Article

Discovery Medicine 2024; 36(186): 1363–1369
https://doi.org/10.24976/Discov.Med.202436186.126

Methylation Analysis of Colitis-Associated Colorectal Carcinomas

Franziska Haumaier[1], Theresa Dregelies[1], William Sterlacci[1], Raja Atreya[2], Michael Vieth[1,*]

[1]Institute for Pathology, Friedrich-Alexander University Erlangen-Nürnberg, Klinikum Bayreuth, D-95445 Bayreuth, Germany
[2]Medical Department 1, Friedrich-Alexander University Erlangen-Nürnberg, D-91054 Erlangen, Germany
*Correspondence: michael.vieth@fau.de (Michael Vieth)
Published: 20 July 2024

Background: Ulcerative colitis is a well-known inflammatory bowel disease. Patients have an increased risk of developing colitis associated carcinoma (CAC). It is important for patient management to be able to distinguish between ulcerative colitis associated carcinoma and sporadic carcinoma (sCRC). However, this distinction is frequently very challenging. It is not readily possible to differentiate this histologically. However, the diagnosis is crucial for the patient's further treatment and follow-up. An attempt was therefore made to develop a diagnostic regime that would enable a reliable distinction between sCRC and CAC.
Methods: We screened 96 patients analyzing more than 850,000 methylation hotspots, to detect distinct epigenetic patterns between both types of carcinomas. Patients with sporadic carcinoma and colitis-associated carcinoma as well as patients with normal colon and patients with confirmed ulcerative colitis without neoplasia were used for the analysis. By extensively filtering the results, methylation sites relevant to distinguish between CAC and sCRC were identified.
Results: After the results were filtered, three methylation sites relevant to distinguish between CAC and sCRC were identified. For this purpose, methylation limit values were defined, which favor the samples as CAC or sCRC up to a certain methylation value of the methylation sites. The combination of three methylation sites allows a correct assignment to CAC or sCRC in 94.5% of the cases.
Conclusion: The results show that these three methylation sites are promising markers in the diagnosis of CAC vs sCRC. Nevertheless, the diagnosis should always be made in conjunction with histomorphological analyses.

Keywords: ulcerative colitis; ulcerative colitis-associated carcinoma; methylation analysis

Introduction

Ulcerative colitis (UC) is one entity of inflammatory bowel diseases (IBD). In most cases, the inflammation spreads continuously from the rectum toward proximal colon segments [1]. Rarely, a backwash ileitis or isolated cecal involvement can be seen, which is otherwise limited to the colon [2]. It is known that genetic susceptibility, environmental factors, and the gut microbiome conspire to initiate an immune-mediated inflammatory cascade that results in intestinal tissue damage and disease-related complications. Compared to the second most common type of IBD, Crohn's disease, the inflammation in UC is mostly limited to the mucosa [2]. The primary onset of UC is at teenage or young adult age [3], however, it is noteworthy that UC can manifest at any age.

In long standing UC, the risk of malignant transformation increases with the duration, severity and extent of the inflammation [1]. The diagnosis is made in combination with clinical symptoms, endoscopic findings and histology [4]. If a colorectal carcinoma (CRC) is present in UC, this neoplasm has either arisen due to the chronic inflammation of the disease itself, a so-called colitis-associated carcinoma or has developed due to sporadic somatic mutations. Until now, the underlying mechanisms of carcinogenesis are widely unknown. It is known, that concomitant primary sclerosing cholangitis (PSC) additionally increases the risk of malignant transformation [5]. For the patient, the information, on whether the tumor represents a colitis associated carcinoma (CAC) or a sporadic carcinoma (sCRC) is of importance, as the required treatments can differ dramatically. In the early carcinogenesis stages of low grade dysplasia (LGD) and adenomas, guidelines allow local endoscopic resection. In the precancerous stage of high grade dysplasia (HGD), or for carcinoma itself, guidelines foresee a total colectomy, which may have a large impact on the patient's quality of life instead of a limited resection in sCRC. Until now, there is no general procedure to distinguish a colitis-associated from a sporadic lesion. In theory, sCRCs show early Adenomatous polyposis coli (APC) and late p53 gene mutations, whereas CACs show early p53 and late as well as rare APC mutations [1,6,7].

Besides the molecular characterization of gene mutations, epigenetics may be involved in CAC carcinogenesis, also. Especially DNA methylation within promoter regions is of particular interest. Here, methylation of 5'-

C-phosphate-G3' (CpG) islands lead to an inactivation of genes, which occurs by transcriptional silencing [8]. These events are associated with the development of cancer. By the activation of oncogenes, but also by the inactivation of tumor suppressor genes, uncontrolled cell growth is prompted [9]. So far, little is known about DNA methylation patterns in CACs, especially if they differ significantly from those in sCRC. Therefore, our group aimed to identify a signature of methylated genes that may facilitate the differentiation between CAC and sCRC.

For this purpose, a patient cohort consisting of four different groups was used. Patients with sCRC, patients with CAC, patients with UC without neoplasia and a control group of patients with normal colon. A methylation analysis of over 850K methylation sites was performed on all of these samples. The very extensive results were then filtered. First, the average values of all methylation sites within patient groups were calculated. Subsequently, all methylation sites with a difference of less than 20% between sCRC and CAC, UC without neoplasia and CAC, as well as normal colon and CAC were filtered out. Likewise, all positions that showed a high standard deviation within the respective group were removed from the results. The individual samples from the CAC and sCRC groups were then analyzed again in order to establish threshold values for distinguishing between the aforementioned groups and to detect defectors.

Materials and Methods

Sample Collection and DNA Extraction

Formalin fixed and paraffin embedded colon tissue (FFPE) samples of 96 patients (48 male and 48 female) were provided by the Institute for Pathology, Klinikum Bayreuth GmbH. All diagnostic results were confirmed by two independent pathologists with consistent outcomes. The area of tissue required for analysis was marked by the pathologist and then microdissected and isolated by trained molecular biologists. Genomic DNA was isolated from all patients, which were placed into four different groups. All individuals were consecutively chosen and were all in Mayo Score 1 (Nancy Score 0, D2C1A0) or 2 (Nancy Score 1, D2C1A1) [10,11] (when suffering from UC and none had PSC). Group 1 consisted of patients with normal colon with no signs of inflammation (n = 20); group 2 of ulcerative colitis patients (Mayo Scores 1 and 2) with no neoplastic transformation (n = 21); group 3 of patients with sporadic colorectal cancer (n = 32); group 4 of patients with colitis associated colorectal cancer (n = 23). DNA was isolated with the Maxwell® 16 LEV Blood DNA Kit, Promega (Mannheim, Germany), #CAT: AS1290. The experimental protocol was established, according to all ethical guidelines and suggestions of the Helsinki Declaration. The study was approved by the ethics committee of the University Bayreuth (Az. 0 1305/1 - GB). All patients signed a written consent with the contract at time of admission to the hospital that data and all tissue can be used based on the ethical committee's decision for a specific study.

Infinium HD FFPE QC Assay

The DNA quality of all patient samples was assessed with the Infinium HD FFPE QC Assay, Illumina (San Diego, CA, USA), Ref: WG-321-1001, according to the instruction manual.

EZ-96 DNA Methylation Kit

Bisulfite conversion of all samples was carried out with the EZ-96 DNA Methylation™ Kit, ZYMO RESEARCH (Freiburg im Breisgau, Germany), #CAT: D5004.

Infinium HD FFPE Restore Kit

To restore degraded FFPE DNA, the samples were processed by the Infinium HD FFPE Restore Protocol, Illumina (San Diego, CA, USA), Ref: WG-321-1002.

Infinium MethylationEPIC BeadChip

To analyze over 850K methylation sites, all samples were loaded on the Infinium MethylationEPIC BeadChip, Illumina (San Diego, CA, USA), Ref: 20087709.

Data Analysis

All methylation analysis and thereafter the results of all samples were examined by the Genome Analysis Center, Helmholtz Zentrum, Munich, Germany regarding their quality. The results were then further evaluated at the Institute of Pathology, Klinikum Bayreuth. Raw data were further processed using GenomeStudio from Illumina (San Diego, CA, USA). Normalization and background subtraction were carried out.

Statistical Analysis

The statistical analyses were performed with SPSS (IBM SPSS Inc., Armonk, NY, USA, Version 23.0). Two-sided independent *t*-test was conducted and p-values < 0.05 were considered statistically significant.

Results

After processing the data with GenomeStudio, the degree of methylation of 865,918 CpG sites of all 96 samples was evaluated. The results were allocated to the four different patient groups (Table 1).

Filtering the Results

The results of the individual groups were exported from GenomeStudio and further processed using Excel 2016, Microsoft (Remond, WA, USA) In a first step, mean values and standard deviations of the individual methylation sites of the samples within a group were calculated. To

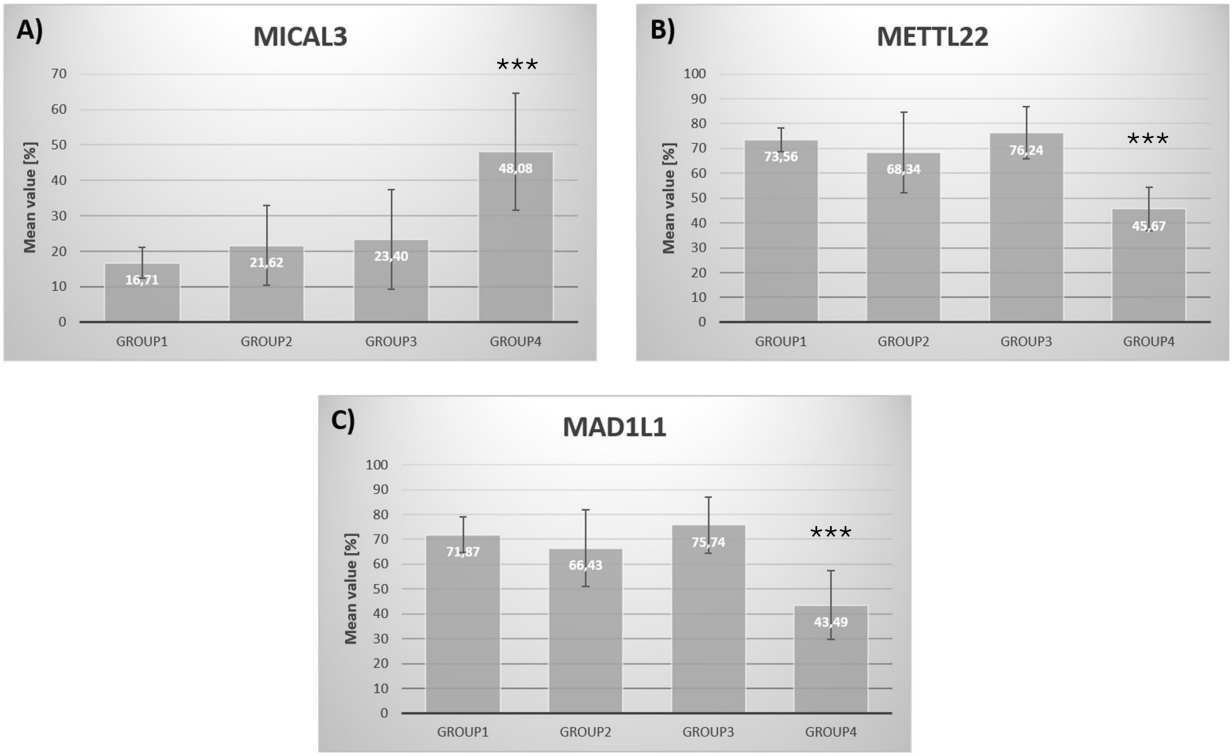

Fig. 1. Mean value of methylation and standard deviation. Groups 1 (Controls with no colonic inflammation), 2 (UC), 3 (sCRC) and 4 (CAC) of CpG-sites of (A) MICAL3; (B) METTL22; (C) MAD1L1 are shown. *** $p \leq 0.001$.

Table 1. Used samples for analysis.

Group	Diagnosis	Number of samples
1	Controls with no colonic inflammation	20
2	UC	21
3	sCRC	23
4	CAC	32

UC, Ulcerative colitis; sCRC, sporadic carcinoma; CAC, colitis associated carcinoma.

define methylation sites relevant for CACs, the results of group 4 were further filtered by removing all CpG sites with a standard deviation greater than 20%. In the next step, differences between colitis patients without (group 2) and with CAC (group 4) were analyzed. In order to achieve this, the degree of methylation in group 2 was subtracted from the filtered results of group 4. Then all results were displayed in which the differences between the two groups were $\geq$20% or $\leq$–20% resulting in 7114 CpG sites. Next, all methylation sites with a standard deviation greater than 20% in group 2 were removed, resulting in 7096 CpG sites. In the next step, differences at these 7096 sites between colitis associated and sporadic carcinomas were analyzed. For this purpose, the differences between the corresponding degrees of methylation in groups 4 and 3 were noted. Differences of $\geq$20% or $\leq$–20% were then displayed and sites with standard deviations greater than 20% in group 3 were removed, resulting in 661 CpG sites. To further reduce the number of CpG sites, the differences between the mean values of the two groups were sorted according to their size. Only differences >20% and <–28% were considered. This resulted in 39 CpG sites. A list of the mean values of groups 4 and 3 of these 39 sites as well as the difference between these groups are listed in **Supplementary Table 1**.

Setting a Limit for Mapping to CAC or sCRC

The degree of methylation of groups 4, 3 and 1 was then compared. While there are statistically significant differences between groups 4 and 3 and between groups 4 and 1, the values of groups 3 and 1 are very similar (**Supplementary Table 2**)

To develop a method to distinguish sporadic from colitis associated carcinomas, the individual patient samples of the 39 CpG sites were considered and not the mean value of the groups. With the mean value and the standard deviation from group 4 (CAC), a threshold was determined for each location, which decides whether a sample is seen as colitis associated or sporadic carcinoma. The threshold shown in Table 2 indicates which degree of methylation is assigned to CAC. This threshold was then used to calculate both the absolute and the relative frequency with which the samples from groups 4 and 3 could be correctly assigned. All these values are listed in **Supplementary Table 3**.

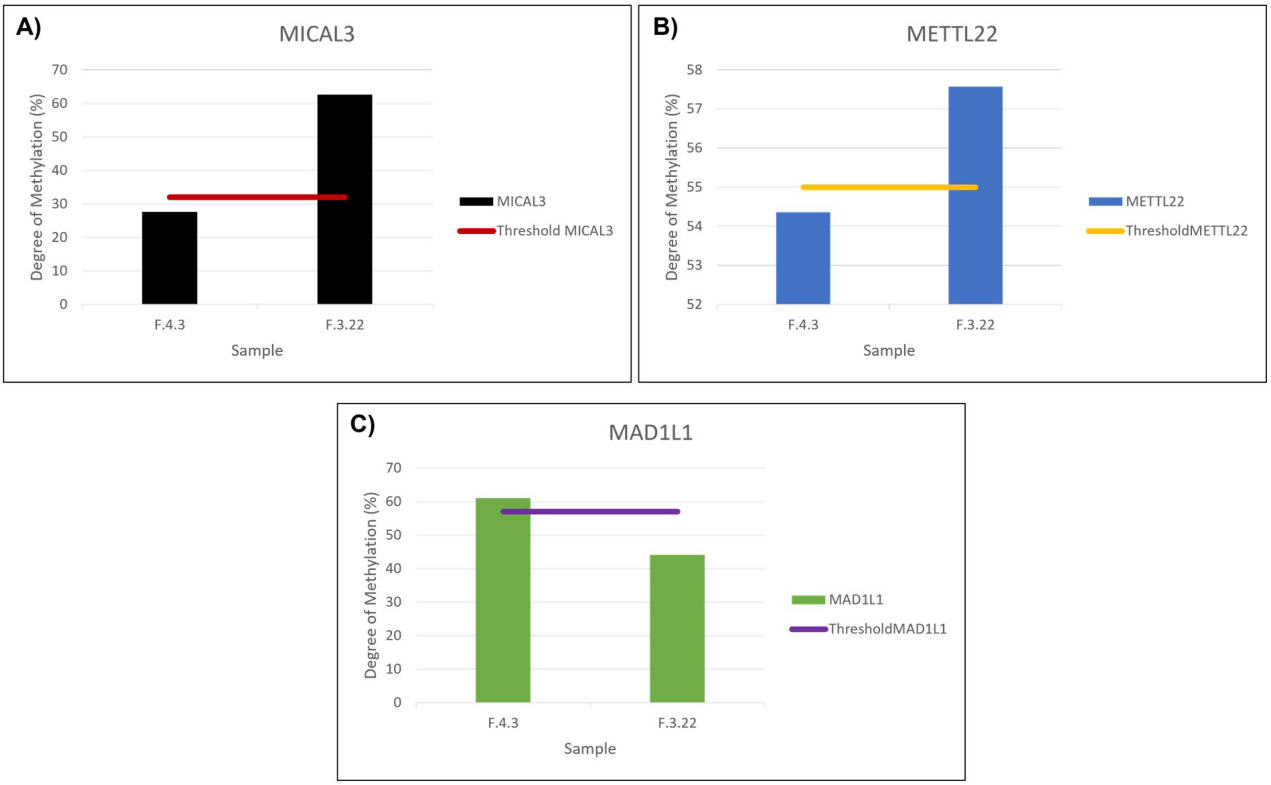

Fig. 2. Graphical analysis of two patient samples. The degree of methylation of samples F.4.3 and F.3.22 of the CpG sites MICAL3 (A), METTL22 (B) and MAD1L1 (C) is shown. In addition the threshold of methylation to indicate which degree of methylation is assigned to CAC is shown as a line for each position. (A) MICAL3: with a degree of methylation greater than 32%, the sample is counted as associated with colitis. (B) METTL22: with a degree of methylation less than 55%, the sample is counted as CAC. (C) MAD1L1: with a degree of methylation less than 57%, the sample is counted as being associated with UC.

Combining Different Methylation Sites to Increase Correct Assignment

Four specific CpG motifs enabled to assign of the sample to the correct patient-group in 90.9% of the cases (**Supplementary Table 3**). Furthermore, an attempt was made to determine whether there is a combination of CpG sites that would enable a higher correct allocation rate. One sample was assigned to the CAC group, if at least one of three markers indicated the presence of colitis-associated, different combinations of CpG sites were analyzed. The combination of three of the 90.9 % correctly assigned CpG sites – MICAL3 (cg05367846), METTL22 (cg11298899) and MAD1L1 (cg15543199) – enabled a correct assessment of the patient samples in 94.5% of the cases. Fig. 1 shows the mean values of the individual groups of these three CpG sites. As indicated in the Fig. 1, there are hardly any differences in the degree of methylation between groups 1, 2 and 3, while group 4 differs significantly from these values. The *p*-values of these CpG sites are shown in Table 3.

Table 2 shows the mean value of the methylation of groups 4 and 3 as well as the threshold which indicates whenever a sample is assigned to colitis-associated carcinoma.

Example of Assigning Two Samples

The graphic representation of the assignment of one sample from group 4 and one sample from group 3 is shown in Fig. 2. The assignment of all samples of groups 4 and 3 is shown in **Supplementary Figs. 1,2**.

Documentation of Incorrect Assignments

As already described, a sample was assigned to the CAC group, if at least one of three markers indicated the presence of a CAC. In sample F.4.3 (Fig. 2), the CpG site of METTL22 corresponds to a CAC. The other markers (MICAL3, MAD1L1) would correspond to sporadic carcinoma. Since one maker (METTL22) indicates CAC, the sample is assigned to this group. Since this is a CAC, it was correctly assigned. Fig. 2 also shows a sample from group 3 (F.3.22) – a sCRC. However, since two markers (MICAL3, MAD1L1) indicate a CAC, this sample was possibly assigned to the wrong group.

During the evaluation, three of the 55 tumor patients could not be assigned correctly. One sample from the CAC group (F.4.1) would have been detected as sCRC (Fig. 3A), while two samples from the sporadic group (F.3.22, F.3.26) would have been detected as CAC (Fig. 3B,C).

Table 2. Calculated threshold to indicate which degree of methylation is assigned to CAC.

UCSC_REFGENE_NAME	Mean value group 4 (%)	Standard deviation group 4 (%)	Threshold group 4	Absolute frequency	Relative frequency (%)
MICAL3	48.08	16.45	≥32	50	90.91
METTL22	45.67	76.24	≤55	50	90.91
MAD1L1	43.49	75.74	≤57	50	90.91

Table 3. Statistical comparison of the individual groups.

p-value/CpG Site	MICAL3	MAD1L1	METTL22
Group 1/Group 4	$p = 0.000/p = 5.9379 \times 10^{-9}$	$p = 0.000/p = 2.8111 \times 10^{-8}$	$p = 0.000/p = 6.5224 \times 10^{-12}$
Group 2/Group 4	$p = 0.000/p = 7.1544 \times 10^{-7}$	$p = 0.000/p = 0.00002$	$p = 0.000/p = 0.000024$
Group 3/Group 4	$p = 0.000/p = 6.6009 \times 10^{-9}$	$p = 0.000/p = 1.3977 \times 10^{-8}$	$p = 0.000/p = 5.8181 \times 10^{-15}$

The histological differentiation between CAC and sCRC may be very difficult or even impossible since both tumors can show similar histomorphological characteristics. Although colitis-associated dysplasia can often be of the sporadic, conventional type, other distinctive, non-conventional morphologies are being recognized in patients with inflammatory bowel disease. Non-conventional types mostly appear as hypermucinous, signet ring cell, serrated or micropapillary, or may show a mixed histology [12,13]. These non-conventional patterns, or just the fact that the lesion is not of the typical sporadic type, assist the pathologist in favouring a diagnosis of CAC versus sCRC.

Discussion

In this study, the degree of methylation of over 850K CpG sites of a patient cohort of different subgroups was analyzed. The aim was to distinguish between CAC and sCRC at the level of DNA-methylation. Here, it was important to understand, that the differences do not only exist between both types of carcinomas, but also between patients with normal colon as well as patients with UC with no dysplasia. By analyzing methylation sites, 39 potential CpG sites were identified that meet the aforementioned criteria. To increase the probability of the correct assignment to CAC or sCRC, a combination of three CpG sites was chosen, which enables correct assignment in 94.5% of cases. In Fig. 1 the clear difference in the level of methylation between groups 1, 2 and 3 compared to group 4 is shown.

The samples that were "incorrectly" assigned in this analysis, were again examined by a pathologist. Even after repeated microscopy, the original diagnosis remains unchanged. This shows the need for the interaction of histology and molecular analysis. In these cases, the histology leaves no doubt about the diagnosis. Obviously, the morphology represents a spectrum and there may be a few cases that cannot be assigned unequivocally solely based on the methylation status but together with histological morphology.

The CpG sites detected here are of the genes MICAL3 (Microtubule Associated Monooxygenease), MAD1L1 (Mitotic Spindle Assembly Checkpoint Protein) and METTL22 (Methyltransferase-Like Protein). It is known that MAD1L1 may play a role in cell cycle control

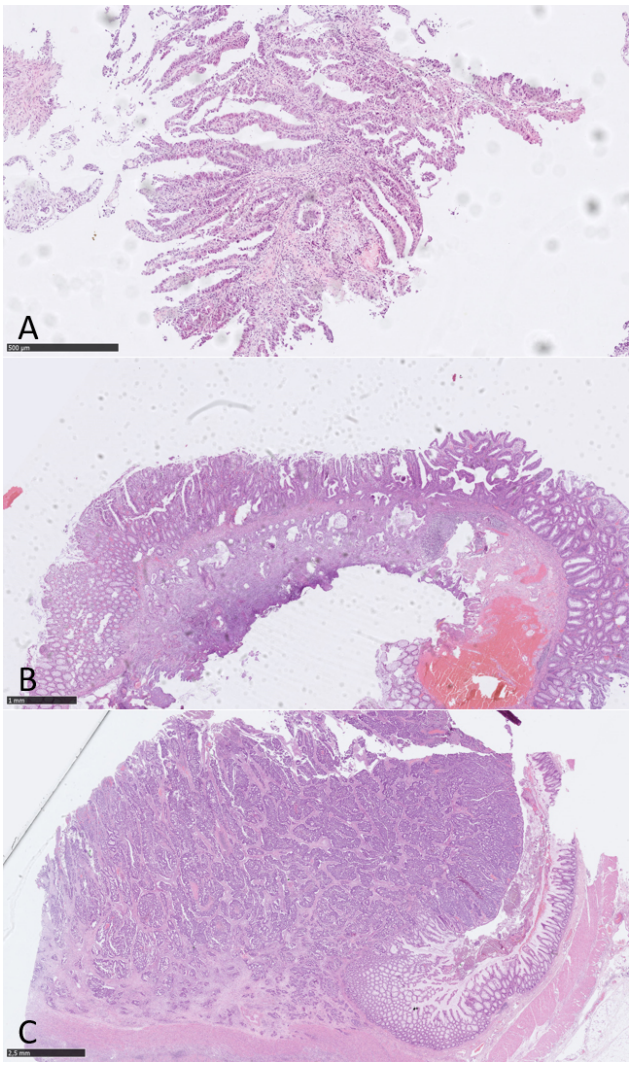

Fig. 3. Hematoxylin-Eosin (HE) staining of the 3 neoplasms from patients who were incorrectly assigned by means of methylation analysis. (A) CAC; (B) sCRC; (C) sCRC. Scale bar: 500 µm (A), 1 mm (B), 2.5 mm (C).

and tumor suppression (GeneCards, The Humane Gene Database). To the best of our knowledge, a connection between these genes and UC has not been described yet.

As a result, it can be assumed that the markers identified are colitis-carcinoma specific sites. However, it was also apparent that no 100% correct classification would be possible based solely on methylation analyses. A crosstalk between clinical, histological and genetic analyses must be ensured in the future.

To validate the results found in this study, a larger group of patients needs to be analyzed specifically for these three methylation sites in future analyses. Furthermore, it will be interesting to learn whether these differences between sporadic and colitis-associated methylation sites can already be found in precursor lesions such as low-grade intraepithelial neoplasia, to further validate our results. One result is that we were able to identify a small set of lesions that were possibly assigned to the wrong group of either colitis-associated or sporadic neoplasia. Further validation in larger cohorts is needed here as well. The main problem is that such lesions are rare, and the cohorts must be rather large to enable sufficient validation. The limitations of the study are the formalin-fixed nature of the tissue, as this can lead to DNA double-strand breaks and thus to tissue that cannot be evaluated, as well as the time required and the smaller size of the patient cohort.

There are already other studies that analysed hypermethylated genes in UC [14,15]. However, the neoplastic UC samples were only compared to normal colon. There was no comparison to sporadic carcinomas. Another study also searched for methylated genes in UC using database research [16]. However, the genes detected being considered as relevant were filtered out in our analyses due to the lack of differences between the different groups we used.

Conclusion

Our results illustrate how relevant control groups used in the study are for distinguishing between sporadic and UC-associated carcinomas and that morphological analysis outrules sole methylation status data. The results show that the three methylation sites are promising markers in the diagnosis of colitis-associated carcinoma. Nevertheless, the diagnosis should always be made in conjunction with histomorphology. Upon further validation, these molecular markers could therefore be applied to facilitate differentiation between CAC and sCRC prospectively and within larger scale studies.

Availability of Data and Materials

The datasets used or analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

FH, TD, WS, RA, MV: conception, writing, proofreading, interpretation of data, planning of lab experiments; FH: lab experiments; TD: statistical analysis. All authors contributed significantly to editorial changes of important content. All authors read and approved the final manuscript. All authors have participated sufficiently in the work and agreed to be accountable for all aspects of the work.

Ethics Approval and Consent to Participate

The experimental protocol was established, according to all ethical guidelines and suggestions of the Helsinki Declaration. The study was approved by the ethics committee of the University Bayreuth (Az. 0 1305/1 - GB). All patients signed a written consent with the contract at time of admission to the hospital that data and all tissue can be used based on the ethical committee's decision for a specific study.

Acknowledgment

We thank the associate editor and the reviewers for their useful feedback that improved this paper.

Funding

The project was financed by the Forschungskommission of Klinikum Bayreuth GmbH for Franziska Haumaier (Fund number: 50010).

Conflict of Interest

The authors declare no conflict of interest.

Supplementary Material

Supplementary material associated with this article can be found, in the online version, at https://doi.org/10.24976/Discov.Med.202436186.126.

References

[1] Rogler G. Chronic ulcerative colitis and colorectal cancer. Cancer Letters. 2014; 345: 235–241.

[2] Porter RJ, Kalla R, Ho GT. Ulcerative colitis: Recent advances in the understanding of disease pathogenesis. F1000Research. 2020; 9: F1000 Faculty Rev-294.

[3] Cosnes J, Gower-Rousseau C, Seksik P, Cortot A. Epidemiology and natural history of inflammatory bowel diseases. Gastroenterology. 2011; 140: 1785–1794.

[4] Kucharzik T, Dignass A, Atreya R, Bokemeyer B, Esters P, Herrlinger K, *et al*. Aktualisierte S3-Leitlinie Colitis ulcerosa (Version 6.1)–Februar 2023–AWMF-Registriernummer: 021-009. Zeitschrift für Gastroenterologie. 2023; 61: 1046–1134.

[5] Shah SC, Itzkowitz SH. Colorectal Cancer in Inflammatory

Bowel Disease: Mechanisms and Management. Gastroenterology. 2022; 162: 715–730.e3.

[6] Markowitz SD, Bertagnolli MM. Molecular origins of cancer: Molecular basis of colorectal cancer. The New England Journal of Medicine. 2009; 361: 2449–2460.

[7] Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, *et al*. Genetic alterations during colorectal-tumor development. The New England Journal of Medicine. 1988; 319: 525–532.

[8] Costello JF, Plass C. Methylation matters. Journal of Medical Genetics. 2001; 38: 285–303.

[9] Moarii M, Boeva V, Vert JP, Reyal F. Changes in correlation between promoter methylation and gene expression in cancer. BMC Genomics. 2015; 16: 873.

[10] Lang-Schwarz C, Agaimy A, Atreya R, Becker C, Danese S, Fléjou JF, *et al*. Maximizing the diagnostic information from biopsies in chronic inflammatory bowel diseases: recommendations from the Erlangen International Consensus Conference on Inflammatory Bowel Diseases and presentation of the IBD-DCA score as a proposal for a new index for histologic activity assessment in ulcerative colitis and Crohn's disease. Virchows Archiv. 2021; 478: 581–594.

[11] Lang-Schwarz C, Angeloni M, Agaimy A, Atreya R, Becker C, Dregelies T, *et al*. Validation of the 'Inflammatory Bowel Disease-Distribution, Chronicity, Activity [IBD-DCA] Score' for Ulcerative Colitis and Crohn's Disease. Journal of Crohn's & Colitis. 2021; 15: 1621–1630.

[12] Gui X, Iacucci M, Ghosh S, Ferraz JGP, Lee S. Revisiting the distinct histomorphologic features of inflammatory bowel disease-associated neoplastic precursor lesions in the SCENIC and post-DALM Era. Human Pathology. 2020; 100: 24–37.

[13] Harpaz N, Goldblum JR, Shepherd NA, Riddell RH, Rubio CA, Vieth M, *et al*. Colorectal dysplasia in chronic inflammatory bowel disease: a contemporary consensus classification and interobserver study. Human Pathology. 2023; 138: 49–61.

[14] Kang K, Bae JH, Han K, Kim ES, Kim TO, Yi JM. A Genome-Wide Methylation Approach Identifies a New Hypermethylated Gene Panel in Ulcerative Colitis. International Journal of Molecular Sciences. 2016; 17: 1291.

[15] Karatzas PS, Mantzaris GJ, Safioleas M, Gazouli M. DNA methylation profile of genes involved in inflammation and autoimmunity in inflammatory bowel disease. Medicine. 2014; 93: e309.

[16] Zhang D, Yan P, Han T, Cheng X, Li J. Identification of key genes and biological processes contributing to colitis associated dysplasia in ulcerative colitis. PeerJ. 2021; 9: e11321.