RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

Check for updates

Cite this: *RSC Chem. Biol.*, 2024, **5**, 500

Advances in the joint profiling technologies of 5mC and 5hmC

Bo He,[ab] Haojun Yao[c] and Chengqi Yi *[ade]

Received 2nd February 2024,
Accepted 21st March 2024

DOI: 10.1039/d4cb00034j

rsc.li/rsc-chembio

DNA cytosine methylation, a crucial epigenetic modification, involves the dynamic interplay of 5-methylcytosine (5mC) and its oxidized form, 5-hydroxymethylcytosine (5hmC), generated by ten-eleven translocation (TET) DNA dioxygenases. This process is central to regulating gene expression, influencing critical biological processes such as development, disease progression, and aging. Recognizing the distinct functions of 5mC and 5hmC, researchers often employ restriction enzyme-based or chemical treatment methods for their simultaneous measurement from the same genomic sample. This enables a detailed understanding of the relationship between these modifications and their collective impact on cellular function. This review focuses on summarizing the technologies for detecting 5mC and 5hmC together but also discusses the limitations and potential future directions in this evolving field.

1. Introduction

5-Methylcytosine (5mC), often termed the fifth base, plays crucial biological roles like regulating tissue-specific gene expression and silencing retroviral elements. The formation of 5mC occurs when a methyl group is added to the 5th carbon of cytosine. This process is facilitated by DNA methyltransferases (DNMTs) and predominantly takes place within CpG dinucleotides.[1,2] The ten-eleven translocation (TET) enzymes are known for their ability to iteratively oxidize DNA, leading to demethylation.[3–5] This process results in the production of 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC).[6] The formation of 5fC and 5caC allows for their active removal by thymine DNA glycosylase (TDG) through the base excision repair (BER) pathway, facilitating active DNA demethylation.[7] Recently, 5hmC has gained significant attention, being recognized as the sixth base in the mammalian genome.[8] As the second most abundant DNA modification following 5mC, 5hmC is now acknowledged to be a distinct epigenetic mark with functions that contrast with those of 5mC.[9–11]

Several base-conversion chemistries are used to differentiate unmodified cytosine (C) from its epigenetic variants. These methods include: (1) bisulfite-based methods like whole-genome bisulfite sequencing (WGBS), where bisulfite treatment converts C to U, but not 5mC or 5hmC;[12] and (2) bisulfite-free techniques such as enzymatic methyl-sequencing (EM-seq), using Tet2 to oxidize 5mC and 5hmC to 5caC, followed by A3A treatment, which converts C to U, excluding the newly generated 5caC.[13] This method can be optimized for single-cell 5mC detection.[14] In TET-assisted pyridine borane sequencing (TAPS), Tet2 oxidizes 5mC and 5hmC to 5caC, and then borane reduction converts 5caC to DHU.[15] However, traditional methods for measuring 5mC often fail to distinguish between 5mC and 5hmC. Thus, several base-conversion techniques have been developed to separate 5mC and 5hmC, such as oxidative bisulfite sequencing (oxBS-seq) and Tet-assisted bisulfite sequencing (TAB-seq) under bisulfite-based methods,[16,17] as well as APOBEC-coupled epigenetic sequencing (ACE-seq),[18] chemical-assisted C-to-T conversion of 5hmC sequencing (hmC-CATCH),[19] chemical-assisted pyridine borane sequencing (CAPS) in bisulfite-free approaches,[20] and single-step deamination sequencing (SSD-seq).[21] OxBS-seq employs potassium perruthenate for the oxidation of 5hmC to 5fC, followed by bisulfite treatment.[16] TAB-seq uses beta-glucosyltransferase (β-GT) to safeguard 5hmC, while Tet2 oxidizes 5mC to 5caC.[17] ACE-seq also utilizes β-GT for 5hmC protection, then applies A3A treatment to deaminate C and 5mC.[18] hmC-CATCH involves potassium ruthenate oxidation of 5hmC to 5fC, followed by indanedione labeling.[19] Similarly, CAPS uses potassium ruthenate for 5hmC oxidation to 5fC and then employs borane reduction to convert it to DHU.[20] SSD-seq utilized a screened engineered A3A protein (eA3A-v10) to selectively

[a] *Peking-Tsinghua Center for Life Sciences, Academy for Advanced Interdisciplinary Studies, Peking University, Beijing, China. E-mail: chengqi.yi@pku.edu.cn*
[b] *Peking University Chengdu Academy for Advanced Interdisciplinary Biotechnologies, Chengdu, China*
[c] *College of Chemistry and Chemical Engineering, Hunan University, Changsha, China*
[d] *State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China*
[e] *Department of Chemical Biology and Synthetic and Functional Biomolecules Center, College of Chemistry and Molecular Engineering, Peking University, Beijing, China*

© 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

deaminate C and 5mC, but not 5hmC.[21] Each method provides unique approaches for accurate differentiation of these epigenetic marks. These approaches are crucial in detecting and differentiating various DNA modifications, enhancing our understanding of epigenetic regulation.

5mC and 5hmC are different types of DNA modifications that play important roles in epigenetic regulation. 5mC is concentrated in promoter regions and is associated with repressive gene expression.[2,22] However, 5hmC, concentrated in specific genomic regions, plays a unique role in gene expression and cellular functions.[23,24] An example of its importance is observed with the methyl-CpG binding protein 2 (MECP2), where mutations cause Rett syndrome.[25] MECP2 binds strongly to 5mC but not to 5hmC in the CG context, impacting epigenetic regulation through its differential affinity.[26,27] Therefore, distinguishing between 5mC and 5hmC is crucial for accurately measuring cell-type-specific epigenomic profiles in health and disease contexts. Thus, this review aims to summarize methods for simultaneously profiling 5mC and 5hmC based on next-generation sequencing methods and to discuss the challenges, limitations, and potential prospects.

2. Detection methods based on restriction-enzymes

These methods utilize various types of restriction enzymes, capitalizing on their different responses to specific DNA modifications. This approach enables selective detection by leveraging the unique activities of these enzymes towards distinct epigenetic marks.

DARESOME

The DARESOME method achieves simultaneous profiling of 5mC and 5hmC by using a sequence of enzymatic reactions and DNA tagging.[28] This process involves three key steps: (1) initial digestion and tagging: the DNA is first digested using the Hpa II enzyme, which targets unmodified CCGG sites, and then tagged with specific adapters (U-tags). (2) Secondary digestion and tagging: the remaining CCGG sites, which may contain 5mC, or 5hmC, are digested with Msp I enzyme and tagged with H-tags. (3) Glycosylation and final tagging: 5hmC bases are glycosylated to form β-glucosyl-5-hydroxymethyl-cytosine (5gmC), protecting them from further digestion. A second Msp I digestion is then performed, followed by tagging with M-tags for fragments containing 5mC. These steps result in tagged DNA fragments, which can be sequenced to identify the modification states of CCGG sites across the genome. The use of different tags for unmodified cytosine, 5mC, and 5hmC allows for their simultaneous detection and differentiation. However, this method primarily captures CCGG sites, representing only about 10% of all CG sites, thus limiting its scope. Furthermore, the reliance on restriction-enzyme cutting restricts the analysis to only one site per DNA fragment, making it challenging to analyze multiple neighboring 5mC and 5hmC sites within short ranges. This highlights a significant limitation in the method's ability to provide comprehensive epigenetic profiling.

Dyad-seq

Dyad-seq combines enzymatic detection of modified cytosines with traditional nucleobase conversion techniques[29] (Fig. 1). It is designed to quantify all combinations of 5mC and 5hmC at individual CpG dyads. Dyad-seq utilizes the different methods to detect the DNA modifications on different strand. Firstly, it uses two restriction enzymes, MspJI and AbaSI, to digest DNA. MspJI is a unique restriction endonuclease recognized 5mC on the top strand.[30] It distinctively cleaves DNA at a fixed distance from a 5mC site. AbaSI is a specialized modification-dependent restriction endonuclease that targets DNA containing 5hmC and 5gmC on the top strand.[31] Following the digestion of DNA with MspJI or AbaSI, the bottom strand of the fragmented DNA molecules is captured through ligation to a well-designed adapter, which is specific liagted to the bottom strand. This adapter contains a random overhang, a sample barcode, a unique molecule identifier (UMI), and a sequence for PCR amplification. Next, to identify methylated cytosines on the opposing DNA strand, the samples are treated in one of two ways. They are either enzymatically processed with Tet2, β-GT and A3A or they undergo sodium bisulfite treatment, which specifically converts unmodified cytosines into uracil, while methylated cytosines remain read as C (M-M-dyad-seq and H-M-dyad-seq). Moreover, a minor adjustment to the enzymatic conversion reaction, involving β-GT and A3A, specifically allows for the detection of 5hmC on the opposite DNA strand (M-H-dyad-seq and H-H-dyad-seq). Each combination uniquely targets and identifies specific cytosine modifications, facilitating comprehensive analysis. The M-M-dyad-seq method was adopted for single-cell methylome detection and further integrated with single-cell RNA-seq. This method, termed scDyad &T-seq, enables the simultaneous analysis of both the transcriptome and the methylome at the single-cell level. This method offers multiple assays for simultaneously detecting 5mC or 5hmC, and it can simultaneously quantify genome-wide methylation levels and mRNA from the same cells. However, this method, like DARESOME, depends on restriction-enzyme cutting, which limits its ability to detect only one site per DNA fragment. This restriction prevents the analysis of multiple neighboring 5mC and 5hmC sites in close proximity, representing a significant limitation in understanding the complex patterns of these epigenetic marks. Additionally, the presence of natural DNA breaks in the sample can lead to false positive results, presenting a challenge for accurate interpretation of the data.

3. Detection methods based on chemical processing

Innovative chemical reactions play a crucial role in developing new epigenome sequencing tools. This section highlights chemical methods for simultaneously detecting 5mC and 5hmC, emphasizing the importance and potential of these

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

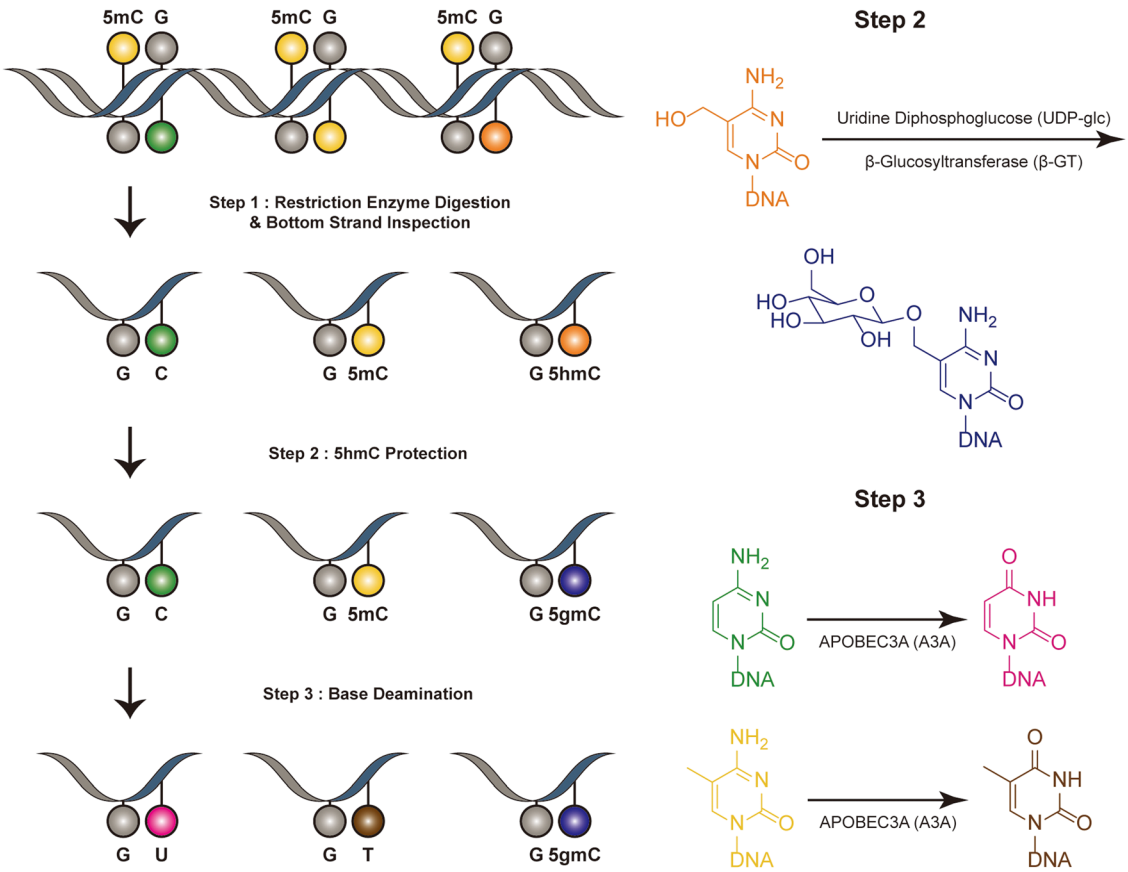

Fig. 1 Overview of dyad-seq. This figure illustrates the schematic and sequencing features of the dyad-seq for combined detection of 5mC and 5hmC.

techniques in advancing our understanding of the epigenome and its implications in various biological processes and diseases.

Six-letter-seq

The six-letter seq workflow, designed to resolve the four genetic bases and the epigenetic modifications 5mC and 5hmC, involves a series of sophisticated steps[32] (Fig. 2). It starts with the fragmentation of DNA, followed by ligation with synthetic DNA hairpin adapters. The strands are separated, and a complimentary copy strand, lacking epigenetic modifications, is synthesized by a synthetic hairpin. A pivotal aspect of the six-letter seq method is its ability to accurately distinguish between 5mC and 5hmC, while simultaneously ensuring precise genetic base calling within the same DNA fragment. DNA methylation modifications are added to the CpG position of newly synthesized DNA strands that do not contain modification information, utilizing DNA methyltransferase 5 (DNMT5) for its specificity in copy methylation.[33] Concurrently, 5hmC is safeguarded from this copying process through glycosylation, executed by beta-glycosyltransferase.[34] This selective glycosylation effectively prevents the replication of 5hmC modifications onto the copy strand, enabling a distinct differentiation between 5mC and 5hmC in the sequencing analysis. Additionally, the 5mC is enzymatically oxidized by a TET2 mutant[35] and subsequently uses β-GT to protect all 5hmCs. Furthermore, the cytosines are deaminated to uracil by APOBEC3A(A3A).[36] Meanwhile, UvrD helicase is added to facilitate the generation of a single-stranded DNA substrate, which is necessary for the efficient action of A3A.[37] The DNA constructs are sequenced using a paired-end format. This involves read 1, primed by P7, representing the original DNA strand, and read 2, primed by P5, representing the synthesized copy strand. These reads are then aligned pairwise, ensuring that read 1 is matched with its complementary read 2. This alignment is critical for accurately reconstructing the sequence and identifying both genetic and epigenetic information from the DNA sample. Computational alignment generates a resolved read, which is then aligned to the reference genome for genetic variant and methylation analysis. This method ensures accurate differentiation between genetic and epigenetic information in DNA analysis. This method can distinguish between four bases and their epigenetic information, using the complementary DNA strand for correction. However, it is not suitable for single-cell detection due to its complex procedural steps. This limitation highlights a trade-off between the method's comprehensive detection capabilities and its practical applicability in more streamlined or single-cell contexts.

Joint-snhmC-seq

The methods used in Joint-snhmC-seq involve a combination of technologies and processes to simultaneously profile 5mC and 5hmC in single cells[38] (Fig. 3). First, lysed individual cells/

 © 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

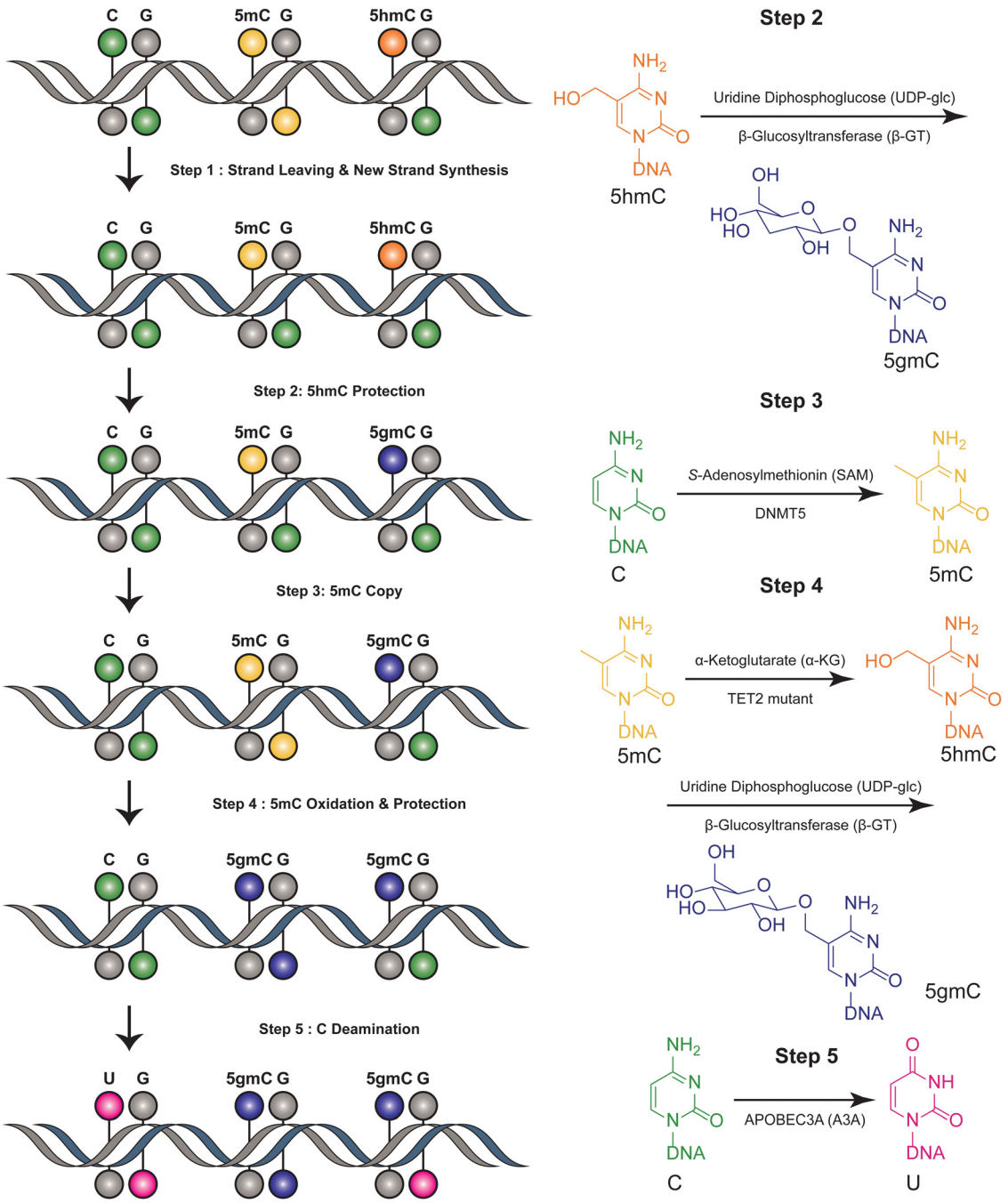

Fig. 2 Overview of six-letter-seq. This figure illustrates the joint profiling of genetics and epigenetics information.

nuclei were treated with bisulfite, which can potentially achieve chemical protection of 5hmC through the formation of cytosine-5-methylenesulfonate (CMS) and can chemically modify cytosine to uracil.[39] The converted single-stranded DNA (ssDNA) is divided into two portions. One part is used directly to detect signals of 5mC and 5hmC(snmC-seq2). The other part undergoes an optimized enzymatic deamination process, typically using A3A. This enzymatic treatment is efficient in converting 5mC to thymine (T), facilitating the differentiation and analysis of these epigenetic marks. After enzymatic deamination using A3A, the treated single-stranded DNA (ssDNA) can be effectively captured through random priming or post-deamination adapter tagging(snhmC-seq2). This step is crucial for enabling low-input bulk or single-cell 5-hydroxymethylome sequencing analysis, enhancing the method's efficiency and applicability for detailed epigenetic studies. The analysis of APOBEC3A-deaminated ssDNA is concurrently conducted by two separate methods: snhmC-seq2 for mapping 5hmC and snmC-seq2 for identifying true 5mC, achieved by subtracting 5hmC signals from the combined 5hmC and 5mC signals. This dual approach,

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

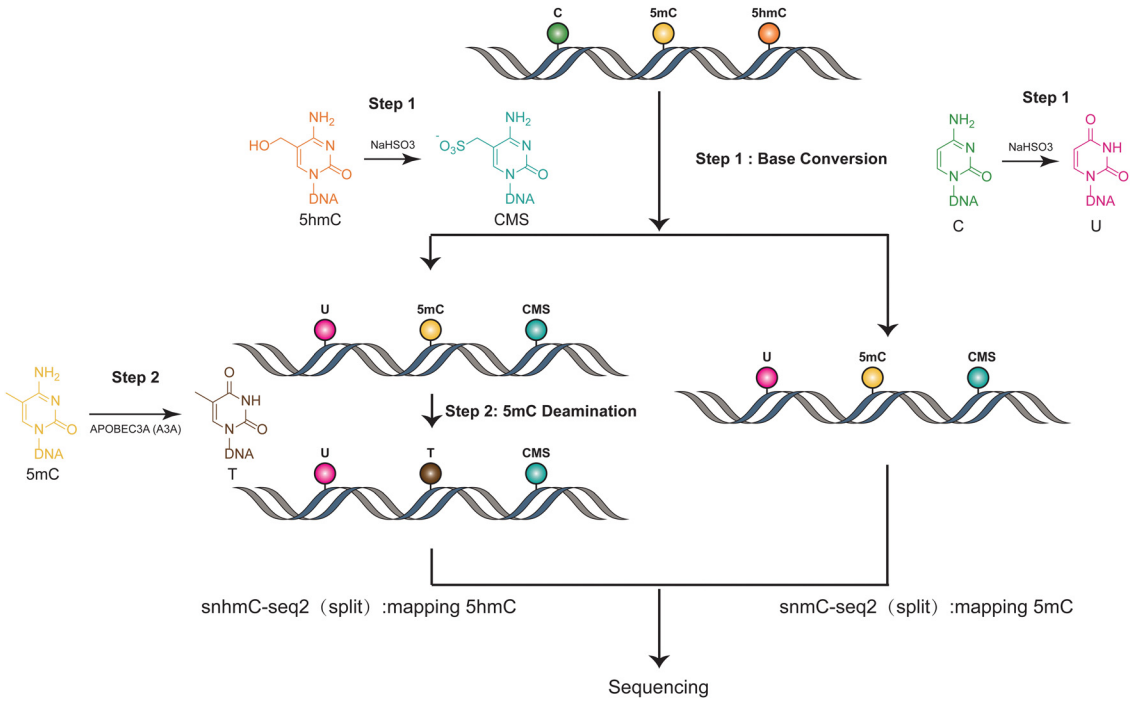

Fig. 3 Overview of Joint-snhmC-seq. The figure demonstrates the concurrent profiling of 5mC and 5hmC at the single-cell level.

by physically linking two distinct epigenetic modalities—5hmC and true 5mC—from the same cell, effectively bypasses the challenge of cross-modal computational integration. This is a significant advancement for reconstructing single-cell 5hmC and 5mC profiles. However, this method's indirect approach in detecting 5mC and 5hmC can lead to false positives if enzyme efficiency is not optimal. Additionally, the conversion of all cytosines to uracil reduces sequence complexity, which may result in a lower mapping ratio. Moreover, this technique necessitates dividing genomic materials into two separate parts for analysis, adding complexity to the process. These factors underscore the need for careful consideration of methodological limitations in epigenetic studies.

SIMPLE-seq

SIMPLE-seq is a high-throughput method for the simultaneous detection of 5mC and 5hmC in single cells.[40] In this process, ruthenate (VI) oxidizes 5hmC to 5fC, which is then labeled by malononitrile[19,41] (Fig. 4). This generates a specific "20C-to-T" signal at 5hmC sites after PCR amplification, while leaving C and 5mC sites unchanged. Primer extension is used to record this "5hmC-to-T" transition on the complementary strand. The subsequent step involves TET-mediated oxidation, which transforms 5mC in the original DNA template into 5caC. This 5caC is then subjected to borane reduction, resulting in the formation of DHU.[15] This process is crucial as it generates a second "C-to-T" signal specifically at the 5mC sites within the same DNA molecule, facilitating the differentiation and identification of these sites. Unmodified cytosines and other bases remain unaffected throughout this process. To differentiate between 5mC and 5hmC, both of which produce "C-to-T" mutations, a specific primer pre-deposited with a 5caC base is designed to record the 5mC signals in the extension products. During the subsequent reaction targeting 5mC, this 5caC base is converted to a "T" signal. This transformation is key in distinguishing between the amplification products derived from 5mC and 5hmC, allowing for accurate identification and analysis of these epigenetic modifications in the DNA template. SIMPLE-seq is scalable and offers base-resolution analysis of both modifications. This allows the identification of the types and locations of modifications from the same DNA molecule in single cells. However, in scenarios where endogenous 5fC is labeled and undergoes a C-to-T transition, it is incorrectly identified as 5hmC, leading to potential false positive results.

4. Discussion

Various studies suggest that increased DNA methylation in promoter regions typically correlates negatively with gene expression.[42,43] However, single-cell multi-omics sequencing indicates that this negative association between promoter methylation levels and gene expression is only evident in a small percentage of cases.[44–46] A potential reason for this discrepancy could be the inability of previous detection methods to differentiate between 5mC and 5hmC. This limitation might lead to mixed signals, thereby obscuring the true relationship between 5mC levels and gene expression. Thus, distinguishing between 5hmC and true 5mC enhances our understanding of gene regulation. Furthermore, the dynamic changes of 5mC and 5hmC during cell fate transition can well explain how these two modifications vary in the process of cell fate change. The dual-modality epigenetic sequencing approach provides clarity

 © 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

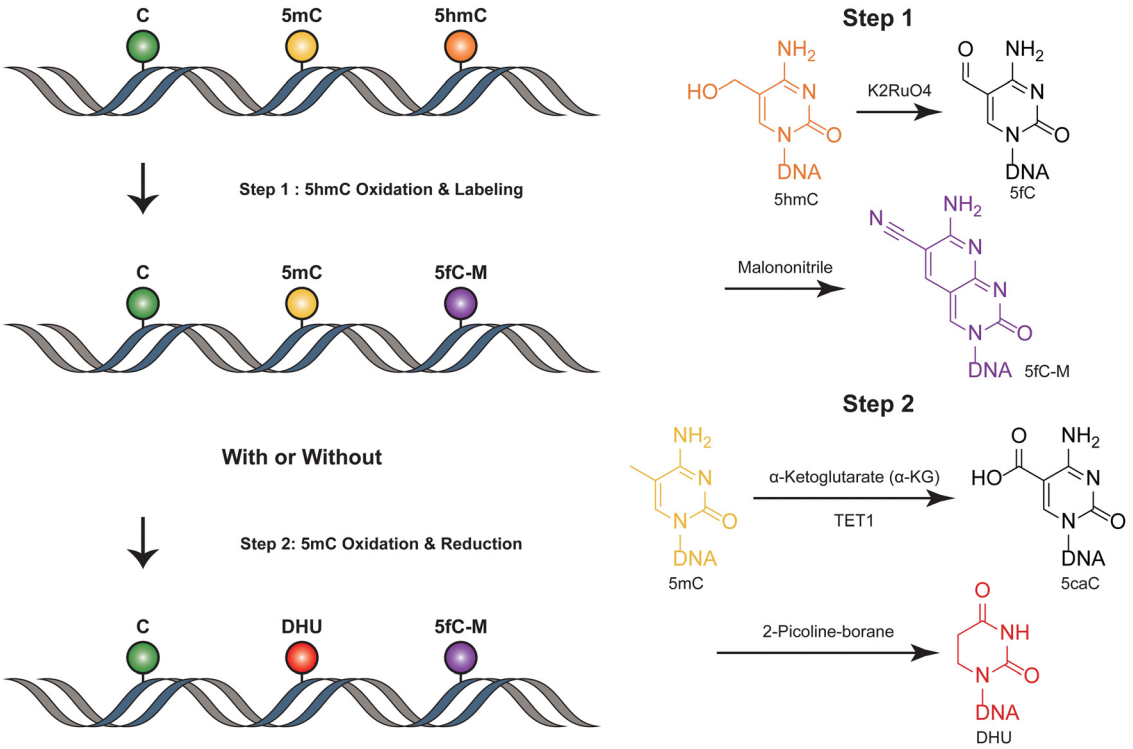

Fig. 4 Overview of SIMPLE-seq. The figure depicts the concurrent profiling of 5mC and 5hmC within a single molecule.

in distinguishing between 5mC and 5hmC, addressing the challenges posed by traditional DNA methylome sequencing methods that often conflate these two modifications. This enhanced resolution in identifying true 5mC and 5hmC profiles significantly improves the accuracy of integrating multimodal data, thereby offering a more precise understanding of epigenetic modifications and their implications. This advancement marks a significant step forward in the field of epigenetics.

Recently, the emergence of spatial omics has significantly enhanced our understanding of disease mechanisms through the spatial mapping of cells within tissues.[47] Nonetheless, the scope of current spatial technologies predominantly focuses on gene expression levels, leaving other aspects like the epigenome less explored. This highlights the necessity for further advancements in spatial omics, aiming to incorporate a broader range of omics data such as epigenomic information, for a more comprehensive and accurate profiling of tissue structures.

Moreover, these methods focusing on 5mC and 5hmC could be integrated with single-molecule multiplexed detection methods for different modifications. With further modification, it has the potential for extensive single-cell multimodal integration, including transcriptome, 3D genome structure, chromatin states, and protein abundances.[48–59] This approach opens up possibilities for comprehensive analysis and understanding of cellular processes at the single-cell level by monitoring various molecular layers simultaneously. Progress in the field of multi-omics will significantly contribute to the development of sophisticated therapeutic strategies. Additionally, it will enable the creation of comprehensive atlases that encompass various omics layers and temporal scales, enhancing our understanding of health and disease.

Furthermore, in recent years, an increasing number of studies have indicated that 5mC and 5hmC can serve as biomarkers for early cancer screening.[60–67] 5mC exhibits significant up-regulation, particularly at certain key oncogenes, while 5hmC experiences a notable decrease in tumors.[66] Accurately distinguishing the signal changes of 5mC and 5hmC can significantly improve the precision of tumor diagnosis. These methods provide a precise and effective technique for future disease detection.

Conflicts of interest

The authors declare that they have no conflicts of interest in this work.

Acknowledgements

We thank Dongsheng Bai and Jinmin Yang for discussion. This study is supported by Beijing Natural Science Foundation (no. Z220013), the National Natural Science Foundation of China (91953201, 22207003, 22107006 and 32200467), and Ministry of Science and Technology of China (2023 YFC3402200). The authors apologize for not being able to cite all the publications related to this topic owing to space constraints of the journal.

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

References

1 A. Bird, DNA methylation patterns and epigenetic memory, *Genes Dev.*, 2002, **16**(1), 6–21.
2 P. A. Jones, Functions of DNA methylation: islands, start sites, gene bodies and beyond, *Nat. Rev. Genet.*, 2012, **13**(7), 484–492.
3 M. Tahiliani, *et al.*, Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1, *Science*, 2009, **324**(5929), 930–935.
4 Y. F. He, *et al.*, Tet-Mediated Formation of 5-Carboxylcytosine and Its Excision by TDG in Mammalian DNA, *Science*, 2011, **333**(6047), 1303–1307.
5 S. Ito, *et al.*, Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine, *Science*, 2011, **333**(6047), 1300–1303.
6 X. Wu and Y. Zhang, TET-mediated active DNA demethylation: mechanism, function and beyond, *Nat. Rev. Genet.*, 2017, **18**(9), 517–534.
7 H. Wu and Y. Zhang, Reversing DNA methylation: mechanisms, genomics, and biological functions, *Cell*, 2014, **156**(1–2), 45–68.
8 T. F. J. Kraus, V. Guibourt and H. A. Kretzschmar, 5-Hydroxymethylcytosine, the "Sixth Base", during brain development and ageing, *J. Neural Transm.*, 2015, **122**(7), 1035–1043.
9 G. P. Pfeifer, S. Kadam and S. G. Jin, 5-hydroxymethylcytosine and its potential roles in development and cancer, *Epigenet. Chromatin*, 2013, **6**, 10–18.
10 B. He, *et al.*, Tissue-specific 5-hydroxymethylcytosine landscape of the human genome, *Nat. Commun.*, 2021, **12**(1), 4249.
11 E. Stoyanova, *et al.*, 5-Hydroxymethylcytosine-mediated active demethylation is required for mammalian neuronal differentiation and function, *eLife*, 2021, **10**, e66973.
12 M. Frommer, *et al.*, A Genomic Sequencing Protocol That Yields a Positive Display of 5-Methylcytosine Residues in Individual DNA Strands, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**(5), 1827–1831.
13 R. Vaisvila, *et al.*, Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA, *Genome Res.*, 2021, **31**(7), 1280–1289.
14 Y. Cao, *et al.*, Single-cell bisulfite-free 5mC and 5hmC sequencing with high sensitivity and scalability, *Proc. Natl. Acad. Sci. U. S. A.*, 2023, **120**(49), e2310367120.
15 Y. B. Liu, *et al.*, Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution, *Nat. Biotechnol.*, 2019, **37**(4), 424–429.
16 M. J. Booth, *et al.*, Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution, *Science*, 2012, **336**(6083), 934–937.
17 M. Yu, *et al.*, Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome, *Cell*, 2012, **149**(6), 1368–1380.
18 E. K. Schutsky, *et al.*, Nondestructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase, *Nat. Biotechnol.*, 2018, 1083–1090.
19 H. Zeng, *et al.*, Bisulfite-Free, Nanoscale Analysis of 5-Hydroxymethylcytosine at Single Base Resolution, *J. Am. Chem. Soc.*, 2018, **140**(41), 13190–13194.
20 Y. Liu, *et al.*, Subtraction-free and bisulfite-free specific sequencing of 5-methylcytosine and its oxidized derivatives at base resolution, *Nat. Commun.*, 2021, **12**(1), 618.
21 N. B. Xie, *et al.*, Whole-Genome Sequencing of 5-Hydroxymethylcytosine at Base Resolution by Bisulfite-Free Single-Step Deamination with Engineered Cytosine Deaminase, *ACS Cent. Sci.*, 2023, **9**(12), 2315–2325.
22 M. V. C. Greenberg and D. Bourc'his, The diverse roles of DNA methylation in mammalian development and disease, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**(10), 590–607.
23 B. M. Colquitt, *et al.*, Alteration of genic 5-hydroxymethylcytosine patterning in olfactory neurons correlates with changes in gene expression and cell identity, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**(36), 14682–14687.
24 H. Wu and Y. Zhang, Charting oxidized methylcytosines at base resolution, *Nat. Struct. Mol. Biol.*, 2015, **22**(9), 1–6.
25 R. E. Amir, *et al.*, Rett syndrome is caused by mutations in X-linked encoding methyl-CpG-binding protein 2, *Nat. Genet.*, 1999, **23**(2), 185–188.
26 B. Kinde, *et al.*, Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**(22), 6800–6806.
27 D. R. Connolly and Z. L. Zhou, Genomic insights into MeCP2 function: A role for the maintenance of chromatin architecture, *Curr. Opin. Neurobiol.*, 2019, **59**, 174–179.
28 R. Viswanathan, *et al.*, DARESOME enables concurrent profiling of multiple DNA modifications with restriction enzymes in single cells and cell-free DNA, *Sci. Adv.*, 2023, **9**(37), eadi0197.
29 A. Chialastri, *et al.*, Combinatorial quantification of 5mC and 5hmC at individual CpG dyads and the transcriptome in single cells reveals modulators of DNA methylation maintenance fidelity, *bioRxiv*, 2023, preprint, DOI: **10.1101/2023.05.06.539708**.
30 D. Cohen-Karni, *et al.*, The MspJI family of modification-dependent restriction endonucleases for epigenetic studies, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**(27), 11040–11045.
31 J. R. Horton, *et al.*, Structure of 5-hydroxymethylcytosine-specific restriction enzyme, AbaSI, in complex with DNA, *Nucleic Acids Res.*, 2014, **42**(12), 7947–7959.
32 J. Fullgrabe, *et al.*, Simultaneous sequencing of genetic and epigenetic bases in DNA, *Nat. Biotechnol.*, 2023, **41**(10), 1457–1464.
33 J. C. Wang, *et al.*, Structural insights into DNMT5-mediated ATP-dependent high-fidelity epigenome maintenance, *Mol. Cell*, 2022, **82**(6), 1186–1198.
34 S. Moréra, *et al.*, T4 phage β-glucosyltransferase:: Substrate binding and proposed catalytic mechanism, *J. Mol. Biol.*, 1999, **292**(3), 717–730.
35 M. Y. Liu, *et al.*, Mutations along a TET2 active site scaffold stall oxidation at 5-hydroxymethylcytosine, *Nat. Chem. Biol.*, 2017, **13**(2), 181–187.
36 E. K. Schutsky, *et al.*, APOBEC3A efficiently deaminates methylated, but not TET-oxidized, cytosine bases in DNA, *Nucleic Acids Res.*, 2017, **45**(13), 7655–7665.

 © 2024 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 05 April 2024. Downloaded on 5/1/2026 5:33:43 PM.
This article is licensed under a Creative Commons Attribution 3.0 Unported Licence.

37 L. Manelyte, *et al.*, The unstructured C-terminal extension of UvrD interacts with UvrB, but is dispensable for nucleotide excision repair, *DNA Repair*, 2009, **8**(11), 1300–1310.
38 E. B. Fabyanic, *et al.*, Joint single-cell profiling resolves 5mC and 5hmC and reveals their distinct gene regulatory effects, *Nat. Biotechnol.*, 2023, DOI: 10.1038/s41587-023-01909-2.
39 Y. Huang, *et al.*, The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing, *PLoS One*, 2010, **5**(1), e8888.
40 D. Bai, *et al.*, Simultaneous single-cell analysis of 5mC and 5hmC with SIMPLE-seq, *Nat. Biotechnol.*, 2024, DOI: 10.1038/s41587-024-02148-9.
41 C. X. Zhu, *et al.*, Single-Cell 5-Formylcytosine Landscapes of Mammalian Early Embryos and ESCs at Single-Base Resolution, *Cell Stem Cell*, 2017, **20**(5), 720–731.
42 A. Meissner, *et al.*, Genome-scale DNA methylation maps of pluripotent and differentiated cells, *Nature*, 2008, **454**(7205), 766–770.
43 T. S. Mikkelsen, *et al.*, Dissecting direct reprogramming through integrative genomic analysis, *Nature*, 2008, **454**(7200), 49-U1.
44 C. Angermueller, *et al.*, Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity, *Nat. Methods*, 2016, **13**(3), 229–232.
45 S. J. Clark, *et al.*, scNMT-seq enables joint profiling of chromatin accessibility DNA methylation and transcription in single cells, *Nat. Commun.*, 2018, **9**, 781.
46 R. Argelaguet, *et al.*, Multi-omics profiling of mouse gastrulation at single-cell resolution, *Nature*, 2019, **576**(7787), 487–491.
47 D. Bressan, G. Battistoni and G. J. Hannon, The dawn of spatial omics, *Science*, 2023, **381**(6657), eabq4964.
48 C. Angermueller, *et al.*, Parallel single-cell sequencing links transcriptional and epigenetic heterogeneity, *Nat. Methods*, 2016, **13**(3), 229–232.
49 M. Stoeckius, *et al.*, Simultaneous epitope and transcriptome measurement in single cells, *Nat. Methods*, 2017, **14**(9), 865–868.
50 J. Cao, *et al.*, Joint profiling of chromatin accessibility and gene expression in thousands of single cells, *Science*, 2018, **361**(6409), 1380–1385.
51 D. S. Lee, *et al.*, Simultaneous profiling of 3D genome structure and DNA methylation in single human cells, *Nat. Methods*, 2019, **16**(10), 999–1006.
52 G. Li, *et al.*, Joint profiling of DNA methylation and chromatin architecture in single cells, *Nat. Methods*, 2019, **16**(10), 991–993.
53 H. Chung, *et al.*, Joint single-cell measurements of nuclear proteins and RNA in vivo, *Nat. Methods*, 2021, **18**(10), 1204–1212.
54 Y. Hao, *et al.*, Integrated analysis of multimodal single-cell data, *Cell*, 2021, **184**(13), 3573–3587 e29.
55 E. P. Mimitou, *et al.*, Scalable, multimodal profiling of chromatin accessibility, gene expression and protein levels in single cells, *Nat. Biotechnol.*, 2021, **39**(10), 1246–1258.
56 C. Zhu, *et al.*, Joint profiling of histone modifications and transcriptome in single cells from mouse brain, *Nat. Methods*, 2021, **18**(3), 283–292.
57 A. F. Chen, *et al.*, NEAT-seq: simultaneous profiling of intranuclear proteins, chromatin accessibility and gene expression in single cells, *Nat. Methods*, 2022, **19**(5), 547–553.
58 C. Luo, *et al.*, Single nucleus multi-omics identifies human cortical cell regulatory genome diversity, *Cell Genomics*, 2022, **2**(3), 100107.
59 B. Zhang, *et al.*, Characterizing cellular heterogeneity in chromatin state with scCUT&Tag-pro, *Nat. Biotechnol.*, 2022, **40**(8), 1220–1230.
60 S. Guo, *et al.*, Identification of methylation haplotype blocks aids in deconvolution of heterogeneous tissue samples and tumor tissue-of-origin mapping from plasma DNA, *Nat. Genet.*, 2017, **49**(4), 635–642.
61 R. H. Xu, *et al.*, Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma, *Nat. Mater.*, 2017, **16**(11), 1155–1161.
62 H. Zeng, *et al.*, Liquid biopsies: DNA methylation analyses in circulating cell-free DNA, *J. Genet. Genomics*, 2018, **45**(4), 185–192.
63 W. Li, *et al.*, 5-Hydroxymethylcytosine signatures in circulating cell-free DNA as diagnostic biomarkers for human cancers, *Cell Res.*, 2017, **27**(10), 1243–1257.
64 C. X. Song, *et al.*, 5-Hydroxymethylcytosine signatures in cell-free DNA provide information about tumor types and stages, *Cell Res.*, 2017, **27**(10), 1231–1242.
65 X. Tian, *et al.*, Circulating tumor DNA 5-hydroxymethylcytosine as a novel diagnostic biomarker for esophageal cancer, *Cell Res.*, 2018, **28**(5), 597–600.
66 P. D. Yousefi, *et al.*, DNA methylation-based predictors of health: applications and statistical considerations, *Nat. Rev. Genet.*, 2022, **23**(6), 369–383.
67 S. Y. Shen, *et al.*, Sensitive tumour detection and classification using plasma cell-free DNA methylomes, *Nature*, 2018, **563**(7732), 579–583.