A Journal of the Gesellschaft Deutscher Chemiker
Angewandte
GDCh
International Edition
Chemie
www.angewandte.org

Accepted Article

Title: Pairwise Proximity-differentiated Visualization of Single-cell DNA Epigenetic Marks

Authors: Jing Xue, Feng Chen, Li Su, Xiaowen Cao, Min Bai, Yue Zhao, Chunhai Fan, and Yongxi Zhao

To be cited as: *Angew. Chem. Int. Ed.* 10.1002/anie.202011172

Link to VoR: https://doi.org/10.1002/anie.202011172

WILEY-VCH

10.1002/anie.202011172

WILEY-VCH

Pairwise Proximity-differentiated Visualization of Single-cell DNA Epigenetic Marks

Jing Xue,[+] Feng Chen,[+] Li Su,[+] Xiaowen Cao, Min Bai, Yue Zhao, Chunhai Fan, and Yongxi Zhao*

Abstract: Spatial positioning and proximity of relevant biomolecules such as DNA epigenetic marks are fundamental to a deeper understanding of life process. However, it remains poorly explored and technically challenging. Here we report pairwise proximity-differentiated visualization of single-cell 5-formylcytosine (5fC) and 5-hydroxymethylcytosine (5hmC). These two marks on chromatin in fixed cells are successively labeled and crosslinked with their DNA primer probes via click chemistry. Based on a pairwise proximity-differentiated mechanism, proximal 5fC/5hmC sites firstly and residual 5fC or 5hmC sites then are encoded with respective circularized barcodes. These barcodes are simultaneously amplified for multiplexed single-molecule imaging. Thus we demonstrate differentiated visualization of 5fC or 5hmC spatial positioning and their pairwise proximity in single cells. Such multi-level subcellular information may provide insights into regulation functions and mechanisms of chromatin modifications, and the spatial proximity can expose the potential crosstalk or interaction between their reader proteins.

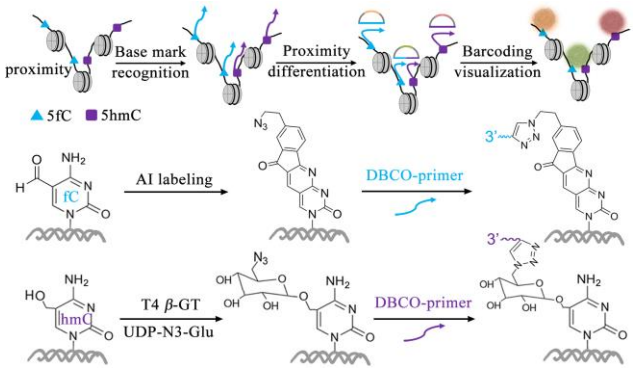
Intracellular positioning of relevant biomolecules and their nanoscale colocalization or proximity are important to the fate and function of both single cells and living organisms. For example, a serial of DNA epigenetic marks exist in mammalian genomes. They have profound influence in gene expression and chromatin architecture, and are implicated in pathological associations including cancers.[1] The best-known ones are 5-methylcytosine (5mC) and its oxidized derivatives 5-hydroxymethylcytosine (5hmC) and 5-formylcytosine (5fC).[2] Several deep sequencing strategies have been well developed for genome-wide analysis of these modifications, revealing their important functions in normal biological processes and diseases.[3] Yet these methods rely on cell population averaging, and are disabled to visualize the subcellular distribution. The spatial positioning of epigenetic marks in single cells is fundamental to understanding their functions.[4] Furthermore, accumulated evidence has verified the co-occurrence of two epigenetic marks at the same gene regions and suggested their spatial proximity on the basis of chromatin interactions.[2a, 5] Exploring the proximity of different epigenetic marks may reveal their potential crosstalk or interaction of their reader proteins. Unfortunately, single-cell visualization of both spatial positioning and proximity of epigenetic marks has not been explored.

Traditional fluorescence *in situ* hybridization-based cell imaging is limited to detect sequences of interest and inadequate for the assessment of base modifications.[6] Immunoassays use antibodies to bind to base modifications such as 5hmC only in naked DNA.[7] The base sites covered by DNA-binding proteins or in condensed chromatins are sterically excluded from antibodies with relatively large sizes. Such large sizes also disturbed the exploration of spatial proximity of different marks. Chemoenzymatic- or chemical-based methods have been well developed for the covalent labeling of 5hmC or 5fC coupled with bioorthogonal click tags.[2a, 3a, 3d] These methods provide considerable improvement over antibody-based ones as they ensure more accurate and comprehensive labeling of DNA epigenetic sites. But the recognition and visualization of proximity between these epigenetic marks remains limited. DNA nanotechnology may be a promising solution ascribed to predictable and programmable assembly with nanoscale precision.[8] We and others have developed various nanostructured DNA probes for molecule positioning, proximity recognition, assembly or disassembly.[9] In particular, DNA walkers or motors are designed based on assembly-disassembly cycles to continuously recognize adjacent substrate molecules.[10] And proximity DNA ligation assay can detect the pairwise proximity of two molecules such as proteins to investigate their spatial distances and interactions.[11] In a typical proximity ligation assay, both pairwise proximity target sites and residual target sites were labeled with the DNA recognition probes. These probes then participated in hybridization and ligation reaction. However, this method only detected the pairwise proximity sites and can't record the residual target sites. Moreover, the probes labeled at residual target sites have been still occupied by ligation probes for proximity assay, thus they were difficult to be detected differently from pairwise proximity target sites. Therefore, differentiated visualization of spatial positioning and proximity of two different DNA epigenetic marks remains challenging.

Herein, we report pairwise proximity-differentiated visualization of single-cell 5fC and 5hmC (Scheme 1 and S1). These two oxidized cytosines are selected as representative epigenetic marks due to their well-performed studies. We require to synthesize an specific 5fC labeling probe azido derivative of 1,3-indandione (AI)[2a]. And we report its practicability in the complex intracellular environment against the interferences from various analogous such as 5-formyluracil (5fU). In addition, 5hmC is routinely labeled with click tags by T4 phage β-glucosyltransferase (β-GT). Both 5fC and 5hmC on chromatins in fixed cells are sequentially labeled and crosslinked with their DNA primer probes via click chemistry. Based on a pairwise proximity-differentiated mechanism, our method can successively encode proximal 5fC/5hmC sites and residual 5fC or 5hmC sites with respective circularized DNA barcodes. In brief, a pair of splint ligation probes captured at 5fC/5hmC

[a] J. Xue, Prof. F. Chen, L. Su, X. W. Cao, M. Bai, Dr. Y. Zhao, Prof. Y. X. 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, China
E-mail: yxzhao@mail.xjtu.edu.cn
Prof. C. H. Fan
Institute of Molecular Medicine, Renji Hospital, School of Medicine and School of Chemistry and Chemical Engineering, Shanghai Jiao Tong University, Shanghai 200127, China
[+] These authors contributed equally to this work.

Supporting information for this article is given via a link at the end of the document.



Scheme 1. Workflow of pairwise proximity-differentiated visualization of single-cell DNA epigenetic marks. The cells are fixed before the recognition and differentiation of target sites.

proximity sites form two nicks, and then they are ligated as one circularized DNA barcode. In contrast, molecules of these probes at residual, unpaired modification sites are released by DNA strand displacement. The restored DNA primer probes can hybridize with the circularized DNA barcodes for 5fC or 5hmC. These three barcodes will work as templates for rolling circle amplification (RCA). The long single-stranded RCA products (RCPs) contain periodically repeated sequences to capture fluorescent DNA probes for signal amplification with single-molecule sensitivity. Thus we realize the differentiated visualization of 5fC or 5hmC spatial positioning and 5fC/5hmC proximity in single cells. Such multi-level spatial information may promote deeper investigations of their regulation functions and mechanisms. Notably, our method is designed to detect hetero-proximity (5fC/5hmC) instead of homo-proximity (5fC/5fC or 5hmC/5hmC). And it can be easily applied to the differentiated visualization of monovalent or bivalent histone modifications and to the detection of protein-protein interactions.

To demonstrate the proposed method generally, we designed DNA origami target substrates containing two kinds of target sites (A and B) with various distances (Figure 1A and Table S1). The sets of DNA probes used here mainly include primer probes for target stie labeling, a pair of splint ligation probes for pairwise proximity (A1/B1, ~10 nm) RCA, two pre-circularized DNA barcodes for respective targrt site (A2 or B2) barcoding RCA, two displacement probes and three fluorescence probes. The displacement probes are designed to be complementary to splint ligation probes for toehold-mediated strand displacement. Their squence lengths can significantly affect the thermodynamics and performance of pairwise proximity-differentiated recognition and visualization. Thus we firstly predicted the standard free energies of different displacement probes hybridizing to the ligated splint ligation probes (Figure S1). And as shown in the bottom panel of Figure 1A, we demonstrated that only the displacement probes with an appropriate squence length can

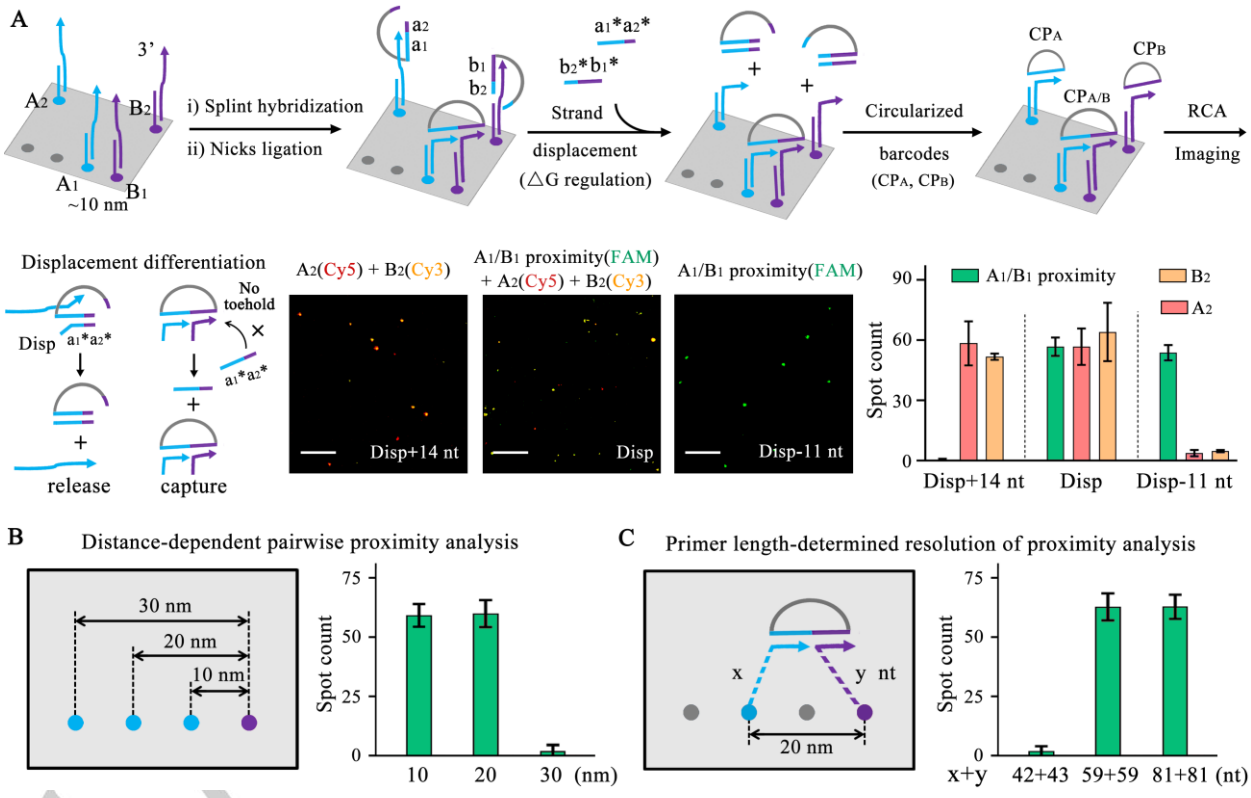


Figure 1. Pairwise proximity-differentiated visualization on DNA origami. (A) Principle of the method. Influence of strand displacement reaction on differentiated recognition and visualization was investigated. Disp, Disp-11 and Disp+14 indicate the displacement probes (a1*a2*) with different sequence lengths. The process of b2*b1* is similar to that of a1*a2*. Cy5, Cy3 and FAM indicate the fluorescence channels. (B) Distance-dependent proximity visualization. (C) The effect of primer probe length on proximity visualization. The scale bars in the images are 15 μm. 10 images are randomly selected for each statistical analysis.

distinguish A1/B1 pairwise proximity sites versus two kinds of residual sites (A2 or B2), which is in agreement with the energy prediction. In contrast, overlong displacement probes caused the loss of fluorescence signals of A1/B1 pairwise proximity sites, and too short displacement probes were disabled to trigger the fluorescence signals of A2 or B2 sites. Furthermore, we investigated the effect of different separation or distances (~10, ~20 and ~30 nm) between two target sites on the pairwise proximity analysis (Figure 1B). We observed very low fluorescence signals when the distance of two sites reached ~30 nm. And we also demonstrated that the detection distance or spatial resolution of pairwise proximity sites was mainly determined by the length of two kinds of DNA primer probes (Figure 1C). Notably, the distances of about 10-20 nm match the study of two protein interactions or protein complexes. Overall, these results verified the pairwise proximity-differentiated visualization analysis by our proposed method.

For the analysis of 5hmC and 5fC, we first confirmed the synthesis of AI, 5fC labeling by AI and 5hmC labeling by β-GT using UDP-N_3-glucose (Figure S2-S4). We also demonstrated the pairwise proximity visualization of 5hmC/5fC at the same dsDNA substrates *in vitro* (Figure S5), which is similar to above experiments of DNA origami substrates. Notably, the quantitative assessment of labeling these marks with DNA primer is of great significance but difficulty, and it may be realized by ultrasensitive analytical mass spectrometry and DNA quantitative techniques such as qPCR or unique molecular identifier-used next generation sequencing. Then the spatial positioning of intracellular 5hmC and 5fC was investigated. We demonstrated that no potential uncrosslinked 5fC-AI adduct was detected by the set of DNA probes for 5hmC analysis (Figure S6-S8). Considering the diversity of chromatin modifications including histone modifications,[12] we explored the colocalization relationship between 5hmC and 5fC, and between 5hmC or 5fC and one of five model histone modifications by routine fluorescence colocalization analysis (Figure 2A and S9). These colocalization values presented remarkable single-cell heterogeneity (Figure 2B and 2C).[13] Especially, a low colocalization level of 5hmC with γH2AX is observed. Previous studies also reported their low colocalization level in untreated cells but a high value after DNA damage treatment.[7] However, the spatial resolution for colocalization analysis by traditional fluorescence microscopes is only about 250 nm in the lateral dimension.[14] It can't explore the nanoscale proximity (e.g., ~10 nm) between 5fC and 5hmC, which may be associated with their protein interaction, gene regulation and other functions.

Finally, we demonstrated differentiated visualization of 5fC or 5hmC spatial positioning and their nanoscale proximity in single cells by our proposed method. As described above, barcoding amplification was performed to generate three kinds of RCPs after pairwise proximity- differentiated recognition. These RCPs can be observed as bright spots (Figure 3 and S10). And the spot count of individual cells were routinely extracted using Image J. Both fluorescence intensity and spot count information are obtained to evaluate the distribution features. Fluorescence intensity of individual cells was used to roughly depict the global or total abundance of labeled sites, and spot count was roughly hypothesized as the number of representative detection sites. Notably, one bright spot with big sizes may cover several RCP

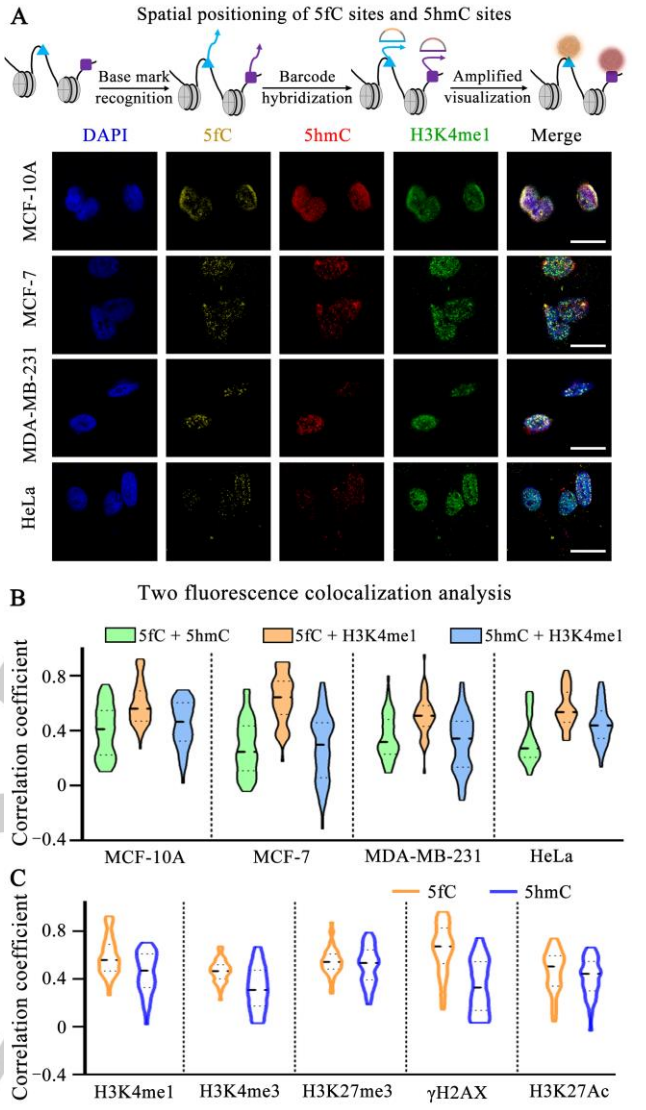


Figure 2. Spatial positioning of single-cell 5fC and 5hmC. (A) Cell images of 5fC, 5hmC and H3K4me1 in different cell lines. Scale bar, 10 μm. (B) Colocalization analysis of 5fC+5hmC, 5fC+H3K4me1 and 5hmC+H3K4me1 in different cell lines (N=50). (C) Colocalization analysis of 5fC or 5hmC with five histone modifications in MCF-10A cells (N=50). N: the number of cells.

molecules or detection sites which are localized closely. Negative controls were also performed to verify the assay specificity (Figure S10). We confirmed the spatial proximity between 5fC and 5hmC in cells for the first time.

In summary, we demonstrate differentiated visualization of 5fC or 5hmC spatial positioning and their proximity in single cells. A pairwise proximity-differentiated mechanism is proposed for successive recognition and amplification barcoding of proximal 5fC/5hmC sites and residual, unpaired 5fC or 5hmC sites. More comprehensive subcellular distribution information is provided for deeper epigenetics studies. Especially, the spatial proximity can expose the potential crosstalk or interaction between their reader proteins and may suggest new function in related

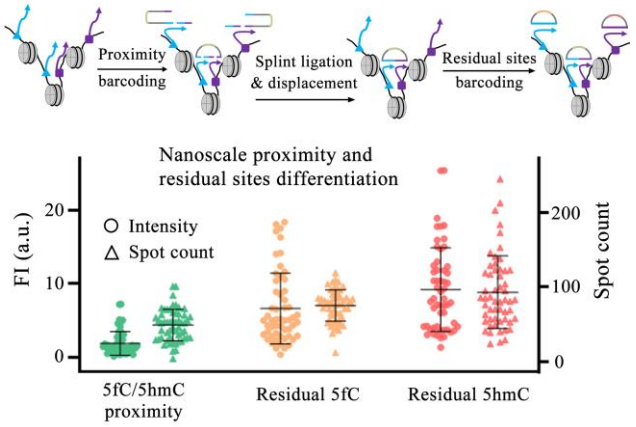


Figure 3. Pairwise proximity-differentiated visualization of 5fC and 5hmC in single cells. Top: schematic description of the design; bottom: statistical analysis of single-cell fluorescence intensity and RCP spot counts for each channel (N=55).

pathways. Notably, we can't determine whether these spatial proximity sites exist in the same or between two gene regions or chromatins. Furthermore, the developed method is not limited to the analysis of DNA epigenetic marks, and can be easily expanded to detect other pair combinations of histone modifications or proteins as mentioned above. It only requires to replace chemical probes for DNA modifications as specific antibodies for histone modifications or proteins whithout the change of DNA probe sets. Multiplexed visualization of large numbers of pairwise proximities in individual cells may also be realized by using sequential hybridization-image-erase barcoding cycles,[15] which increases the detection throughput beyond hundreds of target molecules. On the other hand, the two DBCO-primer probes of our method are occupied after the hybridization with splint ligation probes. It means that each copy of DBCO-primer probes can't be reused for the detection of another target site. If there are residual 5fC or 5hmC sites neighboring the sites of 5fC/5hmC proximity, they will induce fluorescence signals of residual sites rather than those of pairwise proximity sites. In addition, our method detects only an individual site of pairwise proximity per set of probe. To study a comprehensive set of proximities or interactions between multiple components, we should recycle the DNA probes for continuous and repeated recognition of any potential target sites.

Acknowledgements

This research was funded by the National Natural Science Foundation of China (grant number 31671013, 21705124 and 21874105), the China Postdoctoral Science Foundation [grant number 2017M613102, 2018T111032 and 2019M663658], the Natural Science Foundation of Shaanxi Province (grant number 2018JC-001), Innovation Capability Support Program of Shaanxi (grant number 2018PT-28 and 2019PT-05), the Fundamental Research Funds for the Central Universities and "Young Talent Support Plan" of Xi'an Jiaotong University. We thank Miss Hao and Miss Lu at Instrument Analysis Center of Xi'an Jiaotong University for the assistance with confocal fluorescence imaging and LCMS/MS analysis.

Keywords: DNA nanotechnology • epigenetic marks• pairwise proximity • differentiated recognition • single-cell imaging

[1] a) C. X. Song, K. E. Szulwach, Q. Dai, Y. Fu, P. Liu, L. Li, G. L. Xu, P. Jin, C. He, *Cell* **2013**, *153*, 678-691; b) M. Berney, J. F. McGouran, *Nat. Rev. Chem.* **2018**, *2*, 332-348.

[2] a) B. Xia, D. Han, X. Lu, Z. Sun, X. Jiang, W. Xie, C. He, C. Yi, *Nat. Methods* **2015**, *12*, 1047-1050; b) M. J. Booth, G. Marsico, M. Bachman, D. Beraldi, S. Balasubramanian, *Nat. Chem.* **2014**, *6*, 435; c) Q.-Y. Li, N.-B. Xie, J. Xiong, B. F. Yuan, Y. Q. Feng, *Anal. Chem.* **2018**, *90*, 14622-14628; d) M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik, S. Balasubramanian, *Science* **2012**, *336*, 934-937.

[3] a) H. Zeng, B. He, B. Xia, D. Bai, H. Guo, C. He, Q. Dai, C. Yi, *J. Am. Chem. Soc.* **2018**, *140*, 13190-13194; b) C. Zhao, H. Wang, B. Zhao, C. Li, R. Yin, M. Song, B. Liu, Z. Liu, G. Jiang, *Nucleic Acids Res.* **2014**, *42*, e81-e81; c) C. Liu, G. Zou, S. Peng, Y. Wang, W. Yang, F. Wu, Z. Jiang, X. Zhang, X. Zhou, *Angew. Chem. Int. Edit.* **2018**, *57*, 9689-9693; d) L. Hu, Y. Liu, S. Han, L. Yang, W. Sheng, S. Gao, X. He, C. He, *J. Am. Chem. Soc.* **2019**, *141*, 8694-8697; e) C. Zhu, Y. Gao, H. Guo, K. Kee, F. Tang, C. Yi, *Cell Stem Cell* **2017**, *20*, 720-731. e725.

[4] B. Z. Stanton, E. J. Chory, G. R. Crabtree, *Science* **2018**, *359*, eaao5902.

[5] a) D. N. Weinberg, S. Papillon-Cavanagh, H. Chen, Y. Yue, X. Chen, A. Djedid, A. S. Harutyunyan, N. Jabado, B. A. Garcia, H. Li, C. D. Allis, J. Majewski, C. Lu, *Nature* **2019**, *573*, 281-286; b) S. A. Quinodoz, N. Ollikainen, B. Tabak, A. Palla, J. M. Schmidt, M. Jovanovic, A. Chow, L. Cai, P. McDonel, M. Guttman, *Cell* **2018**, *174*, 744-757; c) B. Lai, Q. Tang, W. Jin, G. Hu, D. Wangsa, Y. Ding, M. Zhao, S. Liu, J. Song, T. Ried, K. Zhao, *Nat. Methods* **2018**, *15*, 741-747.

[6] F. Chen, M. Bai, X. Cao, Y. Zhao, J. Xue, Y. Zhao, *Nucleic Acids Res.* **2019**, *47*, e145-e145.

[7] a) G. R. Kafer, X. Li, T. Horii, I. Suetake, S. Tajima, I. Hatada, P. M. Carlton, *Cell rep.* **2016**, *14*, 1283-1292; b) S. Zhong, Z. Li, T. Jiang, X. Li, H. Wang, *Anal. Chem.* **2017**, *89*, 5702-5706.

[8] a) A. V. Pinheiro, D. Han, W. M. Shih, H. Yan, *Nat. Nanotechnol.* **2011**, *6*, 763; b) F. Zhang, S. Jiang, S. Wu, Y. Li, C. Mao, Y. Liu, H. Yan, *Nat. Nanotechnol.* **2015**, *10*, 779; c) P. Zhan, T. Wen, Z. g. Wang, Y. He, J. Shi, T. Wang, X. Liu, G. Lu, B. Ding, *Angew. Chem. Int. Edit.* **2018**, *57*, 2846-2850; d) H. Li, M. Wang, T. Shi, S. Yang, J. Zhang, H. H. Wang, Z. Nie, *Angew. Chem. Int. Edit.* **2018**, *57*, 10226-10230.

[9] a) H. Gu, J. Chao, S.-J. Xiao, N. C. Seeman, *Nature* **2010**, *465*, 202-205; b) W. Fu, L. Tang, G. Wei, L. Fang, J. Zeng, R. Zhan, X. Liu, H. Zuo, C. Z. Huang, C. Mao, *Angew. Chem. Int. Edit.* **2019**, *131*, 16557-16562; c) L. Xu, Y. Gao, H. Kuang, L. M. Liz - Marzán, C. Xu, *Angew. Chem. Int. Edit.* **2018**, *57*, 10544-10548; d) S. Li, Y. Liu, L. Liu, Y. Feng, L. Ding, H. Ju, *Angew. Chem. Int. Edit.* **2018**, *130*, 12183-12187.

[10] a) F. Chen, J. Xue, M. Bai, J. Qin, Y. Zhao, *Chem. Sci.* **2019**, *10*, 3103-3109; b) C. Jung, P. Allen, A. Ellington, *Nat. Nanotechnol.* **2016**, *11*, 157; c) X. Yang, Y. Tang, S. D. Mason, J. Chen, F. Li, *Acs Nano* **2016**,

10, 2324-2330; d) M. Xiao, K. Zou, L. Li, L. Wang, Y. Tian, C. Fan, H. Pei, *Angew. Chem. Int. Edit* . **2019**, *131*, 15594-15600.

[11] a) K. Zhang, R. Deng, X. Teng, Y. Li, Y. Sun, X. Ren, J. Li, *J. Am. Chem. Soc.* **2018**, *140*, 11293-11301; b) S. Fredriksson, M. Gullberg, J. Jarvius, C. Olsson, S. M. Gústafsdóttir, A. Östman, U. Landegren, *Nat. Biotechnol.* **2002**, *20*, 473-477; c) M. Gullberg, S. M. Gústafsdóttir, E. Schallmeiner, J. Jarvius, C. Betsholtz, U. Landegren, S. Fredriksson, *Proc. Nat. Acad. Sci.* **2004**, *101*, 8420-8424; d) O. Söderberg, M. Gullberg, M. Jarvius, K. Ridderstråle, F. Bahram, L.-G. Larsson, U. Landegren, *Nat. Methods* **2006**, *3*, 995-1000.

[12] H. Cedar, Y. Bergman, *Nat. Rev. Genet.* **2009**, *10*, 295-304.

[13] X. Chen, Y. Shen, W. Draper, J. D. Buenrostro, J. A. Doudna, W. J. Greenleaf, J. T. Liphardt, H. Y. Chang, *Nat. Methods* **2016**, *13*, 1013.

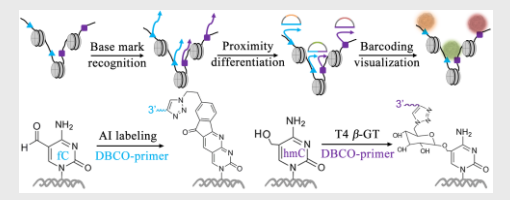
[14] a) K. A. Willets, A. J. Wilson, V. Sundaresan, P. B. Joshi, *Chem. Rev.* **2017**, *117*, 7538-7582; b) A. V. Diezmann, Y. Shechtman, W. E. Moerner, *Chem. Rev.* **2017**, *117*, 7244−7275; c) L. E. Weiss, Y. S. Ezra, S. Goldberg, B. Ferdman, O. Adir, A. Schroeder, O. Alalouf, Y. Shechtman, *Nat. Nanotechnol.* **2020**, *15*, 500-506.

[15] a) K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, X. Zhuang, *Science* **2015**, *348*, aaa6090; b) E. Lubeck, A. F. Coskun, T. Zhiyentayev, M. Ahmad, L. Cai, *Nat. Methods* **2014**, *11*, 360.

10.1002/anie.202011172

WILEY-VCH

COMMUNICATION



*Jing Xue, Feng Chen, Li Su, Xiaowen Cao, Min Bai, Yue Zhao, Chunhai Fan, and Yongxi Zhao**

Page No. – Page No.

Pairwise Proximity-differentiated Visualization of Single-cell DNA Epigenetic Marks

Pairwise proximity-differentiated visualization in single cells has been developed. This method achieves the differentiated recognition and barcoding amplification visualization of both 5fC/5hmC proximity sites and residual sites of 5fC or 5hmC on chromatins. It can be used to probe the valence of histone modifications and even protein interactions in single cells.