

Genome-Wide Analysis of DNA Methylation in Gastrointestinal Cancer

Kiichi Sugimoto^{1,2}, Hirotaka Momose², Tomoaki Ito^{1,3}, Hajime Orita³, Koichi Sato³, Kazuhiro Sakamoto², Malcolm V. Brock¹

¹ Department of Surgery, The Sidney Kimmel Comprehensive Cancer Center, Johns Hopkins University School of Medicine ² Department of Coloproctological Surgery, Juntendo University Faculty of Medicine ³ Department of Surgery, Juntendo University Shizuoka Hospital

Corresponding Author

Kiichi Sugimoto

ksugimo@juntendo.ac.jp

Citation

Sugimoto, K., Momose, H., Ito, T., Orita, H., Sato, K., Sakamoto, K., Brock, M.V. Genome-Wide Analysis of DNA Methylation in Gastrointestinal Cancer. *J. Vis. Exp.* (163), e61355, doi:10.3791/61355 (2020).

Date Published

September 18, 2020

DOI

10.3791/61355

URL

jove.com/video/61355

Introduction

Epigenetics refers to heritable changes in gene function without alteration of the sequence of DNA¹. Such changes may be due to DNA methylation, in which methyl groups on a DNA base may alter the gene expression through changes in chromatin packing. Cancer development and progression may occur if this effect results in altered expression of tumor suppressor genes². Aging and chronic inflammation are both causes of cancer and the main reasons for changes in DNA methylation in humans^{3, 4, 5}. Consequently, this allows

the utilization of DNA methylation as a biomarker in cancer diagnosis, and as a target for treatment and prevention. For early detection and cancer prognosis, DNA methylation are being measured in tumor, blood, urine, and stool samples⁶, while demethylating agents are now being used to treat leukemias such as myelodysplastic syndrome⁷.

Genome-wide DNA methylation analysis using an array platform for complex evaluation of DNA methylation at an individual CpG locus in the human genome can be utilized

Abstract

DNA methylation is an important epigenetic change that is biologically meaningful and a frequent focus of cancer research. Genome-wide DNA methylation is a useful measure to provide an accurate analysis of the methylation status of gastrointestinal (GI) malignancies. Given the multiple potential translational uses of DNA methylation analysis, practicing clinicians and others new to DNA methylation studies need to be able to understand step by step how these genome-wide analyses are performed. The goal of this protocol is to provide a detailed description of how this method is used for the biomarker identification in GI malignancies. Importantly, we describe three critical steps that are needed to obtain accurate results during genome-wide analysis. Clearly and concisely written, these three methods are often lacking and not noticeable to those new to epigenetic studies. We used 48 samples of a GI malignancy (gastric cancer) to highlight practically how genome-wide DNA methylation analysis can be performed for GI malignancies.

to examine the methylation status of more than 450,000 CpG sites in genomic DNA^{8,9}, which permits exploration of cancer epigenetics (see **Table of Materials**). Whole genome bisulfite sequencing (WGBS) technologies have changed our approaches in the field of epigenetics^{10,11}. However, there are some disadvantages to the technologies in terms of a substantial cost and processing time for epigenetic analysis of a large number of samples^{10,11}. Therefore, the array platform is more feasible for complex evaluation of DNA methylation in the human genome. The availability of approaches for genome-wide methylation analyses has improved in the past few years and allows us to expand our knowledge of how DNA methylation contributes to cancer development and progression¹². Recent progresses in microarray-platform approaches provide us the rationale for genome-wide methylation analysis to identify a novel epigenetic signature in gastrointestinal cancers¹³. The goal of this protocol is to provide a detailed description of how this method is used for biomarker identification in GI malignancies.

Protocol

All procedures followed were in accordance with the ethical standards of the institutions' human research ethics committee. The study was approved by the Institutional Review Board at the Juntendo University Shizuoka Hospital, and written informed consent was waived because of the retrospective design.

1. Washing the slides

1. Prepare 10 μm of unstained formalin-fixed paraffin-embedded (FFPE) sections.

2. Place slides in a glass slide holder: use about 3–5 largest cross-sectional slides unless the tissue is minimal, and more slides are needed.
3. Fill the slide holder with xylene, and make sure that all tissue on the slide is submerged. Allow to sit for 15 min.
4. After 15 min, pour out the xylene with the slides held with a pipette tip so that the slides do not fall out.
5. Pour in more xylene to the same level as before. Allow to sit for another 15 min.
6. After 15 min, pour out the xylene again.
7. Fill the slide holder with 100% ethanol (EtOH), and make sure that all tissue on the slide is fully submerged. Allow to sit for 3 min.
8. Pour off the EtOH while carefully holding the slides. Refill to the same level with more EtOH. Allow to sit for 2 min.
9. Pour off the EtOH again and remove the slide. Carefully place them face up on a clean paper towel to dry. Allow to sit for 10 min.

2. Scraping the slides

1. Prepare digestion/lysis buffer with 650 mL of diethylpyrocarbonate (DEPC)-treated water, 100 mL of Ethylenediaminetetraacetic acid (EDTA), 50 ml of Tris hydrochloride (Tris-HCL) 2 M pH 8.8, and 200 mL of 10% sodium dodecyl sulfate (SDS) (see **Table of Materials**).
2. Fill a 1.5 mL single-use polypropylene tube with 80 μL of digestion/lysis buffer (see **Table of Materials**).
3. Put a clean pipette tip into the buffer.
4. Identify cancer tissues of the tumor area most suitable for macrodissection according to the appropriate H&E stained section.

NOTE: The tumor area for macrodissection should be identified by preferably two qualified pathologists.

5. Macrodissect cancer tissue based on the marked H&E stained section.

1. Take a clean razor blade and gently scrape the cancer tissue off the slide, trying to keep it in one piece so that it is easier to work with.

2. Take the pipette tip out of single-use polypropylene tube containing buffer and use it to transfer the scrapped tissue into the buffer vial (see **Table of Materials**).

NOTE: The wet tip should attract the tissue, which makes the transfer easier.

6. Repeat step 2.5 with the rest of the slides.

7. After all the tissue is in single-use polypropylene tube, use the tip to make sure the tissue is fully submerged and is not stuck to the wall of the tube.

8. Add 20 μL of a subtilisin-related serine protease to the vial and gently flick to mix (see **Table of Materials**).

9. Place the vial in a 55 $^{\circ}\text{C}$ heat block for at least 4 h or overnight. Make sure to slightly vortex after 2 h.

3. Bisulfite treatment

1. Use 45 μL of the digested tissue solution as the sample.

2. Perform the bisulfite treatment using reagents in a bisulfite conversion kit according to the manufacturer's instructions (see **Table of Materials**).

1. Add 5 μL of dilution buffer to the DNA sample and incubate at 37 $^{\circ}\text{C}$ for 15 min (see **Table of Materials**).

2. While the samples are incubating, prepare the bisulfite conversion reagent by adding 750 μL of dH_2O and 210 μL of dilution buffer to one tube of CT

conversion reagent (see **Table of Materials**). Mix the tubes by vortexing for 10 min.

3. After the 15 min incubation, add 100 μL of the prepared CT conversion reagent to each sample and mix by inversion.

4. Incubate the samples in the dark at 50 $^{\circ}\text{C}$ for 12 to 16 h.

5. After the incubation, remove the samples and place on ice for 10 min.

6. Add 400 μL of binding buffer and mix each sample by pipetting up and down (see **Table of Materials**). Load each sample into a spin column and place the column into a 2 mL collection tube (see **Table of Materials**).

7. Centrifuge each sample at full speed (10,000 $\times g$) for 1 min and discard the flow-through.

8. Add 200 μL of wash buffer to each column and spin at full speed for 1 min, discarding the flow-through (see **Table of Materials**).

9. Add 200 μL of desulfonation buffer to each column and allow the column to stand at room temperature for 15 min (see **Table of Materials**). After the incubation, spin the columns at full speed for 1 min and discard the flow-through.

10. Add 200 μL of wash buffer to each column and spin at full speed for 1 min (see **Table of Materials**).

11. Repeat this step one more time.

12. Add 46 μL of dH_2O to each column and place each column in a new sterile 1.5 mL single-use polypropylene tube (see **Table of Materials**). Spin each tube for 2 min to elute the DNA.

- Remove each spin column from the single-use polypropylene tube and discard (see **Table of Materials**).

The DNA is now ready for the analysis.

4. Array platform for evaluation of DNA methylation at a CpG locus in the human genome

- Assess the quality of DNA (quality check: QC) using FFPE QC assay on a real-time PCR amplification and detection instrument, with subsequent data analysis performed according to the manufacturers' instructions (see **Table of Materials**).

NOTE: Samples with ΔCq values less than the recommended 5.0 are further processed¹⁴.

- Analyze samples with an array platform for complex evaluation of DNA methylation to assess the methylation status of >450,000 CpG sites in the genome (see **Table of Materials**). Product information sheet, data sheet, and product files for the array platform are available¹⁵.

NOTE: The assay is performed according to the manufacturer's instructions for FFPE material¹⁴.

- Calculate high and low locus methylation in a data analysis tool for an array platform as a β -value, with a range from 0 to 1, respectively (see **Table of Materials**). Use commercially available software (see **Table of Materials**) to calculate a β -value.

- Import data generated on the array platform for complex evaluation of DNA methylation platform into the R software environment (R v.2.15.1) and process them using a tool for analyzing methylation arrays^{16, 17}.

NOTE: On the heatmap, which is generated using a data analysis tool, columns are ordered by unsupervised clustering, whereas the order of rows is based on the

decreasing significance of the t statistic for differential methylation from top to bottom.

- Divide the heatmap into high and low methylation groups using the first differentiator in the unsupervised clustering.

NOTE: For validation with quantitative methylation-specific PCR (qMSP), choose genes based on a larger β -value in relation to CpG islands in the promoter region, and due to their suitability for primer and probe design for qMSP.

5. Quantitative Methylation-Specific PCR (qMSP)

- Use the bisulfite-modified DNA from step 3.3 as a template for fluorescence-based real-time PCR in qMSP to evaluate methylation of the promoter region in each gene analysis.

- Perform qMSP using a 96-well Real-Time PCR instrument (see **Table of Materials**).

- Check for the promoter methylation status of the target gene on the bisulfite-modified DNA using 200 nM forward primer, 200 nM reverse primer, and 80 nM probes. Prepare the master mix with 16.6 mM $(NH_4)_2SO_4$, 67 mM Tris pH 8.8, 10 mM β -mercaptoethanol, 10 nM fluorescein, 0.166 mM of each deoxynucleotide triphosphate, and 0.04 U/ μ l of DNA polymerase (see **Table of Materials**). The final reaction volume in each well for all assays is 25 μ L.

- Perform cycling of qMSP as follows: 95 °C for 5 mins, followed by 55 cycles of 95 °C for 15 s, 60 °C for 1 min, and 72 °C for 1 min.

NOTE: Target gene should be chosen based on the criteria of having larger beta values, being related to CpG islands in the promoter region, and being suitable for primer and probe design for qMSP.

3. Use Human genomic treated with CpG Methylase (M.SssI) as a positive methylation control (see **Table of Materials**).

NOTE: The final quantification of methylation is defined as the relative methylation value (RMV), and calculated as $2^{-\Delta\Delta Ct}$ for each methylation detection replicate compared to the mean Ct for β -Actin (ACTB)¹⁸. Primer and probe sequences are shown in **Table 1**. A Ct of 100 is used for undetected replicates, which gives a value of $2^{-\Delta\Delta Ct}$ close to zero. The following formula is used: $\text{mean } 2^{-\Delta\Delta Ct} \text{ (RMV)} = (2^{-\Delta\Delta Ct_{\text{replicate}_1}} + 2^{-\Delta\Delta Ct_{\text{replicate}_2}} + 2^{-\Delta\Delta Ct_{\text{replicate}_3}}) / 3$ ¹⁸.

Representative Results

The characteristics of 48 patients with gastric cancer in the training cohort are as follows (**Table 2**): the median age of patients was 74 years (52–89 years), and the cohort included 38 males (79.2%), and 10 females (20.8%). There were 35 patients (72.9%) with primary gastric cancer and in 13 patients (27.1%) with remnant gastric cancer (primary gastric cancer: first occurrence of a non-metastatic malignancy in the stomach; remnant gastric cancer: cancer in the remnant stomach that developed more than 5 years after distal gastrectomy, regardless of the reason for the original resection¹⁹). There were 23 patients (47.9%) with lymph node metastasis and 25 patients (52.1%) without. First, all 48 samples (the training cohort) were loaded for identification of outliers (**Figure 1A**). Two samples gave peaks that were greater than two standard deviations

displaced from the others, and these samples were removed (**Figure 1B**). Therefore, 46 samples were clustered by DNA promoter hypermethylation. The resultant heatmap was divided into two groups based on high and low methylation (**Figure 2**). This heatmap allows visualization of the top 50 probes within 1,500 bp of the transcriptional start site (TSS) in the differential methylation analysis. The high and low methylation groups differed in clinicopathological factors related to an aggressive malignant phenotype. That is, the type of cancer (primary gastric cancer: PGC) ($p = 0.01$, odds ratio = 9.09 (1.67–50.00)) and presence of lymph node metastasis (positive) ($p = 0.03$, odds ratio = 6.82 (1.16–40.08)) emerged as significant independent predictive factors when the clinicopathological factors were used as covariates in multivariate analysis (**Table 3**). Finally, we identified the EPB41L3 gene^{20, 21} (primer and probe sequences shown in **Table 1**) to be strongly associated with codifying the training cohort into high and low methylation groups in the microarray analysis. Using qMSP, the results of the microarray analysis for EPB41L3 in the test cohort (126 samples) were validated. The characteristics of the patients in the test cohort are shown in **Table 4**. RMVs of EPB41L3 in PGC tissues were significantly higher than those in remnant gastric cancer (RGC) in univariate analysis ($p = 0.01$) (**Figure 3A**). Similarly, RMVs in samples with lymph node metastasis were significantly higher than those without lymph node metastasis ($p = 0.03$) (**Figure 3B**). In this way, DNA methylation genome-wide analysis can help us to identify specific genes to characterize certain clinical status in patients with GI malignancies.

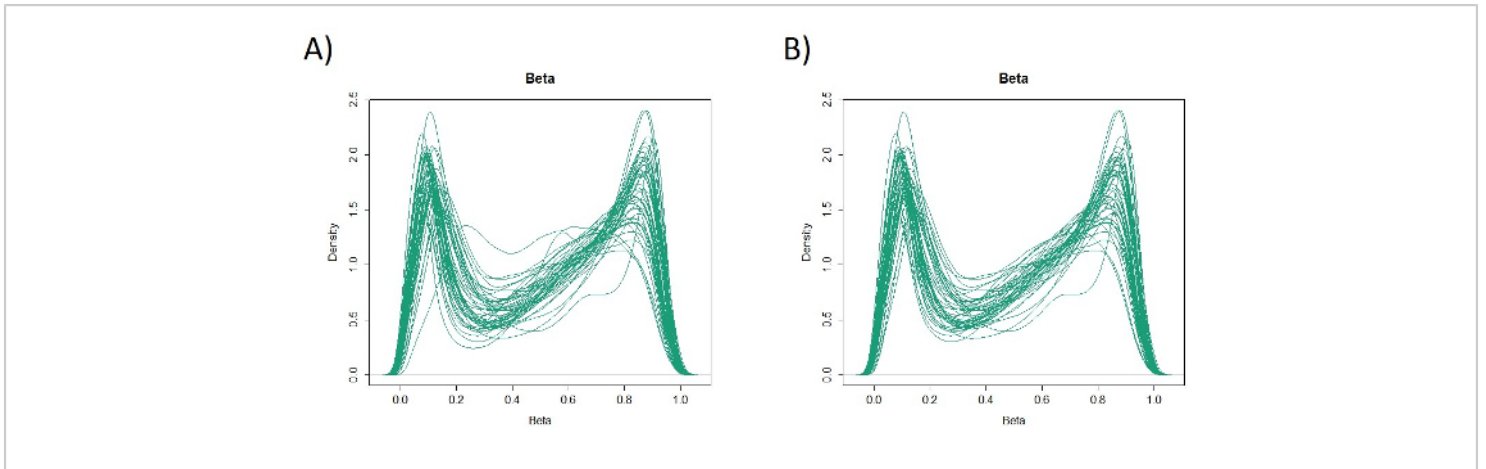


Figure 1: Beta values in 48 samples (training cohort). All 48 samples (training cohort) were loaded and outliers were examined (A). Two samples had peaks that were outliers, and these were removed (B). [Please click here to view a larger version of this figure.](#)

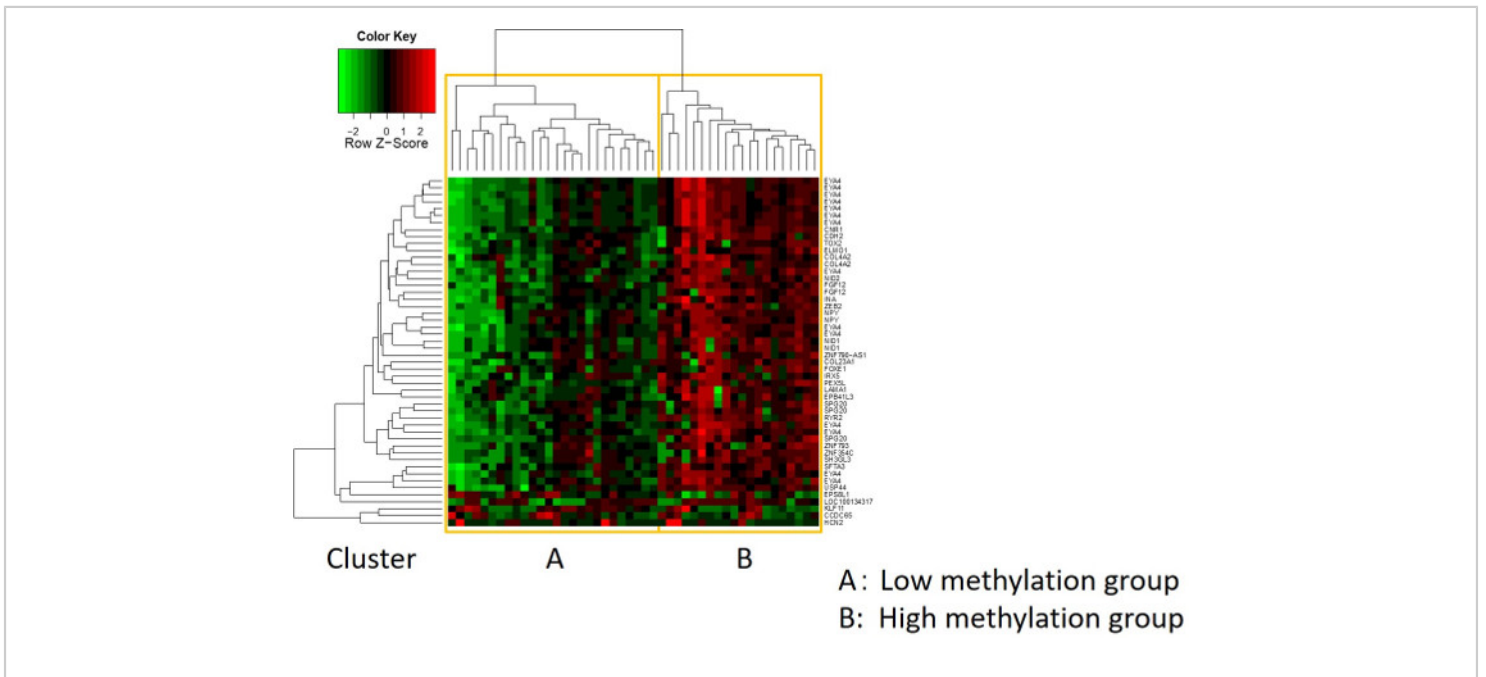


Figure 2: The resultant heatmap. The remaining 46 samples were clustered by DNA promoter hypermethylation. The heatmap was divided into high and low methylation groups. This heatmap allows visualization of the top 50 probes within 1,500 bp of the transcriptional start site (TSS) in the differential methylation analysis. [Please click here to view a larger version of this figure.](#)

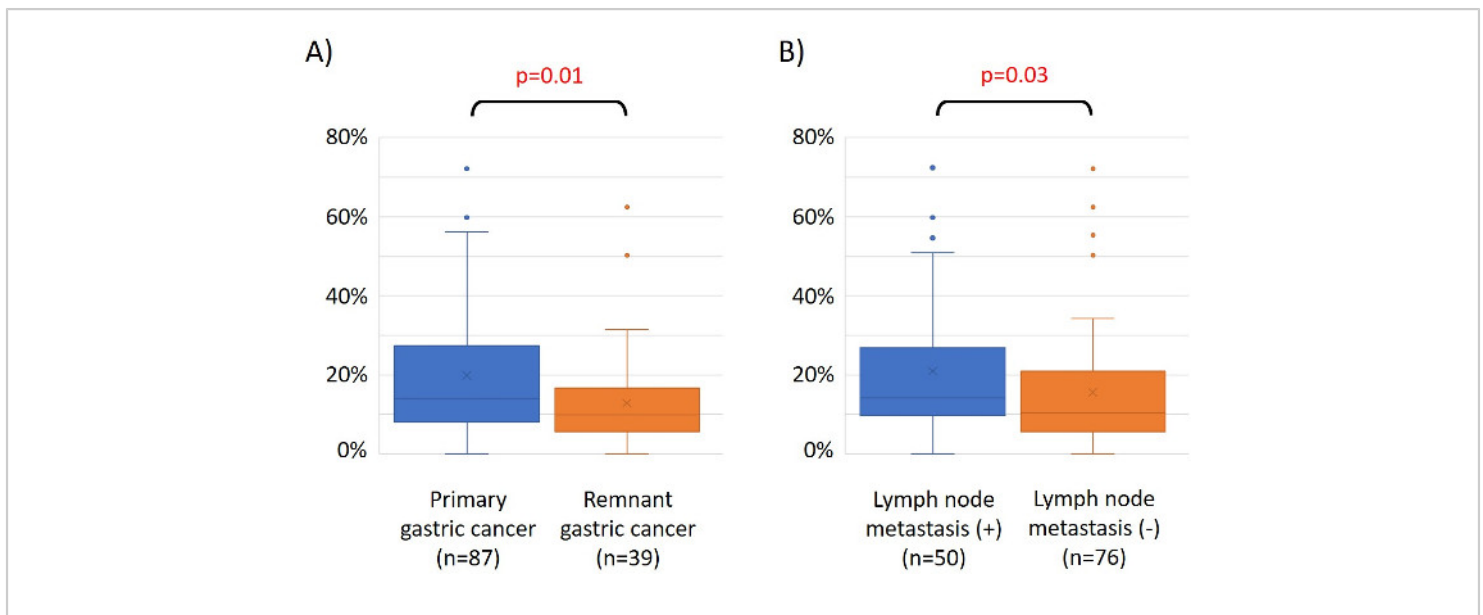


Figure 3: Relative methylation values (RMVs) for EPB41L3 in primary gastric cancer (PGC) vs. remnant gastric cancer (RGC), and in cases with and without lymph node metastasis. The results of microarray analysis for EPB41L3 in the test cohort (126 samples) were validated using qMSP. **(A)** In univariate analysis, RMVs of EPB41L3 in PGC tissues were significantly higher than those in RGC ($p = 0.01$). **(B)** Similarly, RMVs in samples with lymph node metastasis were significantly higher than in those without lymph node metastasis ($p = 0.03$). [Please click here to view a larger version of this figure.](#)

Gene	Forward 5' - 3'	Reverse 5' - 3'	Probe
B-ACTIN	TAG GGA GTA TAT AGG TTG GGG AAG TT	AAC ACA CAA TAA CAA ACA CAA ATT CAC	\56-FAM\ TGT GGG GTG \ZEN\ GTG ATG GAG GAG GTT TAG \3IABkFQ\
EPB41L3	GGG ATA GTG GGG TTG ACG C	ATA AAA ATC CCG ACG AAC GA	AAA TTC GAA AAA CCG CGC GAC GCC GAA ACC A

Table 1: Primer and probe sequences.

Clinicopathological factors	Variables	
Age		74 (52 - 89) *
Gender	Male / Female	38 (79.2%) / 10 (20.8%)
Type	PGC / RGC	35 (72.9%) / 13 (27.1%)
Lymph node metastasis	(+) / (-)	23 (47.9%) / 25 (52.1%)
PGC: Primary gastric cancer, RGC: Remnant gastric cancer		
* Median (minimum-maximum)		

Table 2: The characteristics in 48 patients with gastric cancer in the training cohort.

	P-value	Odds ratio	95% Confidence interval
Type (PGC)	0.01	9.09	1.67 – 50.00
Lymph node metastasis (+)	0.03	6.82	1.16 – 40.08
PGC: Primary gastric cancer			

Table 3: Predictive factors for the high methylation group (Cluster B).

Clinicopathological factors	Variables	
Age		71 (33 - 86) *
Gender	Male / Female	96 (76.2%) / 30 (23.8%)
Type	PGC / RGC	87 (69.0%) / 39 (31.0%)
Lymph node metastasis	(+) / (-)	50 (39.7%) / 76 (60.3%)
PGC: Primary gastric cancer, RGC: Remnant gastric cancer		
* Median (minimum-maximum)		

Table 4: The characteristics in 126 patients with gastric cancer in the test cohort.

Discussion

There are three critical steps in obtaining accurate results from DNA methylation genome-wide analysis. The first is macrodissection of a tumor area by preferably two qualified

pathologists based on representative H&E stained sections. Inaccurate macrodissection can cause contamination with adjacent non-cancerous tissues, which engenders unreliable results; thus, careful macrodissection is required. The second

is assessment of the DNA quality (quality check: QC). Samples which fail the QC ($\Delta Cq > 5.0$) may give poor quality data. Therefore, samples with $\Delta Cq > 5.0$ should be removed and others used. The third step is calculation of the β -value, which is determined by a data analysis tool for the array platform software as the methylated signal / the total (methylated + unmethylated) signal¹⁷. The β -value ranges from 0–1 (or 0%–100%), which is simple to interpret biologically¹⁷. The main problem with this value is its poor statistical properties, since its high heteroscedasticity implies that variance across samples at methylation range extremes ($\beta = 0$ or $\beta = 1$) is highly reduced¹⁷. In addition, due to poor sample quality, β -values may not show reproducible biphasic peaks²², and samples without such peaks should be excluded from further study. In addition, target gene should be chosen based on the criteria of having larger beta values, being related to CpG islands in the promoter region, and being suitable for primer and probe design for qMSP.

Evaluation of DNA methylation at a CpG locus in the human genome is performed using microarray-based technology with a fixed number of probes to survey specific genomic loci. It is the most widely used method in epigenome-wide association studies (EWAS) due to its low cost, small amount of DNA required, and markedly shorter sample processing time, which allows high-throughput processing of many clinical samples²³. However, an array platform for complex evaluation of DNA methylation at an individual CpG locus is limited by the number and specificity of probes for epigenetically altered loci, which prevents exploration of some genomic regions. WGBS is generally viewed as the gold standard method due to its wider spectrum of genomic coverage^{10, 11}. However, this method has a substantial cost and a relatively long processing time for analysis of a large number of samples^{10, 11}. Thus, it is not always feasible. In

comparison, the array platform for complex evaluation of DNA methylation at an individual CpG locus in the human genome is reasonable for use in terms of cost and genomic coverage. Recently, the latest upgraded bead chips have gotten ready to use²⁴. These assays can help us analyze nearly doubled measured CpG sites, which can achieve ideal genome-wide association study (GWAS) for large sample populations.

In summary, DNA methylation genome-wide analysis with the array platform for complex evaluation of DNA methylation at an individual CpG locus in the human genome can provide important information on epigenetic biomarkers in gastrointestinal cancer. Compared with WGBS, this method is cost effective and reduces sample-processing time. Therefore, this method for detection of DNA methylation at a CpG locus is likely to be widely used in epigenetic biomarker research.

Disclosures

The authors have nothing to disclose.

Acknowledgments

We are especially grateful to all members of the Department of Surgery, The Sidney Kimmel Comprehensive Cancer Center at the Johns Hopkins University School of Medicine for useful discussions and technical support. We also thank Kristen Rodgers for generous technical guidance on the procedures for bisulfite treatment and qMSP.

References

1. Qiu, J. Epigenetics: unfinished symphony. *Nature*. **441** (7090), 143–145 (2006).

2. Okugawa, Y., Grady, W. M., Goel, A. Epigenetic alterations in colorectal cancer: emerging biomarkers. *Gastroenterology*. **149** (5), 1204-1225 (2015).
3. Ahuja, N., Li, Q., Mohan, A. L., Baylin, S. B., Issa, J. P. Aging and DNA methylation in colorectal mucosa and cancer. *Cancer Research*. **58** (23), 5489-5494 (1998).
4. Hsieh, C. J. et al. Hypermethylation of the p16INK4a promoter in colectomy specimens of patients with long-standing and extensive ulcerative colitis. *Cancer Research*. **58**, 3942-3945 (1998).
5. Ushijima, T., Okochi-Takada, E. Aberrant methylations in cancer cells: where do they come from? *Cancer Science*. **96** (4), 206-211 (2005).
6. Coppedè, F., Epigenetic biomarkers of colorectal cancer: Focus on DNA methylation. *Cancer Letters*. **342** (2), 238-247 (2014).
7. Vasilatou, D., Papageorgiou, S. G., Dimitriadis, G., Pappa, V. Epigenetic alterations and microRNAs: new players in the pathogenesis of myelodysplastic syndromes. *Epigenetics*. **8** (6), 561-570 (2013).
8. Ma, X., Wang, Y. W., Zhang, M. Q., Gazdar, A. F. DNA methylation data analysis and its application to cancer research. *Epigenomics*. **5** (3), 301-316 (2013).
9. Morris, T. J., Beck, S. Analysis pipelines and packages for Infinium HumanMethylation450 BeadChip (450k) data. *Methods (San Deigo, California)*. **72**, 3-8 (2015).
10. Rauluseviciute, I., Drabløs, F., Rye, M. B. DNA methylation data by sequencing: experimental approaches and recommendations for tools and pipelines for data analysis. *Clinical Epigenetics*. **11** (1), 193 (2019).
11. Cazaly, E. et al. Making sense of the epigenome using data integration approaches. *Frontiers in Pharmacology*. **10**, 126 (2019).
12. Leal, A., Sidransky, D., Brait, M. Tissue and cell-free DNA-based epigenomic approaches for cancer detection. *Clinical Chemistry*. **66** (1), 105-116 (2020).
13. Song, W., Ren, J., Wang, W. J., Wang, C. T., Fu, T. Genome-wide methylation and expression profiling identify a novel epigenetic signature in gastrointestinal pan-adenocarcinomas. *Epigenomics*. **12** (11), (2020).
14. Wong, E. M. et al. Tools for translational epigenetic studies involving formalin-fixed paraffin-embedded human tissue: applying the Infinium HumanMethylation450 Beadchip assay to large population-based studies. *BMC Research Notes*. **8**, 543 (2015).
15. *Illumina*. https://jp.support.illumina.com/downloads/infinium_humanmethylation450_product_files.html (2020).
16. Fackler, M. J. et al. Genome-wide methylation analysis identifies genes specific to breast cancer hormone receptor status and risk of recurrence. *Cancer Research*. **71** (19), 6195-6207 (2011).
17. Dedeurwaerder, S. et al. A comprehensive overview of Infinium HumanMethylation450 data processing. *Briefings in Bioinformatics*. **15** (6), 929-941 (2014).
18. Hulbert, A. et al. Early detection of lung cancer using DNA promoter hypermethylation in plasma and sputum. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*. **23** (8), 1998-2005 (2017).

19. Shimada, H., Fukagawa, T., Haga, Y., Oba, K. Does remnant gastric cancer really differ from primary gastric cancer? A systematic review of the literature by the Task Force of Japanese Gastric Cancer Association. *Gastric Cancer: Official Journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association*. **19** (2), 339-349 (2016).
20. Eijnsink, J. J. et al. Detection of cervical neoplasia by DNA methylation analysis in cervico-vaginal lavages, a feasibility study. *Gynecologic Oncology*. **120** (2), 280-283 (2011).
21. Sugimoto, K. et al. DNA methylation genome-wide analysis in remnant and primary gastric cancers. *Gastric Cancer: Official Journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association*. **22** (6), 1109-1120 (2019).
22. Wang, Z., Wu, X., Wang, Y. A framework for analyzing DNA methylation data from Illumina Infinium HumanMethylation450 BeadChip. *BMC Bioinformatics*. **19** (Suppl 5), 115 (2018).
23. Teh, A. L. et al. Comparison of methyl-capture sequencing vs. Infinium 450K methylation array for methylome analysis in clinical samples. *Epigenetics*. **11** (1), 36-48 (2016).
24. Zhou, W., Laird P. W., Shen, H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation BeadChip probes. *Nucleic Acids Research*. **45** (4), e22 (2017).