Carcinogenesis vol.31 no.6 pp.1158–1163, 2010
doi:10.1093/carcin/bgq077
Advance Access publication April 15, 2010

Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa

Nigel J.Belshaw*, Nandita Pal[1], Henri S.Tapp, Jack R.Dainty, Michael P.N.Lewis[1], Mark R.Williams[2], Elizabeth K.Lund and Ian T.Johnson

Institute of Food Research, Colney Lane, Norwich NR4 7UA, UK, [1]Department of Surgery, Norfolk and Norwich University Hospital, Colney Lane, Norwich, NR4 7UY, UK and [2]School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK

*To whom correspondence should be addressed. Tel: +44 1603 255237; Fax: +44 1603 507723;
Email: nigel.belshaw@bbsrc.ac.uk

Methylation of CpG islands (CGIs) in the promoter regions of tumour suppressor genes is common in colorectal cancer and occurs also in an age-dependent manner in the morphologically normal colorectal mucosa. In this study, we quantified the level of methylation of six genes associated with the Wnt signalling pathway (*adenomatous polyposis coli*, *DKK1*, *WIF1*, *SFRP1*, *SFRP2* and *SFRP5*) together with long-interspersed nuclear element-1 as a surrogate for global methylation. DNA methylation was analysed in 260 individual colorectal crypts obtained from eight female patients with no evidence of colorectal disease and five with colorectal cancer. Significant variation in methylation levels for each of the six genes existed between crypts from the same biopsy. The variation in both global and gene-specific CGI methylation between crypts from the same individual was significantly less than that between individuals. Bisulphite sequencing provided insight into the mechanism of aberrant methylation showing that CGI methylation occurs in an 'all-or-none' manner by the directional spreading of methylation from further upstream. Univariate statistical analyses revealed that there were significant differences in crypt-specific methylation associated with both aging and disease status. A multivariate statistical modelling approach was able to distinguish both subject age and health status based on crypt-specific methylation profiles. Our results indicate that the differential methylation of genes associated with the Wnt signalling pathway affecting individual morphologically normal crypts may contribute to the age-dependent generation of the colonic field defect and, in combination with mutations, to the stepwise development of colorectal neoplasia.

Introduction

Armitage and Doll (1) were the first to propose a multistage theory of carcinogenesis, in which epithelial cancers develop as a consequence of the progressive acquisition of mutations by a small population of asymmetrically dividing stem cells. Colorectal carcinoma is among the most common and thoroughly studied of human cancers, and it is well established that the emergence of both precancerous polyps and malignant tumours is associated with the acquisition of mutations affecting regulatory pathways involved in the growth and development of crypt epithelial cells. Robust models for the contribution of such mutations to the adenoma–carcinoma phenotype have emerged and continue to be refined (2,3). A further development in our understanding of the molecular origins of cancer has come with the realization that heritable epigenetic abnormalities are also common both in tumour cells and in early precancerous lesions *in vivo*. It is well established that aberrant CpG island (CGI) methylation occurs in colorectal carcinomas and that these abnormalities can silence gene expression, but the precise relationships, both temporal and functional, between these epigenetic events and the acquisition of somatic mutations are poorly understood (4). It has been proposed for example both that cancer progenitor cells become, in a sense, 'addicted' to defective genetic pathways associated with epigenetic defects (5) and that aberrant CGI methylation of certain genes can itself serve as the primary trigger for the emergence of mutations (4). In either case, there is now ample evidence that epigenetic events occur at a very early stage in tumour development.

In a previous study, we measured levels of CGI methylation for 18 relevant genes in the morphologically normal mucosa of patients with or without colorectal tumours and presented evidence for subtle changes in the methylation status of the mucosal field in patients harbouring a tumour (6). However, one problem with this approach is that it fails to take account of the fact that the mucosa is a complex tissue comprising epithelial and mesenchymal cell types. The epithelium is organized into a continuous array of glandular crypts, each of which is a discrete proliferative unit containing a small number of stem cells. There is strong evidence that individual tumours develop from a single crypt (7). However, for any CGI, the methylation value derived from a whole mucosal biopsy is an average for DNA derived from several hundred different crypts. Moreover, the DNA from the epithelial cells in such a sample will be 'diluted' by DNA from the range of other mucosal cell types. To gain a better understanding of the spatial distribution and possible functional significance of aberrant CGI methylation in the mucosal field, it is of great interest to compare the methylation levels of individual crypts. In the present study, we isolated crypts from female patients with and without colorectal cancer and used a quantitative assay combined bisulphite and restriction analysis to measure CGI methylation for six genes associated with the Wnt signalling pathway, as well as long-interspersed nuclear element-1 (LINE-1) as a surrogate for global methylation.

The Wnt signalling pathway is critical for the regulation of crypt renewal and homeostasis, and abnormalities affecting it are now known to lie at the heart of the deregulation of crypt homeostasis that initiates colorectal carcinogenesis (8,9). In the normal mucosa, the adenomatous polyposis coli (APC) protein binds the transcription factor β-catenin in the cytoplasm and facilitates its ubiquitin-mediated destruction. The absence of functional APC causes increased expression of β-catenin regulated genes and upregulates crypt cell proliferation. Loss of both *APC* alleles is thought to be sufficient to facilitate the growth of an adenoma (10). Secreted frizzled-related proteins are a family of secreted inhibitors of the Frizzled receptor, which activates Wnt signalling. Epigenetic silencing of *SFRP* genes has been observed in a variety of cancers (11), including colorectal cancer (12). Similarly, the *WIF1* and *DKK1* products, Wnt inhibitory factor-1 and Dickkopf-1, respectively, are secreted antagonists that bind to and inhibit Wnt proteins directly. Epigenetic silencing of both *WIF1* and *DKK1* is implicated in several types of gastrointestinal and other cancers (13,14), and interactions between epigenetic and genetic silencing events affecting the Wnt pathway are likely to be crucial for the initiation of the adenoma–carcinoma sequence (15).

Abbreviations: APC, adenomatous polyposis coli; CGI, CpG island; LINE-1, long-interspersed nuclear element-1; PBS, phosphate-buffered saline; PCR, polymerase chain reaction; PLS, partial least squares; PLS-LDA, partial least squares–linear discriminant analysis.

Materials and methods

Tissue collection and crypt isolation

Volunteers were recruited from the gastroenterology outpatient and surgical lists of the Norfolk and Norwich University Hospital, Norwich, Norfolk, UK. Female volunteers were chosen to avoid gender biases in subsequent statistical analyses, and all were either patients with previously diagnosed colorectal cancer or outpatients with no known gastrointestinal pathology, who were

undergoing colonoscopy as a diagnostic procedure, typically for investigation of non-specific symptoms, such as abnormal bowel habit or unexplained rectal bleeding. Ethical approval for the project was received from the Norfolk Research Ethics Committee (Project reference 97/124) and consent was obtained in advance of endoscopy or surgery. Experimental biopsies were collected from the colonic mucosa of endoscopy patients, and for cancer patients, samples of normal mucosa (>10 cm from tumour margin) were collected in theatre, immediately after surgery. Samples of mucosa were collected from 13 female patients at either colonoscopy (8 patients without any gross morphological pathology) or during resection of large bowel tumour (5 patients) into ~1 ml phosphate-buffered saline (PBS). All tissue samples were obtained from the distal sigmoid/rectosigmoid colon.

Crypts were isolated by an adaptation of the method described by Whitehead *et al.* (16). In brief, individual biopsies were washed two further times in PBS before incubation with gentle mixing in 1 ml PBS containing 3 mM ethylenediaminetetraacetic acid for 60 min. They were then removed and placed in PBS and shaken vigorously to allow detachment of the crypts. Following removal of any 'large' pieces of biopsy tissue, individual normal-appearing cylindrical crypts were collected under a dissecting microscope (Leitz Wetzler—×4 objective, Leica Microsystems Ltd, Milton Keynes, UK). Aberrant crypt foci and branched crypts were not collected. Twenty individual crypts were used from a single biopsy from each subject. The length of each crypt was measured against a stage graticule (scale 100 in 0.01 mm; Graticules Ltd, Edenbridge, Kent, UK).

DNA methylation analyses

DNA was extracted from individual crypts by treatment with proteinase K in lysis buffer (0.2 M Tris–HCl, pH8.5, 0.25 M NaCl, 25 mM ethylenediaminetetraacetic acid and 0.5% sodium dodecyl sulphate) at 60°C for 2 h. Carrier DNA (1 μg salmon sperm DNA) was added prior to ethanol precipitation. The extracted crypt DNA was bisulphite modified as described previously (17). The CGI methylation status of the genes *APC*, *DKK1*, *WIF1*, *SFRP1*, *SFRP2* and *SFRP5* were determined using combined bisulphite and restriction analysis assays (18). CGI fragments were amplified from bisulphite-modified crypt DNA using gene-specific primers as described previously (6). Polymerase chain reaction (PCR) products were incubated with restriction enzymes (Sau3AI for *APC*, FauI for *SFRP1* and *DKK1*, HinfI for *SFRP2* and *WIF1* and BsiEI for *SFRP5*) according to manufacturer's (New England Biolabs Ltd, Hitchin, UK) instructions prior to electrophoresis in 6% polyacrylamide gel electrophoresis gels in 1× tris-borate-ethylenediaminetetraacetic acid buffer. Gels were incubated in 1× tris-borate-ethylenediaminetetraacetic acid buffer containing SYBR green II (Invitrogen Ltd, Paisley, UK) and fluorescent DNA fragments visualized at 530 nm following excitation at 488 nm using a Pharos FXplus molecular imager (Bio-Rad Laboratories Ltd, Hemel Hempstead, UK). Band intensities were quantified using TotalLab software (Nonlinear Dynamics, Newcastle upon Tyne, UK). LINE-1 methylation, a surrogate for global DNA methylation status (19), was measured using a quantitative PCR approach described previously (20).

For bisulphite sequencing, the *WIF1* CGI was PCR amplified from bisulphite-modified DNA from six randomly selected crypts from five randomly selected subjects and cloned (Topo-TA cloning kit; Invitrogen). Multiple (four to eight) clones from each crypt were sequenced. The methylation patterns of the 40 CpG sites per amplicon were considered as a 40 digit binary string representing the presence or absence of methylation at each CpG and analysed for diversity (epigenetic distance between molecules) between individual crypts as described by Yatabe *et al.* (21). A Hamming distance matrix was calculated to allow comparisons of the mean intra- and inter-crypt distance for each subject and comparison of the mean intra- and inter-subject distance. Significance levels were computed from permutation tests using 100 000 permutations.

Statistical analyses

The effect on DNA methylation of 'group' (fixed factor), 'age' (covariate) and 'subject' (random factor) were examined using mixed-effects models. The models took into account the non-independence of each individual's crypt-specific data and allowed for the lack of balance in the dataset to be correctly analysed. Data transformations ($y \rightarrow y^{\lambda}$; $0.1 \leq \lambda \leq 0.5$) were necessary to satisfy the statistical assumptions that mixed-effects models are based on. To search for differences in patterns of DNA methylation among the patient groups, we used binomial logistic regression models fitted by setting the response variable as patient group (normal or cancer) and initially having the measured global (LINE-1) and gene-specific CGI (*APC*, *DKK1*, *WIF1*, *SFRP1*, *SFRP2* and *SFRP5*) methylation status as well as subject age as the predictor variables. These variables were then 'pruned' using both backwards elimination using analysis of variance-type tests and an automated stepwise procedure for optimizing the Akaike Information Criterion. 'Leave-one-out' cross-validation was used to estimate the classification error rate. This was compared with the expected rate given by the proportional chance criterion using an exact binomial test (one sided), to test the null hypothesis that the given success rate of classification was no better than chance. These statistical analyses were carried out using 'R' (R Core Development team, http://www.R-project.org).

All multivariate modelling was performed using Matlab (Mathworks Ltd, Cambridge, UK). Age was treated as a categorical variable using a young–old threshold of 70 years old. Four-group classification (young-normal, young-cancer, old-normal and old-cancer) was modelled using partial least squares–linear discriminant analysis (PLS-LDA) (22) using unit-variance scaling. Single cross-validation (23) assignment success rates were calculated for both the overall classification success rate (four groups) and in terms of the predictions success in classifying age or health status individually. Significance levels were calculated using *y*-scrambling permutation tests (23). Models were visualized using biplots (24) of the coefficients of the first two partial least squares (PLS) 2 (25) loading vectors 'W' and their corresponding non-orthogonal scores 'T'. Regression vectors and corresponding $y = 0$ classification boundaries were calculated by considering the analogous problem of using PLS2 regression for simultaneously predicting two two-class categorical variables. Double cross-validation (23), using permutations of the membership of leave-block-out cross-validation segments was used to assess the ability of the four-group classification models to generalize to new data originating from either within the current cohort or from new individuals. Univariate analysis of the correlation matrix (constructed from pairwise comparisons of each crypt's methylation profile) compared within-individual with within-group correlations using two tests: Wilcoxon rank sum tests for equal medians and two-sample Kolmogorov–Smirnov goodness-of-fit hypothesis tests.

Results

We isolated 20 crypts per subject from biopsies obtained from 13 women, 8 of whom had no identified pathology at endoscopy and 5 had colorectal cancer. Analyses of global and gene-specific CGI methylation by LINE-1 quantitative PCR and combined bisulphite

Table I. Crypt lengths, gene-specific CGI and global (LINE-1) methylation in crypts from the eight disease-free (N1–N8) and five cancer (C1–C5) subjects

Subject	Age (years)	Crypt length (mm)	LINE-1 (%)	*APC* (%)	*DKK1* (%)	*WIF1* (%)	*SFRP1* (%)	*SFRP2* (%)	*SFRP5* (%)
N1	63	0.3 (0.3–0.4)	75.8 (68.4–93.3)	0 (0–15)	0 (0)	28 (0–58)	27 (2–67)	42 (0–58)	0 (0–48)
N2	80	0.5 (0.4–0.7)	68.8 (58.1–74.8)	0 (0–58)	0 (0)	18 (0–31)	10 (0–65)	2 (0–46)	35 (0–85)
N3	72	0.4 (0.3–0.5)	57.8 (49.2–67.1)	0 (0–34)	0 (0–100)	12 (0–59)	62 (0–100)	10 (0–58)	0 (0–100)
N4	77	0.4 (0.2–0.5)	67.4 (58.6–77.6)	4 (0–52)	0 (0–54)	30 (0–59)	0 (0–39)	0 (0–54)	0 (0)
N5	67	0.3 (0.2–0.4)	83.6 (75.9–89.3)	34 (6–72)	0 (0)	0 (0–86)	77 (61–100)	0 (0–100)	0 (0–21)
N6	60	0.3 (0.3–0.5)	76.5 (71.3–86.0)	0 (0–6)	0 (0)	16 (0–36)	62 (28–68)	18 (3–58)	0 (0–18)
N7	79	0.5 (0.4–0.5)	70.6 (63.3–79.7)	25 (0–50)	0 (0–80)	29 (5–69)	49 (14–68)	38 (0–67)	61 (5–93)
N8	76	0.5 (0.3–0.5)	61.4 (51.8–65.4)	0 (0–8)	0 (0–43)	78 (45–100)	0 (0–28)	0 (0–30)	14 (0–100)
C1	66	N.D.	68.8 (60.2–75.7)	6 (0–26)	0 (0–36)	34 (11–53)	1 (0–26)	30 (6–65)	0 (0–14)
C2	80	N.D.	66.3 (58.2–74.7)	30 (0–100)	0 (0–74)	35 (0–100)	28 (0–76)	2 (0–28)	0 (0–56)
C3	77	0.5 (0.3–0.7)	51.8 (39.9–68.2)	3 (0–9)	0 (0–100)	27 (0–63)	0 (0–100)	7 (0–30)	0 (0–13)
C4	60	0.4 (0.3–0.6)	59.5 (36.1–69.6)	32 (0–74)	34 (0–100)	89 (0–100)	87 (38–91)	21 (0–34)	0 (0–100)
C5	65	0.6 (0.4–0.7)	70.8 (62.7–81.7)	0 (0–15)	0 (0–69)	14 (0–26)	15 (2–62)	23 (0–56)	4 (0–55)

Values are median levels and ranges from 20 crypts from each subject. N.D., not determined.

and restriction analysis, respectively, were carried out on all 260 individual crypts. In some cases, one or more analyses were lost, usually because of a failure of PCR amplification. Methylation data for all six genes and LINE-1 were available for a total of 201 individual crypts and are summarized in Table I. For all six genes, there was significant within-subject variation in methylation levels between crypts. Table II summarizes the results from the univariate analysis carried out with the mixed-effects model. Both *SFRP1* and *SFRP2* showed decreased methylation with subject age ($P = 0.0449$ and $P = 0.0363$, respectively). The methylation status of the other genes was not significantly associated with age in these subjects. LINE-1 methylation decreased significantly in the cancer patients relative to the healthy subjects (63.4 ± 0.9% versus 70.4 ± 0.8%, $P = 0.0477$). The lengths of each crypt, shown in Table I, were positively associated with subject age ($P = 0.0411$) and were significantly higher in the cancer patients than in the subjects free of disease (0.507 ± 0.014 mm versus 0.404 ± 0.007 mm, $P = 0.0069$).

The nature of the within-subject variation in crypt methylation was investigated further by using bisulphite sequencing of *WIF1* CGI amplicons from a random selection of crypts from several subjects. The PCR amplified the region from −436 to −21 bp upstream of the transcription start site and consisted of 40 CpG sites. Of particular note was the bimodal distribution of the methylation status of individual amplicons with 36% containing up to four methylated sites and 33% with 36 or more methylated sites (Figure 1) demonstrating the 'all-or-none' methylation of this gene that we observed for *ESR1* previously (26). Also of interest was the gradient of preferential methylation from upstream towards the transcription start site (Figure 2) suggesting a directional 'spreading' of methylation from further upstream.

Bisulphite sequencing has been used previously to explore both clonality within crypts and relatedness between them (21). In the present study, most crypts contained both hypomethylated and hypermethylated *WIF1* sequences. The crypts were quasi-clonal with respect to their *WIF1* methylation profiles containing similar numbers of unique profiles per crypt and there were no significant differences between average numbers of unique profiles between individuals (data not shown). Hamming distances between the methylation profiles of individual pairs of molecules were calculated as the number of differences at each CpG site. Mean intra-crypt distances were significantly ($P < 0.05$) less than the mean inter-crypt distances for all five subjects analysed and the mean intra-subject distance was significantly ($P < 1 \times 10^{-5}$) less than the mean inter-subject distance. These results indicate that methylation patterns are more similar within crypts than between them and that crypts from the same individual are more related than those between individuals. This is entirely consistent with the observations and conclusions of Yatabe *et al.* (21).

To explore the relationship between crypt methylation and the presence or absence of colorectal cancer in more depth, logistic regression models were constructed, with disease status as the predicted variable. The predictive variables included in the initial model were percent methylation for each of the six genes and LINE-1 together with subject age. Using this approach, age together with methylation of LINE-1, *APC*, *SFRP1* and *SFRP5* were selected as significant discriminatory variables. However, age in this study is essentially a reflection of recruitment bias, caused by the fact that younger patients presenting for routine endoscopy are less likely to have cancer than older patients. Excluding age from the model gave LINE-1, *APC* and *WIF1* methylation as significant predictors of the group from which the crypt was obtained, and the model correctly identified 87.3% of the crypts from normal subjects ($P < 0.0001$) but only 54.2% from cancer patients (not significant).

Further multivariate statistical analysis of the crypt-specific methylation profiles, treating age as a categorical variable, with 'young' defined as <70 years, required balancing of the dataset by the random removal of crypt-specific data from the analysis. This strategy enabled us to define four separate groups based on age and health status, each with similar numbers of crypts. PLS-LDA, using leave-one-out cross-validation and unit-variance scaling with this balanced dataset, achieved high assignment success rates using only two PLS factors: 75.0% for classifying the disease status of the individual from which each crypt originated, 80.1% for classifying subject age and 62.8% for classifying both age and health status (Figure 3). Based on *y* scrambling using 2000 permutations, all three success rates are significant ($P < 5 \times 10^{-4}$). In Figure 3, the projections of the loadings onto the regression line are proportional to the regression coefficients for each variable. For classifying health status (Figure 3B), this indicates that the classification of cancer subjects is associated with increased methylation rates of *APC*, *DKK1* and *WIF1*, whereas the classification of healthy subjects is associated with higher methylation rates for *SFRP5* and LINE-1. Similarly, the methylation rates of *SFRP1* and *SFRP2* have only a small direct influence in the classification of health status. Figure 3A illustrates that the vast majority of the crypts misclassified on the basis of age were either 'young with cancer' misclassified as 'old' or 'old without cancer' misclassified as young. Similarly, those misclassified on the basis of health status (Figure 3B) tended to be either old without cancer classified as cancer or young with cancer classified as 'healthy'. These results demonstrate the complex interrelationship that exists between aging and cancer in driving aberrant DNA methylation.

PLS-LDA using leave-one-out cross-validation treats the data from individual crypts as independent. The bisulphite sequencing anaylsis of *WIF1* methylation strongly suggests that this may not be a valid approach. To explore this further, a correlation matrix was calculated for all the methylation data by pairwise calculations of the correlation between the seven methylation values for one crypt with the seven values in each other crypt (supplementary Figure 1 is available at *Carcinogenesis* Online). Two statistical tests were used to assess whether the within-subject correlations were similar to the within-group correlations. In both cases, significant differences were found, suggesting that leave-subject-out cross-validation should be used to assess the classification assignment success rate. This indicates that although there was considerable variation in the levels of methylation for each gene in crypts from the same individual, this variation was significantly lower than that between subjects and that the conclusions from bisulphite sequencing regarding the relatedness of crypt-specific *WIF1* methylation can probably be expanded to include the other genes. In other words, despite the evident diversity, these middle-aged and elderly individuals have acquired their own individual patterns of crypt methylation.

Applying PLS-LDA using leave-subject-out cross-validation reduced the assignment success rates. The optimum model used three PLS factors and achieved rates of 59.6% for classifying health status ($P = 0.016$), 60.3% for classifying age ($P = 0.0105$) and 39.1% for classifying both age and health status ($P = 0.0002$). These success rates are statistically significant even after applying a Bonferroni multiple test correction. Here, the significance levels were based on *y*-scrambling using 10 000 permutations. One plausible weakness of this analysis using single cross-validation is that it may introduce bias because the same data used to determine the optimum model are also

Table II. *P*-values from the univariate analysis of differences in DNA methylation and crypt length between groups (normal versus cancer) and with age using mixed-effects models

	n	Group (*P*-value)	Age (*P*-value)
LINE-1	251	0.0477	NS
APC	259	NS	NS
DKK1	252	NS	NS
WIF1	257	NS	NS
SFRP1	249	NS	0.0449
SFRP2	259	NS	0.0363
SFRP5	216	NS	NS
Crypt length	200	0.0069	0.0411

The number of crypts analysed in each case is represented by *n*. NS, not significant.

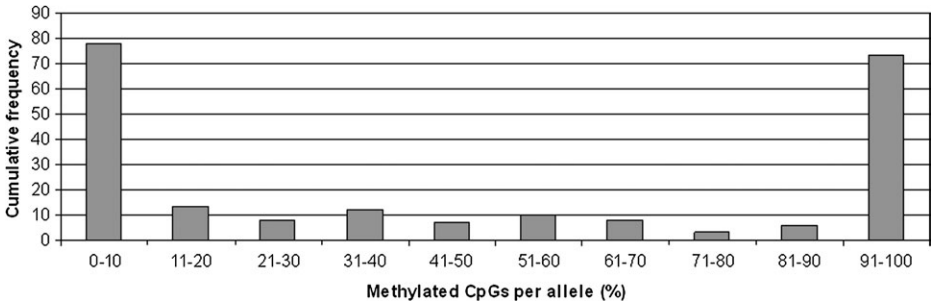

Fig. 1. The cumulative frequency of alleles with specified numbers of methylated CpGs, expressed as a percentage of the total (40), as determined by bisulphite sequencing.

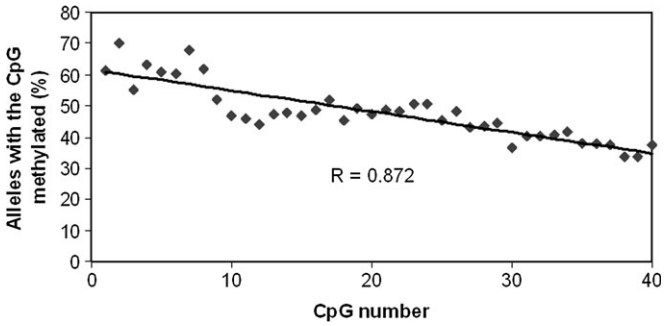

Fig. 2. The frequency of methylation at each CpG within the region analysed by bisulphite sequencing as a percentage of the total number of sequenced alleles. CpG number 1 is furthest upstream from the transcription start site.

used to assess its performance. Reassessing the assignment success rate using double cross-validation led to four-group classification models (both age and health status) with poor predictive ability for classifying both age and health status, with a success rate of 30.1%, close to the 25% rate expected by chance. This was further investigated using resampled 'leave-block-out' double cross-validation, where the number and sizes of the cross-validation segments used in the leave-subject-out analysis were retained. One hundred permutations of the membership of the 13 cross-validation segments were performed, where the permutations were restricted to be only between crypts belonging to the same group. Here, minimum assignment success rate was found to be 54.5% with a median rate of 60.9% comparable with the rate found using leave-one-out single cross-validation. From this analysis, we can conclude that the crypt-specific methylation profiles are significantly different between the disease-free subjects and the cancer patients and between the young and old subjects, and these differences account for their successful discrimination. However, although the model for assigning age and health status is valid within this small sample of subjects, it is not generally applicable and could not reliably predict age or cancer status from the methylation profiles of a set of crypts obtained from another individual.

Discussion

Several previous studies have described hypermethylation of CGIs in the morphologically normal human colorectal mucosa. Ahuja *et al.* (27) used a quantitative assay to establish that CGIs associated with *ESR1*, *MYOD1* and *N33* were all partially methylated in normal mucosa from cancer patients and that the level of methylation increased progressively with age. Similar results were obtained more recently by Kawakami *et al.* (28), and in our own previous work, we observed age-related methylation of *APC*, *AXIN2*, *DKK1*, *HPP1*, *N33*, *p16*, *SFRP1*, *SFRP2* and *SFRP4* (6) in a pooled series of patients with cancer, adenomatous polyps or no evidence of colorectal neoplasia. Thus, partial methylation of CGIs in the promoter regions of genes strongly implicated in the development of colorectal cancer appears to be an age-related phenomenon affecting healthy individuals as well as cancer patients. Ahuja *et al.* proposed that gene silencing as a consequence of aberrant CGI methylation constitutes an epigenetic field defect that increases the vulnerability of the colon to cancer with advancing age, but this hypothesis remains largely untested.

A general feature of previous studies on the CGI methylation of cancer-related genes in the colorectal mucosa is that DNA has usually been extracted from intact biopsies containing large numbers of crypts. It is now widely accepted that the vast majority of colorectal carcinomas arise from preexisting adenomatous polyps and that each polyp has its origins in a single genetically abnormal crypt (29). In order to explore the hypothesis that aberrant gene silencing caused by CGI methylation may contribute to multistage colorectal carcinogenesis in a manner equivalent to mutation, it is essential to quantify methylation at crypt level. Yatabe *et al.* (21) used bisulphite sequencing of DNA isolated from individual human crypts to determine the methylation signatures of *MYOD1*, cardio-specific homeobox (*CSX*) and biglycan (*BGN*). They concluded that human crypts are long lived, derived from the clonal expansion of progeny from a few stem cells and prone to random acquisition of aberrant methylation within CGIs. In the present study, we have used a similar approach to explore the aberrant CGI methylation of six genes directly involved in the Wnt signalling pathway. Our results are generally similar to those of Yatabe *et al.* (21) and Kim and Shibata (30) in that we observe considerable variation in the methylation of our target genes in individual crypts. At first sight, this seems surprising, given that our crypts were obtained from a single biopsy and therefore must have been located within 1 or 2 mm of each other, but it is consistent with the conclusion of Kim and Shibata (30) that the methylation patterns of adjacent crypts are no more similar to one another than to those of crypts from more distant locations. Predictably, based on the crypt niche model (21), our results do indicate that the methylation patterns within crypts, though highly variable, are more closely related than between crypts from the same individual, which in turn are more similar to one another than to crypts from other individuals. This implies that genetic factors, or environmental exposures, have contributed to the aberrant methylation patterns characteristic of our volunteers. Moreover, subtle differences in the patterns of crypt-specific methylation between subjects were able to predict whether the individual concerned harboured a carcinoma elsewhere in the colon. This confirms that despite the substantial variation between crypts, the subtle epigenetic variation that can distinguish between the normal mucosa of disease-free and cancer-bearing individuals (6) is conserved at the crypt level.

The previous study of Yatabe *et al.* (21) suggested that methylation of individual CpGs within the CGIs of *MYOD1*, *CSX* and *BGN* resembled a random process and that methylation of one CGI did not predict the methylation of another. Our results, in contrast, suggest that, at least for *WIF1* and probably for the other genes studied here, allele-specific methylation occurs through a non-random process in which the entire CGI is rapidly hypermethylated by directional spreading from a region upstream of that analysed here. This suggests

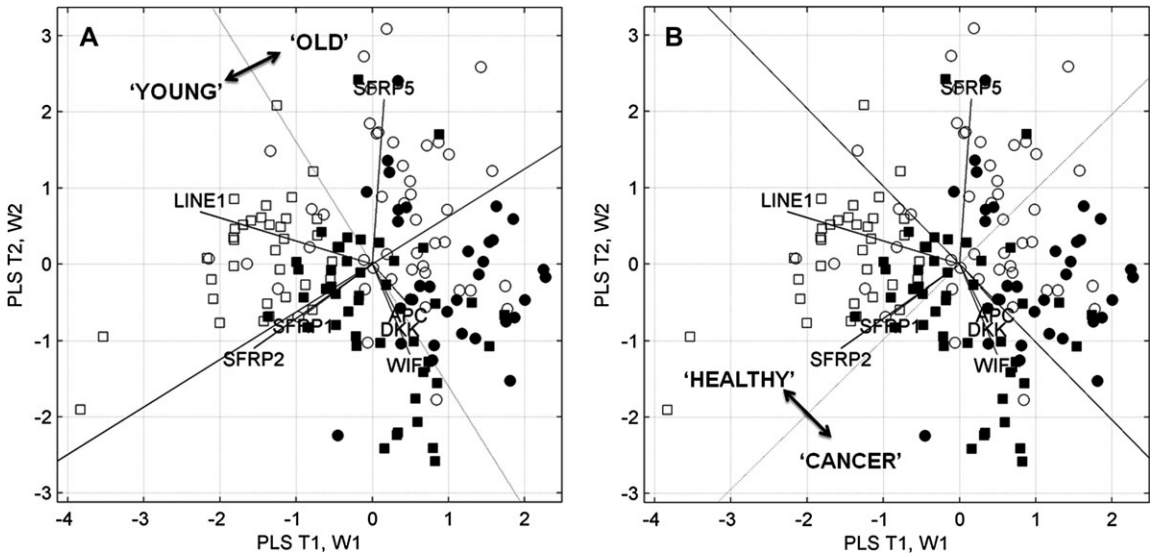

Fig. 3. Biplots showing the first and second PLS scores and the relative contribution of the DNA methylation rates for each of the genes and LINE-1 to the corresponding loading vectors for individual crypts from young-healthy (open squares), young-cancer (closed squares), old-healthy (open circles) and old-cancer (closed circles) subjects. The diagonal regression lines for predicting age (**A**) and cancer (**B**) are indicated by (thick lines). The corresponding $y = 0$ lines (dotted lines) can be interpreted as lines discriminating between young and old subjects (A) or between healthy and cancer (B). Positive values (to the right of the $y = 0$ lines) predict each crypt as belonging to the old groups (A) or the cancer groups (B).

a potential mechanism involving an as yet unidentified factor that protects the CGI from aberrant methylation, loss of which allows for the catastrophic spread of methylation into the CGI.

The levels of gene-specific methylation determined here in isolated crypts are significantly higher than we measured previously for the same genes in whole mucosal biopsies (6), even accounting for methodological differences. This supports the previous observation for the *ESR1* gene in colon tissue that methylation is predominantly a phenomenon of the epithelium and the presence of stroma and other cell types dilutes the epithelial methylation signal (27).

Yatabe *et al.* (21) have argued that CGI methylation drift in the human colon is an essentially random process occurring during stem cell division, with an error rate ~10 000 times greater than the spontaneous mutation rate per cell division. However, as Kim and Shibata (31) acknowledge, the acquisition of CGI methylation cannot be an entirely random process because different genes become methylated to widely differing extents. Methylation patterns acquired by individual stem cells undergo somatic inheritance and can become dominant within crypts through periodic succession 'bottlenecks' at which the total number of stem cells is reduced. The result is that a mosaic-like distribution develops as the methylation patterns of individual crypts diverge from one another (21).

The obvious question that arises from these studies relates to the functional significance for crypt homeostasis, and consequently for disease risk, of the identified patterns of aberrant CGI methylation in crypt epithelial cells, particularly in relation to *APC*, which is often regarded as the key 'gatekeeper' gene associated with initiation of colorectal neoplasia. It is well established that aberrant methylation of CGIs leads to chromatin remodelling and gene silencing. In principle, therefore, complete methylation of both *APC* alleles could initiate colorectal neoplasia in the same way as two *APC* mutations. In practice, however, there is little evidence that this is a frequent occurrence. The great majority of sporadic colorectal cancers are reported to carry *APC* mutations, and of those that do not, virtually all exhibit a replication error phenotype (RER+) (32). Indeed, a recent study suggests that the effects of *APC* hypermethylation results in only a subtle effect on *APC* expression and does not result in complete gene inactivation (33). Other studies have, however, correlated CGI methylation of the Wnt-related genes with reduced gene expression (12–14). Clearly, further studies are needed to clarify the functional consequences for maintaining homeostasis within a crypt harbouring methylation of *APC* and/or other Wnt-related genes. A noteworthy observation in this context is the significant increases in crypt length associated with both age and cancer in these subjects. Although increased crypt length in the normal mucosa of patients at high risk for colorectal carcinoma, or with a manifest carcinoma, has been reported (34), to our knowledge an association with age has not been previously demonstrated in humans. Interestingly, increased crypt length has been noted in very young rats compared with adults and this was related to increased proliferation (35). Indeed, it has been suggested that increased crypt length may be induced by the upwards expansion of the proliferative zone within the crypt, a frequent observation in the colonic mucosa, associated with both neoplasia and aging (36,37). This is supported in the mouse intestine where constitutive activation of the Wnt pathway following *APC* deletion leads to enlarged crypts (9,38). However, a direct association between DNA methylation and crypt morphology remains to be established.

There have been many attempts to refine the multistage model of carcinogenesis in the years since the original paper of Armitage and Doll (1) and a particular area of uncertainty relates to the number of mutational events necessary to establish an invasive cancer (39). In a reexamination of this problem, as it relates to colorectal cancer, Luebeck and Moolgavkar proposed a model in which the emergence of an initiated stem cell requires two rare genetic events, which the authors ascribe to two mutations, each affecting one of the *APC* loci in the cell, and a third high-frequency abnormality. The authors acknowledged that this might be attributable to a methylation event but went on to suggest a 'positional' event in which the normal asymmetric division of a stem cell is disrupted by its translocation from the crypt base to the proliferative zone, where the proliferative restraints associated with the basal microenvironment are lost. Our results raise the intriguing possibility that both conjectures may be correct. The critical microenvironment of the crypt base is believed to be maintained by the production of Wnt signalling ligands that interact with, among others, Frizzled receptors on stem cells located within the basal epithelium (7). Our results suggest that the transcription of key components of the Wnt signalling pathway may be progressively altered in the aging colon, a phenomenon that could disrupt communication between stem cells and their environment, and hence contribute to the spatial translocation postulated by Luebeck and Moolgavkar (29). Studies to explore this possibility are currently in progress in our laboratory. Further research, on the relationship

between methylation in apparently normal crypts and in tumours from the same individual, may shed light on the role of aberrant gene methylation in the transition to cancer.

Supplementary material

Supplementary Table 1 and Figure 1 can be found at http://www.chemse.oxfordjournals.org/.

Funding

Core Strategic Grant of the Biotechnology and Biological Sciences Research Council (BBSRC); Boston Leukaemia and Cancer Research Fund to M.R.W.

Acknowledgements

Conflict of Interest Statement: None declared.

References

1. Armitage,P. *et al*. (1954) The age distribution of cancer and a multi-stage theory of carcinogenesis. *Br. J. Cancer*, **8**, 1–12.
2. Vogelstein,B. *et al*. (2004) Cancer genes and the pathways they control. *Nat. Med.*, **10**, 789–799.
3. Phelps,R.A. *et al*. (2009) A two-step model for colon adenoma initiation and progression caused by APC loss. *Cell*, **137**, 623–634.
4. Sawan,C. *et al*. (2008) Epigenetic drivers and genetic passengers on the road to cancer. *Mutat. Res.*, **642**, 1–13.
5. Baylin,S.B. *et al*. (2006) Epigenetic gene silencing in cancer—a mechanism for early oncogenic pathway addiction? *Nat. Rev. Cancer*, **6**, 107–116.
6. Belshaw,N.J. *et al*. (2008) Profiling CpG island field methylation in both morphologically normal and neoplastic human colonic mucosa. *Br. J. Cancer*, **99**, 136–142.
7. Humphries,A. *et al*. (2008) Colonic crypt organization and tumorigenesis. *Nat. Rev. Cancer*, **8**, 415–424.
8. Powell,S.M. *et al*. (1992) APC mutations occur early during colorectal tumorigenesis. *Nature*, **359**, 235–237.
9. Sansom,O.J. *et al*. (2004) Loss of Apc *in vivo* immediately perturbs Wnt signaling, differentiation, and migration. *Genes Dev.*, **18**, 1385–1390.
10. Lamlum,H. *et al*. (2000) APC mutations are sufficient for the growth of early colorectal adenomas. *Proc. Natl Acad. Sci. USA*, **97**, 2225–2228.
11. Cheng,Y.Y. *et al*. (2007) Frequent epigenetic inactivation of secreted frizzled-related protein 2 (SFRP2) by promoter methylation in human gastric cancer. *Br. J. Cancer*, **97**, 895–901.
12. Suzuki,H. *et al*. (2004) Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. *Nat. Genet.*, **36**, 417–422.
13. Taniguchi,H. *et al*. (2005) Frequent epigenetic inactivation of Wnt inhibitory factor-1 in human gastrointestinal cancers. *Oncogene*, **24**, 7946–7952.
14. Aguilera,O. *et al*. (2006) Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer. *Oncogene*, **25**, 4116–4121.
15. Samuel,M.S. *et al*. (2009) Elevated Dnmt3a activity promotes polyposis in Apc(Min) mice by relaxing extracellular restraints on Wnt signaling. *Gastroenterology*, **137**, 902–913, 913 e1–11.
16. Whitehead,R.H. *et al*. (1993) Establishment of conditionally immortalized epithelial cell lines from both colon and small intestine of adult H-2Kb-tsA58 transgenic mice. *Proc. Natl Acad. Sci. USA*, **90**, 587–591.
17. Belshaw,N.J. *et al*. (2004) Use of DNA from human stools to detect aberrant CpG island methylation of genes implicated in colorectal cancer. *Cancer Epidemiol. Biomarkers Prev.*, **13**, 1495–1501.
18. Xiong,Z. *et al*. (1997) COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.*, **25**, 2532–2534.
19. Weisenberger,D.J. *et al*. (2005) Analysis of repetitive element DNA methylation by MethyLight. *Nucleic Acids Res.*, **33**, 6823–6836.
20. Iacopetta,B. *et al*. (2007) Methylation levels of LINE-1 repeats and CpG island loci are inversely related in normal colonic mucosa. *Cancer Sci.*, **98**, 1454–1460.
21. Yatabe,Y. *et al*. (2001) Investigating stem cells in human colon by using methylation patterns. *Proc. Natl Acad. Sci. USA*, **98**, 10839–10844.
22. Kemsley,E.K. (1996) Discriminant analysis of high-dimensional data: a comparison of principal components analysis and partial least squares data reduction methods. *Chemom. Intell. Lab. Syst.*, **33**, 47–61.
23. Westerhuis,J.A. *et al*. (2008) Assessment of PLSDA cross validation. *Metabolomics*, **4**, 81–89.
24. Ergon,R. (2002) PLS score-loading correspondence and a bi-orthogonal factorization. *J Chemom*, **16**, 368–373.
25. Martens,H. *et al*. (1989) *Multivariate Calibration*. Wiley, Chichester, UK.
26. Belshaw,N.J. *et al*. (2005) Methylation of the ESR1 CpG island in the colorectal mucosa is an 'all or nothing' process in healthy human colon, and is accelerated by dietary folate supplementation in the mouse. *Biochem. Soc. Trans.*, **33**, 709–711.
27. Ahuja,N. *et al*. (1998) Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Res.*, **58**, 5489–5494.
28. Kawakami,K. *et al*. (2006) DNA hypermethylation in the normal colonic mucosa of patients with colorectal cancer. *Br. J. Cancer*, **94**, 593–598.
29. Luebeck,E.G. *et al*. (2002) Multistage carcinogenesis and the incidence of colorectal cancer. *Proc. Natl Acad. Sci. USA*, **99**, 15095–15100.
30. Kim,K.M. *et al*. (2004) Tracing ancestry with methylation patterns: most crypts appear distantly related in normal adult human colon. *BMC Gastroenterol.*, **4**, 8.
31. Kim,K.M. *et al*. (2002) Methylation reveals a niche: stem cell succession in human colon crypts. *Oncogene*, **21**, 5441–5449.
32. Rowan,A.J. *et al*. (2000) APC mutations in sporadic colorectal tumors: a mutational ''hotspot'' and interdependence of the ''two hits''. *Proc. Natl Acad. Sci. USA*, **97**, 3352–3357.
33. Segditsas,S. *et al*. (2008) Promoter hypermethylation leads to decreased APC mRNA expression in familial polyposis and sporadic colorectal tumours, but does not substitute for truncating mutations. *Exp. Mol. Pathol.*, **85**, 201–206.
34. Sandforth,F. *et al*. (1991) Identification of patients at high risk for colorectal carcinoma from biopsy studies of the apparently normal colorectal mucosa. A multivariate analysis. *Eur. J. Clin. Invest.*, **21**, 295–302.
35. Goodlad,R.A. *et al*. (1992) Colonic cell proliferation and growth fraction in young, adult and old rats. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.*, **61**, 415–417.
36. Ponz de Leon,M. *et al*. (1988) Pattern of epithelial cell proliferation in colorectal mucosa of normal subjects and of patients with adenomatous polyps or cancer of the large bowel. *Cancer Res.*, **48**, 4121–4126.
37. Roncucci,L. *et al*. (1988) The influence of age on colonic epithelial cell proliferation. *Cancer*, **62**, 2373–2377.
38. Sansom,O.J. *et al*. (2007) Myc deletion rescues Apc deficiency in the small intestine. *Nature*, **446**, 676–679.
39. Hornsby,C. *et al*. (2007) What can we learn from the population incidence of cancer? Armitage and Doll revisited. *Lancet Oncol.*, **8**, 1030–1038.

Received November 9, 2009; revised March 23, 2010; accepted April 10, 2010