Published in final edited form as:
Epigenomics. 2011 October ; 3(5): 521–523. doi:10.2217/epi.11.74.

The hunt for 5-hydroxymethylcytosine: the sixth base

Chun-Xiao Song and
Department of Chemistry & Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL, USA

Chuan He
Department of Chemistry & Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL, USA

Keywords

5-hydroxymethylcytosine; detection; method; sequencing

"5-hmC is now widely accepted as the sixth base in the mammalian genome, following its precursor, 5-methylcytosine (5-mC), the fifth base."

After its rediscovery in 2009, 5-hydroxymethylcytosine (5-hmC) and its natural creators, the Tet dioxygenases, have received a tremendous amount of attention from the epigenetics and other related communities [1,2]. 5-hmC is now widely accepted as the sixth base in the mammalian genome, following its precursor, 5-methylcytosine (5-mC), the fifth base. To elucidate the exact biological roles of this base modification, the development of new analytical technologies, including detection and sequencing methods for 5-hmC, has become essential.

The first two reports in 2009 used ^{32}P-radioactive labeling 1D or 2D thin layer chromatography to detect 5-hmC [1,2]. With a detection limit of approximately 0.08% of total nucleotide, the thin layer chromatography method is able to detect 5-hmC in genomic DNA from brain tissue and embryonic stem cells; however, it fails to yield a signal with genomic DNA isolated from other cultured cells, especially cancer cells. More recently, a more accurate HPLC-mass spectroscopy method has been developed with isotope-labeled 5-hmC as an internal standard [3]. This method allows a complete quantification of the 5-hmC in all mouse tissues; however, the detection limit still needed to be improved. Since then anti-5-hmC antibodies have become available and are widely used to detect 5-hmC [4–7]. However, this approach is not very quantitative owing to its density-dependent recognition of 5-hmC, as expected [8]. Anti-5-methylenesulfonate – generated from 5-hmC after a bisulfite treatment – antibodies have been demonstrated to be more quantitative than anti-5-hmC antibodies [8]. However, these antibody-based methods cannot improve the detection limit.

Correspondence to: Chuan He.

Author for correspondence: Tel.: +1 773 702 5061 Fax: +1 773 702 0805 chuanhe@uchicago.edu.

Financial & competing interests disclosure
The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.
No writing assistance was utilized in the production of this manuscript.

In 2010, enzymatic labeling of 5-hmC began to emerge [9]. T4-β-glucosyltransferase (β-GT) is the first enzyme of choice, since it naturally transfers a glucose from UDP-glucose to 5-hmC in T4 bacteriophages. Two groups independently reported the use of β-GT to detect and quantify 5-hmC. Szwagierczak *et al.* used ^{3}H radioactive glucose to label and quantify 5-hmC by scintillation counting [10]. Song *et al.* used β-GT to transfer an azide-labeled glucose to 5-hmC and subsequently attached a biotin group through the azido group via copper-free click reaction [11]. This strategy allows for the detection and quantification of 5-hmC in a simple dot-blot assay using avidin-horseradish peroxidase. The approach of β-GT labeling greatly improved the detection limit of 5-hmC from 0.004–0.006% of total nucleotide, and displayed a linear response to 5-hmC since the β-GT-based method is able to label every single 5-hmC in the genome. With these highly sensitive methods, extremely low levels of 5-hmC in other cultured cells, especially cancer cells, can be measured [11]. This is of particular interest since Tet enzymes may play important roles in various cancers; however, cancer cell lines tend to have low levels of 5-hmC [11,12]. The β-GT-based labeling approaches represent promising diagnostic tools for specific cancers.

Compared to simple 5-hmC detection, sequencing methods amenable for whole-genome, next-generation sequencing of 5-hmC are much more desirable, especially if 5-hmC-containing genomic DNA fragments can be enriched first. We generated the first genome-wide 5-hmC distribution map using the same β-GT/azide–glucose/biotin-labeling method [11]. The high specificity of the click chemistry and the highly selective and strong biotin-avidin/streptavidin interaction enabled the robust and efficient pull-down of 5-hmC-containing DNA for deep sequencing. The biotin–avidin/streptavidin enrichment approach proved superior to the antibody-based pull down approach, which has a density-dependent capture bias for 5-hmC [8]. Besides the azido group, other functional groups such as aldehyde and ketone, which can be connected to a biotin via hydroxylamine–aldehyde/ketone condensation, can also be introduced through the β-GT-based labeling of 5-hmC [8,13]. The specificity of the hydroxylamine–aldehyde/ketone condensation may introduce high background since h ydroxylamine is known to react with cytosine.

In an alternative approach, Robertson *et al.* demonstrated that J-binding protein 1 (JBP-1), which is known to recognize glucosylated 5-hydroxymethyluracil, can also bind to glucosylated 5-hmC [14]. Therefore, following β-GT treatment, 5-hmC-containing DNA can be isolated by JBP1 and immobilized on magnetic beads. However, the specificity of this JBP1-based pull-down of glucosylated 5-hmC needs to be compared with the antibody- or chemical labeling-based methods.

Another enzymatic labeling strategy emerged from an interesting discovery made by Liutkevičiūtė *et al.* who demonstrated that bacterial methyltransferases are able to label 5-hmC with a sulfur- or selenium-containing moiety, which can be further functionalized with a biotin group [15]. However, before this strategy can be applied to genomic DNA samples, the sequence-specificity issue of bacterial methyltransferasesneeds to be solved and the labeling yield needs to be improved.

The next challenge is to obtain whole genome, single-base resolution sequencing of 5-hmC. To date, most pull down methods enable whole-genome determination of 5-hmC distribution. In principle, some methods promise single-base resolution sequencing of 5-hmC using either polymerase extension [11] or exonuclease digestion [13]. To this end, recent advances in third-generation sequencing technologies, such as the single-molecule real-time (SMRT) sequencing [16] and nanopore sequencing [17] have demonstrated their ability to distinguish between 5-mC and 5-hmC on synthetic DNA templates. Further technology development is required before these methods can be applied to real biological

samples. Perhaps a combination of the above mentioned 5-hmC labeling methods and the third-generation sequencing technologies will provide an attractive solution.

Besides genome-wide sequencing, methods that precisely detect 5-hmC at specific sites with single-base resolution are also highly desirable; the most cost efficient way is to develop an enzymatic or chemical transformation that is specific to 5-hmC and can be used for site-specific high-resolution detection of 5-hmC, as what bisulfite treatment is for 5-mC [18]. Unfortunately, bisulfite treatment cannot distinguish between 5-mC and 5-hmC [18]. Certain restriction enzymes or endonucleases can now be used to detect 5-hmC at specific loci with the help of β-GT-mediated glycosylation; however, this approach is highly limited owing to the sequence dependence of those enzymes [19,101]. Alternative chemical transformations specific to 5-hmC have yet to be reported, and should be a focus of research in the future. Without single-base resolution information of 5-hmC in the genome, particularly in comparison to distributions of 5-mC revealed by bisulfite-based sequencing, the exact roles of 5-hmC cannot be ascertained.

There still remains a lot to be learned regarding the biology of 5-hmC [20]. In the future development of this newly born area of study, 5-hmC detection and sequencing methods will play key roles to dramatically advance the field. Recently developed methods have already demonstrated their impact and will continue to drive the development of this young research field, especially if single-base resolution detection of 5-hmC can be accomplished and made available to biologists in their everyday techniques in the near future.

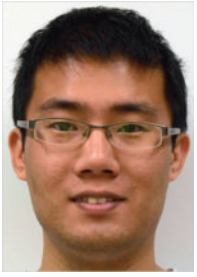
Acknowledgments

The authors thank Sarah Frank Reichard for help with e diting the manuscript.

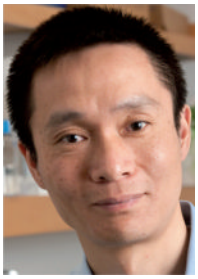
This work was partially supported by the NIH (GM 071440 to Chuan He). A provisional patent application on 5-hmC labeling using glucose and modified glucose by β-*GT was officially filed in April 2010.*

Biography

Chun-Xiao Song



Chuan He



Bibliography

1. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009; 324(5929):929–930. [PubMed: 19372393]
2. Tahiliani M, Koh KP, Shen Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009; 324(5929):930–935. [PubMed: 19372391]
3. Munzel M, Globisch D, Bruckl T, et al. Quantification of the sixth DNA base hydroxymethylcytosine in the brain. Angew. Chem. Int. Ed. Engl. 2010; 49(31):5375–5377. [PubMed: 20583021]
4. Jin SG, Wu X, Li AX, Pfeifer GP. Genomic mapping of 5-hydroxymethylcytosine in the human brain. Nucleic Acids Res. 2011; 39(12):5015–5024. [PubMed: 21378125]
5. Wu H, D'Alessio AC, Ito S, et al. Genome-wide ana lysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. Genes Dev. 2011; 25(7):679–684. [PubMed: 21460036]
6. Williams K, Christensen J, Pedersen MT, et al. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. Nature. 2011; 473(7347):343–348. [PubMed: 21490601]
7. Ficz G, Branco MR, Seisenberger S, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011; 473(7347):398–402. [PubMed: 21460836]
8. Pastor WA, Pape UJ, Huang Y, et al. Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. Nature. 2011; 473(7347):394–397. [PubMed: 21552279]
9. Höbartner C. Enzymatic Labeling of 5-hydroxymethylcytosine in DNA. Angew. Chem. Int. Ed. 2011; 50(19):4268–4270.
10. Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H. Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res. 2010; 38(19):e181. [PubMed: 20685817]
11. Song C-X, Szulwach KE, Fu Y, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat. Biotechnol. 2011; 29(1):68–72. [PubMed: 21151123]
12. Ko M, Huang Y, Jankowska AM, et al. Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2. Nature. 2010; 468(7325):839–843. [PubMed: 21057493]
13. Song C-X, Sun Y, Dai Q, et al. Detection of 5-hydroxymethylcytosine in DNA by transferring a keto-glucose by using T4 phage β-glucosyltransferase. Chembiochem. 2011; 12(11):1682–1685. [PubMed: 21656634]
14. Robertson AB, Dahl JA, Vagbo CB, et al. A novel method for the efficient and selective identification of 5-hydroxymethylcytosine in genomic DNA. Nucleic Acids Res. 2011; 39(8):e55. [PubMed: 21300643]
15. Liutkevičiūtė Z, Kriukienė E, Grigaitytė I, Masevičius V, Klimašauskas S. Methyltransferase-directed derivatization of 5-hydroxymethylcytosine in DNA. Angew. Chem. Int. Ed. 2011; 50(9): 2090–2093.
16. Flusberg BA, Webster DR, Lee JH, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat. Methods. 2010; 7(6):461–465. [PubMed: 20453866]
17. Wanunu M, Cohen-Karni D, Johnson RR, et al. Discrimination of methylcytosine from hydroxymethylcytosine in DNA molecules. J. Am. Chem. Soc. 2010; 133(3):486–492. [PubMed: 21155562]
18. Huang Y, Pastor WA, Shen Y, et al. The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. PLoS ONE. 2010; 5(1):e8888. [PubMed: 20126651]
19. Song C-X, Yu M, Dai Q, He C. Detection of 5-hydroxymethylcytosine in a combined glycosylation restriction analysis (CGRA) using restriction enzyme Taq(α)I. Bioorg. Med. Chem. Lett. 2011; 21(17):5075–5077. [PubMed: 21524909]
20. Munzel M, Globisch D, Carell T. 5-hydroxymethylcytosine, the sixth base of the genome. Angew. Chem. Int. Ed. Engl. 2011; 50(29):6460–6468. [PubMed: 21688365]
101. Davis T, Vaisvila R. High sensitivity 5-hydroxymethylcytosine detection in Balb/C brain tissue. J. Vis. Exp. 2011; 48:pii, 2661. www.jove.com/details.php?id=2661.

"Recently developed methods have already demonstrated their impact and will continue to drive the development of this young research field, especially if single-base resolution detection of 5-hmC can be accomplished and made available to biologists in their everyday techniques in the near future."