FULL PAPER

BJC

British Journal of Cancer (2015) 113, 1035–1045 | doi: 10.1038/bjc.2015.303

Keywords: DNA methylation; CpG methylator phenotype; CIMP; rectal cancer; epigenetic biomarker

Potential of DNA methylation in rectal cancer as diagnostic and prognostic biomarkers

Ruth Exner[1,5], Walter Pulverer[2,5], Martina Diem[3], Lisa Spaller[3], Laura Woltering[3], Martin Schreiber[4], Brigitte Wolf[1], Markus Sonntagbauer[2], Fabian Schröder[2], Judith Stift[3], Fritz Wrba[3], Michael Bergmann[1], Andreas Weinhäusel[2] and Gerda Egger*[,3]

[1]*Department of Surgery, Medical University of Vienna, Waehringer Guertel 18–20, A-1090 Vienna, Austria;* [2]*AIT—Austrian Institute of Technology GmbH, Health & Environment Department, Molecular Diagnostics, Muthgasse 11/2, A-1190 Vienna, Austria;* [3]*Clinical Institute of Pathology, Medical University of Vienna, Waehringer Guertel 18–20, A-1090 Vienna, Austria and* [4]*Department of Obstetrics and Gynecology, Medical University of Vienna, Waehringer Guertel 18–20, A-1090 Vienna, Austria*

Background: Aberrant DNA methylation is more prominent in proximal compared with distal colorectal cancers. Although a number of methylation markers were identified for colon cancer, yet few are available for rectal cancer.

Methods: DNA methylation differences were assessed by a targeted DNA microarray for 360 marker candidates between 22 fresh frozen rectal tumour samples and 8 controls and validated by microfluidic high-throughput and methylation-sensitive qPCR in fresh frozen and formalin-fixed paraffin-embedded (FFPE) samples, respectively. The CpG island methylator phenotype (CIMP) was assessed by MethyLight in FFPE material from 78 patients with pT2 and pT3 rectal adenocarcinoma.

Results: We identified and confirmed two novel three-gene signatures in fresh frozen samples that can distinguish tumours from adjacent tissue as well as from blood with a high sensitivity and specificity of up to 1 and an AUC of 1. In addition, methylation of individual CIMP markers was associated with specific clinical parameters such as tumour stage, therapy or patients' age. Methylation of *CDKN2A* was a negative prognostic factor for overall survival of patients.

Conclusions: The newly defined methylation markers will be suitable for early disease detection and monitoring of rectal cancer.

Rectal cancer comprises one-third of colorectal cancers (CRCs) and is characterised by its specific anatomic localisation in the small pelvis, which makes operative resection with clear margins more difficult but allows for irradiation as an effective neoadjuvant treatment. Molecularly, distal tumours show a higher frequency of chromosomal instability (CIN) and p53 mutations but a lower frequency of microsatellite instability (MSI) and CpG island methylator phenotype (CIMP) compared with proximal tumours (Iacopetta, 2002). Genome-wide genetic and epigenetic analyses revealed an association of hypermutated tumours (mutation rate >12 per 10^6 bases) with high DNA methylation levels, MSI and defects in mismatch repair-related genes as well as epigenetic MLH1 silencing (Network CGA, 2012). Generally, these hypermutated tumours were more likely to be found in the ascending colon and showed higher DNA methylation levels. Thus, molecular differences seem to follow a spatial linearity from proximal to distal locations along the colon rather than a dichotomy of proximal and distal subsites (Yamauchi *et al*, 2012; Bae *et al*, 2013).

CIMP has long been recognised in a subgroup of CRC with extraordinarily high levels of DNA methylation of CpG islands (Toyota *et al*, 1999) and has been associated with different clinical and molecular features such as patient age, gender, tumour localisation or *BRAF* mutation status (Weisenberger *et al*, 2006). In addition, a specific DNA methylation subgroup associated with *KRAS* mutations has been observed in CRC (Ogino *et al*, 2006; Shen *et al*, 2007; Yagi *et al*, 2010), which is now commonly accepted as the CIMP-low group (Hinoue *et al*, 2012). In-depth genome-scale analyses of DNA methylation in CRC revealed four

*Correspondence: Dr G Egger; E-mail: gerda.egger@meduniwien.ac.at
[5]These authors contributed equally to this work.

Received 28 March 2015; revised 17 June 2015; accepted 30 July 2015; published online 3 September 2015

distinct classes (CIMP-high, CIMP-low, cluster 3 and cluster 4) and suggested biological differences between the four subgroups (Hinoue *et al*, 2012). Intriguingly, rectal cancers were mainly associated with a distinct cluster of tumours with low cancer-associated DNA methylation.

Tumour-specific DNA methylation signatures have emerged as promising targets for biomarker development for several malignancies including CRC and might be used for early detection, prognosis and therapy response prediction in the future (Draht *et al*, 2012; Lange and Laird, 2013). However, discovery of markers specific for rectal cancers has been hampered by the lack of studies addressing rectal tumours as an independent entity and studies using limited numbers of candidate genes. A set of five hypermethylated tumour suppressor genes was identified in early-stage rectal cancer using a candidate gene approach and was associated with localised disease (Leong *et al*, 2011).

A large Asian population-based study investigating nine different bowel subsites of CRC revealed a linear association of tumour location and clinicopathological characteristics (Bae *et al*, 2013). In this study, CIMP was an independent poor prognostic factor for disease-free survival (DFS) and overall survival (OS) in rectal cancers but not for tumours from proximal and distal colon locations. Analysis of a set of methylated-in-tumour (MINT)-specific regions revealed a two-marker panel, which was predictive of distant recurrence in early, node-negative rectal cancers and could also identify a group of patients with increased risk of local tumour recurrence, suggesting a value of this marker panel for patient stratification for neoadjuvant radiation therapy (de Maat *et al*, 2008, 2010). Furthermore, *KRAS2* mutation and CIMP were associated with a worse 5-year survival in a population-based study of 990 rectal cancers (Samowitz *et al*, 2009) and *KRAS* mutation and *CDKN2A* promoter methylation were suggested to indicate more aggressive tumours with worse prognosis (Kohonen-Corish *et al*, 2014). DNA methylation levels of the repetitive element LINE-1 can be used to assess genome-wide DNA methylation changes. A recent report identified LINE-1 hypomethylation in state I–II rectal cancers as a surrogate marker associated with a higher chance of tumour recurrence and decreased survival time (Benard *et al*, 2013).

In a recent analysis in patients with locally advanced rectal cancer CIMP positivity was associated with poor response to preoperative radiochemotherapy and significantly reduced DFS (Jo *et al*, 2012). *MGMT* and *TFAP2E* methylation were found predictive of response to radiochemotherapy with *MGMT* hypermethylation being associated with responsiveness and *TFAP2E* with resistance, respectively (Lange *et al*, 2012; Sun *et al*, 2013).

On the basis of these studies, we hypothesised that rectal cancers harbour aberrant tumour-specific DNA methylation, which might serve as a biomarker for clinical assessment of these cancers. We screened a set of 22 primary rectal tumours after short-time preoperative radiation therapy for their methylation status of 360 known tumour-specific methylated genes. We included eight normal peripheral blood mononuclear cell (PBMC) samples in the analyses. The rationale for using PBMCs was the identification of markers, which might be suitable for non-invasive tumour diagnostics in plasma of patients. The initial screen was followed by technical and biological validation of our findings and resulted in the discovery of a three-gene marker panel that was suitable for the classification of rectal tumour tissue. Furthermore, we evaluated the CIMP status of 78 rectal tumours isolated from formalin-fixed paraffin-embedded (FFPE) material and correlated methylation of CIMP-specific markers with clinicopathological parameters. The workflow of our analyses is depicted in Figure 1.

MATERIALS AND METHODS

Patient selection. Patients with pT2/pT3 rectal cancer, who received surgery with curative intention (R0) in the years 2000–2008 at the Department of Surgery at the Medical University of Vienna were included in the analyses. For the initial screen and validation fresh frozen samples ($n = 22$) were used, whereas as second larger cohort of 78 FFPE samples was used for marker confirmation and CIMP marker evaluation. Demographic data for patients from the fresh frozen screening population and the FFPE-CIMP population are listed in Table 1.

DNA isolation and bisulphite conversion. DNA from frozen and FFPE tissues was isolated after assessment by a pathologist. FFPE tissues were deparaffinised using xylol and rehydrated. DNA was isolated using commercially available kits according to the standard protocols (QIAamp DNA Mini Kit; Qiagen, Hilden, Germany). Bisulphite conversion was performed using commercially available kits (EpiTect; Qiagen, Hilden, Germany) according to the manufacturer's protocol.

CIMP marker testing. We employed MethyLight analyses on DNA isolated from FFPE patient samples according to the

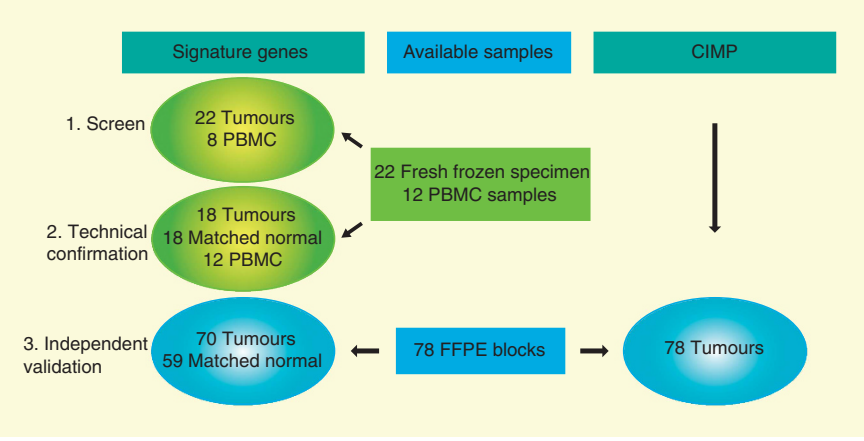

Figure 1. Workflow. Two sets of samples were available for analyses. (i) A total of 22 fresh frozen tumour specimen from patients with rectal cancer, 18 matched fresh frozen adjacent tissues and 12 PBMC samples from blood donors. These samples were used for the initial screen and for the technical confirmation of markers. (ii) A total of 78 FFPE tissue blocks from patients with rectal cancer were used for independent validation of markers and for CIMP marker analyses.

Table 1. Demographic data of FFPE and native sample groups

	Patiens of the FFPE study group ($n=78$)		Patients of the native study group ($n=22$)	
	n	%	*n*	%
Sex				
Male	53	68	15	68
Female	25	32	7	32
Age at randomisation				
Mean	65		64	
Range	31–86		62–86	
Therapy				
Short-time radiation	59	76	22	
Radiochemotherapy	8	10	0	
No therapy	6	8	0	
ND	5	6	0	
Stage				
T2	33	42	0	
T3	45	45	22	
Node involved				
N0	47	60	9	41
N1	19	24	7	32
N2	12	15	6	27
Dukes				
A	24	31	0	
B	22	28	9	41
C	28	36	13	59
D	4	5	0	
Tumour location				
Rectum	70	90	22	
Rectosigmoid	8	10		
Recurrence				
Local	5	6	0	
Distant	15	19	8	36
ND	34	44	4	18
5-year OS				
Survived	26	33	11	50
Not survived	19	24	6	27
ND	33	42	5	23
5-year DFS				
Recurrence	19	24	8	36
No recurrence	24	31	10	45
ND	35	45	4	18

Abbreviations: DFS = disease-free survival; ND = not determined; OS = overall survival.

published protocols (Campan *et al*, 2009) using two different CIMP panels (classical CIMP panel: *CDKN2A*, *MINT1*, *MINT2*, *MINT31* and *MLH1*; new CIMP panel: *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*) (Toyota *et al*, 1999; Weisenberger *et al*, 2006). Data normalisation and percentage of methylated ratio (PMR) value calculations were done as previously described (Campan *et al*, 2009). A PMR that was PMR + 10 above a non-methylated control DNA was assessed as positive methylation (Weisenberger *et al*, 2006).

Targeted CpG-360 DNA methylation arrays. Total DNA (600 ng) from fresh frozen tumour tissue ($n=22$) and PBMCs ($n=8$, 4 male and 4 female) was subjected to methylation-specific restriction enzymes (MSRE) to cleave unmethylated DNA. Methylated DNA remains uncleaved and can be subjected to screening using targeted DNA microarrays for 360 methylation marker candidates (targeting CpG islands and human gene promoters). The main principle of the methodology was published previously (Pulverer *et al*, 2012). For the experiments of the present study we used the MSREs HpaII (cut site: CCGG; Fermentas, St. Leon-Rot, Germany), Hin6I (cut site: GCGC; Fermentas, St. Leon-Rot, Germany), AciI (cut site: CCGC; NEB, Frankfurt am Main, Germany) and HpyCH4IV (cut site: ACGT; NEB, Frankfurt am Main, Germany). A unit of 3 U (0.3 μl) of each enzyme were used per digestion reaction and incubated at 37 °C for 16 h. The digested DNA samples were amplified in 16 multiplex PCR reactions amplifying in total 360 methylation marker candidates using biotinylated reverse primers. Pooled amplicons were detected on a targeted DNA microarray via streptavidin–Cy3 conjugate. Subsequently, significant markers were identified applying statistical tests for class comparison and class prediction. Primer sequences are available on request (Weinhäusel *et al*, 2009; Weinhäusel and Pulverer, 2013).

Validation of markers by MSRE-coupled qPCR. To confirm the differentially methylated gene regions from the microarray experiment DNA of tumour tissue ($n=18$) for which adjacent normal sample material was available together with 12 PBMCs were subjected to a microfluidic high-throughput qPCR (μHT–qPCR) system (Fluidigm's Biomark, San Francisco, CA, USA). Before microfluidic qPCR, a pre-amplification was carried out. For the pre-amplification the samples were digested with the four MSREs as mentioned above. The primer mixes for the pre-amplification contained 41 primer pairs at a final concentration of 200 nM each primer. Subsequently, qPCR was performed using EvaGreen (Biotium, Hayward, CA, USA)-based detection and the following cycling conditions: 50 °C for 2 min, 95 °C for 10 min, 45 cycles at 95 °C for 15 s and 65 °C for 1 min. Ct values are inversely correlated to methylation levels. In order to infer a direct association between Ct value and methylation status we calculated the ΔCt value (number of executed cycles ($n=45$)–Ct of sample). Primer sequences are available on request.

Validation of the identified methylation signature in FFPE-derived DNA. Quantitative methylation-specific PCR (qMSP) assays were designed for the methylated allele of 5′UTR regions of *TMEFF2*, *TWIST1* and *PITX2* and performed in analogy to the MethyLight analyses using SYBR green detection. DNA was isolated and bisulphite converted as described above. Relative methylation was normalised to input DNA and to 100% methylated control DNA as previously described (Campan *et al*, 2009). Primer sequences are available on request.

Bioinformatics and statistics. For the statistical analyses of the generated array data BRB Array tools Version 4.2.1 (https://linus.nci.nih.gov/BRB-ArrayTools.html) and R statistical software Version 2.14.2 (https://cran.r-project.org/bin/windows/base/old/2.14.2/) were used. To invert the indirect association between Ct values and methylation values, the ΔCt value was used (45 − Ct) for the calculations. Data derived from MSRE-based experiments were scale normalised and methylation differences between the sample groups were identified by comparative analyses using a random-variance *t*-test.

We used models based on different algorithms to predict the class of future samples. The models were based on the Diagonal Linear Discriminant Analysis (DLDA) (Dudoit *et al*, 2002), the nearest centroid method, k-nearest-neighbour classification (kNN) (Dudoit *et al*, 2002), support vector machines (SVM) (Ramaswamy *et al*, 2001) and (Bayesian) compound covariate predictor (CCP) (Radmacher *et al*, 2002). The DLDA is a two-class classification method based on the assumption of two Gaussian class distributions with a common covariance matrix. As a parametric model, it is highly efficient when the assumptions are met. However, the maximum likelihood estimators, which are usually plugged into the model are highly sensitive to violations from the assumptions, such as non-Gaussian distributed class distributions (e.g., outliers) or

even when the two classes show strong differences in the covariance structure.

The nearest centroid method is an extension of the DLDA. It embeds a filter selection method by using a soft-threshold parameter to eliminate most non-contributing features. It is a fast algorithm but the choice of features is model dependent.

The kNN is a non-parametric multi-class classifier. It performs well when classes are heterogeneous possibly consisting of several subgroups. KNN uses the Euclidean distance to assign an instance to a class, thus it does not perform well, when there are strong (linear) dependencies between the covariates.

The SVM is non-probabilistic binary linear classifier that finds the optimal hyper plane separating the sample space into disjoint regions. The performance of the classifier depends highly on the distribution of the classes in the sample space, for example, whether it is possible to separate the classes with a linear functional. Therefore, the SVM can be extended to efficiently perform non-linear classification.

The CCP is a weighted linear combination of log-ratios. By specifying a more stringent significance level, fewer genes are included in the multivariate predictor. CCP is implicitly based on the assumption that the prior probabilities are each 0.5. Bayesian CCP selects the differentially methylated loci for distinguishing two classes in a cross-validated training set using weighted average of the log methylation values, with the weights being the *t*-statistics of differential methylation in the training set. The values of the compound covariate scores of samples in each class in the training set are considered to have a Gaussian distribution.

The models incorporated genes that were differentially methylated between the defined groups assessed by the random-variance *t*-test. The prediction error was estimated using the leave-one-out cross-validation (LOOCV) for each model. Cross-validation error rates were estimated by repeating the LOOCV process with randomly permutated class labels. As a threshold for statistical significance $P<0.001$ and a false discovery rate (FDR) <0.01 was used for the microarray data. $P<0.05$ and FDR <0.01 was considered as statistically significant for the qPCR-based data. Correlation of CIMP with clinicopathological characteristics was performed with GraphPad Prism 6 software using a Spearman correlation matrix. Statistical significance was confirmed by χ^2-test or Fisher's exact test. Kaplan–Meier statistics were used to determine the prognostic value of the CIMP marker for OS and DFS. *P*-values to Kaplan–Meier curves were calculated with log-rank tests. Multivariate analyses of the OS and DFS were performed by using a Cox proportional hazard model, unadjusted or adjusted for lymph node status and grading. Cox proportional hazard models and Kaplan–Meier plots were computed with the survival package of R version 2.15.1, an open-source language and environment for statistical computing (Team, 2009). We used $P<0.05$ and 95% confidence intervals for assumption of statistical significance.

Ethical and legal aspects. The project was conducted in accordance with the latest revision of the Declaration of Helsinki and the requirements of Good Clinical Practice of the European Community (CPMP/ICH/135/95). The study protocol was approved by the institutional review board ('Ethikkommission') of the Medical University of Vienna (EK-No 136/2010).

RESULTS

Identification of DNA methylation changes in cancer *vs* control tissue. We subjected 22 fresh frozen rectal cancer DNA samples together with 8 control samples (PBMC from healthy volunteers) to microarray-based methylation analysis interrogating 360 promoters of genes known to be hypermethylated in different cancers (Weinhäusel *et al*, 2009). The differences in methylation intensities of the two groups (cancer *vs* control) were identified using a random-variance *t*-test (class comparison). A total of 53 out of 360 (15%) gene regions with differential methylation between the tumour tissue and the normal control samples were identified (Figure 2A). Full information for the 53 differentially methylated regions (DMRs) including gene symbol, chromosomal position, *P*-value, FDR and log fold change can be found in the Supplementary Information (Supplementary Table 1). Two-thirds ($n=35$) of the identified DMRs showed hypermethylation in the tumour tissue. The average fold change of the microarray-derived signal intensities of the hypermethylated DMRs was 3.82 ± 1.85 (range, 1.2–8.35). The remaining 18 out of 53 DMRs were found hypomethylated in the tumour tissue. Their average fold change was -1.36 ± 0.78 (range, -0.79 to -4.06). Principle component analysis also showed clustering of tumour and normal samples into two major distinct groups (Figure 2B).

Hierarchical clustering allowed for almost perfect distinction between the two sample groups (tumour *vs* PBMC) based on the methylation signatures of the 53 DMRs (Figure 2C). Thus, our data suggest that differential methylation of a selected number of loci might allow for distinction between rectal tumour tissue and normal PBMCs.

Identification of cancer-specific methylation signatures. Following the MSRE array experiment we chose 41 DMRs for technical replication using high-throughput MSRE qPCR using the Fluidigm Biomark platform in tumor *vs* PBMC DNA as well as tumour *vs* matched adjacent normal tissue. The DMRs contained the top 22 significant DMRs identified with the array experiment (20 hypermethylated in tumour *vs* PBMC with average fold change of 4.63 ± 1.99: *TJP2*, *TFPI2*, two different regions for *PITX2* (the two different regions target two different CpG islands of the *PITX2* gene. The exact location is listed in Supplementary Table 1. The *PITX2* gene has two alternative transcripts with independent promoters), *SPARC*, *GDNF*, *PENK*, *ZNF502*, *RARB*, *CDX1*, *SERPINB2*, *SALL3*, *TMEFF2*, *BOLL*, *SFRP2*, *DCC*, *ESR1*, *SRGN*, *MYOD1* and *WT1*; and two hypomethylated DMRs with an average fold change in methylation of -2.81 ± 1.77: *S100A2* and *HLA-G*). In addition we chose 19 well-known cancer-associated genes that were included in the screen that showed high fold changes in some of the samples but did not reach statistical significance in the array experiment (*SEZ6L*, *CHFR*, *CXADR*, *GATA4*, *IL1B*, *ZNF256*, *CALCA*, *DAPK1*, *CLIC4*, *RHOXF1*, *S100A8*, *CD24*, *PTGS2*, *MSH4*, *TWIST1*, *FMR1*, *NKX2-1*, *THBD* and *XIST*). For qPCR confirmation we had 18 out of 22 of the initially tested fresh frozen tumour samples plus the respective adjacent tissues ($n=18+18$) available. In addition 12 PBMC samples were included, which contained the 8 samples initially tested on the CpG-360 array. The analysis confirmed differential methylation between tumour and PBMC DNA for 15 of the 22 significant markers derived from the microarray experiments and 7 of the 19 additional candidate genes between tumour and PBMC DNA (Figure 3 and Supplementary Table 2). Using the different classification algorithms described in the 'Material and Methods' section, we were able to define a three-gene signature, comprising *TFPI2*, *DCC* and *PTGS2*. The combined information of these three genes allowed a perfect discrimination between tumour samples and peripheral blood with a classification success of 100% (sensitivity: 1; specificity: 1; AUC: 1) (Figure 3 and Supplementary Table 3).

Further, we detected differential methylation of 17 out of 41 DMRs between the tumour and their adjacent normal tissue samples (Figure 4A and Supplementary Table 4). Three genes including *TMEFF2*, *TWIST1* and *PITX2* were able to classify 97–100% of the samples to the correct group dependent on the classification algorithm used (sensitivity: 0.89–1; specificity: 0.94–1;

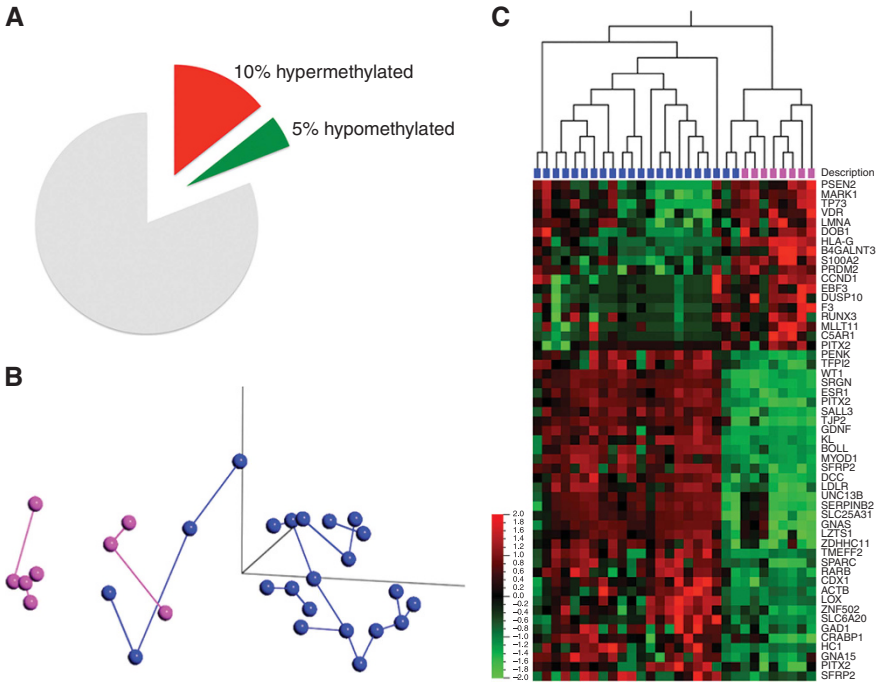

Figure 2. Microarray-based methylation screening of rectal cancers. (**A**) Percentage of differentially methylated gene targets identified by screening fresh frozen rectal tumours (red colour indicates regions hypermethylated in rectal cancers relative to PBMCs; green colour indicates hypomethylated regions). (**B**) Principal component analysis reveals separation of normal and tumour samples based on their methylation profile (purple circles represent PBMCs, blue circles represent rectal tumour samples). The two mis-clustered patients are not differing from the other patients in terms of age, gender, tumour stage/grade or preoperative therapy. The fact that these patients cluster rather with PBMCs is due to a different methylation signature of some of the genes, which might indicate that these tumours are different in terms of their molecular biology. Importantly, our identified signature gene, do classify these tumours correctly. (**C**) Hierarchical clustering based on the 53 top differentially methylated genes group normal and tumour samples in two distinct clusters (blue, tumour samples; purple, PBMCs; red, hypermethylated genes; green, hypomethylated genes).

AUC: 1) (Figure 4A and Supplementary Table 5). These three genes showed also promising performance to dissect the tumours from non-tumourous tissue (adjacent tissue and PBMC samples in one group) with classification values between 96 and 100% (sensitivity: 0.89–1; specificity: 0.83–1; AUC: 0.99) (Figure 4B and Supplementary Table 6).

Thus, we discovered two three-gene signatures, which were able to discriminate with high specificity and sensitivity between rectal cancer and PBMC as well as rectal cancer and adjacent normal tissue, respectively.

Confirmation of methylation signature in FFPE tissue. The three-gene signature, which allowed for a perfect discrimination between tumour tissue and matched adjacent normal tissue (*TMEFF2*, *TWIST1* and *PITX2*), underwent an additional round of confirmation using DNA isolated from 70 FFPE tumour tissues and 59 FFPE adjacent normal tissues of patients with rectal cancer. All three genes showed frequent and high levels of DNA methylation in tumours relative to adjacent normal tissues with high significance, which confirmed our initial analyses in fresh frozen tumour samples (Figure 5).

CIMP analyses. In order to reassess the clinical value of CIMP for rectal cancer, we evaluated two different CIMP marker panels on DNA isolated from FFPE tumour material from patients undergoing neoadjuvant radiation therapy compared with non-radiated patients (Figure 6). We determined the CIMP status for 78 tumour samples. Application of the classical CIMP marker panel comprised of *MINT1*, *MINT2*, *MINT31*, *CDKN2A* and *MLH1* revealed 15 out of 78 (19.23%) as CIMP positive ($\geqslant$3 out of 5 markers methylated). Application of the more stringent CIMP panel comprised of *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1* identified only three CIMP-high cases (3 out of 5 markers methylated) within the 78 tumour samples (3.8%). Methylation of one or two CIMP markers of this panel was detected in 26 samples (33.3%), which might represent the previously described CIMP-low class of tumours (Hinoue *et al*, 2012).

To determine the associations of differential methylation of individual genes in the CIMP panels or of different CIMP classifiers with clinicopathological data we calculated a correlation matrix between the different variables using Spearman rho correlation (Figure 7A). Exposure to radiation therapy was negatively correlated with CIMP classifiers based on the classic CIMP panel and individual methylation of *MLH1*. Tumour stage was negatively correlated with *MLH1* methylation and positively correlated with *MINT2* methylation. *MLH1* methylation was correlated to the age of the patient at time of surgery.

To define the prognostic value of our data we computed the association of methylation status with the OS and DFS using Kaplan–Meier analyses. These analyses identified methylation of *CDKN2A* as a negative prognostic factor for OS of patients with rectal cancer (Figure 7B).

Next, we used Cox proportional hazards models to determine the prognostic value of *CDKN2A* in univariate and multivariate analyses together with nodal status and tumour stage, which are both important prognostic clinical markers (Table 2). Our data revealed that *CDKN2A* was a negative prognostic factor for OS in both univariate ($P = 0.045$) and in multivariate analysis ($P = 0.017$).

Generally, DNA methylation of selected CIMP markers correlates with clinicopathological characteristics and disease prognosis. Relating to the stringent CIMP panel (here referred to

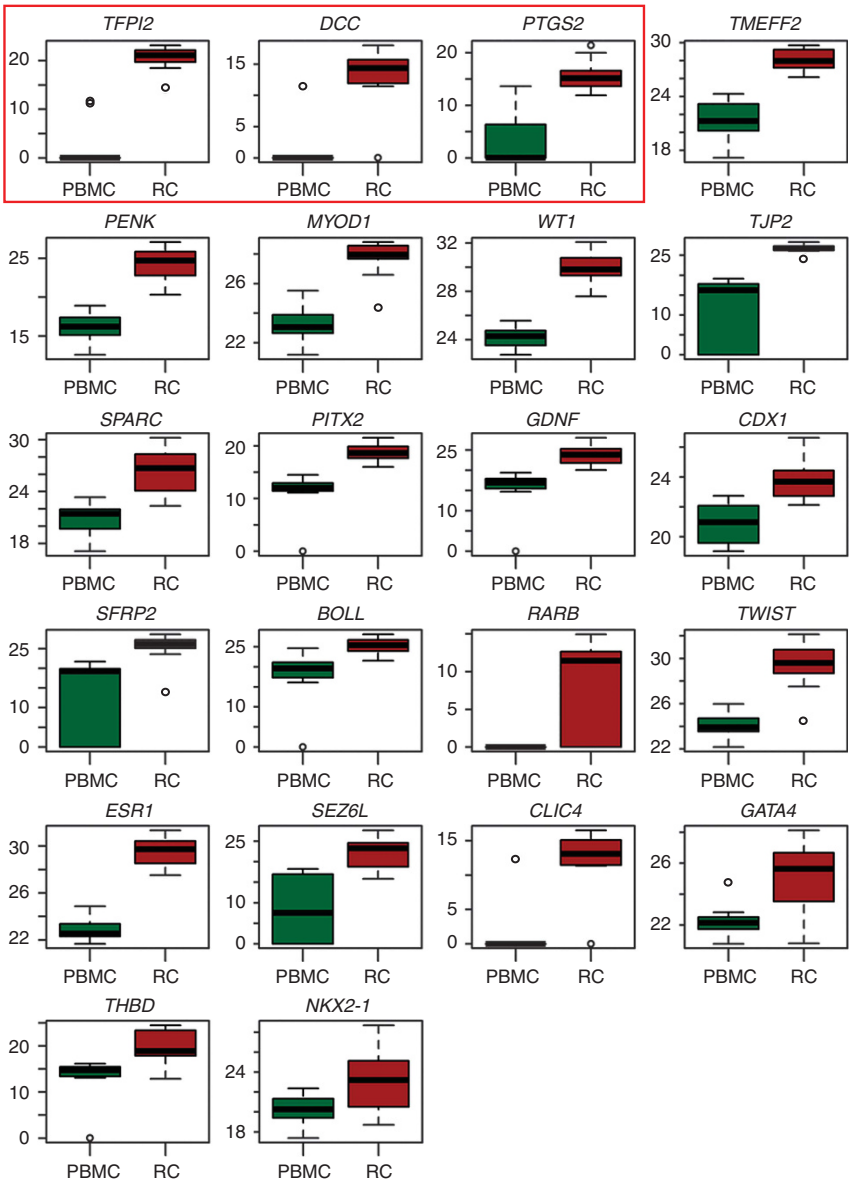

Figure 3. qPCR confirmation of a classification signature between tumour and PBMCs. Boxplots indicating the methylation differences of 22 markers, which were selected from the top differentially methylated genes on the microarray and from candidate genes, by MSRE-based qPCR. Ct values were inverted (45 − Ct) to reflect high and low methylation levels, respectively. *TFPI2*, *DCC* and *PTGS2* (marked by a red frame) were determined as the best markers to distinguish between tumour and PBMCs by class prediction (RC, rectal carcinoma). Whiskers depict the 25 and 75% quartile of the data.

as CIMP new), our data confirm previous findings indicating that CIMP high is a rare event in rectal cancer.

DISCUSSION

Our study aimed at identifying potentially diagnostic and/or prognostic signatures of primary human rectal carcinoma based on DNA methylation. We performed a targeted DNA methylation screen interrogating 360 genes for cancer-associated marker candidates in 22 rectal tumour samples in comparison to PBMCs identifying 53 genes with differential methylation ($P<0.001$). Validation of the 22 selected genes of the screen and 19 additional candidate markers, frequently hypermethylated in diverse cancers, analysed by class prediction identified two novel diagnostic marker sets consisting of three genes each, which were able to discriminate between tumour and PBMC as well as between tumour and adjacent tissue or normal tissue, respectively. Interestingly, the genes were not shared between the two signatures, indicating that for diagnostic applications (e.g., DNA methylation testing in cfDNA) different marker panels might be more sensitive to detect methylated tumour DNA in serum (tumour/PBMC signature) as compared with stool (tumour/adjacent normal signature).

The signature differentiating tumours from adjacent tissue and PBMCs, which we also confirmed in DNA isolated from FFPE samples using an independent technology, comprised *TMEFF2* (transmembrane protein with EGF-like and two follistatin-like domains 2), *TWIST1* (twist basic helix-loop-helix transcription factor1) and *PITX2* (paired-like homeodomain 2). *TMEFF2* methylation was first reported in bladder and colon cancer (Liang *et al*, 2000) and was meanwhile established as a marker for non-invasive testing of serum, urine or stool for different cancer entities (Lee *et al*, 2012; Elliott *et al*, 2013; Monteiro-Reis

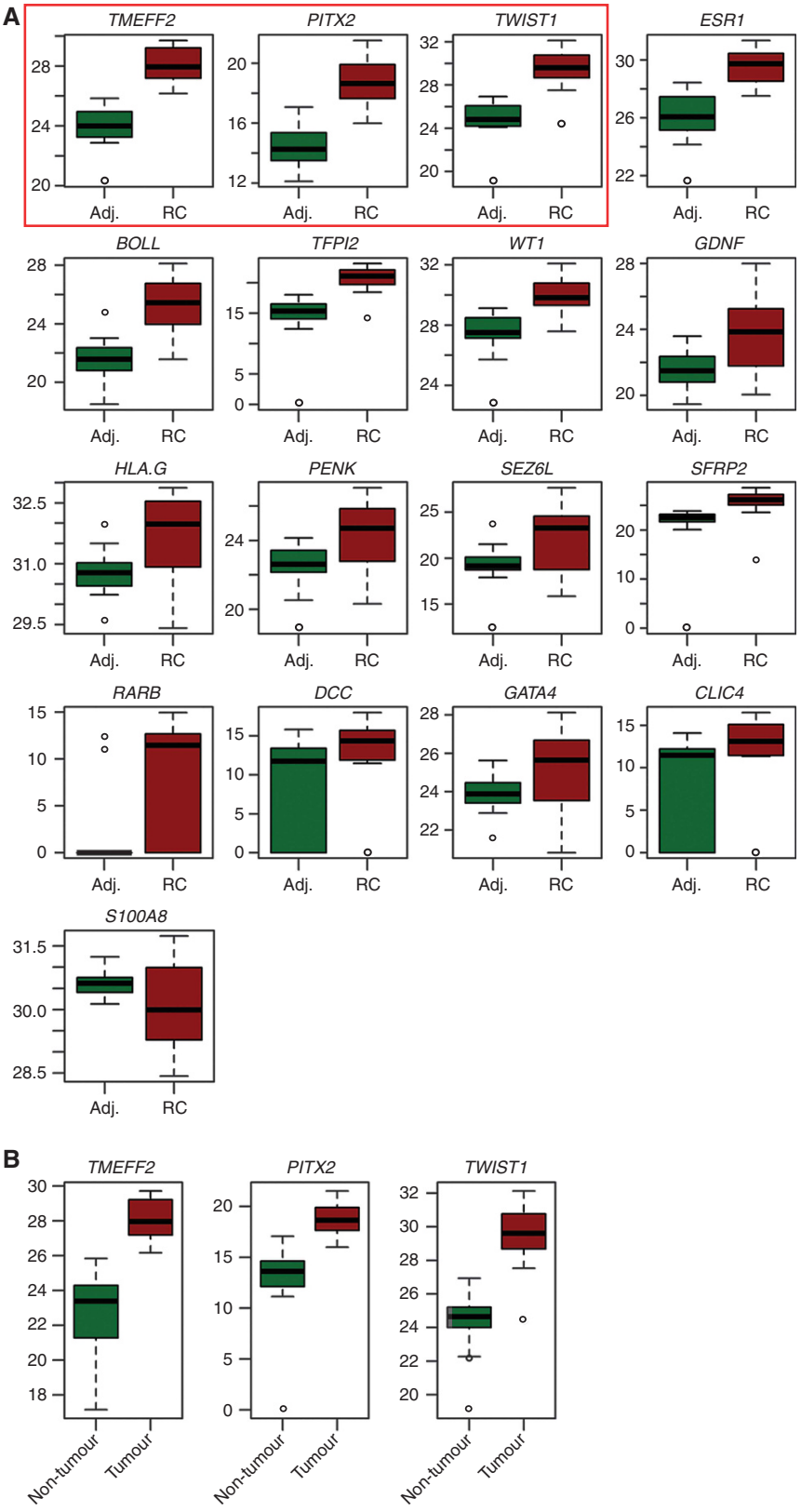

Figure 4. qPCR confirmation of a classification signature between normal and tumour. (**A**) Boxplots indicating the 17 markers, which were significantly differentially methylated between normal adjacent and tumour tissue as determined by MSRE-coupled qPCR. Ct values are depicted as (45 – ct) to reflect the methylation levels, directly (adj., adjacent normal; RC, rectal carcinoma). *TMEF2*, *PITX2* and *TWIST* (marked by a red frame) were determined as the best markers to distinguish normal adjacent from tumour tissue by the classification algorithm. (**B**) Boxplots showing classification signature genes that best distinguish non-tumour (adjacent normal and PBMC) from tumour tissues as determined by class prediction. Whiskers depict the 25 and 75% quartile of the data.

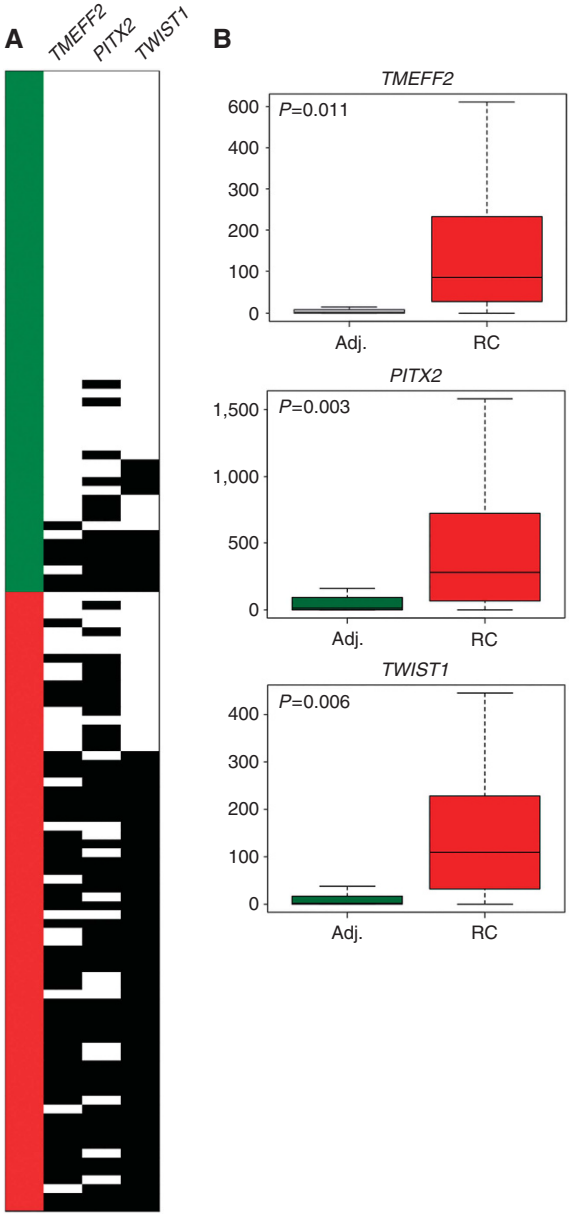

Figure 5. Confirmation of classification signature in a FFPE sample cohort. (**A**) Dichotomous heatmap indicating methylation frequency of *TMEF2*, *PITX2* and *TWIST* (columns) in normal adjacent (green) and tumour (red) samples (rows) isolated from FFPE material using ms-qPCR (black, methylated; white, unmethylated). Samples were defined as methylated if their PMR value was >PMR of an unmethylated control DNA + 10. (**B**) Boxplots indicating relative methylation of tumour and adjacent normal samples analysed as in **A**. Significance was calculated using unpaired *t*-tests. Whiskers depict the 25 and 75% quartile of the data.

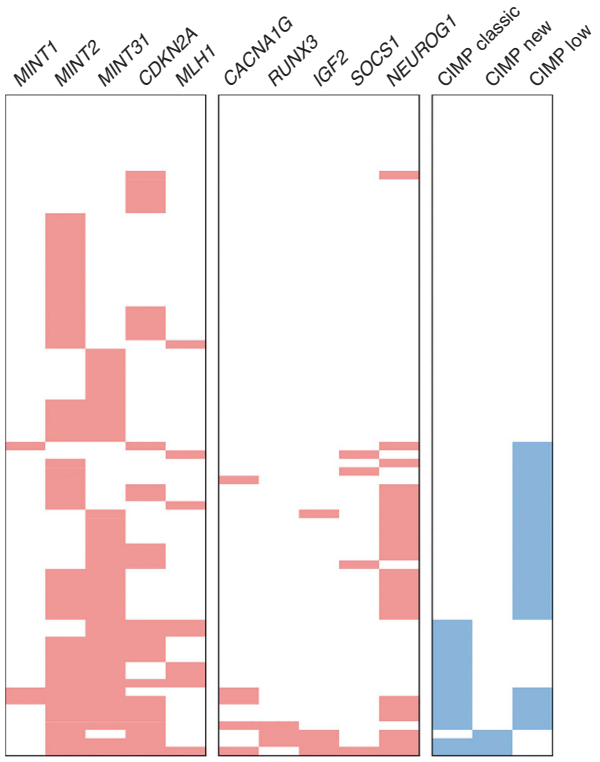

Figure 6. CIMP analysis of rectal tumours from a FFPE sample cohort. Methylation of indicated CIMP markers was performed in 78 rectal tumour samples using MethyLight. The dichotomous heatmap on the left indicates methylated (red) or unmethylated (white) markers in each sample (rows). Methylated regions were integrated into different CIMP panels shown in the histograms on the right (blue) including: CIMP classic (3 out of 5 of the classic makers *MINT1*, *MINT2*, *MINT31*, *CDKN2A* and *MLH1* methylated); CIMP new (3 out of 5 or the new CIMP markers *CACNA1G*, *RUNX3*, *IGF2*, *SOCS1* and *NEUROG1* methylated); and CIMP low (one or two markers of the new panel methylated).

et al, 2013). The transcription factors *TWIST1* and *PITX2* are both hypermethylated in different cancers and were suggested as potential biomarkers. Especially *PITX2* methylation has been well established as a prognostic biomarker for breast and prostate cancer (Mikeska *et al*, 2012). Together, our data suggest that genes commonly methylated in cancers might be useful as markers for rectal cancer diagnostics and it will be interesting to study their prognostic and/or predictive value in a larger cohort in the future.

The signature differentiating tumours from PBMCs contained *PTGS2* (prostaglandin-endoperoxide synthase 2), *DCC* (deleted in colorectal carcinoma) and *TFPI2* (tissue factor pathway inhibitor 2). PTGS2 also known as COX-2 (cyclooxygenase-2) is the rate-limiting enzyme in the prostaglandin biosynthesis. Although overexpression of COX-2 has been reported for gastrointestinal tumours and COX inhibitors are promising antitumour drugs (Jendrossek, 2013), methylation of *COX-2* was detected in 13% of sporadic colorectal cancers and has been associated with CIMP-positive tumours (Toyota *et al*, 2000). Our data imply that *COX2* methylation might also be a marker for rectal cancers, suggesting that COX inhibitors might not be efficacious in rectal tumours.

DCC is a tumour suppressor and both deletion and promoter methylation of *DCC* are involved in its frequent downregulation in colorectal cancer (Tanaka *et al*, 1991; Derks *et al*, 2009).

TFPI2 belongs to a family of Kunitz-type serine protease inhibitors, implicated in tissue remodelling and proliferation. *TFPI2* was suggested to act as a tumour suppressor and it was found methylated in different tumour types including pancreatic, cervical, gastric or prostate carcinoma (Matsubayashi *et al*, 2006; Sova *et al*, 2006; Jee *et al*, 2009; Ribarska *et al*, 2010). Furthermore, *TFPI2* methylation was identified as a specific and sensitive marker for CRC detection in stool and serum (Glockner *et al*, 2009; Hibi *et al*, 2011).

Regarding CIMP occurrence, we identified only 3.8% of tumour samples taken from a retrospective cohort of 78 rectal cancer patients as CIMP positive according to the described marker panel comprised of *CACNA1G*, *IGF2*, *NEUROG1*, *RUNX3* and *SOCS1*

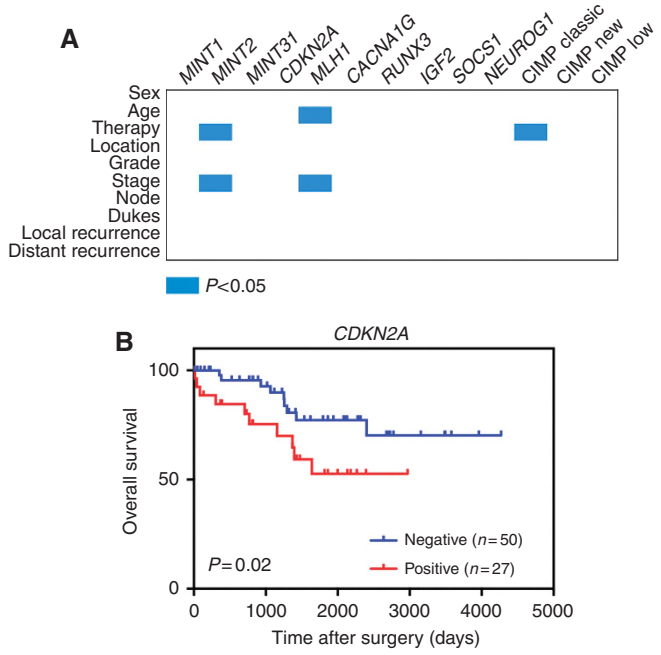

Figure 7. Correlation of methylation data with clinicopathological characteristics. (**A**) Spearman rho correlation matrix of methylation data (horizontal) and clinical variables (vertical). Blue coloured squares indicate significant correlations. (**B**) Kaplan–Meier survival statistics showing significant association between methylation of *CDKN2A* and OS of patients (negative, unmethylated *CDKN2A*; positive, methylated *CDKN2A*).

Table 2. Univariable and multivariable analyses of the overall survival using a Cox proportional hazards model

		Univariable		Multivariable	
Variable	**Subcategory**	**HR**	***P*-value**	**Covariables**	***P*-value**
CDKN2A	Pos = 0, Neg = 1	2.54	0.045		0.017
Nodal status	pN0 = 0, pN + = 1	2.28	0.074	Treated as covariable 1	
Stage	pT2 = 2, pT3 = 3	4.48	0.052	Treated as covariable 2	

Abbreviations: HR = hazard ratio; Neg = negative; Pos = postive.

(Weisenberger *et al*, 2006). This corresponds well to previous data describing CIMP as being rare in rectal cancer and increasing linearly from the rectum to the ascending colon (Jo *et al*, 2012; Yamauchi *et al*, 2012). However, application of the classic CIMP marker panel including *CDKN2A*, *MINT1*, *MINT2*, *MINT31* and *MLH1* revealed a positivity of 19.23%. Importantly, *CDKN2A* methylation was of prognostic relevance in Kaplan–Meier and Cox regression analyses.

The clinical impact of CIMP in rectal cancer has recently been investigated and poor DFS was found in patients with CIMP-positive rectal cancers (Jo *et al*, 2012) and colorectal cancers (Barault *et al*, 2008; Dahlin *et al*, 2010). *KRAS* mutation and *CDKN2A* methylation was found to have an adverse effect on survival and predicts recurrence of rectal cancer in a recent study investigating a cohort of 381 rectal cancers (Kohonen-Corish *et al*, 2014). In contrast to previous studies, we included mainly patients undergoing neoadjuvant radiation therapy or radiochemotherapy in our sample collective. We believe that this might be important, as available tissue for the biomarker assessment of rectal cancer might frequently be only available after such neoadjuvant treatment. Intriguingly, we detected an association of radiation therapy and loss of methylation in the classic CIMP markers. This might imply that radiation therapy interferes with DNA methylation directly or that CIMP-positive cells are more sensitive to radiation therapy and thus selected for in the heterogeneous tumour. There is controversial data relating to alterations of DNA methylation after irradiation of cancer cells (Kim *et al*, 2013). *In vitro* experiments measuring DNA methylation after radiation of MCF7 breast cancer cells observed no significant overall changes of methylation but a delayed hypomethylation of candidate genes identified by a methylation screen ~2 weeks after exposure to radiation (Kuhmann *et al*, 2011). A recent genome-wide methylation study of breast cancer tissue of patients before and after radiation therapy revealed a dose-dependent change in DNA methylation in a set of genes involving both hyper- and hypomethylation (Halvorsen *et al*, 2014). In addition the authors found that methylation of a panel of five genes before radiation was associated with response to radiation therapy. Moreover, a link between DNA methylation and radiation therapy is evident from studies suggesting that pretreatment and DNA hypomethylation of colon cancer cells amongst others with DNA methyltransferase inhibitors can radiosensitise tumour cells (Cho *et al*, 2009; De Schutter and Nuyts, 2009; Hofstetter *et al*, 2010). Furthermore, methylation of genes implicated in DNA repair mechanisms might be surrogates for radiation sensitivity. For example, methylation or loss of *MLH1*, which is implicated in DNA mismatch repair during DNA replication and interacts with apoptotic pathways, was suggested to predict an improved response to radiation (Shin *et al*, 2013). However, other reports suggested that MSI could not predict response to neoadjuvant radiotherapy in advanced rectal cancer (Du *et al*, 2013) or that MLH1-positive tumours show a higher response rate to neoadjuvant chemoradiotherapy (Bertolini *et al*, 2007). Our study included only a limited number of non-irradiated tumours, which might have biased our results. Thus, it will be instrumental to confirm our findings in a larger set of tumours and to further investigate the correlation of mismatch repair deficiency and radiosensitivity in more detail.

In summary, our data underline the relevance of DNA methylation in rectal tumours. The identification of a new set of methylation markers now provides the basis to apply and investigate the potential of these for diagnosis and prognosis for rectal carcinoma.

ACKNOWLEDGEMENTS

GE was funded by the Austrian Science Foundation (FWF) (P 27616 and V 102). Selected analyses were performed using BRB-ArrayTools developed by Dr Richard Simon and BRB-ArrayTools Development Team. We are thankful to Tobias Reinstadler for technical support and conducting qPCR analyses on FFPE samples.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Bae JM, Kim JH, Cho NY, Kim TY, Kang GH (2013) Prognostic implication of the CpG island methylator phenotype in colorectal cancers depends on tumour location. *Br J Cancer* **109**(4): 1004–1012.

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 (2008) Hypermethylator phenotype in sporadic colon cancer: study on a population-based series of 582 cases. *Cancer Res* **68**(20): 8541–8546.

Benard A, van de Velde CJ, Lessard L, Putter H, Takeshima L, Kuppen PJ, Hoon DS (2013) Epigenetic status of LINE-1 predicts clinical outcome in early-stage rectal cancer. *Br J Cancer* **109**(12): 3073–3083.

Bertolini F, Bengala C, Losi L, Pagano M, Iachetta F, Dealis C, Jovic G, Depenni R, Zironi S, Falchi AM, Luppi G, Conte PF (2007) Prognostic and predictive value of baseline and posttreatment molecular marker expression in locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy. *Int J Radiat Oncol Biol Phys* **68**(5): 1455–1461.

Campan M, Weisenberger DJ, Trinh B, Laird PW (2009) MethyLight. *Methods Mol Biol* **507**: 325–337.

Cho HJ, Kim SY, Kim KH, Kang WK, Kim JI, Oh ST, Kim JS, An CH (2009) The combination effect of sodium butyrate and 5-Aza-2′-deoxycytidine on radiosensitivity in RKO colorectal cancer and MCF-7 breast cancer cell lines. *World J Surg Oncol* **7**: 49.

Dahlin AM, Palmqvist R, Henriksson ML, Jacobsson M, Eklof V, Rutegard J, Oberg A, Van Guelpen BR (2010) The role of the CpG island methylator phenotype in colorectal cancer prognosis depends on microsatellite instability screening status. *Clin Cancer Res* **16**(6): 1845–1855.

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 (2010) Identification of a quantitative MINT locus methylation profile predicting local regional recurrence of rectal cancer. *Clin Cancer Res* **16**(10): 2811–2818.

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 (2008) Quantitative analysis of methylation of genomic loci in early-stage rectal cancer predicts distant recurrence. *J Clin Oncol* **26**(14): 2327–2335.

De Schutter H, Nuyts S (2009) Radiosensitizing potential of epigenetic anticancer drugs. *Anticancer Agents Med Chem* **9**(1): 99–108.

Derks S, Bosch LJ, Niessen HE, Moerkerk PT, van den Bosch SM, Carvalho B, Mongera S, Voncken JW, Meijer GA, de Bruine AP, Herman JG, van Engeland M (2009) Promoter CpG island hypermethylation- and H3K9me3 and H3K27me3-mediated epigenetic silencing targets the deleted in colon cancer (DCC) gene in colorectal carcinogenesis without affecting neighboring genes on chromosomal region 18q21. *Carcinogenesis* **30**(6): 1041–1048.

Draht MX, Riedl RR, Niessen H, Carvalho B, Meijer GA, Herman JG, van Engeland M, Melotte V, Smits KM (2012) Promoter CpG island methylation markers in colorectal cancer: the road ahead. *Epigenomics* **4**(2): 179–194.

Du C, Zhao J, Xue W, Dou F, Gu J (2013) Prognostic value of microsatellite instability in sporadic locally advanced rectal cancer following neoadjuvant radiotherapy. *Histopathology* **62**(5): 723–730.

Dudoit S, Fridlyand J, Speed TP (2002) Comparison of discrimination methods for the classification of tumors using gene expression data. *J Am Stat Assoc* **97**(457): 77–87.

Elliott GO, Johnson IT, Scarll J, Dainty J, Williams EA, Garg D, Coupe A, Bradburn DM, Mathers JC, Belshaw NJ (2013) Quantitative profiling of CpG island methylation in human stool for colorectal cancer detection. *Int J Colorectal Dis* **28**(1): 35–42.

Glockner SC, Dhir M, Yi JM, McGarvey KE, Van Neste L, Louwagie J, Chan TA, Kleeberger W, de Bruine AP, Smits KM, Khalid-de Bakker CA, Jonkers DM, Stockbrugger RW, Meijer GA, Oort FA, Iacobuzio-Donahue C, Bierau K, Herman JG, Baylin SB, Van Engeland M, Schuebel KE, Ahuja N (2009) Methylation of TFPI2 in stool DNA: a potential novel biomarker for the detection of colorectal cancer. *Cancer Res* **69**(11): 4691–4699.

Halvorsen AR, Helland A, Fleischer T, Haug KM, Grenaker Alnaes GI, Nebdal D, Syljuasen RG, Touleimat N, Busato F, Tost J, Saetersdal AB, Borresen-Dale AL, Kristensen V, Edvardsen H (2014) Differential DNA methylation analysis of breast cancer reveals the impact of immune signaling in radiation therapy. *Int J Cancer* **135**(9): 2085–2095.

Hibi K, Goto T, Shirahata A, Saito M, Kigawa G, Nemoto H, Sanada Y (2011) Detection of TFPI2 methylation in the serum of colorectal cancer patients. *Cancer Lett* **311**(1): 96–100.

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 (2012) Genome-scale analysis of aberrant DNA methylation in colorectal cancer. *Genome Res* **22**(2): 271–282.

Hofstetter B, Niemierko A, Forrer C, Benhattar J, Albertini V, Pruschy M, Bosman FT, Catapano CV, Ciernik IF (2010) Impact of genomic methylation on radiation sensitivity of colorectal carcinoma. *Int J Radiat Oncol Biol Phys* **76**(5): 1512–1519.

Iacopetta B (2002) Are there two sides to colorectal cancer? *Int J Cancer* **101**(5): 403–408.

Jee CD, Kim MA, Jung EJ, Kim J, Kim WH (2009) Identification of genes epigenetically silenced by CpG methylation in human gastric carcinoma. *Eur J Cancer* **45**(7): 1282–1293.

Jendrossek V (2013) Targeting apoptosis pathways by Celecoxib in cancer. *Cancer Lett* **332**(2): 313–324.

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 (2012) CpG island methylator phenotype infers a poor disease-free survival in locally advanced rectal cancer. *Surgery* **151**(4): 564–570.

Kim JG, Park MT, Heo K, Yang KM, Yi JM (2013) Epigenetics meets radiation biology as a new approach in cancer treatment. *Int J Mol Sci* **14**(7): 15059–15073.

Kohonen-Corish MR, Tseung J, Chan C, Currey N, Dent OF, Clarke S, Bokey L, Chapuis PH (2014) KRAS mutations and CDKN2A promoter methylation show an interactive adverse effect on survival and predict recurrence of rectal cancer. *Int J Cancer* **134**(12): 2820–2828.

Kuhmann C, Weichenhan D, Rehli M, Plass C, Schmezer P, Popanda O (2011) DNA methylation changes in cells regrowing after fractioned ionizing radiation. *Radiother Oncol* **101**(1): 116–121.

Lange CP, Campan M, Hinoue T, Schmitz RF, van der Meulen-de Jong AE, Slingerland H, Kok PJ, van Dijk CM, Weisenberger DJ, Shen H, Tollenaar RA, Laird PW (2012) Genome-scale discovery of DNA-methylation biomarkers for blood-based detection of colorectal cancer. *PLoS One* **7**(11): e50266.

Lange CP, Laird PW (2013) Clinical applications of DNA methylation biomarkers in colorectal cancer. *Epigenomics* **5**(2): 105–108.

Lee SM, Park JY, Kim DS (2012) Methylation of TMEFF2 gene in tissue and serum DNA from patients with non-small cell lung cancer. *Mol Cells* **34**(2): 171–176.

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

Liang G, Robertson KD, Talmadge C, Sumegi J, Jones PA (2000) The gene for a novel transmembrane protein containing epidermal growth factor and follistatin domains is frequently hypermethylated in human tumor cells. *Cancer Res* **60**(17): 4907–4912.

Matsubayashi H, Canto M, Sato N, Klein A, Abe T, Yamashita K, Yeo CJ, Kalloo A, Hruban R, Goggins M (2006) DNA methylation alterations in the pancreatic juice of patients with suspected pancreatic disease. *Cancer Res* **66**(2): 1208–1217.

Mikeska T, Bock C, Do H, Dobrovic A (2012) DNA methylation biomarkers in cancer: progress towards clinical implementation. *Expert Rev Mol Diagn* **12**(5): 473–487.

Monteiro-Reis S, Leca L, Almeida M, Antunes L, Monteiro P, Dias PC, Morais A, Oliveira J, Henrique R, Jeronimo C (2013) Accurate detection of upper tract urothelial carcinoma in tissue and urine by means of quantitative GDF15, TMEFF2 and VIM promoter methylation. *Eur J Cancer* **50**(1): 226–233.

Network CGA (2012) Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**(7407): 330–337.

Ogino S, Kawasaki T, Kirkner GJ, Loda M, Fuchs CS (2006) CpG island methylator phenotype-low (CIMP-low) in colorectal cancer: possible associations with male sex and KRAS mutations. *J Mol Diagn* **8**(5): 582–588.

Pulverer W, Wielscher M, Panzer-Grumayer R, Plessl T, Kriegner A, Vierlinger K, Weinhausel A (2012) The stem cell signature of CHH/CHG methylation is not present in 271 cancer associated 5′UTR gene regions. *Biochimie* **94**(11): 2345–2352.

Radmacher MD, McShane LM, Simon R (2002) A paradigm for class prediction using gene expression profiles. *J Comput Biol* **9**(3): 505–511.

Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang CH, Angelo M, Ladd C, Reich M, Latulippe E, Mesirov JP, Poggio T, Gerald W, Loda M, Lander ES, Golub TR (2001) Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci USA* **98**(26): 15149–15154.

Ribarska T, Ingenwerth M, Goering W, Engers R, Schulz WA (2010) Epigenetic inactivation of the placentally imprinted tumor suppressor gene TFPI2 in prostate carcinoma. *Cancer Genomics Proteomics* **7**(2): 51–60.

Samowitz WS, Curtin K, Wolff RK, Tripp SR, Caan BJ, Slattery ML (2009) Microsatellite instability and survival in rectal cancer. *Cancer Causes Control* **20**(9): 1763–1768.

Shen L, Toyota M, Kondo Y, Lin E, Zhang L, Guo Y, Hernandez NS, Chen X, Ahmed S, Konishi K, Hamilton SR, Issa JP (2007) Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. *Proc Natl Acad Sci USA* **104**(47): 18654–18659.

Shin JS, Tut TG, Yang T, Lee CS (2013) Radiotherapy response in microsatellite instability related rectal cancer. *Korean J Pathol* **47**(1): 1–8.

Sova P, Feng Q, Geiss G, Wood T, Strauss R, Rudolf V, Lieber A, Kiviat N (2006) Discovery of novel methylation biomarkers in cervical carcinoma by global demethylation and microarray analysis. *Cancer Epidemiol Biomarkers Prev* **15**(1): 114–123.

Sun W, Sun Y, Zhu M, Wang Z, Zhang H, Xin Y, Jiang G, Guo X, Zhang Z, Liu Y (2013) The role of plasma cell-free DNA detection in predicting preoperative chemoradiotherapy response in rectal cancer patients. *Oncol Rep* **31**(3): 1466–1472.

Tanaka K, Oshimura M, Kikuchi R, Seki M, Hayashi T, Miyaki M (1991) Suppression of tumorigenicity in human colon carcinoma cells by introduction of normal chromosome 5 or 18. *Nature* **349**(6307): 340–342.

Team RDC (2009) *A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria.

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

Toyota M, Shen L, Ohe-Toyota M, Hamilton SR, Sinicrope FA, Issa JP (2000) Aberrant methylation of the Cyclooxygenase 2 CpG island in colorectal tumors. *Cancer Res* **60**(15): 4044–4048.

Weinhäusel A, Pulverer W (2013) DNA methylation biomarkers for identification of colorectal cancer. European Patent Application EP2886659.4. Available from http://worldwide.espacenet.com/publicationDetails/biblio?DB=worldwide.espacenet.com&II=1&ND=3&adjacent=true&locale=en_EP&FT=D&date=20150625&CC=WO&NR=2015091979A2&KC=A2.

Weinhäusel A, Pichler R, Noehammer C (2009) Methylation Assay. European Patent Application EP 09450020.4. Available from: http://worldwide.espacenet.com/publicationDetails/biblio?DB=worldwide.espacenet.com&II=4&ND=3&adjacent=true&locale=en_EP&FT=D&date=20111124&CC=US&NR=2011287968A1&KC=A1.

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

Yagi K, Akagi K, Hayashi H, Nagae G, Tsuji S, Isagawa T, Midorikawa Y, Nishimura Y, Sakamoto H, Seto Y, Aburatani H, Kaneda A (2010) Three DNA methylation epigenotypes in human colorectal cancer. *Clin Cancer Res* **16**(1): 21–33.

Yamauchi M, Morikawa T, Kuchiba A, Imamura Y, Qian ZR, Nishihara R, Liao X, Waldron L, Hoshida Y, Huttenhower C, Chan AT, Giovannucci E, Fuchs C, Ogino S (2012) Assessment of colorectal cancer molecular features along bowel subsites challenges the conception of distinct dichotomy of proximal versus distal colorectum. *Gut* **61**(6): 847–854.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)