



# Selective Chemical Labeling and Sequencing of 5-Hydroxymethylcytosine in DNA at Single-Base Resolution

Xiaogang Li<sup>1,2,3</sup>, Xinxin Shi<sup>4</sup>, Yin Gong<sup>5</sup>, Wenting Guo<sup>2</sup>, Yuanrui Liu<sup>6\*</sup>, Chunwei Peng<sup>4\*</sup> and Yingchun Xu<sup>1,3\*</sup>

<sup>1</sup>Department of Clinical Laboratory, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China, <sup>2</sup>Medical Research Center, State Key Laboratory of Complex Severe and Rare Diseases, Peking Union Medical College Hospital, Chinese Academy of Medical Science and Peking Union Medical College, Beijing, China, <sup>3</sup>Beijing Key Laboratory for Mechanisms Research and Precision Diagnosis of Invasive Fungal Diseases, Beijing, China, <sup>4</sup>Gastrointestinal Surgery Department of the First Affiliated Hospital of Anhui Medical University, Hefei, China, <sup>5</sup>School of Electronics Engineering and Computer Science, Peking University, Beijing, China, <sup>6</sup>Characteristic Medical Center of the Chinese People's Armed Police Force, Tianjin, China

## OPEN ACCESS

### Edited by:

Mojgan Rastegar,  
University of Manitoba, Canada

### Reviewed by:

Bi-Feng Yuan,  
Wuhan University, China  
Alexey Fomenkov,  
New England Biolabs, United States

### \*Correspondence:

Yingchun Xu  
xycpumch@139.com  
Chunwei Peng  
pcwayfy@163.com  
Yuanrui Liu  
lbekeyl@qq.com

### Specialty section:

This article was submitted to  
Epigenomics and Epigenetics,  
a section of the journal  
Frontiers in Genetics

Received: 29 July 2021

Accepted: 08 October 2021

Published: 17 November 2021

### Citation:

Li X, Shi X, Gong Y, Guo W, Liu Y,  
Peng C and Xu Y (2021) Selective  
Chemical Labeling and Sequencing of  
5-Hydroxymethylcytosine in DNA at  
Single-Base Resolution.  
Front. Genet. 12:749211.  
doi: 10.3389/fgene.2021.749211

5-Hydroxymethylcytosine (5hmC), the oxidative product of 5-methylcytosine (5mC) catalyzed by ten-eleven translocation enzymes, plays an important role in many biological processes as an epigenetic mediator. Prior studies have shown that 5hmC can be selectively labeled with chemically modified glucose moieties and enriched using click chemistry with biotin affinity approaches. Besides, DNA deaminases of the AID/APOBEC family can discriminate modified 5hmC bases from cytosine (C) or 5mC. Herein, we developed a method based on embryonic stem cell (ESC) whole-genome analysis, which could enrich 5hmC-containing DNA by selective chemical labeling and locate 5hmC sites at single-base resolution with enzyme-based deamination. The combination experimental design is an extension of previous methods, and we hope that this cost-effective single-base resolution 5hmC sequencing method could be used to promote the mechanism and diagnosis research of 5hmC.

**Keywords:** 5-hydroxymethylcytosine, single-base resolution, chemical labeling, cost-effective, AID/APOBEC family

## INTRODUCTION

5-Hydroxymethylcytosine (5hmC), which exists in various mammalian tissues and cell types, is an oxidative product of 5-methylcytosine (5mC) catalyzed by ten-eleven translocation (TET) enzymes (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). Many researches show that 5hmC not only is an intermediate of DNA demethylation but also plays an important role in many biological processes and human diseases as an epigenetic mediator (Szulwach et al., 2011; Mellén et al., 2012; Colquitt et al., 2013; Taylor et al., 2016). The recent development of high-throughput sequencing technology has enabled whole-genome sequencing of 5hmC in mammalian systems. Generally, there are two strategies, which are the selective enrichment-based profiling and deamination-based single-base resolution sequencing methods (Song et al., 2011; Pastor et al., 2012; Tan et al., 2013; Thomson et al., 2013; Cui et al., 2014; Petterson et al., 2014; Sun et al., 2015; Han et al., 2016; Hu et al., 2019; Liu et al., 2019; Gibas et al., 2020; Wang et al., 2021). While the current application of these methods provides key information about the distribution and functional insights of 5hmC, there is major shortage for both strategies. The lack of single-base resolution information of the profiling strategy limited its application in detailed 5hmC

location context, while the single-base resolution method is limited by its high sequencing cost. Therefore, a method that has both advantage of single-base resolution and enrichment would be highly valued in broad biological and clinical studies. A recent study developed a bisulfite-free strategy to detect 5hmC in single-base resolution using AID/APOBEC family DNA deaminase enzyme as deamination reagent, in which 5hmC is protected by a glucose motif catalyzed by T4 beta-glucosyltransferase (T4- $\beta$ GT) (Liu et al., 2019; Vaisvila et al., 2019; Sun et al., 2021a). This strategy has also been reported in a well-studied chemical selective profiling method back to year 2011 using a modified UDP-glucose (Song et al., 2011).

Here, we introduce a cost-effective single-base resolution 5hmC sequencing method, which allows genome-wide chemical labeling and enrichment of 5hmC and a bisulfite-free single-base resolution detection of 5hmC. Our strategy (DIP-CAB-Seq) has three steps: 1) label and enrich the 5hmC containing fragments using the N<sub>3</sub>-UDP-glucose, click chemistry, and biotin-avidin interaction; 2) deaminate unprotected cytosines and 5mC using a deaminase; and 3) construct sequencing library in a single-stranded manner. This strategy is compatible with the labeling chemistry, enzyme-based deamination, and single-stranded DNA library construction. To demonstrate the superiority and effectiveness of this method, we have applied this approach to compare detected 5hmC signals among our new methods (DIP-CAB-Seq), 5hmC-Seal and ACE-Seq (Song et al., 2011; Liu et al., 2019). We found that we can detect reliable single-base 5hmC signals using DIP-CAB-Seq method with limited sequencing depth.

## METHODS

### Cell Culture and DNA Preparation

The v6.5 mouse embryonic stem cells (mESCs) are the same cell line from a previous study (Wray et al., 2011; Guo et al., 2021). Cell culture and genomic DNA was prepared as previously described. Briefly, mESCs were cultured on 0.1% gelatin-coated plates in 2i + leukemia inhibitory factor (2i + LIF) media, which consist of N2B27 (DMEM)/F-12, neurobasal, N-2 supplement, and B27 supplement, supplemented with 1  $\mu$ mol/L of PD0325901 (PZ0162, Sigma-Aldrich, St. Louis, MO, USA), 3  $\mu$ mol/L of CHIR99021 (SML1046, Sigma-Aldrich), and 1,000 U/ml LIF (PMC9484; Gibco, Grand Island, NY, USA). The genomic DNA was extracted by sodium dodecyl sulfate (SDS)/proteinase K digestion, phenol/chloroform extraction, and ethanol precipitation.

### Library Construction

#### Nano-ACE-Seq Library Construction of 5-Hydroxymethylcytosine-Containing Genomic DNA (DIP-CAB-Seq)

The mESC genomic DNA was sheared with KAPA Frag Kit (KK8600, Kapa Biosystems, Wilmington, MA, USA) to average 200 base pair and purified using the manufacturer's protocol (Ring et al., 2017). Fragmented mESC genomic DNA measuring 500 ng (average 200 bp) was treated with  $\beta$ -GT (New England Biolabs, Ipswich, MA, USA; Catalog #M0357S) in the presence of UDP-6-N<sub>3</sub>-Glu and labeled with cyclooctyne-biotin. Subsequently, 5hmC-containing fragments were enriched using Streptavidin beads. The

enriched fragments were mixed with control DNAs, CPG methylated pUC19, and unmethylated lambda and subjected to denaturation with NaOH and enzymatic deamination using APOBEC enzyme (New England Biolabs, Catalog #E7120S) to transform C/mC to U but not hmC. Then the converted DNA was tagged with Illumina compatible adapter and amplified to an appropriate library concentration using Accel-NGS Methyl-Seq DNA Library Kit (Swift BioSciences, Ann Arbor, MI, USA) combined with NEB Next Multiplex Oligos for Illumina. Libraries were checked for quality and quantified using Agarose Gel Electrophoresis and Qubit3.0 individually.

#### 5-Hydroxymethylcytosine-Seal Library Construction of 5-Hydroxymethylcytosine-Containing Genomic DNA

5hmC-Seal Library was prepared as previously described (Song et al., 2011). Briefly, the fragmented mESC genomic DNA (average 200 bp) was treated with UDP-6-N<sub>3</sub>-Glu in the presence of  $\beta$ -GT to form chemical modification, followed by labeling with DBCO-PEG<sub>4</sub>-Biotin via click reaction. The 5hmC-containing DNA fragments were captured by C1 Streptavidin beads. The beads with enriched DNA fragments were resuspended in water and amplified with 12–17 cycles of PCR using an enzyme mixture in the Nextera kit. The PCR products were purified using AMPure XP beads. Sequencing was performed on the NextSeq instrument.

#### ACE-Seq Library Construction of 5-Hydroxymethylcytosine-Containing Genomic DNA

ACE-Seq Library was prepared as previously described (Liu et al., 2019). The mESC genomic DNA was sheared with KAPA Frag Kit (KK8600, Kapa Biosystems, Wilmington, MA, USA) to an average of 200 base pair and purified using the manufacturer's protocol (Ring et al., 2017). DNA fragment was prepared to a concentration of 15–20 ng/ $\mu$ l. The reaction was assembled in a total volume of 50  $\mu$ l using the table below as per reaction. If multiple samples were processed, making a master mix of everything except the sample should be recommended.

The reaction was placed in a thermocycler and incubated at 37°C for 1 h. Subsequent reaction was cleaned separately by DNA Clean & Concentrator-5 Kit (Zymo Research, Irvine, CA, USA) and eluted in 25  $\mu$ l of EB buffer. Glycosylated dsDNA was treated with a fresh 0.1 N NaOH solution and incubated in a thermocycler at 50°C for 10 min for denaturation. After that, the sample was transferred into an ice box and was left to stand at least 5 min to keep the DNA in its single-stranded form. Subsequently, the ssDNA was subjected to enzymatic deamination using APOBEC enzyme to transform C/mC to U but not hmC with NEB Next<sup>®</sup> Enzymatic Methyl-seq Kit (New England Biolabs, Catalog #E7120S) (Liu et al., 2019; Vaisvila et al., 2019; Sun et al., 2021a). Then the converted DNA was tagged with Illumina compatible adapter and amplified to an appropriate library concentration using Accel-NGS<sup>®</sup> Methyl-Seq DNA Library Kit (Swift BioSciences) coupled with NEB Next<sup>®</sup> Multiplex Oligos for Illumina<sup>®</sup> (New England Biolabs). Libraries were checked for quality and quantified using Agarose Gel Electrophoresis and Qubit3.0 individually.

## Sequence Alignment and Peak Identification

Sequencing reads were trimmed by trimmomatic with the trim tail option set true and mapped to the mouse genome (mm10) by bowtie2. The peak analysis was conducted by MACS2 call peaks, under the paired-end mode with bam file as input format. Afterward, the peaks were annotated by HOMER annotate Peak, and mm10 was used as the reference (**Supplementary Table S1**).

## Single-Base Resolution Analysis

The numbers as well as site information of converted and unconverted cytosines in samples were recognized by the software Bismark, with the criterion that the C site cutoff is greater than zero; that is, the conversion was valid as long as any read indicates modification of the site. The process of running Bismark can be split into three steps. First, conduct bisulfite conversion on the genome of interest and build its index to allow bowtie2 alignments. Second, run Bismark alignment step, which means the actual bisulfite alignment and methylation calling part. Finally, extract the methylation information from the Bismark alignment output (**Supplementary Table S2**). Validation of genomic 5hmC at the CH sites was done on *AbaSI*-digested glucosylated ESC genomic DNA followed by real-time PCR (Sun et al., 2013; Sun et al., 2021b). Real-time PCR of digested DNA was done with iQ SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA, USA: 170-8882). All real-time PCR primers and the method are listed in **Supplementary Table S3** for CH sites.

## Definition of Enhancer Subgroups and Motif Analysis

Enhancers with evidence showing their interaction with distal genomic regions were considered as the most active enhancer subgroup (interacting enhancers). The genomic locations of these enhancers were obtained from published ChIA-PET dataset. The active and poised enhancers were defined by using histone modification markers. Both H3K4me1- and H3K27ac-enriched regions were obtained from previous publication. Liftover was used to convert the genomic location from mm8 to mm10. Regions that overlapped with interacting enhancers were discarded. Regions enriched with H3K27ac were considered as active enhancers. Regions only enriched with H3K4me1 but not H3K27ac were considered as poised enhancers. To detect TF motifs around 5hmC-modified cytosines, we performed *de novo* motif analysis with HOMER find Motifs Genome around the 5hmC sites in mESCs.

## RESULTS

### DIP-CAB-Seq Can Generate Deaminated DNA Library From the Pull-Down DNA Fragments

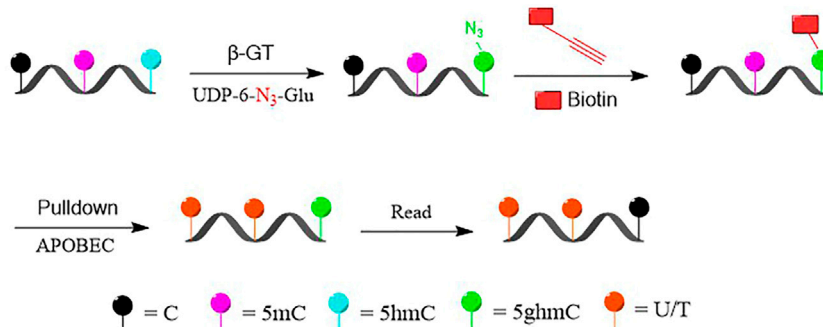
We present here a selective chemical labeling coupled with bisulfite-free sequencing (DIP-CAB-Seq) approach to generate

the genome-wide, single-base resolution maps for 5hmC. In the current approach, we first selectively labeled and enriched 5hmC-containing DNA fragments by using the glucosyltransferase and azido-UDP-glucose-based DNA profiling prior to the subsequent deamination (**Figure 1**). We then performed the deamination on the enriched fragments by using deaminase-based bisulfite-free reaction. The normal cytosine and 5mC can be deaminated, while 5hmC remained as glucosylated 5hmC (N<sub>3</sub>-5gmC). Therefore, 5hmC will be read as C after PCR amplification, while normal C and 5mC will be read as T. After the single-stranded library construction and PCR amplification, the 5hmC containing fragments will be selectively amplified, while the 5hmC can be read in single-base resolution and generate the precise genomic locations of 5hmC.

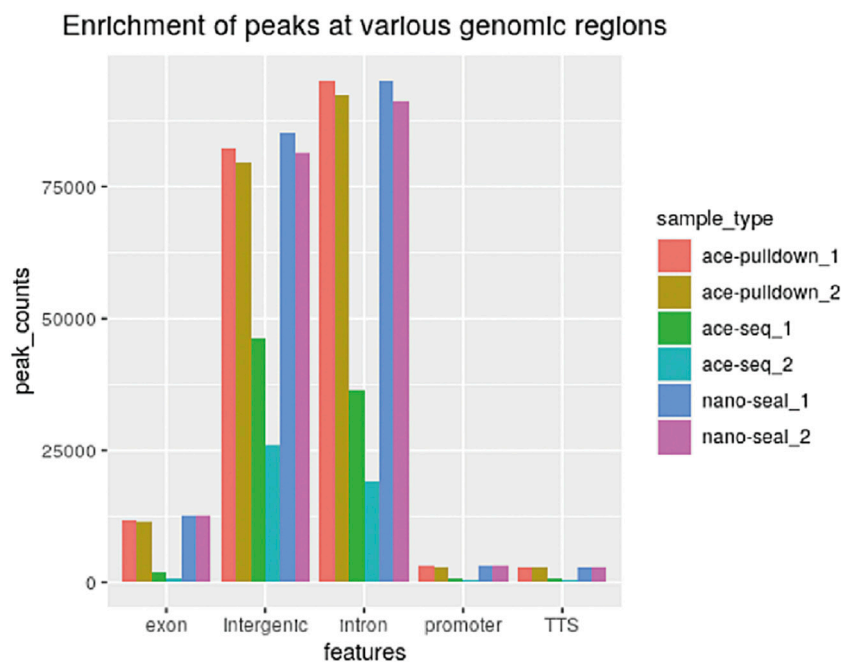
The first challenge is whether we can generate high-quality next-generation sequencing library from the trace amount pull-down DNA. We validated that approach by generating 5hmC library in genomic DNA of mESCs. We constructed ACE-Seq libraries and DIP-CAB-Seq libraries from two replicated mESCs gDNA. The output data size was determined by library input. After getting the DNA library, we examined its concentration used Qubit. The first normalization was conducted through dilution between library's concentration and validated by quantitative PCR. The second normalization was conducted base on the quantitative-PCR data. A library mixture was prepared by pooling the same amount of diluted library. The library mixture was sequenced on Nextseq500 device by following the manufacturer's instruction. We observed 99.89% deamination of control DNA. The ACE-Seq library generated 35,067,558 reads with 68% mapping ratio, while the DIP-CAB-Seq library generated 32,769,029 reads with 61% mapping ratio, not far from the former result. The DIP-CAB-Seq methods have slightly lower mapping ratio than ACE-Seq method, which is fully acceptable considering the ultra-low input DNA amount in the library construction step in DIP-CAB-Seq method. Both DIP-CAB-Seq libraries and regular ACE-Seq libraries have similar total sequencing reads, which guaranteed us a fair comparison between the ability of the two methods in detecting the 5hmC sites in the genome. These data indicate that we can successfully construct the 5hmC single-base resolution library from the enriched DNA fragments.

### DIP-CAB-Seq Can Selectively Enrich 5-Hydroxymethylcytosine-Containing DNA Fragments

Using the DIP-CAB-Seq strategy, we performed both regular hmC-Seal profiling and single-base resolution mapping of 5hmC on mESC genomic DNA. The profiling analysis indicated that 5hmC from both libraries accumulate at the intergenic and intron regions in mESCs (**Figure 2**). This observation is consistent with previous findings (Song et al., 2011). We barely observed any abnormal 5hmC signals, nor did we observe any noticeable changes of 5hmC at these genomic element regions between the two methods. Therefore, the DIP-CAB-Seq can totally inherit



**FIGURE 1** | Schematic diagram of selectively labeling and deaminase-based sequencing of 5-hydroxymethylcytosine (5hmC) in DNA (DIP-CAB-Seq).



**FIGURE 2** | Enrichment of peaks at various genomic regions (ACE-pulldown: DIP-CAB-Seq method).

the profiling information that the traditional hmC-Seal method carried.

### DIP-CAB-Seq Can Detect Single-Base Information With Limited Sequencing Depth

We next analyzed 5hmC enriched regions using the base-resolution data in ACE-Seq and DIP-CAB-Seq so as to evaluate the repeatability of both methods. We observed 5,936,200 5hmC sites in CpG context in the union set of two ACE-Seq samples, while there are only 291,527 5hmC sites, that is, only 4.9% of the union set in the intersection set of both samples. On the other hand, we found 7,821,453 5hmC sites in the CpG context in the union set of the DIP-CAB-Seq samples, and there are 1,273,953 5hmC sites in the intersection set of both samples, accounting for 16.3% of the

union set. These data indicate that the DIP-CAB-Seq method has higher detection ability and consistency than ACE-Seq with the same sequencing depth. Considering that the total sequencing depth is much lower than that of a regular expensive whole-genome 5hmC sequencing, the DIP-CAB-Seq method shows much better application potential in large-scale sample studies. We also detect 5hmC signals in CHH and CHG, which may have 5hmC sites according to the previous study. As expected, DIP-CAB-Seq methods showed higher consistency than ACE-Seq. We observed many CHH and CHG sites from the ACE-Seq library, which may come from the false-positive signal of the ACE-Seq due to the non-perfect deamination. The amount of 5hmC sites in CHH and CHG is shown in the figure below (**Table 1**). Therefore, the enrichment of the 5hmC fragments can also decrease the false-positive signals even if the deamination reagents are not perfect in a bisulfite-free system.

**TABLE 1** | The amount of 5hmC sites in CHH and CHG.

Methods	Sample_num	CpG	CHG	CHH
ACE-Seq	ACE_1	6,342,475	20,163,184	62,652,700
ACE-Seq	ACE_2	1,183,130	3,235,097	11,574,624
DIP-CAB-Seq	DIP-CAB_1	5,427,793	1,418,759	4,081,655
DIP-CAB-Seq	DIP-CAB_2	5,818,168	2,586,082	7,293,592

Note. 5hmC, 5-hydroxymethylcytosine.

We further analyzed the single-base resolution signals from both methods in genomic elements. The following four pie charts (Figure 3) show the distribution of 5hmC sites in samples of ACE-Seq and DIP-CAB-Seq in different feature regions. The proportions are quite different between samples of ACE-Seq, while those of samples of DIP-CAB-Seq are almost the same.

Based on the above findings, we observed much better consistency from DIP-CAB-Seq method than ACE-Seq, indicating the DIP-CAB-Seq method could have better understanding of the biology with limited sequencing depth (Figure 4).

### DIP-CAB-Seq Can Detect Detailed Preferential Occurrences of 5-Hydroxymethylcytosine

With single-base resolution information available, we continued to investigate the detailed preferential occurrences of 5hmC sites

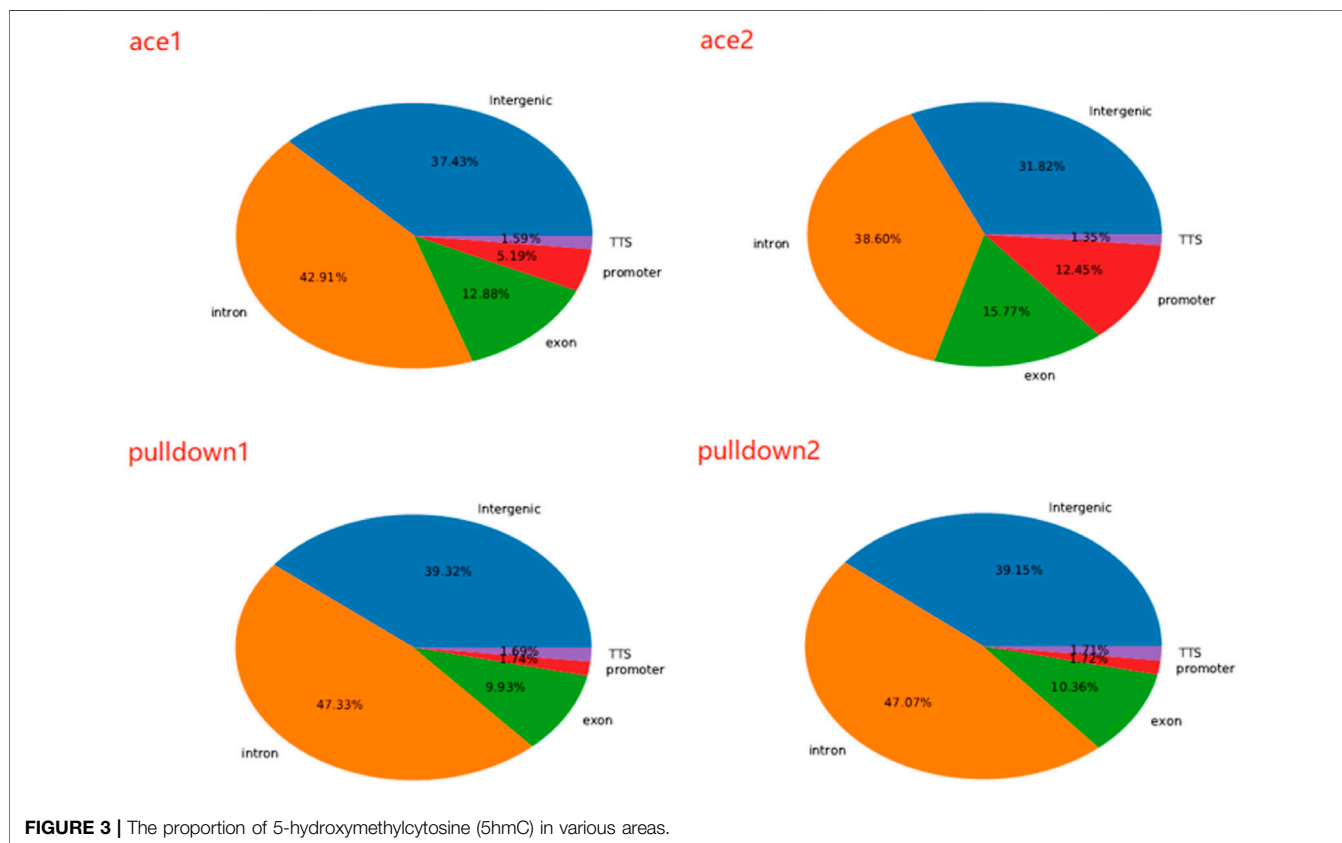
in the genome to understand the potential of this method in biology study.

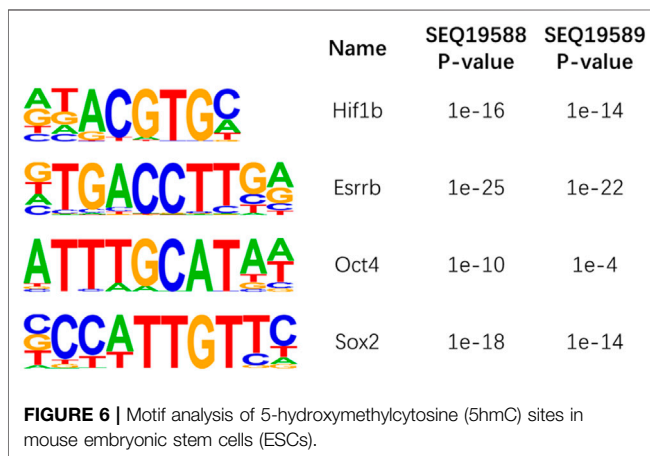
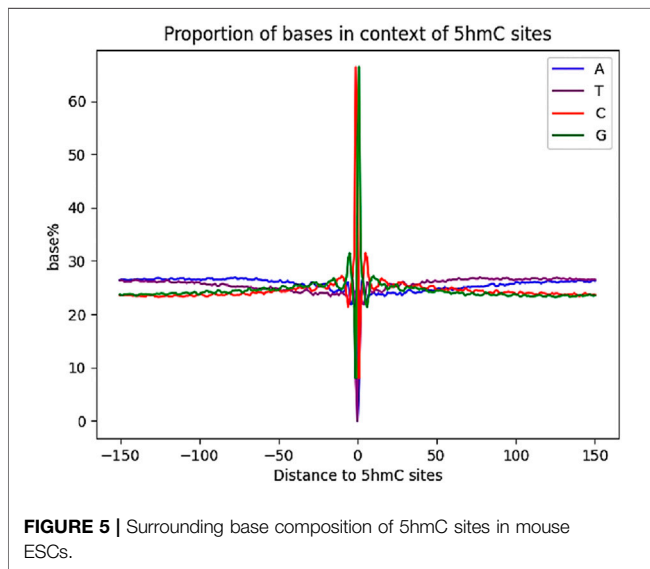
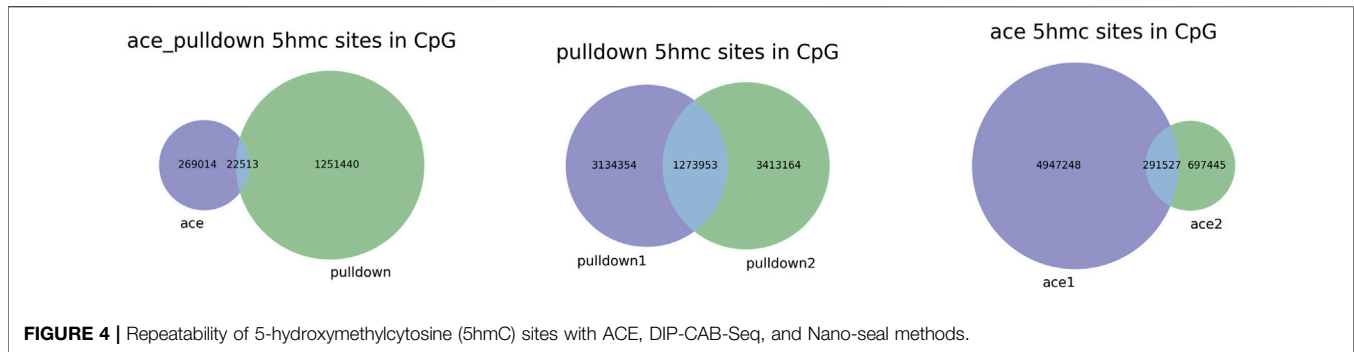
The previous base-resolution mapping of 5hmC allowed for the determination of the surrounding base composition of 5hmC sites in mESCs (Ring et al., 2017). We aligned our 5hmC sites in CG context and examined the base compositions (Figure 5). Compared with reported results, our data possess a similar local sequence context, having increased guanine abundance with a depletion of thymine.

We then performed motif analysis with HOMER32 at  $\pm 100$ -bp region around the 5hmC sites. We studied the well-known 5hmC containing motif Klf4, Oct4, Hif1a, Esrrb, and Sox2; our motif analysis successfully identified these motifs around the 5hmC sites in mESCs (Figure 6). Our data indicated that our method can obtain these detailed 5hmC signals with very low sequencing depth as compared with the reported method.

## DISCUSSION

The current dilemma between cost and sequencing resolution has hampered many biological and clinical studies of 5hmC. Generally, 5hmC has 10 times lower genome percentage than 5mC, and 5hmC is distributed broadly across the genome, which requires very deep sequencing for detecting low abundant 5hmC sites. Although the profiling methods provide a cheap manner to map the 5hmC peak, it is much more difficult to quantify the 5hmC signals than the single-





base resolution method. The DIP-CAB-Seq method we developed can solve this problem; although we cannot quantify the 5hmC abundance for a specific 5hmC site, it provides much better resolution than the regular profiling method. We showed that the base-resolution data from the very limited sequencing depth can provide detailed 5hmC loci information. It is well known that 5hmC

profiling has the potential to be a clinical tool in different clinical questions. The current profiling can only calculate the 5hmC score based on 5hmC read density, which needs much adjustment between the samples and cohorts. To further study the role of 5hmC in clinical application, our method can provide better quantification information and could be more reliable than the profiling method.

## CONCLUSION

Herein, we developed a method based on ESC whole-genome analysis, which could enrich 5hmC-containing DNA by selective chemical labeling and locate 5hmC sites at single-base resolution with enzyme-based deamination. The combination experimental design is an extension of previous methods, and we hope that this cost-effective single-base resolution 5hmC sequencing method could be used to promote the mechanism and diagnostic research of 5hmC.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. These data can be found here: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA732757.

## AUTHOR CONTRIBUTIONS

I) Conception and design: XL. II) Administrative support: YX. III) Provision of study materials: WG and XL. IV) Collection and assembly of data: XL, XS, and YG. V) Data analysis and interpretation: XL and XS. VI) Article writing: all authors. VII) Final approval of article: all authors.

## FUNDING

This research was supported by the Beijing Natural Science Foundation (7194312 to XL), Beijing Key Clinical Specialty for Laboratory Medicine-Excellent Project (No. ZK201000),

National Key Research and Development Program of China (2017YFC1601502) and Beijing Municipal Science and Technology Project (Z181100001618015), and Beijing Key Laboratory for Mechanisms Research and Precision Diagnosis of Invasive Fungal Diseases.

## REFERENCES

- Colquitt, B. M., Allen, W. E., Barnea, G., and Lomvardas, S. (2013). Alteration of Genic 5-hydroxymethylcytosine Patterning in Olfactory Neurons Correlates with Changes in Gene Expression and Cell Identity. *Proc. Natl. Acad. Sci. USA* 110, 14682–14687. doi:10.1073/pnas.1302759110
- Cui, L., Chung, T. H., Tan, D., Sun, X., and Jia, X.-Y. (2014). JBP1-seq: a Fast and Efficient Method for Genome-wide Profiling of 5hmC. *Genomics* 104, 368–375. doi:10.1016/j.ygeno.2014.08.023
- Gibas, P., Narmontė, M., Staševskij, Z., Gordevičius, J., Klimašauskas, S., and Kriukienė, E. (2020). Precise Genomic Mapping of 5-hydroxymethylcytosine via Covalent Tether-Directed Sequencing. *Plos Biol.* 18 (4), e3000684. doi:10.1371/journal.pbio.3000684
- Guo, W., Wang, S., Zhang, X., Shi, M., Duan, F., Hao, J., et al. (2021). Acidic pH Transiently Prevents the Silencing of Self-Renewal and Dampens MicroRNA Function in Embryonic Stem Cells. *Sci. Bull.* 66, 1319–1329. doi:10.1016/j.scib.2021.03.005
- Han, D., Lu, X., Shih, A. H., Nie, J., You, Q., Xu, M. M., et al. (2016). A Highly Sensitive and Robust Method for Genome-wide 5hmC Profiling of Rare Cell Populations. *Mol. Cell* 63, 711–719. doi:10.1016/j.molcel.2016.06.028
- Hu, L., Liu, Y., Han, S., Yang, L., Cui, X., Gao, Y., et al. (2019). Jump-seq: Genome-wide Capture and Amplification of 5-Hydroxymethylcytosine Sites. *J. Am. Chem. Soc.* 141 (22), 8694–8697. doi:10.1021/jacs.9b02512
- Kriaucionis, S., and Heintz, N. (2009). The Nuclear DNA Base 5-hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* 324, 929–930. doi:10.1126/science.1169786
- Liu, Y., Szejka-Zielińska, P., Velikova, G., Bi, Y., Yuan, F., Tomkova, M., et al. (2019). Bisulfite-free Direct Detection of 5-methylcytosine and 5-hydroxymethylcytosine at Base Resolution. *Nat. Biotechnol.* 37 (4), 424–429. doi:10.1038/s41587-019-0041-2
- Mellén, M., Ayata, P., Dewell, S., Kriaucionis, S., and Heintz, N. (2012). MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell* 151, 1417–1430. doi:10.1016/j.cell.2012.11.022
- Pastor, W. A., Huang, Y., Henderson, H. R., Agarwal, S., and Rao, A. (2012). The GLIB Technique for Genome-wide Mapping of 5-hydroxymethylcytosine. *Nat. Protoc.* 7, 1909–1917. doi:10.1038/nprot.2012.104
- Petterson, A., Chung, T. H., Tan, D., Sun, X., and Jia, X.-Y. (2014). RRHP: a Tag-Based Approach for 5-hydroxymethylcytosine Mapping at Single-Site Resolution. *Genome Biol.* 15, 456. doi:10.1186/s13059-014-0456-5
- Ring, J. D., Sturk-Andreaggi, K., Peck, M. A., and Marshall, C. (2017). A Performance Evaluation of Nextera XT and KAPA HyperPlus for Rapid Illumina Library Preparation of Long-Range Mitogenome Amplicons. *Forensic Sci. Int. Genet.* 29, 174–180. doi:10.1016/j.fsigen.2017.04.003
- Song, C.-X., Szulwach, K. E., Fu, Y., Dai, Q., Yi, C., Li, X., et al. (2011). Selective Chemical Labeling Reveals the Genome-wide Distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* 29, 68–72. doi:10.1038/nbt.1732
- Sun, Z., Dai, N., Borgaro, J. G., Quimby, A., Sun, D., Corrèa, I. R., et al. (2015). A Sensitive Approach to Map Genome-wide 5-hydroxymethylcytosine and 5-formylcytosine at Single-Base Resolution. *Mol. Cell* 57, 750–761. doi:10.1016/j.molcel.2014.12.035
- Sun, Z., Terragni, J., Borgaro, J. G., Liu, Y., Yu, L., Guan, S., et al. (2013). High-resolution Enzymatic Mapping of Genomic 5-hydroxymethylcytosine in Mouse Embryonic Stem Cells. *Cel Rep.* 3 (2), 567–576. doi:10.1016/j.celrep.2013.01.001
- Sun, Z., Terragni, J., Zhu, Z., Zheng, Y., and Pradhan, S. (2021). High-Resolution Enzymatic Mapping of Genomic 5-Hydroxymethylcytosine. *Methods Mol. Biol.* 2272, 13–27. doi:10.1007/978-1-0716-1294-1\_2

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.749211/full#supplementary-material>

- Sun, Z., Vaisvila, R., Hussong, L.-M., Yan, B., Baum, C., Saleh, L., et al. (2021). Nondestructive Enzymatic Deamination Enables Single-Molecule Long-Read Amplicon Sequencing for the Determination of 5-methylcytosine and 5-hydroxymethylcytosine at Single-Base Resolution. *Genome Res.* 31 (2), 291–300. doi:10.1101/gr.265306.120
- Szulwach, K. E., Li, X., Li, Y., Song, C.-X., Wu, H., Dai, Q., et al. (2011). 5-hmC-mediated Epigenetic Dynamics during Postnatal Neurodevelopment and Aging. *Nat. Neurosci.* 14, 1607–1616. doi:10.1038/nn.2959
- Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in Mammalian DNA by MLL Partner TET1. *Science* 324, 930–935. doi:10.1126/science.1170116
- Tan, L., Xiong, L., Xu, W., Wu, F., Huang, N., Xu, Y., et al. (2013). Genome-wide Comparison of DNA Hydroxymethylation in Mouse Embryonic Stem Cells and Neural Progenitor Cells by a New Comparative hMeDIP-Seq Method. *Nucleic Acids Res.* 41, e84. doi:10.1093/nar/gkt091
- Taylor, S. E., Li, Y. H., Smeriglio, P., Rath, M., Wong, W. H., and Bhutani, N. (2016). Stable 5-Hydroxymethylcytosine (5hmC) Acquisition Marks Gene Activation during Chondrogenic Differentiation. *J. Bone Miner. Res.* 31, 524–534. doi:10.1002/jbmr.2711
- Thomson, J. P., Hunter, J. M., Nestor, C. E., Dunican, D. S., Terranova, R., Moggs, J. G., et al. (2013). Comparative Analysis of Affinity-Based 5-hydroxymethylation Enrichment Techniques. *Nucleic Acids Res. Techniques. Nucleic Acids Res.* 41, e206. doi:10.1093/nar/gkt1080
- Vaisvila, R., Ponnaluri, V. K. C., Sun, Z., Langhorst, B. W., Saleh, L., Guan, S., et al. (2019). EM-seq: Detection of DNA Methylation at Single Base Resolution from Picograms of DNA. *bioRxiv* 12, doi:10.1101/2019.12.20.884692
- Wang, T., Loo, C. E., and Kohli, R. M. (2021). Enzymatic Approaches for Profiling Cytosine Methylation and Hydroxymethylation. *Mol. Metab.* 8, 101314. doi:10.1016/j.molmet.2021.101314
- Wray, J., Kalkan, T., Gomez-Lopez, S., Eckardt, D., Cook, A., Kemler, R., et al. (2011). Inhibition of Glycogen Synthase Kinase-3 Alleviates Tcf3 Repression of the Pluripotency Network and Increases Embryonic Stem Cell Resistance to Differentiation. *Nat. Cell Biol.* 13, 838–845. doi:10.1038/ncb2267

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Li, Shi, Gong, Guo, Liu, Peng and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.