surgeon in the form of digital examination provides important anatomical information such as fixity to the bowel wall, distance from the anal verge, size and location that may aid surgical decision-making. The imaging modalities currently used to stage rectal cancers in the UK are computed tomography (CT), magnetic resonance imaging (MRI) and endorectal ultrasound (EUS). Positron emission tomography (PET) with and without CT fusion is not routinely used in the UK. A combination of these modalities is required to achieve complete staging.

1.3.1.1 The Tumour stage (T-stage)/extramural depth of invasion

EUS and MRI offer accurate assessment of the depth of tumour growth into the rectal wall. Studies that have directly compared EUS and MRI in T-staging have shown similar accuracies between these two modalities in the region of 77% - 85%[14-16]. The use of conventional CT protocols with low spatial and contrast resolution, in the past, has resulted in poor accuracy rates of between 52%-72% for T-staging[14, 17]. This accuracy has since been improved to over 80%, following the introduction of multislice CT[18, 19]. A meta-analysis of the three imaging modalities in rectal cancer have found that EUS and MRI are very sensitive (94%) in the staging of T1 and T2 tumours with the specificity of EUS being superior to that of MRI (86% vs. 69%)[20]. Their accuracy, however, diminishes in the staging of advanced T3 tumours.

There is a heterogeneous survival range in T3 tumours, which still make up a large proportion of rectal tumours seen in clinical practice. It has been shown that survival is inversely related to the depth of extramural spread, independent of other prognostic factors[21, 22]. The prospective multicentre MERCURY study has shown that thin-slice high resolution MRI with phased array coils was highly accurate (92.5%) in the prediction of extramural depth of invasion. Preoperative MRI prediction of extramural spread was equivalent to histopathological assessment to within 0.5mm[23].

1.3.1.2 The Nodal stage (N-stage)

The significance of the information gained from preoperative assessment of lymph node status in rectal cancer could not be overemphasised. The presence and

location of involved lymph nodes defines the treatment strategy. Firstly, tumours with involved lymph nodes, regardless of their T-stage, are not suitable for organ-conservation strategy. Secondly, tumours with metastatic lymph nodes encroaching the mesorectal fascia, which forms the surgical circumferential resection margin, will require downstaging with neoadjuvant chemoradiation prior to surgery due to the increase risk of local recurrence following surgery[24].

Despite advances in imaging, the ability to predict lymph node involvement in rectal cancer remains a challenge. The main issue is that many metastatic lymph nodes are small (<5mm) and indistinguishable from normal or reactive nodes due to a considerable size overlap between benign and malignant nodes[25, 26]. Two meta-analyses have shown that EUS, CT and MRI are poor in predicting lymph node involvement in rectal cancer[20, 27]. Using intranodal signal heterogeneity on MRI, the prediction of lymph node involvement could achieve a sensitivity of 85% and specificity of 97%[25, 28] but this has not been reproducible by other groups. Novel imaging such as MRI enhanced with ultrasmall superparamagnetic iron oxide (USPIO) has failed to demonstrate any improvement in lymph node prediction[29]. To date, the inability to detect microscopic lymph node metastasis suggests that a radiological N0 staging should not be recommended to select patients for organ conservation strategy as a curative treatment.

1.3.1.3 Circumferential resection margin (CRM)

Involvement of the lateral or, more recently termed, circumferential resection margin (CRM) is the most important risk factor for local recurrence after rectal

cancer surgery[30]. Patients preoperatively staged to have threatened CRMs will require neoadjuvant chemoradiotherapy rather than undergoing surgery directly. A meta-analysis has shown that MRI is highly accurate in the prediction of CRM involvement with sensitivity of 94% and specificity of 85%[31]. The multicentre MERCURY study also confirmed the accurate prediction of CRM status using MRI with a specificity of 92.4% and a negative predictive value of 94%[32]. Consequently, MRI is now the imaging modality of choice to predict CRM involvement.

1.3.2 Surgery for early rectal cancer

Surgery for rectal cancer is more challenging than that for colonic cancer due to the difficulty in accessing the tumour, which is located deep in the narrow pelvis. Its close proximity to the anal sphincter complex makes this a major consideration when deciding the appropriate surgical approach.

1.3.2.1 Total mesorectal excision (TME)

Radical surgery with total mesorectal excision (TME) surgery, described over 20 years ago by Heald[33], is regarded as the 'gold standard' surgery for rectal cancer. TME involves the sharp dissection along the mesorectal plane, leading to the removal of the mesorectum enveloping the rectal tumour, mesorectal lymph nodes and vascular supply. The Dutch TME trial showed that the introduction of TME reduced the local recurrence rate from 16% to 9% and that TME was an independent predictor of overall survival[34]. Reports from the combined MRC CR07 and NCIC-CTG CO16 randomised trials found that the plane of rectal cancer surgery had a significant impact on local recurrence. At 3 years, the estimated local recurrence

rates were 4% for mesorectal plane, 7% for intramesorectal plane, and 13% for muscularis propria plane[35]. For low rectal cancer, an extralevator abdominoperineal excision (ELAP), which incorporates the resection of the levator muscles to avoid a "coning effect" deep in the pelvis, may be required to achieve a clear CRM and improved local recurrence rates[36].

While radical TME surgery for early rectal cancer offers high rates of cure, it comes at a cost. 6-month mortality figures following radical TME is 4.6% for those aged 65-74 and rise to 13.4% in patients aged 75-84[37]. Permanent stoma rates of up to 37% for low anterior resection have been reported[38, 39]. The Dutch TME trial reported clinical anastomotic leaks in 16% of non-irradiated patients[34]. Over half of all patients suffer some form of faecal incontinence following TME and a third experience daily symptoms of urgency, incomplete emptying and stool frequency[40, 41]. Three prospective cohort studies have demonstrated poor health-related quality of life scores following rectal cancer surgery[42-44]. Therefore, it is questionable whether this level of surgical morbidity and mortality is acceptable in exchange for the satisfactory treatment of early rectal cancer, particularly in the elderly population. An organ preserving local approach may provide significantly less morbidity without substantially compromising oncological outcomes.

1.3.2.2 Local excision for early rectal cancer

1.3.2.2.1 Endoscopic mucosal resection (EMR)

Endoscopic mucosal resection has been evolved to treat early rectal cancer following its success in the treatment of early gastric cancer, with reported 5-year disease-free

survival rates of 97%[39]. This technique is appealing as it can be performed under sedation using an endoscope and does not require expensive equipment. During EMR, the tumour is lifted away from the underlying muscle by infiltrating its base with normal saline or colloid before being resected with a diathermy loop. Its biggest drawback, as its name suggests, is that the maximal depth of excision is in the submucosal layer. A full histological assessment of the depth of invasion of the excised tumour is, therefore, not possible and consequently, the use of EMR has largely been reserved for premalignant adenomas. Furthermore, removal of lesions greater than 2 cm often requires piecemeal resection that increases the risk of recurrence[45, 46]. In a selected group of early rectal cancer, however, EMR has been shown to be safe with minimal mortality and is associated with a low recurrence rate[46].

1.3.2.2.2 Transanal Endoscopic Microsurgery

Early rectal cancers can be locally excised via the anus with low morbidity and mortality using Transanal Endoscopic Microsurgery (TEMS)[47, 48]. Its advantage over EMR is that a full thickness local excision is possible, enabling histopathological assessment on the depth of invasion. Similar to EMR, TEMS allows for preservation of the rectum but risks neglecting potential microscopic lymph node metastases *in situ*, as the mesorectum is left behind. As previously discussed, current imaging modalities are unable to reliably detect microscopic lymph node metastases. Current predictive histopathological features of local failure lack precision and are only able to identify a minority of early rectal cancers suitable for local excision[47], although over half of early rectal cancers may be adequately treated using an organ

preservation strategy[49]. Highly accurate predictive markers are needed to further stratify for these tumours.

1.3.3 Non-operative options for early rectal cancer

1.3.3.1 Contact radiotherapy

Contact radiotherapy delivers radiation via an endoluminal technique using an X-ray tube that produces a 50-kV beam. This technique was popularised for clinical use in rectal cancer by Papillon from Lyon in the 1950s[50]. The primary advantage of this technique is the ability to deliver high doses of radiotherapy directly to the tumour with minimal radiation exposure to adjacent normal tissues. Papillon has shown that this approach is feasible for T1-T2 N0 rectal tumours, with a reported 5-year overall survival of 74% and a local failure rate of only 5%[50]. Gerard *et. al.* subsequently reported in a series of 101 T1-T2 rectal tumours that a complete response was achievable following contact radiotherapy and selective brachytherapy, with a local failure rate of 14% and a 5-year survival rate of 83%[51]. In the UK, contact radiotherapy is only provided at single centre in Clatterbridge. Sun Myint *et. al.* use contact radiotherapy as part of a multimodal treatment approach. Tumours that do not respond well to initial contact radiotherapy are offered either external beam radiotherapy or chemoradiotherapy. They reported a local failure rate of 7% and a 5-year overall survival rate of 70%[52]. There is still a lack of randomised trial evidence to support the use of contact radiotherapy as an alternative curative treatment to radical surgery for early rectal cancer and published results have thus far come from single institutions.

1.3.3.2 External beam radiotherapy

Results from a contemporary series of operable distal rectal cancer treated with neoadjuvant chemoradiation by Habr-Gama *et. al.* suggest that this non-operative treatment is a viable option[53]. In their study of 265 patients comprising 4.5% T2 disease, 86.5% T3 and 9%T4, 72 patients (27%) achieved a complete clinical response. With a mean follow up of 57 months (range 18-156 months), two patients developed local recurrence and were successfully salvaged by surgery or brachytherapy. The rate of pathological complete response (pCR) following radical surgery in those with an incomplete clinical response was 8% (22 patients). The 5-year disease-free survival was similar for the non-operated and the operated groups (92% vs. 83%). While such results are encouraging, it has not been reproducible by other groups.

1.3.4 Multimodal organ preservation strategies for early rectal cancer

Evidence thus far supports the use of local excision or non-operative options alone for a carefully selected minority of patients with early rectal cancer or for those who are not fit for radical surgery. For the remainder of patients with early rectal cancer, a combination of these approaches may be required to achieve similar oncological results as TME surgery but without the morbidity and mortality associated with radical surgery.

1.3.4.1 Local excision and adjuvant chemoradiation therapy

There are no randomised controlled trials that have evaluated the efficacy of adjuvant therapy after local excision. A retrospective study by Chakravarti *et. al.*,

found that adjuvant chemoradiation following local excision of T1-2 rectal cancer did not confer any benefit in terms of local recurrence and disease free survival[54]. The Radiation Oncology Therapy Group study found that adjuvant chemoradiation therapy after local excision did not improve loco-regional control[55]. Evidence to date does not support the use of adjuvant chemoradiation following local excision of early rectal cancer.

1.3.4.2 Preoperative radiotherapy and local excision

Given than four large randomised controlled trials have shown that the addition of preoperative radiotherapy to radical TME surgery reduces the incidence of local recurrence[56-59], numerous studies have subsequently evaluated the efficacy of neoadjuvant chemoradiotherapy as part of an organ preservation strategy. A meta-analysis of seven studies on neoadjuvant chemoradiotherapy followed by local excision to treat 237 patients with T2-3 rectal cancers, reported a pCR of 22% with no recurrence found in this cohort[60]. Downstaged ypT1 and ypT2 had local recurrences of only 2% and 6%, respectively. More recently, a randomised trial of neoadjuvant chemoradiotherapy followed by either local resection using TEM or laparoscopic total mesorectal excision for T2 rectal cancer found similar local recurrence rates between the two groups (8% vs. 6%)[61]. While this trial result is promising, it is not applicable to the UK population, since the current treatment for T2N0 rectal cancer in the UK is TME surgery alone without neoadjuvant therapy. Results from the Transanal Endoscopic Microsurgery and Radiotherapy in Early Rectal Cancer (TREC) trial should provide meaningful insight into the efficacy of this

organ preservation strategy compared to the gold standard treatment of radical TME surgery alone for early rectal cancer.

There is limited evidence to guide the choice of either short course preoperative radiotherapy (SCPRT) or neoadjuvant chemoradiotherapy (nCRT) as part of an organ preservation strategy. Bujko *et. al.* randomised 47 patients with predominantly T1 and T2 rectal tumours (borderline T3 allowed) to either SCPRT or nCRT followed by delayed local excision and found that pCR was 35% (11/31) following SCPRT and 54% (7/13) following nCRT[62]. Grades I–II acute radiation toxicity was seen in 35% (11/31) of patients after SCPRT and 64% (9/14) in patients after nCRT. Grade III toxicity was seen in one patient in nCRT group and none in the SCPRT group. It has to be emphasised that the results of this study should be interpreted with caution since the authors failed to recruit the necessary number of patients. Other studies that have evaluated the two different radiotherapy schedules appear to favour SCPRT over nCRT due to the lower incidence of acute severe toxicity following SCPRT[56, 58, 63]. Another issue of contention is the timing of surgery following neoadjuvant therapy. There are suggestions that SCPRT may downstage locally advanced rectal cancers if surgery is delayed[64, 65]. This possibility is currently being prospectively evaluated as part of the Stockholm III study.

1.3.5 Summary

The gold standard radical TME surgery offers high rates of cure for rectal cancer but is associated with significant morbidity, mortality and poor quality of life. An organ preservation strategy of neoadjuvant therapy and local excision for early rectal

cancer provides an attractive alternative that is low in morbidity, mortality and without substantially compromising oncological outcomes. The efficacy of this approach, however, is dependent on two important factors: the accurate preoperative staging of nodal disease and the identification of rectal cancers that will respond to radiotherapy. Current imaging modalities and conventional histopathological features are not sufficiently sensitive to identify microscopic nodal metastasis and there are no predictive markers that can select for radiosensitive tumours. This clinical need remains unmet and there is an urgency to develop accurate predictive markers that can stratify for likelihood of microscopic mesorectal nodal metastasis and response to radiotherapy in rectal cancer.

1.4 Molecular markers in clinical practice

A biomarker is 'any molecular characteristic that may be used to stratify patients for treatment benefit within clinical trials, prognosticate patient outcome or predict and/or monitor response to therapy'[66]. The main aim of biomarker discovery is to identify the appropriate patient group to treat and the distinct treatment modality to offer that will result in improved patient outcome compared to standard or no treatment. Over the years, numerous biomarker discovery programmes have led to the introduction of these biomarkers in clinical practice. For example, testing for the HER2 receptor status in patients with breast cancer for Trastuzumab (Herceptin®) treatment is now routinely performed. EGF receptor antagonists are offered selectively for patients with colorectal cancer expressing the wild-type *KRAS*. The National Institute of Clinical Excellence recommends the routine testing for EGF receptor tyrosine kinase mutations in patients with locally advanced or metastatic

non-small cell lung cancer for consideration of Erlotinib, an EGF receptor tyrosine kinase inhibitor. More recently, studies have shown that somatic loss of MLH1 or MSH2 expression in colorectal cancers correlates with resistance to fluoropyrimidine-based chemotherapy[67, 68].

To facilitate the successful transfer of biomarkers from 'bench to clinic', Cancer Research UK (CRUK) has developed well-structured biomarker roadmaps for scientists and clinicians[69]. These roadmaps describe the process from initial biomarker discovery, assay development, initial correlation to clinical outcome and clinical qualification. In the early phase of the roadmap, also known as the rationale stage, the biomarker discovery process should address three important issues; an unmet clinical need, the work is primarily focussed on clinical material and there is a sample acquisition process in place for clinical outcome correlation studies. The work in this thesis was produced in accordance with such roadmap. Firstly, in the absence of reliable clinical markers, there is an unmet need to develop biomarkers associated with disease progression in early stage rectal cancer to guide surgical treatment options. Secondly, the molecular markers identified in this study are performed on clinical samples. Thirdly, there is an existing biobank in the University of Birmingham, which contains hundreds of fresh frozen and formalin-fixed paraffin-embedded (FFPE) colorectal tissues for clinical outcome correlation studies. For an accurate selection of clinically important biomarkers, an in-depth understanding of the relevant molecular biology of colorectal cancer is central.

1.5 Molecular biology of colorectal cancer

Colorectal cancer is the product of a multistep accumulation of genetic and epigenetic aberrations that lead to the transformation of normal colonic epithelium into adenocarcinoma. This sequential process is widely believed to drive the initiation and progression of benign adenomas to malignant carcinomas as these genetic alterations often affect key signalling pathways that control the hallmarks of tumour behaviour[70, 71]. For many years, genetic alterations and how they lead to genomic instability have been the main focus in scientific research with the study of epigenetics taking a 'backseat'. Our appreciation of the epigenetic landscape has greatly expanded in recent years, partly due to the growing awareness of epigenetic inheritance is critical in many cellular processes such as gene transcription and differentiation and more importantly, of how deregulation of epigenetics can lead to cancer. In addition, rapid advances in technology have enabled us to investigate epigenetic alterations at a greater depth. Emerging evidence shows that epigenetic changes are involved in every step of tumourigenesis[72] and may have a clinical role in the diagnosis, monitoring of therapy and prognosis of cancer.

The current molecular landscape of colorectal cancer divides tumourigenesis into three distinct pathways; epigenetic, chromosomal and microsatellite instability pathways, which is an oversimplification as evidence has shown that these pathways often overlap. In the next section, these individual pathways and how they interact mechanistically to drive tumourigenesis are discussed. There is growing evidence that epigenetic alterations are key players in tumourigenesis and often influence genetic changes during tumour development. An understanding of the mechanisms

of epigenetic deregulation in colorectal cancer could provide important insight into the clinical relevance of this pathway.

1.5.1 Chromosomal Instability Pathway (CIN)

Loss of genetic stability is a frequent and key mechanism in most solid tumour formation[73]. Genomic instability can take the shape of two forms; at the chromosomal level, resulting in gains or losses of whole or large parts of the chromosomes (aneuploidy) and at the nucleotide level and results in base deletions, substitutions and insertions. Tumours exhibiting alterations at the chromosomal level are thought to arise via the chromosomal instability pathway (CIN). This pathway is by far the most common genomic instability reported in colorectal cancer; observed in up to 85% of tumours[74]. Various mechanisms responsible for chromosomal instability have been postulated and are discussed below.

1.5.1.1 Proposed mechanisms responsible for CIN

A) Mutations in mitotic spindle checkpoint genes

Gene products of *MAD* and *BUB* act as checkpoint sensors and signal transducers that regulate sister chromatid separation at the metaphase-anaphase transition. They ensure correct chromosome segregation by delaying the start of anaphase until duplicated chromatids are aligned on the metaphase plate. Activation of these proteins leads to cell cycle arrest through inhibition of the anaphase-promoting complex/C (APC/C). The active state of APC/C results in degradation of securin, which interacts with separase. Separase, in turn, regulates cohesin, which creates physical links with sister chromatids. Defects in this cascade checkpoint results in

chromosome missegregation, leading to subsequent aneuploidy. Colon cancer cell lines with mutations in *BuBR1*/*MAD3* and *BUB1* have been shown to fail cycle arrest when spindles were depolarized[75]. Other genes involved in the mitotic spindle checkpoint such as *MAD1* and *MAD2* have also been found mutated in leukaemia and breast cancers[76, 77] but as yet, not in colorectal cancer.

B) Centrosome deregulation

A strong association between abnormal centrosome number and aneuploidy exists in cancers[78, 79]. Centrosomes provide anchorage for the reorganization of microtubules into a mitotic spindle apparatus. It has been postulated that centrosome amplification lead to multiple spindle poles during mitosis and subsequent chromosome missegregation[80]. Evidence for this was provided in two studies that found overexpression of *STK15*, a centrosome-associated serine-threonine kinase, in breast cancer and colorectal cancer cell lines[81, 82]. This was further supported by other studies, which found overexpression of members of the polo family (*PLK1, PLK2* and *PLK4*) implicated in aneuploidy and abnormal centrosome function in colorectal cancer[83, 84].

C) Mutations in DNA damage repair genes

The DNA damage response system protects cells from exogenous and endogenous genotoxic stress by arresting cell cycle to allow adequate time for damage repair or in the case whereby DNA damage is perceived as irreparable, by inducing senescence and apoptosis. Inactivating mutations of genes involved in DNA repair predisposes individuals to syndromes with increased susceptibility to cancer.

Examples of these syndromes include hereditary breast-ovarian cancer (*BRCA1* and *BRCA2* mutations), ataxia telangiectasia (*ATM* mutation) and Li-Fraumeni (*TP53* mutation). An early indication linking replication checkpoint errors and genomic instability was derived from the observation that genomic rearrangements seen in primary cancers were similar to that seen in yeast when DNA repair genes were inactivated[85]. To date, *TP53* has been the only checkpoint genes directly implicated in colorectal cancer with at least a permissive role in CIN development. Wang *et. al.* performed mutation analysis of 100 genes to identify regulators of CIN in colorectal cancer and found 19 mutations in 5 genes representing three distinct pathways. One of these genes was *MRE11*, whose product is involved in double-stranded DNA repair[86], providing evidence of an association between mutation of DNA damage repair genes and CIN.

D) Telomere dysfunction

Telomeres are regions of repetitive sequence (TTAGGG in humans) situated at the end of chromosomes and protect these regions from degradation and fusion from adjacent chromosomes. In somatic cells, a segment of telomeric DNA is lost during each replication round due to the inability of DNA polymerase to completely synthesize the chromosomal ends. When a critical short length is reached, cellular checkpoints similar to those activated by DNA damage are activated to trigger apoptosis[87]. Cells that managed to avoid apoptosis continue to divide and enter a 'cellular crisis' state of massive cell death. Cells that survive this crisis have a protective mechanism whereby telomerase, a ribonucleoprotein complex is activated to maintain telomere length. Telomere loss in human cancer has been

shown to initiate the breakage-fusion-bridge cycles leading to gene amplification and prolonged chromosomal instability[88].

There is evidence that during the adenoma:carcinoma sequence, colonocytes undergo a period of telomeric shortening before regaining mechanisms to maintain telomere length[89-91]. Increased telomerase activity is more frequently observed advanced colorectal tumours and is an independent factor of poor prognosis[92].

E) Tumour suppressor genes, oncogenes and aneuploidy

CIN tumours have been shown to possess a distinct set of mutations of tumour suppressor genes and oncogenes. It remains unclear how the accumulation of these mutations lead to CIN but proposed mechanisms for some of these genes are discussed below.

i) *APC*

The human Adenomatous Polyposis Coli (*APC*) tumour suppressor gene is located on the long arm of chromosome 5 between positions 21 and 22. It encodes a 312 KDa protein with multiple functional domains that regulates differentiation, adhesion, migration and chromosomal segregation[93]. Germline mutations in the *APC* gene are known to result in familial adenomatous polyposis (FAP), an autosomal dominant condition characterised by the development of multiple adenomas in the gastrointestinal tract, most notably in the colon and rectum. Undiagnosed and untreated, colorectal cancer invariably develops in FAP patients by the early forties. Germline mutations are generally spread uniformly between codons 200 and 1600, with two hotspots at codons 1061 and 1309[94, 95]. In contrast, over 60% of somatic

mutations are clustered in the mutation cluster region between codons 1286 and 1513[96]. Somatic mutations have been detected in as early as aberrant crypt foci (5%), precursors of colorectal polyps and its mutational frequency increases (over 80%) in the latter stages of malignant transformation[97].

A suggested role for *APC* in chromosomal stability was initiated following the discovery that during mitosis, APC protein localizes to kinetochores, centrosomes and astral microtubules[98]. APC cooperates with EB1, a plus-end microtubule-binding protein to regulate mitotic spindle dynamics and chromosome alignment[99]. Mutation in *APC* causes defect in microtubule plus-end attachments during mitosis, leading to chromosomal missegregration[100]. A much better characterized role of APC is as a key inhibitor of the canonical Wnt signaling. Wildtype APC forms a destruction complex with β-catenin (CTNNB1), glycogen syntase kinase-3β (GSK3β), casein kinase 1 (CK1) and Actin. GSK3β and CK1 phosphorylate β-catenin and consequently, phosporylated β-catenin is targeted by β-TrCP, a component of the dedicated E3 ubiquitin ligase complex for ubiquitination and subsequent proteosomal degradation[101]. Mutation in *APC* discrupts the formation of the destruction complex and leads to an accumulation of cytosolic β-catenin, which then translocate to the nucleus to drive transcription of genes implicated in tumour growth and invasion through its interaction with T-cell factor/lymphoid enhancer factor of transcription factors[102]. One example of how Wnt can stimulate proliferation is by upregulating the key transcription factor c-myc, which in turn, downregulates the cell cycle inhibitor p21, locking cells in the stem cell fate[103].

ii) *KRAS*

v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*), a member of the *RAS* family of proto-oncogene is mutated in 30-50% of colorectal cancers[104, 105]. Point mutations in codons 12 and 13 and to a lesser degree in codon 61, lock the KRAS protein in the guanosine triphosphate (GTP)-bound activated conformation by impairing the hydrolysis of GTP to GDP by GTPase. This leads to a constitutively active KRAS that regulates multiple cellular functions via downstream effectors of the RAF-MEK-MAPK and the PI3K pathways. It has been suggested that *KRAS*'s role in chromosomal instability is indirect[106] and possibly act via these two pathways.

In vitro studies have shown an association between centrosome amplification and chromosome misalignment and *RAS* activation. These changes were thought to be due to altered interactions between chromosomes and microtubules by unregulated MAPK activation[107]. The most prominent downstream target of the RAS-RAF-MEK-MAPK pathway is cyclin D1, a subunit of CDK4 or CDK6, whose activity is essential for cell cycle G1/S transition. Cyclin D1 overexpression and amplification have been demonstrated in various cancers and it is thought to stimulate cell proliferation by inactivating G1-maintaining function of the Retinoblastoma protein[108]. In the *PIK3* pathway, *RAS* binds with the catalytic subunit of type 1 phosphotidylinositol 3-kinases, leading to activation of the lipid kinase via its translocation to the plasma membrance[109]. AKT/PKB is an important downstream target of this pathway and has a strong anti-apoptotic function by inactivating BAD and Forkhead transcription factors[110].

iii) *TP53*

The *TP53* gene encodes a 393 amino acid transcription factor, p53 that acts as a tumour suppressor gene. This protein is activated by cellular stress such as DNA damage, oxidative stress and oncogenic deregulation[111]. Some of its downstream targets include the cell cycle inhibitor, p21; GADD45 and 14-3-3, which contribute to G2 arrest; and BAX, which is involved in apoptosis[112, 113]. Thus, loss of p53 causes genomic instability that most often results in aneuploidy[114]. The loss of p53 is a common event in tumourigenesis; reported in 25% of adenomas, 50% in adenomas with invasive foci and up to 75% in colorectal cancers[115] and this is most frequently due missense mutations at five hotspot codons (175, 245, 248, 273 and 282)[116].

1.5.2 Microsatellite Instability Pathway (MSI)

Microsatellites are short DNA sequences such as $(CA)_n$ and A_n (n denoting the number of bases) that are tandemly repeated in the genome. Microsatellites are intrinsically unstable and are susceptible to mutation via a mechanism called DNA slippage during replication[117]. The mismatch repair system (MMR), comprising of at least seven different proteins (MLH1, MLH2, MSH2, MSH3, MSH6, PMS1 and PMS2), recognises and repairs errors caused by nucleotide mispairing or by slippage at microsatellite sequences in newly replicated DNA, thereby maintaining replicative fidelity[74].

Unlike CIN, the mechanisms underpinning the microsatellite instability pathway (MSI) are relatively well understood and involve inactivation of the genes in the DNA mismatch repair system. MSI occurs in over 95% of colon cancers arising in

individuals with Lynch syndrome (Hereditary Non-polyposis Colorectal Cancer)[74, 118, 119], an autosomal dominant cancer syndrome characterized by an increased risk of colonic and extracolonic cancers. The mechanism of gene inactivation in this cancer syndrome is by germline mutation in one of the DNA repair genes. Interestingly, no *hMSH3* germline mutation has been identified in individuals with Lynch syndrome[120]. MSI is reported in about 15% of sporadic colorectal tumours but, unlike the familial syndrome, the mechanism of transcriptional silencing is by biallelic methylation of the *hMLH1*[121, 122] promoter.

A defective mismatch repair system induces genomic instability by allowing mutations at microsatellite sequences to accumulate during subsequent DNA replication. Mutated microsatellite loci occur in coding regions of critical genes of cell proliferation (*TGF-βR2*, *IGF2R, ACVR2*) and cell cycle/apoptosis (*BAX,* caspase-5), producing a frameshift mutation that renders their proteins inactive[123]. A repeat of eight adenines (A_8) of the mismatch repair gene, *hMsH3* undergoes a frameshift mutation[124] that leads to an increasingly defective repair system due to the cumulative loss of components on the system.

The degree of stability or the lack of, is determined using a panel of mono-, di-, tri and tetranucleotide markers (BAT-25, BAT-26, BAT-40, D5S346, D2S123, D17S250, D18S58, D18S69, D18S69 and *TGF-βRII*) selected at a National Cancer Institute consensus conference[125]. High frequency microsatellite instability (MSI-H) is defined as the presence of at least 40% of the markers screened while low frequency microsatellite instability (MSI-L) is defined as the presence of one to three of the ten

markers. Tumours are considered microsatellite stable (MSS) if no markers are detected. MSI-H tumours are more often proximal, mucinous and have a better prognosis than MSI-L and MSS tumours[126].

1.5.3 The Epigenetic Pathway

Epigenetics refers to molecular factors or processes that influence gene expression, heritable through mitosis and meiosis, without alterations in the DNA sequence[127, 128]. Regulation of gene expression by epigenetics is fundamental in normal embryonic development; the two noteworthy examples are epigenetic silencing of the X chromosome in mammalian females and in imprinted alleles to ensure monoallelic expression[129]. The epigenetic machinery can be divided into DNA methylation, histone modifications and small non-coding miRNAs. Deregulation of the epigenetic machinery in cancer is common and how this contributes to tumourigenesis is discussed below, with an emphasis on DNA methylation.

1.5.3.1 DNA methylation

Among the epigenetic mechanisms, DNA methylation is perhaps the most extensively studied and consequently the best understood. It involves the addition of a methyl group from the donor S-adenosyl methionine (SAM) to the carbon 5 of the cytosine ring of a CpG dinucleotide. This reaction is catalysed by DNA methyltransferases (DNMTs). Three DNMTs have been identified; DNMT1, DNMT3a and DNMT3b. DNMT1 is regarded as the 'maintenance methyltransferase' as it copies the pre-existing methylation patterns onto the new DNA strand during DNA replication[130]. DNMT3a and DNMT3b are 'de novo' methyltransferases and are

responsible for introducing cytosine methylation to previously unmethylated CpG sites[130]. CpG dinucleotides are not equally distributed in the human genome and instead, there are CpG-rich regions termed CpG islands that are clustered in promoters of about 50% of genes. In a normal differentiated cell, CpGs across the genome are methylated while most promoter CpG islands are kept in an unmethylated state.

Two patterns of DNA methylation have been observed in cancer: global hypomethylation across the genome and localised regions of DNA hypermethylation at CpG islands of gene promoters. Feinberg and Vogelstein first described global hypomethylation changes in colorectal cancer[131]. The proposed mechanisms by which global hypomethylation drives tumourigenesis are thought to involve the generation of chromosomal instability, reactivation of transposable elements and loss of imprinting[132]. Since then, many studies have described frequent regional hypermethylation in colorectal cancer[133, 134]. Site-specific hypermethylation usually occurs in CpG islands located near or within promoters of tumour suppressor genes and is often associated with transcriptional silencing[135-137].

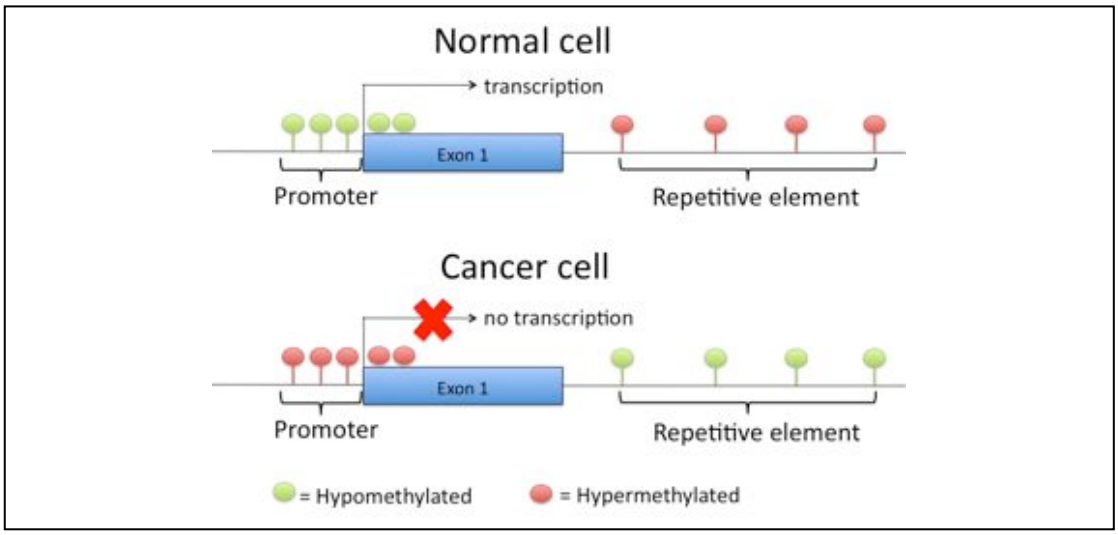


Figure 1.1. The methylation patterns in a normal and cancer cell. In a normal cell, global (represented by the repetitive elements) hypermethylation and promoter CpG island hypomethylation is observed. In a cancer cell, global hypomethylation and promoter CpG island hypermethylation, leading to transcriptional silencing is observed.

1.5.3.1.1 DNA methylation and transcriptional silencing

One of the early-described *in vitro* models of transcriptional repression by DNA methylation is the direct inhibition of transcriptional activators from binding to cognate sequences[138-140]. It has been suggested that in such a circumstance, DNA methylation acts as a 'mutation' within the factor-binding site[138], thereby diminishing the ability of transcription factors to recognised their target sites.

As already discussed, DNA methylation is catalysed by DNMTs. DNMTs can be recruited to endogenous genes by interaction with site-specific repressor proteins. For example, PML-RAR, an oncogenic transcription factor found in acute promyelocytic leukaemia, aberrantly recruits DNMT1 and DNMT3a to the promoter of *RARB*, one of PML-RAR's target genes, and causes hypermethylation and gene repression[141]. Interestingly, evidence has shown that DNA methyltransferases themselves might be involved in gene repression in addition to their catalytic

properties. DNMT-mediated gene silencing appears to rely on biochemical interactions of DNMTs with modifiers of chromatin, such as histone methyltransferases and histone deacetylases[142-144].

It has been proposed that DNA methylation can affect transcriptional elongation in addition to its inhibitory effect on transcriptional activation. Indeed, a large portion of genomic methylation is found in both introns and exons of genes. Plasmid-based reporter gene experiments indicated that methylation over the length of the gene can result in gene repression[145]. It has been proposed that intragenic methylation induced-gene repression is achieved by the formation of a chromatin structure that reduces the efficiency of the RNA polymerase II (POL II) elongation[146].

The forth and final model of methylation-associated gene silencing is by recruitment of methyl-binding domain (MBD) proteins. The human MBD proteins comprise of MECP2, MBD1, MBD2, MBD3 and MBD4. Each of the proteins, with the exception of MBD3, recognises and binds to methylated CpGs. MBD proteins can mediate gene silencing by targeting chromatin remodelling corepressor complexes such as histone deacetylases and histone methyltransferases to methylated DNA regions[147-149].

1.5.3.2 Post-translational histone modification

Histones are proteins that package and order DNA into nucleosomes, which bundle together to form chromatin. Each nucleosome comprises a tetramer of two histone 2A (H2A) and two histone 2B (H2B) molecules and is flanked by H3 and H4 dimers[127]. Remodelling of chromatin leading to changes in gene expression is linked with

specific post-translational modifications at N-terminal tails of histones. Two forms of chromatin have been described so far; euchromatin is lightly packed and tends to favour transcription whereas heterochromatin has a more condensed conformation and is associated with transcriptional silencing[150].

Examples of post-translational modifications of histones include acetylation, methylation, phosphorylation, ubiquitination, sumoylation and ADP ribosylation; the most studied of which are acetylation and methylation of specific lysine residues on histones H3 and H4. Acetylation and deacetylation are regulated by the enzymes histone acetyltransferases (HAT) and histone deacetylases (HDAC), respectively. Generally, histone acetylation is found more frequently in euchromatin whilst histone deacetylation is associated with heterochromatin[151].

In contrast, histone methylation can be associated with either transcriptional activation or repression, depending on the amino acid affected. For example, transcriptionally active genes are enriched with di- and trimethylation of histone H3 lysine 4[127, 151, 152] (H3K4me2 and H3K4me3) while trimethylation of histone H3 on lysines 9 and 27 (H3K9me3 and H3K27me3) is associated with gene repression. H3K27 trimethylation is a key activity of the Polycomb Repressor Complex 2 (PRC2), consisting of the Polycomb group proteins EZH1, EZH2, SUZ12 and EED. It is thought that the H3K4me3 mark results in the initiation of gene repression and then serves as a docking site for the recruitment of Polycomb Repressor Complex 1 (PRC1) to maintain the repressive state[153]. Interestingly, a study by Bracken *et. al.* showed a considerable overlap between genes targeted by the Polycomb group proteins and

those silenced in cancer by promoter hypermethylation[154], suggesting that PRC2 marks specific tumour suppressor genes for silencing by DNA methylation during tumourigenesis.

Figure 1.2. The two different states of chromatin are dependent on the post-transcriptional modifications of the histone tails, including acetylation, methylation and phosphorylation. Taken from http://www.sigmaaldrich.com/life-science/learning-center/biofiles/biofiles-5-6/cancer-research.html.

1.5.3.3 Small, noncoding microRNAs (miRNAs)

The most recently discovered component of the epigenetic machinery, miRNAs are 20 to 22 nucleotides long, post transcriptional regulators that bind to the complementary 3' untranslated regions of their target mRNAs, leading to their degradation and subsequent inhibition of gene expression[155]. They are first synthesized as long, noncoding RNAs and then cleaved by the enzyme DROSHA in the nucleus, transported into the cytoplasm as short hairpin RNAs and further cleaved by the enzyme DICER into the final form of double-stranded miRNAs[155]. Since their discovery in 1993, over 1000 human miRNAs have been identified and they may target over 60% of the mammalian genes[156]. The study of miRNAs has become of great interest especially since the discovery of their roles in a host of

cellular processes such as cell cycle progression, apoptosis, differentiation and tumourigenesis[157]. miRNAs also have an important role in the chromatin structure control by interacting with the Polycomb group of proteins. MiR-101 and miR-26a have been found to target the 3'-UTR of EZH2, one of the PRC2 subunits, through which they exert their function[158]. By regulating EZH2 expression at the transcriptional and post-transcriptional levels, miRNAs can affect H3K27 methylation and consequently contribute to gene silencing in tumourigenesis[159].

1.5.3.4 Epigenetic alterations influence genetic changes in tumourigenesis

DNA hypermethylation at CpG islands near promoters of tumour suppressor genes with associated gene silencing is the most well recognised epigenetic change in tumours. It has been suggested that the loss of function of classic tumour suppressor genes by promoter hypermethylation is more common than by mutation[160]. There is an increasing list of candidate tumour suppressor genes silenced by promoter hypermethylation in certain cancers, which do not appear to be frequently mutated. Examples include RASSF1A[161, 162], which encodes a protein of unknown function that can bind to the RAS oncogene; cyclin-dependent kinase inhibitor 2B (CDKN2B), which encodes p15, a cell-cycle regulator[163]; and O6-methylguanine-DNA methyl-transferase (MGMT)[164], which encodes a critical DNA-repair gene. There is convincing evidence that epigenetic events occur at early stages of tumour development and precede genetic changes[160, 165]. Consistent with this notion, Feinberg *et. al.* proposes a new cancer development model whereby the first step involves an epigenetic disruption of progenitor cells in a given organ or system, which leads to a polyclonal precursor population of neoplasia-ready cells. This is

followed by the second step, which is an initiating mutation within the subpopulation of epigenetically disrupted progenitor cells. The third and final step involves genetic and epigenetic plasticity[165]. Various methylation-related mechanisms that trigger genetic changes and contribute to tumourigenesis have been described and are discussed below.

1.5.3.4.1 Increased mutability of methylcytosines

5-methylcytosine (5-mC) is intrinsically unstable and has an enhanced susceptibility to mutational events compared to cytosine[166]. 5-mC is prone to spontaneous hydrolytic deamination into thymine under physiological conditions and transition mutation (mostly C to T) is thought to arise as a result of this process[167]. Transition mutations (C to T and G to A) from methylated CpG sites are an important mechanism underlying different cancer types as these mutations occur at hotspots located in numerous tumour suppressor genes[168]. For example, five of the six mutation hotspots in *TP53* are found at CpG sites containing 5-mCs[168, 169]. These sites show C to T and G to A transitions. In addition, point mutations in the retinoblastoma gene, *RB1*, are also C to T transitions at CpG sites, resulting in premature termination of translation[170]. These findings suggest that 5-mCs at CpG sites is a major source of endogenous mutation.

1.5.3.4.2 Silencing of DNA repair genes

The mismatch repair gene, h*MLH1* recognises and repairs errors caused by nucleotide mispairing or by slippage at microsatellite sequences in newly replicated DNA, thereby maintaining replicative fidelity[74]. A defect in the mismatch repair

system induces genomic instability by allowing mutations at microsatellite sequences to accumulate during subsequent DNA replication. Microsatellite instability is reported in about 15% of sporadic colorectal tumours and the mechanism of transcriptional silencing is by biallelic hypermethylation of the *hMLH1*[121, 122] promoter.

MGMT is another DNA repair gene that is silenced by promoter hypermethylation in colorectal cancer[164]. The MGMT protein removes the carcinogenic O^6-methylguanine adducts from DNA, which result in G to A transition mutations if left unrepaired. Tumours with silenced MGMT alleles are predisposed to transition mutation in key genes, such as *TP53*[171] and *KRAS*[172]. Similar to that observed in *hMLH1,* promoter hypermethylation of *MGMT* precedes genetic events by occurring in premalignant adenomas that do not have gene mutations[171, 172].

1.5.3.4.3 Enhanced carcinogen binding

Products of sunlight exposure and tobacco smoke have an enhanced binding preference to 5-mCs compared to unmethylated cytosines[173, 174]. The major UV mutagen, cyclobutane pyrimidine dimer, has a strong association with the methylated 5'TCG sites, conferring an over 10-fold increase in pyrimidine dimer formation after exposure to sunlight[175]. 5-mCs are also the preferred target sites for benzo[*a*]pyrene, a carcinogenic agent found in cigarette smoke, leading to G to T transversion mutations in lung cancer[176].

1.5.3.5 The CpG island methylator phenotype

In 1999, Toyota *et. al.* pioneered the concept 'CpG island methylator phenotype' (CIMP) to describe a group of colonic tumours with distinct clinicopathological and molecular features[177]. They analysed 26 novel loci, termed MINT (methylated-in-tumour), all of which contained CpG islands. The methylation pattern of these newly cloned loci fell into two groups; age-related (Type A) and cancer-specific (Type C). Type A loci were found methylated in 'normal' ageing epithelial cells as well as in colonic tumours while Type C loci were exclusively methylated in colonic tumours. The CIMP+ definition was arbitrarily used to encompass tumours exhibiting methylation of 3 or more of the Type C loci (MINT1, 2, 12, 17, 25, 27 and 31). In the same study, CIMP+ tumours were found more frequently in the proximal colon and had excellent concordance with methylation of *p16* (*CDKN2A*), *THBS1* and *hMLH1*. Their subsequent work found that CIMP+ tumours were associated with *KRAS* mutation and wild-type *TP53*, independent of MSI status[178].

Large population-based studies have subsequently evaluated the CIMP concept in colorectal cancer and confirm the existence of this phenotype[179-181]. Despite this, there has been much debate about this concept. It remains unknown whether CIMP tumours represent a biologically distinct group of colorectal cancers or are merely an arbitrary selected group from the continuum of tumours showing varying degrees of methylation at particular loci. The use of different types and numbers of candidate genes to classify CIMP, leading to varying clinicopathological and molecular features have only fuelled the debate further[179-183]. A consensus on the exact groups of candidate genes that defines CIMP and a standardised methodology to measure

methylation are clearly required to elucidate the biological significance of the CIMP concept.

1.5.3.6 Prognostic relevance of DNA methylation in rectal cancer

A biomarker that is strongly associated with prognosis is likely to be useful in the treatment algorithm of rectal cancer. Most studies, however, have investigated DNA methylation in cohorts consisting of either colon cancers only or colon and rectal cancers together. It is unknown whether the results for colon cancers can be extrapolated into rectal cancers.

Small cohort studies that have evaluated the prognostic value of DNA methylation in colorectal cancer have yielded conflicting results. Wang *et. al.* reported that in a cohort of 85 tumours, methylation of *CDH13* and *FLBN3* was found more frequently in advanced stage tumours and consequently was associated with poor overall survival[184]. This pattern of methylation was also observed in another study, where, higher levels of *CHFR* methylation were associated with a lower recurrence-free survival in locally advanced colon cancer[185]. In contrast, Dallo *et. al.* found that methylation of polycomb group target genes (*SFRP1*, *MYOD1*, *HIC1*, and *SLIT2*) in colorectal cancer confers a favourable prognosis[186]. Large population-based DNA methylation studies have failed to resolve these conflicting results. Studies by Barault *et. al.*[179] and Ward *et. al.*[187] suggested that CIMP had an adverse effect on survival in microsatellite stable tumours while Ogino *et. al.*[180] showed that CIMP was independently associated with low cancer-specific mortality. The conclusion that can be drawn from these studies is that the prognostic impact is determined by

individual genes that undergo aberrant DNA methylation and not by the process itself.

To date, there are only two groups who have investigated the role of DNA methylation specifically in rectal cancer. de Maat *et. al.* described a methylation profile consisting of hypermethylation of MINT3 and hypomethylation of MINT1, 12 and 17 that can predict for distant recurrence, cancer-specific survival, and overall survival in node-negative rectal cancers[188]. Interestingly, in a separate study, the authors refined this panel to MINT3 hypermethylation and MINT17 hypomethylation and found that this signature was associated with a lower probability of local recurrence[189]. The authors, however, did not provide any explanations as to why in one study, the described methylation profile was an adverse prognostic feature but in another study, it was associated with a good outcome. In a recent study, Jo *et. al.* evaluated the prognostic role of CIMP in 150 advanced rectal cancers and found that CIMP+ tumours were associated with a worse three and five year survival. The results, however, should be interpreted with caution given that only 15 patients in their study were classified as CIMP+ tumours[190].

1.5.3.7 Summary

In the absence of sensitive radiological and other clinical parameters, the use of molecular markers to predict the risk of mesorectal nodal metastasis in rectal cancer is appealing. Emerging evidence suggests that DNA methylation has a prognostic role in colorectal carcinogenesis but has not been evaluated in rectal cancer in great detail. Identification of specific genes that undergo aberrant DNA methylation may

unravel a unique and clinically relevant methylation signature that is associated with disease progression in rectal cancer. This can be used to guide decision-making at the multidisciplinary meetings.

1.6 DNA methylation platforms

DNA methylation is an attractive clinical biomarker because it is preserved during routine DNA extraction, can be quantitatively measured in archived tissues and has negligible temporal variations[191]. Various DNA methylation platforms exist and for discussion purposes, they can be divided into those that interrogate methylation at single CpG loci or short sequences (candidate gene approach) and those that interrogate methylation at a global level (genome-wide approach). Some of the commonly used platforms are summarised below.

1.6.1 Platforms that interrogate methylation at single CpG loci or short sequences

1.6.1.1 Methylation–specific PCR (MSP)/Quantitative MSP (Q-MSP)

The principle of MSP relies on the chemical conversion of DNA by bisulphite treatment. Bisulphite treatment protocols preferentially convert unmethylated cytosine residues into uracil with preservation of methyl-cytosine. Subsequent PCR reaction amplifies uracil as thymidine while methyl-cytosine remains unaffected. MSP takes advantage of the sequence differences arising from this process.

For MSP analysis, primer pairs for both methylated and unmethylated DNA are designed to determine the methylation status of the loci of interest. One PCR is performed with methylation specific primers and another PCR with primers specific

for unmethylated DNA. Methylated and unmethylated primer sets differ only in the CG position of the bisulfite converted primary sequence: methylated primer sets contain CG dinucleotides and unmethylated primers contain TG at these positions[192]. Results can be visualised on agarose gel electrophoresis. The biggest drawback of this method is that it only provides qualitative data and risks producing false positive results. In addition, it cannot distinguish low from high levels of methylation. Quantitative measurement of methylation is important because low levels of methylation (below the threshold of transcriptional silencing) may not be biologically important[193, 194].

A quantitative MSP (Q-MSP) assay based on Taqman technology, called MethyLight has been developed[195]. Bisulfite-converted genomic DNA is amplified using PCR primers flanking an oligonucleotide probe with a 5′ fluorescent reporter dye (6FAM) and a 3′ quencher dye (TAMRA). The exonuclease activity of the DNA polymerase cleaves the probe and releases the reporter, whose fluorescence can be detected by a sequence detection system. After exceeding a predetermined threshold, the PCR amplification results in a light signal proportional to the amount of PCR product generated[195].

1.6.1.2 Combined Bisulphite Restriction Analysis (COBRA)

Similar to MSP, COBRA relies on the chemical conversion of DNA by bisulphite treatment. This platform utilises a restriction endonuclease whose recognition sequence includes the CpG of interest only (e.g. BstUI cleaving CGCG). Bisulphite converted DNA is amplified using primers flanking the recognition site of the

endonuclease. Incubation of the methylated DNA (CGCG sequence preserved during PCR) with the endonuclease results in cleavage of the PCR product while the unmethylated DNA will not be cleaved due to the recognition sequence being converted to TGTG during PCR. Relative quantification of methylation can be estimated based on the signal intensity of the separated fragments on gel electrophoresis. This platform has the advantage of being simple, inexpensive and provides a quantitative readout but is limited to loci containing sequences that are recognised by endonucleases.

1.6.1.3 Bisulphite genomic sequencing

Bisulphite sequencing follows the principle of Sanger sequencing. The main difference is that bisulphite sequencing analyses bisulphite-modified DNA rather than genomic DNA. It is capable of interrogating methylation at single nucleotide resolution in up to 500 bp DNA fragments. Methylation status can be determined either by direct sequencing (detection of average methylation status) or by sub-cloning sequencing (detection of methylation of a single nucleotide of an allele)[196]. To improve its sensitivity, a large number of clones is usually required and this process is time-consuming and labour-intensive. Newer technique, such as bisulphite pyrosequencing can simultaneously quantify methylation of multiple sequences and has superseded bisulphite sequencing as a high-throughput alternative.

1.6.1.4 Bisulphite pyrosequencing

Bisulphite pyrosequencing was developed as an alternative to bisulphite sequencing[197] and quantifies methylation in bisulphite-modified DNA. This technique

is based on the detection and quantification of pyrophosphate (PPi), released from incorporated nucleotides by DNA polymerase during elongation. PPi molecules undergo a series of enzymatic processes to enable a luciferase-dependent final reaction emitting a light signal, the intensity of which is recorded by a camera. At a CpG locus, both dCTP and dTTP are injected consecutively. The signal ratio obtained from these injections gives a readout that correlates with the amount of methylation.

The major advantage of the pyrosequencing method compared to methylation-specific PCR is that the data are actual sequences rather than fluorescence data from PCR-based amplification. Unlike MSP, pyrosequencing does not rely on primer/probe annealing for specificity. In addition, pyrosequencing can detect partially methylated sequences that are flanked by the priming sites, whereas MSP can only detect sequences that are completely complimentary to the primer and probe sequences[198]. Another strength of pyrosequencing is its ability to simultaneously quantify multiple sequential CpG sites in a single run. The main disadvantage of this platform is the low sequencing distance it can read (100 to 150 bp)[199]. In addition, it requires dedicated and expensive equipment for analysis.

1.6.1.5 High-resolution melting analysis (HRM)

High-resolution melting analysis (HRM) is a real-time PCR-based technique using bisulphite modified DNA[200]. HRM relies on the precise monitoring of the change of fluorescence as DNA melts. The temperature range over which melting occurs and the shape of the curve is a function of the length, sequence and GC content of the

product[201]. Products derived from DNA template containing methylated cytosines will have a higher GC content, and thus, a higher melting and touch down temperature than those of DNA with unmethylated cytosines. Methylation can be estimated by comparing the melting profiles of unknown samples with the profiles of fully methylated and unmethylated references amplified after bisulphite modification. Unlike pyrosequencing, this platform assesses methylation in the amplified region as a whole rather than at specific CpG sites.

1.6.2 Genome-wide approaches

Recent advances in technology have allowed us to interrogate global methylation at an affordable and high-throughput level. This approach will unravel many novel loci that may serve as useful markers for early detection of disease, monitoring of therapy and disease progression. Currently, several microarray and next generation sequencing technologies are available for commercial use. Some of these promising approaches are discussed below.

1.6.2.1 Immunoprecipitation-based DNA microarray

This technology utilises a protein, most commonly an antibody raised against methyl-cytosine to isolate methylated DNA. The antibody-methyl-cytosine complex is immunoprecipitated using magnetic beads conjugated to anti-mouse IgG. The methylated fraction is compared with the input DNA, which is not enriched for DNA methylation. The enriched DNA is labelled with a fluorescent dye that is different from the input DNA and the labelled DNA samples are then co-hybridised onto a microarray spotted with probes corresponding to particular regions of interest such

as CpG island libraries or promoter arrays. Downstream bioinformatics analysis is performed to estimate DNA methylation level using the relative signal intensity of the labelled DNA. The accuracy of technique is dependent on the efficient binding of the protein to the methylated DNA and the CpG content of the loci of interest. The readout provides the average methylation status of a region and not individual CpGs.

1.6.2.2 DNA BeadChip microarray

Similar to bisulphite sequencing and pyrosequencing, this method utilizes bisulphite treated DNA and quantifies methylation levels at specific loci within the genome. The latest microarray from illumina™ using this BeadChip technology interrogates over 450 000 methylation sites in a single reaction. This technology utilises two different fluorescent-labelled bead types for each locus of interest; one bead type is designed to hybridise to the locus if it is methylated and the other will hybridize to the same locus if it is unmethylated. The fluorescence intensity ratios between the two bead types for each locus are calculated and correspond to the methylation percentage. The main advantage of the BeadChip microarray over other technologies for analysing the methylome is that the data is quantitative at single nucleotide resolution.

1.6.2.3 High-throughput sequencing

The latest and most promising technology for genome-wide methylation analysis is high-throughput sequencing. The aim of high-throughput sequencing is to generate vast amount of sequence information faster and at a lower cost than Sanger sequencing, without the need for cloning. To put this into context, the Roche/454

Life Science genome sequencer, the first commercially available high-throughput sequencing system, can produce the same sequence data at 10% of the cost and in 0.9% of the time that a single ABI3730 Sanger sequencer can. The other commercially available systems are the Solexa/Illumina Genome Analyser and the SOLiD sequencing system.

These high-throughput sequencing platforms follow three important steps: DNA sample preparation, immobilisation and sequencing. Preparation of a DNA sample for sequencing involves the addition of defined sequences, known as 'adapters', to the terminals of fragmented DNA and this is referred to as the 'sequencing library'. Adapters anchor the DNA fragments of the sequencing library to a solid surface and define the site in which the sequencing reactions begin. Amplification of the sequencing library is performed to form spatially distinct and detectable sequencing features. Sequencing is then performed using either DNA polymerase synthesis for fluorescent nucleotides or the ligation of fluorescent oligonucleotides.

High-throughput sequencing can also be used as an alternative to microarrays for analysing DNA methylation. Enriched and control DNA can be sequenced directly rather than being labeled and hybridised to the oligonucleotide array. This has the advantage of providing a quantitative measure of methylation rather than a qualitative readout by some microarrays[202]. In addition, this method is not affected by the efficiency of probe-DNA hybridisation.

1.7 Aims

Aberrant DNA methylation is an important mechanism of tumourigenesis and the identification of specific genes that undergo this process may be useful in the risk stratification of local recurrence in early rectal cancer.

The aims of this study are:

- To determine site-specific DNA methylation changes that occur in rectal cancer, adjacent rectal tissue and normal tissue.
- To correlate methylation changes with histopathological features of disease progression.
- To develop a unique prognostic methylation signature for rectal cancer.
- To determine the impact of tissue enrichment techniques on methylation yield.
- To evaluate the relationship between DNA methylation and protein expression.
- To correlate protein expression changes with histopathological features of disease progression.
- To determine genome-wide methylation changes that occur during premalignant transformation to established cancer and subsequent progression to advanced disease.

Chapter 2

Materials and Methods

2.1 List of Reagents and Suppliers

SALSA MS-MLPA ME001 Tumour Suppressor-1 kit, **MRC-Holland, the Netherlands**.

Proteinase K, Tris-HCl, Phenol:Chloroform:Isoamyl Alcohol 25:24:1, ethanol, chloroform, sodium acetate, **Sigma-Aldrich, Dorset, UK**

Sodium dodecyl sulphate (SDS), **Applied Biosystems, Warrington, UK.**

Ethylenediaminetetraacetic acid (EDTA), **Fisher Scientific, Leicestershire, UK.**

DNeasy Blood and Tissue Kit , Epitect Plus FFPE Bisulphite Kit , Epitect Bisulphite Kit, PyroMark CpG assays, RNase A, PyroMark Binding buffer, PyroMark Annealing buffer, PyroMark Denaturation buffer, PyroMark Wash Buffer, PyroMark Gold Q96 Reagents, PyroMark CpG assays *ESR1, UNC5C, POU3F3, TIAM1, GLI3, TTC16, RAPGEF3, PTCH1, TMEM106A* and *PGBD2*, **QIAGEN, Crawley, UK.**

Unmethylated and methylated genomic DNA, **Millipore, Watford, UK**

Oligonucleotides for *CDH13, CHFR, CXCL12, DAPK1, MINT3, MINT17, CDH1, SNAP25, SOX7* were synthesized by **biomers.net**.

IMMOLASE DNA polymerase, **Bioline, London, UK.**

Streptavidin Sepharose High Performance beads, **GE Healthcare Life Sciences, Buckinghamshire, UK**

APC (c-20) rabbit polyclonal antibody, **Insight Biotechnology, Wembley, UK**

DAPK1 rabbit polyclonal antibody, **Source Bioscience Lifesciences, Nottingham, UK**

CXCL12/SDF-1 mouse monoclonal antibody, **R&D Systems Europe, Abingdon, UK**

RARB (c-19) rabbit polyclonal antibody, **Insight Biotechnology, Wembley, UK**

UNC5C mouse monoclonal antibody, **R&D systems Europe, Abingdon, UK**

E-cadherin mouse monoclonal antibody **BD Biosciences, Oxford, UK**

ESR-α rabbit monoclonal antibody, **Thermoscientific, Northumberland, UK**

CDH13/T-cadherin rabbit polyclonal antibody, **Insight Biotechnology, Wembley, UK**

CHFR mouse monoclonal antibody, **Abcam, Cambridge, UK**

2.2 Clinical samples

Colorectal cancer tissues were obtained from a tissue bank consisting of over 200 flash-frozen colorectal cancer specimens from patients treated at the University Hospitals of Birmingham NHS Trust between 2001 and 2010. Samples were taken from the core of the resected tumour by the operating surgeon immediately after surgery and flash-frozen in liquid nitrogen. Rectal surgery consisted of total mesorectal excision (TME). Histologically confirmed tumour-free matched adjacent tissues (taken 5 cm or more from tumour edge) were also available for paired analysis. Histologically confirmed normal rectal tissues were obtained from patients who underwent normal colonoscopy or flexible sigmoidoscopy for rectal bleeding, change in bowel habit or incontinence. Formalin-fixed paraffin embedded (FFPE) rectal adenocarcinomas were obtained from the Department of Histopathology, University Hospitals of Birmingham NHS Trust. Patients were treated in the same institution between 2005 and 2010. Patients who underwent neoadjuvant chemoradiotherapy, have a family history or previous history of colorectal cancer or inflammatory bowel disease were excluded.

2.3 DNA extraction of fresh frozen tissues for MS-MLPA and bisulphite pyrosequencing analyses

Approximately 30-60 mg of frozen tissue was added to 0.5ml of DNA digestion buffer (50mM Tris-HCl pH8, 100mM EDTA pH8.0, 200mM NaCl, 1% SDS) containing

0.5mg/ml proteinase K. Samples were incubated overnight at 37°C with gentle shaking, to aid complete disruption of the tissues. 0.7 ml of phenol/chloroform/isoamyl alcohol mixture (Sigma-Aldrich, Dorset, UK) was added to the samples which were then centrifuged at 13000 rpm for 5 mins. 0.5 ml of the upper phase was recovered and 0.7 ml of chloroform was added before samples were mixed and centrifuged at 13000 rpm for a further 5 mins. The upper phase was recovered and 0.7 ml of chloroform was again added. Samples were mixed and centrifuged at 13000 rpm for another 5 mins. A tenth volume and twice the volume (of the upper phase) of 3M sodium acetate and 100% ethanol respectively were added to the upper phase. The samples were centrifuged at 13000 rpm for 5 mins and the supernatant was carefully removed and discarded. 1 ml of 70% cold ethanol was added and the samples were centrifuged at 13000 rpm for 5 mins after which the supernatant was carefully removed and discarded. Samples were air dried, resuspended in 200 µl of TE buffer and added with 0.5 µl of RNase A (100mg/ml). Samples were stored at -20°C.

2.4 DNA extraction of fresh frozen tissues for Nimblegen Methylation microarray

Extraction was performed using the DNeasy® Blood and Tissue Kit (QIAGEN, Crawley, UK) according to manufacturer's instructions. Prior to extraction, 25ml of 100% ethanol was added to 19ml of concentrated Buffer AW1 and 30 ml of 100% ethanol was added to 13ml of concentrated Buffer AW2. 180 µl of Buffer ATL and 20 µl of proteinase K were added to approximately 10-25mg of frozen tissue. Samples were vortexed and then incubated overnight at 56°C. 4 µl of RNaseA (100mg/ml) was added and the samples were incubated at 37°C for 30 mins. 200 µl of Buffer AL and

200 µl of 100% ethanol were then added to the samples which were then vortexed for 15s. Samples were transferred into the DNeasy Mini spin columns and centrifuged at 8000 rpm for 1 min. The flow-through was discarded. 500 µl of Buffer AW1 was added and the samples were centrifuged at 8000 rpm for 1min. The flow-through was discarded. 500 µl of Buffer AW2 was added and samples were centrifuged at 14000 rpm for 3 mins. The flow-through was once gain discarded. 200 µl of Buffer AE was added directly onto the DNeasy membrane. Samples were incubated at room temperature for 1 min and then centrifuged at 8000 rpm for 1 min. Samples were stored at -20°C.

2.5 DNA quantification and quality analysis

Spectrophotometry was used to quantify the amount and analyse the quality of DNA extracted. Spectrophotometer readings were taken at wavelengths 260nm and 280nm. Nucleic acids have a maximum absorption at 260nm whereas contaminates such as proteins and single stranded DNA and/or RNA absorb maximally at 280nm. The ratio of the Optical Density (OD) at 260nm/280nm provides an indication of the purity of the sample. Clean DNA has an OD_{260}/OD_{280} between 1.8 – 2.0. An OD_{260} of 1 corresponds to approximately 50 µg/ml of double-stranded DNA. DNA was quantified using the following formula:

DNA concentration (µg/ml) = OD_{260} x dilution factor x 50 µg/ml

Spectrophotometry was only performed on DNA extracted from fresh frozen tissues.

2.6 Bisulphite treatment of fresh frozen DNA

Incubation of target DNA with sodium bisulphate results in conversion of unmethylated cytosine residues into uracil, leaving methylated cytosines unchanged. Bisulphite treatment of fresh frozen DNA was performed using the Epitect® Bisulphite Kit (QIAGEN,Crawley,UK) according to manufacturer's instructions. Prior to bisulphite treatment, 30 ml of 100% ethanol was added to concentrated Buffer BW and 27 ml of 100% was added to concentrated Buffer BD. Approximately 150 ng - 500 ng of DNA was added with RNase-free water to make up a total volume of 20 µl. 85 µl of Bisulphite Mix (dissolved in 800 µl of RNAse-free water) and 35 µl of DNA protect buffer were added to the samples, following which the samples were treated according to the following conditions:

Step	Time (mins)	Temperature (°C)
Denaturation	5	95
Incubation	25	60
Denaturation	5	95
Incubation	85	60
Denaturation	5	95
Incubation	175	60
Hold	Overnight	20

Table 2.1. Bisulphite treatment steps and conditions.

560 µl of Buffer BL containing 10 µg/ml of carrier RNA was added to the samples and were then vortexed and centrifuged for 20s. The mixtures were transferred to the EpiTect spin columns and were centrifuged at 13000 rpm for 1 min. The flow-through was discarded and 500 µl of Buffer BW was added to the samples which were then centrifuged at 13000 rpm for 1 min. The flow-through was discarded and 500 µl of Buffer BD was added to the samples. Samples were incubated at room temperature for 15 min and then centrifuged at 13000 rpm for 1 min. The flow-through was discarded and 500 µl

of Buffer BW was added to the samples which were then centrifuged at 13000 rpm for 1 min. This step was repeated. Samples were centrifuged at 13000 rpm for 1 min to remove any residual liquid and were in 1.5 ml microcentrifuged tubes. 20 µl of Buffer EB was added directly onto the Epitect membrance and samples were centrifuged at 12000 rpm for 1 min. This step was repeated to obtain 40 µl of bisulphite treated DNA. Samples were immediately stored at -20°C.

2.7 Macrodissection and laser-capture microdissection of formalin-fixed paraffin embedded tissues

Paraffin-embedded blocks were sectioned at 5µm thick and mounted on glass slides. At least 10 sections were made for each tissue block. A reference slide for each sample was stained with haematoxylin and eosin and was reviewed by Dr Philippe Taniere, Consultant Histopathologist. Areas containing over 80% of tumour tissues were marked with an indelible pen and used to guide the sampling of representative regions from subsequent, non-H&E stained sections. For macrodissection, approximately 1 x 1 cm of adjacent sections were scraped off the glass slides using fresh scalpel blades and were placed in 500 µl tubes for DNA extraction.

Areas to be microdissected were marked on reference slides. These corresponded to the same areas that were macrodissected, allowing for direct comparison. Slides were deparaffinised and stained with haematoxylin and eosin. Cover slips were not placed. Areas marked on slides were laser-captured microdissected using the PALM CombiSystem (Carl Zeiss Microscopy GmbH, Cambridge) and placed in 500 µl tubes for DNA extraction.

2.8 DNA extraction and bisulphite treatment of formalin-fixed paraffin embedded (FFPE) tissues

Bisulphite treatment of formalin-fixed paraffin embedded tissues was performed using the Epitect® Plus FFPE Bisulphite Kit (QIAGEN,Crawley,UK) according to manufacturer's instructions. This kit has the advantage of bisulphite treating DNA without prior DNA isolation. Prior to bisulphite treatment, 30 ml of 100% ethanol was added to concentrated Buffer BW and 27 ml of 100% was added to concentrated Buffer BD. 310 µl of 1 µg/µl lyophilized carrier RNA was added to buffer BL. 150 µl of Deparaffinisation solution was added to dissected FFPE tissues and then vortexed until all paraffin is dissolved. A master mix consisting of 20 µl distilled water, 15 µl Lysis Buffer FTB and 5 µl proteinase K was added to each mixture. The mixtures were then incubated at 56 °C for 30 mins followed by 95 °C for 1 hour. Deparaffinisation solution was removed prior to bisulphite treatment.

2.9 Methylation-Specific Multiplex Ligation Probe Analysis (MS-MLPA®)

MS-MLPA (MRC-Holland, Amsterdam, The Netherlands), a high-throughput, semi-quantitative, methylation specific enzyme-based PCR assay, was performed according to manufacturer's instructions. The SALSA MS-MLPA ME001 Tumour Suppressor-1 kit, containing probes that detect methylation status of 24 promoters of tumour suppressor genes found frequently methylated in tumours [203, 204] was used. Each probe corresponded to a single gene apart from *MLH1* and *RASSF1* which contained 2 probes each. To calculate methylation percentage, each sample required two MLPA reactions; one contained the methylation sensitive enzyme, *HhaI* and the other without. Briefly, 100-150 ng of samples were diluted to 5 µl with TE and heated at 98

°C for 40 minutes in a thermocycler. Following the addition of MLPA probes, samples were incubated at 60 °C for 16 hours to allow hybridisation. For ligation and digestion reactions, each sample was split into 2 vials. Ligase-65 mix (containing Ligase-65 buffer, Ligase-65 enzyme and water) was added to the first vial whilst Ligase-Digestion mix (Ligase-65 buffer, Ligase-65 enzyme, *HhaI* enzyme (Promega, Southampton, UK) and water) was added to the second vial. Samples were then incubated at 49 °C for 30 minutes. The ligase enzyme was inactivated by heating at 98 °C for 5 minutes. PCR buffer, dNTPs and Taq polymerase were added to the samples. The PCR reaction was performed in a 72 °C preheated thermocycler under these conditions: 35 cycles at 95 °C for 30 s, 60 °C for 30 s and 72 °C for 60 s. Final incubation was at 72 °C for 20 minutes. Amplification products were analysed on an ABI-3730 DNA Analyzer (Applied Biosystems, Warrington, UK). Internal validation was performed using unmethylated and methylated genomic DNA (Millipore, Watford, UK). A methylation percentage for each gene was obtained using the following calculation:

$$\text{Methylation\%} = \frac{(\text{Peak height of a given probe/Mean height of control probes})_{\text{with HhaI}}}{(\text{Peak height of a given probe/Mean height of control probes})_{\text{without HhaI}}} \times 100$$

2.10 Bisulphite pyrosequencing

2.10.1 Primer design and PCR conditions

The *APC* and *GSTP1* primer sequences and PCR conditions were obtained from the PyroMark Assay Database (http://techsupport.pyrosequencing.com). The *RARB* assay has been designed and optimised by Yuk-Ting Ma at the University of Birmingham. The LINE-1 primer sequences and PCR conditions were obtained from a published study[205]. *ESR1, UNC5C, POU3F3, TIAM1, GLI3, TTC16, RAPGEF3, PTCH1, TMEM106A*

and *PGBD2* assays were purchased as PyroMark CpG assays from QIAGEN, Crawley, UK. Primer design for *CDH13, CHFR, CXCL12, DAPK1, MINT3, MINT17, CDH1, SNAP25* and *SOX7* was performed using an online software (http://www.urogene.org/methprimer/index1.html). 3 µl of each PCR products was analysed on a 1% agarose gel before pyrosequencing.

The PCR conditions were as follows:

Component	CpG Assays	In-house design assays	MINT3, SOX7
10x buffer 50 mmol/l	2.5 ul	5 ul	2.5 ul
MgCl2	1 ul	2.5 ul	1 ul
dNTPs	2.5 mmol each	2.5 mmol each	2.5 mmol each
Primers	0.2 ul	5 pmol of each	5 pmol of each
Immolase	0.1 ul	0.2 ul	0.1 ul
Sterile water	17.9 ul	38.3 ul	18.9 ul
DNA	1 ul	1ul	1ul
PCR conditions			
95 °C	15 min	15 min	15 min
3-step cycling			
95 °C	20s	20s	20s
Tm °C	20s	30s	30s
72 °C	20s	30s	30s
No of cycles	45	45	45
95 °C	5 min	5 min	5 min

Table 2.2. PCR conditions of assays prior to bisulphite pyrosequencing.

2.10.2 Preparation for methylation analysis

Immoblisation and preparation of PCR products for pyrosequencing was performed on the PyroMark Q96 Vacuum WorkStation (QIAGEN), according to manufacturer's instructions. A mastermix consisting of 3 µl streptavidin beads (GE Healthcare Life Sciences, Buckinghamshire, UK) and 37 µl of PyroMark Binding buffer (QIAGEN) was added to each PCR products. The mixture was placed on a shaker for 5 mins at 13000

rpm. Another mastermix consisting of 1.5 µl of sequencing primer (10 pmol/µl) and 38.5 µl of PyroMark Annealing buffer (QIAGEN) was added into the wells of the sequencing plate. Each troughs on the workstation was filled with the following reagents:

a) 180 ml 70% ethanol

b) 120 ml PyroMark Denaturation buffer (QIAGEN)

c) 180 ml PyroMark Wash Buffer (1:10 dilution)

d) 180 ml water

e) 180 ml water in parking position

Washing of the filter probes on the hedgehog was performed by lowering it into water in the parking position for 20s on vacuum. The streptavidin-bound PCR products were captured by lowering the hedgehog into the 96-well plate containing the PCR products. The hedgehog was then placed in the trough containing 70% ethanol for 10s, followed by the trough with PyroMark Denaturation buffer (QIAGEN) for 10s and finally into the trough with PyroMark Wash Buffer for 10s. The vacuum was turned off before lowering the hedgehog into the sequencing plate to release the enriched streptavidin-bound PCR products. The sequencing plate was then placed on a preheated heat block at 80 °C for 2 mins and allowed to cool. Appropriate amount of PyroMark Gold Q96 Reagents (QIAGEN), dependent on the assays analysed, were added into the PyroMark Q96 cartridge (QIAGEN). Analysis was performed using the PyroMark CpG software.

Genes		Primers	Tm (°C)	No. of CpG sites
RARB	Forward	GAGTTGTTTGAGGATTGGGATGT	54	10
	Reverse	(B)CCAAAAAAATCCCAAATTCTC		
	Sequencing	AGTTGTTTGAGGATTGG		
CDH13	Forward	TTGGAAAAGTGGAATTAGTTGGTAT	56	9
	Reverse	(B)ACCAAAACCAATAACTTTACAAAAC		
	Sequencing	AAAGAAGTAAATGGGATGTTATTTT		
CHFR	Forward	TAGAATTTTTGGGGTTTTTAATT	56	10
	Reverse	(B)ACCATCTTTAATCCTAACCAAAC		
	Sequencing	TAGAATTTTTGGGGTTTTTAATT		
CXCL12	Forward	GGGATTAATTTGTTTGTTTTTTATTG	58	7
	Reverse	(B)ACCTTTAACCTTCTCAAACTCC		
	Sequencing	TGTTTTTTATTGGTTTTTATTTAGTTT		
DAPK1	Forward	TTTTGGAGGTGGGAAAGTTG	55	10
	Reverse	(B)AAAAACACCCTTTATTAAAACTAAAC		
	Sequencing	GGGTATGTGTGTAGAGAAAGGGGA		
MINT3	Forward	TGATGGTGTATGTGATTTTGTGTT	60	10
	Reverse	(B)ACCCCACCCCTCACAAAC		
	Sequencing	TGATGGTGTATGTGATTTTGTGTT		
MINT17	Forward	AGGGGTTAGGTTGAGGTTGTT	58	4
	Reverse	(B)TCTACCTCTTCCCAAATTCCA		
	Sequencing	AGGGGTTAGGTTGAGGTTGTT		
CDH1	Forward	TTTTAGTAATTTTAGGTTAGAGGGTTAT	56	10
	Reverse	(B)TAACTACAACCAAATAAACCCC		
	Sequencing	TTTTAGTAATTTTAGGTTAGAGGGTTAT		
CASP8	Forward	TTAATAGGAAGTGAGGTTATGGAGG	58	4
	Reverse	(B)TAAACCAAACAACACCCAAAAATAT		
	Sequencing	TTAATAGGAAGTGAGGTTATGGAGG		
TIMP3	Forward	TGGTTTGGGTTAGAGATATTTAGTG	56	5
	Reverse	(B)CCCCCTCAAACCAATAACAA		
	Sequencing	TGGTTTGGGTTAGAGATATTTAGTG		
SNAP25	Forward	GATTAGTATAAGTAGTTTTTGTTTTTT	56	6
	Reverse	(B)TACTAAATATCTTAAAAATACCCCAC		

	Sequencing	GATTAGTATAAGTAGTTTTTGTTTTTT		
SOX7	Forward	TTTATAGTTGTTAAGGTGTATAAATG	52	6
	Reverse	(B)TAATTTCTTACTTTTCCTCTACAAC		
	Sequencing	TTTATAGTTGTTAAGGTGTATAAATG		

Table 2.3. Primer sequences, PCR annealing temperatures, Tm of pyrosequencing assays and the number of CpG sites examined. (B)=biotinylated.

2.11 Immunohistochemistry

Immunohistochemistry was performed using the Envision™ FLEX+ detection system (Dako, Ely, UK) at the Department of Histopathology, University Hospital Birmingham NHS Foundation Trust. Paraffin tissues were first incubated at 37^{o}C for 30 mins. Deparaffinisation, rehydration and heat-induced epitope retrieval (HIER) was performed on paraffin-embedded tissue sections according to the 3-in-1 specimen preparation procedure protocol using the PT Link system. Briefly, slides were immersed in 1.5 litre of diluted 1:50 FLEX Target Retrieval Solution, preheated to 65^{o}C. Slides were then incubated for 15 mins at 97^{o}C and allowed to cool to 65^{o}C. For staining with the APC antibody, slides were first deparaffinised in xylene for 5 mins and washed with 70% ethanol followed by water. Epitope retrieval was performed in a pressure cooker for 2 mins. Slides were transferred onto the Dako Autostainer instrument and washed with diluted 1:20 FLEX Wash Buffer. Primary antibodies were diluted accordingly with the FLEX Antibody Diluent. The staining steps and incubation times were preprogrammed according to the FLEX_200 (200 µL dispense volume) protocol and were as follows:

a) FLEX Peroxidase-Blocking Reagent for 5 mins

b) Diluted primary antibody for 60 mins

c) Flex/HRP for 30 mins

d) Diluted 1:20 Flex Wash Buffer for 5 mins

e) Flex DAB+ chromogen in Flex Substrate Buffer for 10 mins

f) Diluted 1:20 Flex Wash Buffer for 5 mins

Slides were counterstained in a bath of Haematoxylin for 30 s and in Scott's water for 15s. Dehydration was performed in ethanol series (70% ethanol x1, 100% ethanol x2 for 1 min each). Further deparaffinisation was performed in xylene for 1 min and slides were mounted with non-aqueous permanent mounting medium. Slides were analysed using light microscopy. Each run was performed with known positive control for that particular antibody. Scoring of staining intensity was classified as absent (nil stained), weak (<10% cells were stained), moderate (10-50% cells stained) and strong (>50% cells stained). Slides were scored by two independent observers (Kai Leong and Dr. Rahul Hejmadi). Discrepancies in scoring were discussed and a consensus result was derived.

2.12 Methylation profiling using MeDIP

2.12.1 Preparation of samples

The MeDIP assay was performed by the manufacturer as previously described[206]. Briefly, genomic DNA was sheared into 300–1200-bp fragments by sonication and approximately 5 ug were incubated with 5ug of monoclonal mouse anti-5-methyl-cytidine mAb overnight at 4°C. Antibody-bound DNA was precipitated with 50 mL of Dynabeads (M-280 sheep antibodies to mouse IgG, Dynal Biotech) at 4°C for 2 hours on a rotating wheel. Bound DNA and sonicated input DNA were differentially labeled with fluorescent dyes (Cy3 and Cy5 respectively) and hybridised to promoter tiling array. Scanning of the array was performed using the Nimblegen MS 200 Microarray Scanner.

2.12.2 Array design, data processing and probe annotation of the array for detection of genome-wide promoter methylation

The NimbleGen Human DNA Methylation 385K RefSeq Promoter Array was designed based on the HG18 genome release. The array contains 385 019 probes with lengths of between 50-75 bp with a median spacing of 100 bp along the upstream promoter regions of well-characterized RefSeq genes. Signal intensity data was extracted from the scanned images of each array using NimbleScan, Nimblegen's data extraction software. Each feature on the array has a corresponding scaled $\log_2$-ratio. This is the ratio of the signal intensities for the enriched and input samples that were co-hybridized to the array. The $\log_2$-ratio was computed and scaled to centre the ratio data around zero. Scaling was performed by subtracting the bi-weight mean for the $\log_2$-ratio values for all features on the array from each $\log_2$-ratio value. Significant

differences between log2-ratio values between groups were identified using Linear Models for Microarray Data (LIMMA) with cutoff thresholds of Benjamini and Hockberg adjusted *P* value < 0.05 and fold change > 1.5. Data analysis was performed with the help of Dr Wenbin Wei.

Results

Chapter 3

Screening for novel methylation markers associated with disease progression in rectal cancer

3. 1 Introduction

Promoter hypermethylation is associated with tumour invasion and metastasis. The objective of this chapter was to determine whether the methylation status of a group of tumour suppressor genes (TSG), commonly methylated in cancer, was associated with histopathological markers of disease progression in radically excised rectal cancer. Methylation status was correlated with the presence of microscopic mesorectal lymph node metastasis, as confirmed by a pathologist during routine histopathological analysis. Occult loco-regional metastasis is a major clinical concern when considering organ preserving treatment strategies for early-stage rectal cancer. A high throughput, quantitative assay that can simultaneously screen for methylation at multiple sites would provide an excellent exploratory study. The Methylation-Specific Multiplex Ligation Probe Amplification was selected for this purpose (MS-MLPA®, MRC-Holland, Amsterdam, The Netherlands).

3.1.1 Methylation-Specific Multiplex Ligation Probe Amplification (MS-MLPA®)

MS-MLPA is a semi-quantitative methylation-specific enzyme-based PCR assay that can analyse up to 40 sequences in a single reaction. Each MS-MLPA probes targets and hybridises to a specific sequence containing the restriction site of the methylation-sensitive endonuclease, *HhaI*. Probes consist of two separate oligonucleotides: a short synthetic oligonucleotide and a long M13-derived oligonucleotide. These oligonucleotides contain a pair of universal primers. In addition, the M13-derived oligonucleotide also contains a stuffer sequence that varies in length. It is these probes that are amplified rather than the target sequences, a feature unique to MLPA. Probes that hybridise to unmethylated target

sequences are digested by the *HhaI* endonuclease and not subsequently amplified during PCR. In contrast, probes hybridised to methylated sequences are protected against *HhaI* digestion and subsequent amplification yields signals that are detected by capillary electrophoresis (Figure 3.1).

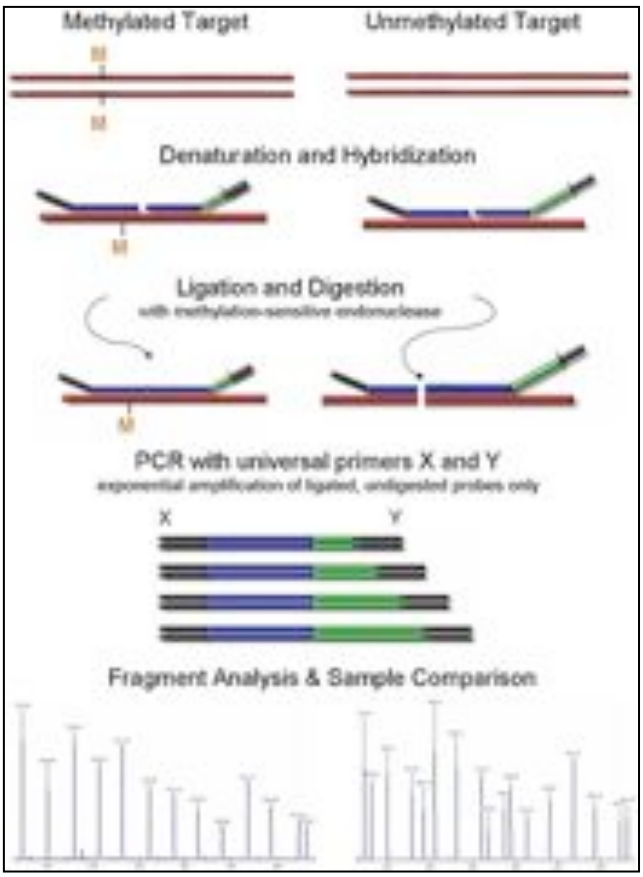


Figure 3.1. Diagrammatic representation of the MS-MLPA reaction.

An advantage of this assay is that the probes can be customised, allowing the development of a single assay containing multiple discriminatory methylation sites. Initial analysis used an 'off the shelf' tumour suppressor gene kit. The SALSA MS-MLPA ME001 Tumour Suppressor-1 kit contains probes that detect methylation status of 24 promoters of tumour suppressor genes that are frequently methylated in solid tumours [203, 204]. Each gene corresponds to a single probe apart from *MLH1* and

RASSF1, which have two probes. The list of tumour suppressor genes analysed is shown in Table 3.1. Results generated by MS-MLPA were subsequently validated using bisulphite pyrosequencing (QIAGEN, West Sussex, UK).

Genes	**Name**
TIMP3	Tissue inhibitor of metalloproteinases-3
APC	Adenomatosis polyposis coli
CDKN2A	Cyclin-dependent kinase inhibitor 2A; p16
MLH1	mutL homologue 1
ATM	Ataxia telangiectasia mutated
RARB	Retinoic acid receptor,beta
CDKN2B	Cyclin-dependent kinase inhibitor 2B; p15
HIC1	Hypermethylated in cancer 1
CHFR	Checkpoint with forkhead and ring finger domains
CASP8	Caspase 8
CDKN1B	Cyclin-dependent kinase inhibitor 1B; p27
PTEN	Phosphatase and tensin homologue deleted on chromosome 10
BRCA1	Breast cancer 1
BRCA 2	Breast cancer 2
CD44	CD44 molecule
RASSF1A	RAS association domain family 1A
DAPK1	Death-associated protein kinase 1
VHL	von Hippel – Lindau
ESR1	Oestrogen receptor 1
TP73	Tumour protein p73
FHIT	Fragile histidine triad gene
IGSF4	Immunoglobulin superfamily, member 4
CDH13	Cadherin 13
GSTP1	Glutathione S-transferase pi

Table 3.1 List of tumour suppressor genes analysed using the SALSA MS-MLPA ME001 Tumour Suppresor-1 kit.

3.2 Results

3.2.1 Demographics and clinicopathological characteristics of samples used

51 fresh frozen sporadic rectal adenocarcinomas, resected using the principles of total mesorectal excision (TME) were identified from our database. Subjects had no family or personal history of colorectal cancer or inflammatory bowel disease. No subjects received neoadjuvant therapy. Among the 51 tumours, 35 had histologically confirmed tumour-free adjacent tissues (taken 5 cm or more from tumour edge) for paired analysis. 19 'normal' controls were taken from patients following colonoscopy or flexible sigmoidoscopy for a range of non-specific symptoms. The clinical and pathological data of patients are summarised in Table 3.2.

Demographics/Clinicopathological characteristics	Normal n=19	Adjacent tissue n=35	Cancer n=51
Age (years)			
Median	63	71	71
Range	36 - 83	33 - 89	33 - 89
Sex			
Male	13	25	33
Female	6	10	18
Staging			
Duke's A			13
Duke's B			14
Duke's C			15
Duke's D			9
Tumour size (mm)			
Median			40
Range			13 - 110
Lymphovascular invasion			
No			34
Yes			17
Degree of Differentiation			
Well + Moderate			46
Poor			5

Table 3.2. Demographics and clinicopathological characteristics of samples used.

3.2.2 Internal validation of MS-MLPA using methylated and unmethylated controls.

The ability of MS-MLPA to discriminate fully methylated from unmethylated samples was examined using known methylated and unmethylated controls (Millipore). The methylation percentages of the two controls for the genes examined are shown in Table 3.3. In the case of the unmethylated control, methylation levels were very low (<10%) in all genes apart from *ESR1* (11.3%) and *CDH13* (12.3%). Conversely, for the methylated control, methylation levels for all genes were high (>90%) apart from two

genes; *APC* (88.8%) and *MLH1* (87.4% and 76.7%). As mentioned previously, the MS-MLPA kit used in this study contained two probes for the *MLH1* gene. The overall results provided adequate confidence that MS-MLPA could be used to detect and quantify methylation in clinical samples.

Gene	Methylation percentage (%)	
	Unmethylated control	Methylated control
TIMP	2.41	100.00
APC	5.19	88.83
CDKN2A	2.72	100.00
MLH1	2.99	87.35
ATM	0.51	100.00
RARB	5.79	100.00
CDKN2B	7.86	100.00
HIC1	1.50	100.00
CHFR	0.00	100.00
BRCA1	1.72	97.06
CASP8	2.96	95.78
CDKN1B	1.23	100.00
PTEN	4.82	100.00
BRCA2	0.98	98.64
CD44	1.88	100.00
RASSF1	1.00	97.55
DAPK1	1.99	90.69
VHL	1.75	100.00
ESR1	11.29	100.00
RASSF1	1.23	95.78
TP73	2.51	100.00
FHIT	1.62	100.00
IGSF4	1.52	100.00
CDH13	12.25	100.00
GSTP1	4.10	100.00
MLH1	1.39	76.74

Table 3.3. Methylation percentages for unmethylated and methylated controls of tumour suppressor genes using the SALSA MS-MLPA ME001 Tumour Suppressor Kit-1 kit.

3.2.3. Quantitative comparison of methylation changes in rectal cancer and matched adjacent tissue

Methylation levels in cancer were analysed by MS-MLPA and compared to those in adjacent tissue. Figure 3.2 shows examples of capillary electrophoresis patterns for cancer and adjacent tissue analysed with MS-MLPA. Of the 24 tumour suppressor genes, five were significantly methylated in cancer; *ESR1* ($P<0.001$), *CDH13* ($P<0.001$), *CHFR* ($P<0.001$), *APC* ($P<0.001$) and *RARB* ($P=0.023$).

3.2.4 Quantitative comparison of methylation changes in rectal cancer and normal tissue from individuals without neoplasia

When methylation levels in cancer were compared to normal controls from individuals without colorectal neoplasia the significance of these five genes was preserved. The methylation levels of individual cancers, adjacent tissues and normal controls for these five genes are shown in Figure 3.2.

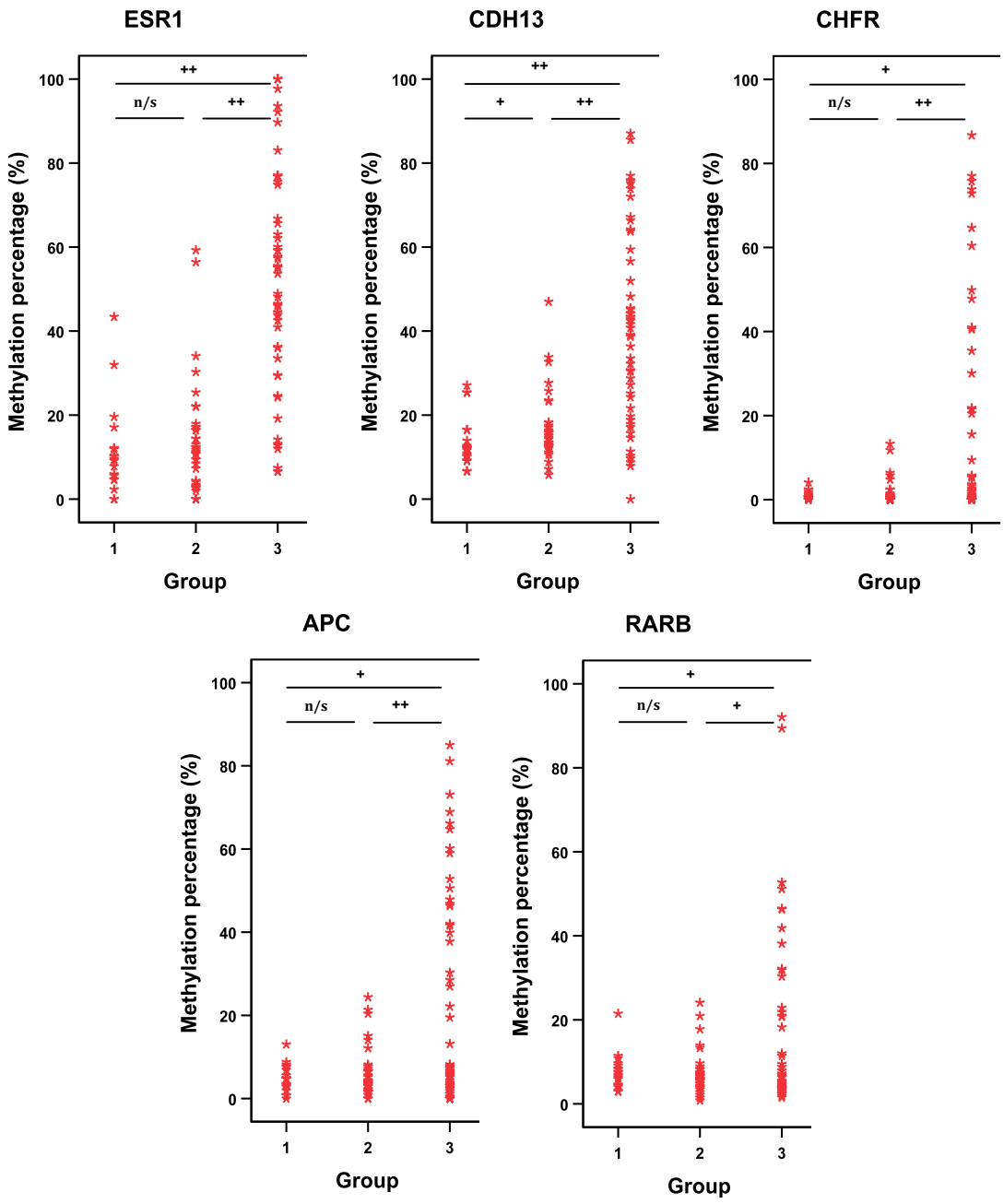


Figure 3.2. Methylation percentage of ESR1, CDH13, CHFR, APC and RARB in normal controls (Group 1, n=19), matched adjacent tissue (Group 2, n=35) and rectal cancer (Group 3, n=51). $^{++}$ *P<0.001,* $^{+}$*P<0.05, n/s = not significant. P values between matched adjacent tissues and cancers were derived using Wilcoxon Signed Rank Test (2-tailed). All other P values were derived using Mann-Whitney U Test (2-tailed).*

3.2.5 Analysis of methylation as a categorical variable in rectal cancer

A sample was considered hypermethylated when the methylation percentage exceeded a value derived from normal controls (Mean control value + 2x standard

deviation). Using this definition all 24 genes were hypermethylated in at least one cancer. The proportion of samples in each group (rectal cancer, matched adjacent tissue, normal control) exhibiting categorical hypermethylation across all five genes (derived from the initial quantitative analysis) was determined; *ESR1* (80% vs. 14% vs. 5%, *P*<0.001), *CDH13* (71% vs. 6% vs. 5%, *P*<0.001), *CHFR* (45% vs. 14% vs. 5%, *P*<0.001), *APC* (43% vs. 6% vs. 5%, *P*<0.001) and *RARB* (31% vs. 9% vs. 5%, P=0.009). The results are summarised in Tables 3.4 to 3.6.

The total number of hypermethylated genes for each tissue sample was also determined (Table 3.7). Rectal cancers possessed more hypermethylated genes compared to matched adjacent tissue and normal control {median 3 (range 0-24) vs. 1 (0-15) vs. 0 (0-8), *P*<0.001). Only 4% (2 of 51) of cancers did not exhibit any gene hypermethylation, in marked contrast to adjacent tissue (49%, 17 of 35) and normal control (68%, 13 of 19).

	ESR1	CDH13	CHFR	APC	RARB	CDKN2A
Cancer (n=51)						
Median methylation (%)	53.7	40.8	1.9	7.7	6.4	2.2
Range (%)	6.5 - 100	0 - 87.0	0 - 86.7	0 - 84.9	1.5 - 92.0	0 - 37.1
No. of samples hypermethylated	41	36	23	22	16	2
Median methylation (%) of hypermethylated samples	57.7	45.1	35.4	47.3	35.1	24.5
Range (%)	29.3 - 100	28.7 - 87.0	3.2 - 86.7	19.4 - 84.9	18.2 - 92.0	11.8 - 37.1
Adjacent tissue (n=35)						
Median methylation (%)	10.7	12.7	0.7	4.7	4.9	2.8
Range (%)	0 - 59.2	6.5 - 46.9	0 - 13.2	0 - 24.3	0 - 21.4	0 - 14.1
No. of samples hypermethylated	5	2	5	2	3	4
Median methylation (%) of hypermethylated samples	43.4	40.4	6.4	22.9	20.9	13.6
Range (%)	30.2 - 59.2	33.7 - 46.9	4.0 - 13.2	21.3 - 24.3	17.7 - 21.4	10.7 - 14.1
Normal (n=19)						
Median methylation (%)	11.2	14.6	0	4.1	4.8	1.7
Range (%)	0 - 34.0	0 - 32.6	0 - 4.7	0 - 20.4	0 - 24.1	0 - 12.0
No. of samples hypermethylated	1	1	1	1	1	2
Median methylation (%) of hypermethylated samples	34.0	32.7	4.8	20.4	24.1	11.3
Range (%)	-	-	-	-	-	10.5-12.0
P value (Cancer vs. Adjacent)*	**<0.001**	**<0.001**	**<0.001**	**<0.001**	**0.023**	0.599
P value (Cancer vs. Normal)^	**<0.001**	**<0.001**	**0.001**	**0.011**	**0.028**	0.552
P value (Adjacent vs. Normal) ^	0.379	**0.025**	0.262	0.978	0.431	**0.049**

*Table 3.4. Quantitative and categorical methylation values for ESR1, CDH13, CHFR, APC and RARB in rectal cancer, matched adjacent tissue and normal controls (without neoplasia). P values in bold are significant. *Wilcoxon Signed Rank Test (2-tailed). ^Mann-Whitney U test (2-tailed).*

	GSTP1	TIMP3	TP73	FHIT	ATM	CDKN2B	BRCA1	CDKN1B	PTEN	BRCA2
Cancer (n=51)										
Median methylation (%)	6.2	2.6	2.9	1.3	0.9	5.0	1.1	1.5	3.1	1.2
Range (%)	0 - 56.8	0 - 89.3	0 - 92.0	0 - 21.7	0 - 50.7	0 - 52.5	0 - 50.0	0 - 53.3	0 - 100	0 - 31.8
No. of samples hypermethylated	5	8	6	7	5	6	6	5	5	6
Median methylation (%) of hypermethylated samples	23.4	28.6	15.0	8.9	19.5	36.9	12.8	16.6	20.5	14.8
Range (%)	20.8 - 56.8	20.4 - 89.3	11.2 - 92.0	7.2 - 21.7	12.3 - 50.7	29.6 - 52.5	9.1 - 50.0	13.8 - 53.3	14.3 - 100	9.5 - 31.8
Adjacent tissue (n=35)										
Median methylation (%)	5.5	4.3	2.6	2.5	1.3	5.1	1.3	1.6	3.0	1.1
Range (%)	0 - 32.1	0 - 30.9	0 - 14.9	0 - 23.2	0 - 27.8	0 - 86.9	0 - 15.7	0 - 36.3	0 - 23.6	0 - 35.1
No. of samples hypermethylated	1	3	3	8	4	6	3	5	3	5
Median methylation (%) of hypermethylated samples	32.2	28.4	13.6	11.2	17.3	44.5	14.0	23.9	22.9	17.5
Range (%)	-	26.9 - 30.9	11.6 - 14.9	6.5 - 23.2	10.6 - 27.8	27.9 - 86.9	13.4 - 15.7	15.1 - 36.3	14.1 - 23.6	8.9 - 35.1
Normal (n=19)										
Median methylation (%)	6.0	4.8	3.6	1.7	1.1	6.2	1.4	1.9	3.1	1.2
Range (%)	0 - 2.7	1.8 - 24.0	0 - 12.3	0 - 7.09	0 - 15.7	0 - 26.5	0 - 11.2	0 - 17.7	0 - 18.8	0 - 11.8
No. of samples hypermethylated	1	2	1	1	1	1	1	1	1	1
Median methylation (%) of hypermethylated samples	25.7	22.7	12.4	7.1	15.8	26.6	11.3	17.7	18.9	11.9
Range (%)	-	-	-	-	-	-	-	-	-	-
P value (Cancer vs. Adjacent)*	0.106	0.659	0.216	0.561	0.561	0.561	0.193	0.789	0.678	0.294
P value (Cancer vs. Normal)^	0.601	0.058	0.102	0.941	0.302	0.345	0.399	0.593	0.932	0.736
P value (Adjacent vs. Normal) ^	0.409	0.606	0.074	0.121	0.906	0.400	0.710	0.690	0.332	0.221

*Table 3.5. Quantitative and categorical methylation values for GSTP1, TIMP3, TP73, FHIT, ATM, CDKN2B, BRCA1, CDKN1B, PTEN and BRCA2 in rectal cancer, matched adjacent tissue and normal controls (without neoplasia). *Wilcoxon Signed Rank Test (2-tailed). ^Mann-Whitney U test (2-tailed).*

	HIC1	CD44	DAPK1	CASP8	VHL	IGSF4	MLH1(A)	MLH1(B)	RASSF1(A)	RASSF1(B)
Cancer (n=51)										
Median methylation (%)	1.5	1.7	4.2	1.6	0	3.4	3.1	0.0	0.9	1.6
Range (%)	0 - 41.1	0 - 86.6	0 - 52.7	0 - 100	0 - 25.3	0 - 100	0 - 100	0 - 55.3	0 - 66.0	0 - 100
No. of samples hypermethylated	6	4	8	6	2	9	1	5	9	6
Median methylation (%) of hypermethylated samples	17.6	27.3	31.6	20.8	20.3	39.3	100.0	7.5	26.0	47.9
Range (%)	13.7 - 41.1	21.2 - 86.6	17.1 - 52.7	11.1 - 100	15.2 - 25.3	22.1 - 100	-	5.0 - 55.3	13.8 - 66.0	18.4 - 100
Adjacent tissue (n=35)										
Median methylation (%)	1.5	2.4	5.0	1.9	0.7	5.1	3.3	0	1.3	2.3
Range (%)	0 - 24.2	0 - 25.2	0 - 35.3	0 - 20.0	0 - 21.6	0 - 44.9	0 - 17.2	0 - 14.7	0 - 26.5	0 - 23.4
No. of samples hypermethylated	4	2	5	3	3	2	0	5	4	3
Median methylation (%) of hypermethylated samples	18.5	24.7	23.1	18.8	13.7	43.8	0.0	4.7	14.2	23.1
Range (%)	12.3 - 24.2	24.1 - 25.2	19.0 - 35.3	17.4 - 20.0	11.4 - 21.6	42.7 - 44.9	-	4.2 - 14.7	10.1 - 26.5	18.1 - 23.4
Normal (n=19)										
Median methylation (%)	1.5	2.1	5.1	2.1	0	5.4	2.9	0	0.6	1.6
Range (%)	0 - 19.6	0 - 28.1	0 - 18.5	0 - 14.0	0 - 13.5	0 - 31.2	1.0 - 30.9	0 - 5.4	0 - 15.4	0 - 22.5
No. of samples hypermethylated	1	2	1	1	1	1	1	1	1	1
Median methylation (%) of hypermethylated samples	19.6	24.1	18.0	14.0	11.9	31.2	30.9	5.4	15.4	22.5
Range (%)	-	20.0-28.1	-	-	-	-	-	-	-	-
P value (Cancer vs. Adjacent)*	0.663	0.262	0.087	0.244	0.438	0.167	0.172	0.600	0.144	0.369
P value (Cancer vs. Normal)^	0.726	0.931	0.319	0.345	0.184	0.053	0.838	0.154	0.263	0.629
P value (Adjacent vs. Normal) ^	0.380	0.191	0.574	0.351	0.078	0.289	0.474	0.582	0.540	0.618

*Table 3.6. Quantitative and categorical methylation values for HIC1, CD44, DAPK1, CASP8, VHL, IGSF4, MLH1 and RASSF1 in rectal cancer, matched adjacent tissue and normal controls (without neoplasia). MLH1 and RASsF1 each contained two sites. *Wilcoxon Signed Rank Test (2-tailed). ^Mann-Whitney U test (2-tailed).*

Tissue Types	No. of hypermethylated loci						*P* values
	0	1 - 2	3 - 4	5 - 10	11 - 15	≥16	
Cancer (n=51)	2	15	18	11	0	5	
Adjacent tissue (n=35)	17	14	1	1	2	0	<0.001*
Normal control (n=19)	13	4	0	2	0	0	0.251^

*Table 3.7. Distribution of the number of hypermethylated genes in cancers, adjacent tissues and normal controls. P values were derived using Fisher's Exact Test (2-tailed). *Cancer vs. adjacent tissue and cancer vs. normal control. ^Adjacent tissue vs. normal control.*

3.2.6 Association of methylation with histopathological indices of disease progression

The five genes previously selected (*ESR1, CDH13, CHFR, APC* and *RARB)* with significantly diferential methylation percentages in cancer compared to controls were collectively analysed according to Dukes' stage to determine whether there were significant variations with tumour progression (Figure 3.3A). A trend of decreasing methylation values with advancing Dukes' stage was observed. The methylation difference between Dukes' A (TNM Stage I) and Dukes' D (TNM Stage IV) cancers was statistically significant (*P*=0.013).

Individual sites were analysed to determine any association between methylation percentage and tumour progression, using the presence of mesorectal lymph node involvement and distant metastasis as markers of tumour progression. *GSTP1* exhibited increased methylation percentages in node-negative compared to node-positive cancers [median 8.1% (IQR 4.1-11.3%) vs. 5.8% (IQR 0–7.1%) *P*=0.006]. *RARB* methylation percentages were higher in non-metastatic cancers compared to those

with distant metastasis [median 7.5% (IQR 4.3-28.5%) vs. 3.9% (IQR 3.8–4.8%), *P*=0.008] [Figure 3.3(B)]. Categorical analysis, using the cutoff defined in section 3.2.5, found that *RARB* was hypermethylated in 16/42 (38%) non-metastatic cancers and 0/9 cancers with distant metastasis (*P*=0.043).

The correlation between the total number of hypermethylated loci in each cancer tissue and tumour progression was also examined. Our results showed that cancers with six or more hypermethylated loci were more frequently node-negative (*P*=0.015). A methylation signature was constructed to identify cancers that are associated with localised and hence favourable disease phenotype. Concomitant hypermethylation of two or more of *CASP8*, *APC, RARB, TIMP3* and *GSTP1* was seen more frequently in node-negative (p=0.002) and non-metastatic cancers (*P*=0.044) (Table 3.8). In contrast, hypermethylation of none or one of these genes was associated with advanced T stage, pT3-4 (*P*=0.047). The calculated sensitivity and specificity of this methylation signature in predicting nodal disease in rectal cancer was 46.4% (95% CI +/- 13.7%) and 91.3% (95%CI +/- 7.7%). Figure 3.4 illustrates the methylation levels of these five genes in normal controls, adjacent tissues and different stages of cancers.

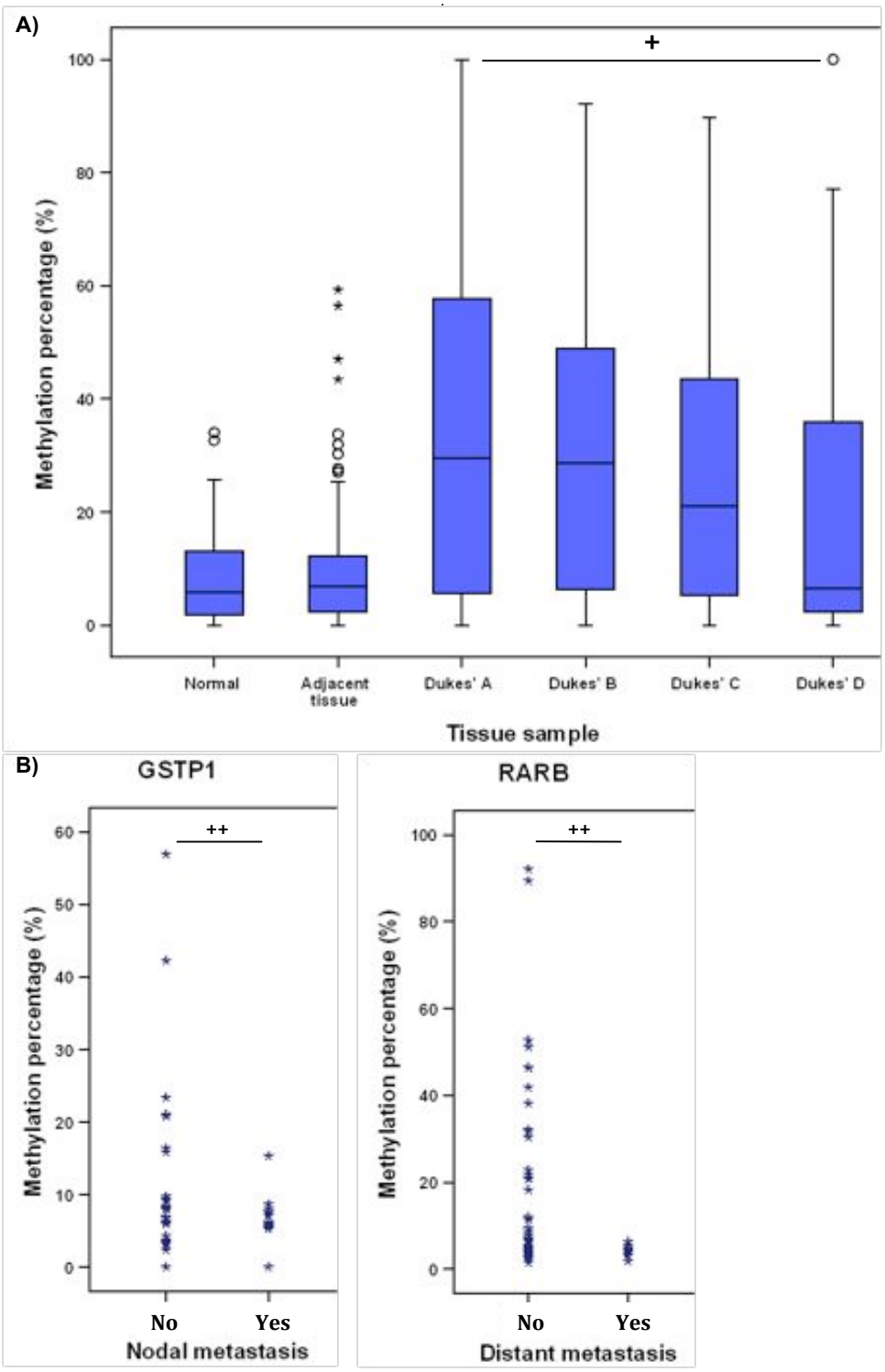


*Figure 3.3. A) Boxplot demonstrating the combined median methylation levels of ESR1, CDH13, CHFR, APC, RARB in normal controls, adjacent tissues and cancers according to stage. ○ = normal outlier, * = extreme outlier, $^{+}P<0.05$. B) Methylation levels of GSTP1 and RARB in cancers according to nodal and distant metastasis status respectively. $^{++}P<0.01$. All P values were derived using Mann-Whitney U Test (2-tailed).*

Clinicopathological	Methylation of *CASP8*, *APC, RARB*, *TIMP3*, *GSTP1*		
Features	≤ 1 gene	≥ 2 genes	*P* values*
Age (years)			
< median	18	6	0.554
≥ median	18	9	
Sex			
Male	22	11	0.527
Female	14	4	
Depth of invasion			
pT1 and pT2	8	8	**0.047**
pT3 and pT4	28	7	
Nodal metastasis			
No	15	13	**0.002**
Yes	21	2	
Distant metastasis			
No	27	15	**0.044**
Yes	9	0	
Lymphovascular invasion			
No	22	12	0.527
Yes	13	4	
Tumour size			
< Median	16	7	1.00
>Median	20	8	
Degree of differentiation			
Well + Moderate	32	14	1.00
Poor	4	1	
Total	36	15	

*Table 3.8. Correlation of clinicopathological features with methylation of APC, RARB, TIMP8, CASP8 and GSTP1 in cancers. P values in bold are significant. *Fisher's Exact Test (2-tailed).*

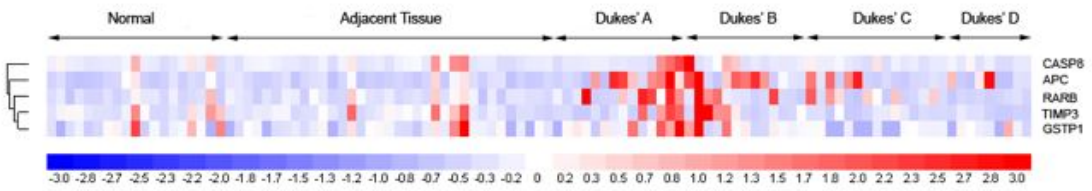


Figure 3.4. Heat map illustrating the methylation levels of CASP8, APC, RARB, TIMP3 and GSTP1 in normal controls, adjacent tissues and cancers according to the Dukes' stage. This heat map was generated using dChip (http://www.dchip.org). On the left is the dendrogram showing gene clustering using 1-correlation as distance metric and centroid linkage method. Methylation levels of each gene across samples were linearly scaled to have a mean of 0 and a standard deviation of 1. Hypermethylation of two or more of CASP8, APC, RARB, TIMP3 and GSTP1 was seen more frequently in Dukes' A and B than in Dukes' C and 'D' cancers.

3.2.7 Methylation levels in adjacent tissues and normal controls

Differences in methylation values between macroscopically normal colonic mucosa located adjacent to cancer and tumour-free bowel mucosa have been reported. It is postulated that such differences may be due to a field change, that incorporates bowel beyond the tumour itself[207]. Methylation analysis between macroscopically normal tissue taken at least 5 cm from the tumour edge and bowel mucosa from tumour-free individuals was performed to determine whether a field effect was present in this cohort. The median age for both groups was over 60, although adjacent tissue had a higher median age. The difference, however, did not reach statistical significance (*P*=0.055). The male to female ratio was similar in both groups (Table 3.2).

Overall, subtle differences in methylation levels were seen between adjacent tissue and normal control with an almost equal number of sites differentially hypermethylated in each group (Tables 3.4 – 3.6). 10 genes showed higher methylation values in tissue adjacent to the tumour, with *CDKN2A* reaching statistical

significance (*P*=0.049). 13 genes displayed higher methylation values in normal tissue. Of these *CDH13* was statistically significant (*P*=0.025). Adjacent tissue had a higher number of hypermethylated loci but the difference was not statistically significant (*P*=0.133) (Table 3.7). Two adjacent tissues had 11 and 15 hypermethylated loci. Their matched cancers, both node-negative tumours, showed greater numbers of hypermethylated loci; 17 and 24 respectively.

3.3 Discussion

This study identified increased site-specific methylation in rectal cancer compared to both adjacent tissue and normal rectal mucosa. Of the 24 tumour suppressor genes, only 5 (*ESR1, CDH13, CHFR, APC* and *RARB*) were significantly hypermethylated in rectal cancer. Interestingly, the prevalence of these methylation events decreased significantly as tumour stage increased. When the methylation percentage of individual genes was correlated with histological features, only *GSTP1* and *RARB* showed any statistical significance. This low 'hit-rate' (2/24 genes) is likely to be due to the constraints of using an 'off the shelf' methylation platform which contains a predetermined set of tumour suppressor genes. Nevertheless, the purpose of the work in this chapter is largely exploratory and the results obtained have provided some valuable insight into methylation patterns in rectal cancer.

In light of such findings, a new analysis incorporating multiple methylated genes was performed to identify a panel of potentially useful prognostic markers. This study found that concomitant hypermethylation of ≥2 of *CASP8, APC, RARB, TIMP3* and *GSTP1* was strongly associated with histopathologically localised disease (TxN0M0). We accept that the selected endpoint of lymph node disease is artificial, since it misses patients with low volume disease and a small proportion of patients with distant metastasis but it remains an important independent predictor for locoregional recurrence[56]. Furthermore, any information regarding lymph node status is particularly important in local excision of rectal cancer, as this surgical technique allows the depth of invasion to be assessed, but provides no data regarding the presence of local micrometastasis.

Only one other group has investigated methylation exclusively in rectal cancer, examining the association between prognosis and methylation of seven genomic loci previously found methylated in colorectal cancer[177]. Using samples from the Dutch TME trial, they found that a subgroup of patients with relative hypomethylation of MINT17 had a significant risk of distant recurrence, shorter cancer-specific survival and shorter overall survival[188]. Their more recent work showed that hypermethylation of MINT17 was associated with a lower local recurrence rate following TME surgery[189]. The present study has provided further evidence that locus-specific hypermethylation could be used to predict disease progression in rectal cancer.

Methylation levels in tumours were compared with both adjacent tissue and mucosa from normal individuals as, although the adjacent tissue is inherently age- and sex-matched, it could provide an unrepresentative result due to a field change in the whole bowel. Only small differences were found between normal and adjacent tissues for the target loci, however, indicating that the degree of field change for the loci analysed is minor. Variation in methylation even among individual crypts from the same biopsy has been detected[208] and this would suggest that analysis at the crypt level is necessary to identify the subtle differences between the two groups.

It is evident from this study that MS-MLPA is a useful platform to screen for candidate markers. It, however, provides less flexibility in assay design than bisulphite pyrosequencing. This is largely due to the fact that most commercially available MS-MLPA kits are predesigned and therefore, probes contained in these kits

are not custom-made for individual studies. Furthermore, the design for methylation sites is entirely reliant on the presence of restriction site for the *HhaI* endonuclease. Bisulphite pyrosequencing offers information on sequential CpG sites and its usefulness in identifying discriminatory sites will be evaluated subsequently in this thesis.

One criticism of this study is the use of non-macrodissected tumour tissues. Tumour specimens will invariably contain some amount of non-tumourous tissues, which could dilute the true methylation measurement. Therefore, the effect of tissue enrichment with macrodissection and laser-capture microdissection on methylation yield will be evaluated in the following chapter.

Results

Chapter 4

Assessment of validity and reproducibility of methylation assays and the impact of enrichment techniques on methylation yield

4.1 Introduction

This chapter results consists of two parts. In the first part, the methylation levels measured by MS-MLPA were correlated with bisulphite pyrosequencing, to explore whether measurement at a single site would be representative of the methylation status of adjacent CpGs. MS-MLPA evaluates CpG sites that are only recognised by the methylation-sensitive restriction enzyme, *HhAI*. This usually translates to quantification of methylation at only one or two CpG sites for any given locus. In contrast, bisulphite pyrosequencing technology is capable of quantifying multiple, sequential CpG sites. This could potentially identify the most discriminatory sites between two different phenotypic groups of tumours. The methylation levels of *ESR1, CDH13, CHFR, APC* and *RARB,* measured by MS-MLPA were correlated with that measured by pyrosequencing. These five genes were selected as their methylation levels were significantly higher in cancer compared to adjacent tissue and normal controls.

In the second part, the reproducibility of methylation results in formalin-fixed paraffin embedded (FFPE) tissues, which have been archived for at least 4 years, was investigated. Although fresh frozen tissues provide the gold standard tissue quality for nucleic acid research, most hospitals lack the infrastructure to support collection and storage of fresh frozen tissues. The current practice of tissue processing and archiving in most pathology departments involves formalin fixation and paraffin embedding tissues. The clinical application of any tissue-based molecular assays will remain limited unless the results are reproducible in FFPE clinical samples. One major problem with formalin fixation is the extensive cross-linking of proteins and

nucleic acids that results in DNA fragmentation and low DNA quality[209]. Many methylation assays such as pyrosequencing, real-time MSP and COBRA incorporate a bisulfite modification step, which further degrades DNA[210], making assessment of methylation in FFPE tissue challenging. This has been overcome in our laboratory using an optimised protocol of DNA extraction and bisulphite treatment.

One problem common to both fresh frozen and FFPE samples is tissue heterogeneity. Any cancer tissue biopsy will invariably contain some stromal and adjacent morphologically normal cells. These, in theory, could 'contaminate' the readout, either giving abnormally high or low results in quantitative assays. To circumvent this problem, tumour cells can be isolated under direct microscopy. Manual macrodissection is a relatively easy and quick method of enriching for tumour cells by excluding the majority of non-tumour cells. Laser-capture microdissection offers higher resolution and could isolate relatively pure individual cells[211]. This state-of-the-art technology, however, is expensive and is not available in many clinical laboratories. Whether such high level of cell purity is necessary for epigenetic and genetic testing in colorectal cancer remains unknown. Therefore, the difference in methylation results between manual macrodissected and laser capture microdissected DNA was evaluated using the *APC* and *LINE-1* pyrosequencing assays. *APC* and *LINE-1* were selected as they represent locus-specific and global methylation respectively[212, 213].

4.1.1 Bisulphite pyrosequencing

Bisulphite pyrosequencing uses bisulphite treatment, PCR and sequencing to differentiate methylated cytosine from unmethylated cytosine. A major strength of this technique is its ability to quantitatively measure methylation at sequential CpGs without a dependence upon a restriction site, a feature of MS-MLPA. Furthermore, it is high-throughput, relatively low cost and do not require labelled primers, labelled nucleotides and gel electrophoresis. These features make bisulphite pyrosequencing more attractive than conventional Sanger sequencing[214]. Prior to pyrosequencing, CpG sites of genomic DNA are first chemically converted by bisulfite treatment. Following incubation with sodium bisulfite, cytosine forms 5,6-dihydro-6-sulphonate, which then undergoes hydrolytic deamination to form 5,6-dihydrouracil-6-sulphonate. Subsequent desulphonation with alkaline treatment results in conversion to deoxyuracil (U). The methyl group on 5-methylcytosine protects the amine group from deamination during this process. Therefore, bisulfite treatment gives rise to different DNA sequences for methylated and unmethylated DNA (Table 4.1). During PCR, deoxyuracil is amplified as thymine (T) whilst methylated cytosine remains as cytosine (C).

	Original sequence	After bisulfite treatment
Unmethylated DNA	N-C-G-N-C-G-N-C-G-N	N-U-G-N-U-G-N-U-G-N
Methylated DNA	N-**C**-G-N-**C**-G-N-**C**-G-N	N-**C**-G-N-**C**-G-N-**C**-G-N

Table 4.1. Differences in DNA sequence following bisulfite treatment. Taken from the Epitect® 96 Bisulfite Handbook, September 2009 (Qiagen).

Following PCR, a sequencing primer is hybridized to the single-stranded PCR amplicon and then incubated with enzymes DNA polymerase, ATP sulfurylase, luciferase and apyrase and substrates adenosine 5' phosphosulfate (APS) and luciferin. The incorporation of complementary deoxyribo-nucleotide triphosphate (dNTP) into the DNA strand is catalysed by DNA polymerase. Each incorporation is accompanied by release of pyrophosphate (PPi), which is then converted to ATP by ATP sulfurylase in the presence of adenosine 5' phosphosulfate (APS). ATP is used to drive the luciferase-mediated conversion of luciferin to oxyluciferin that generates light in amounts that are proportional to the amount of ATP. The light produced is detected and seen as a peak in the raw data output, called Pyrogram. Apyrase, a nucleotide-degrading enzyme, continuously degrades unincorporated nucleotides and ATP. When degradation is complete, another nucleotide is added. Addition of dNTPs is performed sequentially (Figure 4.1). In the pyrogram, methylated cytosine and unmethylated cytosine are represented as C and T peaks, respectively. The peak height ratio of C and T is used to calculate the methylation percentage at each site. Every pyrosequencing assay contains an internal control for bisulfite treatment. Cytosine not followed by guanine is usually not methylated and should be fully converted to thymine following bisulfite treatment and PCR. Full bisulfite conversion is confirmed if template shows thymine and no cytosine at this position.

Figure 4.1. Diagrammatic representation of the principle of pyrosequencing technology. Taken from http://www.pyrosequencing.com/DynPage.aspx?id=8726&mn1=1366.

4.2 Results

4.2.1 Internal validation and reproducibility of pyrosequencing assays

Internal validation of the *ESR1, CDH13, CHFR, APC, RARB* and *LINE-1* assays was carried out using fully methylated and unmethylated genomic DNA. All assays generally showed good concordance with both controls, having mean methylation percentage >90% for methylated controls and <15% for unmethylated controls, apart from *LINE-1* (Table 4.2) To assess precision and reproducibility of each assay, bisulfite treatment and PCR of a control DNA (fresh frozen tissue of a rectal adenocarcinoma) were performed and repeated on four different days. This was followed by pyrosequencing of each sample on five separate occasions. The results showed very little variation in the level of methylation detected at individual CpG sites for all assays. The maximum standard deviations detected among the sites for *ESR1, CDH13, CHFR, APC, RARB* and *LINE-1* were 2.00%, 0.89%, 1.30%, 1.82%, 2.07% and 0.96% respectively. The minimum standard deviation detected for the six genes were 1.14%, 0.54%, 0.84%, 0.45%, 1.14% and 0.50% respectively.

4.2.2 Validation of MS-MLPA methylation using pyrosequencing

40 samples, consisting of rectal cancers and adjacent tissues were reanalysed using pyrosequencing for methylation of *ESR1, CDH13, CHFR, APC* and *RARB*. These samples had been previously analysed by MS-MLPA and represented a wide range of methylation for these five genes. Methylation levels were found to be concordant between the two methylation platforms across all CpG sites, indicating that methylation level at a single CpG site is generally reflective of the methylation status of surrounding CpGs for these five genes. The Spearman Rank correlation coefficient,

ρ, for each CpG site is shown in Table 4.2. Representative graphs demonstrating the correlation between these two platforms are shown in Figure 4.2.

4.2.3 Analysis of methylation levels across consecutive CpG sites

Variations in methylation levels for *ESR1, CDH13, CHFR, APC* and *RARB* were observed across the CpG sites. The distribution of methylation levels across CpG sites for these five genes is shown in Figure 4.3. The coefficient of variation, which is a normalised measure of dispersion (standard deviation/mean) provides a more accurate assessment of dispersion than standard deviation. For each CpG, the coefficient of variation ranged from 0.209 to 0.692 for *ESR1*, from 0.061 to 0.713 for *CDH13*, 0.160 to 2.65 for *CHFR*, from 0.082 to 1.67 for *APC* and from 0.045 to 0.636 for *RARB*.

Assay		Control	CpG sites										
			Site-1	Site-2	Site-3	Site-4	Site-5	Site-6	Site-7	Site-8	Site-9	Site-10	Mean
ESR1 (5 sites)	Methylation percentage (%)	Unmethylated	10	9	6	7	-	-	-	-	-	-	8.00
		Methylated	92	100	91	95	-	-	-	-	-	-	94.50
	Correlation coefficient, ρ		0.80	0.82	0.73	0.67							0.87
CDH13 (9 sites)	Methylation percentage (%)	Unmethylated	10	7	12	15	11	13	13	9	10	-	11.11
		Methylated	91	95	95	99	92	100	90	94	100	-	95.11
	Correlation coefficient, ρ		0.84	0.81	0.81	0.77	0.8	0.72	0.76	0.78	0.74		0.82
CHFR (10 sites)	Methylation percentage (%)	Unmethylated	0	5	11	0	0	8	3	3	0	0	3.00
		Methylated	91	100	93	90	91	100	99	92	91	100	94.70
	Correlation coefficient, ρ		0.81	0.74	0.70	0.70	0.75	0.74	0.70	0.64	0.60	0.66	0.72
APC (10 sites)	Methylation percentage (%)	Unmethylated	3	6	7	9	0	6	4	5	0	8	4.80
		Methylated	98	90	100	100	92	100	91	90	92	100	95.30
	Correlation coefficient, ρ		0.88	0.91	0.90	0.81	0.90	0.84	0.89	0.89	0.87	0.88	0.89
RARB (10 sites)	Methylation percentage (%)	Unmethylated	6	7	0	5	5	9	0	9	6	12	5.90
		Methylated	97	100	91	96	100	100	93	100	95	100	97.20
	Correlation coefficient, ρ		0.93	0.91	0.90	0.90	0.92	0.92	0.92	0.94	0.94	0.93	0.95
LINE-1 (6 sites)	Methylation percentage (%)	Unmethylated	29	36	36	32	35	33	-	-	-	-	33.50
		Methylated	79	67	73	75	88	78	-	-	-	-	76.67

Table 4.2. The methylation percentages for unmethylated and methylated controls of individual CpG sites for ESR1, CDH13, CHFR, APC RARB and LINE-1 pyrosequencing assays. The number of CpG sites evaluated for each assay is shown in parentheses. Correlation coefficient, ρ, between individual pyrosequencing sites and MS-MLPA are all statistically significant (P<0.05). The MS-MLPA assay does not contain any LINE-1 CpG site, therefore, correlation coefficient for LINE-1 was not calculated.

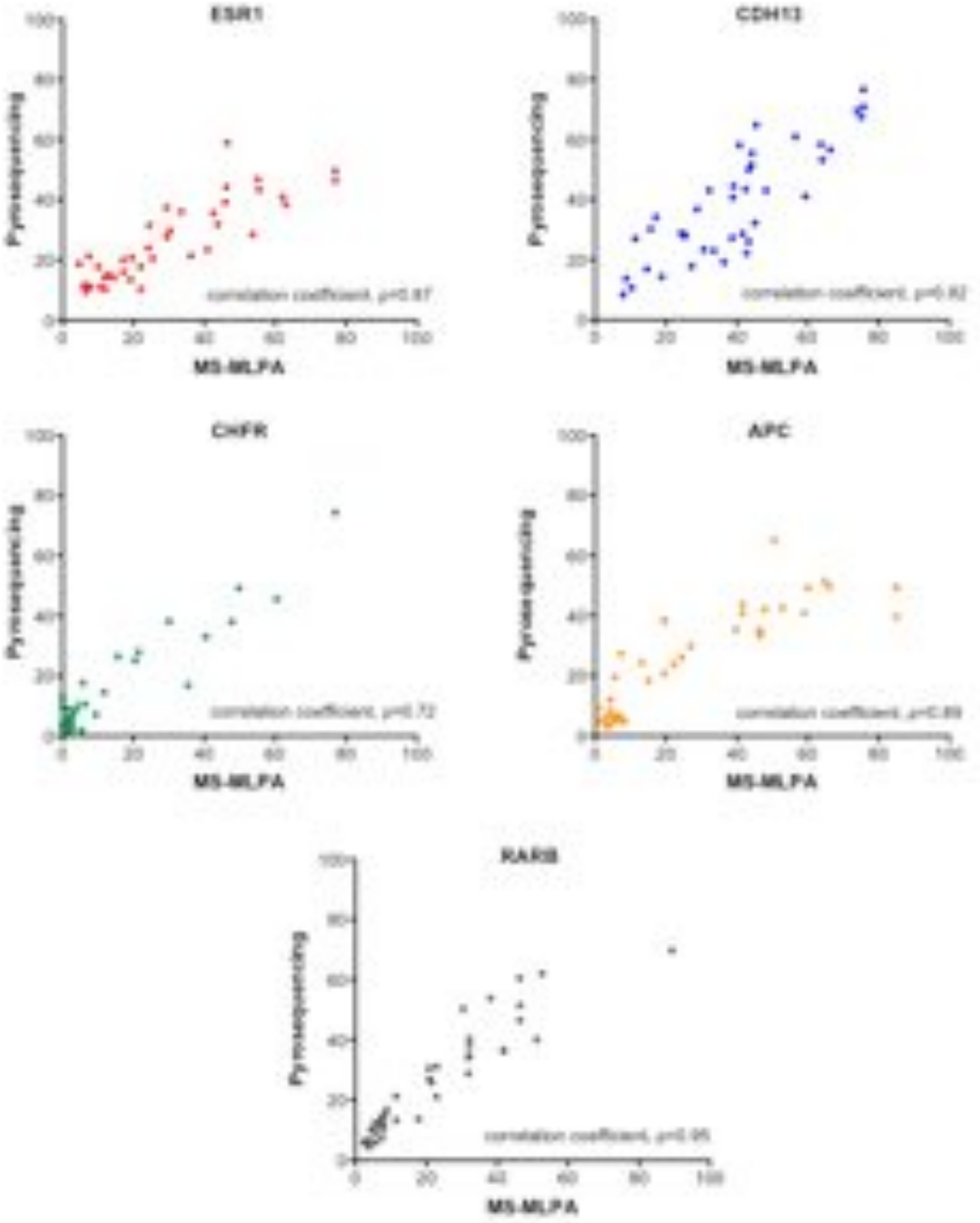


Figure 4.2. Correlation of mean methylation levels of ESR1, CDH13, CHFR, APC and RARB between pyrosequencing and MS-MLPA. Spearman Rank correlation coefficients, ρ, for these five genes are all statistically significant (P<0.05).

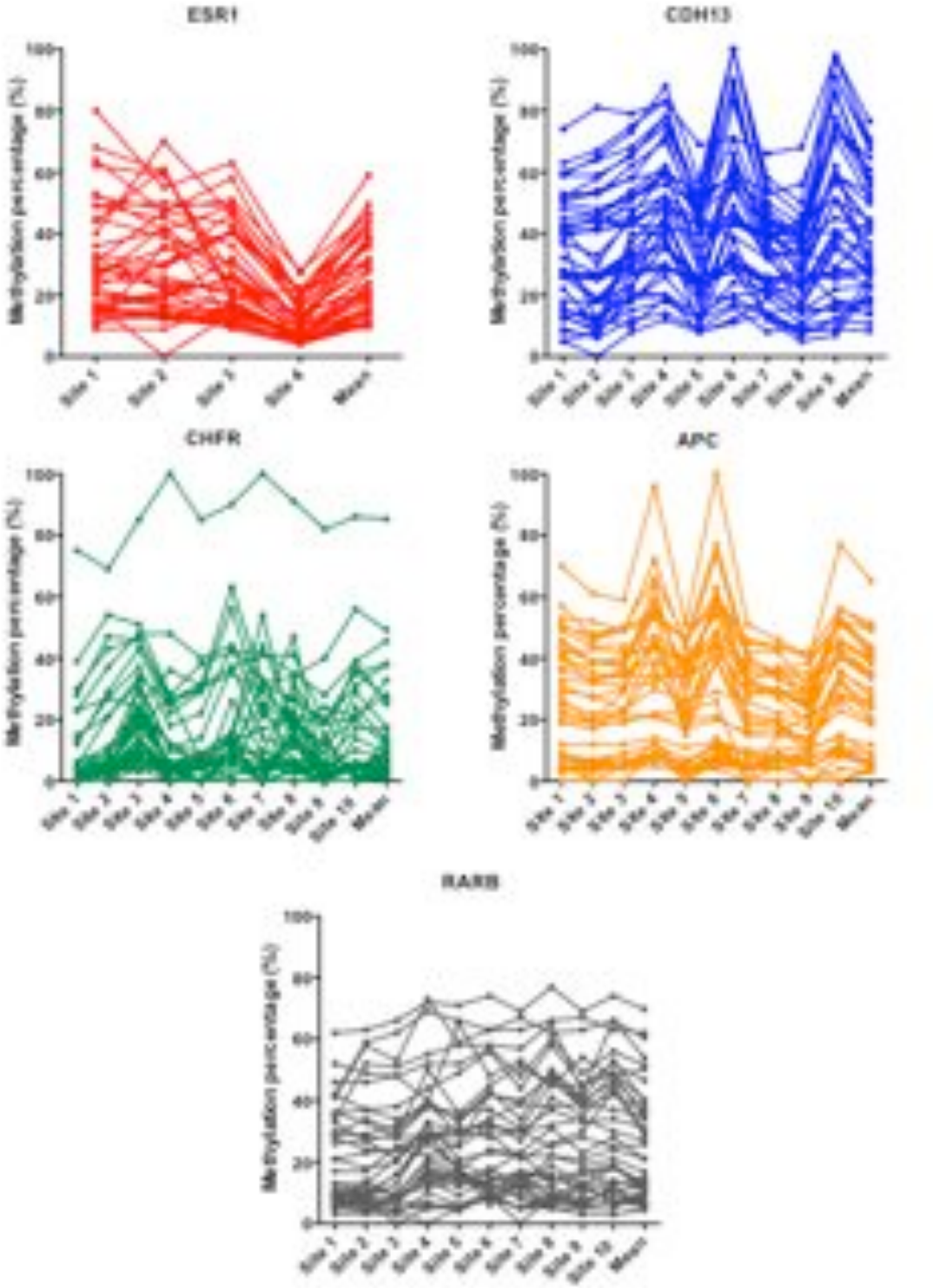


Figure 4.3. Distribution of methylation levels across consecutive CpG sites for ESR1, CDH13, CHFR, APC and RARB in 40 rectal tumour samples.

4.2.4 Impact of tissue enrichment techniques on methylation yield

Tissue enrichment techniques such as manual macrodissection and laser capture microdissection can increase the proportion of cells of interest for DNA extraction but will reduce the overall amount of DNA available for extraction. This is most problematic for FFPE tissues when the starting material consists of poor DNA quality. To determine the minimal amount of cells required for methylation analysis using bisulphite pyrosequencing, laser capture microdissection was performed on four different FFPE slides taken from the same rectal cancer sample that had been archived for four years. Representative sections for microdissection are shown in Figure 4.4. Four different amounts of cells were captured; 1 000, 5 000, 10 000 and 15 000 cells. These are an approximation, calculated from a X10 microscopic view. DNA extraction and bisulphite treatment of the four populations of cells were performed according to our optimised protocol described in Chapter 2. Following PCR using primers for the *APC* pyrosequencing assay, 5 μl of PCR products were run on a 1% agarose gel to determine the presence and intensity of the amplicon bands. As a positive control, DNA extracted from a 1cm X 1cm area of FFPE slide from the same sample was used.

PCR using *APC* primers on samples with approximately 1 000 bisulphite-treated cells did not produce any discernable amplicon bands. The band intensity for the remaining three groups (5 000, 10 000 and 15 000 bisulphite-treated cells) was similar (Figure 4.5). Subsequent pyrosequencing analysis only produced results for samples with 5 000, 10 000 and 15 000 bisulphite-treated cells (Figure 4.6). This result indicates that 5 000 cells are sufficient for successful pyrosequencing of *APC*.

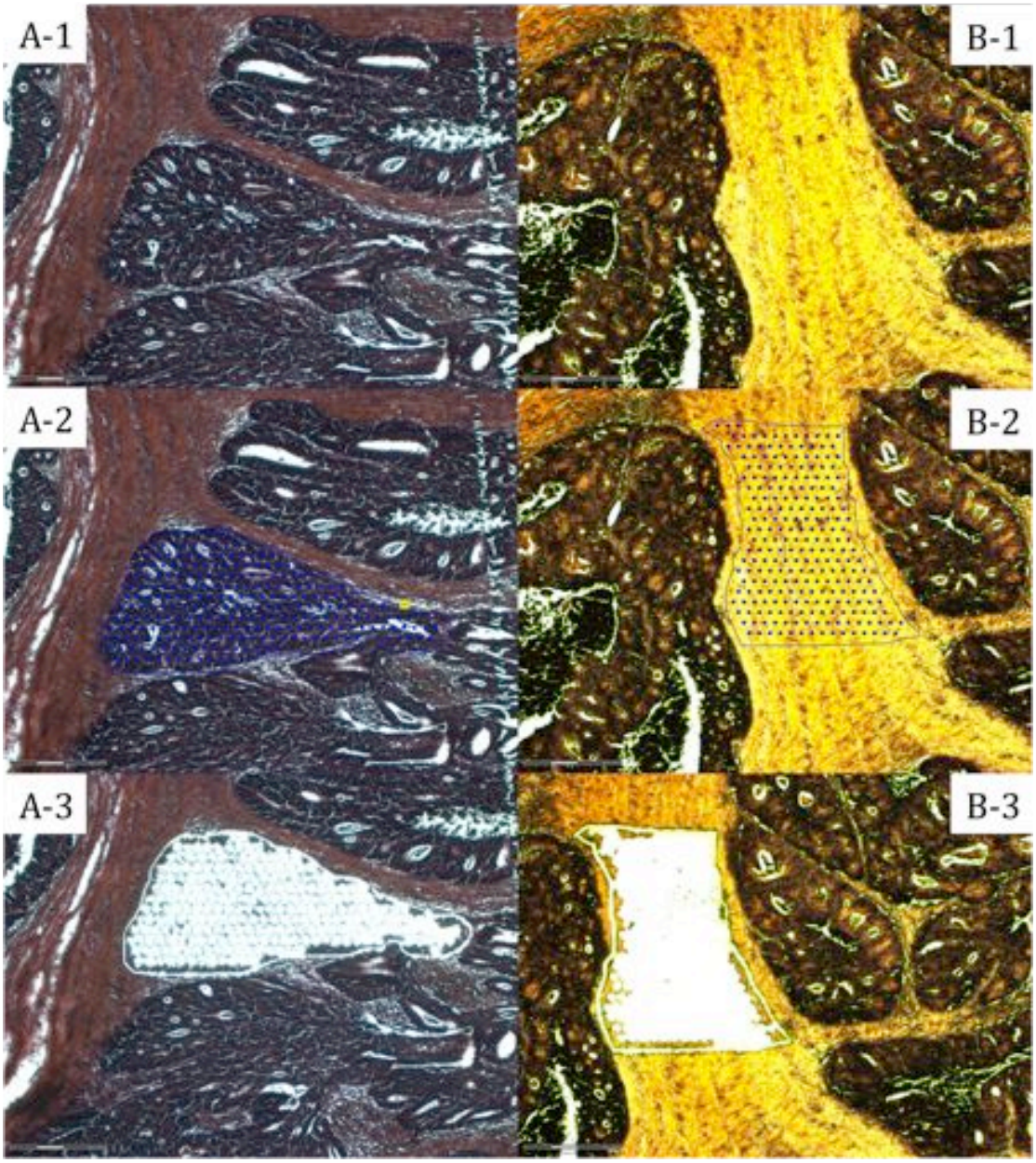


Figure 4.4. A-1) and B-1) Two representative FFPE sections of rectal tumour. Areas of tumour cells (A-2) and stromal tissue (B-2) marked for laser capture microdissection. Residual areas of tumour cells (A-3) and stromal tissue (B-3) following microdissection. Magnification X10.

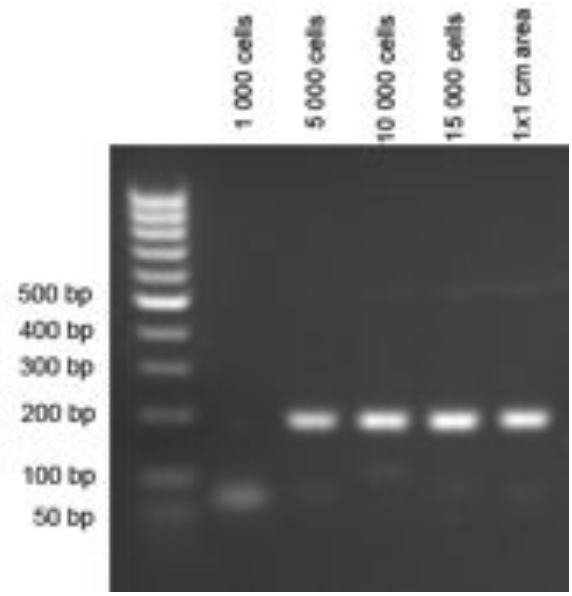


Figure 4.5. APC amplicons (195 bp) of microdissected samples containing approximately 1 000, 5 000, 10 000 and 15 000 cells, taken from the same FFPE rectal tissue, are visualised on a 1% agarose gel. A 1 x 1 cm macrodissected area of the same tissue was used as positive control.

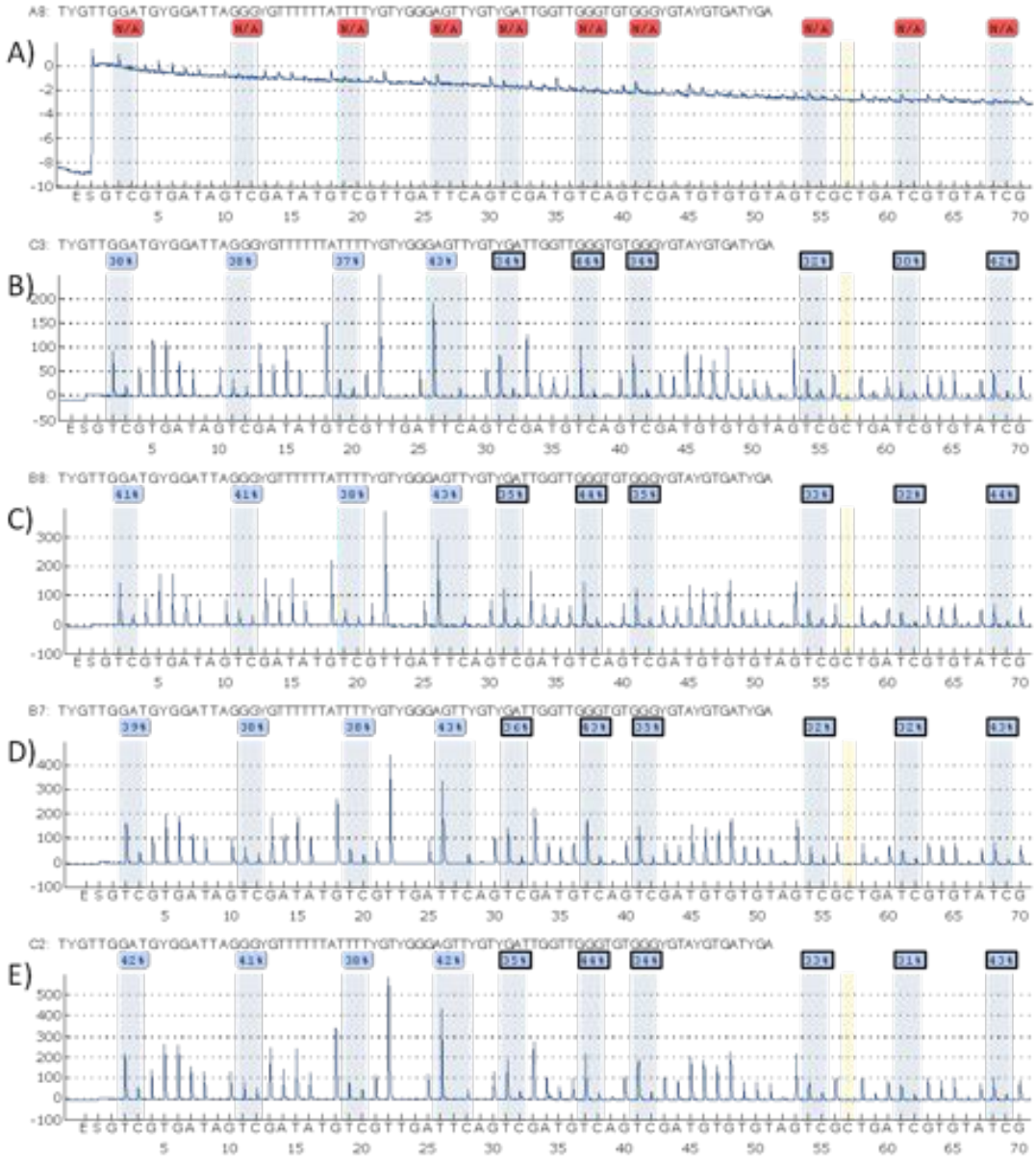


Figure 4.6. Pyrogram of APC generated from samples with A) 1 000 cells, B) 5 000 cells, C) 10 000 cells, D) 15 000 cells and E) 1 x 1 cm area, taken from the same rectal cancer tissue. The sample with approximately 1 000 cells failed pyrosequencing. The sequence in the upper part of each Pyrogram represents the sequence under investigation. The sequence below the Pyrogram indicates the sequentially added nucleotides. The gray regions highlight the analyzed C/T sites, with percentage values for the respective cytosine above them. Yellow parts highlight the positions where a cytosine was added to verify the complete conversion from unmethylated cytosine to thymine.

4.2.5 Comparison of methylation in matched fresh frozen, macrodissected FFPE and laser capture microdissected FFPE tissues

Ten matched fresh tissues, macrodissected FFPE and laser capture microdissected FFPE rectal cancers were analysed for *APC* and *LINE-1* methylation. Their clinicopathological details are summarised in Table 4.3. The two assays selected represent two different methylation phenomena that occur in cancer; promoter hypermethylation (*APC*) and global hypomethylation (*LINE-1*). For the purpose of this analysis, the average methylation percentage across all CpG sites for each assay was used. Any difference in methylation of less than 0.61% for *APC* and 0.47% for *LINE-1* (these values correspond to the standard deviation of the average methylation of *APC* and *LINE-1* obtained by taking readings from five different days) was considered a consequence of the run-to-run variation of the assay and therefore would not reflect the true difference in methylation due to enrichment techniques.

Sample	Age	Sex	Tumour stage			Maximum diameter (mm)	Vascular invasion	Degree of differentiation
			T	N	M			
A	63	M	1	0	0	25	no	well
B	70	F	1	0	0	25	no	well
C	76	F	2	0	0	35	no	well
D	59	F	2	0	0	40	no	moderate
E	78	M	2	0	0	45	no	moderate
F	86	F	2	0	0	60	no	moderate
G	75	M	3	0	0	60	yes	moderate
H	70	M	3	1	0	50	yes	moderate
I	73	F	3	1	0	50	yes	moderate
J	76	F	4	0	1	40	no	moderate

Table 4.3. Clinicopathological details of samples used. M=male, F=female.

Tissue enrichment techniques with either macrodissection or laser capture microdissection did not result in any significant increase in *APC* methylation. Some

macro and microdissected samples showed higher or lower methylation levels, when compared to their matched fresh tissues, but the differences were only marginal. Similar results were also noted for *LINE-1* methylation. Again, tissue enrichment techniques did not yield any significant improvements in *LINE-1* methylation levels.

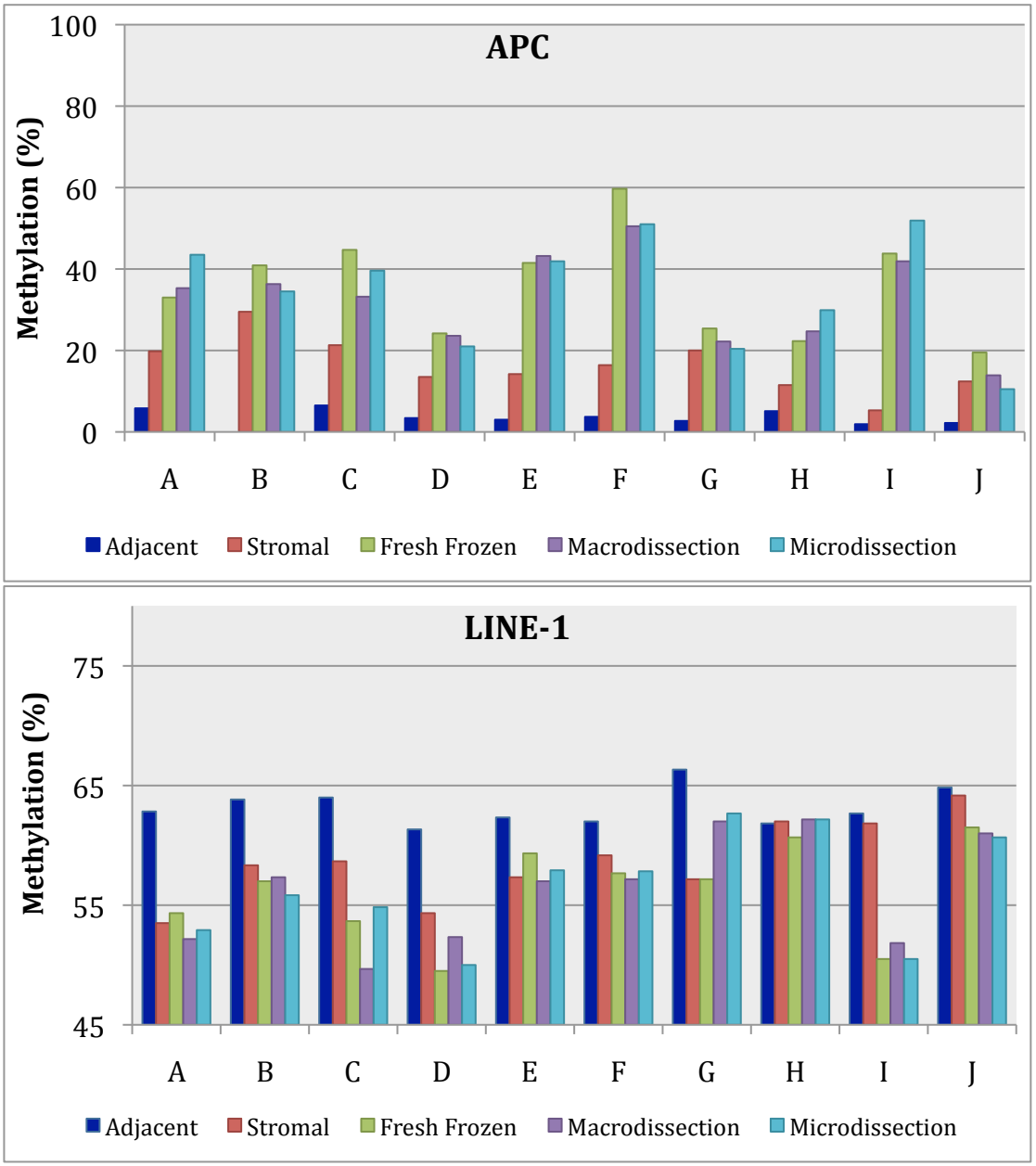


Figure 4.7. Methylation levels of APC and LINE-1 for adjacent tissue, stromal tissue, fresh frozen tissue, manual macrodissected tissue and laser capture microdissected tissue of samples A to J.

4.2.6 Comparison of methylation in tumour adjacent 'normal' tissues, laser capture microdissected stromal and cancer tissues

Laser capture microdissected tissues contained a greater proportion of cancer cells compared to matched macrodissected tissues, although no significant increase in promoter methylation (as detected in promoter of *APC*) or significant decrease in global methylation (as detected using *LINE-1*) was observed as a consequence. It is possible that the connective tissues that provide a supportive framework around cancer cells also undergo methylation changes and could explain why methylation values in macrodissected tissues, containing both connective tissues and cancer cells, did not differ much compared to relatively pure laser capture microdissected tissues. To examine this hypothesis, stromal tissues adjacent to cancer cells were microdissected from the same ten cancer samples described above and were subjected to pyrosequencing analysis using the *APC* and *LINE-1* assays. Their methylation levels were evaluated against that of matched adjacent macroscopically normal tissues taken at least 5 cm away from the tumour edge and microdissected cancer tissues.

For *APC*, significant methylation changes were observed in stromal tissues. Overall methylation levels of stromal tissues were significantly higher than that of adjacent 'normal' tissues [median (interquartile range) 15.3% (13.2-19.9) vs. 3.4% (2.6-4.1%), $P<0.0001$] but were still less than those observed in microdissected cancer tissues [38.2% (28.8-43.8%) $P=0.0003$]. Individually, the methylation levels of all ten stromal tissues were at least two-fold higher than those seen in tumour adjacent tissues (range 2.3-7.4). When compared to microdissected cancer tissues, the methylation

levels of 3 stromal tissues (B, G and J) were comparable to that seen in their matched cancer tissues (Figure 4.7).

Methylation levels for *LINE-1* in stromal tissue were significantly lower than those seen in tumour adjacent tissue [58.5% (57.2-61.2%) vs. 62.8% (62.1-64.0%) *P*=0.0046] and were comparable to the low levels seen in cancer tissue [56.5% (53.1-57.9%) *P*=0.13]. Only one stromal tissue (H) showed higher methylation compared to matched adjacent tissues. Interestingly, for sample H, there was no appreciable difference between adjacent and cancer tissues.

4.3 Discussion

Similar to the commercially available MS-MLPA, in-house design of pyrosequencing assays offers high sensitivity for quantification of methylation. In addition, the pyrosequencing assays in this study were highly reproducible, showing very little run-to-run variation. This provided confidence that pyrosequencing technology could be used to corroborate the results obtained from MS-MLPA.

Unlike many established quantitative methylation approaches, such as real-time methylation-specific polymerase chain reaction (MethyLight) and combined bisulphite restriction analysis (COBRA) which only examine methylation at one CpG site, pyrosequencing technology provides quantitative analysis of multiple sequential sites, making it a more informative methylation platform. It is only with this unique function that we were able to show that although there is a general concordance, variation in methylation levels among multiple sequential CpG sites exists.

Reliable pyrosequencing results can be obtained from archived laser capture microdissected tissues, provided sufficient cells were harvested for analysis. Laser capture microdissection, however, is expensive and time-consuming and for these reasons, would be difficult to be incorporated into routine clinical laboratory. Manual macrodissection is a cheaper alternative tissue enrichment technique but receives criticism for harvesting less 'pure' samples compared to laser capture microdissection. Only one study has examined the molecular differences between matched macrodissected and microdissected cancer tissues. de Bruin *et al*[215] observed minor differences in gene expression profiling in cancers dissected by

these two techniques. The impact of different enrichment techniques on methylation profiling has not been explored, until recently. Our results, using *APC* and *LINE-1* pyrosequencing assays, showed that laser capture microdissected samples only produced minor and insignificant methylation differences compared to macrodissected samples.

The main difference between macrodissected and microdissected tissues, in terms of tissue composition, is the higher proportion of stromal cells in macrodissected tissues. The fact that there were only minor differences in methylation between macrodissected and microdissected tissues led us to investigate the impact of stromal tissues on methylation. The results indicate that, compared to adjacent macroscopically 'normal' mucosa, stromal tissues have undergone aberrant methylation changes but not to an extent that is comparable to cancer tissues. Our results are consistent with other studies, which have found gene-specific methylation of stromal tissues in breast and prostate cancer[216, 217]. The mechanism by which stromal tissues acquire gene-specific methylation remains unclear but several hypotheses exist. One such hypothesis suggests that methylated stromal tissues are derived from the transition of tumour epithelium into cells with a mesenchymal phyenotype[217]. These transdifferentiated cells are associated with increased cell motility and migration and therefore, have a potential for malignant transformation.

Of interest, four macrodissected rectal cancers had lower *APC* methylation levels compared to their matched undissected fresh tissues but the difference was only

minimal. This is unlikely to be due to the run-to-run variation, as the standard deviation calculated from the repeated experiments was less than the difference in methylation between macrodissected and fresh tissues. It is possible that the difference could be due to intratumoural heterogeneity. Discordant results on HER2 testing due to intratumoural heterogeneity have been described[218]. Although taken from the same resected rectal specimen, fresh tissues were taken from the core of the tumour while the FFPE tissues used could have been sampled at varying distances from the tumour core. It is, however, worth emphasising that the difference in methylation between undissected fresh and macrodissected tissues is minimal.

Three important conclusions can be drawn from the results of this chapter. Firstly, pyrosequencing revealed variations in methylation levels among CpGs within a locus. These differences may allow us to identify the most discriminatory sites when used to correlate with conventional histopathological criteria. This will be evaluated in the following chapter. Secondly, the minor difference in methylation between macrodissected and microdissected tissues does not support the routine use of laser capture microdissection to enrich for cancer cells in paraffin embedded sections. Thirdly, the continued use of undissected fresh rectal tissues in our studies is appropriate due to the relatively low impact of tissue enrichment.

Results

Chapter 5

Site-specific methylation changes and their association with disease progression in rectal cancer

5.1 Introduction

Earlier analysis (Chapter 3) found that hypermethylation of specific loci was associated with early stage rectal cancer. One of the aims of the current chapter was to refine this molecular signature by investigating the methylation profile of surrounding CpG sites of these loci. This was accomplished using bisulphite pyrosequencing. Bisulphite pyrosequencing provides a greater in-depth methylation analysis compared to MS-MLPA as it can examine multiple sequential CpG sites for any DNA sequence of interest. Assay design for pyrosequencing is also flexible and can be performed in-house; two important factors for exploratory analysis. This chapter also aimed to determine whether the examined CpG sites were clinically relevant for colonic cancer.

The methylation sites chosen for this study can be broadly divided into three categories. The first category consists of genes (*APC, RARB, GSTP1, TIMP3* and *CASP8*) whose concomitant hypermethylation has already been shown, in Chapter 3, to be associated with early-stage rectal cancer. The rationale for their selection in this chapter is to determine whether the results obtained by MS-MLPA are reproducible using bisulphite pyrosequencing.

The second category consists of genes whose methylation has been reported in the literature to be associated with prognosis in gastrointestinal tumours. These include *CXCL12*[219], *UNC5C*[220], *CDH1*[221], *MINT3*[188, 189], *MINT17*[188, 189] and *LINE-1*[213, 222]. CXCL12 is a small chemotactic cytokine, and via its interactions with CXCR4 receptor[223], has a role in gut vascularisation, a fundamental process in intestinal mucosal homeostasis

and immunity[224]. DNA hypermethylation is associated with downregulation of CXCL12 expression in colorectal cancer cells and re-expression results in reduced *in vivo* metastatic potential[219].

UNC5C belongs to a family of type 1 transmembrance receptors that induce apoptosis when not engaged with their ligand, Netrin-1, but mediate signals for proliferation, differentiation or migration when ligand-bound[225, 226]. There have been conflicting reports on the methylation status of *UNC5C* in different stages of colorectal cancer. Hibi *et. al.*[227] found that methylation of *UNC5C* was more frequently found in advanced colonic cancers but Shin *et. al.*[220] showed a stepwise reduction in the proportion of colorectal tumours with *UNC5C* methylation from adenomas to Duke's 'D' tumours.

The *CDH1* gene encodes E-Cadherin, a transmembrane protein found on epithelial cells and is responsible for cell-cell adhesion[228]. Loss of cadherin-dependent anchorage may promote detachment of tumour cells with subsequent invasion to adjacent cells and dissemination to distant organs. In diffuse-type gastric[229] and lobular breast cancer[230], the mechanism responsible for E-cadherin down-regulation is frequently due to gene mutation but in colorectal cancer, this mechanism is infrequent and instead, transcriptional silencing of E-cadherin is associated with promoter methylation[221].

Methylated-in-tumour (MINT) loci are CpG-rich areas located in nonprotein-encoding DNA regions and have been found methylated in colorectal[231], gastric[232]

and malignant melanoma[233]. Among these loci, MINT3 and MINT17 were most informative in predicting local and distant recurrence in rectal cancer[188, 189].

Long interspersed nucleotide element (LINE-1) repeats are autonomous retrotransposons that are ubiquitous in the human genome. In somatic cells, LINE-1 repeats are heavily methylated, which restricts the activities of retrotransposal elements[234]. These activities include inducing genomic deletions[235], chromosome breaks[236] and translocations[237] leading to enhanced genomic instability of the cell. LINE-1 hypomethylation has been found in a range of human malignancies, including colorectal cancer and its methylation status has been suggested to represent global methylation status[213]. Progressive LINE-1 hypomethylation has been observed in colorectal tumours with increasing Dukes' stage[222]. In a large cohort study, LINE-1 hypomethylation is linearly associated with an increase in colon cancer–specific and overall mortality[213].

The third category consists of genes (*ESR1, CDH13, CHFR, and DAPK1)* whose methylation levels, as analysed by MS-MLPA in Chapter 3, bear no correlation with histopathological features. As already described, MS-MLPA only examines a single CpG site for each locus and, as such, may miss clinically important surrounding CpG sites. Therefore, the reason for their inclusion is to determine whether analysis of CpGs surrounding these MS-MLPA sites would yield a different result that may be clinically relevant.

Methylation analysis was initially performed in an exploratory cohort consisting of 64 fresh frozen rectal cancers. The results were subsequently validated in a second independent cohort consisting of 69 formalin-fixed paraffin-embedded rectal cancers. To determine whether our results are also applicable to colonic cancers, the methylation analysis was performed in a third cohort consisting of 68 fresh frozen colonic cancers.

5.2 Results

5.2.1 Internal validation and reproducibility of pyrosequencing assays

Similar to the internal validation already performed for the *ESR1, CDH13, CHFR, APC*, *RARB* and *LINE-1* assays in the previous chapter, validation for the *DAPK1, CXCL12, UNC5C, CDH1, MINT3*, *MINT17*, *GSTP1, CASP8* and *TIMP3* pyrosequencing assays was carried out using fully methylated and unmethylated genomic DNA. All assays generally showed relatively high sensitivity in detecting low and high levels of methylation (Table 5.1). To assess the precision and reproducibility of each assay, bisulfite treatment and PCR of a control DNA (fresh frozen tissue of a rectal adenocarcinoma) were performed and repeated on four different days. This was followed by pyrosequencing of each sample on five separate occasions. The results showed very little variation in the level of methylation detected at individual CpG sites for all assays. The maximum standard deviation detected among the sites for *DAPK1, CXCL12, UNC5C, CDH1, MINT3*, *MINT17*, *GSTP1, CASP8* and *TIMP3* were 2.16%, 2.28%, 2.40%, 1.52%, 2.30%, 1.34%, 1.14% 1.84% and 2.36% respectively. The minimum standard deviation detected for the nine genes were 1.30%, 1.22%, 1.30%, 0.45%, 0.89%, 0.54%, 0.45%, 1.13% and 0.51% respectively.

Assay		Control	CpG sites										
			Site-1	Site-2	Site-3	Site-4	Site-5	Site-6	Site-7	Site-8	Site-9	Site-10	Mean
DAPK1	Methylation	Unmethylated	6	7	4	14	3	3	0	0	4	5	4.60
(10 sites)	percentage (%)	Methylated	88	90	90	99	90	86	92	94	97	100	92.60
CXCL12	Methylation	Unmethylated	8	3	2	6	7	6	8	-	-	-	5.71
(7 sites)	percentage (%)	Methylated	89	93	88	87	87	91	86	-	-	-	88.71
UNC5C	Methylation	Unmethylated	2	3	1	0	10	-	-	-	-	-	3.20
(5 sites)	percentage (%)	Methylated	86	89	92	84	93	-	-	-	-	-	88.80
CDH1	Methylation	Unmethylated	2	4	2	4	4	3	2	3	7	0	3.1
(10 sites)	percentage (%)	Methylated	93	97	89	96	92	96	97	100	91	97	94.8
MINT3	Methylation	Unmethylated	3	5	4	6	5	7	5	6	4	6	5.10
(10 sites)	percentage (%)	Methylated	86	89	91	80	84	88	90	85	86	97	87.60
MINT17	Methylation	Unmethylated	4	6	3	2	-	-	-	-	-	-	3.75
(4 sites)	percentage (%)	Methylated	90	89	87	89	-	-	-	-	-	-	88.75
GSTP1	Methylation	Unmethylated	2	4	2	4	4	3	2	3	-	-	3.00
(8 sites)	percentage (%)	Methylated	93	97	89	96	92	86	87	100	-	-	92.50
CASP8	Methylation	Unmethylated	0	2	4	0	-	-	-	-	-	-	1.43
(4 sites)	percentage (%)	Methylated	91	96	90	96	-	-	-	-	-	-	93.25
TIMP3	Methylation	Unmethylated	0	4	0	2	0	-	-	-	-	-	1.27
(5 sites)	percentage (%)	Methylated	91	91	88	88	89	-	-	-	-	-	89.40

Table 5.1. The methylation percentages for unmethylated and methylated controls of individual CpG sites for DAPK1, CXCL12, UNC5C, MINT3, MINT17, CASP8 and TIMP3 pyrosequencing assays. The number of CpG sites evaluated for each assay is shown in parentheses.

5.2.2 Demographics and clinicopathological characteristics of samples used

The samples used in this study consisted of two rectal and one colonic cancer cohorts. The first cohort of rectal cancer was identified from our tissue bank consisting of colorectal cancer specimens from patients treated at the University Hospitals of Birmingham NHS Trust between 2001 and 2004. Samples were taken from the core of the resected tumour by the operating surgeon immediately after surgery and flash-frozen in liquid nitrogen. The set comprise 64 rectal adenocarcinomas treated with total mesorectal excision. In addition, all 64 cancers have histologically confirmed tumour-free matched adjacent tissues (taken 5 cm or more from tumour edge) for paired analysis. The second cohort of rectal cancer comprises 69 radically excised formalin-fixed paraffin embedded (FFPE) rectal adenocarcinomas treated at the same institution between 2005 and 2010. The third, colonic set comprised 68 colonic adenocarcinomas treated in the same institution between 2001 and 2004. Patients who underwent neoadjuvant chemoradiotherapy, have a family history or previous history of colorectal cancer or inflammatory bowel disease were excluded. Patient demographics and clinicopathological characteristics are summarised in Table 5.2.

Clinicopathological features	Adjacent tissue (n=64)	Rectal cancer cohort 1 (n=64)	Rectal cancer cohort 2 (n=69)	Colonic cancer cohort (n=68)
Age (years)*	71 (33-89)	71 (33-89)	70 (31-92)	72(41-89)
Gender				
Male	42	42	36	31
Female	22	22	33	37
Dukes' stage				
A		17	15	5
B		15	28	36
C		22	19	12
D		10	7	15
Tumour size (mm)*		45 (13-110)	38.5 (9-70)	48 (15-120)
Vascular invasion				
No		44	43	37
Yes		20	26	31
Degree of differentiation				
Well		8	4	5
Moderate		52	63	48
Poor		4	2	15

*Table 5.2. Demographics and clinicopathological characteristics of samples used. *Values are median (range).*

5.2.3 Quantitative comparison of methylation changes in rectal cancer and matched adjacent, macroscopically normal rectal mucosa

All 15 loci except *CDH1*, *GSTP1, CASP8* and *TIMP3* showed rectal cancer-specific methylation changes. The mean methylation levels of *APC, CDH13, CHFR, RARB, ESR1, CXCL12, DAPK1, UNC5C, MINT3* and *MINT17* were significantly higher in rectal cancer compared to matched adjacent rectal mucosa while the methylation levels of *LINE-1* was significantly lower in rectal cancer compared to adjacent mucosa (Figures 5.1 and 5.2). No significant differences in methylation levels were observed for these 15 loci between the two rectal cancer cohorts.

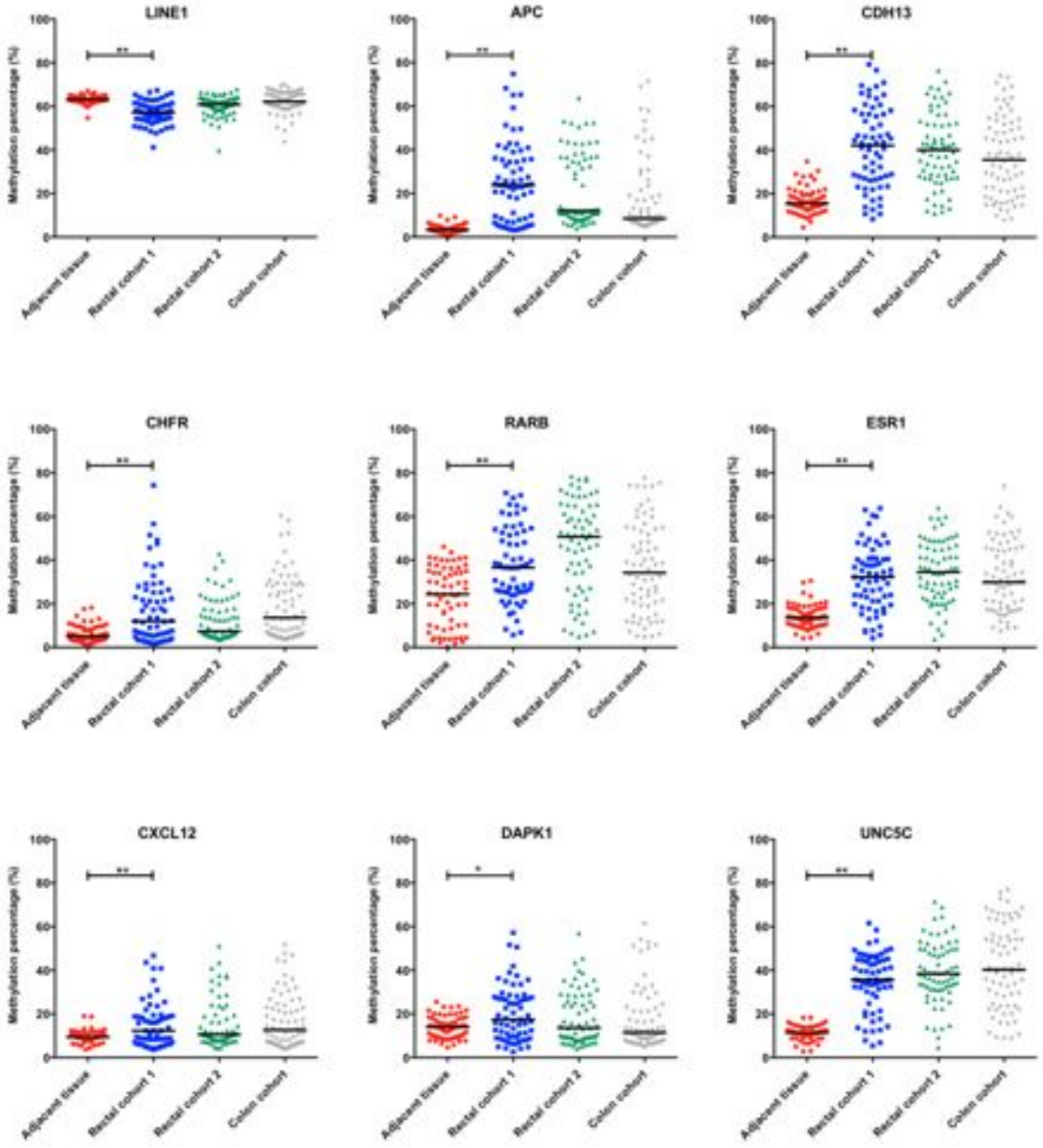


*Figure 5.1. Mean methylation levels of LINE1, APC, CDH13, CHFR, RARB, ESR1, CXCL12, DAPK1 and UNC5C in rectal cancer cohort 1 (n=64) and their matched adjacent tissue (n=64), rectal cancer cohort 2 (n=69) and colon cancer cohort (n=68). The horizontal bars represent the median methylation levels. **P<0.001, *P<0.05 (2-tailed Wilcoxon signed rank test).*

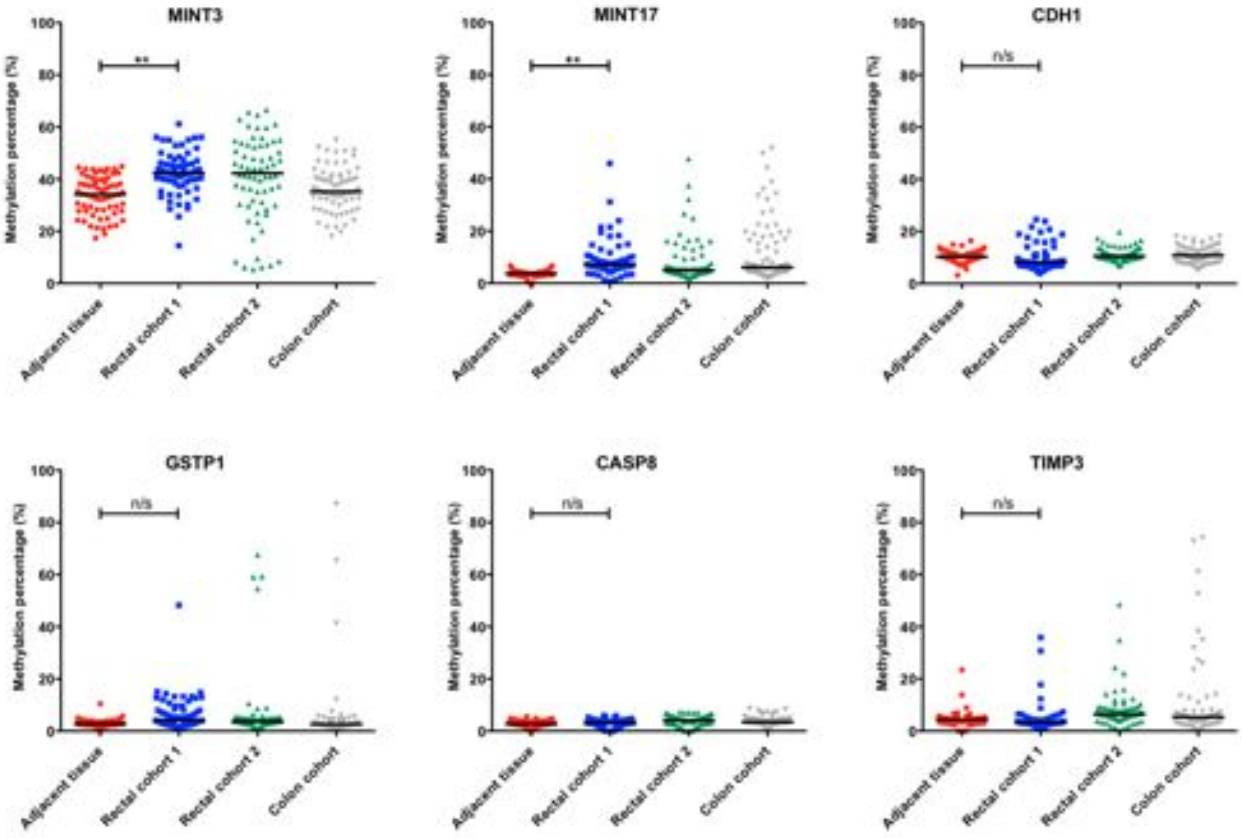


*Figure 5.2. Mean methylation levels MINT3, MINT17, CDH1, GSTP1, CASP8 and TIMP3 in rectal cancer cohort 1 (n=64) and their matched adjacent tissue (n=64), rectal cancer cohort 2 (n=69) and colon cancer cohort (n=68). The horizontal bars represent the median methylation levels. **P<0.001, n/s=not significant (2-tailed Wilcoxon signed rank test).*

5.2.4 Quantitative comparison of methylation changes between rectal and colonic cancer

Differences in molecular and histopathological features of lesions arising from the right side of the colon and of those from the left have been documented[238]. To determine whether such molecular changes would include methylation in our panel of genes, the methylation profile of the two rectal cancer cohorts were pooled (n=133) and compared with that of 68 colonic cancers. Both rectal and colonic cohorts were matched for age (*P*=0.715), gender (*P*=0.127), tumour size (*P*=0.613),

vascular invasion (*P*=0.171) and Dukes' stage (A/B vs. C/D, *P*= 0.652). The colonic cohort, however, has significantly greater number of poorly differentiated adenocarcinomas (*P*<0.001). Higher methylation levels of *APC* [median, (interquartile range, IQR) 19.5% (8.5-35.9%) vs. 8.4% (6.5-19.3%); *P*=0.0013] and *MINT3* [42.1% (35.2-48.6%) vs. 35.2% (31.3-40.8%); *P*=0.0002] were seen in the rectal compared to the colonic cohort. No other methylation differences were observed between the two groups.

5.2.5 Association of rectal cancer methylation with clinicopathological features

5.2.5.1 Evaluation of mean methylation levels as a prognostic marker

The mean methylation levels, calculated as the average methylation levels of all individual CpGs examined for any given gene, of these 15 genes were stratified according to conventional histopathological indices to identify markers associated with disease progression (Figure 5.3). In the first rectal cohort, methylation of *RARB* and *MINT3* were significantly higher in superficial (T1-2) compared to deep (T3-4) tumours while *CHFR* methylation was higher in T3-4 tumours. The mean methylation levels of *CXCL12* and *DAPK1* were significantly higher in node-negative tumours (N0) compared to those with lymph node metastasis (N1). Higher levels of *RARB* methylation were seen in non-metastatic cancers (M0) than those with distant metastasis (M1). When stratified according to Dukes' stage, the mean methylation levels of *RARB*, *DAPK1* and *CXCL12* were significantly higher in 'early' (Dukes' A and B) compared to 'advanced' (Dukes' C and 'D') cancers.

Pyrosequencing of our 15-gene panel was repeated in an independent, second cohort of 69 radically excised rectal cancers. Consistent with the results seen in the previous cohort, the mean methylation levels of *RARB* were also found significantly higher in T1-2 compared to T3-4 tumours. Higher methylation levels of *CDH13* were seen in cancers without lymph node metastasis (N0) compared to cancers with metastatic lymph nodes (N1) but this was not observed in the previous set. Higher methylation levels of *RARB* (also seen in the first cohort) and *CDH13* were seen in tumours without distant metastasis (M0) in comparison to those already metastasised to distant organs (M1).

5.2.5.2 Evaluation of individual CpG methylation levels as a prognostic marker

The previous chapter (Chapter 4) has shown that variation exists in methylation levels across consecutive CpG sites and taking an average value could mask potentially discriminatory individual CpGs. In the present study, a total of 112 CpG sites, representing the 15-genes were analysed individually in the first rectal cohort and 19 CpG sites were found to have significantly higher methylation levels in 'early' compared to 'advanced' cancers. These included 4/10 sites in *RARB,* 7/10 in *DAPK1,* 6/7 in *CXCL12,* 1/5 in *UNC5C* and 1/10 in *CHFR*.

The significance in methylation levels of the same four *RARB* sites and one of the DAPK1 sites identified in the first cohort was preserved in the second rectal cohort. In addition, all nine *CDH13* CpG sites in the second cohort were significantly higher in 'early' cancer. The absence of vascular invasion was associated with higher site-specific methylation levels in *RARB* and *CHD13* but this was only seen in the second

rectal cohort. No other association between methylation levels and the remaining prognostic features was observed.

5.2.6 Association of colonic cancer methylation with clinicopathological features

5.2.6.1 Evaluation of mean methylation levels as a prognostic marker

When the methylation profile of colonic cancers was stratified according to histopathological features, a similar methylation pattern was noted in colonic cancer. The mean methylation levels of *RARB* and *CDH13* were also higher in cancers with favourable histopathological features i.e. higher in N0 compared to N1 cancers and in M0 compared to M1 cancers.

5.2.6.2 Evaluation of individual CpG methylation levels as a prognostic marker

Four sites in *RARB*, eight in *CDH13* and one in *CHFR* identified in either of the two rectal cohorts as having higher methylation levels in 'early' cancer were also higher in 'early' compared to 'advanced' colonic cancers. The methylation levels of all twenty *RARB* and *CDH13* CpG sites were significantly higher in colonic cancers without lymphovascular invasion.

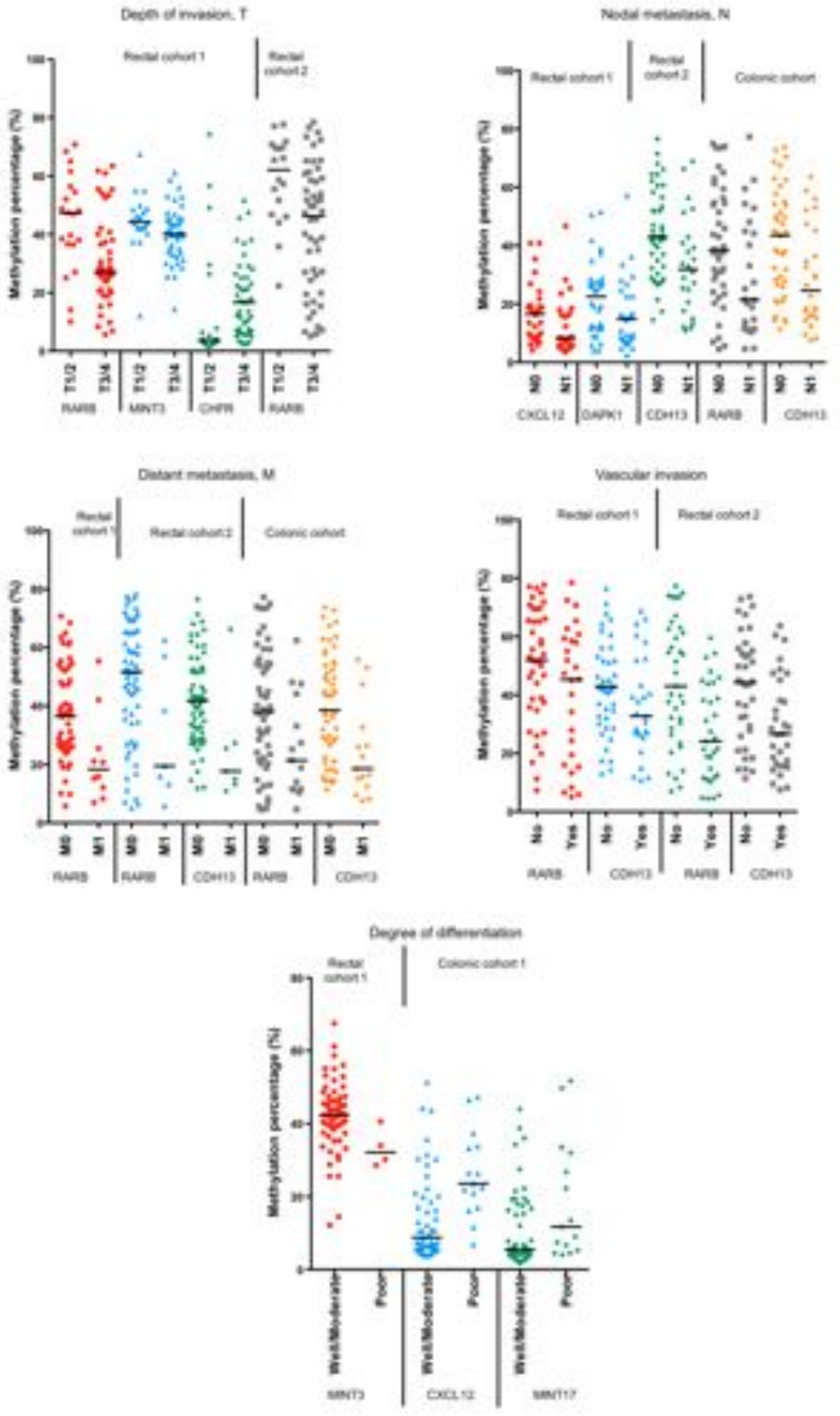


Figure 5.3. Mean methylation levels stratified according to histopathological features of disease progression in rectal and colonic cancers. Differential methylation of all genes shown in this figure is statistically significant (P<0.05, 2-tailed Mann-Whitney U Test).

5.2.7 A prognostic methylation signature in rectal and colonic cancer

Results from our earlier study using MS-MLPA have identified a novel methylation signature whereby concomitant hypermethylation of two or more of *APC, RARB, GSTP1, CASP8* and *TIMP3* is strongly associated with nodal disease in rectal cancer. To determine whether this result was reproducible using a different methylation platform, the methylation status of these five genes was reanalysed using bisulphite pyrosequencing in two different rectal cancer cohorts. The threshold used to consider a sample hypermethylated was the same as that used in Chapter3. Consistent with our previous study, the pyrosequencing results showed that concomitant hypermethylation of two or more of these five genes were associated with nodal status in both of the rectal cancer cohorts (Table 5.2). This methylation signature was also evaluated in the colonic cancer cohort and was also found to be associated with nodal disease (Table 5.2).

The remaining ten genes were added either individually or in combination to the methylation signature above to determine whether the statistical significance of the above methylation signature could be improved. Different permutations were taken into consideration and our results found that concomitant hypermethylation of three or more of *APC, RARB, GSTP1, CASP8, TIMP3, CXCL12* and *DAPK1* showed a statistically stronger association with nodal status in rectal cancer. This new methylation signature appeared to be unique to rectal cancer and not colonic cancer (Table 5.3).

5.2.8 Sensitivity and specificity of methylation signature in predicting nodal disease in rectal cancer

To determine the overall sensitivity and specificity in predicting nodal disease, the methylation results from the two rectal cancer cohorts were pooled. The sensitivity and specificity of the original methylation signature consisting of *APC*, *RARB, GSTP1, CASP8* and *TIMP3* was 68.0% (95% confidence interval (CI) +/- 8.0%) and 63.7% (95% CI +/- 8.0%) respectively. The addition of *CXCL12* and *DAPK1* into the original methylation signature improved the specificity to 75.8% (95% CI +/- 7.0%) while maintaining the sensitivity. Although there is an improvement in the specificity, the sensitivity of predicting nodal disease is still relatively low and is unlikely to have any clinical impact. For any prognostic biomarker to have any clinical utility, its sensitivity and specificity should be in the region of at least 90%. Further work to identify more accurate biomarkers will be required.

	Rectal cancer cohort1			Rectal cancer cohort 2			Colonic cohort		
	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**
	≤ 1 gene	≥2 genes		≤ 1 gene	≥2 genes		≤ 1 gene	≥2 genes	
Overall	31	33		30	39		45	23	
Age (years)									
Below median	13	19	0.317	14	20	0.809	22	12	1.000
Median and above	18	14		16	19		23	11	
Gender									
F	11	11	1.000	14	19	1.000	25	12	0.803
M	20	22		16	20		20	11	
Depth of invasion									
pT1 and pT2	8	12	0.086	3	15	**0.012**	2	3	0.327
pT3 and pT4	23	11		27	24		43	20	
Nodal metastasis									
No	10	22	**0.006**	14	29	**0.025**	23	18	**0.038**
Yes	21	11		16	10		22	5	
Distant metastasis									
No	24	30	0.178	26	36	0.440	33	20	0.235
Yes	7	3		4	3		12	3	

*Table 5.3. Association of hypermethylation of APC, RARB, GSTP1, CASP8 and TIMP3 with clinicopathological features. pT= pathological tumour category, *Two-tailed Fisher's exact test. Values in bold are statistically significant.*

	Rectal cancer cohort1			Rectal cancer cohort 2			Colonic cohort		
	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**	Hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3*		*P**
	≤ 1 gene	≥2 genes		≤ 1 gene	≥2 genes		≤ 1 gene	≥2 genes	
Vascular invasion									
No	20	24	0.592	17	26	0.457	21	16	0.122
Yes	11	9		13	13		24	7	
Tumour size									
Below median	6	16	**0.019**	12	21	0.332	25	9	0.305
Median and above	25	17		18	18		20	14	
Tumour grade									
Well/Moderate	28	32	0.347	30	37	0.501	38	15	0.120
Poor	3	1		0	2		7	8	

*Table 5.3 (continue). Association of hypermethylation of APC, RARB, GSTP1, CASP8 and TIMP3 with clinicopathological features. *Two-tailed Fisher's exact test. Values in bold are statistically significant.*

	Rectal cancer cohort1			Rectal cancer cohort 2			Colonic cohort		
	Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*			Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*			Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*		
	≤ 2 genes	≥3 genes	*P**	≤ 1 gene	≥2 genes	*P**	≤ 1 gene	≥2 genes	*P**
Overall	35	29		33	36		48	20	
Age									
Below median	17	14	1.000	16	18	1.000	24	10	1.000
Median and above	18	15		17	18		24	10	
Gender									
F	14	8	0.428	14	19	0.472	25	12	0.602
M	21	21		19	17		23	8	
Depth of invasion									
pT1 and pT2	11	9	1.000	6	12	0.179	2	3	0.147
pT3 and pT4	24	20		27	24		46	17	
Nodal metastasis									
No	9	23	**<0.001**	15	28	**0.007**	25	16	0.055
Yes	26	6		18	8		23	4	
Distant metastasis									
No	27	27	0.097	29	33	0.702	35	18	0.199
Yes	8	2		4	3		13	2	

*Table 5.4. Association of hypermethylation of APC, RARB, GSTP1, CASP8, TIMP3, CXCL12 and DAPK1 with clinicopathological features. pT= pathological tumour category, *Two-tailed Fisher's exact test. Values in bold are statistically significant.*

	Rectal cancer cohort1			Rectal cancer cohort 2			Colonic cohort		
	Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*			Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*			Hypermethylation of *APC, RARB, GSTP1 CASP8, TIMP3, CXCL12* and *DAPK1*		
	≤ 2 genes	≥3 genes	*P**	≤ 1 gene	≥2 genes	*P**	≤ 1 gene	≥2 genes	*P**
Vascular invasion									
No	24	20	1.000	20	23	0.808	23	14	0.115
Yes	11	9		13	13		25	6	
Tumour size									
Below median	11	18	**0.023**	14	19	0.472	27	7	0.183
Median and above	24	11		19	17		21	13	
Tumour grade									
Well/Moderate	32	28	0.620	32	35	1.000	40	13	0.116
Poor	3	1		1	1		8	7	

Table 5.4 (continue). Association of hypermethylation of *APC, RARB, GSTP1, CASP8, TIMP3, CXCL12* and *DAPK1* with clinicopathological features.
*Two-tailed Fisher's exact test. Values in bold are statistically significant.

5.3 Discussion

This study has identified specific methylation changes between rectal cancers and matched adjacent tissues. Consistent with the results in Chapter 3, the methylation levels of *APC, CDH13, CHFR, RARB* and *ESR1*, examined in this chapter with pyrosequencing, were higher in rectal cancer compared to adjacent mucosa. In addition, five other loci namely, *CXCL12, DAPK1, UNC5C, MINT3* and *MINT17* showed cancer-specific hypermethylation. The methylation level of *LINE-1* which is a marker of global methylation, was lower in rectal cancer compared to adjacent tissue and this is consistent with the current literature[213, 222].

When the methylation profiles between rectal and colonic cancer were compared, only minor differences were found. Our results could not be directly compared with published reports suggesting that proximal colon cancers had higher frequencies of methylated loci compared to distal colorectal cancers. The hypothesis from these studies is that methylation changes within the colorectum are related to their embryological origin. The caecum up to the splenic flexure are derived from the midgut whereas the descending colon to the rectum are derived from the hindgut. In this study, the colorectal tract was divided into colon and rectum on the basis of their different treatment modalities rather than their embryological differences. The purpose for such a division is to identify molecular markers that can guide choice of surgery. Furthermore, the panel of genes investigated in this study were different from those studies.

Even though a different methylation platform was utilised compared to the one used in Chapter 3, the results were similar. Specifically, this study also found that higher methylation levels of individual genes; *RARB, DAPK1, CXCL12* and *CDH13* were associated with favourable histopathological features in rectal cancers. This pattern of methylation was also observed in colonic cancer, whereby higher levels of *RARB* and *CDH13* methylation were seen in tumours without lymph node or distant metastasis. The unique methylation signature of concomitant hypermethylation of *APC, RARB, GSTP1, CASP8* and *TIMP3,* identified previously, was reproducible in this the first rectal cancer cohort which included the same 51 cancers that were examined in Chapter 3. This result was successfully validated in another independent cohort of rectal cancers. Interestingly, this methylation signature as a prognostic marker was also applicable to colonic cancers.

The addition of *DAPK1* and *CXCL12* into the existing methylation signature showed a stronger association with nodal disease in both the rectal cancer cohorts but not in colonic cancers. In addition, the new signature incorporating these two genes improved the overall specificity in identifying nodal metastasis in rectal cancer while preserving its sensitivity. This provided us with optimism that incorporating new candidate markers into the panel could improve the detection rates of rectal cancers with occult nodal metastasis.

While it has been shown that site-specific methylation changes are associated with favourable histopathological indices of disease progression, candidate approaches such as MS-MLPA and bisulphite pyrosequencing could not interrogate global

methylation changes that occur during cancer progression. The ability to map out genome-wide methylation changes at different stages of colorectal cancer progression could define the epigenetic landscape, enabling novel diagnostic, predictive and prognostic biomarkers to be identified. This will be investigated in the Chapter 7.

Results

Chapter 6

Association of protein expression with methylation changes and clinicopathological features

6.1 Introduction

Promoter CpG island methylation is a known epigenetic modification associated with gene silencing[136]. This phenomenon has been extensively examined and validated in many *in vitro* studies[239-242] but only a few have evaluated the relationship between DNA methylation and protein expression in patients with cancer[243, 244]. While pyrosequencing and MS-MLPA have shown promise as platforms that can be introduced into the clinical setting, immunohistochemistry (IHC) as a method to detect protein expression has been in routine use in most clinical laboratories and will remain so in the foreseeable future. If a relationship exists between DNA methylation and protein expression, IHC can then be used as a surrogate marker of methylation changes and this would have an immediate clinical utility. To the best of our knowledge, such a relationship has not been evaluated in patients with rectal cancer. The aims of this chapter are, therefore:

a) To evaluate the relationship between DNA methylation and protein expression of candidate genes that have been investigated in the previous chapters.
b) To determine the association between protein expression changes and clinicopathological features.

6.2 Results

6.2.1 Demographics and clinicopathological characteristics of samples used

A subset of 64 formalin-fixed paraffin-embedded (FFPE) rectal adenocarcinoma tissues, examined in the previous chapter was immunostained for APC, DAPK1, CXCL12, RARB, UNC5C, E-cadherin, ESR-α, CDH13 and CHFR protein expression. Only 40 of these slides contained adequate adjacent normal rectal cells for comparison. Patient demographics and clinicopathological characteristics are summarised in Table 6.1.

Clinicopathological features	Rectal cancer
Age (years)*	71 (33-89)
Sex	
Male	39
Female	25
Dukes' stage	
A	16
B	18
C	22
D	8
Tumour size (mm)*	45 (13-110)
Vascular invasion	
No	40
Yes	24
Degree of differentiation	
Well	7
Moderate	54
Poor	3

*Table 6.1. Demographics and clinicopathological characteristics of samples used. *Values are median (range).*

6.2.2 Optimisation of antibodies

The choice of positive controls, antigen retrieval conditions, initial concentrations of primary antibodies and duration of incubation were guided by previously published reports for these antibodies[219, 245-248]. These parameters were titrated specifically for rectal tissue so that good staining with minimal background was achieved. The results are summarised in Table 6.2. The CHFR antibody failed optimisation and was therefore excluded from analysis.

Antibody	Positive control	Antigen retrieval condition	Optimal concentration	Antibody incubation period (mins)
APC	Breast	High pressure, low pH	1 in 1000	60
DAPK1	Small intestine	High pH	1 in 50	60
CXCL12	Adjacent rectal epithelium	Low pH	1 in 50	60
RARB	Breast	Low pH	1 in 50	60
UNC5C	Pancreas	Low pH	1 in 50	60
E-cadherin	Adjacent rectal epithelium	High pH	1 in 2000	30
ESR-α	Breast	High pH	1 in 50	30
CDH13	Endothelial cells of blood vessels in liver	Low pH	1 in 1000	60

Table 6.2. The positive controls, antigen retrieval conditions, optimal antibody concentrations and duration of antibody incubation for APC, DAPK1, CXCL12, RARB, UNC5C, E-cadherin, ESR-α and CDH13 antibodies.

6.2.3 Comparison of protein expression in rectal cancer and adjacent mucosa

Difference in staining patterns between rectal cancer and adjacent normal epithelial cells was observed for APC, DAPK1, CXCL12, RARB and UNC5C and are summarised in Tables 6.4 and 6.5. Examples of the staining intensities for each antibody are illustrated in Figure 6.1. Strong cytoplasmic APC immunoreactivity was detected in

normal epithelial cells but was predominantly weak or absent (95.3%) in cancer cells. All normal epithelial cells showed strong cytoplasmic DAPK1 immunostaining but in cancer, staining intensity was generally reduced, with 9.4%, 46.9% and 25.0% of rectal cancers showing absent, weak and moderate staining intensity, respectively. For CXCL12, strong immunostaining was only detected in the cytoplasm of apical cells of normal epithelium. Over two thirds of rectal cancer samples showed weak or absent cytoplasmic CXCL12 staining while twenty rectal cancer (31.3%) samples showed moderate to strong CXCL12 immunostaining. RARB was either weakly or not expressed in the cytoplasm and nuclear compartments of adjacent normal epithelium and in approximately two-thirds of rectal cancer. UNC5C cytoplasmic staining was absent in adjacent epithelial cells and the majority of cancer samples (87.5%) showed weak to absent UNC5C staining. Strong membranous E-cadherin staining was detected in both adjacent epithelial and cancer cells while ESR-α and CDH13 staining was absent in both normal and cancer cell types.

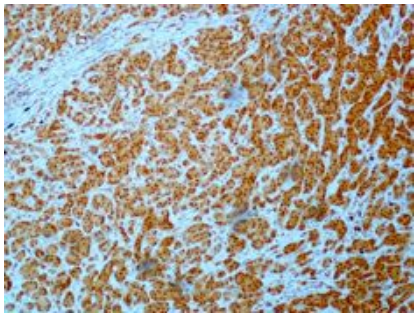
Samples	Staining intensity															
	APC				DAPK1				CXCL12				RARB			
	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong
Adjacent (n=40)	2	4	1	33	0	0	0	40	0	0	10	30	31	9	0	0
Cancer (n=64)	52	9	3	0	6	30	16	12	7	37	7	13	6	34	22	2

Table 6.3. APC, DAPK1, CXCL12 and RARB protein expression in adjacent normal tissue and rectal cancer.

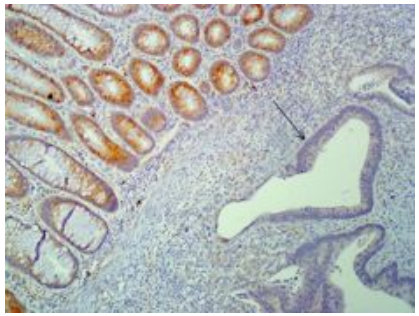
Samples	Staining intensity															
	UNC5C				E-cadherin				ESR-α				CDH13			
	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong	Absent	Weak	Moderate	Strong
Adjacent (n=40)	40	0	0	0	0	0	0	40	40	0	0	0	40	0	0	0
Cancer (n=64)	26	30	8	0	0	0	0	64	64	0	0	0	64	0	0	0

Table 6.4. UNC5C, E-cadherin, ESR-α and CDH13 protein expression in adjacent normal tissue and rectal cancer.

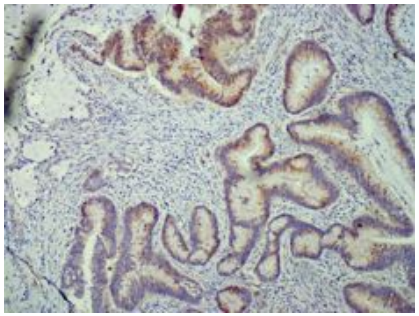
A) APC



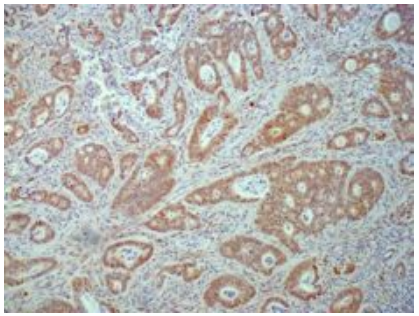
Control – breast tissue



Adjacent epithelium – Strong
Rectal cancer (arrow) – Absent

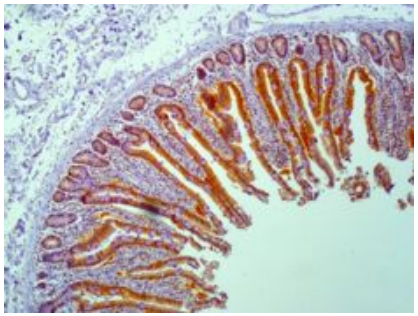


Rectal cancer – Weak

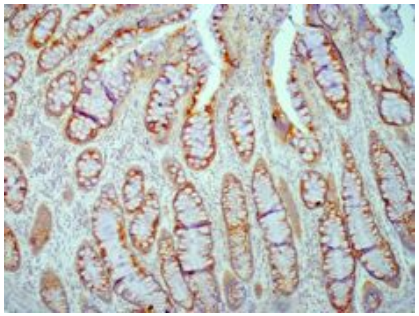


Rectal cancer – Strong

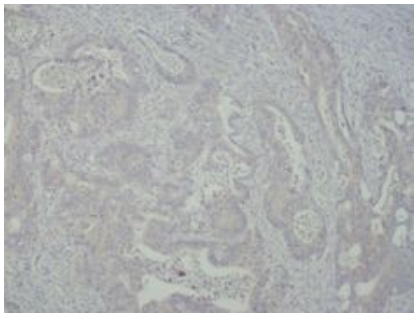
B) DAPK1



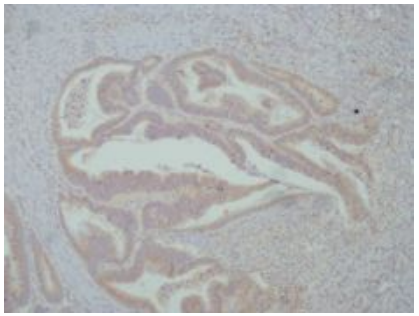
Control – small intestine



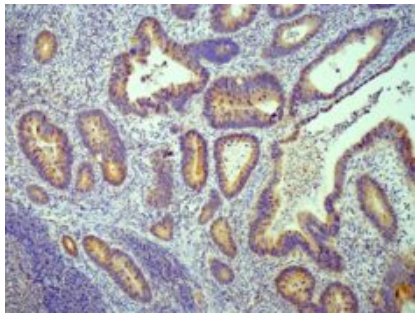
Adjacent epithelium – Strong



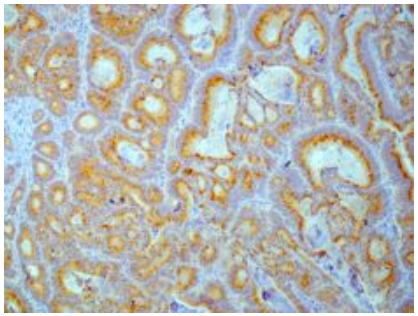
Rectal cancer – Absent



Rectal cancer – Weak

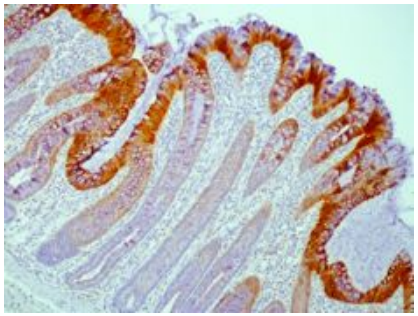


Rectal cancer – Moderate

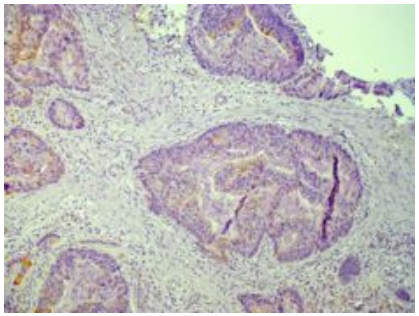


Rectal cancer - Strong

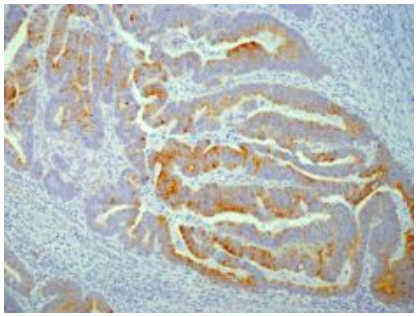
C) CXCL12



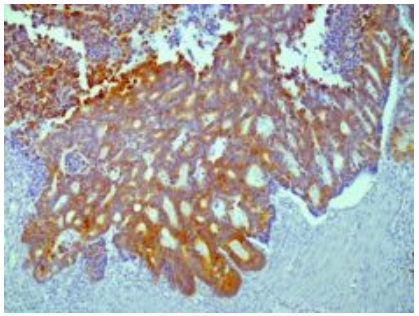
Control – adjacent epithelium



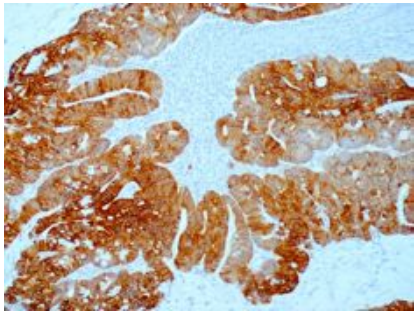
Rectal cancer – Absent



Rectal cancer – Weak

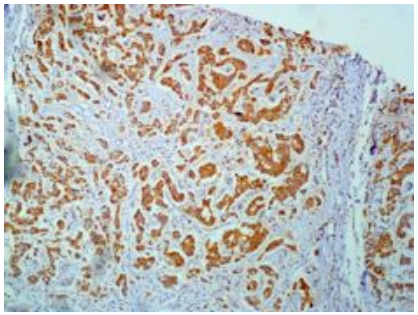


Rectal cancer – Moderate

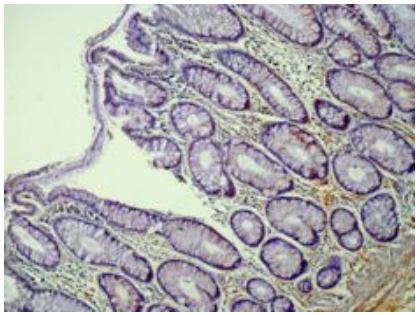


Rectal cancer – Strong

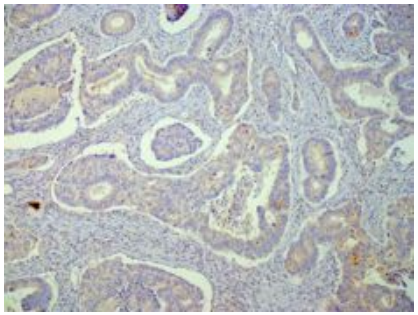
D) RARB



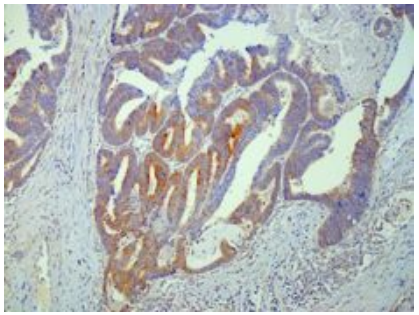
Control – breast



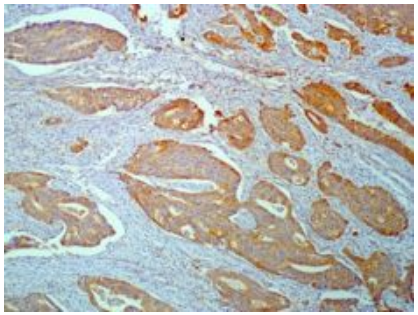
Adjacent epithelium – Absent



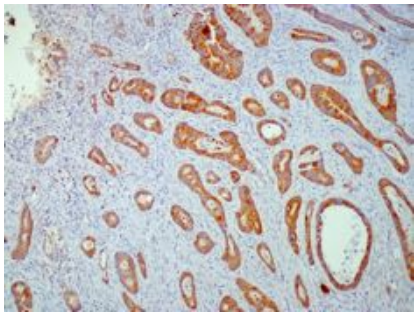
Rectal cancer – Absent



Rectal cancer – Weak

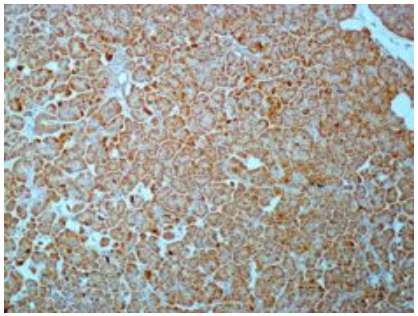


Rectal cancer – Moderate

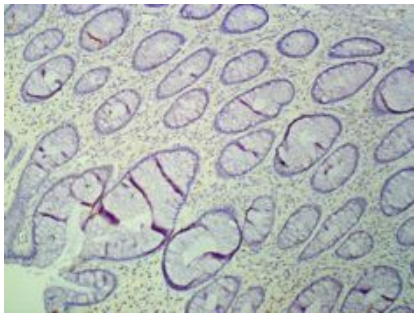


Rectal cancer – Strong

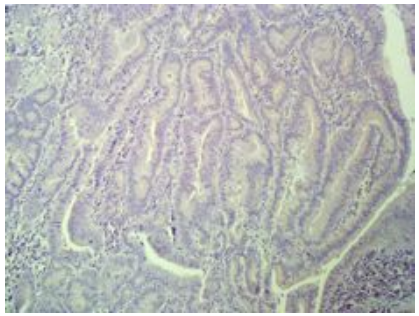
E) UNC5C



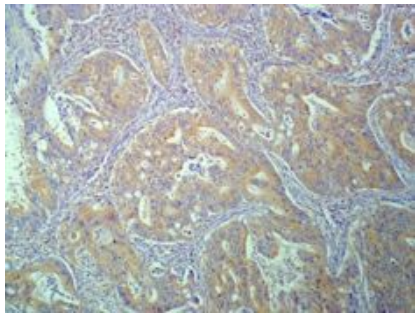
Control - pancreas



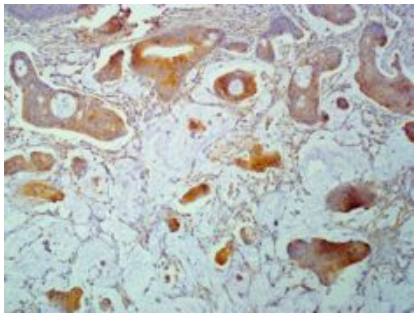
Adjacent epithelium - Absent



Rectal cancer – Absent

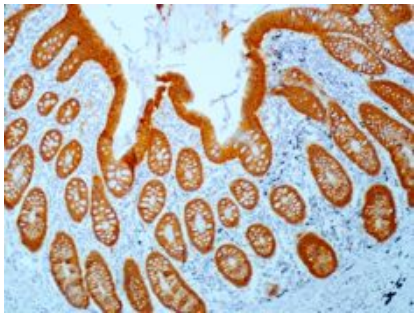


Rectal cancer – Weak

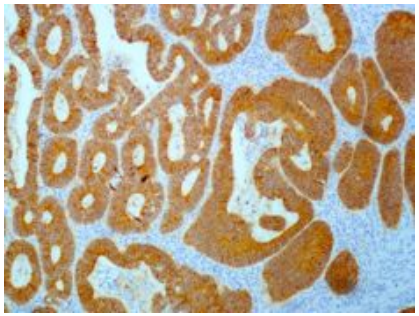


Rectal cancer – Moderate

F) E-cadherin

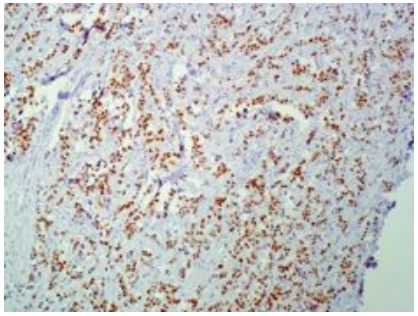


Control – Adjacent epithelium

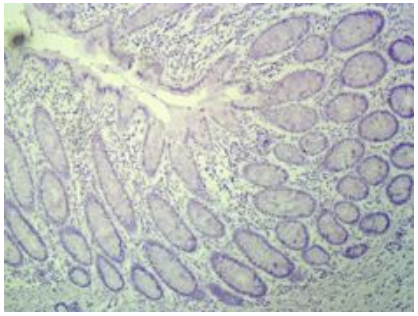


Rectal cancer – Moderate/Strong

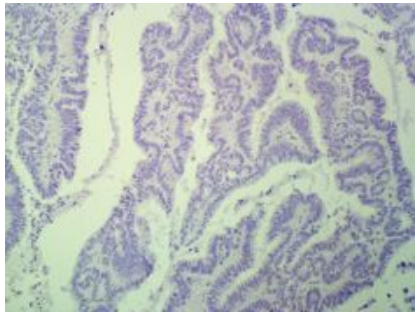
G) ESR-α



Control – breast

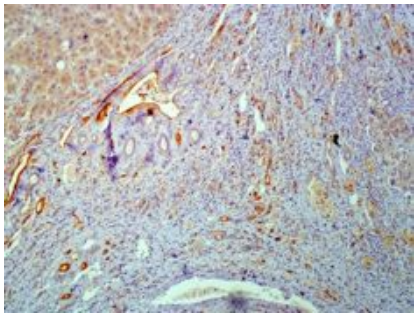


Adjacent epithelium – Absent

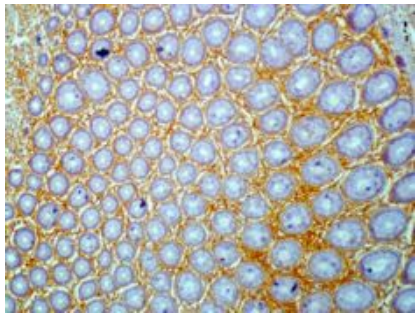


Rectal cancer – Absent

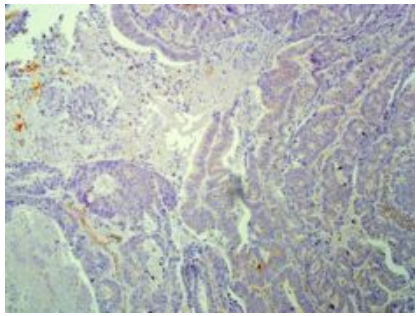
H) CDH13



Control – endothelial cells of blood vessels in liver



Adjacent epithelium – Absent



Rectal cancer – Absent

Figure 6.1. Representative staining patterns and intensities of controls, adjacent epithelial cells and cancers for APC, DAPK1, CXCL12, RARB, UNC5C, E-cadherin, ESR-α and CDH13. All magnification at x10.

6.2.4 Association of protein expression with methylation changes

To determine whether any association exists between protein expression and methylation, the protein expression of APC, CXCL12, RARB and UNC5C for the 64 rectal cancers was compared with their respective methylation profiles. ESR-α, CDH13 and E-cadherin were excluded from analysis, as all 64 rectal cancers were negative for ESR-α and CDH13 and positive for E-cadherin. Two analyses were performed: The first compared the methylation profiles of cancers with no protein staining against those stained positive (weak, moderate and strong staining samples were grouped together) and the second analysis compared the methylation profiles of cancers with either absent or weak staining against those that stained either moderate or strong. Our results did to identify any association between methylation changes and protein expression. In both analyses, the median methylation percentages of those cancers with less staining intensity (either absent in the first analysis or absent/weak in the second analysis) were generally higher than their counterparts, apart from UNC5C, but these differences were not statistically significant (Figure 6.2).

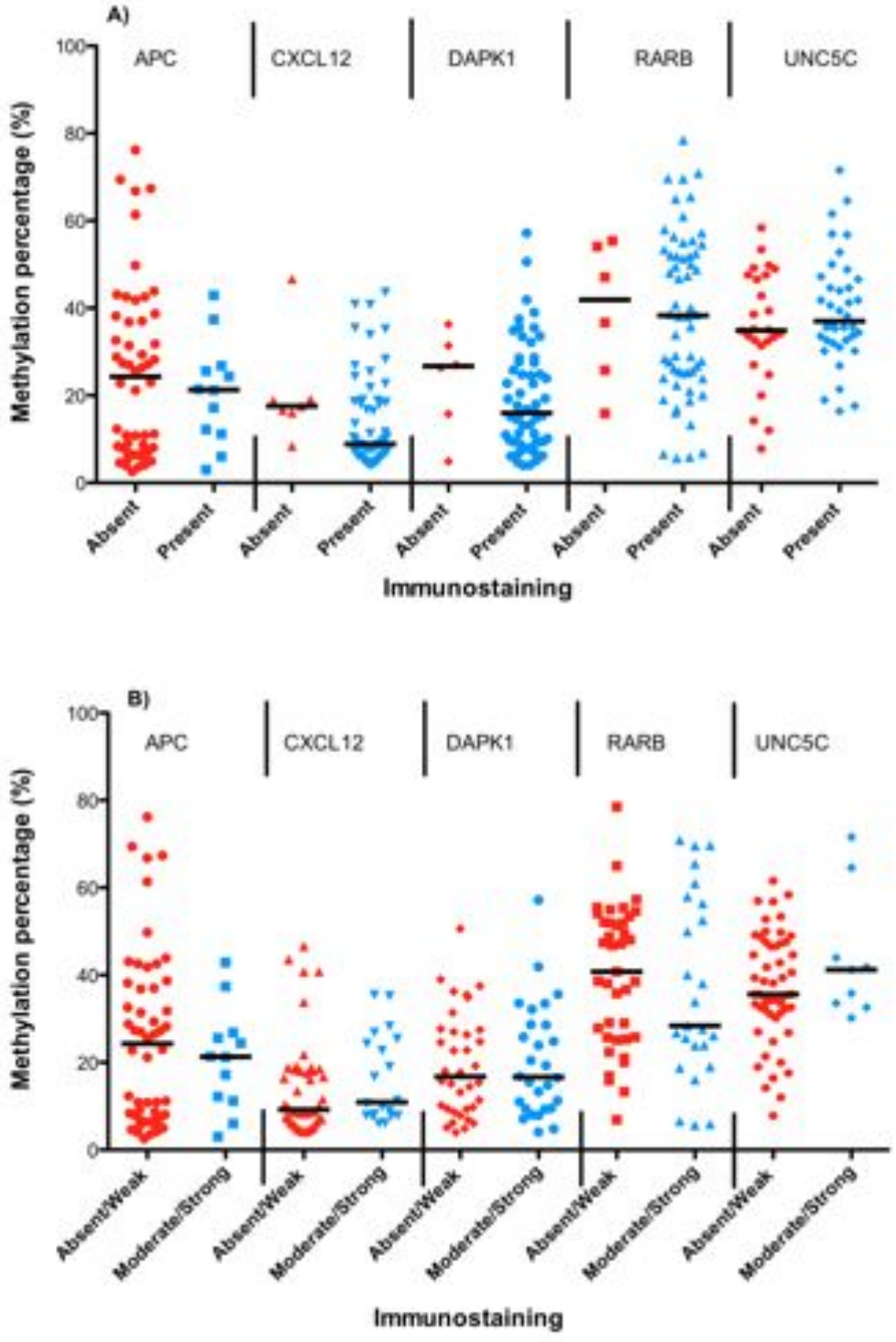


Figure 6.2. Association between methylation changes and protein expression for APC, CXCL12, DAPK1, RARB and UNC5C. A) Cancers with weak, moderate or strong staining were grouped collectively as 'Present' and compared with cancers with 'Absent' staining. B) Cancers with either absent or weak staining were compared against those with either moderate or strong protein staining. Horizontal bars represent median methylation (%). No statistical significance (P>0.05) was found between groups.

6.2.5 Association of protein expression with clinicopathological features in rectal cancer

Immunohistochemical staining patterns of APC, CDH13, CHFR, RARB, ESR-α, CXCL12, DAPK1, UNC5C and E-cadherin for the 64 rectal cancers were stratified against conventional indices of disease progression and the results are summarised in Tables 6.6 and 6.7. Only UNC5C expression was found to be significantly associated with lymph node status. The proportion of node-positive tumours with positive UNC5C expression was significantly higher than those without UNC5C expression (*P*=0.0429). In addition, there is an association between tumour grade and UNC5C and DAPK1 expression but the results should be interpreted with caution given the small number of poorly differentiated tumours in this cohort. Subgroup analysis of cancers with weak, moderate and strong expression was performed to identify staining patterns that may be associated with clinicopathological features but no significant association was found.

Clinicopathological features	UNC5C			DAPK1			Ecadherin			APC		
	Absent	Present	*P**	Absent	Present	*P**	Absent	Present	*P**	Absent	Present	*P**
Age			0.799			0.672			1.000			0.750
Median or below	14	18		4	28		0	32		25	7	
Above median	12	20		2	30		0	32		27	5	
Gender			0.123			1.000			1.000			0.747
F	7	19		2	23		0	26		22	4	
M	18	20		3	35		0	38		30	8	
Depth of invasion			0.593			0.656			1.000			1.000
pT1 and pT2	7	13		1	19		0	20		16	4	
pT3 and pT4	19	25		5	39		0	46		36	8	
Nodal metastasis			**0.043**			0.676			1.000			1.000
No	18	16		4	30		0	34		28	6	
Yes	8	22		2	28		0	30		24	6	
Distant metastasis			0.456			1.000			1.000			1.000
No	24	32		6	50		0	56		46	10	
Yes	2	6		0	8		0	8		6	2	
Lymphovascular invasion			0.794			0.177			1.000			1.000
No	16	25		2	39		0	34		33	8	
Yes	10	13		4	19		0	30		19	4	
Tumour size			1.000			1.000			1.000			0.190
Median or below	10	15		2	23		0	25		18	7	
Above median	16	23		4	35		0	39		34	5	
Degree of differentiation			**0.032**			**0.007**			1.000			0.574
Well and Moderate	36	24		6	54		0	34		49	11	
Poor	0	4		3	1		0	30		3	1	

*Table 6.5. Association of UNC5C, DAPK1, Ecadherin and APC protein expression with clinicopathological features. pT= pathological tumour category, *Two-tailed Fisher's exact test. Values in bold are statistically significant.*

Clinicopathological features	CXCL12			RARB			ESR-α			CDH13		
	Absent	Present	*P**	Absent	Present	*P**	Absent	Present	*P**	Absent	Present	*P**
Age			0.426			0.672			1.000			1.000
Median or below	2	30		4	28		32	0		32	0	
Above median	5	27		2	30		32	0		32	0	
Gender			1.000			1.000			1.000			1.000
F	3	23		2	24		26	0		26	0	
M	4	34		4	34		38	0		38	0	
Depth of invasion			0.419			1.000			1.000			1.000
pT1 and pT2	1	19		2	18		20	0		20	0	
pT3 and pT4	6	38		4	30		46	0		46	0	
Nodal metastasis			0.433			1.000			1.000			1.000
No	5	29		3	31		34	0		34	0	
Yes	2	28		3	27		30	0		30	0	
Distant metastasis			1.000			0.159			1.000			1.000
No	6	50		4	52		56	0		56	0	
Yes	1	7		2	6		8	0		8	0	
Lymphovascular invasion			0.240			1.000			1.000			1.000
No	3	38		4	37		34	0		34	0	
Yes	4	19		2	21		30	0		30	0	
Tumour size			1.000			0.671			1.000			1.000
Median or below	3	22		3	22		25	0		25	0	
Above median	4	35		3	36		39	0		39	0	
Degree of differentiation			0.378			0.332			1.000			1.000
Well and Moderate	6	54		5	55		60	0		60	0	
Poor	1	3		1	3		4	0		4	0	

*Table 6.6. Association of CXCL12, RARB, ESR-α and CDH13 protein expression with clinicopathological features. pT= pathological tumour category, *Two-tailed Fisher's exact test.*

6.3 Discussion

While methylation changes of the genes investigated in the previous chapters have been well documented in the literature, their protein expression changes, apart from APC, in colorectal cancer are less well known. We have identified cancer-specific changes in protein expression in a subset of markers investigated. The moderate to strong expression of APC, DAPK1 and CXCL12 seen in normal epithelium was reduced in rectal cancer. In contrast, RARB and UNC5C proteins were absent from normal epithelial cells but present in varying degrees in rectal cancer cells. To the best of our knowledge, we are the first group to characterise expression of these proteins specifically in rectal cancer.

No significant association between methylation and protein changes were identified in our panel of markers. This finding is not surprising given that other studies have also not been able to demonstrate such an association *in vivo*[249, 250]. This can be explained by a number of reasons. Firstly, the degree of methylation that is required to influence expression is unknown. Secondly, the sensitivity and specificity of the assays to study both methylation and protein expression are also not known. Thirdly, DNA methylation is not the only mechanism that influences gene expression and most methylation assays do not detect the presence of deletions or mutations. It is possible, however, that the association between protein expression and DNA methylation may occur in a gene-specific manner as proposed by Bueso-Ramos *et al*[244]. They investigated the relationship between protein expression and methylation

of *p73, p57* and *p53* in acute lymphocytic leukaemia and only found a significant inverse association for *p73*[244]. The combined use of genome-wide methylation and protein arrays could identify genes whose protein expressions are specifically associated with methylation in rectal cancer. In addition, using these two powerful platforms together could streamline our search for molecular markers that can predict disease progression but is beyond the scope of this thesis.

Studies that have found an association between methylation and expression have predominantly examined mRNA rather than protein expression[241, 251]. This is largely due to the fact that sensitive and semiquantitative mRNA expression assays such as reverse transcription PCR (RT-PCR) are widely available but we have deliberately avoided using assays that study RNA. This is because human tissue samples obtained from surgical procedures have been routinely formalin-fixed and paraffin-embedded (FFPE) for long-term storage in most clinical laboratories and this poses major limitations for gene expression studies for three reasons: RNA can degrade prior to fixation; formalin fixation produces significant modification to RNA; and over time, RNA continues to degrade and fragment in FFPE states. These reasons have, thus far, restricted the routine use of RNA in the clinical setting but advances in RNA extraction technology from FFPE tissues are showing promise[252, 253].

Some studies have found an inverse association between methylation and gene expression but this does not necessarily translate to a similar association between

methylation and protein expression. Many studies have failed to demonstrate a genome-wide correlation between gene and protein expression[254-256]. These authors attribute the lack of universal relationship to a number of factors: limitations of the experimental designs and assays used; biological processes such as transcriptional and post-transcriptional splicing, translational modifications, translational regulation and protein complex formation; different mRNA and protein degradation rates; and different mRNA secondary structures that might result in different protein translation efficiencies[254].

Independent of methylation, this study has identified a significant difference in the protein expression of UNC5C in node-negative and node-positive tumours, with a greater number of node-positive tumours being UNC5C positive. In addition, we have also identified an association between the expression of UNC5C and DAPK1 and the degree of tumour differentiation but the results should be interpreted with caution given the small number of poorly differentiated tumours in the cohort. IHC is a robust technique in determining protein expression but is not without its disadvantages. It relies heavily on the availability of good quality antibodies and different optimised conditions are required for various tissue types. Unlike pyrosequencing, evaluation of staining intensity is subjective and therefore, introduces intra- and inter-observer variability. Furthermore, IHC is not sufficiently sensitive to detect small changes in expression for proteins with low basal expression. These issues can be overcome with the use of an automated assessment of IHC staining software such as the OncoMark

(NovaUCD, Dublin) but its cost and its relatively novel technology still limits its use in the clinical laboratories. Overall, pyrosequencing offers greater flexibility, which is important during the discovery of prognostic and predictive molecular markers and perhaps, IHC may then have an important role once these markers have been introduced into the clinical arena.

Results

Chapter 7

Genome-wide promoter and CpG island methylation alterations in multistage colorectal tumourigenesis

7.1 Introduction

Aberrant CpG island hypermethylation in premalignant adenomas[257-259] and in aberrant crypt foci[260], the earliest precursor of colorectal cancer, is well documented, indicating that DNA methylation is not only frequent but also occurs very early during colorectal tumourigenesis. To date, the prevalence and functional significance of these events at different stages of colorectal tumourigenesis remained largely unexplored and therefore, poorly understood.

Recent advances in technology, such as DNA tiling arrays and next generation sequencing, coupled with robust bioinformatics analyses have allowed us to interrogate epigenetic events at a genome-wide level. Irizarry *et. al.*[261] provided an interesting and comprehensive demonstration of the existence of differentially methylated regions between colon cancer and non-cancerous tissue in the human genome. This has since been followed up by Oster *et. al.*[262] who identified differentially methylated patterns between normal mucosa and adenoma and between tumours with different microsatellite stability status.

Epigenomic studies have not explored the global changes in methylation patterns that occur during premalignant transformation to established cancer with subsequent progression to advanced disease. Mapping these events could increase our understanding of disease progression from premalignant adenomas as well as providing novel biomarkers of tumour behaviour and perhaps future targets for

therapy. Importantly, markers that provide prognostic information could guide surgical decision-making, particularly, in rectal cancer, which is the focus of this thesis.

With this in mind, methylation changes in four distinct groups of colorectal tissues (normal mucosa, adenoma, early and advanced cancer), representing different stages of tumourigenesis were explored using the methyl-DNA immunoprecipitation approach (MeDIP) combined with hybridisation of enriched and total DNA fragments to a promoter and CpG island array. Data analysis was performed using a modified version of the linear analysis for microarray data (LIMMA), which also incorporates an empirical Bayes model. This analysis not only provides a global overview of the methylation patterns occurring during the classical adenoma:carcinoma sequence but also methylation changes during progression of early to advanced cancer, which has not been reported.

7.1.1 Methyl-DNA immunoprecipitation (MeDIP) hybridised to promoter and CpG island array

The NimbleGen Human DNA Methylation 385K RefSeq Promoter Array, designed based on the HG18 genome release was used. The array contains 385 019 probes with lengths of between 50-75 bp spaced in 100 bp steps along the upstream promoter regions of well-characterized RefSeq genes. Genomic DNA is extracted, purified and sonicated into 300-1200 bp fragments. DNA fragments are denatured

and incubated with monoclonal mouse anti-5-methyl-cytidine antibody before immunoprecipitation with magnetic beads conjugated to anti-mouse-IgG. A fraction of the input DNA obtained after sonication is labelled with cyanine-3 (Cy3; green) while the methylated DNA, enriched following immunoprecipitation, is labelled with cyanine-5 (Cy3; red). The labelled DNA samples are co-hybridised on the array. Signal intensity data is extracted from the scanned images using NimbleScan. Each feature on the array has a corresponding scaled $\log_2$-ratio, which is the ratio of the signal intensities for the enriched and input samples that were co-hybridised to the array. The $\log_2$-ratio was computed and scaled to centre the ratio data around zero. Significant differences between log2-ratio values between groups were identified using Linear Models for Microarray Data (LIMMA) with cutoff thresholds of Benjamini and Hockberg adjusted *P* value < 0.05 and fold change > 1.5.

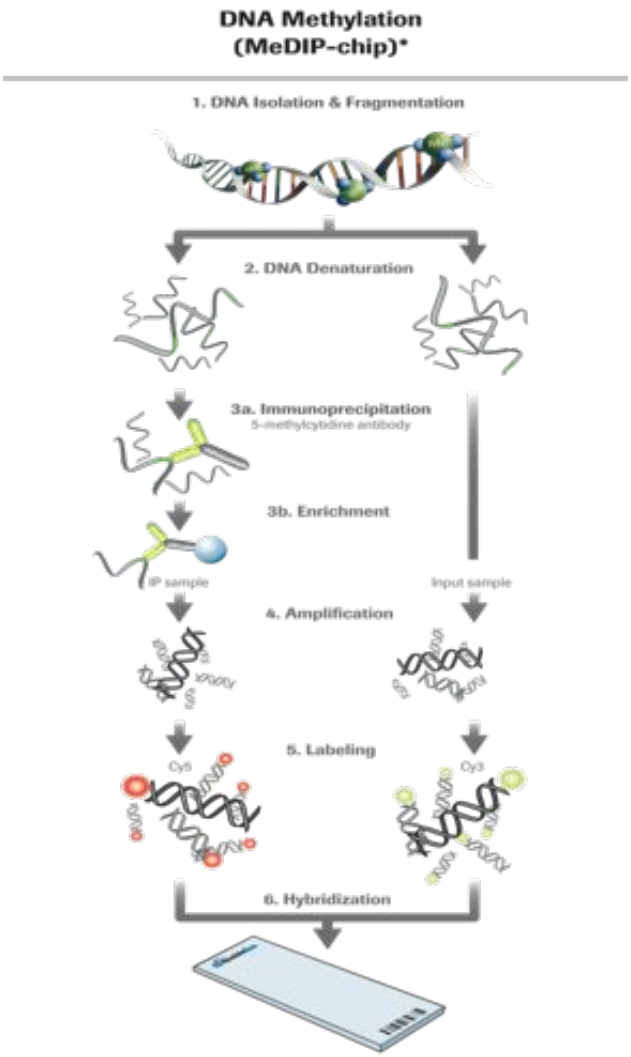


Figure 7.1. Schematic diagram of the MeDIP procedure followed by array hybridisation. Taken from http://www.nimblegen.com/products/lit/methylation/index.html.

7.2 Results

7.2.1 Demographics and clinicopathological characteristics of samples used

A subset of 20 fresh frozen rectal adenocarcinoma tissues, consisting of 10 early (T1-2N0M0) and 10 advanced cancers (TxN1M0), examined in the previous chapters were selected. 10 adenomatous tissues that were considered high-risk of progression to cancer i.e. the possession of any of these features: villous morphology, moderate to high grade dysplasia, diameter greater than 1 cm and the presence of more than 3 adenomas in the colon at the time of investigation were selected. 10 normal controls obtained from patients who underwent normal colonoscopy or flexible sigmoidoscopy for a range of non-specific symptoms were also included. Validation of a subset of potentially discriminatory genes by pyrosequencing was performed in an independent cohort of 64 radically excised rectal cancers. The clinical and pathological data of patients are summarised in Table 7.1.

Clinicopathological features	Screening set				Validation set
	Normal	Adenoma	Early Cancer	Advanced Cancer	Cancer
Male:Female	4:6	3:7	6:4	6:4	42:22
Age (range)	69.5 (48-82)	71.5 (61-87)	73.5 (51-86)	71.5 (56-86)	71.5 (33-89)
High risk features					
>1 cm in diameter only		0			
>3 polyps only		1			
Villous morphology only		0			
Moderate/High grade dysplasia only		1			
>1 of the above features		8			
Dukes' Stage					
A			10	0	10
B			0	0	22
C			0	10	22
D			0	0	10
Tumour size (mm)*			28 (13-60)	50 (35-55)	40 (13-110)
Vascular invasion					
No			9	5	44
Yes			1	5	20
Differentiation					
Well			4	2	7
Moderate			6	8	53
Poor			0	0	4

Table 7.1. Demographics and clinicopathological characteristics of samples used.
**Values are median (range).*

7.2.2 Correlation of differentially methylated loci with bisulphite pyrosequencing

Ten loci showing differential methylation between groups; five between tumour and normal and five between early and advanced cancer were selected for pyrosequencing analysis to verify whether that the log2 ratio used in the array was reflective of the samples' methylation levels. Therefore, the tumour samples used for the Nimblegen methylation array were reanalyzed by pyrosequencing. A total of 200 paired analyses were performed. There was a moderate positive correlation between log2 ratio values and methylation percentages as quantified by pyrosequencing, Spearman rank correlation coefficient, r=0.518 (95% confidence interval 0.410-0.611, *P*<0.001)(Figure 7.2). While accepting that only a fraction of the number of differentially methylated loci identified in the array underwent pyrosequencing analysis, the correlation provided reasonable confidence that the methylation array results are interpretable.

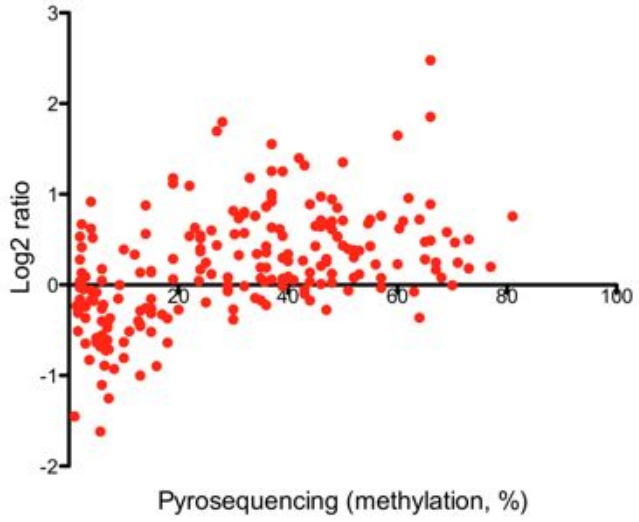


Figure 7.2. Correlation between signal intensities, expressed as Log2 ratio and methylation percentage (%) by pyrosequencing. Spearman correlation coefficient, r=0.518 (P<0.001).

7.2.3 Genome-wide methylation changes from normal mucosa to adenoma, early and advanced colorectal cancer

a) Normal mucosa vs. adenoma

208 loci showed higher methylation levels in adenoma compared to normal mucosa. Having excluded redundancy, i.e. multiple hypermethylated loci encoded by the same gene and genes containing both hyper- and hypomethylated sites, 148 genes were found to be hypermethylated in adenoma compared to normal mucosa. In contrast, only one locus, encoding the *IGF2* gene, was found to have higher methylation levels in normal mucosa than in adenoma, suggesting that DNA hypomethylation is a rare event at the pre-malignant stage.

b) Normal mucosa vs. early cancer

In contrast to adenomas, there was a significantly larger number of loci hypermethylated in early cancer when compared to normal mucosa. A total of 1926 loci, which translated to 807 genes, had higher methylation levels in early cancer compared to normal mucosa. In contrast, 33516 loci, representing 9974 genes, had higher methylation levels in normal mucosa compared to early cancer.

c) Normal mucosa vs. advanced cancer

Compared to adenoma and early cancer, advanced cancer had the largest number of hypermethylated loci and genes relative to normal mucosa. A total of 19041 loci, representing 5203 genes had higher methylation levels in advanced cancer compared

to normal mucosa. In comparison, 31053 loci, representing 6250 genes showed higher methylation levels in normal mucosa than in advanced cancer.

d) Normal mucosa vs. adenoma and cancer

Our study has identified two intriguing methylation patterns in colorectal tumourigenesis. Firstly, there is a global accumulation of aberrantly methylated loci at different stages of tumourigenesis, already occurring in pre-malignant disease and increasing sharply in early carcinoma and subsequently in advanced disease. Of the 148 genes that were hypermethylated in adenoma relative to normal mucosa, 50 of those were also hypermethylated in early and advanced cancer. Genes that are hypermethylated from the outset of premalignant adenoma and remained so in latter stages of cancer could potentially be important predictive biomarkers of disease progression (Figure 7.3). Secondly, DNA hypomethylation (with reference to normal mucosa) occurs at a later stage of tumourigenesis but its prevalence in established cancer is far greater than that of DNA hypermethylation. The majority of genes that were hypomethylated in cancer were present in both early and advanced stages.

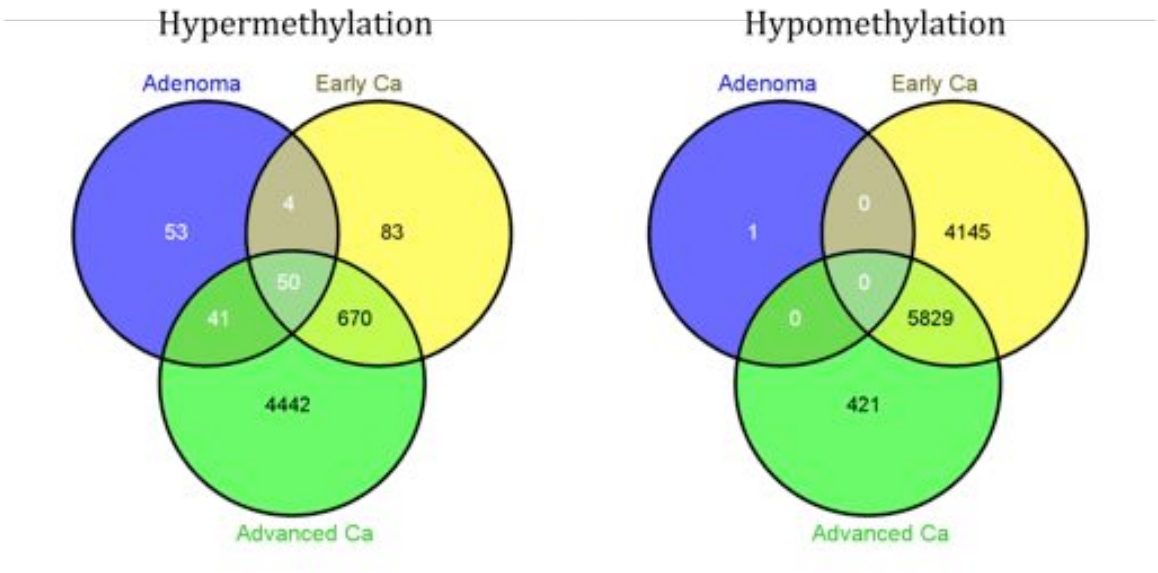


Figure 7.3. Venn diagrams showing the number of genes hypermethylated (left) and hypomethylated (right) in adenoma, early and advanced cancer relative to normal mucosa. Early Ca= Early cancer, Advanced Ca= Advanced Cancer.

d) Early vs. advanced cancer

A genome-wide search for differentially methylated genes between early and advanced cancers could provide novel prognostic markers that complement candidate genes already identified in previous chapters. 16435 loci were found differentially methylated between the two cancer groups, of which 16299 had higher methylation levels in advanced cancer and 136 loci with higher levels in early cancer. They represented 7103 and 116 genes respectively. 55 genes were found to contain both hyper- and hypomethylated loci and were therefore excluded, leaving 7049 genes being hypermethylated exclusively in advanced cancer and 63 genes hypermethylated in early cancer. The top 20 ranking genes for each cancer group are shown in Table 7.2.

Higher levels in advanced cancer	Higher levels in early cancer
C21orf56	*C1orf70*
FAM179A	*C6orf70*
FBXL16	*CHPT1*
FOXJ1	*DNMT3B*
GALNT9	*FGF23*
GALNT9	***GLI3***
GTPBP1	*INO80D*
HSP90AA1	*MARCH6*
JAK3	***MYL12A***
KRT3	*NFYC*
L1TD1	*OR4A5*
PGBD2	*POU3F3*
PTCH1	*RSPO4*
PTPRN2	*SEMG1*
RAPGEF3	***SNAP25***
RUNDC3B	***SOX7***
SNURF	***TIAM1***
TMEM106A	*TUBB2A*
TTC16	*ZBBX*
UAP1	*ZNF599*

Table 7.2. The top 20 ranking genes that have higher methylation levels in advanced cancer (left column) and in early cancer (right column). Genes in bold were reanalysed using pyrosequencing.

7.2.4 Chromosomal distribution of methylation changes

To gain a better understanding of how these methylation changes are distributed in the human genome, differentially methylated loci were mapped to individual chromosomes. Of the three tumour groups, adenoma has the lowest prevalence of hypermethylated loci. The 208 hypermethylated loci identified in adenoma were distributed equally, albeit at low frequency, among the chromosomes. Chromosome 21 contained the highest frequency of hypermethylated loci in adenoma but overall, they only accounted for <0.2% of the loci examined in the array (Figure 7.4).

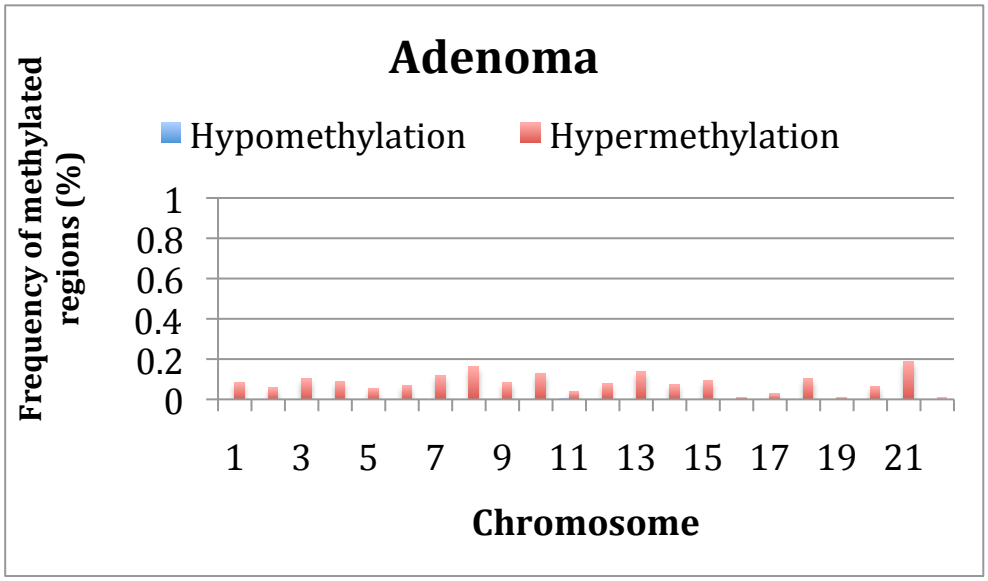


Figure 7.4. Chromosomal distribution of hypermethylated and hypomethylated loci (relative to normal mucosa) in adenoma. The frequency of methylated loci indicates the number of hyper or hypomethylated loci/the total number of array probes in the chromosome X100.

In early cancer, however, the distribution pattern was more striking. The prevalence of hypomethylated loci was far greater than that of hypermethylated loci on all chromosomes. Chromosomes 4 and 6 contained the highest frequency of hypomethylated loci in early cancer. The lowest frequency of hypomethylated loci was seen in chromosome 22 (3%). In comparison, the frequency of hypermethylated loci in early cancer was very low across all chromosomes. Less than 1% of loci in each chromosome were found to be hypermethylated in early cancer (Figure 7.5).

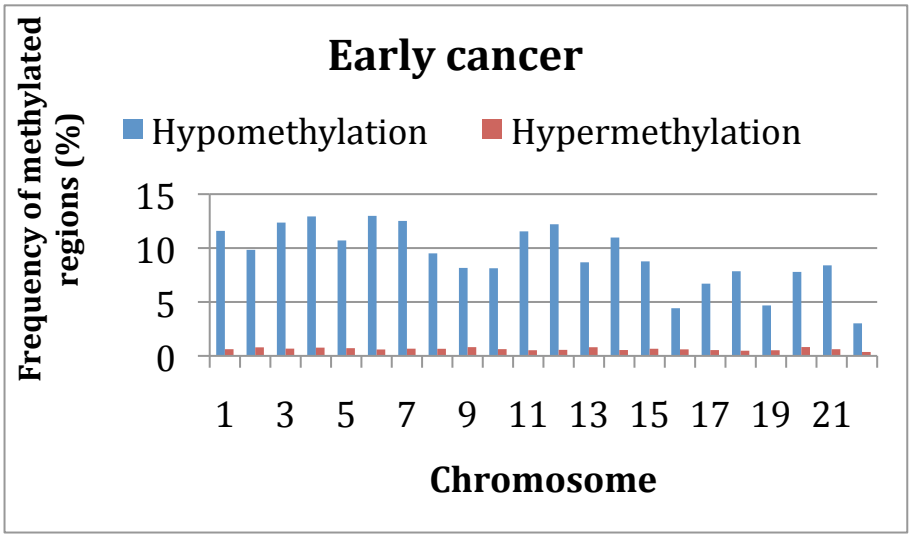


Figure 7.5. Chromosomal distribution of hypermethylated and hypomethylated loci (relative to normal) early cancer. The frequency of methylated loci indicates the number of hyper or hypomethylated loci/the total number of array probes in the chromosome X100.

In advanced cancer, there were more differentially methylated (hypo- and hypermethylated) loci compared to in adenoma and early cancer. Hypermethylation was more prevalent in some chromosomes e.g. chromosomes 1-7, 11, 12, 14 and 15 while hypomethylation was more frequent in chromosomes 16, 19, 20 and 22 (Figure 7.6).

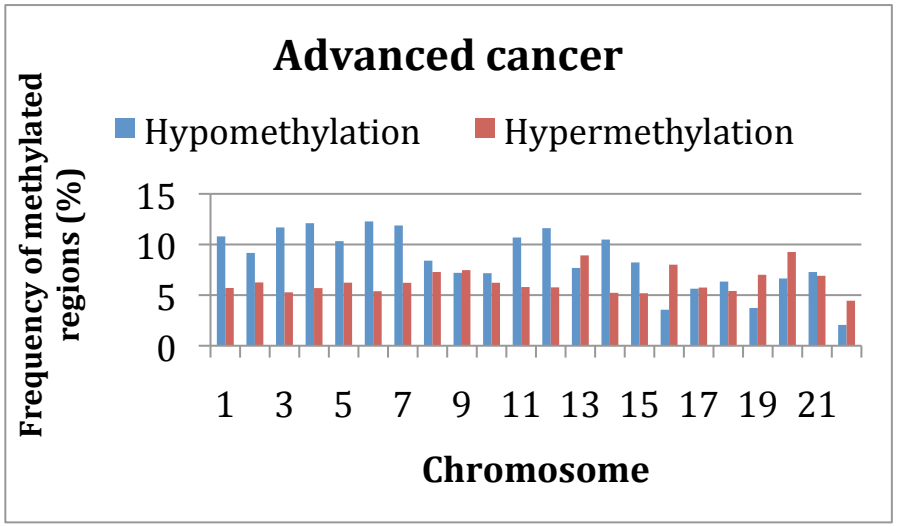


Figure 7.6. Chromosomal distribution of hypermethylated and hypomethylated loci (relative to normal mucosa) in advanced cancer. The frequency of methylated loci indicates the number of hyper or hypomethylated loci/the total number of array probes in the chromosome X100.

7.2.5 Validation of candidate genes using bisulphite pyrosequencing

Ten genes differentially methylated between early and advanced cancer were selected for validation using bisulphite pyrosequencing. The loci assessed by pyrosequencing were designed such that they either overlapped with the array probes or were contained within the same CpG island. A pragmatic approach was employed, whereby, pyrosequencing primers for these loci were either commercially available or were easily designed using the methprimer software. From the array, *SNAP25, SOX7, POU3F3, TIAM1* and *GLI3* were identified to have higher methylation in early campared to advanced cancer and were selected for validation. *TTC16, GTPBP1, TMEM106A, RAPGEF3* and *PGBD2* were selected from the array as they were determined to have higher levels in advanced compared to early cancer.

7.2.5.1 Internal validation of pyrosequencing assays

Internal validation of pyrosequencing assays were performed using fully methylated and unmethylated genomic DNA (Millipore,). All assays generally showed relatively high sensitivity in detecting low and high levels of methylation, with the *POU3F3* assay showing the least sensitivity for low levels of methylation (Tables 7.3 and 7.4). To assess precision and reproducibility of each assay, bisulfite treatment and PCR of a control DNA (fresh frozen tissue of a rectal adenocarcinoma) were performed and repeated on four different days. This was followed by pyrosequencing of each sample on five separate occasions. The results showed very little variation in the level of methylation detected at individual CpG sites for all assays; all with standard deviations of less than 4%.

Assay		Control	CpG sites								
			Site-1	Site-2	Site-3	Site-4	Site-5	Site-6	Site-7	Site-8	Mean
SNAP25	Methylation percentage	Unmethylated	15	17	14	9	17	11	-	-	13.83
(6 sites)	(%)	Methylated	82	100	100	77	100	78	-	-	89.50
SOX7 (6	Methylation percentage	Unmethylated	7	3	7	6	9	5	-	-	6.17
sites)	(%)	Methylated	83	89	86	90	91	85	-	-	87.33
POU3F3	Methylation percentage	Unmethylated	23	27	22	29	-	-	-	-	25.25
(4 sites)	(%)	Methylated	100	86	90	98	-	-	-	-	94.70
TIAM1 (8	Methylation percentage	Unmethylated	2	2	2	2	1	2	3	2	2.00
sites)	(%)	Methylated	97	98	98	98	98	88	99	100	97.00
GLI3	Methylation percentage	Unmethylated	3	4	50	7	-	-	-	-	16.00
(4 sites)	(%)	Methylated	98	98	100	91	-	-	-	-	96.75

Table 7.3. The methylation percentages for unmethylated and methylated controls of individual CpG sites for SNAP25, SOX7, POU3F3, TIAM1 and GLI3 pyrosequencing assays. The number of CpG sites evaluated for each assay is shown in parentheses.

Assay		Control	CpG sites						
			Site-1	Site-2	Site-3	Site-4	Site-5	Site-6	Mean
TTC16	Methylation percentage	Unmethylated	1	1	1	3	4	-	2.00
(5 sites)	(%)	Methylated	90	96	94	100	100	-	96.00
RAPGEF3	Methylation percentage	Unmethylated	1	2	2	6	3	-	2.80
(5 sites)	(%)	Methylated	89	86	90	95	93	-	90.60
PTCH1	Methylation percentage	Unmethylated	5	5	9	7	-	-	6.50
(4 sites)	(%)	Methylated	90	90	94	92	-	-	91.50
TMEM106A	Methylation percentage	Unmethylated	1	2	1	2	5	-	2.20
(5 sites)	(%)	Methylated	88	85	89	85	90	-	87.40
PGBD2	Methylation percentage	Unmethylated	1	1	2	2	4	2	2.00
(6 sites)	(%)	Methylated	90	90	93	89	86	88	89.33

Table 7.4. The methylation percentages for unmethylated and methylated controls of individual CpG sites for TTC16, RAPGEF3, PTCH1, TMEM106A and PGBD2 pyrosequencing assays. The number of CpG sites evaluated for each assay is shown in parentheses.

7.2.5.2 Methylation levels of candidate genes in the screening cohort of rectal cancers

To determine whether the results of the candidate genes selected from the array were reproducible, the 10 early and 10 advanced rectal cancers used in the array were reanalysed using bisulphite pyrosequencing. Consistent with the array findings, pyrosequencing confirms that the median methylation levels of *SNAP25, SOX7, POU3F3, TIAM1* and *GLI3* were higher in early than in advanced cancers (Figure 7.7). Pyrosequencing, however, did not show any appreciable difference in methylation for the other five genes, which the array had identified as having higher methylation levels in advanced compared to early cancer, apart from *TMEM106A* (Figure 7.7).

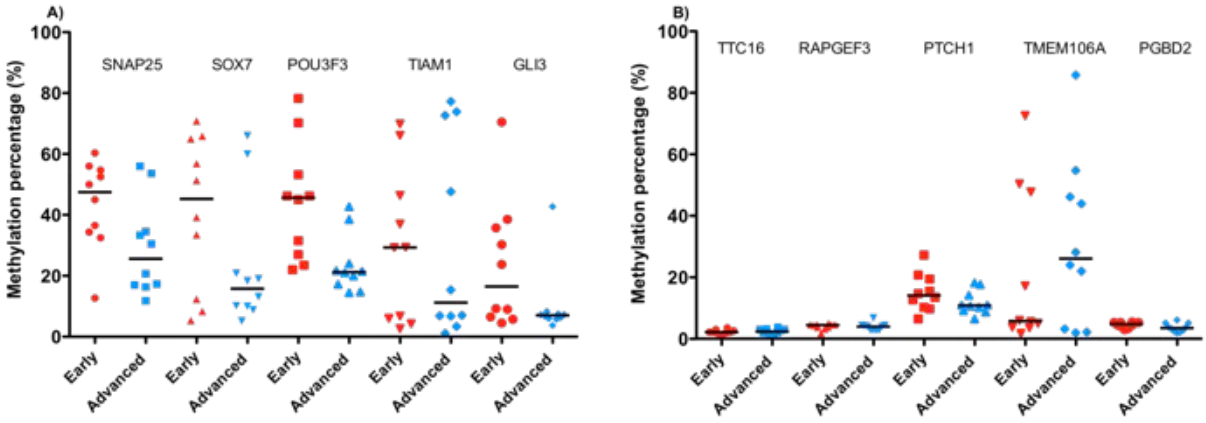


Figure 7.7. Pyrosequencing results of A) five candidate genes from the array found to have higher methylation levels in early cancer and B) five candidate genes from the array found to have higher methylation levels in advanced cancer. Bold horizontal line represents the median methylation percentage.

7.2.6 Association of candidate gene methylation with clinicopathological features

The methylation profiles of the 10 candidate genes were examined by pyrosequencing in a larger independent cohort of 64 radically excised rectal cancers and their methylation levels were associated with clinicopathological features of disease progression.

7.2.6.1 Methylation profile of the five 'early' candidate genes

Only *SNAP25* was found to have significantly higher methylation levels in early Dukes' A/B compared to advanced Dukes' C/'D' cancers (*P*=0.010). When stratified according to individual clinicopathological features, the methylation levels of *SNAP25* were higher in superficial (T1/2, *P*=0.011), node negative (N0, *P*=0.010) and non-metastatic (M0, *P*=0.006) cancers (Figure 7.8). *TIAM1* had higher methylation levels in cancers without distant metastasis (M0, *P*=0.020) while *SOX7* showed higher methylation levels in cancers without lymphovascular invasion (*P*=0.033). No other significant difference in methylation levels for the other clinicopathological features were found.

7.2.6.2 Methylation profile of the five 'advanced' candidate genes

Pyrosequencing results did not show any significant difference in methylation for *TTC16, GTPBP1, TMEM106A, RAPGEF3* and *PGBD2* when stratified according to conventional clinicopathological features. Apart from *TMEM106A*, the methylation levels of these genes were generally low in most of the cancers.

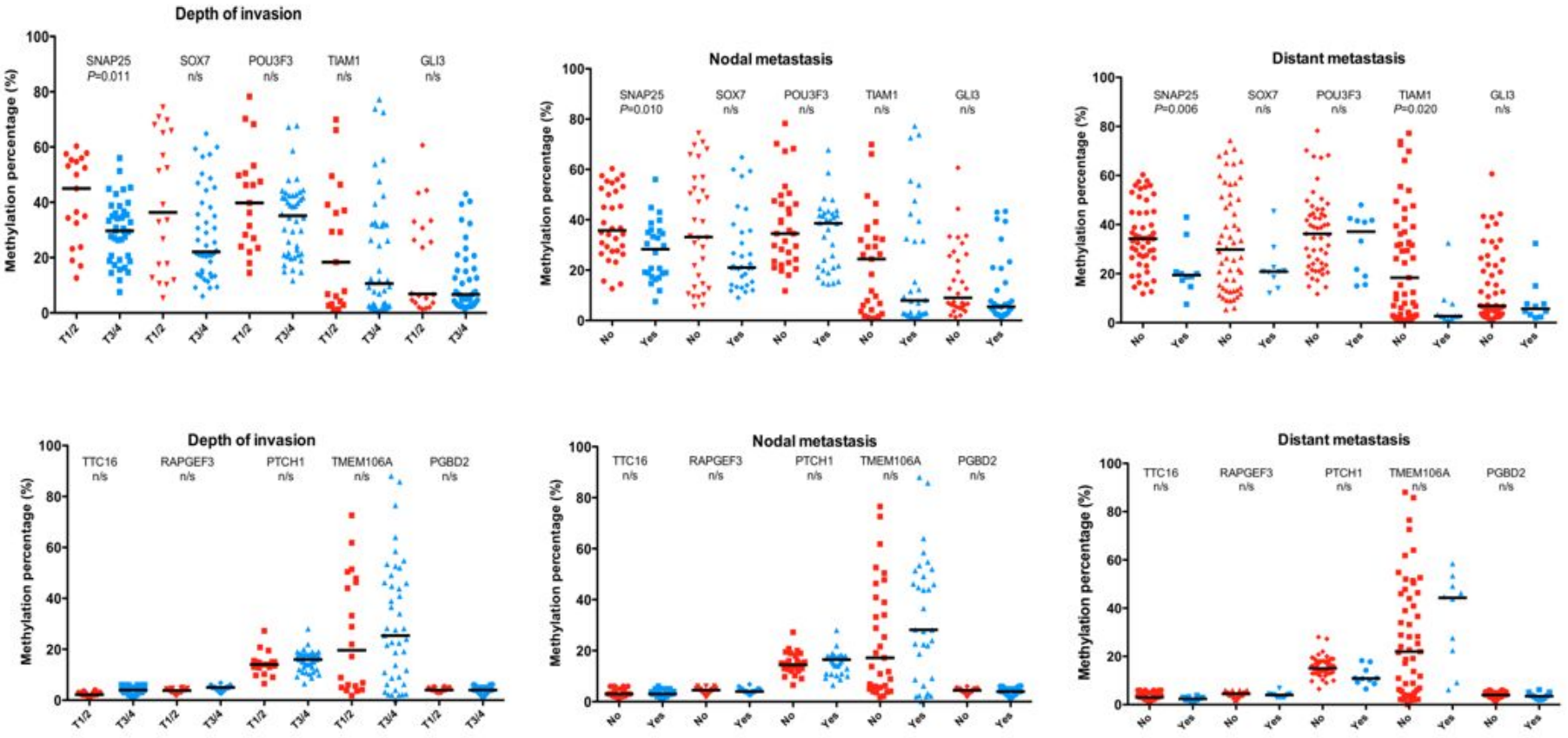


Figure 7.8. Pyrosequencing results of SNAP25, SOX7, POUF3F, TIAM1, GLI3, TTC16, RAPGEF3, PTCH1, TMEM106A and PGBD2, stratified according to depth of invasion, nodal metastasis and distant metastasis. n/s = not significant. (2-tailed Mann-Whitney U Test).

7.2.7 Discussion

The use of a genome-wide promoter and CpG island methylation array has allowed us to interrogate the global methylation changes at different stages of rectal cancer. This study has shown that aberrant DNA methylation occurs early in tumourigenesis, i.e. in premalignant adenomas, and its prevalence increases in latter stages of disease progression. Hypomethylation appeared to be a later event, occurring mainly in cancer and rather infrequently in adenoma. Its prevalence in cancer, however, is far greater than that of hypermethylation. Only until recently, global hypomethylation has been thought to occur predominantly in repetitive elements and many studies have used methylation of *LINE-1* and *Alu* as surrogate markers of global methylation[263, 264]. This study, however, has shown that hypomethylation in cancer also occurs frequently in gene promoters and CpG islands.

This study has identified a set of 50 genes that are hypermethylated in all stages of tumourigenesis. Hypermethylation of 13 of these genes has been reported in numerous epithelial cancers such as colorectal (*ADAMTS5, ALX4, CDH2, RGMA and TWIST1)*, lung (*HOXC9, PAX9, ZEB2)*, renal (*WNT3A)*, bladder (*GATA2)*, breast (*LHX5)*, pancreas (*SLC31A1)* and malignant melanoma (*TLX2)*. The methylation profile of these genes could have significant clinical utility. For example, these genes could be used as an adjunct to current bowel cancer screening programmes to improve detection rates. In addition to faecal occult blood testing, stool-extracted DNA can be tested for this methylation profile and any positive cases will be invited for screening colonoscopy. This prospect remains theoretical, as a great deal of work to develop such profile into a screening biomarker is still required.

Pyrosequencing validation of a subset of differentially methylated genes between early and advanced cancer has identified three genes (*SNAP25, TIAM1* and *SOX7)* that are associated with favourable histological indices. *SNAP25* encodes a presynaptic plasma membrane protein involved in the regulation of neurotransmitter release. Increased immunoreactivity of SNAP25 in neuroendocrine tumours such as pituitary adenomas has been described[265] but its role in other solid tumours is unknown. TIAM1 is a guanine nucleotide exchange factor for Rac and activates guanosine 5′-diphosphate–bound Rac for active GTP-bound Rac. This activates downstream pathways, leading to increased invasion and migration[266]. Overexpression of TIAM1 has been associated with a more aggressive phenotype in human colon, breast and prostate carcinomas[267-269]. More recently, Jin *et. al*.[270] found that methylation of *TIAM1* was inversely associated with expression but unlike our study, they did not find an association with any clinicopathological features. *SOX7* encodes a transcription factor that can enhance and inhibit transcription[271]. It interacts with β-catenin, mediating the degradation of active β-catenin and resulting in the suppression of β-catenin/T-cell factor–regulated transcription[272]. Aberrant hypermethylation of this gene with downregulation of its expression have been demonstrated in colorectal and prostate cancer[272, 273].

This array platform has been previously employed and validated in colon cancer and melanoma by three independent groups[206, 261, 274]. When our data was compared with that of other studies that have employed different genome-wide methylation approaches in colorectal cancer, a reasonable overlap on the number of hyper and

hypomethylated genes was found. In a study by Estecio *et. al.*[275], 78 (out of 174) genes found hypermethylated in colorectal cancer were also observed in our study. Similarly, 52 (out of 112) hypermethylated and 65 (out of 119) hypomethylated genes reported by Hinoue *et. al.*[276] were also present in our study.

Pyrosequencing validation on a subset of genes in this study, however, was less convincing. This could be explained by the fact that the loci interrogated by pyrosequencing did not overlap with the loci identified from the array, although, they lie within the same CpG island. One of the drawbacks of pyrosequencing is that it relies on the availability of suitable PCR primers. To avoid bias amplification of either methylated or unmethylated alleles during PCR, primers are designed in regions not affected by bisulphite treatment, i.e. in CpG-poor sequences. This poses a problem in long stretches of CpG-rich regions, such as those found in the array. As a compromise, primers designed in this study were located about 100-200 bp away from the loci of interest. It is possible, that given more time, primers and PCR conditions, such as a nested-approach, could be optimised to interrogate the loci of interest. In addition, the pre-processing normalisation performed by Nimblegen may not be optimal for this dataset, leading to a number of genes that failed validation. The scaling normalisation method employed by Nimblegen does not consider any region or intensity-dependent effects and therefore may not sufficiently filter out background noise. Some of the more sophisticated normalisation techniques such as the Lowess or Quantile methods may be more appropriate and may yield better discrimination.

Chapter 8

Discussion

8.1 Discussion

Aberrant DNA methylation is a frequent event in rectal cancer. This has been shown in this study using three different methylation platforms; MS-MLPA, bisulphite pyrosequencing and the Nimblegen CpG island and promoter methylation array. In addition, a unique methylation signature was found to be associated with early stage disease. This signature was reproducible in an independent cohort of rectal cancer. Genome-wide methylation analysis identified over 7000 differentially methylated genes between early and advanced rectal cancer. Validation work confirmed that the methylation of a subset of these genes was associated with favourable histological indices. This study also found that tissue enrichment techniques did not produce any significant improvement in methylation yield.

Using MS-MLPA and bisulphite pyrosequencing, this study found an association between site-specific hypermethylation and favourable histopathological stage. The methylation levels of specific CpGs in *GSTP1, CXCL12, DAPK1* and *CDH13* were higher in rectal cancers without nodal metastasis while higher methylation levels of *RARB* and *CDH13* were observed in rectal cancers without distant metastasis. A novel prognostic methylation signature for rectal cancer was identified using MS-MLPA, whereby concomitant hypermethylation of two or more of *APC, RARB, GSTP1, TIMP3* and *CASP8* was associated with histopathologically localised disease (TxN0M0). Extensive validation work demonstrated that this prognostic signature was reproducible using bisulphite pyrosequencing not only in the same rectal cancer cohort but also in two independent cohorts of rectal and colonic cancers. The

addition of *DAPK1* and *CXCL12* into the existing methylation signature improved its prognostic significance in nodal disease for rectal cancer but not for colonic cancers. The new methylation signature consisting of concomitant hypermethylation of three or more of *APC*, *RARB, GSTP1, TIMP3, CASP8, DAPK1* and *CXCL12* has an improved overall specificity in identifying nodal metastasis in rectal cancer while preserving its sensitivity.

These results are consistent with studies that have correlated methylation in single or multiple genes with clinical stage and outcome[277, 278]. Toyota *et al.* first pioneered the concept of the 'CpG island methylation phenotype' (CIMP), originally referring to a subgroup of colorectal cancers with multiple methylated loci possessing distinct clinicopathologic and molecular features such as association with proximal tumour location and methylation of *p16, THBS1* and *hMLH1*[177]. Subsequent large population-based studies in colon cancer confirmed the existence of this phenotype and in addition, showed CIMP positive tumours to be associated with *KRAS* and *BRAF* mutations and older age[179, 181]. When stratified according to microsatellite instability status (MSI), CIMP positive tumours that were also microsatellite stable had a shorter five-year survival. One of the main criticisms of these studies is that methylation was measured qualitatively, giving rise to positive results in tumours with low or insignificant levels of methylation[180]. Using a quantitative methylation approach, Ogino *et al*[180] showed that CIMP was associated with lower colon cancer-specific mortality, independent of other molecular changes. It is worthy to note that CIMP is a concept and there is no global consensus on the genes that make up the CIMP phenotype. With advances in molecular technology, it is envisaged that more

novel methylation markers will be discovered and they could outperform the initial set of loci that was first used to describe the CIMP phenotype.

In Chapter 3, the MS-MLPA assay was selected as it has been optimised to simultaneously quantify methylation levels of multiple genes in a single run. This provided a useful insight into the methylation patterns of rectal cancer and allowed exploratory analysis to be performed. MS-MLPA is not without drawbacks. Firstly, the MS-MLPA kit used contained a predetermined set of tumour suppressor genes, of which only a handful were relevant in rectal cancer. There is, however, an option for customising probes of interest into a single kit, which could prove useful in the future when sensitive and specific prognostic genes have been identified. Secondly, MS-MLPA relies on the methylation sensitive endonuclease, *HhaI*, which only recognises the 5'-GCGC-3' sequence, thereby limiting the number of CpG sites within an island that can be examined. Thirdly, MS-MLPA does not provide a global methylation landscape, which is useful in the diagnostic, prognostic and therapeutic aspects of tumourigenesis.

Brakensiek *et. al.* found specific CpG sites within an island with high discriminatory significance in myelodysplastic syndrome and myeloid leukemia[279]. This highlights the importance of the analysis of multiple CpG sites within a locus, which is possible using either bisulphite pyrosequencing or bisulphite genomic sequencing but not with other methylation platforms such as MSP, COBRA or MS-MLPA. Bisulphite genomic sequencing requires cloning of PCR products prior to sequencing and is, therefore, very labour intensive and impractical in the clinical setting. To overcome

these limitations, bisulphite pyrosequencing was used instead. There are abundant reports that have demonstrated the reliability of pyrosequencing for the accurate quantification of methylation levels at a single nucleotide level[280-282] and shown very good correlation with other methylation platforms such as PCR/LDR/Universal array assay[283], COBRA[197], SNaPmeth[284] and mass spectrometry[285]. Indeed, our experience with pyrosequencing has also shown good correlation with MS-MLPA. Variations in methylation levels among consecutive CpGs were found in our study, a result consistent with the published literature[279, 286, 287]. These differences were exploited in this study to optimise single site testing, similar to that offered by MS-MLPA.

A genome-wide assessment of methylation in colorectal tumourigenesis was carried out using the Nimblegen CpG island and promoter methylation array to expand our understanding of the global methylation patterns that occur during progression from normal mucosa to adenoma and subsequently to adenocarcinoma. The experiments found two intriguing methylation changes in colorectal tumourigenesis. Firstly, there is a global accumulation of aberrantly methylated loci at different stages of tumourigenesis, already occurring in pre-malignant adenoma and increasing sharply in early carcinoma and subsequently in advanced disease. Secondly, DNA hypomethylation is rare in adenoma but its prevalence in established cancer is far greater than that of DNA hypermethylation.

DNA hypomethylation in colorectal cancer was first described by Feinberg and Vogelstein[131]. The proposed mechanisms by which global hypomethylation drives tumourigenesis are thought to involve the generation of chromosomal instability,

reactivation of transposable elements and loss of imprinting[132]. Many researchers, past and present, have investigated methylation of *LINE-1* and *ALU* as surrogate markers of global methylation[288-290]. New technology has now allowed us to move away from surrogate markers of global methylation but instead quantitatively interrogate methylation of individual loci at a genome-wide level. Despite this, there has been very limited evidence in the literature reporting on the global methylation changes during colorectal tumourigenesis[262], making these findings unique. There is consistency between these results and those of others but the application of their results are limited, as they have only used a candidate gene approach in their methodology[291-293].

Genome-wide methylation assessment has allowed us to identify new candidate markers that can distinguish indolent rectal cancers from those with an aggressive phenotype, complementing the novel methylation signature that has already been described. These preliminary results indicate that over 7000 genes are differentially methylated between early and advanced rectal cancers with the majority of them being hypermethylated in advanced cancer. Due to time constraints, only 10 genes were validated by pyrosequencing. Within this pool of genes, *SNAP25, TIAM1* and *SOX7* were successfully identified as candidate genes that could be incorporated into the methylation signature but would require further validation in other cohorts of rectal cancers.

The lack of supporting evidence from our validation work with the other 7 genes identified from the array could be ascribed to a few factors. Firstly, the

pyrosequencing assays did not interrogate the same CpG sites as that found on the array. One of the drawbacks of pyrosequencing is that it relies on the availability of suitable PCR primers. Primers are preferentially designed within CpG-poor regions to avoid amplification bias toward either methylated or unmethylated alleles during PCR. As a consequence, pyrosequencing assays are unable to cover long stretches of CpG-rich regions, such as those found in the array. Secondly, the pre-processing normalisation performed by Nimblegen may not be optimal for this dataset. Normalisation is essential for microarray work to take into account non-biological variations arising from the experiments such as dye bias and probe hybridisation efficiency. Nimblegen uses a scaling method to centre the log2 ratio around zero. This is a simple normalisation method but does not consider any region or intensity-dependent effects. Other normalisation methods such as the Lowess and the Quantile methods are more sophisticated and correct for other effects. It is possible that normalisation of the dataset with either of these two methods will yield better discriminating genes for validation.

The impact of tissue enrichment techniques on methylation yield was also investigated in this study. There have been conflicting reports about the value of tissue enrichment techniques in molecular analysis. Most of these studies are RNA- rather than methylation-based. In addition to tumour cells, clinical specimens also contain stromal cells, inflammatory cells and blood vessels, which may influence molecular profiling[294]. Kim *et. al.* found differential gene expression profiles between macro and microdissected gastric tumours[294]. Their work is supported by Inoue *et. al.* who suggested an association between disease recurrence and gene expression

from microdissected but not in undissected rectal cancers[295]. In contrast, de Bruin *et. al.* found that there are only minor differences in gene expression between micro and macrodissected rectal tumours and subsequently showed that the involvement of stromal cells in gene expression was only minor. This study showed only minor variations in methylation levels between the two different types of dissected tumours and is consistent with Irihara *et. al.*'s study who also showed insignificant difference in *LINE-1* methylation between macro and microdissected colonic tumours[296]. Therefore, the results do not support the routine use of microdissected tissues for methylation-based analysis.

Promoter CpG island methylation is a known epigenetic modification associated with gene silencing[136]. Various methylation platforms have been used in this study to measure methylation changes but arguably, measuring the downstream effect of epigenetics i.e. on protein expression could have the most immediate clinical utility. MS-MLPA and pyrosequencing have only recently been introduced into the clinical laboratories (Dr P. Taniere, personal communication), but IHC to detect protein expression has been performed routinely in most clinical pathology departments for many years. IHC identified an association between UNC5C protein expression and nodal status in rectal cancer, independent of its methylation status. These experiments also identified an association between UNC5C and DAPK1 expression and tumour grade but the results should be interpreted with caution given the small number of poorly differentiated tumours in the cohort.

IHC is a robust technique for some proteins but its performance relies substantially on the availability of good quality antibodies. Unlike quantitative platforms such as MS-MLPA and pyrosequencing, IHC readout is subjective and is liable to intra- and inter-observer variability. Furthermore, our experience has shown that IHC is not sufficiently sensitive to detect small changes in expression for proteins with low basal expression. On balance, pyrosequencing offers greater flexibility, which is essential during the discovery stage and IHC may have an important role in the latter stages once discriminatory molecular markers have been identified and introduced into the clinical arena.

This work has provided significant insight into the methylation patterns of rectal cancer. The exploratory studies have identified novel biomarkers that are worthy of further investigations and refinement. The principles and application of these results extend beyond rectal cancer and can be incorporated into the management of colonic cancers and any other gastrointestinal malignancies. Below, I discuss some of the future work that can be taken forward as a result of the work that has been performed in this study.

8.2 Future work

It would not be practical to validate every differentially methylated locus between early and advanced rectal cancer indentified in the array. The relatively low validation success rate (3 of 10) could be due to the selection of suboptimal candidates for validation. Data analysis should be repeated after renormalisation with either the Lowess or Quantile method. Alternatively, the results could be

validated using a different methylation array, e.g. the Illumina Human Methylation 450 Beadchip.

The search for prognostic markers in rectal cancer could be expanded beyond DNA methylation *per se* to investigate the underlying epigenetic mechanisms. Some studies have suggested that the specific epigenetic mechanism at work may determine tumour phenotype[297, 298]. Three mechanisms for epigenetic gene silencing operate in colorectal cancers. Overexpression of DNA methyltransferases (DNMTs) influences DNA methylation directly, while over-expression of the class III HDAC Sirtuin1 (SIRT1) and activation of the Polycomb repressor complex (PRC) initiate silencing through histone modifications, leading to DNA methylation[299, 300]. An assessment of DNMTs and SIRT1 by IHC and histone marks (H3K27me3, H3K27me3) by carrier chromatin immunoprecipitation (cCHIP) may identify a specific epigenetic silencing pathway that preferentially drives advanced cancers. In addition, *KRAS* and *BRAF* mutation and microsatellite instability status should be characterised in our cohort of rectal cancers, as there have been reports to suggest a prognostic role of these markers in colorectal cancer[301-305].

Rectal cancers exhibiting a good prognostic molecular signature may be stratified for organ preserving surgery in the form of TEMs. TEMs in isolation, however, will not be a sufficient curative treatment for many early rectal cancers. The addition of radiotherapy prior to TEMS is an attractive novel organ-preservation strategy and is being evaluated in the CRUK TREC study. Radiotherapy may combat one potential source of relapse; microscopic loco-regional metastasis lying undetected outside of

the local excision field. Studies have shown that a favourable response to radiotherapy in the primary tumour is independently associated with low risk of recurrence, even if no surgery is performed[306]. One of the novel areas where our molecular markers may have a role is in predicting early rectal cancers that will respond to radiotherapy. Studies have shown that sensitivity to radiotherapy is, in part, also regulated by epigenetics and novel methylation markers associated with radiosensitivity have been identified in glioma and lung cancer[307, 308]. Stratification of patients with early-stage rectal cancer for probability of microscopic tumour dissemination and sensitivity to radiotherapy would be highly valuable for planning individualised treatments that avoid unnecessary morbidity and mortality. Our panel of biomarkers will be further refined and evaluated prospectively as part of the CRUK TREC study to determine whether it reproducibly predicts mesorectal nodal metastasis following radical excision, response to neoadjuvant radiotherapy and local recurrence following a strategy of organ preservation.

Appendix A

Published work

List of References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2008;**127**(12): 2893-2917.

2. Cancer Research UK Bowel Cancer Incidence statistics. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/bowel/incidence/ [Accessed 31 October 2012.

3. Cancer Research UK Bowel Cancer Survival Statistics. http://www.cancerresearchuk.org/cancer-info/cancerstats/types/bowel/survival/ [Accessed 31 October 2012.

4. NHS Bowel Cancer Screening Programme. http://www.cancerscreening.nhs.uk/bowel/age-extension-bowel-cancer-screening.html [Accessed 31 October 2012.

5. Imperiale TF, Ransohoff DF, Itzkowitz SH, Turnbull BA, Ross ME. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *N Engl J Med* 2004;**351**(26): 2704-2714.

6. Wong CK, Fedorak RN, Prosser CI, Stewart ME, van Zanten SV, Sadowski DC. The sensitivity and specificity of guaiac and immunochemical fecal occult blood tests for the detection of advanced colonic adenomas and cancer. *Int J Colorectal Dis* 2012.

7. Hewitson P, Glasziou P, Watson E, Towler B, Irwig L. Cochrane systematic review of colorectal cancer screening using the fecal occult blood test (hemoccult): an update. *Am J Gastroenterol* 2008;**103**(6): 1541-1549.

8. Atkin WS, Edwards R, Kralj-Hans I, Wooldrage K, Hart AR, Northover JM, Parkin DM, Wardle J, Duffy SW, Cuzick J. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010;**375**(9726): 1624-1633.

9. Segnan N, Armaroli P, Bonelli L, Risio M, Sciallero S, Zappa M, Andreoni B, Arrigoni A, Bisanti L, Casella C, Crosta C, Falcini F, Ferrero F, Giacomin A, Giuliani O, Santarelli A, Visioli CB, Zanetti R, Atkin WS, Senore C. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial--SCORE. *J Natl Cancer Inst* 2011;**103**(17): 1310-1322.

10. Schoen RE, Pinsky PF, Weissfeld JL, Yokochi LA, Church T, Laiyemo AO, Bresalier R, Andriole GL, Buys SS, Crawford ED, Fouad MN, Isaacs C, Johnson CC, Reding DJ, O'Brien B, Carrick DM, Wright P, Riley TL, Purdue MP, Izmirlian G, Kramer BS, Miller AB, Gohagan JK, Prorok PC, Berg CD. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 2012;**366**(25): 2345-2357.

11. Hoff G, Grotmol T, Skovlund E, Bretthauer M. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. *BMJ* 2009;**338**: b1846.

12. Results of the first round of a demonstration pilot of screening for colorectal cancer in the United Kingdom. *BMJ* 2004;**329**(7458): 133.

13. Steele RJ, McClements PL, Libby G, Black R, Morton C, Birrell J, Mowat NA, Wilson JA, Kenicer M, Carey FA, Fraser CG. Results from the first three rounds of the Scottish demonstration pilot of FOBT screening for colorectal cancer. *Gut* 2009;**58**(4): 530-535.

14. Kim NK, Kim MJ, Yun SH, Sohn SK, Min JS. Comparative study of transrectal ultrasonography, pelvic computerized tomography, and magnetic resonance imaging in preoperative staging of rectal cancer. *Dis Colon Rectum* 1999;**42**(6): 770-775.

15. Gualdi GF, Casciani E, Guadalaxara A, d'Orta C, Polettini E, Pappalardo G. Local staging of rectal cancer with transrectal ultrasound and endorectal magnetic resonance imaging: comparison with histologic findings. *Dis Colon Rectum* 2000;**43**(3): 338-345.

16. Bianchi P, Ceriani C, Palmisano A, Pompili G, Passoni GR, Rottoli M, Cappellani A, Montorsi M. A prospective comparison of endorectal ultrasound and pelvic magnetic resonance in the preoperative staging of rectal cancer. *Ann Ital Chir* 2006;**77**(1): 41-46.

17. Zerhouni EA, Rutter C, Hamilton SR, Balfe DM, Megibow AJ, Francis IR, Moss AA, Heiken JP, Tempany CM, Aisen AM, Weinreb JC, Gatsonis C, McNeil BJ. CT and MR imaging in the staging of colorectal carcinoma: report of the Radiology Diagnostic Oncology Group II. *Radiology* 1996;**200**(2): 443-451.

18. Kulinna C, Scheidler J, Strauss T, Bonel H, Herrmann K, Aust D, Reiser M. Local staging of rectal cancer: assessment with double-contrast multislice computed tomography and transrectal ultrasound. *J Comput Assist Tomogr* 2004;**28**(1): 123-130.

19. Filippone A, Ambrosini R, Fuschi M, Marinelli T, Genovesi D, Bonomo L. Preoperative T and N staging of colorectal cancer: accuracy of contrast-enhanced multi-detector row CT colonography--initial experience. *Radiology* 2004;**231**(1): 83-90.

20. Bipat S, Glas AS, Slors FJ, Zwinderman AH, Bossuyt PM, Stoker J. Rectal cancer: local staging and assessment of lymph node involvement with endoluminal US, CT, and MR imaging--a meta-analysis. *Radiology* 2004;**232**(3): 773-783.

21. Cawthorn SJ, Parums DV, Gibbs NM, A'Hern RP, Caffarey SM, Broughton CI, Marks CG. Extent of mesorectal spread and involvement of lateral resection margin as prognostic factors after surgery for rectal cancer. *Lancet* 1990;**335**(8697): 1055-1059.

22. Willett CG. Technical advances in the treatment of patients with rectal cancer. *Int J Radiat Oncol Biol Phys* 1999;**45**(5): 1107-1108.

23. MERCURY. Extramural depth of tumor invasion at thin-section MR in patients with rectal cancer: results of the MERCURY study. *Radiology* 2007;**243**(1): 132-139.

24. Adam IJ, Mohamdee MO, Martin IG, Scott N, Finan PJ, Johnston D, Dixon MF, Quirke P. Role of circumferential margin involvement in the local recurrence of rectal cancer. *Lancet* 1994;**344**(8924): 707-711.

25. Brown G, Richards CJ, Bourne MW, Newcombe RG, Radcliffe AG, Dallimore NS, Williams GT. Morphologic predictors of lymph node status in rectal cancer with use of high-spatial-resolution MR imaging with histopathologic comparison. *Radiology* 2003;**227**(2): 371-377.

26. Gunther K, Dworak O, Remke S, Pfluger R, Merkel S, Hohenberger W, Reymond MA. Prediction of distant metastases after curative surgery for rectal cancer. *J Surg Res* 2002;**103**(1): 68-78.

27. Lahaye MJ, Engelen SM, Nelemans PJ, Beets GL, van de Velde CJ, van Engelshoven JM, Beets-Tan RG. Imaging for predicting the risk factors--the

circumferential resection margin and nodal disease--of local recurrence in rectal cancer: a meta-analysis. *Semin Ultrasound CT MR* 2005;**26**(4): 259-268.

28. Brown G, Radcliffe AG, Newcombe RG, Dallimore NS, Bourne MW, Williams GT. Preoperative assessment of prognostic factors in rectal cancer using high-resolution magnetic resonance imaging. *Br J Surg* 2003;**90**(3): 355-364.

29. Koh DM, George C, Temple L, Collins DJ, Toomey P, Raja A, Bett N, Farhat S, Husband JE, Brown G. Diagnostic accuracy of nodal enhancement pattern of rectal cancer at MRI enhanced with ultrasmall superparamagnetic iron oxide: findings in pathologically matched mesorectal lymph nodes. *AJR Am J Roentgenol* 2010;**194**(6): W505-513.

30. Quirke P, Durdey P, Dixon MF, Williams NS. Local recurrence of rectal adenocarcinoma due to inadequate surgical resection. Histopathological study of lateral tumour spread and surgical excision. *Lancet* 1986;**2**(8514): 996-999.

31. Purkayastha S, Tekkis PP, Athanasiou T, Tilney HS, Darzi AW, Heriot AG. Diagnostic precision of magnetic resonance imaging for preoperative prediction of the circumferential margin involvement in patients with rectal cancer. *Colorectal Dis* 2007;**9**(5): 402-411.

32. MERCURY. Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: prospective observational study. *BMJ* 2006;**333**(7572): 779.

33. Heald RJ. The 'Holy Plane' of rectal surgery. *J R Soc Med* 1988;**81**(9): 503-508.

34. Kapiteijn E, Putter H, van de Velde CJ. Impact of the introduction and training of total mesorectal excision on recurrence and survival in rectal cancer in The Netherlands. *Br J Surg* 2002;**89**(9): 1142-1149.

35. Quirke P, Steele R, Monson J, Grieve R, Khanna S, Couture J, O'Callaghan C, Myint AS, Bessell E, Thompson LC, Parmar M, Stephens RJ, Sebag-Montefiore D. Effect of the plane of surgery achieved on local recurrence in patients with operable rectal cancer: a prospective study using data from the MRC CR07 and NCIC-CTG CO16 randomised clinical trial. *Lancet* 2009;**373**(9666): 821-828.

36. Stelzner S, Koehler C, Stelzer J, Sims A, Witzigmann H. Extended abdominoperineal excision vs. standard abdominoperineal excision in rectal cancer--a systematic overview. *Int J Colorectal Dis* 2011;**26**(10): 1227-1240.

37. Rutten HJ, den Dulk M, Lemmens VE, van de Velde CJ, Marijnen CA. Controversies of total mesorectal excision for rectal cancer in elderly patients. *Lancet Oncol* 2008;**9**(5): 494-501.

38. Lindgren R, Hallbook O, Rutegard J, Sjodahl R, Matthiessen P. What is the risk for a permanent stoma after low anterior resection of the rectum for cancer? A six-year follow-up of a multicenter trial. *Dis Colon Rectum* 2011;**54**(1): 41-47.

39. Rutkowski A, Chwalinski M, Zajac L, Nowecki ZI, Nowacki MP. Risk of permanent stoma after resection of rectal cancer depending on the distance between the tumour lower edge and anal verge. *Pol Przegl Chir* 2011;**83**(11): 588-596.

40. Wallner C, Lange MM, Bonsing BA, Maas CP, Wallace CN, Dabhoiwala NF, Rutten HJ, Lamers WH, Deruiter MC, van de Velde CJ. Causes of fecal and urinary incontinence after total mesorectal excision for rectal cancer based on cadaveric surgery: a study from the Cooperative Clinical Investigators of the Dutch total mesorectal excision trial. *J Clin Oncol* 2008;**26**(27): 4466-4472.

41. Temple LK, Bacik J, Savatta SG, Gottesman L, Paty PB, Weiser MR, Guillem JG, Minsky BD, Kalman M, Thaler HT, Schrag D, Wong WD. The development of a validated instrument to evaluate bowel function after sphincter-preserving surgery for rectal cancer. *Dis Colon Rectum* 2005;**48**(7): 1353-1365.

42. Engel J, Kerr J, Schlesinger-Raab A, Eckel R, Sauer H, Holzel D. Quality of life in rectal cancer patients: a four-year prospective study. *Ann Surg* 2003;**238**(2): 203-213.

43. Grumann MM, Noack EM, Hoffmann IA, Schlag PM. Comparison of quality of life in patients undergoing abdominoperineal extirpation or anterior resection for rectal cancer. *Ann Surg* 2001;**233**(2): 149-156.

44. Wilson TR, Alexander DJ. Clinical and non-clinical factors influencing postoperative health-related quality of life in patients with colorectal cancer. *Br J Surg* 2008;**95**(11): 1408-1415.

45. Chua TC, Chong CH, Liauw W, Morris DL. Approach to rectal cancer surgery. *Int J Surg Oncol* 2012;**2012**: 247107.

46. Lee SH, Jeon SW, Jung MK, Kim SK, Choi GS. A comparison of transanal excision and endoscopic resection for early rectal cancer. *World J Gastrointest Endosc* 2009;**1**(1): 56-60.

47. Bach SP, Hill J, Monson JR, Simson JN, Lane L, Merrie A, Warren B, Mortensen NJ. A predictive model for local recurrence after transanal endoscopic microsurgery for rectal cancer. *Br J Surg* 2009;**96**(3): 280-290.

48. Cataldo PA, O'Brien S, Osler T. Transanal endoscopic microsurgery: a prospective evaluation of functional results. *Dis Colon Rectum* 2005;**48**(7): 1366-1371.

49. Bach SP. Developments in early rectal cancer treatment. *Eur J Cancer* 2009;**45 Suppl 1**: 464-465.

50. Papillon J. Present status of radiation therapy in the conservative management of rectal cancer. *Radiother Oncol* 1990;**17**(4): 275-283.

51. Gerard JP, Ayzac L, Coquard R, Romestaing P, Ardiet JM, Rocher FP, Barbet N, Cenni JL, Souquet JC. Endocavitary irradiation for early rectal carcinomas T1 (T2). A series of 101 patients treated with the Papillon's technique. *Int J Radiat Oncol Biol Phys* 1996;**34**(4): 775-783.

52. Sun Myint A, Grieve RJ, McDonald AC, Levine EL, Ramani S, Perkins K, Wong H, Makin CA, Hershman MJ. Combined modality treatment of early rectal cancer: the UK experience. *Clin Oncol (R Coll Radiol)* 2007;**19**(9): 674-681.

53. Habr-Gama A, Perez RO, Nadalin W, Sabbaga J, Ribeiro U, Jr., Silva e Sousa AH, Jr., Campos FG, Kiss DR, Gama-Rodrigues J. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg* 2004;**240**(4): 711-717; discussion 717-718.

54. Chakravarti A, Compton CC, Shellito PC, Wood WC, Landry J, Machuta SR, Kaufman D, Ancukiewicz M, Willett CG. Long-term follow-up of patients with rectal cancer managed by local excision with and without adjuvant irradiation. *Ann Surg* 1999;**230**(1): 49-54.

55. Russell AH, Harris J, Rosenberg PJ, Sause WT, Fisher BJ, Hoffman JP, Kraybill WG, Byhardt RW. Anal sphincter conservation for patients with adenocarcinoma of the distal rectum: long-term results of radiation therapy oncology group protocol 89-02. *Int J Radiat Oncol Biol Phys* 2000;**46**(2): 313-322.

56. Kapiteijn E, Marijnen CA, Nagtegaal ID, Putter H, Steup WH, Wiggers T, Rutten HJ, Pahlman L, Glimelius B, van Krieken JH, Leer JW, van de Velde CJ. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med* 2001;**345**(9): 638-646.

57. Sauer R, Becker H, Hohenberger W, Rodel C, Wittekind C, Fietkau R, Martus P, Tschmelitsch J, Hager E, Hess CF, Karstens JH, Liersch T, Schmidberger H, Raab R. Preoperative versus postoperative chemoradiotherapy for rectal cancer. *N Engl J Med* 2004;**351**(17): 1731-1740.

58. Sebag-Montefiore D, Stephens RJ, Steele R, Monson J, Grieve R, Khanna S, Quirke P, Couture J, de Metz C, Myint AS, Bessell E, Griffiths G, Thompson LC, Parmar M. Preoperative radiotherapy versus selective postoperative chemoradiotherapy in patients with rectal cancer (MRC CR07 and NCIC-CTG C016): a multicentre, randomised trial. *Lancet* 2009;**373**(9666): 811-820.

59. Dahlberg M, Glimelius B, Pahlman L. Improved survival and reduction in local failure rates after preoperative radiotherapy: evidence for the generalizability of the results of Swedish Rectal Cancer Trial. *Ann Surg* 1999;**229**(4): 493-497.

60. Borschitz T, Wachtlin D, Mohler M, Schmidberger H, Junginger T. Neoadjuvant chemoradiation and local excision for T2-3 rectal cancer. *Ann Surg Oncol* 2008;**15**(3): 712-720.

61. Lezoche E, Baldarelli M, Lezoche G, Paganini AM, Gesuita R, Guerrieri M. Randomized clinical trial of endoluminal locoregional resection versus laparoscopic total mesorectal excision for T2 rectal cancer after neoadjuvant therapy. *Br J Surg* 2012;**99**(9): 1211-1218.

62. Bujko K, Richter P, Kolodziejczyk M, Nowacki MP, Kulig J, Popiela T, Gach T, Oledzki J, Sopylo R, Meissner W, Wierzbicki R, Polkowski W, Kowalska T, Stryczynska G, Paprota K. Preoperative radiotherapy and local excision of rectal cancer with immediate radical re-operation for poor responders. *Radiother Oncol* 2009;**92**(2): 195-201.

63. Bujko K, Nowacki MP, Nasierowska-Guttmejer A, Michalski W, Bebenek M, Kryj M. Long-term results of a randomized trial comparing preoperative short-course radiotherapy with preoperative conventionally fractionated chemoradiation for rectal cancer. *Br J Surg* 2006;**93**(10): 1215-1223.

64. Graf W, Dahlberg M, Osman MM, Holmberg L, Pahlman L, Glimelius B. Short-term preoperative radiotherapy results in down-staging of rectal cancer: a study of 1316 patients. *Radiother Oncol* 1997;**43**(2): 133-137.

65. Radu C, Berglund A, Pahlman L, Glimelius B. Short-course preoperative radiotherapy with delayed surgery in rectal cancer - a retrospective study. *Radiother Oncol* 2008;**87**(3): 343-349.

66. Sikorski R, Yao B. Visualizing the landscape of selection biomarkers in current phase III oncology clinical trials. *Sci Transl Med* 2010;**2**(34): 34ps27.

67. Jover R, Zapater P, Castells A, Llor X, Andreu M, Cubiella J, Balaguer F, Sempere L, Xicola RM, Bujanda L, Rene JM, Clofent J, Bessa X, Morillas JD, Nicolas-Perez D, Pons E, Paya A, Alenda C. The efficacy of adjuvant chemotherapy with 5-fluorouracil in colorectal cancer depends on the mismatch repair status. *Eur J Cancer* 2009;**45**(3): 365-373.

68. Sargent DJ, Marsoni S, Monges G, Thibodeau SN, Labianca R, Hamilton SR, French AJ, Kabat B, Foster NR, Torri V, Ribic C, Grothey A, Moore M, Zaniboni A, Seitz

JF, Sinicrope F, Gallinger S. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. *J Clin Oncol* 2010;**28**(20): 3219-3226.

69. CRUK Biomarkers and Imaging Definitions. http://www.cancerresearchuk.org/science/funding/apply/additional-information/biomarkers-imaging-definitions/ [Accessed 21 October 2013.

70. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* 1990;**61**(5): 759-767.

71. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;**100**(1): 57-70.

72. Feinberg AP, Tycko B. The history of cancer epigenetics. *Nat Rev Cancer* 2004;**4**(2): 143-153.

73. Lengauer C, Kinzler KW, Vogelstein B. Genetic instabilities in human cancers. *Nature* 1998;**396**(6712): 643-649.

74. Grady WM, Carethers JM. Genomic and epigenetic instability in colorectal cancer pathogenesis. *Gastroenterology* 2008;**135**(4): 1079-1099.

75. Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, Markowitz SD, Kinzler KW, Vogelstein B. Mutations of mitotic checkpoint genes in human cancers. *Nature* 1998;**392**(6673): 300-303.

76. Jin DY, Spencer F, Jeang KT. Human T cell leukemia virus type 1 oncoprotein Tax targets the human mitotic checkpoint protein MAD1. *Cell* 1998;**93**(1): 81-91.

77. Li Y, Benezra R. Identification of a human mitotic checkpoint gene: hsMAD2. *Science* 1996;**274**(5285): 246-248.

78. D'Assoro AB, Lingle WL, Salisbury JL. Centrosome amplification and the development of cancer. *Oncogene* 2002;**21**(40): 6146-6153.

79. Brinkley BR. Managing the centrosome numbers game: from chaos to stability in cancer cell division. *Trends Cell Biol* 2001;**11**(1): 18-21.

80. Lingle WL, Barrett SL, Negron VC, D'Assoro AB, Boeneman K, Liu W, Whitehead CM, Reynolds C, Salisbury JL. Centrosome amplification drives chromosomal instability in breast tumor development. *Proc Natl Acad Sci U S A* 2002;**99**(4): 1978-1983.

81. Sen S, Zhou H, White RA. A putative serine/threonine kinase encoding gene BTAK on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines. *Oncogene* 1997;**14**(18): 2195-2200.

82. Bischoff JR, Anderson L, Zhu Y, Mossie K, Ng L, Souza B, Schryver B, Flanagan P, Clairvoyant F, Ginther C, Chan CS, Novotny M, Slamon DJ, Plowman GD. A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *EMBO J* 1998;**17**(11): 3052-3065.

83. Macmillan JC, Hudson JW, Bull S, Dennis JW, Swallow CJ. Comparative expression of the mitotic regulators SAK and PLK in colorectal cancer. *Ann Surg Oncol* 2001;**8**(9): 729-740.

84. Takahashi T, Sano B, Nagata T, Kato H, Sugiyama Y, Kunieda K, Kimura M, Okano Y, Saji S. Polo-like kinase 1 (PLK1) is overexpressed in primary colorectal cancers. *Cancer Sci* 2003;**94**(2): 148-152.

85. Kolodner RD, Putnam CD, Myung K. Maintenance of genome stability in Saccharomyces cerevisiae. *Science* 2002;**297**(5581): 552-557.

86. Wang Z, Cummins JM, Shen D, Cahill DP, Jallepalli PV, Wang TL, Parsons DW, Traverso G, Awad M, Silliman N, Ptak J, Szabo S, Willson JK, Markowitz SD, Goldberg ML, Karess R, Kinzler KW, Vogelstein B, Velculescu VE, Lengauer C. Three classes of

genes mutated in colorectal cancers with chromosomal instability. *Cancer Res* 2004;**64**(9): 2998-3001.

87. Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. *Science* 2002;**297**(5581): 565-569.

88. Lo AW, Sabatier L, Fouladi B, Pottier G, Ricoul M, Murnane JP. DNA amplification by breakage/fusion/bridge cycles initiated by spontaneous telomere loss in a human cancer cell line. *Neoplasia* 2002;**4**(6): 531-538.

89. Plentz RR, Wiemann SU, Flemming P, Meier PN, Kubicka S, Kreipe H, Manns MP, Rudolph KL. Telomere shortening of epithelial cells characterises the adenoma-carcinoma transition of human colorectal cancer. *Gut* 2003;**52**(9): 1304-1307.

90. Engelhardt M, Drullinsky P, Guillem J, Moore MA. Telomerase and telomere length in the development and progression of premalignant lesions to colorectal cancer. *Clin Cancer Res* 1997;**3**(11): 1931-1941.

91. Chadeneau C, Hay K, Hirte HW, Gallinger S, Bacchetti S. Telomerase activity associated with acquisition of malignancy in human colorectal cancer. *Cancer Res* 1995;**55**(12): 2533-2536.

92. Tatsumoto N, Hiyama E, Murakami Y, Imamura Y, Shay JW, Matsuura Y, Yokoyama T. High telomerase activity is an independent prognostic indicator of poor outcome in colorectal cancer. *Clin Cancer Res* 2000;**6**(7): 2696-2701.

93. Pino MS, Chung DC. The chromosomal instability pathway in colon cancer. *Gastroenterology* 2010;**138**(6): 2059-2072.

94. Beroud C, Soussi T. APC gene: database of germline and somatic mutations in human tumors and cell lines. *Nucleic Acids Res* 1996;**24**(1): 121-124.

95. Miyoshi Y, Ando H, Nagase H, Nishisho I, Horii A, Miki Y, Mori T, Utsunomiya J, Baba S, Petersen G, et al. Germ-line mutations of the APC gene in 53 familial adenomatous polyposis patients. *Proc Natl Acad Sci U S A* 1992;**89**(10): 4452-4456.

96. Miyoshi Y, Nagase H, Ando H, Horii A, Ichii S, Nakatsuru S, Aoki T, Miki Y, Mori T, Nakamura Y. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum Mol Genet* 1992;**1**(4): 229-233.

97. Rowan AJ, Lamlum H, Ilyas M, Wheeler J, Straub J, Papadopoulou A, Bicknell D, Bodmer WF, Tomlinson IP. APC mutations in sporadic colorectal tumors: A mutational "hotspot" and interdependence of the "two hits". *Proc Natl Acad Sci U S A* 2000;**97**(7): 3352-3357.

98. Rusan NM, Peifer M. Original CIN: reviewing roles for APC in chromosome instability. *J Cell Biol* 2008;**181**(5): 719-726.

99. Green RA, Wollman R, Kaplan KB. APC and EB1 function together in mitosis to regulate spindle dynamics and chromosome alignment. *Mol Biol Cell* 2005;**16**(10): 4609-4622.

100. Green RA, Kaplan KB. Chromosome instability in colorectal tumor cells is associated with defects in microtubule plus-end attachments caused by a dominant mutation in APC. *J Cell Biol* 2003;**163**(5): 949-961.

101. Clevers H. Wnt/beta-catenin signaling in development and disease. *Cell* 2006;**127**(3): 469-480.

102. Mann B, Gelos M, Siedow A, Hanski ML, Gratchev A, Ilyas M, Bodmer WF, Moyer MP, Riecken EO, Buhr HJ, Hanski C. Target genes of beta-catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. *Proc Natl Acad Sci U S A* 1999;**96**(4): 1603-1608.

103. van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP, Tjon-Pon-Fong M, Moerer P, van den Born M, Soete G, Pals S, Eilers M, Medema R, Clevers H. The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* 2002;**111**(2): 241-250.

104. Santini D, Loupakis F, Vincenzi B, Floriani I, Stasi I, Canestrari E, Rulli E, Maltese PE, Andreoni F, Masi G, Graziano F, Baldi GG, Salvatore L, Russo A, Perrone G, Tommasino MR, Magnani M, Falcone A, Tonini G, Ruzzo A. High concordance of KRAS status between primary colorectal tumors and related metastatic sites: implications for clinical practice. *Oncologist* 2008;**13**(12): 1270-1275.

105. Downward J. Targeting RAS signalling pathways in cancer therapy. *Nat Rev Cancer* 2003;**3**(1): 11-22.

106. Castagnola P, Giaretti W. Mutant KRAS, chromosomal instability and prognosis in colorectal cancer. *Biochim Biophys Acta* 2005;**1756**(2): 115-125.

107. Saavedra HI, Knauf JA, Shirokawa JM, Wang J, Ouyang B, Elisei R, Stambrook PJ, Fagin JA. The RAS oncogene induces genomic instability in thyroid PCCL3 cells via the MAPK pathway. *Oncogene* 2000;**19**(34): 3948-3954.

108. Pruitt K, Der CJ. Ras and Rho regulation of the cell cycle and oncogenesis. *Cancer Lett* 2001;**171**(1): 1-10.

109. Pacold ME, Suire S, Perisic O, Lara-Gonzalez S, Davis CT, Walker EH, Hawkins PT, Stephens L, Eccleston JF, Williams RL. Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase gamma. *Cell* 2000;**103**(6): 931-943.

110. Hennessy BT, Smith DL, Ram PT, Lu Y, Mills GB. Exploiting the PI3K/AKT pathway for cancer drug discovery. *Nat Rev Drug Discov* 2005;**4**(12): 988-1004.

111. Molchadsky A, Rivlin N, Brosh R, Rotter V, Sarig R. p53 is balancing development, differentiation and de-differentiation to assure cancer prevention. *Carcinogenesis* 2010;**31**(9): 1501-1508.

112. Vogelstein B, Kinzler KW. Cancer genes and the pathways they control. *Nat Med* 2004;**10**(8): 789-799.

113. Menendez D, Inga A, Resnick MA. The expanding universe of p53 targets. *Nat Rev Cancer* 2009;**9**(10): 724-737.

114. Schmitt CA, Fridman JS, Yang M, Baranov E, Hoffman RM, Lowe SW. Dissecting p53 tumor suppressor functions in vivo. *Cancer Cell* 2002;**1**(3): 289-298.

115. Leslie A, Carey FA, Pratt NR, Steele RJ. The colorectal adenoma-carcinoma sequence. *Br J Surg* 2002;**89**(7): 845-860.

116. Beroud C, Soussi T. The UMD-p53 database: new mutations and analysis tools. *Hum Mutat* 2003;**21**(3): 176-181.

117. Schlotterer C. Evolutionary dynamics of microsatellite DNA. *Chromosoma* 2000;**109**(6): 365-371.

118. Bellizzi AM, Frankel WL. Colorectal cancer due to deficiency in DNA mismatch repair function: a review. *Adv Anat Pathol* 2009;**16**(6): 405-417.

119. Raut CP, Pawlik TM, Rodriguez-Bigas MA. Clinicopathologic features in colorectal cancer patients with microsatellite instability. *Mutat Res* 2004;**568**(2): 275-282.

120. Huang J, Kuismanen SA, Liu T, Chadwick RB, Johnson CK, Stevens MW, Richards SK, Meek JE, Gao X, Wright FA, Mecklin JP, Jarvinen HJ, Gronberg H,

Bisgaard ML, Lindblom A, Peltomaki P. MSH6 and MSH3 are rarely involved in genetic predisposition to nonpolypotic colon cancer. *Cancer Res* 2001;**61**(4): 1619-1623.

121. Kane MF, Loda M, Gaida GM, Lipman J, Mishra R, Goldman H, Jessup JM, Kolodner R. 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 Res* 1997;**57**(5): 808-811.

122. Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD, Markowitz SD. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. *Proc Natl Acad Sci U S A* 1998;**95**(15): 8698-8702.

123. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res* 2002;**62**(9): 2447-2454.

124. Malkhosyan S, Rampino N, Yamamoto H, Perucho M. Frameshift mutator mutations. *Nature* 1996;**382**(6591): 499-500.

125. Boland CR, Thibodeau SN, Hamilton SR, Sidransky D, Eshleman JR, Burt RW, Meltzer SJ, Rodriguez-Bigas MA, Fodde R, Ranzani GN, Srivastava S. 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 Res* 1998;**58**(22): 5248-5257.

126. Ward R, Meagher A, Tomlinson I, O'Connor T, Norrie M, Wu R, Hawkins N. Microsatellite instability and the clinicopathological features of sporadic colorectal cancer. *Gut* 2001;**48**(6): 821-829.

127. Taby R, Issa JP. Cancer epigenetics. *CA Cancer J Clin* 2010;**60**(6): 376-392.

128. Skinner MK, Manikkam M, Guerrero-Bosagna C. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab* 2010;**21**(4): 214-222.

129. Kiefer JC. Epigenetics in development. *Dev Dyn* 2007;**236**(4): 1144-1156.

130. Klose RJ, Bird AP. Genomic DNA methylation: the mark and its mediators. *Trends Biochem Sci* 2006;**31**(2): 89-97.

131. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. *Biochem Biophys Res Commun* 1983;**111**(1): 47-54.

132. Esteller M. Epigenetics in cancer. *N Engl J Med* 2008;**358**(11): 1148-1159.

133. Caldwell GM, Jones C, Gensberg K, Jan S, Hardy RG, Byrd P, Chughtai S, Wallis Y, Matthews GM, Morton DG. The Wnt antagonist sFRP1 in colorectal tumorigenesis. *Cancer Res* 2004;**64**(3): 883-888.

134. Leong KJ, Wei W, Tannahill LA, Caldwell GM, Jones CE, Morton DG, Matthews GM, Bach SP. Methylation profiling of rectal cancer identifies novel markers of early-stage disease. *Br J Surg* 2011;**98**(5): 724-734.

135. Jones PA, Laird PW. Cancer epigenetics comes of age. *Nat Genet* 1999;**21**(2): 163-167.

136. Herman JG, Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. *N Engl J Med* 2003;**349**(21): 2042-2054.

137. Kondo Y, Issa JP. Epigenetic changes in colorectal cancer. *Cancer Metastasis Rev* 2004;**23**(1-2): 29-39.

138. Watt F, Molloy PL. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* 1988;**2**(9): 1136-1143.

139. Griswold MD, Kim JS. Site-specific methylation of the promoter alters deoxyribonucleic acid-protein interactions and prevents follicle-stimulating hormone receptor gene transcription. *Biol Reprod* 2001;**64**(2): 602-610.

140. Comb M, Goodman HM. CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. *Nucleic Acids Res* 1990;**18**(13): 3975-3982.

141. Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C, Minucci S, Pelicci PG. Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. *Science* 2002;**295**(5557): 1079-1082.

142. Geiman TM, Sankpal UT, Robertson AK, Zhao Y, Robertson KD. DNMT3B interacts with hSNF2H chromatin remodeling enzyme, HDACs 1 and 2, and components of the histone methylation system. *Biochem Biophys Res Commun* 2004;**318**(2): 544-555.

143. Fuks F, Burgers WA, Godin N, Kasai M, Kouzarides T. Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. *EMBO J* 2001;**20**(10): 2536-2544.

144. Rountree MR, Bachman KE, Baylin SB. DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. *Nat Genet* 2000;**25**(3): 269-277.

145. Hsieh CL. Stability of patch methylation and its impact in regions of transcriptional initiation and elongation. *Mol Cell Biol* 1997;**17**(10): 5897-5904.

146. Lorincz MC, Dickerson DR, Schmitt M, Groudine M. Intragenic DNA methylation alters chromatin structure and elongation efficiency in mammalian cells. *Nat Struct Mol Biol* 2004;**11**(11): 1068-1075.

147. Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN, Bird A. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. *Nature* 1998;**393**(6683): 386-389.

148. Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, Erdjument-Bromage H, Tempst P, Reinberg D, Bird A. MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. *Nat Genet* 1999;**23**(1): 58-61.

149. Sarraf SA, Stancheva I. Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. *Mol Cell* 2004;**15**(4): 595-605.

150. Wolffe AP, Kurumizaka H. The nucleosome: a powerful regulator of transcription. *Prog Nucleic Acid Res Mol Biol* 1998;**61**: 379-422.

151. Hake SB, Xiao A, Allis CD. Linking the epigenetic 'language' of covalent histone modifications to cancer. *Br J Cancer* 2007;**96 Suppl**: R31-39.

152. van Engeland M, Derks S, Smits KM, Meijer GA, Herman JG. Colorectal cancer epigenetics: complex simplicity. *J Clin Oncol* 2011;**29**(10): 1382-1391.

153. Kohler C, Villar CB. Programming of gene expression by Polycomb group proteins. *Trends Cell Biol* 2008;**18**(5): 236-243.

154. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev* 2006;**20**(9): 1123-1136.

155. Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. *Nat Rev Genet* 2009;**10**(2): 94-108.

156. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* 2009;**19**(1): 92-105.

157. Schickel R, Boyerinas B, Park SM, Peter ME. MicroRNAs: key players in the immune system, differentiation, tumorigenesis and cell death. *Oncogene* 2008;**27**(45): 5959-5974.

158. Cao P, Deng Z, Wan M, Huang W, Cramer SD, Xu J, Lei M, Sui G. MicroRNA-101 negatively regulates Ezh2 and its expression is modulated by androgen receptor and HIF-1alpha/HIF-1beta. *Mol Cancer* 2010;**9**: 108.

159. Benetatos L, Voulgaris E, Vartholomatos G, Hatzimichael E. Non-coding RNAs and EZH2 interactions in cancer: Long and short tales from the transcriptome. *Int J Cancer* 2012.

160. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002;**3**(6): 415-428.

161. Burbee DG, Forgacs E, Zochbauer-Muller S, Shivakumar L, Fong K, Gao B, Randle D, Kondo M, Virmani A, Bader S, Sekido Y, Latif F, Milchgrub S, Toyooka S, Gazdar AF, Lerman MI, Zabarovsky E, White M, Minna JD. Epigenetic inactivation of RASSF1A in lung and breast cancers and malignant phenotype suppression. *J Natl Cancer Inst* 2001;**93**(9): 691-699.

162. Dammann R, Li C, Yoon JH, Chin PL, Bates S, Pfeifer GP. Epigenetic inactivation of a RAS association domain family protein from the lung tumour suppressor locus 3p21.3. *Nat Genet* 2000;**25**(3): 315-319.

163. Herman JG, Civin CI, Issa JP, Collector MI, Sharkis SJ, Baylin SB. Distinct patterns of inactivation of p15INK4B and p16INK4A characterize the major types of hematological malignancies. *Cancer Res* 1997;**57**(5): 837-841.

164. Esteller M, Hamilton SR, Burger PC, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is a common event in primary human neoplasia. *Cancer Res* 1999;**59**(4): 793-797.

165. Feinberg AP, Ohlsson R, Henikoff S. The epigenetic progenitor origin of human cancer. *Nat Rev Genet* 2006;**7**(1): 21-33.

166. Robertson KD, Jones PA. DNA methylation: past, present and future directions. *Carcinogenesis* 2000;**21**(3): 461-467.

167. Sawan C, Vaissiere T, Murr R, Herceg Z. Epigenetic drivers and genetic passengers on the road to cancer. *Mutat Res* 2008;**642**(1-2): 1-13.

168. Rideout WM, 3rd, Coetzee GA, Olumi AF, Jones PA. 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* 1990;**249**(4974): 1288-1290.

169. Magewu AN, Jones PA. Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. *Mol Cell Biol* 1994;**14**(6): 4225-4232.

170. Cowell JK, Smith T, Bia B. Frequent constitutional C to T mutations in CGA-arginine codons in the RB1 gene produce premature stop codons in patients with bilateral (hereditary) retinoblastoma. *Eur J Hum Genet* 1994;**2**(4): 281-290.

171. Esteller M, Risques RA, Toyota M, Capella G, Moreno V, Peinado MA, Baylin SB, Herman JG. Promoter hypermethylation of the DNA repair gene O(6)-methylguanine-DNA methyltransferase is associated with the presence of G:C to A:T transition mutations in p53 in human colorectal tumorigenesis. *Cancer Res* 2001;**61**(12): 4689-4692.

172. Esteller M, Toyota M, Sanchez-Cespedes M, Capella G, Peinado MA, Watkins DN, Issa JP, Sidransky D, Baylin SB, Herman JG. Inactivation of the DNA repair gene O6-methylguanine-DNA methyltransferase by promoter hypermethylation is associated with G to A mutations in K-ras in colorectal tumorigenesis. *Cancer Res* 2000;**60**(9): 2368-2371.

173. Yoon JH, Smith LE, Feng Z, Tang M, Lee CS, Pfeifer GP. Methylated CpG dinucleotides are the preferential targets for G-to-T transversion mutations induced by benzo[a]pyrene diol epoxide in mammalian cells: similarities with the p53 mutation spectrum in smoking-associated lung cancers. *Cancer Res* 2001;**61**(19): 7110-7117.

174. You YH, Li C, Pfeifer GP. Involvement of 5-methylcytosine in sunlight-induced mutagenesis. *J Mol Biol* 1999;**293**(3): 493-503.

175. You YH, Pfeifer GP. Similarities in sunlight-induced mutational spectra of CpG-methylated transgenes and the p53 gene in skin cancer point to an important role of 5-methylcytosine residues in solar UV mutagenesis. *J Mol Biol* 2001;**305**(3): 389-399.

176. Smith LE, Denissenko MF, Bennett WP, Li H, Amin S, Tang M, Pfeifer GP. Targeting of lung cancer mutational hotspots by polycyclic aromatic hydrocarbons. *J Natl Cancer Inst* 2000;**92**(10): 803-811.

177. Toyota M, Ahuja N, Ohe-Toyota M, Herman JG, Baylin SB, Issa JP. CpG island methylator phenotype in colorectal cancer. *Proc Natl Acad Sci U S A* 1999;**96**(15): 8681-8686.

178. Toyota M, Ohe-Toyota M, Ahuja N, Issa JP. Distinct genetic profiles in colorectal tumors with or without the CpG island methylator phenotype. *Proc Natl Acad Sci U S A* 2000;**97**(2): 710-715.

179. Barault L, Charon-Barra C, Jooste V, de la Vega MF, Martin L, Roignot P, Rat P, Bouvier AM, Laurent-Puig P, Faivre J, Chapusot C, Piard F. Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. *Cancer Res* 2008;**68**(20): 8541-8546.

180. Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Meyerhardt JA, Loda M, Giovannucci EL, Fuchs CS. CpG island methylator phenotype, microsatellite instability, BRAF mutation and clinical outcome in colon cancer. *Gut* 2009;**58**(1): 90-96.

181. Samowitz WS, Albertsen H, Herrick J, Levin TR, Sweeney C, Murtaugh MA, Wolff RK, Slattery ML. Evaluation of a large, population-based sample supports a CpG island methylator phenotype in colon cancer. *Gastroenterology* 2005;**129**(3): 837-845.

182. Weisenberger DJ, Siegmund KD, Campan M, Young J, Long TI, Faasse MA, Kang GH, Widschwendter M, Weener D, Buchanan D, Koh H, Simms L, Barker M, Leggett B, Levine J, Kim M, French AJ, Thibodeau SN, Jass J, Haile R, Laird PW. CpG

island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. *Nat Genet* 2006;**38**(7): 787-793.

183. Lee S, Cho NY, Yoo EJ, Kim JH, Kang GH. CpG island methylator phenotype in colorectal cancers: comparison of the new and classic CpG island methylator phenotype marker panels. *Arch Pathol Lab Med* 2008;**132**(10): 1657-1665.

184. Wang Z, Yuan X, Jiao N, Zhu H, Zhang Y, Tong J. CDH13 and FLBN3 gene methylation are associated with poor prognosis in colorectal cancer. *Pathol Oncol Res* 2011;**18**(2): 263-270.

185. Tanaka M, Chang P, Li Y, Li D, Overman M, Maru DM, Sethi S, Phillips J, Bland GL, Abbruzzese JL, Eng C. Association of CHFR promoter methylation with disease recurrence in locally advanced colon cancer. *Clin Cancer Res* 2011;**17**(13): 4531-4540.

186. Dallol A, Al-Maghrabi J, Buhmeida A, Gari MA, Chaudhary AG, Schulten HJ, Abuzenadah AM, Al-Ahwal MS, Sibiany A, Al-Qahtani MH. Methylation of the polycomb group target genes is a possible biomarker for favorable prognosis in colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 2012;**21**(11): 2069-2075.

187. Ward RL, Cheong K, Ku SL, Meagher A, O'Connor T, Hawkins NJ. Adverse prognostic effect of methylation in colorectal cancer is reversed by microsatellite instability. *J Clin Oncol* 2003;**21**(20): 3729-3736.

188. de Maat MF, van de Velde CJ, van der Werff MP, Putter H, Umetani N, Klein-Kranenbarg EM, Turner RR, van Krieken JH, Bilchik A, Tollenaar RA, Hoon DS. Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. *J Clin Oncol* 2008;**26**(14): 2327-2335.

189. de Maat MF, van de Velde CJ, Benard A, Putter H, Morreau H, van Krieken JH, Meershoek Klein-Kranenbarg E, de Graaf EJ, Tollenaar RA, Hoon DS. Identification of a quantitative MINT locus methylation profile predicting local regional recurrence of rectal cancer. *Clin Cancer Res* 2010;**16**(10): 2811-2818.

190. Jo P, Jung K, Grade M, Conradi LC, Wolff HA, Kitz J, Becker H, Ruschoff J, Hartmann A, Beissbarth T, Muller-Dornieden A, Ghadimi M, Schneider-Stock R, Gaedcke J. CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer. *Surgery* 2012;**151**(4): 564-570.

191. Talens RP, Boomsma DI, Tobi EW, Kremer D, Jukema JW, Willemsen G, Putter H, Slagboom PE, Heijmans BT. Variation, patterns, and temporal stability of DNA methylation: considerations for epigenetic epidemiology. *FASEB J* 2010;**24**(9): 3135-3144.

192. Herman JG, Graff JR, Myohanen S, Nelkin BD, Baylin SB. Methylation-specific PCR: a novel PCR assay for methylation status of CpG islands. *Proc Natl Acad Sci U S A* 1996;**93**(18): 9821-9826.

193. Cameron EE, Baylin SB, Herman JG. p15(INK4B) CpG island methylation in primary acute leukemia is heterogeneous and suggests density as a critical factor for transcriptional silencing. *Blood* 1999;**94**(7): 2445-2451.

194. Hsieh CL. Dependence of transcriptional repression on CpG methylation density. *Mol Cell Biol* 1994;**14**(8): 5487-5494.

195. Eads CA, Danenberg KD, Kawakami K, Saltz LB, Blake C, Shibata D, Danenberg PV, Laird PW. MethyLight: a high-throughput assay to measure DNA methylation. *Nucleic Acids Res* 2000;**28**(8): E32.

196. Li Y, Tollefsbol TO. DNA methylation detection: bisulfite genomic sequencing analysis. *Methods Mol Biol* 2011;**791**: 11-21.

197. Colella S, Shen L, Baggerly KA, Issa JP, Krahe R. Sensitive and quantitative universal Pyrosequencing methylation analysis of CpG sites. *Biotechniques* 2003;**35**(1): 146-150.

198. Shames DS, Minna JD, Gazdar AF. Methods for detecting DNA methylation in tumors: from bench to bedside. *Cancer Lett* 2007;**251**(2): 187-198.

199. Ammerpohl O, Martin-Subero JI, Richter J, Vater I, Siebert R. Hunting for the 5th base: Techniques for analyzing DNA methylation. *Biochim Biophys Acta* 2009;**1790**(9): 847-862.

200. Wojdacz TK, Dobrovic A. Methylation-sensitive high resolution melting (MS-HRM): a new approach for sensitive and high-throughput assessment of methylation. *Nucleic Acids Res* 2007;**35**(6): e41.

201. Smith E, Jones ME, Drew PA. Quantitation of DNA methylation by melt curve analysis. *BMC Cancer* 2009;**9**: 123.

202. Zilberman D, Henikoff S. Genome-wide analysis of DNA methylation patterns. *Development* 2007;**134**(22): 3959-3965.

203. Verma M, Srivastava S. Epigenetics in cancer: implications for early detection and prevention. *Lancet Oncol* 2002;**3**(12): 755-763.

204. Laird PW. The power and the promise of DNA methylation markers. *Nat Rev Cancer* 2003;**3**(4): 253-266.

205. Daskalos A, Nikolaidis G, Xinarianos G, Savvari P, Cassidy A, Zakopoulou R, Kotsinas A, Gorgoulis V, Field JK, Liloglou T. Hypomethylation of retrotransposable

elements correlates with genomic instability in non-small cell lung cancer. *Int J Cancer* 2009;**124**(1): 81-87.

206. Koga Y, Pelizzola M, Cheng E, Krauthammer M, Sznol M, Ariyan S, Narayan D, Molinaro AM, Halaban R, Weissman SM. Genome-wide screen of promoter methylation identifies novel markers in melanoma. *Genome Res* 2009;**19**(8): 1462-1470.

207. Belshaw NJ, Elliott GO, Foxall RJ, Dainty JR, Pal N, Coupe A, Garg D, Bradburn DM, Mathers JC, Johnson IT. Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. *Br J Cancer* 2008;**99**(1): 136-142.

208. Belshaw NJ, Pal N, Tapp HS, Dainty JR, Lewis MP, Williams MR, Lund EK, Johnson IT. Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa. *Carcinogenesis* 2010;**31**(6): 1158-1163.

209. Bonin S, Petrera F, Niccolini B, Stanta G. PCR analysis in archival postmortem tissues. *Mol Pathol* 2003;**56**(3): 184-186.

210. Grunau C, Clark SJ, Rosenthal A. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. *Nucleic Acids Res* 2001;**29**(13): E65-65.

211. Fend F, Raffeld M. Laser capture microdissection in pathology. *J Clin Pathol* 2000;**53**(9): 666-672.

212. Leong KJ, Wei W, Tannahill LA, Caldwell GM, Jones CE, Morton DG, Matthews GM, Bach SP. Methylation profiling of rectal cancer identifies novel markers of early-stage disease. *Br J Surg* 2011.

213. Ogino S, Nosho K, Kirkner GJ, Kawasaki T, Chan AT, Schernhammer ES, Giovannucci EL, Fuchs CS. A cohort study of tumoral LINE-1 hypomethylation and prognosis in colon cancer. *J Natl Cancer Inst* 2008;**100**(23): 1734-1738.

214. Ronaghi M. Pyrosequencing sheds light on DNA sequencing. *Genome Res* 2001;**11**(1): 3-11.

215. de Bruin EC, van de Pas S, Lips EH, van Eijk R, van der Zee MM, Lombaerts M, van Wezel T, Marijnen CA, van Krieken JH, Medema JP, van de Velde CJ, Eilers PH, Peltenburg LT. Macrodissection versus microdissection of rectal carcinoma: minor influence of stroma cells to tumor cell gene expression profiles. *BMC Genomics* 2005;**6**: 142.

216. Fiegl H, Millinger S, Goebel G, Muller-Holzner E, Marth C, Laird PW, Widschwendter M. Breast cancer DNA methylation profiles in cancer cells and tumor stroma: association with HER-2/neu status in primary breast cancer. *Cancer Res* 2006;**66**(1): 29-33.

217. Hanson JA, Gillespie JW, Grover A, Tangrea MA, Chuaqui RF, Emmert-Buck MR, Tangrea JA, Libutti SK, Linehan WM, Woodson KG. Gene promoter methylation in prostate tumor-associated stromal cells. *J Natl Cancer Inst* 2006;**98**(4): 255-261.

218. Yang J, Luo H, Li Y, Li J, Cai Z, Su X, Dai D, Du W, Chen T, Chen M. Intratumoral Heterogeneity Determines Discordant Results of Diagnostic Tests for Human Epidermal Growth Factor Receptor (HER) 2 in Gastric Cancer Specimens. *Cell Biochem Biophys* 2011.

219. Wendt MK, Johanesen PA, Kang-Decker N, Binion DG, Shah V, Dwinell MB. Silencing of epithelial CXCL12 expression by DNA hypermethylation promotes colonic carcinoma metastasis. *Oncogene* 2006;**25**(36): 4986-4997.

220. Shin SK, Nagasaka T, Jung BH, Matsubara N, Kim WH, Carethers JM, Boland CR, Goel A. Epigenetic and genetic alterations in Netrin-1 receptors UNC5C and DCC in human colon cancer. *Gastroenterology* 2007;**133**(6): 1849-1857.

221. Garinis GA, Menounos PG, Spanakis NE, Papadopoulos K, Karavitis G, Parassi I, Christeli E, Patrinos GP, Manolis EN, Peros G. Hypermethylation-associated transcriptional silencing of E-cadherin in primary sporadic colorectal carcinomas. *J Pathol* 2002;**198**(4): 442-449.

222. Sunami E, de Maat M, Vu A, Turner RR, Hoon DS. LINE-1 hypomethylation during primary colon cancer progression. *PLoS One* 2011;**6**(4): e18884.

223. Bleul CC, Fuhlbrigge RC, Casasnovas JM, Aiuti A, Springer TA. A highly efficacious lymphocyte chemoattractant, stromal cell-derived factor 1 (SDF-1). *J Exp Med* 1996;**184**(3): 1101-1109.

224. Heidemann J, Ogawa H, Rafiee P, Lugering N, Maaser C, Domschke W, Binion DG, Dwinell MB. Mucosal angiogenesis regulation by CXCR4 and its ligand CXCL12 expressed by human intestinal microvascular endothelial cells. *Am J Physiol Gastrointest Liver Physiol* 2004;**286**(6): G1059-1068.

225. Rabizadeh S, Oh J, Zhong LT, Yang J, Bitler CM, Butcher LL, Bredesen DE. Induction of apoptosis by the low-affinity NGF receptor. *Science* 1993;**261**(5119): 345-348.

226. Bredesen DE, Mehlen P, Rabizadeh S. Receptors that mediate cellular dependence. *Cell Death Differ* 2005;**12**(8): 1031-1043.

227. Hibi K, Mizukami H, Shirahata A, Goto T, Sakata M, Sanada Y. Aberrant methylation of the netrin-1 receptor genes UNC5C and DCC detected in advanced colorectal cancer. *World J Surg* 2009;**33**(5): 1053-1057.

228. Shiozaki H, Oka H, Inoue M, Tamura S, Monden M. E-cadherin mediated adhesion system in cancer cells. *Cancer* 1996;**77**(8 Suppl): 1605-1613.

229. Machado JC, Soares P, Carneiro F, Rocha A, Beck S, Blin N, Berx G, Sobrinho-Simoes M. E-cadherin gene mutations provide a genetic basis for the phenotypic divergence of mixed gastric carcinomas. *Lab Invest* 1999;**79**(4): 459-465.

230. Berx G, Nollet F, van Roy F. Dysregulation of the E-cadherin/catenin complex by irreversible mutations in human carcinomas. *Cell Adhes Commun* 1998;**6**(2-3): 171-184.

231. de Maat MF, Umetani N, Sunami E, Turner RR, Hoon DS. Assessment of methylation events during colorectal tumor progression by absolute quantitative analysis of methylated alleles. *Mol Cancer Res* 2007;**5**(5): 461-471.

232. Toyota M, Ahuja N, Suzuki H, Itoh F, Ohe-Toyota M, Imai K, Baylin SB, Issa JP. Aberrant methylation in gastric cancer associated with the CpG island methylator phenotype. *Cancer Res* 1999;**59**(21): 5438-5442.

233. Tanemura A, Terando AM, Sim MS, van Hoesel AQ, de Maat MF, Morton DL, Hoon DS. CpG island methylator phenotype predicts progression of malignant melanoma. *Clin Cancer Res* 2009;**15**(5): 1801-1807.

234. Hoffmann MJ, Schulz WA. Causes and consequences of DNA hypomethylation in human cancer. *Biochem Cell Biol* 2005;**83**(3): 296-321.

235. Gilbert N, Lutz-Prigge S, Moran JV. Genomic deletions created upon LINE-1 retrotransposition. *Cell* 2002;**110**(3): 315-325.

236. Symer DE, Connelly C, Szak ST, Caputo EM, Cost GJ, Parmigiani G, Boeke JD. Human l1 retrotransposition is associated with genetic instability in vivo. *Cell* 2002;**110**(3): 327-338.

237. Kazazian HH, Jr., Goodier JL. LINE drive. retrotransposition and genome instability. *Cell* 2002;**110**(3): 277-280.

238. Deng G, Kakar S, Tanaka H, Matsuzaki K, Miura S, Sleisenger MH, Kim YS. Proximal and distal colorectal cancers show distinct gene-specific methylation profiles and clinical and molecular characteristics. *Eur J Cancer* 2008;**44**(9): 1290-1301.

239. Gustavsson E, Sernbo S, Andersson E, Brennan DJ, Dictor M, Jerkeman M, Borrebaeck CA, Ek S. SOX11 expression correlates to promoter methylation and regulates tumor growth in hematopoietic malignancies. *Mol Cancer* 2012;**9**: 187.

240. Wang BX, Yin BL, He B, Chen C, Zhao M, Zhang WX, Xia ZK, Pan YZ, Tang JQ, Zhou XM, Yin N. Overexpression of DNA damage-induced 45 alpha gene contributes to esophageal squamous cell cancer by promoter hypomethylation. *J Exp Clin Cancer Res* 2012;**31**(1): 11.

241. Suh ER, Ha CS, Rankin EB, Toyota M, Traber PG. DNA methylation down-regulates CDX1 gene expression in colorectal cancer cell lines. *J Biol Chem* 2002;**277**(39): 35795-35800.

242. Toyooka S, Toyooka KO, Miyajima K, Reddy JL, Toyota M, Sathyanarayana UG, Padar A, Tockman MS, Lam S, Shivapurkar N, Gazdar AF. Epigenetic down-regulation of death-associated protein kinase in lung cancers. *Clin Cancer Res* 2003;**9**(8): 3034-3041.

243. Alvarez C, Tapia T, Cornejo V, Fernandez W, Munoz A, Camus M, Alvarez M, Devoto L, Carvallo P. Silencing of tumor suppressor genes RASSF1A, SLIT2, and WIF1 by promoter hypermethylation in hereditary breast cancer. *Mol Carcinog* 2012.

244. Bueso-Ramos C, Xu Y, McDonnell TJ, Brisbay S, Pierce S, Kantarjian H, Rosner G, Garcia-Manero G. Protein expression of a triad of frequently methylated genes, p73, p57Kip2, and p15, has prognostic value in adult acute lymphocytic leukemia independently of its methylation status. *J Clin Oncol* 2005;**23**(17): 3932-3939.

245. Iwamoto M, Ahnen DJ, Franklin WA, Maltzman TH. Expression of beta-catenin and full-length APC protein in normal and neoplastic colonic tissues. *Carcinogenesis* 2000;**21**(11): 1935-1940.

246. Milne AN, Sitarz R, Carvalho R, Polak MM, Ligtenberg M, Pauwels P, Offerhaus GJ, Weterman MA. Molecular analysis of primary gastric cancer, corresponding xenografts, and 2 novel gastric carcinoma cell lines reveals novel alterations in gastric carcinogenesis. *Hum Pathol* 2007;**38**(6): 903-913.

247. Link BC, Reichelt U, Schreiber M, Kaifi JT, Wachowiak R, Bogoevski D, Bubenheim M, Cataldegirmen G, Gawad KA, Issa R, Koops S, Izbicki JR, Yekebas EF. Prognostic implications of netrin-1 expression and its receptors in patients with adenocarcinoma of the pancreas. *Ann Surg Oncol* 2007;**14**(9): 2591-2599.

248. Yan Q, Zhang ZF, Chen XP, Gutmann DH, Xiong M, Xiao ZY, Huang ZY. Reduced T-cadherin expression and promoter methylation are associated with the development and progression of hepatocellular carcinoma. *Int J Oncol* 2008;**32**(5): 1057-1063.

249. Caldeira JR, Prando EC, Quevedo FC, Neto FA, Rainho CA, Rogatto SR. CDH1 promoter hypermethylation and E-cadherin protein expression in infiltrating breast cancer. *BMC Cancer* 2006;**6**: 48.

250. Chen J, Rocken C, Lofton-Day C, Schulz HU, Muller O, Kutzner N, Malfertheiner P, Ebert MP. Molecular analysis of APC promoter methylation and protein expression in colorectal cancer metastasis. *Carcinogenesis* 2005;**26**(1): 37-43.

251. Sproul D, Nestor C, Culley J, Dickson JH, Dixon JM, Harrison DJ, Meehan RR, Sims AH, Ramsahoye BH. Transcriptionally repressed genes become aberrantly methylated and distinguish tumors of different lineages in breast cancer. *Proc Natl Acad Sci U S A* 2011;**108**(11): 4364-4369.

252. Ribeiro-Silva A, Zhang H, Jeffrey SS. RNA extraction from ten year old formalin-fixed paraffin-embedded breast cancer samples: a comparison of column purification and magnetic bead-based technologies. *BMC Mol Biol* 2007;**8**: 118.

253. Abramovitz M, Ordanic-Kodani M, Wang Y, Li Z, Catzavelos C, Bouzyk M, Sledge GW, Jr., Moreno CS, Leyland-Jones B. Optimization of RNA extraction from FFPE tissues for expression profiling in the DASL assay. *Biotechniques* 2008;**44**(3): 417-423.

254. Guo Y, Xiao P, Lei S, Deng F, Xiao GG, Liu Y, Chen X, Li L, Wu S, Chen Y, Jiang H, Tan L, Xie J, Zhu X, Liang S, Deng H. How is mRNA expression predictive for protein expression? A correlation study on human circulating monocytes. *Acta Biochim Biophys Sin (Shanghai)* 2008;**40**(5): 426-436.

255. Lee PS, Shaw LB, Choe LH, Mehra A, Hatzimanikatis V, Lee KH. Insights into the relation between mrna and protein expression patterns: II. Experimental observations in Escherichia coli. *Biotechnol Bioeng* 2003;**84**(7): 834-841.

256. Chen G, Gharib TG, Huang CC, Taylor JM, Misek DE, Kardia SL, Giordano TJ, Iannettoni MD, Orringer MB, Hanash SM, Beer DG. Discordant protein and mRNA expression in lung adenocarcinomas. *Mol Cell Proteomics* 2002;**1**(4): 304-313.

257. Caldwell GM, Jones CE, Taniere P, Warrack R, Soon Y, Matthews GM, Morton DG. The Wnt antagonist sFRP1 is downregulated in premalignant large bowel adenomas. *Br J Cancer* 2006;**94**(6): 922-927.

258. Kakar S, Deng G, Cun L, Sahai V, Kim YS. CpG island methylation is frequently present in tubulovillous and villous adenomas and correlates with size, site, and villous component. *Hum Pathol* 2008;**39**(1): 30-36.

259. Petko Z, Ghiassi M, Shuber A, Gorham J, Smalley W, Washington MK, Schultenover S, Gautam S, Markowitz SD, Grady WM. Aberrantly methylated CDKN2A, MGMT, and MLH1 in colon polyps and in fecal DNA from patients with colorectal polyps. *Clin Cancer Res* 2005;**11**(3): 1203-1209.

260. Suzuki H, Watkins DN, Jair KW, Schuebel KE, Markowitz SD, Chen WD, Pretlow TP, Yang B, Akiyama Y, Van Engeland M, Toyota M, Tokino T, Hinoda Y, Imai K, Herman JG, Baylin SB. Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat Genet* 2004;**36**(4): 417-422.

261. Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP. The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* 2009;**41**(2): 178-186.

262. Oster B, Thorsen K, Lamy P, Wojdacz TK, Hansen LL, Birkenkamp-Demtroder K, Sorensen KD, Laurberg S, Orntoft TF, Andersen CL. Identification and validation of highly frequent CpG island hypermethylation in colorectal adenomas and carcinomas. *Int J Cancer* 2011.

263. Estecio MR, Gharibyan V, Shen L, Ibrahim AE, Doshi K, He R, Jelinek J, Yang AS, Yan PS, Huang TH, Tajara EH, Issa JP. LINE-1 hypomethylation in cancer is highly

variable and inversely correlated with microsatellite instability. *PLoS One* 2007;**2**(5): e399.

264. Choi IS, Estecio MR, Nagano Y, Kim do H, White JA, Yao JC, Issa JP, Rashid A. Hypomethylation of LINE-1 and Alu in well-differentiated neuroendocrine tumors (pancreatic endocrine tumors and carcinoid tumors). *Mod Pathol* 2007;**20**(7): 802-810.

265. Nishioka H, Haraoka J. Significance of immunohistochemical expression of Rab3B and SNAP-25 in growth hormone-producing pituitary adenomas. *Acta Neuropathol* 2005;**109**(6): 598-602.

266. Minard ME, Kim LS, Price JE, Gallick GE. The role of the guanine nucleotide exchange factor Tiam1 in cellular migration, invasion, adhesion and tumor progression. *Breast Cancer Res Treat* 2004;**84**(1): 21-32.

267. Minard ME, Ellis LM, Gallick GE. Tiam1 regulates cell adhesion, migration and apoptosis in colon tumor cells. *Clin Exp Metastasis* 2006;**23**(5-6): 301-313.

268. Adam L, Vadlamudi RK, McCrea P, Kumar R. Tiam1 overexpression potentiates heregulin-induced lymphoid enhancer factor-1/beta -catenin nuclear signaling in breast cancer cells by modulating the intercellular stability. *J Biol Chem* 2001;**276**(30): 28443-28450.

269. Engers R, Mueller M, Walter A, Collard JG, Willers R, Gabbert HE. Prognostic relevance of Tiam1 protein expression in prostate carcinomas. *Br J Cancer* 2006;**95**(8): 1081-1086.

270. Jin H, Li T, Ding Y, Deng Y, Zhang W, Yang H, Zhou J, Liu C, Lin J. Methylation status of T-lymphoma invasion and metastasis 1 promoter and its overexpression in colorectal cancer. *Hum Pathol* 2011;**42**(4): 541-551.

271. Takash W, Canizares J, Bonneaud N, Poulat F, Mattei MG, Jay P, Berta P. SOX7 transcription factor: sequence, chromosomal localisation, expression, transactivation and interference with Wnt signalling. *Nucleic Acids Res* 2001;**29**(21): 4274-4283.

272. Guo L, Zhong D, Lau S, Liu X, Dong XY, Sun X, Yang VW, Vertino PM, Moreno CS, Varma V, Dong JT, Zhou W. Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells. *Mol Cancer Res* 2008;**6**(9): 1421-1430.

273. Zhang Y, Huang S, Dong W, Li L, Feng Y, Pan L, Han Z, Wang X, Ren G, Su D, Huang B, Lu J. SOX7, down-regulated in colorectal cancer, induces apoptosis and inhibits proliferation of colorectal cancer cells. *Cancer Lett* 2009;**277**(1): 29-37.

274. Weber M, Hellmann I, Stadler MB, Ramos L, Paabo S, Rebhan M, Schubeler D. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet* 2007;**39**(4): 457-466.

275. Estecio MR, Yan PS, Ibrahim AE, Tellez CS, Shen L, Huang TH, Issa JP. High-throughput methylation profiling by MCA coupled to CpG island microarray. *Genome Res* 2007;**17**(10): 1529-1536.

276. Hinoue T, Weisenberger DJ, Lange CP, Shen H, Byun HM, Van Den Berg D, Malik S, Pan F, Noushmehr H, van Dijk CM, Tollenaar RA, Laird PW. Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res* 2011;**22**(2): 271-282.

277. Chen SP, Chiu SC, Wu CC, Lin SZ, Kang JC, Chen YL, Lin PC, Pang CY, Harn HJ. The association of methylation in the promoter of APC and MGMT and the prognosis of Taiwanese CRC patients. *Genet Test Mol Biomarkers* 2009;**13**(1): 67-71.

278. Kim YT, Park JY, Jeon YK, Park SJ, Song JY, Kang CH, Sung SW, Kim JH. Aberrant promoter CpG island hypermethylation of the adenomatosis polyposis coli

gene can serve as a good prognostic factor by affecting lymph node metastasis in squamous cell carcinoma of the esophagus. *Dis Esophagus* 2009;**22**(2): 143-150.

279. Brakensiek K, Wingen LU, Langer F, Kreipe H, Lehmann U. Quantitative high-resolution CpG island mapping with Pyrosequencing reveals disease-specific methylation patterns of the CDKN2B gene in myelodysplastic syndrome and myeloid leukemia. *Clin Chem* 2007;**53**(1): 17-23.

280. Mikeska T, Bock C, El-Maarri O, Hubner A, Ehrentraut D, Schramm J, Felsberg J, Kahl P, Buttner R, Pietsch T, Waha A. Optimization of quantitative MGMT promoter methylation analysis using pyrosequencing and combined bisulfite restriction analysis. *J Mol Diagn* 2007;**9**(3): 368-381.

281. Juhlin CC, Kiss NB, Villablanca A, Haglund F, Nordenstrom J, Hoog A, Larsson C. Frequent promoter hypermethylation of the APC and RASSF1A tumour suppressors in parathyroid tumours. *PLoS One* 2010;**5**(3): e9472.

282. Dejeux E, El abdalaoui H, Gut IG, Tost J. Identification and quantification of differentially methylated loci by the pyrosequencing technology. *Methods Mol Biol* 2009;**507**: 189-205.

283. Cheng YW, Shawber C, Notterman D, Paty P, Barany F. Multiplexed profiling of candidate genes for CpG island methylation status using a flexible PCR/LDR/Universal Array assay. *Genome Res* 2006;**16**(2): 282-289.

284. Uhlmann K, Brinckmann A, Toliat MR, Ritter H, Nurnberg P. Evaluation of a potential epigenetic biomarker by quantitative methyl-single nucleotide polymorphism analysis. *Electrophoresis* 2002;**23**(24): 4072-4079.

285. Dupont JM, Tost J, Jammes H, Gut IG. De novo quantitative bisulfite sequencing using the pyrosequencing technology. *Anal Biochem* 2004;**333**(1): 119-127.

286. Xu Q, Ma JZ, Payne TJ, Li MD. Determination of Methylated CpG Sites in the Promoter Region of Catechol-O-Methyltransferase (COMT) and their Involvement in the Etiology of Tobacco Smoking. *Front Psychiatry* 2010;**1**: 16.

287. Palmieri C, Monteverde M, Lattanzio L, Gojis O, Rudraraju B, Fortunato M, Syed N, Thompson A, Garrone O, Merlano M, Lo Nigro C, Crook T. Site-specific CpG methylation in the CCAAT/enhancer binding protein delta (CEBPdelta) CpG island in breast cancer is associated with metastatic relapse. *Br J Cancer* 2012;**107**(4): 732-738.

288. Fustinoni S, Rossella F, Polledri E, Bollati V, Campo L, Byun HM, Agnello L, Consonni D, Pesatori AC, Baccarelli A, Bertazzi PA. Global DNA methylation and low-level exposure to benzene. *Med Lav* 2012;**103**(2): 84-95.

289. Matsuda Y, Yamashita S, Lee YC, Niwa T, Yoshida T, Gyobu K, Igaki H, Kushima R, Lee S, Wu MS, Osugi H, Suehiro S, Ushijima T. Hypomethylation of Alu repetitive elements in esophageal mucosa, and its potential contribution to the epigenetic field for cancerization. *Cancer Causes Control* 2012;**23**(6): 865-873.

290. Bae JM, Shin SH, Kwon HJ, Park SY, Kook MC, Kim YW, Cho NY, Kim N, Kim TY, Kim D, Kang GH. ALU and LINE-1 hypomethylations in multistep gastric carcinogenesis and their prognostic implications. *Int J Cancer* 2012;**131**(6): 1323-1331.

291. Kim YH, Petko Z, Dzieciatkowski S, Lin L, Ghiassi M, Stain S, Chapman WC, Washington MK, Willis J, Markowitz SD, Grady WM. CpG island methylation of genes

accumulates during the adenoma progression step of the multistep pathogenesis of colorectal cancer. *Genes Chromosomes Cancer* 2006;**45**(8): 781-789.

292. Kim HC, Roh SA, Ga IH, Kim JS, Yu CS, Kim JC. CpG island methylation as an early event during adenoma progression in carcinogenesis of sporadic colorectal cancer. *J Gastroenterol Hepatol* 2005;**20**(12): 1920-1926.

293. Bai AH, Tong JH, To KF, Chan MW, Man EP, Lo KW, Lee JF, Sung JJ, Leung WK. Promoter hypermethylation of tumor-related genes in the progression of colorectal neoplasia. *Int J Cancer* 2004;**112**(5): 846-853.

294. Kim HK, Kim J, Korolevich S, Choi IJ, Kim CH, Munroe DJ, Green JE. Distinctions in gastric cancer gene expression signatures derived from laser capture microdissection versus histologic macrodissection. *BMC Med Genomics* 2011;**4**: 48.

295. Inoue Y, Tanaka K, Yokoe T, Saigusa S, Toiyama Y, Miki C, Kusunoki M. Microdissection is essential for gene expression analysis of irradiated rectal cancer tissues. *Oncol Rep* 2009;**22**(4): 901-906.

296. Irahara N, Nosho K, Baba Y, Shima K, Lindeman NI, Hazra A, Schernhammer ES, Hunter DJ, Fuchs CS, Ogino S. Precision of pyrosequencing assay to measure LINE-1 methylation in colon cancer, normal colonic mucosa, and peripheral blood cells. *J Mol Diagn* 2010;**12**(2): 177-183.

297. Nosho K, Shima K, Irahara N, Kure S, Firestein R, Baba Y, Toyoda S, Chen L, Hazra A, Giovannucci EL, Fuchs CS, Ogino S. SIRT1 histone deacetylase expression is associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. *Mod Pathol* 2009;**22**(7): 922-932.

298. Hahn MA, Hahn T, Lee DH, Esworthy RS, Kim BW, Riggs AD, Chu FF, Pfeifer GP. Methylation of polycomb target genes in intestinal cancer is mediated by inflammation. *Cancer Res* 2008;**68**(24): 10280-10289.

299. Zhang T, Kraus WL. SIRT1-dependent regulation of chromatin and transcription: linking NAD(+) metabolism and signaling to the control of cellular functions. *Biochim Biophys Acta* 2010;**1804**(8): 1666-1675.

300. Paul TA, Bies J, Small D, Wolff L. Signatures of polycomb repression and reduced H3K4 trimethylation are associated with p15INK4b DNA methylation in AML. *Blood* 2010;**115**(15): 3098-3108.

301. Andreyev HJ, Norman AR, Cunningham D, Oates J, Dix BR, Iacopetta BJ, Young J, Walsh T, Ward R, Hawkins N, Beranek M, Jandik P, Benamouzig R, Jullian E, Laurent-Puig P, Olschwang S, Muller O, Hoffmann I, Rabes HM, Zietz C, Troungos C, Valavanis C, Yuen ST, Ho JW, Croke CT, O'Donoghue DP, Giaretti W, Rapallo A, Russo A, Bazan V, Tanaka M, Omura K, Azuma T, Ohkusa T, Fujimori T, Ono Y, Pauly M, Faber C, Glaesener R, de Goeij AF, Arends JW, Andersen SN, Lovig T, Breivik J, Gaudernack G, Clausen OP, De Angelis PD, Meling GI, Rognum TO, Smith R, Goh HS, Font A, Rosell R, Sun XF, Zhang H, Benhattar J, Losi L, Lee JQ, Wang ST, Clarke PA, Bell S, Quirke P, Bubb VJ, Piris J, Cruickshank NR, Morton D, Fox JC, Al-Mulla F, Lees N, Hall CN, Snary D, Wilkinson K, Dillon D, Costa J, Pricolo VE, Finkelstein SD, Thebo JS, Senagore AJ, Halter SA, Wadler S, Malik S, Krtolica K, Urosevic N. Kirsten ras mutations in patients with colorectal cancer: the 'RASCAL II' study. *Br J Cancer* 2001;**85**(5): 692-696.

302. Guerrero S, Casanova I, Farre L, Mazo A, Capella G, Mangues R. K-ras codon 12 mutation induces higher level of resistance to apoptosis and predisposition to

anchorage-independent growth than codon 13 mutation or proto-oncogene overexpression. *Cancer Res* 2000;**60**(23): 6750-6756.

303. Roth AD, Tejpar S, Delorenzi M, Yan P, Fiocca R, Klingbiel D, Dietrich D, Biesmans B, Bodoky G, Barone C, Aranda E, Nordlinger B, Cisar L, Labianca R, Cunningham D, Van Cutsem E, Bosman F. Prognostic role of KRAS and BRAF in stage II and III resected colon cancer: results of the translational study on the PETACC-3, EORTC 40993, SAKK 60-00 trial. *J Clin Oncol* 2010;**28**(3): 466-474.

304. Farina-Sarasqueta A, van Lijnschoten G, Moerland E, Creemers GJ, Lemmens VE, Rutten HJ, van den Brule AJ. The BRAF V600E mutation is an independent prognostic factor for survival in stage II and stage III colon cancer patients. *Ann Oncol* 2010;**21**(12): 2396-2402.

305. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. *J Clin Oncol* 2005;**23**(3): 609-618.

306. Maas M, Nelemans PJ, Valentini V, Das P, Rodel C, Kuo LJ, Calvo FA, Garcia-Aguilar J, Glynne-Jones R, Haustermans K, Mohiuddin M, Pucciarelli S, Small W, Jr., Suarez J, Theodoropoulos G, Biondo S, Beets-Tan RG, Beets GL. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. *Lancet Oncol* 2010;**11**(9): 835-844.

307. Liu ZG, Chen HY, Cheng JJ, Chen ZP, Li XN, Xia YF. Relationship between methylation status of ERCC1 promoter and radiosensitivity in glioma cell lines. *Cell Biol Int* 2009;**33**(10): 1111-1117.

308. Kim EH, Park AK, Dong SM, Ahn JH, Park WY. Global analysis of CpG methylation reveals epigenetic control of the radiosensitivity in lung cancer cell lines. *Oncogene* 2010;**29**(33): 4725-4731.