JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

Differentiated Visualization of Single-cell 5-Hydroxymethylpyrimidines with Microfluidic Hydrogel Encoding

Feng Chen, Jing Xue, Jin Zhang, Min Bai, Xu Yu, Chunhai Fan, and Yongxi Zhao

J. Am. Chem. Soc., **Just Accepted Manuscript** • DOI: 10.1021/jacs.9b11393 • Publication Date (Web): 27 Jan 2020

Differentiated Visualization of Single-cell 5-Hydroxymethylpyrimidines with Microfluidic Hydrogel Encoding

Feng Chen[‡,†], Jing Xue[‡,†], Jin Zhang[†], Min Bai[†], Xu Yu[†], Chunhai Fan[#] and Yongxi Zhao[†,*]

[†]Institute of Analytical Chemistry and Instrument for Life Science, The Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xianning West Road, Xi'an, Shaanxi 710049, P. R. China

[#]Institute of Molecular Medicine, Renji Hospital, School of Medicine and School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Shanghai 200127, P. R. China

ABSTRACT: 5-Hydroxymethyluracil (5hmU) is found in the genomes of a diverse range of organisms as another kind of 5-hydroxymethylpyrimidines except 5-hydroxymethylcytosine (5hmC). The biological function of 5hmU have not been well explored due to lacking both specific 5hmU recognition and single-cell analysis methods. Here we report differentiated visualization of single-cell 5hmU and 5hmC with microfluidic hydrogel encoding (sc5hmU/5hmC-microgel). Single cells and their genomic DNA after cell lysis can be encapsulated in individual agarose microgels. The 5hmU sites are then specifically labeled with thiophosphate for the first time, followed by labeling 5hmC with azide glucose. These labeled bases are each encoded into corresponding DNA barcode primers by chemical crosslinking. *In situ* amplification is triggered for single-molecule fluorescence visualization of single-cell 5hmU and 5hmC. Based on sc5hmU/5hmC-microgel, we reveal cell type-specific molecular signatures of these two bases with remarkable single-cell heterogeneity. Utilizing machine learning algorithms to decode four-dimensional signatures of 5hmU/5hmC, we visualize the discrimination of non-tumorigenic, carcinoma and highly-invasive breast cell lines. This strategy provides a new route to analyze and decode single-cell DNA epigenetic modifications.

INTRODUCTION

Genomes contain chemically modified DNA bases besides the four canonical ones. These modified bases can be generated by endogenous enzymes or exogenous factors.[1] They have the potential to profoundly influence genome function and cellular processes.[1,2] Many modified DNA bases have been discovered. The best-known ones in mammalian genomes are 5-methylcytosine (5mC, methylation of cytosine at the 5-position) and its oxidized derivatives 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxycytosine (5caC). These epigenetic marks have been well demonstrated to play important roles in regulating gene expression.[3-8] Yet, thymine modifications and their biological function remain largely elusive. 5-hydroxymethyluracil (5hmU) is a thymine derivative identified in the genomes of diverse organisms. Direct incorporation of 5hmU into DNA during the replication is well known in bacteriophage. In mammals, 5hmU is formed through post-replication processing mechanisms including thymine hydroxylation by ten-eleven translocation (TET) enzymes or reactive oxygen species (ROS) and 5hmC deamination.[9-11] The majority of 5hmU in mES cells is generated by TET enzymes and very little through 5hmC deamination or ROS pathway.[9] Thus matched 5hmU:A rather than mismatched 5hmU:G is the major existing form of 5hmU in mammalian genome. The levels of modified bases in genomic DNA vary by many factors such as cell or tissue type and disease state of an organism. Unfortunately, 5hmU has been found to comprise very little faction of all thymine residues, for example ppm level in mES cells.[9,12,13] It challenges the reliable analysis of hmU especially in single cells, and the structural similarity between 5hmU and 5hmC also impedes the discrimination of these modifications.

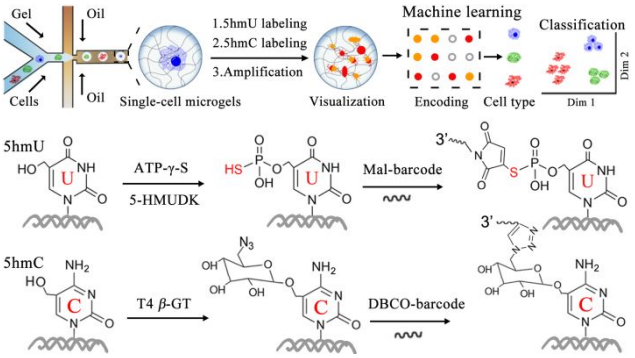
Recently, two chemical methods have been reported to detect genomic 5hmU via transforming 5hmU to 5fU by chemical oxidation. One of both induced T-to-C base change by forming 5fU:G base pairing during the polymerase extension reaction.[14] Yet, the proportion of such base change is lower than 40% even under optimized conditions, and it is not able to label 5hmU with functional tags. The other method labeled the resulting 5fU with biotin by (+)-biotinamideohexanoic acid hydrazide, followed by enrichment.[2] It achieved genome-wide analysis of 5hmU in 5fU-absent eukaryote parasite *Leishmania* with submicrogram DNA input (equivalent to genome of millions of cells). Nevertheless, this hydrazine reagent also presents high reactivity to other aldehyde bearing components existing in genomic

DNA, especially 5fC and abasic sites. Besides chemical method, enzymatic recognition provides another choice to resolve modified DNA bases. We have demonstrated human single-strand selective monofunctional uracil DNA glycosylase (hSMUG1) to excise and release uracil from damage DNA in living cells.[15,16] As we known, hSMUG1 also catalyzes the hydrolysis of 5hmU and 5fU substrates. We may recognize 5hmU in 5fU-absent genomic DNA prior to excising uracil by E. coli UDG,[17] yet it is not appropriate for the mammalian genomes. T4 phage β-glucosyltransferase (β-GT) have been used to label and analyze 5hmC in genomic DNA.[3,7,8,18] It transfers an azide-glucose moiety to 5hmC:G in double-stranded (ds) DNA. As 5hmU and 5hmC are structurally similar, β-GT has been found to label 5hmU residues at mismatched 5hmU:G sites.[19] However, the matched 5hmU:A is not an active substrate for β-GT. In addition, these existing methods required large amounts of DNA materials. Therefore, the specific 5hmU recognition and single-cell analysis of 5hmU/5hmC remain challenging.

Droplet microfluidics offers a promising platform for the manipulation and analysis of single cells. It uses immiscible multiphase flows to generate monodisperse droplets as individual compartmentalized microreactors that contain single cells and reagents.[20-24] Thousands of droplets can be formed within one second, which enables high-throughput single-cell analysis to reveal cell heterogeneity of large populations.[25,26] Due to these intrinsic superiorities, various droplet systems have been reported for the capture, culture, sorting or/and detection of single cells.[27-29] For example, single-cell or single-molecule DNA and RNA analysis have been well achieved by introducing DNA amplification in microfluidic droplets or hydrogel droplets.[29-34] The reaction in picoliter droplets significantly improved the amplification and analysis performance.[35,36] Unfortunately, these existing strategies are not capable of detecting modified bases in single cells. Developing microfluidic single-cell 5hmU/5hmC specific analysis method is an urgent demand.

Herein, we report single-cell 5hmU and 5hmC differentiated visualization in microfluidic hydrogel droplets, termed sc5hmU/5hmC-microgel. We report the specific 5hmU recognition and labeling against its analogues including 5hmC, 5fU and 5fC for the first time. In brief, 5hmU in dsDNA is labeled with a thiophosphate group as 5-thiophosphomethyluracil (5psmU) by 5hmU DNA kinase (5-HMUDK) from Pseudomonas aeruginosa bacteriophage M6. We demonstrate adenosine-5-O-thiotriphosphate (ATP-γ-S) can be utilized as the phosphate donor by 5-HMUDK. The sulfhydryl group of 5psmU allows routine chemical crosslinking for downstream applications. In sc5hmU/5hmC-microgel (Scheme 1), individual cells are encapsulated in picoliter droplets containing low gelling temperature agarose. After microgel formation, emulsion break and cell lysis, the genomic DNA has been still trapped in the microgels, providing a compartmentalized environment for single-cell analysis. Subsequently, 5hmU and 5hmC are successively labeled with thiophosphate and azide glucose. These labeled bases are converted into 5'

Scheme 1. Design and workflow of sc5hmU/5hmC-microgel visualizing single-cell 5hmU and 5hmC.



maleimide or dibenzocyclooctyne (DBCO) modified DNA barcode primers by chemical crosslinking. *In situ* amplification and hybridization of fluorescence probes are then performed for single-molecule visualization of these bases. Their molecular signatures in microgels are encoded into the information of fluorescence intensity and spot count. We further reveal cell type-specific abundance, number and even distribution feature of 5hmU and 5hmC. We also visualize the classification of non-tumorigenic, carcinoma and highly-invasive breast cell lines by using machine learning algorithms to decode four-dimensional signatures of these two bases for the first time.

RESULTS

Specific 5hmU recognition and labeling. We first demonstrated T4 β-GT efficiently transferred azide glucose groups to 5hmU:G instead of 5hmU:A in model dsDNA. The dsDNA samples were first degraded to nucleosides by nuclease and phosphatase after labeling reaction. As shown in Figure 1A and S1, negligible glycosylated 5hmU nucleoside is detected in the sample of 5hmU:A-containing dsDNA by liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS). In contrast, remarkable glycosylated nucleosides were observed in both 5hmU:G and 5hmC:G samples. T4 β-GT recognizes mismatched 5hmU:G rather than matched 5hmU:A probably depending on the binding of G base. Then we investigated whether 5-HMUDK utilized ATP-γ-S to thiophosphorylate 5hmU residue in dsDNA. The dsDNA oligos containing recognition sequence (5'-CCAXGG-3', X=T, U, hmU and fU) for restrict endonuclease NcoI is prepared as the substrate for 5-HMUDK. NcoI can cleave this dsDNA with X:A rather than 5psmU:A in the recognition sequence (Figure 1B and S2). DNA melting curve is used to detect dsDNA break, and the 5hmU:A sample with ATP-γ-S treatment blocks DNA break catalyzed by NcoI. Thus ATP-γ-S is verified to thiophosphorylate 5hmU:A by 5-HMUDK.

Subsequently, the recognition of 5hmU:G and discrimination of 5hmU from analogues such as 5hmC by 5-HMUDK are studied with ATP-γ-S as the phosphate

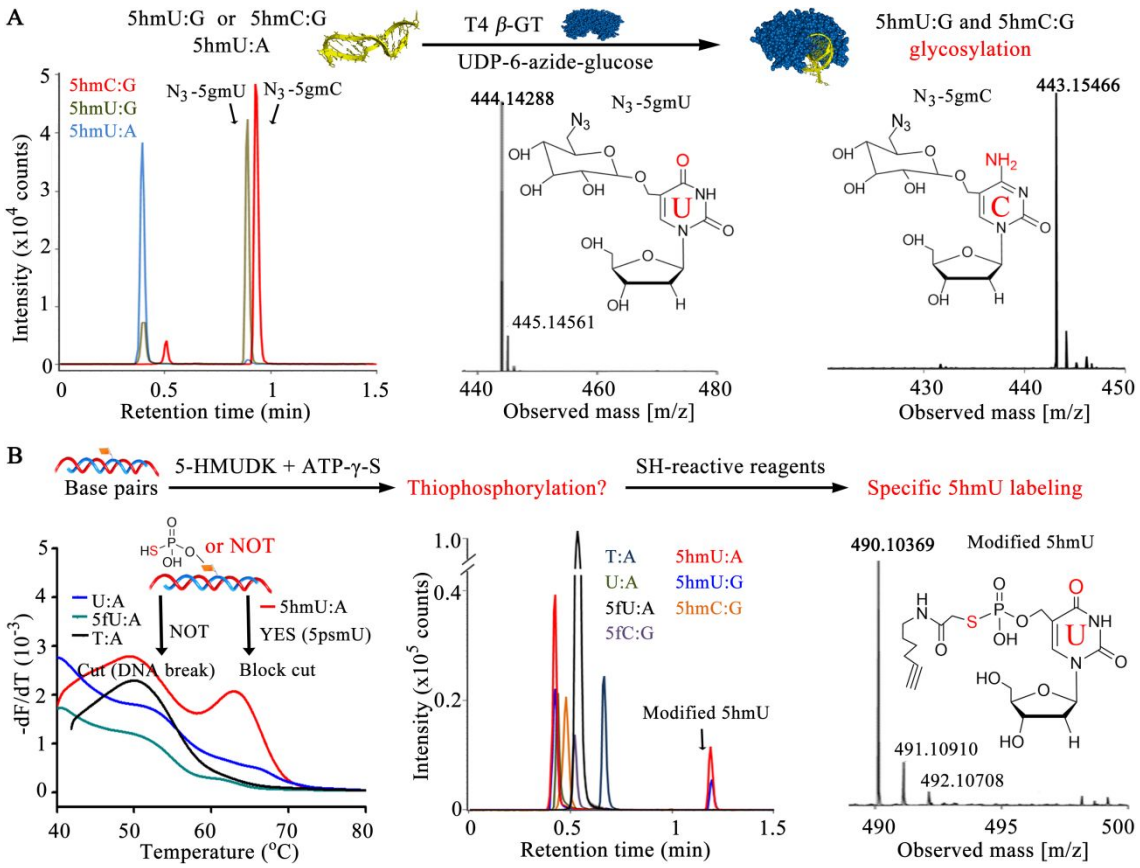


Figure 1. Specific recognition and labeling of 5hmU by 5-HMUDK and SH-reactive reagents. (A) LC-MS characterization of the nucleoside products in labeling reactions with model DNA by T4 β-GT. Left panel, chromatograms of nucleosides; middle panel, the mass spectrum of N_3-5gmU (calculated mass of $[C_{16}H_{22}N_5O_{10}]^-$ = 444.13722, found: 444.14288); right panel, the mass spectrum of N_3-5gmC (calculated mass of $[C_{16}H_{23}N_6O_9]^-$ = 443.15320, found: 443.15466). The complex of purple protein and yellow DNA indicates DNA-bound T4 β-GT. (B) Characterization of the products in labeling reactions with model DNA by 5-HMUDK and SH-reactive reagents. Left panel, melting curves of DNA samples cut by restrict endonuclease NcoI; middle panel, chromatograms of nucleosides; right panel, the mass spectrum of modified 5hmU (calculated mass of $[C_{18}H_{25}N_3O_9PS]^-$ = 490.10546, found: 490.10369).

donor. We design seven dsDNA substrates each containing different base pair (5hmU:A, 5hmU:G, 5hmC:G, 5fC:G, 5fU:A, U:A or T:A) site. After the thiophosphorylation reaction, a sulfhydryl-reactive iodoacetyl reagent is added to react with the sulfhydryl of thiophosphate groups. After DNA purification, the samples were treated with nuclease and phosphatase. The resulting nucleosides were detected by LC-MS/MS. Both 5hmU:A and 5hmU:G were thiophosphorylated and further labeled with the iodoacetyl reagent, while others are not the active substrate for 5-HMUDK (Figure 1B, S3 and S4). Notably, labeling 5psmU with iodoacetyl reagents here is to prevent 5psmU dephosphorylation by phosphatase before LC-MS/MS analysis. Other routine sulfhydryl-reactive reagents such as maleimides can also react with 5psmU. Therefore, we demonstrated the specific recognition and labeling of 5hmU against many analogues such as 5hmC and 5fU for the first time. It settles the differentiated analysis of 5-hydroxymethylpyrimidines (5hmU and 5hmC) in the same genomic DNA sample and even single cells.

Single-cell 5hmU visualization in microfluidic hydrogel droplets. To obtain single cells and label the 5hmU for amplified analysis, they should be physically compartmentalized and isolated. Here we used a two stream co-flow droplet chip to encapsulate single cells in agarose microgels. In this device, the cell suspension stream is merged with the molten agarose stream. Under our condition, about 97% of positive droplets contain individual cell (Figure 2A and Figure S5), though more than 80% of droplets were negative (empty, no cells). The cell loading distribution into droplets or microgels is well consistent with the Poisson distribution. After agarose cooling and gelling, the microgels in oil phase were transferred to water phase. After cell lysis and washing, we then stained the microgels with SYBR green dye, and verified the genomic DNA remained encased in the microgels after several times of washing and purification in 4 days

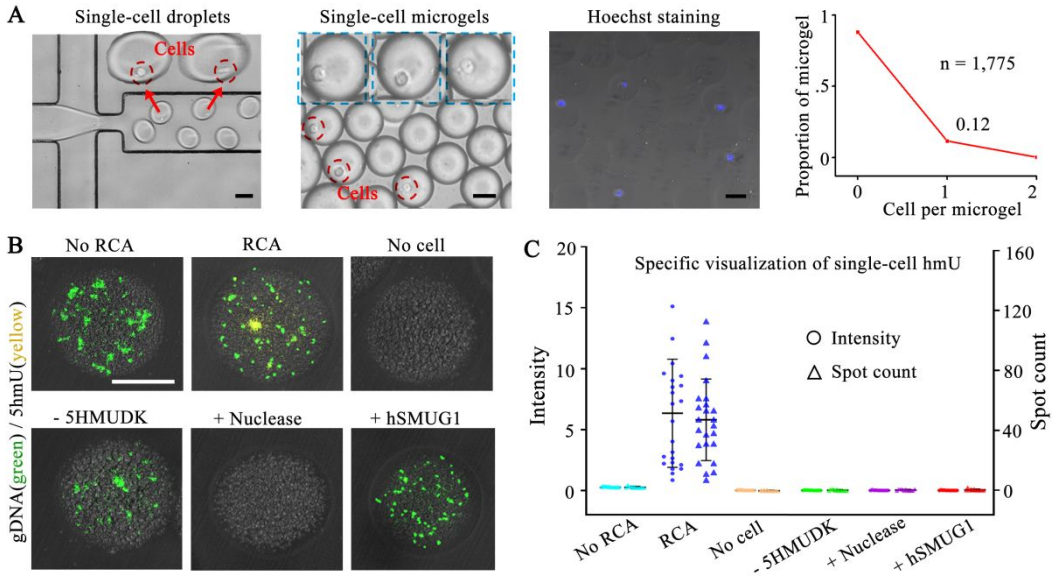


Figure 2. Single-cell 5hmU analysis in microgels. (A) Characterization of single-cell microgels. The first panel: generating hydrogel droplets containing a single cell; the second panel: microgels containing a single cell; the third panel: microgels stained by Hoechst 33342; the fourth panel: cell loading distribution into microgels. (B) Representative fluorescence images (merged: green, gDNA; yellow, 5hmU) of individual microgels with different treatments. (C) Corresponding analysis of single-cell 5hmU fluorescence spot count and signal intensity. MCF-7 cell line is used here. The scale bars are 50 μm in this work.

(Figure S6). DNA leakage and cross-contamination between microgels were not observed. As we know, the pore size of 1.5% agarose is about 130 nm that limits the diffusion of genomic fragments more than tens of kb molecular weight.[37,38] These merits allow us to label, amplify and visualize single-cell 5hmU successively.

For 5hmU labeling reaction, the microgels were incubated with 5-HMUDK and ATP-γ-S. To introduce DNA amplification, the newly generated 5psmU was reacted with 5' chemically modified DNA primer. This primer can induce rolling circle amplification (RCA), producing a long single-stranded DNA concatemer. This RCA product (RCP) contains hundreds of a periodically repeated sequence, allowing the capture of many molecules of fluorescence probe for signal amplification. Thus the RCP can be visualized as a bright fluorescence spot (Figure 2B). The number of 5hmU sites in single cells may be roughly assessed by counting the spots in the micorgels. We observed higher fluorescence intensity (yellow, including lots of spots) in RCA-treated micorgels compared to those without RCA. In order to validate the assay specificity, negative experiments and blank control (empty microgel, no cell) were also performed. Among them, two 5hmU-deficiency samples are prepared by the treatment of hSMUG1 or nuclease prior to the thiophosphorylation reaction. And another negative experiment lacks 5-HMUDK. We observed negligible nonspecific labeling or adsorption background fluorescence in these control samples. Both the fluorescence intensity and spot count in the microgels of positive sample are very much higher than those of all negative controls. These results verified

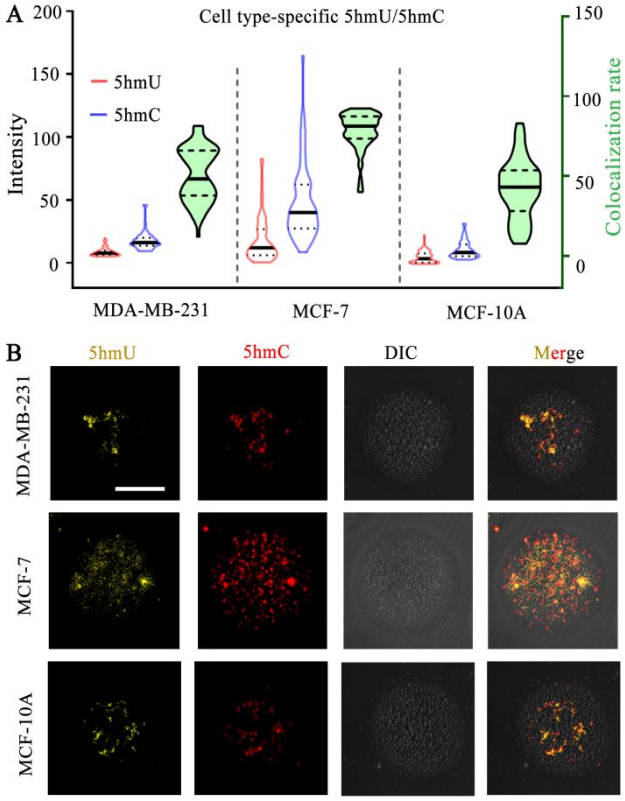


Figure 3. Cell type-specific 5hmU and 5hmC visualized by sc5hmU/5hmC-microgel. (A) Their intensity distribution and colocalization analysis of different cell lines (n= 68 for MDA-MB-231, 69 for MCF-7 and 56 for MCF-10A). (B) Representative microgel images. DIC is the abbreviation of differential interference contrast.

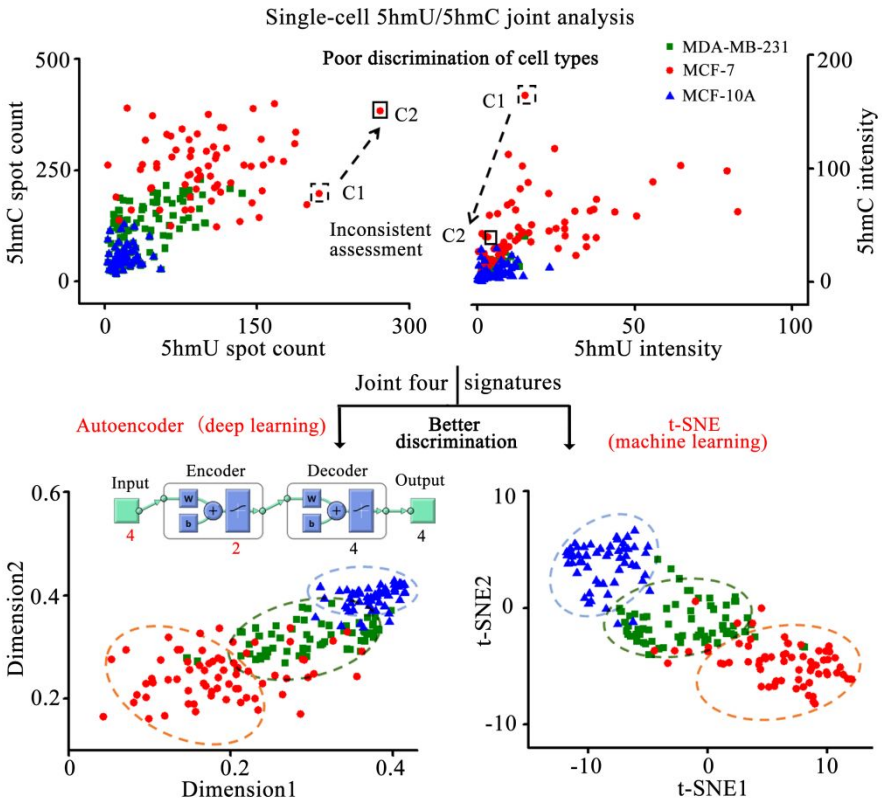


Figure 4. Discrimination of breast cell lines by signatures of both 5hmU and 5hmC. Only by spot numbers (upper left) or (upper right) fluorescence intensities. C1 and C2 indicate two individual cells; Both their locations and relationship in these two figures are not consistent. The four-dimensional signatures (two spot numbers and two fluorescence intensities) are visualized onto a two-dimensional space by Autoencoder (bottom left) and t-SNE (bottom right). n= 68, 69 and 56 for MDA-MB-231, MCF-7 and MCF-10A, respectively.

the assay specificity of 5hmU by our method in single-cell samples.

Sc5hmU/5hmC-microgel revealing cell type-specific molecular signatures. We then developed sc5hmU/5hmC-microgel by introducing T4 β-GT to label 5hmC with azide glucose after 5hmU labeling. The azide group can be crosslinked with 5' DBCO modified DNA primer via click chemistry. Two sets of 5' modified primers and corresponding padlock probes (Table S1) were validated to execute independent RCA reactions without cross interference (Figure S7). The detection specificity of 5hmC in microgels was also confirmed (Figure S8), which is similar to that of 5hmU mentioned above. Joint analysis of single-cell 5hmU and 5hmC in microgels was then carried out. Three human breast epithelial cell lines, non-tumorigenic MCF-10A, carcinoma MCF-7 and highly-invasive MDA-MB-231, were used as cell samples.

As shown in Figure 3, we used fluorescence intensity of individual microgels to roughly depict the abundance of these two bases in different cell lines. It can be seen that their mean values of MCF-7 are each higher than those of MCF-10A and MDA-MB-231, despite the remarkable single-cell heterogeneity. The colocalization analysis of two fluorescence channels was also performed to study the location relationship between 5hmU and 5hmC in genomic DNA. Different and high colocalization rates are observed in these cell lines (Figure 3). It may imply the close location of 5hmU and 5hmC sites in genomic DNA, though most 5hmU are not derived from 5hmC in mammalians.[9] It is worthy to note that the spatial resolution of routine fluorescence microscopy is limited to about 250 nm and even the super-resolution microscopy is not accurate for the colocalization analysis of two DNA bases. Although deep sequencing obtains single-base resolution information, yet simultaneous sequencing of 5hmU and 5hmC in single cells has not been reported. In addition to fluorescence intensity, the fluorescence spot count was also extracted to investigate the number and even distribution of these modified bases (Figure S9). Each spot here is arbitrarily hypothesized as one modified base site or one RCP for statistical analysis, though some spots with large size result from the overlap of multiple base sites or RCPs (Figure 3B). Thus the spot number can't actually represent the site number of modified bases, but it may reveal the distribution (decentralized or concentrated) feature partly combined with the intensity information. These results revealed cell type-specific molecular signatures of both 5hmU and 5hmC with significant single-cell variation.

Furthermore, we explored whether 5hmU and 5hmC signatures (fluorescence intensity and spot number, raw data in Table S2) can be used to discriminate non-tumorigenic cells from carcinoma cells. Based on the raw

data of only one base, no good distinction between cell samples is observed (Figure S10). We then carried out the joint analysis of two bases. As shown in the upper panels of Figure 4, these cell lines are still poorly resolved by only two spot numbers or two fluorescence intensities. And the assessment of the same single cells by spot numbers are not consistent with that by fluorescence intensities, which is ascribed to the decentralized or concentrated distribution of bases as mentioned above. Therefore, both fluorescence intensity and spot number of two bases as four signatures should be taken into consideration for the joint analysis. Several routine dimension reduction algorithms, including deep learning Autoencoder and machine learning t-distributed stochastic neighbor embedding (t-SNE) and Principal Component Analysis (PCA), are used to visualize these four-dimensional raw data on a two-dimensional space, respectively. As shown in the bottom panels of Figure 4, most of the same cell lines are seen to dominate in particular regions by Autoencoder and t-SNE. And t-SNE achieves better cell discrimination performance than Autoencoder. Yet PCA fails to improve cell-type assignments (Figure S11). The different performance of these three dimension reduction methods may depend on their algorithms and the original dataset. Using t-SNE as the representative, we also map the raw data of two signatures (above-mentioned four combinations) onto a two-dimensional space. Their results are shown as corresponding scatter plots in Figure S12. They present somewhat improved cell discrimination compared to using raw data, yet the samples of the same cell types are not well colocalized or to dominate in particular regions. And their discrimination performance are still much lower than that of four signatures by t-SNE. Overall, we demonstrate that the molecular information of 5hmU and 5hmC can be used to differentiate non-tumorigenic, carcinoma and highly-invasive breast cell lines. Notably, higher dimensional signatures of three or more modified bases (e.g., 5mC, 5hmC, 5hmU, 5fC and 5fU) may afford better discrimination of these and more cell types.

DISCUSSION

Although genomic 5hmC has been well detected and studied, the 5hmU specific recognition and labeling remain challenging. It also limits the joint analysis of 5hmU and 5hmC in single cells, which is crucial for the exploration of their signatures and function. In this work, we specifically label 5hmU sites with thiophosphate and follow-up DNA barcodes for the first time. Combining droplet microfluidics and DNA amplification, we demonstrate the differentiated visualization of single-cell 5hmU and 5hmC encapsulated in microgels. Our sc5hmU/5hmC-microgel method allows high-throughput processing of single cells, which is necessary for the reliable description of a heterogeneous cell population and the identification of rare cell types. The DNA barcoding amplification enables simultaneous analysis of multiple modified bases with single-molecule sensitivity. The information of rough abundance, number and distribution feature of both 5hmU and 5hmC is obtained. We reveal the cell type-specific molecular signatures of these bases with single-cell heterogeneity. Moreover, we differentiate non-tumorigenic, carcinoma and highly-invasive breast cell lines with improved performance by machine learning and deep learning algorithms to four-dimensional signatures. In addition to cultured cells, the proposed method may also be applicable to the analysis of clinical cell samples such as fine needle aspiration biopsies. It is notable that the flow cytometer can also be used to measure the fluorescence intensity of microgels, yet it would not provide the information of number and distribution of modified bases.

Despite global analysis of interested bases in single cells, sometimes we only want to evaluate the epigenetic changes at particular genomic sites. To achieve locus-specific detection, DNA-binding probes should be designed and combined with base labeling probes, followed by proximity ligation assay.[39,40] Notably, genome editing tools such as zinc-finger nucleases, transcription activator-like effector nucleases and clustered regularly interspaced short palindromic repeats-associated protein systems, contain sequence-programmable DNA-binding modules. They may be utilized to design promising gene locus-specific probes.[41] This proximity assay is perhaps appropriate for the investigation of proximal base modifications. It is well known that 5fU is easily converted into 5hmU by reductants such as sodium borohydride. Thus, our method is also suitable for single-cell 5fU analysis with the pretreatment of 5hmU blocking and 5fU reduction. Moreover, it may also be used to evaluate RNA modifications such as N_6-methyladenosine if we can capture RNA inside the microgels.

It is notable that our method is unable to measure subcellular distribution patterns of these modified bases. To obtain the spatial information within cells and even tissues, in situ 5hmU imaging assay should be developed. Yet it is challenged by the interference from abundant sulfhydryls inside cells. Therefore, we should replace ATP-γ-S with new ATP analogues modified with orthogonal click tags such as azide groups.[42] Except chemical synthesis of these ATP analogues, it is required to investigate whether 5-HMUDK can recognize and use them as phosphate donors. By in situ cell imaging, we would also explore the modified bases in mitochondria and reveal their interaction with histone post-translational modifications. Furthermore, the specific 5hmU recognition and labeling method proposed in this work may be applied to genome-wide mapping of 5hmU or 5fU by deep sequencing. For instance, 5hmU-containing genomic fragments can be selectively pulled down and enriched by chemical labeling, followed by library construction and sequencing. Such a strategy will fail to identify genomic 5hmU with single-base resolution.[2,14] Furthermore, single-cell genomic 5hmU sequencing remain still a great challenge. Altogether, our strategy provides a new route to analyze and decode single-cell DNA epigenetic modifications. We also expect this conceptual framework of specific 5hmU labeling integrated with microfluidics device to accelerate the

development of high-throughput single-cell epigenome sequencing.

ASSOCIATED CONTENT

Supporting Information.
This material is available free of charge via the Internet at http://pubs.acs.org.
Experimental details and data (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: yxzhao@mail.xjtu.edu.cn.

ORCID

Yongxi Zhao: 0000-0002-1796-7651

Author Contributions

‡These authors contributed equally.

Funding Sources

This research was financially supported by the National Natural Science Foundation of China [grant number 31671013, 21705124 and 21874105], the China Postdoctoral Science Foundation [grant number 2017M613102 and 2018T111032], the Fundamental Research Funds for the Central Universities and "Young Talent Support Plan" of Xi'an Jiaotong University.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

We thank Miss Lu and Miss Hao at Instrument Analysis Center of Xi'an Jiaotong University for their assistance with LC-MS and confocal fluorescence imaging analysis.

REFERENCES

(1) Berney, M.; McGouran, J. F. Methods for detection of cytosine and thymine modifications in DNA. *Nat. Rev. Chem.* **2018**, 2, 332-248.

(2) Kawasaki, F.; Beraldi, D.; Hardisty, R. E.; Mcinroy, G. R.; Delft, P. V.; Balasubramanian, S. Genome-wide mapping of 5-hydroxymethyluracil in the eukaryote parasite Leishmania. *Genome Biol.* **2017**, 18, 23.

(3) Zeng, H.; He, B.; Xia, B.; Bai, D.; Lu, X.; Cai, J.; Chen, L.; Zhou, A.; Zhu, C.; Meng, H. Bisulfite-Free, Nanoscale analysis of 5-hydroxymethylcytosine at single base resolution. *J. Am. Chem. Soc.* **2018**, 140, 13190-13194.

(4) Xia, B.; Han, D.; Lu, X.; Sun, Z.; Zhou, A.; Yin, Q.; Zeng, H.; Liu, M.; Jiang, X.; Xie, W. Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. *Nat. Methods* **2015**, 12, 1047-1050.

(5) Zhao, C.; Wang, H.; Zhao, B.; Li, C.; Yin, R.; Song, M.; Liu, B.; Liu, Z.; Jiang, G. Boronic acid-mediated polymerase chain reaction for gene- and fragment-specific detection of 5-hydroxymethylcytosine. *Nucleic Acids Res.* **2014**, 42, e81.

(6) Zhou, X. 5-Formyluracil as a Multifunctional Building Block in Biosensor Designs. *Angew. Chem.* **2018**, 57, 9689-9693.

(7) Li, Q.; Xie, N.; Xiong, J.; Yuan, B.; Feng, Y. Single-nucleotide resolution analysis of 5-hydroxymethylcytosine in DNA by enzyme-mediated deamination in combination with sequencing. *Anal. Chem.* **2018**, 90, 14622-14628.

(8) Hu, L.; Liu, Y.; Han, S.; Yang, L.; Cui, X.; Gao, Y.; Dai, Q.; Lu, X.; Kou, X.; Zhao, Y. Jump-seq: Genome-Wide Capture and Amplification of 5-Hydroxymethylcytosine Sites. *J. Am. Chem. Soc.* **2019**, 141, 8694-8697.

(9) Pfaffeneder, T.; Spada, F.; Wagner, M.; Brandmayr, C.; Laube, S. K.; Eisen, D.; Truss, M.; Steinbacher, J.; Hackner, B.; Kotljarova, O. Tet oxidizes thymine to 5-hydroxymethyluracil in mouse embryonic stem cell DNA. *Nat. Chem. Biol.* **2014**, 10, 574-581.

(10) Pais, J. E.; Dai, N.; Tamanaha, E.; Vaisvila, R.; Fomenkov, A. I.; Bitinaite, J.; Sun, Z.; Guan, S.; Corrêa, I. R.; Noren, C. J. Biochemical characterization of a Naegleria TET-like oxygenase and its application in single molecule sequencing of 5-methylcytosine. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, 112, 4316-4321.

(11) Nabel, C. S.; Jia, H.; Ye, Y.; Shen, L.; Goldschmidt, H. L.; Stivers, J. T.; Zhang, Y.; Kohli, R. M. AID/APOBEC deaminases disfavor modified cytosines implicated in DNA demethylation. *Nat. Chem. Biol.* **2012**, 8, 751-758.

(12) Iwan, K.; Rahimoff, R.; Kirchner, A.; Spada, F.; Schröder, A. S.; Kosmatchev, O.; Ferizaj, S.; Steinbacher, J.; Parsa, E.; Müller, M. 5-Formylcytosine to cytosine conversion by C-C bond cleavage in vivo. *Nat. Chem. Biol.* **2017**, 14, 72-78.

(13) Bachman, M.; Uribe-Lewis, S.; Yang, X.; Williams, M.; Murrell, A.; Balasubramanian, S. 5-Hydroxymethylcytosine is a predominantly stable DNA modification. *Nat. Chem.* **2014**, 6, 1049-1055.

(14) Fumiko, K.; Sergio, M. C.; Dario, B.; Areeb, M.; Hardisty, R. E.; Mark, C.; Shankar, B. Sequencing 5-Hydroxymethyluracil at Single-Base Resolution. *Angew. Chem.* **2018**, 57, 9694 –9696.

(15) Chen, F.; Bai, M.; Cao, K.; Zhao, Y.; Wei, J.; Zhao, Y. Fabricating MnO2 Nanozymes as Intracellular Catalytic DNA Circuit Generators for Versatile Imaging of Base-Excision Repair in Living Cells. *Adv. Funct. Mater.* **2017**, 27, 1702748.

(16) Chen, F.; Xue, J.; Bai, M.; Qin, J.; Zhao, Y. Programming in situ accelerated DNA walkers in diffusion-limited microenvironments. *Chem. Sci.* **2019**, 10, 3103-3109.

(17) Shu, X.; Liu, M.; Lu, Z.; Zhu, C.; Meng, H.; Huang, S.; Zhang, X.; Yi, C. Genome-wide mapping reveals that deoxyuridine is enriched in the human centromeric DNA. *Nat. Chem. Biol.* **2018**, 14, 680-687.

(18) Song, C.; Szulwach, K. E.; Fu, Y.; Dai, Q.; Yi, C.; Li, X.; Li, Y.; Chen, C.; Zhang, W.; Jian, X. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat. Biotechnol.* **2011**, 29, 68-72.

(19) Yu, M.; Song, C.; He, C. Detection of mismatched 5-hydroxymethyluracil in DNA by selective chemical labeling. *Methods* **2015**, 72, 16-20.

(20) Shang, L.; Cheng, Y.; Zhao, Y. Emerging droplet microfluidics. *Chem. Rev.* **2017**, 117, 7964-8040.

(21) Zhu, Z.; Yang, C. J. Hydrogel droplet microfluidics for high-throughput single molecule/cell analysis. *Acc. Chem. Res.* **2016**, 50, 22-31.

(22) Kaminski, T.; Garstecki, P. Controlled droplet microfluidic systems for multistep chemical and biological assays. *Chem. Soc. Rev.* **2017**, 46, 6210-6226.

(23) Zhang, L.; Feng, Q.; Wang, J.; Sun, J.; Shi, X.; Jiang, X. Microfluidic synthesis of rigid nanovesicles for hydrophilic reagents delivery. *Angew. Chem. Int. Ed.* **2015**, 54, 3952-3956.

(24) Chen, P.; Yan, S.; Wang, J.; Guo, Y.; Dong, Y.; Feng, X.; Zeng, X.; Li, Y.; Du, W.; Liu, B. Dynamic Microfluidic Cytometry for Single-Cell Cellomics: High-Throughput Probing Single-Cell-Resolution Signaling. *Anal. Chem.* **2018**, 91, 1619-1626.

(25) Xiao, M.; Zou, K.; Li, L.; Wang, L.; Tian, Y.; Fan, C.; Pei, H. Stochastic DNA Walker in Droplets for Super-Multiplex Bacteria Phenotype Detection. *Angew. Chem.* **2019**, 131, 15594 – 15600.

(26) Liu, C.; Zhao, J.; Tian, F.; Chang, J.; Zhang, W.; Sun, J. λ-DNA-and Aptamer-Mediated Sorting and Analysis of Extracellular Vesicles. *J. Am. Chem. Soc.* **2019**, 141, 3817-3821.

(27) Kuo, C.-T.; Thompson, A. M.; Gallina, M. E.; Ye, F.; Johnson, E. S.; Sun, W.; Zhao, M.; Yu, J.; Wu, I.-C.; Fujimoto, B. Optical painting and fluorescence activated sorting of single adherent cells labeled with photoswitchable Pdots. *Nat. Commun.* **2016**, 7, 11468.

(28) Mao, S.; Zhang, W.; Huang, Q.; Khan, M.; Li, H.; Uchiyama, K.; Lin, J. M. In Situ Scatheless Cell Detachment Reveals Correlation between Adhesion Strength and Viability at Single - Cell Resolution. *Angew. Chem. Int. Ed.* **2018**, 57, 236-240.

(29) Qin, Y.; Wu, L.; Schneider, T.; Yen, G. S.; Wang, J.; Xu, S.; Li, M.; Paguirigan, A. L.; Smith, J. L.; Radich, J. P. A Self - Digitization Dielectrophoretic (SD - DEP) Chip for High - Efficiency Single - Cell Capture, On - Demand Compartmentalization, and Downstream Nucleic Acid Analysis. *Angew. Chem.* **2018**, 130, 11548-11553.

(30) Yen, G. S.; Fujimoto, B. S.; Schneider, T.; Kreutz, J. E.; Chiu, D. T. Statistical Analysis of Nonuniform Volume Distributions for Droplet-Based Digital PCR Assays. *J. Am. Chem. Soc.* **2019**, 141, 1515-1525.

(31) Novak, R.; Zeng, Y.; Shuga, J.; Venugopalan, G.; Fletcher, D. A.; Smith, M. T.; Mathies, R. A. Single - Cell Multiplex Gene Detection and Sequencing with Microfluidically Generated Agarose Emulsions. *Angew. Chem. Int. Ed.* **2011**, 50, 390-395.

(32) Zhang, H.; Jenkins, G.; Zou, Y.; Zhu, Z.; Yang, C. J. Massively parallel single-molecule and single-cell emulsion reverse transcription polymerase chain reaction using agarose droplet microfluidics. *Anal. Chem.* **2012**, 84, 3599-3606.

(33) Kim, S. C.; Clark, I. C.; Shahi, P.; Abate, A. R. Single-cell RT-PCR in microfluidic droplets with integrated chemical lysis. *Anal. Chem.* **2018**, 90, 1273-1279.

(34) Streets, A. M.; Zhang, X.; Cao, C.; Pang, Y.; Wu, X.; Xiong, L.; Yang, L.; Fu, Y.; Zhao, L.; Tang, F. Microfluidic single-cell whole-transcriptome sequencing. *Proc. Natl. Acad. Sci.* **2014**, 111, 7048-7053.

(35) Fu, Y.; Li, C.; Lu, S.; Zhou, W.; Tang, F.; Xie, X.; Huang, Y. Uniform and accurate single-cell sequencing based on emulsion whole-genome amplification. *Proc. Natl. Acad. Sci.* **2015**, 112, 11923-11928.

(36) Fu, Y.; Chen, H.; Liu, L.; Huang, Y. Single cell total RNA sequencing through isothermal amplification in picoliter-droplet emulsion. *Anal. Chem.* **2016**, 88, 10795-10799.

(37) Stellwagen, N. C. Agarose gel pore radii are not dependent on the casting buffer. *Electrophoresis* **1992**, 13, 601-603.

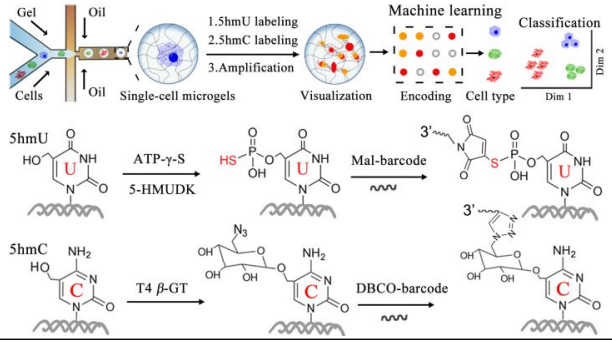
(38) Xiong, J.; Narayanan, J.; Liu, X.; Chong, T.; Chen, S.; Chung, T. Topology Evolution and Gelation Mechanism of Agarose Gel. *J. Phys. Chem. B* **2005**, 109, 5638-5643.

(39) Li, G.; Montgomery, J. E.; Eckert, M. A.; Chang, J. W.; Tienda, S. M.; Lengyel, E.; Moellering, R. E. An activity-dependent proximity ligation platform for spatially resolved quantification of active enzymes in single cells. *Nat. Commun.* **2017**, 8, 1775.

(40) Gomez, D.; Shankman, L. S.; Nguyen, A. T.; Owens, G. K. Detection of histone modifications at specific gene loci in single cells in histological sections. *Nat. methods* **2013**, 10, 171-177.

(41) Zhang, K.; Deng, R.; Teng, X.; Li, Y.; Sun, Y.; Ren, X.; Li, J. Direct visualization of single-nucleotide variation in mtDNA using a CRISPR/Cas9-mediated proximity ligation assay. *J. Am. Chem. Soc.* **2018**, 140, 11293-11301.

(42) Chen, F.; Bai, M.; Cao, X.; Zhao, Y.; Xue, J.; Zhao, Y. Click-encoded rolling FISH for visualizing single-cell RNA polyadenylation and structures. *Nucleic Acids Res.* **2019**, 47, e145.


Gel
Oil
Oil
Cells
Single-cell microgels
1.5hmU labeling
2.5hmC labeling
3.Amplification
Visualization
Machine learning
Encoding
Cell type
Classification
Dim 1
Dim 2
5hmU
ATP-γ-S
5-HMUDK
Mal-barcode
5hmC
T4 β-GT
DBCO-barcode