Tech Note

Global and Locus-specific 5-Methylcytosine Detection

Josh Haimes, Amanda Haas, David Mierzejewski, Melissa Kelley
Thermo Fisher Scientific, Lafayette, CO, USA

Abstract

5-Methylcytosine (mC) is a modified DNA base found in the genomes of many species and has been implicated to function epigenetically in a diverse set of cellular processes and disease states. Presented here are two methods for studying mC in the genome. In one method, genomic DNA (gDNA) is hydrolyzed to single nucleotides by T7 DNA polymerase, and then analyzed by UPLC to find total genomic mC levels as a percent of all cytosine species. This method is utilized to analyze gDNA from human cells, demonstrating small, yet reproducibly discernible differences among samples. Second, a locus-specific method is described that provides percent methylation information within a sequence context of the modification. This is achieved using isoschizomers, or pairs of restriction endonucleases with the same recognition sequence, which have divergent sensitivity to methylation, followed by real-time quantitative PCR (qPCR) using primers flanking the cleavage site. Gene body and promoter sequences are analyzed for locus-specific methylation status to demonstrate the utility of this detection method.

Genomic DNA Hydrolysis

- Incubate 2.5 µg DNA (in 43 µL water) at 95 °C for 10 min
- Add 5 µL of 10x T7 DNA polymerase reaction buffer and 10 U T7 DNA polymerase (Thermo Scientific, Cat #EP0081)
- Incubate reaction overnight at 37 °C
- Inactivate the reaction by incubating at 75 °C for 10 min
- Centrifuge sample for 45 min at > 12,000 × g

UPLC Conditions

- Column and Temperature: UPLC BEH C18 (Waters, Cat #186002352); 10 °C
- Ion Pairing Buffer System:
  - » Buffer A: 50 mM Dimethyl,hexyl-amine acetate (DMHAA, CAS #4385-04-0)
  - » Buffer B: 50 mM DMHAA, 10% acetonitrile
- Sample: 10 µL of 2 ng/µL T7-hydrolyzed gDNA
- Gradient and Flow Rate:

Time (min.)	Flow (mL/min)	Buffer A (%)	Buffer B (%)
Initial	0.075	62	38
30	0.075	62	38
40	0.075	0	100
42	0.075	0	100
42	0.075	62	38
44	0.075	62	38
45	1.01	100	0

- Detection: 254 nm and 280 nm, Acquity Photodiode Array eλ Detector

Box 1. Global mC analysis by UPLC

Introduction

5-Methylcytosine (mC) is a modified DNA base found in the genomes of many species from bacteria to vertebrates. It is the most common epigenetic mark in eukaryotes and, in mammals, this heritable modification is represented in about one percent of all DNA bases, with majority presence at CpG dinucleotides[1]. mC arises from DNA methyltansferase activity on cytosine and is thought to be removed by a process involving ten-eleven-translocation (TET) oxidation followed by base excision repair. mC is involved in several biological processes including transcriptional regulation, splicing, pluripotency maintenance[2] and development[3]; it is also responsible for heterochromosome structure and X chromosome inactivation. Further, mC is implicated in disease states such as cancer, autism[4], schizophrenia[5] and obesity[6]. mC has recently received attention from the social sciences community, resulting from a report that methylation patterns correlate to socio-economic status[7]. mC is an interesting mark to study and given the diversity of its biological function, it is important to consider its epigenetic role in your system of interest. In this technical note, two methods for mC analysis are presented, one for analysis of global genomic methylation - which does not provide sequence context - and one for locus-specific analysis.

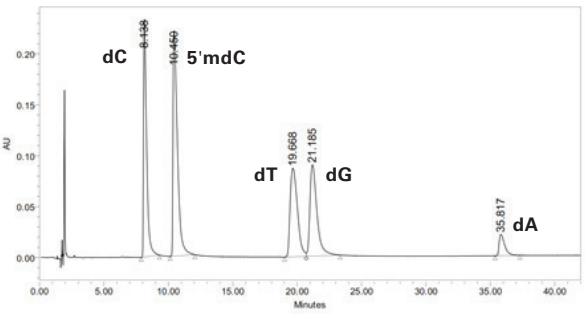
Results

Global mC Detection

The global method utilizes a UPLC system to analyze the percent of all cytosines methylated within a genomic DNA sample that has been hydrolyzed to single nucleotides by 3' to 5' exonuclease activity

Thermo SCIENTIFIC

A.



B.

Sample	Mean Empirical Modification (%)	StdDev	CV (%)	Expected Modification (%)
A	0.0	0.0	0.0	0.0
B	3.7	0.1	2.7	3.3
C	56.6	1.0	1.7	53.3

Figure 1. 5-Methylcytosine is discernible from other nucleotide monophosphates by UPLC and the method yields data that reproducibly correlates to expected values. Five nucleotide monophosphates (dA, dT, dG, dC and 5'mdC) were co-injected for UPLC analysis. Each peak from the trace is labeled with the represented species (A). The mean, standard deviation (StdDev) and coefficient of variation (CV) of five replicate UPLC analyses of hydrolyzed synthetic DNA duplexes with known mC modifications are reported along with the expected modification percent for each sample (B).

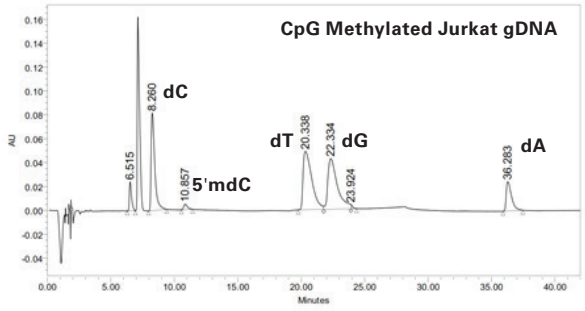


Figure 2. Example UPLC trace used to quantify mC in gDNA. A trace from hydrolyzed, enzymatically methylated gDNA is shown with each species labeled. Peaks appearing before 8.260 minutes result from components of the hydrolysis reaction.

Example Data

Enzyme	Cq Value
No Enzyme	20
MspI	27
HpaII	23

Calculations

- Percent of undigested sample: $(2^{-(27-20)}) \times 100 = 0.78\%$
- Percent modification: $(2^{-(23-20)}) \times 100 = 12.50\%$

Box 2. Example percent modification calculation from qPCR data. Incomplete digestion by MspI is considered to be greater than 4.5% and can result, for example, if the external cytosine of the CCGG recognition sequence is methylated or if the internal cytosine is hydroxymethylated.

using Thermo Scientific™ T7 DNA Polymerase (Cat #EP0081; see Box 1 for a protocol outline). The UPLC protocol was first optimized using each of five nucleotide monophosphates (dA, dT, dG, dC and 5'mdC) individually. Following, all five species were co-injected to confirm reproducible peak retention times and resolution (Figure 1A). Finally, the method was validated for dsDNA polymer using duplexed oligomers synthesized with known mC modifications. Duplexes were hydrolyzed and injected to calculate percent mC using peak areas; these empirical data compare to expected values (Figure 1B). To assess ability to detect methylation differences in genomic DNA (gDNA), Jurkat gDNA (Thermo Scientific, Cat #SD1111) and CpG Methylated Jurkat gDNA (Thermo Scientific, Cat #SD1121) were analyzed five times each (Figure 2). Mean Jurkat gDNA methylation was 3.37% (StdDev 0.37) and CpG Methylated Jurkat gDNA was 4.99% (StdDev 0.42), yielding p-value = 0.0002.

Sequence context of the modifications is lost upon hydrolysis of the sample, but this method retains utility in certain applications. For example, mC levels differ across tissue and cell types[8]. As such, this global analysis method is a simple, efficient and cost-effective initial investigation that prospecting researchers can perform to justify if further exploration of a putative role of mC in their biological system of interest is warranted. Further, given the amenability of this method to an automated platform, total genomic mC could be a read-out for screening response to treatments or for assessing levels in multiple samples in high throughput.

Locus-specific mC Detection

While global methylation data has its value, sequence context information of epigenetic modifications provides important perspective. For example, to propose that mC functions as an epigenetic transcriptional regulator for a gene of interest, global methylation data would not be persuasive evidence, but mC levels in the promoter of that particular gene of interest could indicate such function[2]. One method for elucidating sequence context of methylation is to employ isoschizomers, or pairs of restriction endonucleases with the same recognition sequence, which have divergent sensitivity to methylation. For example, HpaII has the recognition sequence CCGG, but will not cleave methylated DNA; MspI recognizes the same sequence but cleaves regardless of methylation status of the CpG cytosine. TaqI and HpyF30I is another pair of isoschizomers with divergent sensitivity to methylation that can be used to assay mC at the recognition sequence TCGA. The Thermo Scientific™ EpiJET™ DNA Methylation Analysis Kits (Cat #K1441 and K1451) include formulations of these isoschizomer pairs that digest gDNA samples in one hour and include a buffer for the enzymes, as well as methylated and unmethylated DNA controls. Resultant fragments from such digestions can be analyzed for percent modification by qPCR using primers flanking the cleavage site (for example calculations see Box 2).

To demonstrate the utility of these reagents, Jurkat gDNA was digested with the two enzyme pairs, followed by qPCR analysis using primers for the *RASSF1* promoter or *THRB* gene. Figure 3 shows the amplification plots and ΔCq values used to calculate percent methylation at these loci.

Conclusion

With the research community's increased attention to methylation and its role in biology, several methods for investigating epigenetic modifications have emerged recently. See Table 1 for a list of Thermo Scientific™ products useful for epigenetic study. For more information, visit thermoscientificbio.com/molecular-biology-applications/epigenetics

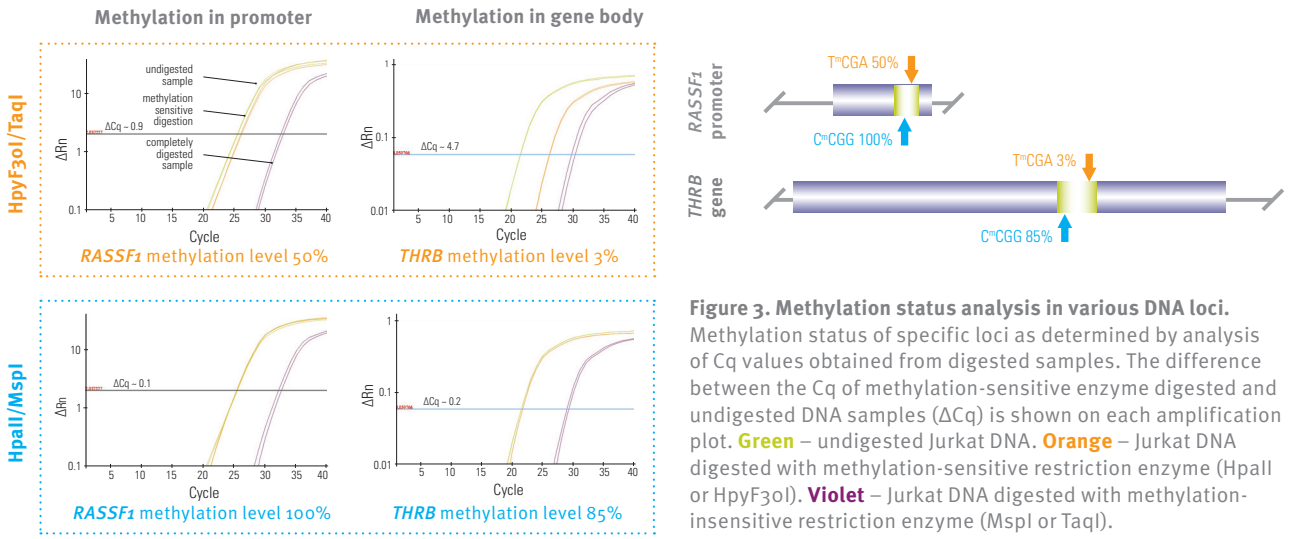


Figure 3. Methylation status analysis in various DNA loci. Methylation status of specific loci as determined by analysis of Cq values obtained from digested samples. The difference between the Cq of methylation-sensitive enzyme digested and undigested DNA samples (ΔCq) is shown on each amplification plot. **Green** – undigested Jurkat DNA. **Orange** – Jurkat DNA digested with methylation-sensitive restriction enzyme (HpaII or HpyF30I). **Violet** – Jurkat DNA digested with methylation-insensitive restriction enzyme (MspI or TaqI).

Table 1. Thermo Scientific™ epigenetic product offerings.

Product Name	Utility	Cat #
EpiJET DNA Methylation Analysis Kit (MspI/HpaII)	Methylation analysis at CCGG sequence	K1441
EpiJET DNA Methylation Analysis Kit (TaqI/HpyF30I)	Methylation analysis at TCGA sequence	K1451
EpiJET 5-hmC Analysis Kit	Hydroxymethylation analysis at CCGG	K1481
EpiJET Bisulfite Conversion Kit DNA	Bisulfite Conversion for methylation analysis	K1461
CpG Methylated Human Genomic DNA	Control DNA for use in methylation studies	SD1131
CpG Methylated Jurkat Genomic DNA	Control DNA for use in methylation studies	SD1121
Jurkat Genomic DNA	Control DNA for use in methylation studies	SD1111
CpG Methyltransferase (M.SssI)	*In vitro* enzymatic CpG methylation of dsDNA	EM0821
SgeI	Restriction endonuclease which cleaves only methylated DNA at recognition sequence m5CNNG(9/13)^	ER2211

References

1. A. Bird, DNA methylation patterns and epigenetic memory. *Genes Dev.* **16**, 6-21 (2002).
2. G. Ficz, M. R. Branco, et al., Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature.* **473**, 398-402 (2011).
3. E. Li, T.H. Bestor, et al., Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell.* **69**, 915-926 (1992).
4. C.C. Wong, E.L. Meaburn, et al., *Nat. Mol. Psych.*, in press (available at nature.com/mp/journal/vaop/ncurrent/full/mp201341a.html).
5. P.A. Melas, M. Rogdaki, et al., Epigenetic aberrations in leukocytes of patients with schizophrenia: association of global DNA methylation with antipsychotic drug treatment and disease onset. *FASEB J.* **26**, 2712-2718 (2012).
6. F.I. Milagro, J. Campión, et al., A dual epigenomic approach for the search of obesity biomarkers: DNA methylation in relation to diet-induced weight loss. *FASEB J.* **26**, 1378-1389 (2011).
7. P. Tehranifar, H. Wu, et al., Early life socioeconomic factors and genomic DNA methylation in mid-life. *Epigenetics.* **8**, 23-27 (2013).
8. M. Ehrlich, M.A. Gama-Sosa, Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* **10**, 2709-2721 (1982).

thermoscientific.com/onebio

Europe

Customer Service
cs.molbio.eu@thermofisher.com

Technical Support
ts.molbio.eu@thermofisher.com

Tel 00800 222 00 888
Fax 00800 222 00 889

United States

Customer Service
cs.molbio@thermofisher.com

Technical Support
ts.molbio@thermofisher.com

Tel 800 235 9880
Fax 800 292 6088

Canada

Customer Service
cs.molbio@thermofisher.com

Technical Support
ts.molbio@thermofisher.com

Tel 800 340 9026
Fax 800 472 8322

Thermo SCIENTIFIC
Part of Thermo Fisher Scientific