



OPEN Dynamic DNA methylation changes during colorectal oncogenesis with insights from adenoma stages

Alexis Overs^{1,2}✉, Chloé Molimard³, Jules Durand², Frédéric Bibeau³, Laurent Arnould⁴, Franck Monnien³, Claire Clavier², Christophe Borg⁵, Michael Guittaut², Jean-Paul Feugeas^{1,2}, Eric Hervouet², Paul Peixoto^{2,7} & Zohair Selmani^{2,6,7}

The dynamics of colorectal epigenetics within the adenoma stages of oncogenesis remain undocumented. In this study, we investigated DNA methylation dynamics in colorectal cancer oncogenesis from non-tumor colon tissue to low-grade, high-grade adenoma and adenocarcinoma. The methylome of 12 low-grade and 19 high-grade colorectal adenomas was determined via the EPIC v1 Human Methylation BeadChip. These methylation profiles were complemented with the methylomes of 206 non-tumor colon and 22 colon adenocarcinoma samples from the GEO and TCGA databases. Differentially methylated CpGs were identified via Student's *t* test and used to monitor the evolution of the colon methylome during oncogenesis. The differentially methylated promoters were used to infer the associated biological process via gene ontology and the evolution of the methylation of 34 described colorectal cancer DNA methylation biomarkers was explored. A total of 11.9% of the colon methylome was significantly altered ($q < 10^{-4}$) during oncogenesis, with half corresponding to DNA demethylation. Of which, 67.4% occurred during the transition from non-tumor colon tissue to low-grade adenoma. A total of 9% of the DNA methylation changes were specific to low-grade and/or high-grade adenomas. The biological pathways related to the sensory perception of odor and stimulus were hypomethylated early, nucleic acid metabolic process were methylated early, post-transcriptional regulation were transiently hypomethylated and mitotic cell cycle were transiently methylated. Twenty-one out of 34 the biomarkers were methylated in low-grade adenomas and 11 out of 34 in high-grade adenomas. This suggests that they could be used to distinguish stages of oncogenesis. This study provides insight into the dynamics of colonic epigenetics during oncogenesis, with early DNA methylation changes in low-grade adenomas associated with transient DNA methylation changes. However, the causality of these changes remains to be elucidated. This study also explores the evolution of known biomarkers and their clinical applications for indirectly asserting the tumor's stage.

Keywords Colon adenoma, DNA methylation, Dynamic, Epigenetics, Colorectal cancer, Biomarker, Oncogenesis

Colorectal cancer is a major health problem in high-income countries, being the second leading cause of cancer-related deaths, with more than 900,000 cases reported worldwide in 2022¹. The oncogenesis of colorectal cancer has been described for most cases as adenocarcinoma arising from colonic dysplasia, from low-grade adenoma to high-grade adenoma to adenocarcinoma². The World Health Organization (WHO) Classification of Tumors of the Digestive System defines conventional colorectal adenoma as a benign, premalignant neoplasm composed of dysplastic epithelium. A two-tiered stratification is used to classify adenomas into low-grade and high-grade adenomas by a pathologist³.

¹Bioinformatique et Big Data Au Service de La Santé, Centre Hospitalier Universitaire de Besançon, 3 Bd Fleming, 25000 Besançon, France. ²INSERM UMR1098 RIGHT, Université Marie et Louis Pasteur, Besançon, France 25000. ³Anatomie et cytologie pathologiques, Centre Hospitalier Universitaire de Besançon, 3 Bd Fleming, 25000 Besançon, France. ⁴Anatomie et cytologie pathologiques, Centre Georges François Leclerc, 1 Rue du Professeur Marion, 21000 Dijon, France. ⁵Oncologie médicale, Centre Hospitalier Universitaire de Besançon, 3 Bd Fleming, 25000 Besançon, France. ⁶Oncobiologie, Centre Hospitalier Universitaire de Besançon, 3 Bd Fleming, 25000 Besançon, France. ⁷These authors jointly supervised this work: Paul Peixoto and Zohair Selmani. ✉email: aovers@chu-besancon.fr

Given the central role of adenomas in colorectal oncogenesis, the study of molecular alterations such as DNA methylation is essential to better understand tumor progression and identify potential biomarkers.

Indeed, DNA methylation has emerged as a hallmark of cancer⁴ and a marker of cellular aging that predicts cancer risk in affected tissues^{5–8}. DNA methylation has also been shown to vary with anatomic location within the colon and sex⁹.

DNA methylation in colorectal adenomas has been studied previously^{10–13}, but few methylome-wide analyses of colorectal oncogenesis have been performed via the Illumina EPIC Methylation Beadchip¹⁴. The published studies did not differentiate between low-grade and high-grade adenomas^{10–14}.

Large-scale DNA methylation changes have been described during colorectal oncogenesis, with global DNA hypomethylation responsible for genomic instability associated with hypermethylation of tumor suppressor gene promoters.

Studying the DNA methylation process may provide a better understanding and identify biomarkers of different stages of colorectal oncogenesis.

The aim of this study was to investigate DNA methylation changes during the early stages of colorectal oncogenesis, with a distinction between low-grade and high-grade adenomas to explore the biological pathways altered during oncogenesis and the evolution of known colorectal cancer biomarkers for clinical applications.

Methods

Samples

Thirty-one colon adenomas were obtained from formalin-fixed paraffin-embedded (FFPE) samples from the Tumorothèque régionale de Franche-Comté (Table 1), including twelve low-grade adenomas and nineteen high-grade adenomas.

Grading of adenoma dysplasia was performed via the 2019 WHO Classification of Tumors of the Digestive System, which uses a two-tiered system to distinguish low-grade adenomas from high-grade adenomas. The difference between high-grade and low-grade adenomas was determined by a digestive pathologist by the presence of complex architectural patterns and cytologic features indicative of high-grade dysplasia³.

In association with these local samples, non-tumor colon and colon adenocarcinoma methylomes were downloaded from GEO.

The non-tumor colon methylomes from the GSE132804 series ($n = 206$) were generated by Wang et al.¹⁵, who analyzed the epigenetic aging of 206 colon tissues in detail. Adenocarcinoma methylomes from the GSE149282 series ($n = 10$) were generated by Muhiddin et al.¹⁶, who performed open chromatin profiling. Twelve additional

	CHU of Besançon		HCMI projet	GEO Database			
	Low-grade adenoma $n = 12$	High-grade adenoma $n = 19$	Colonic adenocarcinoma $n = 12$	GSE149282 Colonic adenocarcinoma $n = 10$	GSE132804 Non-tumor colonic tissues $n = 206$	GSE48684 normal colon $n = 41$, adenoma $n = 42$ adenocarcinoma $n = 64$	GSE42921 normal colon $n = 12$
Beadchip	EPICv1	EPICv1	EPICv1	EPICv1	EPICv1	450k	450k
Age							
Mean	74	71.4	66	66.3	59.5	NA	14.7
Min - Max	60–86	46–91	51–76	59–73	19–85	NA	9–17
Stage							
I	n/a	n/a	1	0	n/a	0	n/a
II	n/a	n/a	2	1	n/a	0	n/a
III	n/a	n/a	5	7	n/a	0	n/a
IV	n/a	n/a	2	2	n/a	0	n/a
NA	n/a	n/a	2	0	n/a	64	n/a
Sex							
Male	5	12	9	5	97	58	8
Female	7	7	3	5	109	89	4
Histology							
Adenocarcinoma type NOS	n/a	n/a	12	10	n/a	0	n/a
NA	n/a	n/a	0	0	n/a	64	n/a
Anatomic site							
Right colon	4	10	2	0	0	27	0
Left colon	8	7	7	0	206	39	0
Rectum	0	0	0	0	0	1	0
NA	0	2	3	10	0	80	12

Table 1. Clinical characteristics of the patients. NA: not available, n/a: not applicable.

adenocarcinoma methylomes were obtained from the Human Cancer Models Initiative (HCMI), excluding organoid methylation data.

Three samples, one low-grade adenoma from the Tumorothèque régionale de Franche-Comté and two adenocarcinomas from GSE149282 had a methylated *MLH1* promoter (Supplementary Fig. 1), suggesting MSI status and were therefore excluded from the study.

DNA methylation assessment

DNA was extracted from FFPE samples via the QIAmp DNA Mini Kit[®] (Qiagen, Netherlands) according to the manufacturer's instructions. DNA quantification was performed via a Qubit[®] fluorometer (Invitrogen, USA). DNA was bisulfited and converted via the Diagenode Premium Bisulfite Kit[®] according to the manufacturer's instructions.

Methylomes were obtained via Illumina Methylation EPIC BeadChip v1 and analyzed via R software v4.4.2¹⁷ and the tidyverse suite¹⁸, minfi¹⁹, missMethyl²⁰, limma²¹ and InfiniumPurify²² packages.

Bioinformatics analysis

Methylation values

The idat files were processed into beta values in the same batch via the minfi package. The beta values were then normalized via functional normalization associated with background and dye bias correction²³. After normalization, the beta values were converted to M values²⁴. For a given CpG, the beta value is computed from the light signal of the BeadChips probes and can be interpreted as the percentage of methylation of the CpG. M values are an arithmetic transformation of beta values and are the logarithm of the ratio of methylated to unmethylated intensity. Negative M values are unmethylated (beta value < 0.5), positive M values are methylated (beta value > 0.5), and M values near zero are intermediate (beta value approximately 0.5). Probes on the X and Y chromosomes ($n = 19090$ and 537 , respectively) were excluded from the analysis.

Colonic tissue content

The content of colon tissue was estimated via the getPurity function of the InfiniumPurity package²², with non-tumor colon samples used as a reference. The tumor cellularity of the adenocarcinomas and the fraction of dysplastic cells in the adenomas were estimated separately. The getPurity function identifies 1,000 differentially methylated CpGs with high variance between non-tumor and tumor samples. These CpGs were then used to estimate the cellularity of each sample via density evaluation of a Gaussian kernel.

PCA

Principal component analysis of the methylation data was performed via the R package FactoMineR²⁵. The number of principal components was set to 17, and the data were not scaled because of the bimodal distribution of M values.

Differentially methylated CpGs

Differential methylation analysis was performed via the limma R package. Linear models were fitted to M values with the lmFit function, and empirical Bayes was applied using eBayes to improve variance estimation.

A differentially methylated CpG (DMC) was considered significant if the q value was less than 10^{-4} and if there was an absolute difference in methylation of at least 2 M values.

Evolution of the methylation status

The evolution of the methylation status was inferred based on the methylation status of different tissue types, assuming an evolution from non-tumor tissue to low-grade adenoma, then to high-grade adenoma, and finally, to adenocarcinoma. “Definitive methylation” changes are defined as modifications in the methylation status between non-tumor colonic tissue and adenocarcinoma. These changes can occur during the transition to low-grade or high-grade adenomas, or during the transition to adenocarcinoma. “Transitory methylation” changes are defined as having the same methylation status between non-tumor tissue and adenocarcinoma, but a different status from low-grade and/or high-grade adenomas.

Gene ontology

Gene ontology enrichment analysis was performed on the Gene Ontology consortium²⁶ via the missMethyl package²⁰. The gene lists were generated via the DMC in the promoter region (TSS200, TSS1500 and 1st exon) compared to the list of genes whose promoters were covered by the methylation BeadChip.

Validation with 450k BeadChip datasets

Methylation changes were verified with two external GEO series acquired with the Infinium Methylation 450k BeadChip. A total of 41 non-tumor colons, 42 colon adenomas and 64 colon adenocarcinomas were obtained from the GSE48684 series, and 12 non-tumor colons were obtained from the GSE42921 series. These datasets did not discriminate between low-grade and high-grade adenomas. For comparison with the external dataset, low-grade and high-grade adenomas from our dataset were pooled.

Results

Methylation changes during colonic oncogenesis

Tumor cellularity

Tumor cellularity was significantly lower in the adenoma samples than in the adenocarcinoma samples ($p < 10^{-5}$). There were no significant differences between low-grade and high-grade adenomas ($p = 0.059$) (Fig. 1).

The ratio of tumor to non-tumor cellularity was evaluated with the InfinumPurify package. Tumor cellularity was significantly higher in adenocarcinoma samples than in adenoma samples ($p < 10^{-4}$, Wilcoxon test) and no difference was observed between low-grade and high-grade adenomas ($p = 0.059$, Wilcoxon test).

Global distribution

The PCA grouped the samples according to their histologic type. The non-tumor group was randomly subsampled ($n = 20$) to match the number of dysplasia and adenocarcinoma samples. The first dimension of the PCA separates non-tumor tissues from adenomas and adenocarcinomas. The second dimension separates adenomas from adenocarcinomas (Fig. 2).

Principal component analysis of non-tumor colonic tissue, adenomas, and adenocarcinomas methylation data revealed segregation of the samples by histology. The non-tumor colonic tissues were randomly downsampled ($n = 20$) to be on the same scale as the number of low-grade adenomas ($n = 12$), high-grade adenomas ($n = 19$) and adenocarcinomas ($n = 22$).

Major epigenetic changes

Overall, 11.9% of the methylome is significantly modified during oncogenesis (linear models and empirical Bayes, $q < 10^{-4}$ and an absolute difference of at least 2 M value). The observed changes are equally divided between hypomethylation and methylation. Approximately 67.4% of DNA methylation changes occur during the transition from non-tumor colon tissue to low-grade adenoma (Fig. 3). Methylation mostly occurred in CpG islands and CpG shores and hypomethylation mostly occurred in CpG shelves and open seas (Supplementary Table 1).

However, we observed that some methylation changes are specific to adenoma samples. These adenoma-specific methylation changes are less common than the definitive changes and represent approximately 9% of the methylation changes during oncogenesis.

A: Mean M values of differentially methylated CpGs across colon tissue types, separated into four groups based on methylation profile distribution: Definitive methylation (32051 CpGs), Transitory methylation (822 CpGs), Definitive unmethylation (27735 CpGs) and Transitory unmethylation (4227 CpGs). **B:** Distribution of differentially methylated CpGs according to the linear model and empirical Bayes ($q < 10^{-4}$ and an absolute difference of at least two M value).

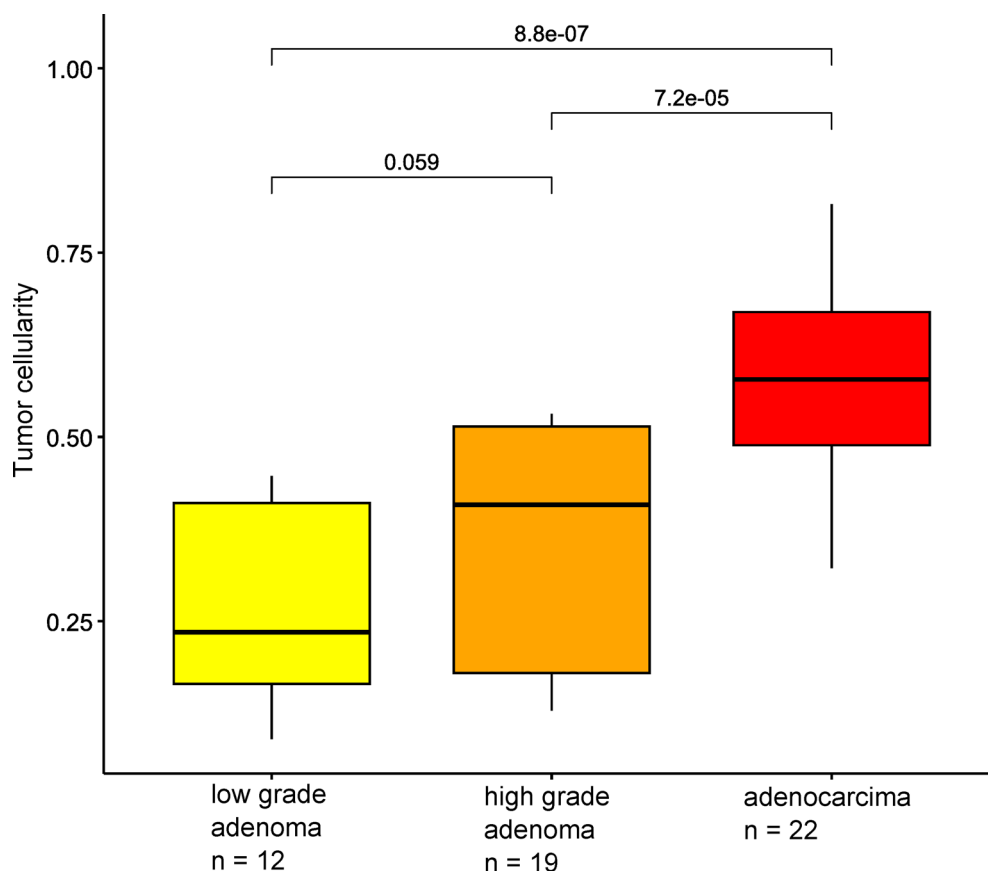


Fig. 1. Tumor cellularity according to histologic types.

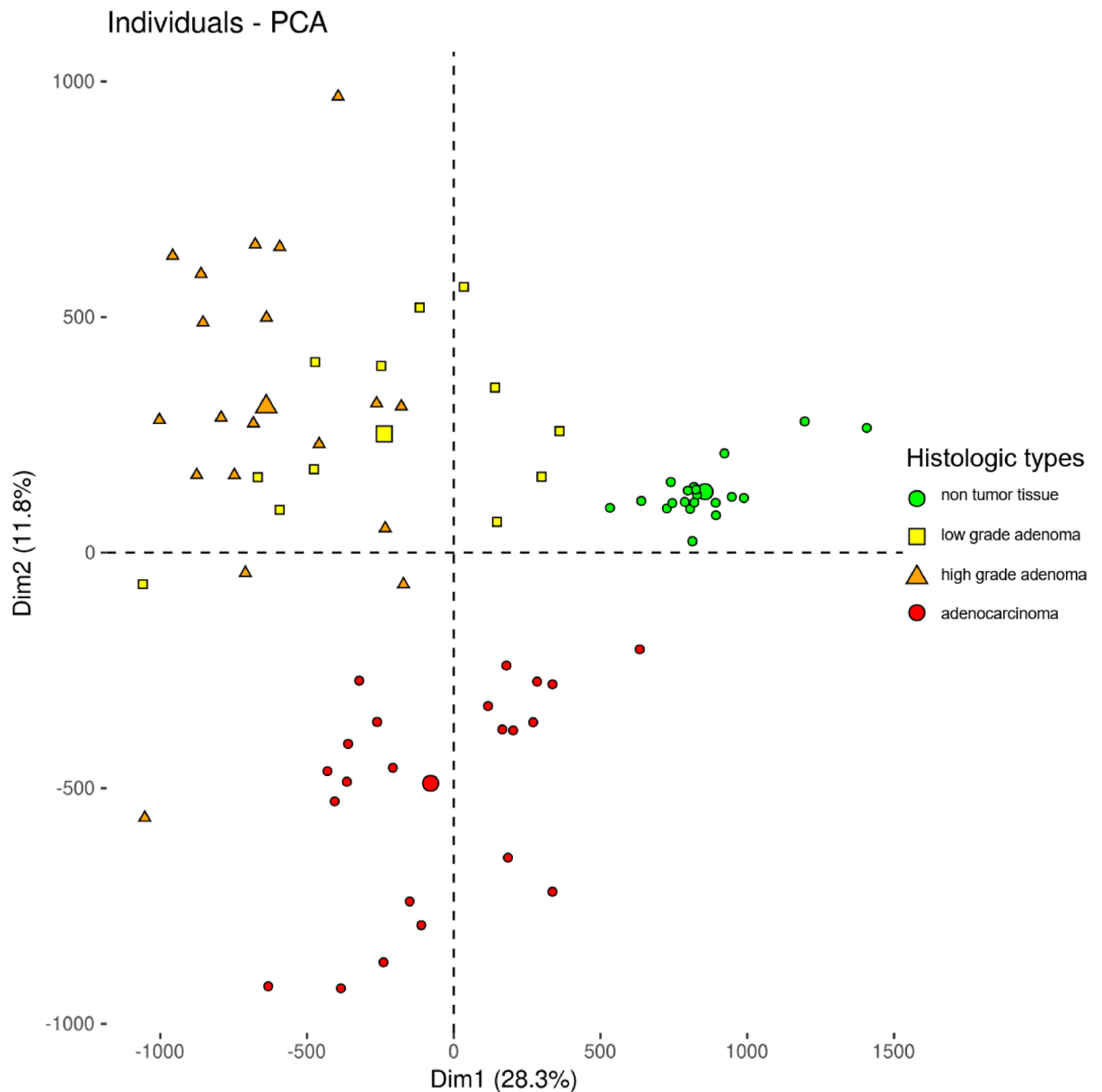


Fig. 2. PCA of the samples by methylation data.

Pathways involved

Biological processes involving genes with methylated promoters during oncogenesis were associated with nucleic acids process (2,215 out of 5,492 genes, $q < 20^{-44}$). Biological processes of genes with unmethylated promoters during oncogenesis were associated with the sensory perception of smell (154 out of 407 genes, $q < 10^{-28}$). Transient methylation during oncogenesis was associated with the mitotic cell cycle process (57 out of 777 genes, $q = 0.012$), and transient unmethylation was associated with post-transcriptional gene silencing (41 out of 615 genes, $q = 0.0015$) (Fig. 4, Additional File 1).

A: Early hypomethylated genes, during the transition from non-tumor colonic tissue to low-grade adenoma and remain hypomethylated through high-grade adenoma and adenocarcinoma. These genes are involved in biological processes related to olfaction. **B:** Early methylated genes, during the transition from non-tumor colonic tissue to low-grade adenoma and remain methylated through high-grade adenoma and adenocarcinoma. These genes are involved in biological processes related to nucleic acids process. **C:** Transiently hypomethylated genes, undergo demethylation from non-tumor colonic tissue to adenoma, then become methylated again in adenocarcinoma. These genes are involved in biological processes related to post-transcriptional regulation. **D:** Transiently methylated genes undergo demethylation from non-tumor colonic tissue to adenoma, then become methylated again in adenocarcinoma. These genes are involved in biological processes related to mitotic cell cycle.

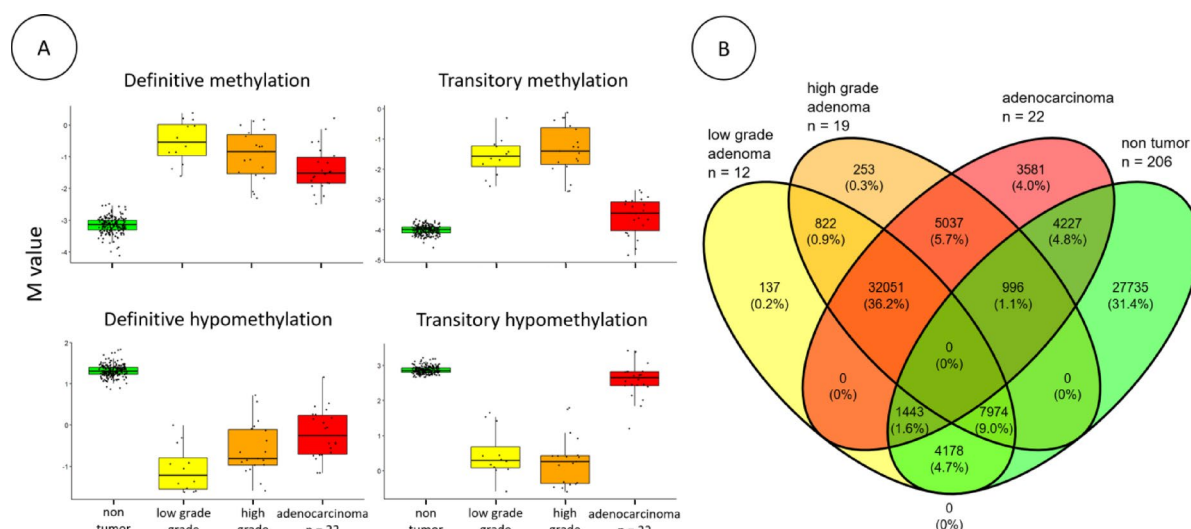


Fig. 3. Boxplot and Venn diagram of CpG methylation across the sample types.

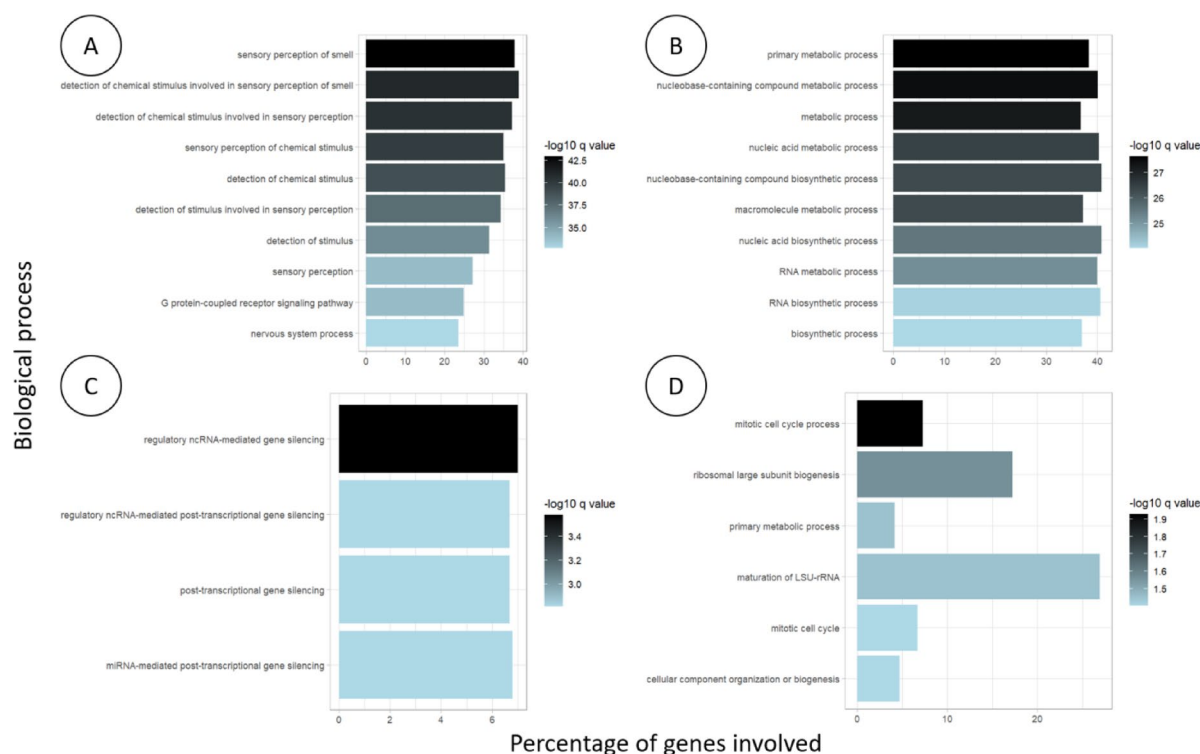


Fig. 4. Ontology of biological processes involved in early and transitory methylation changes.

Potential early colorectal DNA methylation biomarkers

Validated biomarkers are early

DNA methylation biomarkers for colorectal cancer, such as *SEPT9*, *NDRG4* and *BMP3*, which have received FDA approval for blood- or stool-based CRC screening, have been validated. Other biomarkers were tested after a review of published colorectal cancer biomarkers using DNA methylation in plasma or stool (Table 2).

An increase in methylation was observed for all known biomarkers (Fig. 5, Supplementary Table 2).

BMP3 and *KCNJ12* were not methylated in adenocarcinomas. All the other known biomarkers were methylated in adenocarcinomas and in high-grade adenomas. *FBN1* and *SNCA* were not methylated in low-grade adenomas. The *SEPT9* and *VIPR2* promoters were not methylated in low-grade adenomas but were methylated in high-grade adenomas and adenocarcinomas.

Gene	CpG	Location	Biomarker usage	Publication
ADHFE1	cg18065361	TSS200	stool	27
BCAT1	cg02765913	5UTR	plasma	28
BMP3	cg20276585	TSS200	plasma	29
C9orf50	cg18973112	TSS200	stool	30,31
CLIP4	cg09695033	TSS1500	stool	32
CNRIP1	cg11573679	1stExon	stool	33
COL25A1	cg07095995	TSS200	plasma	34
FBN1	cg15385562	TSS1500	stool	35
FNB1	cg15385562	TSS1500	stool	33
FOXF1	cg00314966	1stExon	plasma	36
GATA5	cg16714055	TSS1500	plasma	37
GRIA4	cg04747226	TSS200	stool	38
HAND1	cg03158581	TSS1500	plasma	39
IKZF1	cg23140175	TSS200	plasma	28
KCNJ12	cg27056599	TSS200	plasma	31
KCNQ5	cg09303936	TSS1500	stool	30
LIFR	cg11841722	TSS1500	plasma	40
LINC00473	cg09830769	TSS1500	plasma	41
MAL	cg04804539	TSS1500	stool	33
METAP1D	cg08750504	3UTR	plasma	34
MPPED2	cg11855526	5UTR	plasma	42
NDRG4	cg00687686	TSS1500	stool	43
NPY	cg00355281	TSS200	plasma	44
PPP2R5C	cg00723271	Body	stool	27
SDC2	cg24732574	TSS200	stool	27,43,45
SEPT9	cg17300544	TSS200	plasma	46
SHOX2	cg06759819	Body	plasma	46
SNCA	cg08767460	TSS1500	stool	33,35
SPG20	cg03966514	5UTR	stool	33
TFPI2	cg15649801	TSS1500	stool	45,47
TWIST1	cg09799658	TSS200	plasma	31
VIPR2	cg03976877	1stExon	stool	38
WIF1	cg26733786	5UTR	plasma	44
ZNF132	cg03735888	TSS200	plasma	31
ZNF304	cg21627760	TSS200	plasma	40

Table 2. Colorectal cancer biomarkers described in the literature.

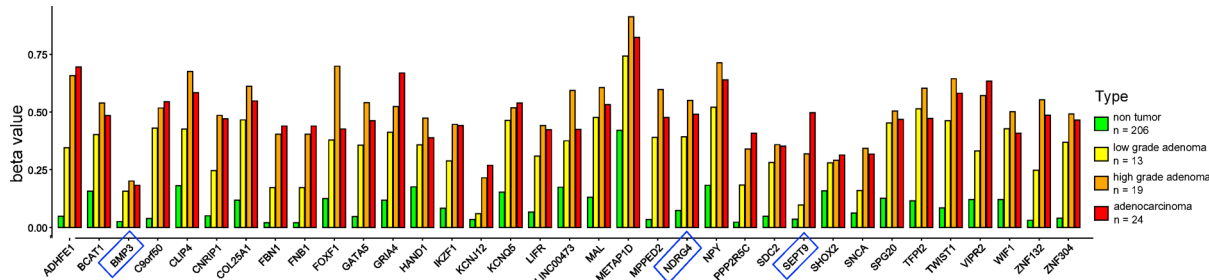


Fig. 5. Methylation of known biomarkers.

Potential and FDA-approved CRC biomarkers (marked with blue rectangles) based on DNA methylation in plasma and stools. The beta value cutoff used to determine methylation was 0.3. All of the biomarkers, except BMP3 and KCN12 are methylated during oncogenesis. ADHFE1, CNRIP1, FNB1, IKZF1, LIFR, PPP2R5C, SDC2, SEPT9, SHOX2, SNCA and ZNF132 are only partially methylated in low-grade adenomas and can be used to differentiate low-grade adenomas from high-grade adenomas or adenocarcinomas.

Comparison with an external dataset

A strong correlation was detected between the beta values of CpG sites from the two BeadChips ($n = 452,034$), with Spearman correlation coefficients consistently exceeding 0.9 ($p < 10^{-4}$), as shown in Supplementary Fig. 2. Adenoma specific methylation changes were also observed in the dataset (Supplementary Fig. 3). Despite the high level of correlation, only 33% concordance was observed between the datasets. The biological processes associated with genes displaying methylation or unmethylation between non tumor colonic tissue and adenoma, and between adenoma and adenocarcinoma were consistent across both datasets (Additional File 2).

Two CpGs that were methylated in high-grade adenomas and adenocarcinomas, as well as two CpGs that were only methylated in adenocarcinomas, were tested using methylation-specific PCR and digital droplet PCR on non-tumor, adenoma, and adenocarcinoma samples from the Tumorothèque de Franche-Comté. The results confirmed the BeadChips data ($p < 0.01$ for each, Spearman's test) (Supplementary Tables 3 and Supplementary Fig. 4).

Discussion

To our knowledge, this study is the first to investigate the evolution of the colonic methylome in the context of distinguishing between low-grade and high-grade adenomas. Notably, we also report, for the first time, the presence of transient methylation changes during this progression. The 450k data revealed a global methylation gain during oncogenesis, whereas the EPIC data revealed a global methylation loss during oncogenesis. This inconsistent observation can be explained by the fact that the 450k focuses on gene promoters and that the EPIC BeadChip targets more intergenic CpGs, which are more affected by global hypomethylation in colorectal oncogenesis. The global methylation changes observed with the EPIC BeadChip are similar to those observed with immunohistochemistry⁴⁸, especially concerning intergenic sequences such as long interspersed elements (LINEs).

We observed that the DNA methylation changes observed during colorectal oncogenesis were DNA evenly distributed between methylation and hypomethylation. The hypomethylation mostly occurs in CpGs in CpG shelves and open seas which has been previously described and associated with chromosomal instability in colorectal cancer⁴⁹. Global DNA hypomethylation accounted for approximately 4% of the methylome, which is lower than the 8–10% reported via liquid chromatography⁵⁰. These inconsistencies may be explained by the q value threshold used to determine the DMC ($q < 10^{-4}$) and by the fact that liquid chromatography is a quantitative measure that does not require the CpG to segregate in DMCs.

During oncogenesis, we noticed that the majority of DNA methylation modifications occurred during the transition between normal colon tissue and low-grade adenoma. This observation is inconsistent with the results of Andrew D Beggs et al.¹¹, who used Illumina 27k BeadChips and reported that the DNA methylation pattern was acquired during the transition from adenoma to adenocarcinoma. However, our observation is consistent with the results of Yanxin Luo et al.¹⁰, who used the BeadChips Illumina 450k and reported that most of the methylation changes occur during the transition from normal colon mucosa to adenoma.

We found that global methylation clustering did not show a continuum from normal colon mucosa to adenocarcinoma but that the adenoma had its own methylation pattern. This effect can be observed in the results of Janssens et al.¹⁴. The magnitude of global methylation changes is the same, with approximately one-third of methylation and two-thirds hypomethylation occurring during the transition from non-tumor colon tissue to adenoma¹⁰.

Gene promoters with the earliest DMCs during oncogenesis were associated with tissular disorganization. The associated pathways are similar to those revealed by Lu YW et al.¹², who used SeqCap-targeted bisulfite sequencing, and involve neuronal activity. This observation could be explained by the presence of mesenteric neurons in normal colon tissue, and an interesting phenomenon of tumor-neuron crosstalk could also explain these observations⁵¹. However, these explanations are not exclusive and need to be verified with additional studies.

The adenoma methylome confirmed the evolution of the methylation profile of 12 out of 15 described CRC biomarkers, and 10 out of 15 of these biomarkers occur early in oncogenesis, with differential methylation in low-grade adenomas, providing promising opportunities for DNA methylation CRC screening, diagnosis and follow-up. Biomarkers with methylation in high-grade adenomas may be relevant for the noninvasive assessment of tumor stage.

No batch effect was found between the two adenocarcinoma datasets (HCIM and GSE149282), but we cannot exclude a batch effect between the other datasets (Besançon adenoma and GSE132804). To limit potential batch effects, the methylation data were normalized. There are several ways to apply normalization to methylation BeadChip data: quantile normalization, Genome Studio normalization, SWAN normalization, functional normalization, and background and dye bias correction. Functional normalization combined with background and dye bias correction was chosen because it is the best fit when global changes are expected, such as in tumor versus non-tumor tissue comparisons, as in the case of this study. This method was developed for the 450k BeadChip and then adapted to the EPIC chip⁵².

Although computational normalization was performed uniformly to limit the batch effect, technical variability, such as differences in DNA extraction, bisulfite treatment, and array hybridization⁵³, as well as biological variability such as genetic ancestry and sociocultural environment, may influence methylation profiles^{54,55}. These factors should be considered when interpreting the results.

Future work may distinguish between different conventional colorectal adenoma subtypes, such as tubular, villous and tubulovillous adenomas based on methylation data. Additionally, exploring colorectal serrated lesions could provide new insight into the evolution of the colorectal methylome during oncogenesis. Our present study contains only a few numbers of samples, preventing a histological analysis. Samples with *MLH1*

promoter methylation were excluded from the analysis, but those with MSI due to a mechanism other than *MLH1* promoter methylation could not be excluded.

The DMCs were not filtered by the absolute methylation difference due to the cellularity of adenoma samples; therefore, the DMCs were not expected to be quantitatively important, and an absolute methylation difference threshold would not be appropriate for the data. To limit the number of artifactual DMCs, a stringent FDR p value of 10^{-4} was applied.

PCA did not identify subgroups of adenomas, but the small sample size is likely to have missed potential rare molecular subtypes of adenomas. The transcriptomic impact of methylation alterations was not investigated. The study of methylation coupled with transcriptomics is challenging due to the small amount of colorectal adenoma material and the high input required by these techniques.

Conclusion

This study is the first to explore the dynamics of DNA methylation during the low-grade and high-grade adenoma stages of colorectal cancer oncogenesis and to suggest the potential reversibility of epigenetic alterations in this context. We examined the behavior of established DNA methylation biomarkers for colorectal cancer, providing insights into their potential utility for indirect tumor staging. Further research is required to validate these findings and assess their translational relevance.

Data availability

The methylome beta values and idat files of the adenomas from Besançon are available at the GEO accession number GSE288652. Other datasets generated and analyzed during this study are available from the corresponding author upon reasonable request.

Received: 20 June 2025; Accepted: 11 November 2025

Published online: 25 November 2025

References

- Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 (2024).
- Fearon, E. R. & Vogelstein, B. A genetic model for colorectal tumorigenesis. *Cell* **61**, 759–767 (1990).
- Nagtegaal, I. D. et al. The 2019 WHO classification of tumours of the digestive system. *Histopathology* **76**, 182–188 (2020).
- Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
- Zhu, T., Tong, H., Du, Z., Beck, S. & Teschendorff, A. E. An improved epigenetic counter to track mitotic age in normal and precancerous tissues. *Nat. Commun.* **15**, 4211 (2024).
- Galamb, O. et al. Aging related methylation influences the gene expression of key control genes in colorectal cancer and adenoma. *World J. Gastroenterol.* **22**, 10325–10340 (2016).
- Luebeck, G. E. et al. Implications of epigenetic drift in colorectal neoplasia. *Cancer Res.* **79**, 495–504 (2019).
- Murcia, O. et al. Effects of somatic methylation in colonic polyps on risk of developing metachronous advanced colorectal lesions. *Cancers (Basel)*. **13**, 246 (2021).
- Kaz, A. M. et al. Patterns of DNA methylation in the normal colon vary by anatomical location, gender, and age. *Epigenetics* **9**, 492–502 (2014).
- Luo, Y. et al. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. *Gastroenterology* **147**, 418–429e8 (2014).
- Beggs, A. D. et al. Whole-genome methylation analysis of benign and malignant colorectal tumours. *J. Pathol.* **229**, 697–704 (2013).
- Lu, Y. W. et al. Early results of the integrative epigenomic-transcriptomic landscape of colorectal adenoma and cancer. *World J. Gastrointest. Oncol.* **16**, 414–435 (2024).
- Druliner, B. R. et al. Molecular characterization of colorectal adenomas with and without malignancy reveals distinguishing genome, transcriptome and methylome alterations. *Sci. Rep.* **8**, 3161 (2018).
- Janssens, K. et al. Epigenome-wide methylation analysis of colorectal carcinoma, adenoma and normal tissue reveals novel biomarkers addressing unmet clinical needs. *Clin. Epigenetics*. **15**, 111 (2023).
- Wang, T. et al. Dysfunctional epigenetic aging of the normal colon and colorectal cancer risk. *Clin. Epigenetics*. **12**, 5 (2020).
- Lendvai, Á. et al. Genome-wide methylation profiling identifies hypermethylated biomarkers in high-grade cervical intraepithelial neoplasia. *Epigenetics* **7**, 1268–1278 (2012).
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2020). <https://www.R-project.org/>. <https://www.r-project.org/>
- Wickham, H. et al. Welcome to the tidyverse. *J. Open. Source Softw.* **4**, 1686 (2019).
- Aryee, M. J. et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics* **30**, 1363–1369 (2014).
- Phipson, B., Maksimovic, J. & Oshlack, A. MissMethyl: an R package for analyzing data from Illumina's HumanMethylation450 platform. *Bioinformatics* **32**, 286–288 (2016).
- Ritchie, M. E. et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).
- Qin, Y. et al. An R package for estimating and accounting for tumor purity in cancer methylation research. *Genes Dis.* **5**, 43 (2018).
- Fortin, J. P. et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol.* **15**, 503 (2014).
- Du, P. et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **11**, 587 (2010).
- Lê, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
- Harris, M. A. et al. The gene ontology (GO) database and informatics resource. *Nucleic Acids Res.* **32**, D258–261 (2004).
- Li, B., Liu, S., Gao, Y., Zheng, L. & Lu, Y. Combined detection of SDC2/ADHFE1/PPP2R5C methylation in stool DNA for colorectal cancer screening. *J. Cancer Res. Clin. Oncol.* **149**, 10241–10253 (2023).
- Pedersen, S. K. et al. A two-gene blood test for methylated DNA sensitive for colorectal cancer. *PLoS One*. **10**, e0125041 (2015).
- Lima, A. B. et al. Combined SEPT9 and BMP3 methylation in plasma for colorectal cancer early detection and screening in a Brazilian population. *Cancer Med.* **12**, 15854–15867 (2023).
- Cao, Y. et al. KCNQ5 and C9orf50 methylation in stool DNA for early detection of colorectal cancer. *Front. Oncol.* **10**, 621295 (2020).

31. Zhang, Y. et al. Sensitive detection of colorectal cancer in peripheral blood by a novel methylation assay. *Clin. Epigenetics*. **13**, 90 (2021).
32. Cao, Y. et al. Feasibility of methylated CLIP4 in stool for early detection of colorectal cancer: A training study in Chinese population. *Front. Oncol.* **11**, 647066 (2021).
33. Lind, G. E. et al. Identification of an epigenetic biomarker panel with high sensitivity and specificity for colorectal cancer and adenomas. *Mol. Cancer*. **10**, 85 (2011).
34. Overs, A. et al. COL25A1 and METAP1D DNA methylation are promising liquid biopsy epigenetic biomarkers of colorectal cancer using digital PCR. *Clin. Epigenetics*. **16**, 146 (2024).
35. Li, W. et al. Detection of SNCA and FBN1 Methylation in the Stool as a Biomarker for Colorectal Cancer. *Disease Markers* 657570 (2015). (2015).
36. Caputo, V. et al. Diagnostic value of liquid biopsy in the era of precision medicine: 10 years of clinical evidence in cancer. *Explor. Target. Antitumor Ther.* **4**, 102–138 (2023).
37. Zhang, X. et al. Combined detection of plasma GATA5 and SFRP2 methylation is a valid noninvasive biomarker for colorectal cancer and adenomas. *World J. Gastroenterol.* **21**, 2629–2637 (2015).
38. Vega-Benedetti, A. F. et al. Colorectal cancer early detection in stool samples tracing CpG Islands methylation alterations affecting gene expression. *Int. J. Mol. Sci.* **21**, 4494 (2020).
39. Shavali, M., Moradi, A., Tahmaseb, M., Mohammadian, K. & Ganji, S. M. Circulating-tumour DNA methylation of HAND1 gene: a promising biomarker in early detection of colorectal cancer. *BMC Med. Genomics*. **17**, 117 (2024).
40. Li, D. et al. Discovery and validation of tissue-specific DNA methylation as noninvasive diagnostic markers for colorectal cancer. *Clin. Epigenetics*. **14**, 102 (2022).
41. Ruiz-Bañobre, J. et al. Noninvasive early detection of colorectal cancer by hypermethylation of the LINC00473 promoter in plasma cell-free DNA. *Clin. Epigenetics*. **14**, 86 (2022).
42. Gu, S. et al. Genome-wide methylation profiling identified novel differentially hypermethylated biomarker MPPED2 in colorectal cancer. *Clin. Epigenetics*. **11**, 41 (2019).
43. Long, L. et al. Significance of SDC2 and NDRG4 methylation in stool for colorectal cancer diagnosis. *Clin. Biochem.* **124**, 110717 (2024).
44. Roperch, J. P. et al. Aberrant methylation of NPY, PENK, and WIF1 as a promising marker for blood-based diagnosis of colorectal cancer. *BMC Cancer*. **13**, 566 (2013).
45. Zhang, W. et al. SDC2 and TFPI2 methylation in stool samples as an integrated biomarker for early detection of colorectal cancer. *Cancer Manag Res.* **13**, 3601–3617 (2021).
46. Semaan, A. et al. SEPT9 and SHOX2 DNA methylation status and its utility in the diagnosis of colonic adenomas and colorectal adenocarcinomas. *Clin. Epigenetics*. **8**, 100 (2016).
47. Glöckner, S. C. et al. Methylation of TFPI2 in stool DNA: A potential novel biomarker for the detection of colorectal cancer. *Cancer Res.* **69**, 4691–4699 (2009).
48. Hernandez-Blazquez, F. et al. Evaluation of global DNA hypomethylation in human colon cancer tissues by immunohistochemistry and image analysis. *Gut* **47**, 689–693 (2000).
49. Rodriguez, J. et al. Chromosomal instability correlates with genome-wide DNA demethylation in human primary colorectal cancers. *Cancer Res.* **66**, 8462–9468 (2006).
50. Feinberg, A. P., Gehrke, C. W., Kuo, K. C. & Ehrlich, M. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res.* **48**, 1159–1161 (1988).
51. Huang, M. et al. Crosstalk between cancer cells and the nervous system. *Med. Adv.* **1**, 173–189 (2023).
52. Fortin, J. P., Triche, T. J. Jr & Hansen, K. D. Preprocessing, normalization and integration of the illumina humanmethylationepic array with Minfi. *Bioinformatics* **33**, 558–560 (2017).
53. Ross, J. P. et al. Batch-effect detection, correction and characterisation in illumina HumanMethylation450 and methylationepic BeadChip array data. *Clin. Epigenetics*. **14**, 58 (2022).
54. Galanter, J. M. et al. Differential methylation between ethnic sub-groups reflects the effect of genetic ancestry and environmental exposures. *Elife* **6**, e20532 (2017).
55. Chan, M. H., Merrill, S. M., Konwar, C. & Kobor, M. S. An integrative framework and recommendations for the study of DNA methylation in the context of race and ethnicity. *Discov Soc. Sci. Health.* **3**, 9 (2023).

Acknowledgements

We thank all patients who agreed to participate in this study, the Canceropole Grand-Est with the “projet émergence” and the Région Bourgogne franche-Comté. We thank Pr Aurelien de Reynies for his advice on bioinformatics approaches involving the methylome.

Author contributions

A.O. performed the bioinformatics analysis, made the figures, analyzed the results and wrote the main manuscript text. C.M., F.B., F.M. and L.A. selected the adenoma FFPE samples, asserted the histologic type and the tumor cellularity. C.C., E.H. and P.P. performed the DNA for the methylation BeadChip and MSP. Z.S., J.D., J.P.F., M.G., and C.B. contributed to the results analysis, and the writing of the manuscript. A.O., Z.S., P.P. and J.P.F. also contributed to the overall supervision of the study. All authors reviewed the manuscript.

Funding

This work was partly funded by the Canceropole Grand-Est.

Declarations

Competing interests

The authors declare no competing interests.

Ethics approval and consent to participate

All procedures performed in studies involving human participants were in accordance with the ethical standards of the national research committee and with the 1964 Helsinki Declaration and its later amendments. In France, this search is considered as a non-interventional study according to European legislation. All patients were individually informed that their data should be used to scientific research. All

experimental protocols were approved by the scientific board of the regional biobank of Franche-Comté, France (registration number BB-0033-00024, Tumorothèque Régionale de Franche-Comté), which ensures patients informed consent.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-28656-5>.

Correspondence and requests for materials should be addressed to A.O.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025