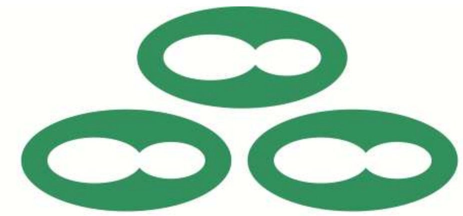


A Novel Method for Locus Specific Detection of 5-Hydroxymethylcytosine

James L. Yen, Stella Choi, Marc Van Eden, & Xi Yu Jia

Zymo Research Corp., Irvine, CA

Background

Epigenetic misregulation leading to diseases, such as cancer, has classically been understood in terms of aberrant DNA methylation patterns. For decades, 5-methylcytosine (5mC) has been the focus of epigenetic DNA modification studies. But recently, a novel epigenetic DNA modification, 5-hydroxymethylcytosine (5hmC), has been shown to be present at high levels in mammalian brain and ES cells.[1,2] It has also been found at appreciable levels in many normal and diseased tissues.

DNA methylation (5-methylcytosine) plays a major role in gene regulation. Promoters of many genes are methylated for regulated gene silencing. When this set DNA methylation pattern is undermined, the risk of acquiring diseases may increase. It has been well documented that many diseases (e.g. cancer, neurodegenerative diseases, etc.) have epigenetic roots. Many of these misregulated epigenetic alterations have proven to be reliable markers of disease or of disease progression.

The role of the newly discovered, "6th base" (5-hydroxymethylcytosine) is less clear. Unlike methylation, 5hmC has been shown to be found predominantly within gene bodies as well as non-promoter regions. Its role in biology has been speculated to be involved in DNA demethylation (active and passive) as well as in poising genes for later activation.

Current tools available to map DNA methylation (e.g. bisulfite conversion) cannot distinguish between 5mC and 5hmC.[3] Here we describe the first method for locus specific interrogation of 5hmC within DNA by use of our Quest 5-hmC Detection Kit™.

Method Highlights

The strength of Quest 5-hmC Detection Kit™ is that a number of glucosyl-5-hydroxymethylcytosine sensitive restriction endonucleases (GSRE) (Table 1) can be used for 5hmC detection thereby allowing for:

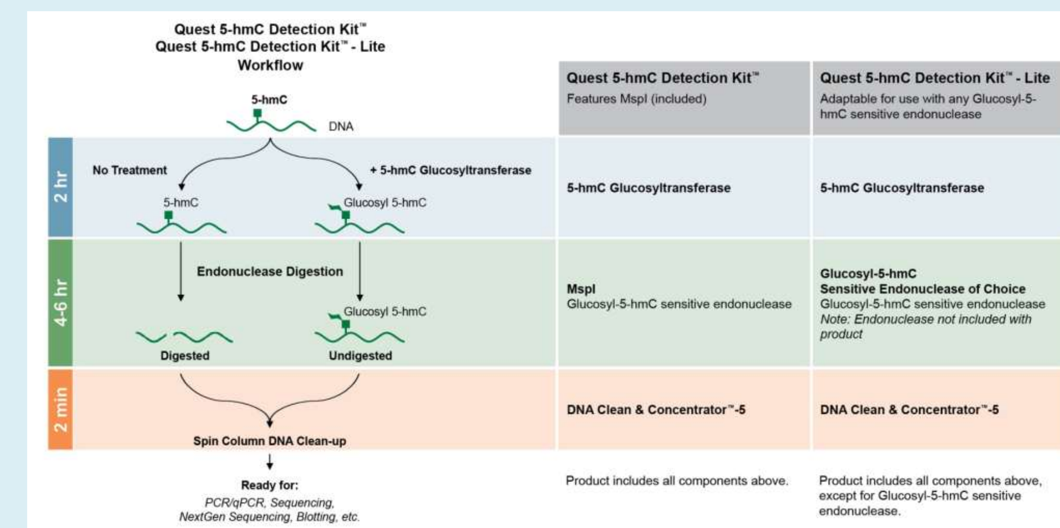
1. Detection at wider ranges of sequence contexts
2. CpG centered (MspI)
3. Non-CpG centered (GTAC, GGCC, GATC, etc.)
4. Hemi-methylation sequence contexts

GSRE	Recognition Sequence
MspI	CCGG
GlaI	GCGC
Csp6I	GTAC
HaeIII	GGCC
Taq$^{\alpha}$I	TCGA
MboI	GATC
McrBC	R^{m}C(N$_{40\text{-}3000}$)R^{m}C

Table 1: $^{Glucosyl\text{-}5hm}$C Sensitive Restriction Endonucleases – GSREs
GSREs cleave DNA when cytosine, 5-methylcytosine, or 5-hydroxymethylcytosine is located within their recognition sequence. Glucosylation of 5hmC (i.e., glucosyl-5-hydroxymethylcytosine) inhibits cleavage by the enzyme. Detection of 5hmC with MspI and HaeIII can only be used for the C indicated (by an underline).

Overview

5hmC residues are specifically modified, making them resistant to cleavage by a special class of restriction endonucleases. This lack of cleavage can then be efficiently quantified by qPCR. Our method is reliable, robust, and highly reproducible.



Methodology

DNA is first treated with a 5hmC glucosyltransferase enzyme that specifically adds a glucose moiety onto preexisting 5hmC residues, yielding glucosyl-5-hydroxymethylcytosine (glu-5-hmC) (Figure 1A). Subsequent digestion of DNA with glu-5-hmC sensitive restriction endonucleases (GSREs) effectively cleaves DNA with cytosine, 5mC, or 5hmC within their recognition site. But when glu-5-hmC is present, DNA is not able to be cleaved (Figure 1B). Exploiting this difference in ability of cleavage, we show that 5hmC can be faithfully detected at any loci of interest by qPCR (Figure 2).

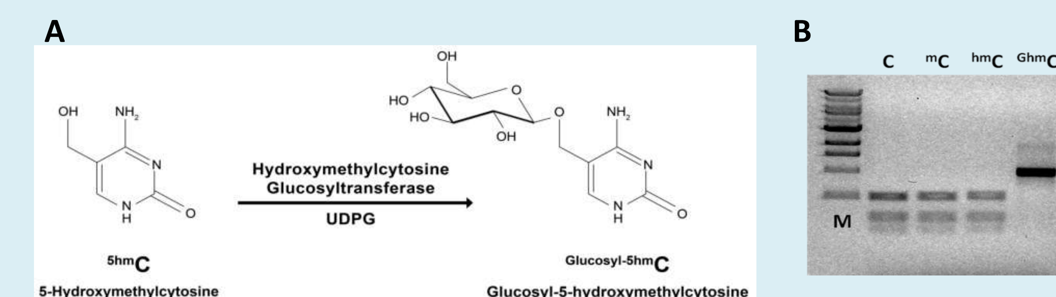


Figure 1: (A) 5-hmC Glucosyltransferase transfers a glucose moiety from uridine diphosphoglucose (UDPG) onto preexisting 5-hydroxymethylcytosines within DNA. **(B)** DNA with all cytosines as: unmodified (C), 5-methylcytosine (mC), 5-hydroxymethylcytosine (hmC), or glucosyl-5-hydroxymethylcytosine (GhmC) was digested with 4 units of Csp6I for 2 hours at 37°C and resolved in a 0.8% w/v agarose/TAE/EtBr gel.

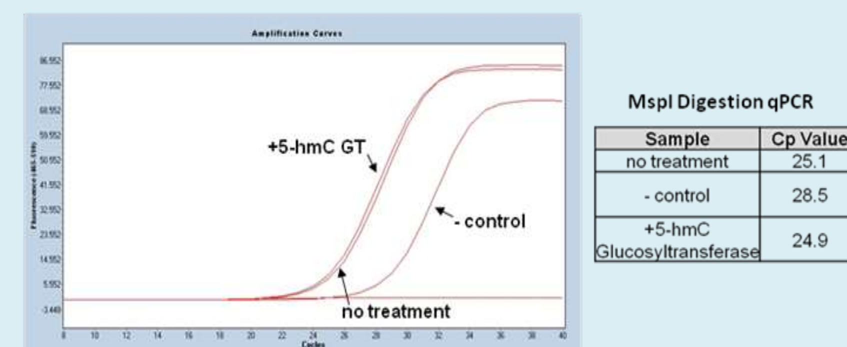


Figure 2: Detection of 5-hmC by qPCR
DNA template with 5-hmC at the inner C position of the MspI site (CCGG) was processed according to protocol (+/- glucosylation and MspI digested). Differences in qPCR amplification efficiencies (Cp values) between "- control" (unglucosylated) and "+5-hmC GT" (glucosylated) samples indicates the presence of 5-hmC. "no treatment" DNA control establishes the level (i.e., Cp value) representative of complete hydroxymethylation at the interrogation site.

Results

5-hydroxymethylcytosine is found in genomic DNA from various mouse organs

In order to identify candidate gDNA samples to study, we first sought to evaluate global levels of DNA methylation (5-methylcytosine and 5-hydroxymethylcytosine). Genomic DNA from various mouse organs were first quantified for global levels of 5hmC and 5mC by ELISA using 5hmC polyclonal antibody (Zymo Research – 5hmC pAb) and 5mC specific monoclonal antibody (Zymo Research – 5mC mAb) (Fig. 2).

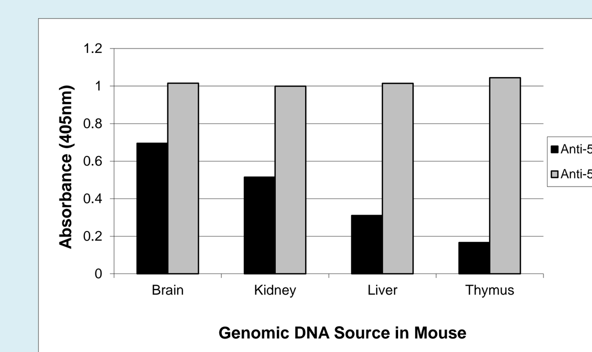


Figure 3: Global 5hmC & 5mC Quantitation of Mouse gDNA via ELISA
Anti-5-methylcytosine mAb (Zymo Research) and anti-5-hydroxymethylcytosine pAb (Zymo Research) were used for ELISA detection.

ELISA results clearly show that 5hmC and 5mC are found in all genomic DNA from all organs tested. Interestingly, global 5mC levels are roughly consistent among samples, while 5hmC levels show drastic differences.

5-Hydroxymethylcytosine is found in varying levels at different loci

The best strategy to verify the efficacy of our 5hmC detection method was to validate select loci previously identified to contain high levels of 5hmC in mouse brain. These loci were previously identified by HMeDIP (hydroxymethylcytosine DNA immunoprecipitation).

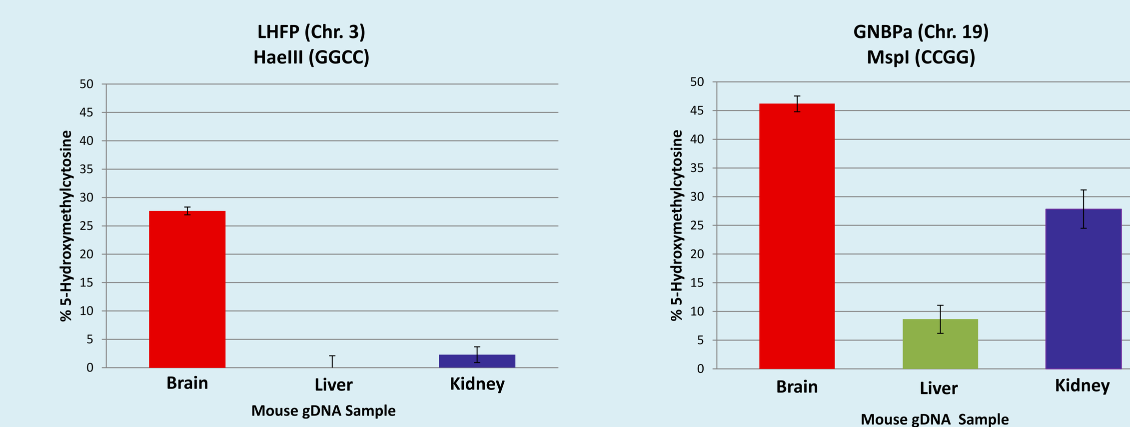


Figure 4: Loci Specific 5-hydroxymethylcytosine Detection
Genomic DNA from Mouse 5hmC & 5mC DNA Set (Zymo Research) was processed according to Quest 5-hmC Detection Kit™ protocol. qPCR results were quantified as percentage of 5-hydroxymethylcytosine at each loci. 5-hydroxymethylcytosine was detected at LHFP and GNBPa loci with the use of HaeIII (GGCC) and MspI (CCGG) , respectively. Use of HaeIII allows for greater coverage as well as detection of 5-hmC in CpG and non-CpG contexts.

Our Quest 5-hmC Detection Kit™ was effective in detecting and quantitating levels of 5hmC at specific loci of interest using both GSREs, HaeIII and MspI.

Presence of 5-hydroxymethylcytosine is not indicative of differential methylation patterns

We next looked at the DNA methylation profile of the LHFP loci from genomic DNA obtained from the brain, kidney, and liver. LHFP in these three tissues were shown to contain different levels of 5hmC (Figure 4).

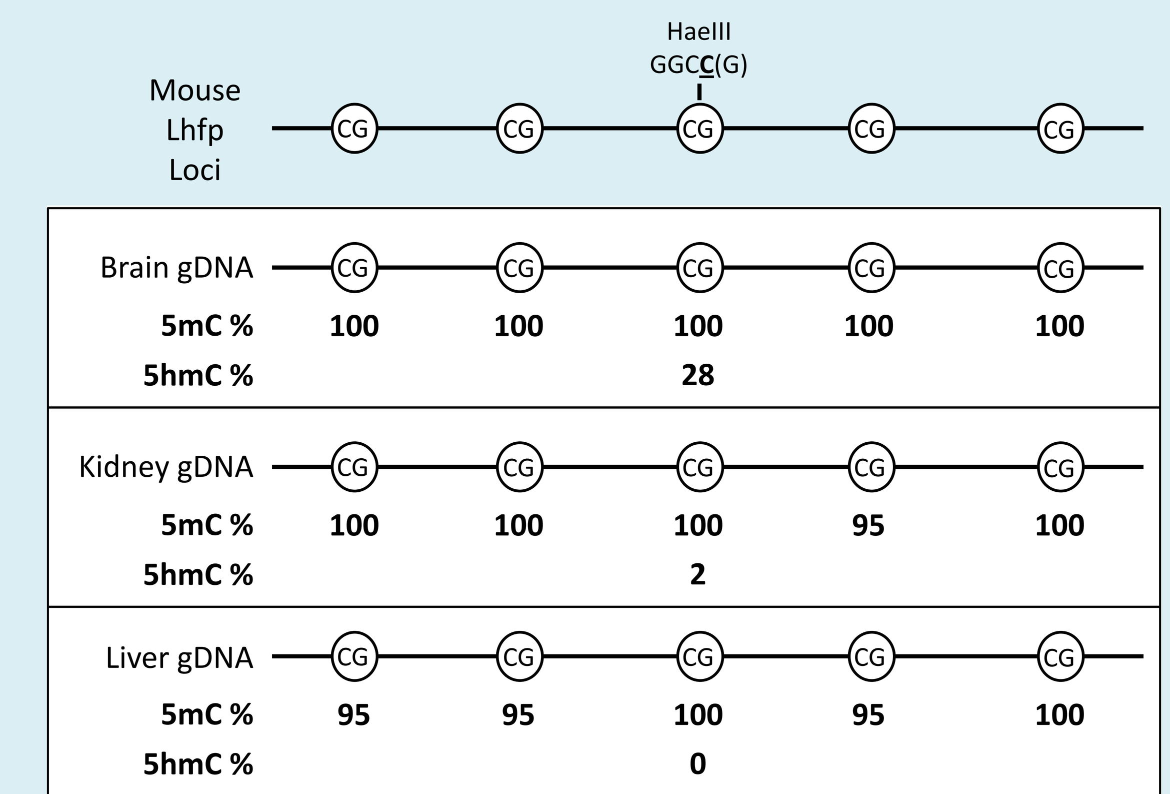


Figure 5: DNA Methylation Profile at Lhfp loci – 5mC and 5hmC
Five CpG residues were inspected by bisulfite conversion to determine DNA methylation signature. 20 independent clones from each organ were sequenced and percent methylation at each CpG site was calculated from these 20 clones. Use of Quest 5hmC Detection Kit™ with HaeIII was used to detect 5hmC at a single CpG residue.

Although bisulfite conversion cannot distinguish between 5mC and 5hmC, our results show the general methylation pattern from gDNA containing high levels of 5hmC (i.e. brain), low levels (i.e. kidney), to undetectable levels of 5hmC (i.e. liver) are virtually identical.

Discussion and Conclusion

The role of 5hmC in biology is still speculative. However, our data adds proof that 5hmC is a bona fide epigenetic mark found at varying levels at different loci from different tissue sources. Our method has proven to effectively detect and quantitate 5hmC at genomic loci of interest. Also, the utility of our method is underscored by the observation that the DNA methylation signatures (as gauged by bisulfite conversion) at particular loci are identical despite varying levels of 5hmC.

References

1. Tahiliani M. *et al.* (2009) "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1" **Science**.
2. Kriaucionis S *et al.* (2009) "The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain" **Science**.
3. Huang Y. *et al.* (2010) "The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing" **PLoS One**.
4. Song CX *et al.* (2011) "Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine" **Nature Biotech.**