RSC Chemical Biology

ROYAL SOCIETY OF CHEMISTRY

REVIEW

View Article Online
View Journal | View Issue

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Check for updates

Cite this: *RSC Chem. Biol.*, 2021, **2**, 1096

# Quantification and mapping of DNA modifications

Yi Dai,[a] Bi-Feng Yuan *[ab] and Yu-Qi Feng [ab]

Received 3rd February 2021,
Accepted 20th May 2021

DOI: 10.1039/d1cb00022e

rsc.li/rsc-chembio

Apart from the four canonical nucleobases, DNA molecules carry a number of natural modifications. Substantial evidence shows that DNA modifications can regulate diverse biological processes. Dynamic and reversible modifications of DNA are critical for cell differentiation and development. Dysregulation of DNA modifications is closely related to many human diseases. The research of DNA modifications is a rapidly expanding area and has been significantly stimulated by the innovations of analytical methods. With the recent advances in methods and techniques, a series of new DNA modifications have been discovered in the genomes of prokaryotes and eukaryotes. Deciphering the biological roles of DNA modifications depends on the sensitive detection, accurate quantification, and genome-wide mapping of modifications in genomic DNA. This review provides an overview of the recent advances in analytical methods and techniques for both the quantification and genome-wide mapping of natural DNA modifications. We discuss the principles, advantages, and limitations of these developed methods. It is anticipated that new methods and techniques will resolve the current challenges in this burgeoning research field and expedite the elucidation of the functions of DNA modifications.

## 1 Introduction

DNA molecules employ four canonical nucleobases, adenine (A), thymine (T), cytosine (C), and guanine (G), to encode genetic information in living organisms.[1] In addition to these four nucleobases, recent years have witnessed the discovery of a variety of modified nucleobases in genomes.[2,3] DNA cytosine methylation (5-methylcytosine, 5mC) is the most extensively characterized epigenetic modification, and it plays sophisticated roles in regulating gene expression in mammalian cells.[4] The formation of 5mC by DNA methyltransferases (DNMTs) and oxidation by ten-eleven translocation (TET) proteins diversify the genomes and shape the chromatin landscape.[5,6] TET proteins catalyze the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) and 5-formylcytosine (5fC), and finally to

[a] *Sauvage Center for Molecular Sciences, Department of Chemistry, Wuhan University, Wuhan 430072, P. R. China. E-mail: bfyuan@whu.edu.cn; Fax: +86-27-68755595; Tel: +86-27-68755595*
[b] *School of Health Sciences, Wuhan University, Wuhan 430071, China*

**Yi Dai**

*Yi Dai is a master's student in the Department of Chemistry, Wuhan University, China. She received her BSc degree in Chemistry from the Beijing Normal University in China. Presently, she is working under the guidance of Prof. Bi-Feng Yuan, focusing on the development and applications of analytical methods for the study of nucleic acid modifications.*

**Bi-Feng Yuan**

*Bi-Feng Yuan studied biochemistry and biophysics at Wuhan University in China, where he received his BSc and PhD degrees in 2001 and 2006, respectively. From 2006–2007 and 2007–2010, he worked as a postdoctoral researcher in the Department of Biological Sciences at the National University of Singapore and in the Department of Chemistry at the University of California Riverside, respectively. He joined Wuhan University in 2011 and is currently a professor of chemistry. His research focuses on the development and applications of new analytical techniques in the investigation of the occurrences, locations, and biological functions of nucleic acid modifications.*


© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

5-carboxycytosine (5caC).[7] The 5fC and 5caC undergo base excision repair to restore the unmodified cytosines.[8] Apart from being intermediates in the active demethylation pathway, 5hmC, 5fC, and 5caC are also considered to function as epigenetic modifications with potential regulatory roles in gene expression.[9]

In addition to DNA cytosine modifications, many other types of modifications also exist in the genomes of living organisms.[10,11] $N^6$-Methyladenosine (6mA) is a natural DNA modification existing in both prokaryotes and eukaryotes.[12] 6mA functions in the restriction–modification (RM) systems and gene regulatory processes in prokaryotes.[13] In 2015, it was reported that 6mA was presented in the genomes of *Chlamydomonas reinhardtii*,[14] *Caenorhabditis elegans*,[15] and *Drosophila melanogaster*.[16] Later on, 6mA was detected in a variety of living organisms of vertebrates and mammals.[17–21] The most recent studies demonstrated that 6mA mainly existed in the mitochondrial DNA[22] and originated from the degraded RNA nucleoside of $N^6$-methyladenosine ($m^6A$), which was further processed through the nucleotide-salvage pathway and incorporated into genomic DNA by DNA polymerases.[23,24]

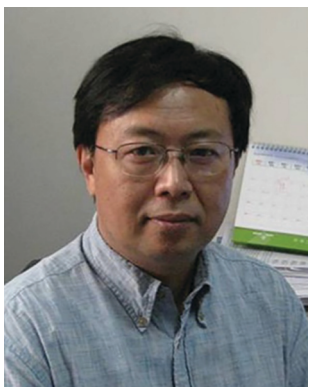
5-Hydroxymethyluracil (5hmU) is a thymine modification existing in the genomes of many living organisms, such as leishmania,[25] bacteriophages,[26] dinoflagellates,[27] and eukaryotes.[28] TET enzymes can oxidize thymine to form 5hmU and 5-formyluracil (5fU) in mammalian genomic DNA.[28] In some parasite genomes, the TET homologue of JBP (J-binding protein) can oxidize thymine to 5hmU, which is further converted to a new nucleobase β-D-glucosyl-5-hydroxymethyluracil (base J) through J-GT (base J glycosyltransferase).[29,30] Recently, another TET homologue of the 5mC-modifying enzyme (CMD1) from the alga *Chlamydomonas reinhardtii* has been reported to catalyze the conversion of 5mC to generate a new modified nucleobase, *i.e.*, 5-glyceryl-methylcytosine (5-glyceryl-mC), which played critical roles in the photoprotective process.[31]

Reactive oxygen species (ROS) produced by cellular aerobic metabolism can oxidize DNA, and 8-oxo-7,8-dihydroguanine (OG) is the main oxidation product.[32] OG could lead to G-to-T transversion mutation.[33,34] OG in genomes can be repaired through the base excision repair (BER) mechanism.[35] However, OG has recently been considered to have regulatory and epigenetic properties in modulating gene expression.[36,37] For example, the increased OG level in genomes was closely related to the increased gene expression through the BER pathway.[38] OG could also activate mRNA synthesis by promoting the formation of G-quadruplex in the DNA promoter.[39]

Phosphorothioate (PT) modification on DNA has been discovered in bacteria and archaea.[40] So far, PT modification has not been discovered in the DNA of eukaryotes. In PT modification, the non-bridging oxygen in the phosphate group of DNA strands is replaced by sulfur. For PT modifications, the *dndABCDE* genes function as part of RM systems.[41] The PT modifications in the DNA of bacteria and archaea involve regulating gene expression and cellular stress response, and maintaining cellular redox homeostasis.[42] In addition, some 7-deazaguanine modifications, such as 2′-deoxy-7-amido-7-deazaguanosine (dADG), 2′-deoxy-7-cyano-7-deazaguanine ($dPreQ_0$) and 2′-deoxy-7-aminomethyl-7-deazaguanine ($dPreQ_1$), were found to be present in the DNA of bacteria and phage.[43,44] These DNA modifications are considered to function in protecting the phage DNA from cleavage by host restriction enzymes.

The innovation of analytical methods has greatly stimulated the research of DNA modifications. In the past few decades, a variety of methods have been developed for the genome-wide quantification and mapping of DNA modifications. With the recent advances in analytical methods and techniques, a series of new endogenous DNA modifications have been discovered. In addition, many previously discovered DNA modifications have been endowed with epigenetic-like properties in regulating biological processes. These methods have enabled the establishment of the correlation between the dynamic changes of DNA modifications and various human diseases, including cancers, imprinting disorders, neurological disorders, cardiovascular diseases, and developmental diseases.[4,45] In addition, the interplay between the epigenetic DNA modifications and metabolism emerges as a new research direction by virtue of these sophisticated analytical methods.[46,47]

Revealing the functions of DNA modifications depends on the sensitive detection, accurate quantification, and mapping of these modifications in genomes.[48–50] Here, we review the recent advances in analytical methods for the quantification and mapping of DNA modifications, including 5mC, 5hmC, 5fC, 5caC, 6mA, 5hmU, 5fU, base J, 5-glyceryl-mC, $N^4$-methylcytosine (4mC), $N^6$-hydroxymethyladenosine (6hmA), OG, PT, and 7-deazaguanine modifications (Fig. 1). We discuss the principles, advantages, and limitations of these established methods.



**Yu-Qi Feng**

*Yu-Qi Feng received his BSc and master's degrees from Lanzhou University in China in 1982 and 1985, respectively. From 1986 to 1991, he worked in Central China Normal University in China. He then studied chemistry at Chiba University in Japan, where he received his PhD degree in 1996. He joined Wuhan University in 1996 and became a full professor in 2000. His research focuses on sample pretreatment and LC/MS-based metabolomics.*

## 2 Methods for quantification of DNA modifications

The reported methods for the detection of overall DNA modifications mainly include liquid chromatography (LC), liquid chromatography/mass spectrometry (LC/MS), gas chromatography/mass spectrometry (GC/MS), chemical derivatization-mass

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

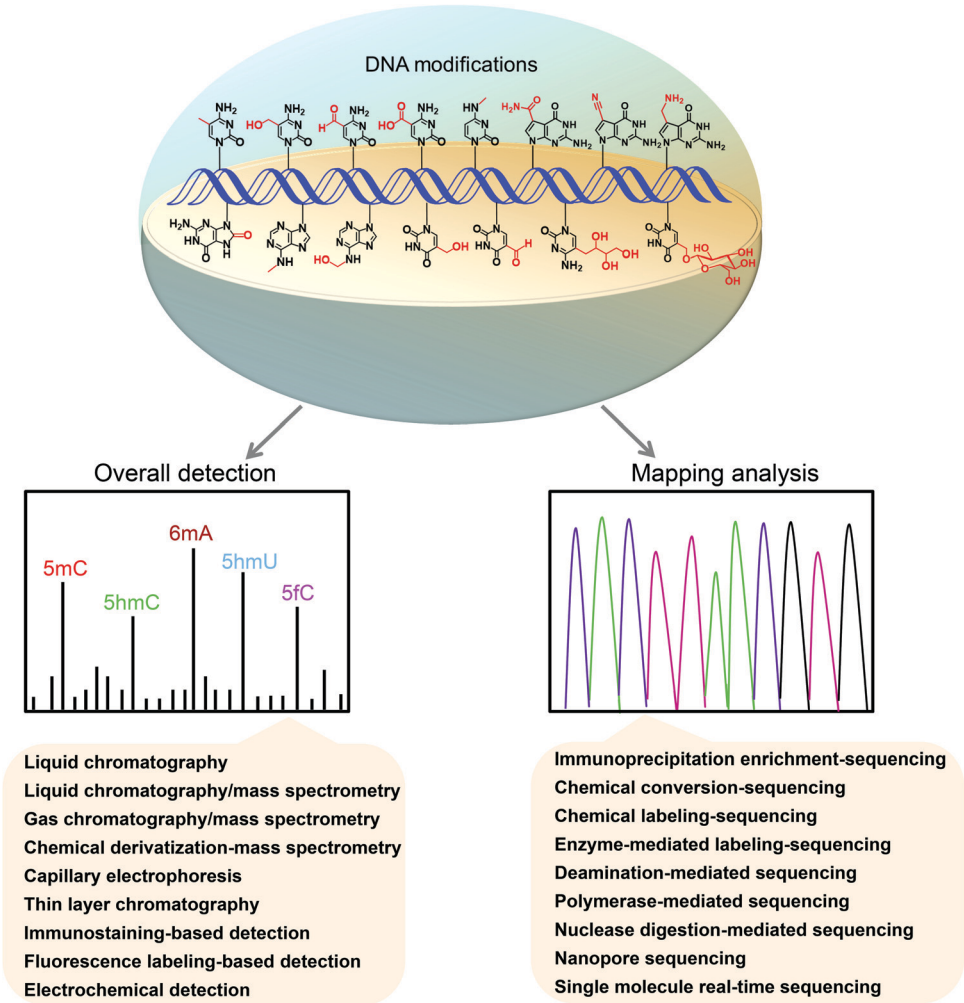



Fig. 1 Schematic illustration for the overall detection and mapping of DNA modifications. The chemical structures shown in DNA strand represent 5mC, 5hmC, 5fC, 5caC, 6mA, 5hmU, 5fU, base J, 5-glyceryl-mC, 4mC, 6hmA, OG, dADG, $dPreQ_0$ and $dPreQ_1$.

spectrometry, capillary electrophoresis (CE), thin layer chromatography (TLC), immunostaining-based detection, fluorescence labeling-based detection, and electrochemical detection.

### 2.1 Liquid chromatography

Determination of DNA modifications by liquid chromatography (LC) is based on the detection of different nucleobases, 2′-deoxynucleosides or 2′-deoxynucleotides, which are obtained from the enzymatic or chemical hydrolysis of DNA. These DNA components are typically separated by reversed-phase LC,[51] and then detected using a photometric or electrochemical detector. Thus, the baseline separation of the DNA components is essential since the determination relies on the chromatographic separation to distinguish analytes.

It should be noted that if DNA is hydrolyzed by formic acid to release nucleobases for LC detection, the complete elimination of RNA contamination in isolated DNA is required because the nucleobases released from RNA, such as 5-methylcytosine and $N^6$-methyladenine, will affect the accurate quantification of the corresponding DNA modifications. As for the LC-based detection of DNA modifications, a relatively large amount of genomic DNA (∼1–50 μg) is typically needed owing to the low detection sensitivity of the method. Moreover, the LC-based detection of DNA modifications is less confirmative than the mass spectrometry-based detection.

### 2.2 Liquid chromatography/mass spectrometry

Mass spectrometry (MS) has been widely used to detect and quantify nucleic acid modifications as MS exhibits high detection sensitivity and good capability to identify compounds.[52–58] LC/MS is the preferred method for the sensitive detection and quantification of the global levels of modifications in genomic DNA. Typically, DNA is enzymatically digested into 2′-deoxynucleosides or 2′-deoxynucleotides and then determined by LC/MS. LC/MS now plays a central role in the discovery, identification and quantification of DNA modifications across different cells, tissues and organisms.[59,60] LC/MS provides the nucleoside composition but not the sequence context of DNA modifications in genomes.

**2.2.1 LC/MS.** LC/MS with the multiple reaction monitoring (MRM) detection mode is frequently employed to detect DNA

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

modifications due to the inherent high sensitivity of this detection mode. LC/MS with MRM detection can usually achieve a sub-femtomole level of limit of detection (LOD) towards DNA modifications. We realized the analysis of 5mC in a single cell using liquid chromatography tandem mass spectrometry (LC-MS/MS) with the MRM detection mode.[61] The PT modifications in prokaryotic and eukaryotic genomes were quantitatively measured by LC-MS/MS. The results showed that the PT modifications were as low as 1 per $10^6$ nt for d($T_{PS}$T) and 2 per $10^8$ nt for d($C_{PS}$T).[62,63] The LC-MS/MS analysis showed that dADG, dPre$Q_0$, and dPre$Q_1$ modifications were in the range of 6–790 (per $10^6$ nt).[44] Ultra-performance liquid chromatography (UPLC) has also been employed in the LC-MS/MS analysis, thus enabling short chromatographic runs in the analysis of 5mC and 6mA in DNA.[18,64,65] Using a stable isotope labeled (SIL) analogue of the nucleoside as an internal standard can calibrate the detection variation caused by the matrix effect and sample pretreatment during the measurements of DNA modifications. Stable isotope labeling coupled with LC/MS methods has been developed for the accurate measurement of DNA modifications, such as 5mC, 5hmC (LOD, 0.056 fmol), 5fC (LOD, 0.098 fmol), 5caC (LOD, 0.14 fmol), 5hmU (LOD, 80 fmol), base J, 6mA (LOQ, 1.6 fmol), PT and 6hmA.[66–71] In addition, stable isotope labeling also benefits the qualitative analysis of DNA modifications.

A vitamin-C-derived DNA modification (5-glyceryl-mC) has been recently discovered in the genome of the green alga *Chlamydomonas reinhardtii* by LC/MS analysis.[31] The measured level of 5-glyceryl-mC was about 10 modifications per $10^6$ dC in the genomic DNA of wild-type *Chlamydomonas reinhardtii*. The discovery of a new modification of 5-glyceryl-mC highlights the structural diversity of DNA associated with environmental adaptation.

**2.2.2 High-resolution mass spectrometry analysis.** High-resolution mass spectrometry (HRMS) can provide accurate mass measurements and plays an important role in deciphering the structural information of compounds. LC coupled with HRMS (LC-HRMS) has been frequently used in the determination of DNA modifications. For example, LC-HRMS enabled the discovery of 5fC and 5caC in the DNA of mouse embryonic stem cells,[72,73] providing confirmatory evidence of the 5mC oxidative demethylation *via* 5fC and 5caC. Using high-resolution quadrupole TOF MS (qTOF-MS), 5mC and 5hmC in genomic DNA can be simultaneously detected,[74] which allowed the sensitive determination of low contents of 5mC and 5hmC with 2 ng of DNA. This study demonstrated that the content of 5hmC in hepatocellular carcinoma (HCC) tumor tissues was decreased compared to that of the tumor adjacent tissues. The dysregulation of 5hmC in cancer tissues has now become a common phenomenon in a variety of cancers.[75–78]

**2.2.3 Inductively coupled plasma mass spectrometry analysis.** The use of external elemental tags and inductively coupled plasma-mass spectrometry (ICP-MS) detection has become a new strategy in analytical chemistry.[79] This method is based on the exceptional features of elemental mass spectrometry in terms of sensitivity, selectivity, and multi-elemental and isotopic capabilities, which can enhance the quantification of labeled biomolecules. The hyphenation of LC with ICP-MS (LC-ICP-MS) has been developed for evaluating DNA methylation.[80] The reversed-phase separation of 5-methyl-2′-deoxycytidine-3′-monophosphate (5mCMP) with specific ICP-MS detection on $^{31}$P was established to detect 5mC.[80] In addition, potassium acid ($K_2OsO_4$) was used to selectively label 5mC in single-stranded DNA (ssDNA) in the presence of strong oxidant $K_3Fe(CN)_6$ and *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (TEMED).[80] Then, size-exclusion chromatography-ICP-MS was used to detect $^{31}$P (proportional to the total amount of DNA) and $^{189}$Os (equivalent to 5mC). These methods were successfully applied in analyzing 5mC from the DNA of salmon testes and commercial oligonucleotides, which provided a proof-of-concept that ICP-MS was a useful and complementary platform to the LC/MS in the detection of DNA modifications.

**2.2.4 Metabolic isotope tracing by mass spectrometry.** Metabolic isotope tracing is a useful strategy that has been widely used in the detection and confirmation of endogenous biomolecules.[81] The *in vivo* metabolic pathways and dynamics of DNA modifications can be revealed by the combination of isotope tracing and LC-MS/MS analysis. It has been known that ATP and L-methionine could be converted to *S*-adenosyl-L-methionine (SAM) by methionine adenosyltransferase, and SAM is a universal reagent for the methylation of nucleic acids. Therefore, by feeding cells with SIL L-methionine, the source of methyl groups for DNA methylation and the dynamics of 5mC and 5hmC were uncovered.[82] Similarly, various isotopologues of modified bases have been utilized to elucidate the mechanisms of the formation and removal of modifications, such as 5fC,[83] 5hmU,[28] 6mA,[84] and 5caC.[85,86]

### 2.3 Gas chromatography/mass spectrometry

Compared to LC/MS, gas chromatography/mass spectrometry (GC/MS) has the advantage of providing better chromatographic separation. Nevertheless, GC/MS requires the conversion of analytes into volatile derivatives before separation.[87] In this regard, nucleobases can only be analyzed after converting them into volatile counterparts through derivatization.

For the GC/MS analysis of 5mC, DNA is typically hydrolyzed using aqueous formic acid (88%), and the resulting nucleobases are then derivatized with appropriate reagents, such as *N*-methyl-*N*-(*tert*-butyldimethylsilyl)trifluoroacetamide with 1% *tert*-butyldimethyl-chlorosilane[88] and *N*,*O*-bis-(trimethylsilyl)-trifluoroacetamide (BSTFA) with 1% cholorotrimethylsilane.[89] The results showed that the LOD value of 5mC was as low as 0.8 pg (6.4 fmol).[89] GC/MS was also used to detect OG from human breast tissues and lymphocytes.[90] However, it was found that the OG content measured by GC/MS was higher than that by other methods. The gas-phase derivatization step could cause the oxidation of guanine, which may lead to increased OG due to the artificially formed OG during sample preparation.

### 2.4 Chemical derivatization-mass spectrometry

DNA modifications typically present in low abundance *in vivo*, and it is a challenging task to detect some rare DNA modifications

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 1 Summary of the detection sensitivities of various DNA modifications analyzed by chemical derivatization-mass spectrometry. LOD, limit of detection

| Derivatization reagents | Modifications | LOD (fmol) | Ref. |
|---|---|---|---|
| 2-Bromo-1-(4-dimethyl aminophenyl)-ethanone (BDAPE) | 5mC | 0.10 | 98 |
| | 5hmC | 0.06 | |
| | 5fC | 0.11 | |
| | 5caC | 0.23 | |
| 2-Bromo-1-(4-diethylaminophenyl)-ethanone (BDEPE) | 5mC | 0.06 | 99 |
| | 5hmC | 0.07 | |
| | 5fC | 0.10 | |
| | 5caC | 0.08 | |
| 2-Bromo-1-(4-1-pyrrolidinylphenyl)-ethanone (BPPE) | 5mC | 0.43 | |
| | 5hmC | 0.69 | |
| | 5fC | 0.79 | |
| | 5caC | 1.21 | |
| 3-Bromoacetonyltrimethylammonium bromide (BTA) | 5mC | 2.20 | |
| | 5hmC | 4.72 | |
| | 5fC | 6.22 | |
| | 5caC | 0.53 | |
| ω-Bromoacetonylpyridinium bromide (BPB) | 5mC | 5.25 | |
| | 5hmC | 2.72 | |
| | 5fC | 16.59 | |
| | 5caC | 4.41 | |
| 4-(Dimethylamino) benzoic anhydride | 5mC | 2.24 | 100 |
| | 5hmC | 2.53 | |
| | 5fC | 2.54 | |
| | 5caC | 1.27 | |
| Uridine diphosphoglucose (UDP-glucose) | 5hmC | 6.7 | 102 |
| | 5gmC | 1.5 | |
| Girard's T reagent | 5fU | 3.3 | 103 |
| | 5fC | 0.09 | 104 |
| | 5caC | 0.77 | |
| | 5fC | 0.08 | 105 |
| | 5fU | 0.14 | |
| Girard's D reagent | 5fC | 0.15 | 104 |
| | 5caC | 0.42 | |
| Girard's P reagent | 5fC | 0.03 | 104 |
| | 5caC | 0.75 | |
| | 5fC | 0.03 | 105 |
| | 5fU | 0.05 | |
| 4-(2-(Trimethylammonio)ethoxy)benzenaminium halide (4-APC) | 5fC | 0.05 | 105 |
| | 5fU | 0.09 | |
| Hydrazino-*s*-triazine based glycan labelling reagents i-$Pr_2N$ | 5fC | 0.01 | 106 |
| | 5caC | 0.025 | |
| Hydrazino-*s*-triazine based glycan labelling reagents $Me_2N$ | 5fC | 0.05 | |
| Hydrazino-*s*-triazine based glycan labelling reagents $Et_2N$ | 5fC | 0.0125 | |
| Dansylhydrazine (DNSH) | 5hmC | 0.04 | 107 |
| 2-Bromoacetophenone | 5mC | 22.7 | 131 |

by direct LC/MS analysis. In this respect, the approaches based on chemical labeling in conjugation with LC/MS analysis were developed to analyze the low levels of compounds.[91–97] The tagged group that was added to the target modifications from labeling reagents can endow the desired properties of DNA modifications, which can be utilized to achieve the improved ionization efficiencies of DNA modifications during mass spectrometry analysis (summarized in Table 1).

We and others previously utilized 2-bromo-1-(4-dimethylaminophenyl)-ethanone (BDAPE), 2-bromo-1-(4-diethylaminophenyl)-ethanone (BDEPE), and 4-(dimethylamino) benzoic anhydride to simultaneously label 5mC, 5hmC, 5fC, and 5caC in DNA,[98–101] which led to a 313-fold increase of the detection sensitivities of DNA modifications. Similarly, some other chemical derivatization strategies were developed to enable the sensitive detection of DNA modifications, such as the glycosylation of 5hmC by T4 β-glucosyltransferase,[102] derivatization of 5fC, 5caC and 5fU by Girard's reagents (Girard D, T, and P)[103–105] and hydrazine reagents ($Me_2N$, $Et_2N$, and i-$Pr_2N$),[106] and derivatization of 5fC by dansylhydrazine (DNSH).[107] In addition, 8-(diazomethyl)quinoline (8-DMQ) was used to derivatize the phosphate group of NTPs and modified NTPs, which resulted in the discovery of a series of new modifications occurring in NTPs.[108]

### 2.5 Capillary electrophoresis

The separation of analytes by capillary electrophoresis (CE) is mainly based on the equilibrium of charged analytes between the stationary phase and the mobile phase. Several detection techniques, including UV-absorbance, MS, and laser-induced fluorescence (LIF) detection, can be coupled with CE to determine the DNA modifications from clinical samples and plant tissues. CE-LIF and CE-MS generally can offer better detection sensitivity than CE-UV in detecting 5mC and 5hmC in genomic

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

DNA. CE-MS usually provides better detection sensitivity than CE-UV, and 5mC and 5hmC could be easily detected using only subnanogram DNA samples.[109] CE-LIF has been used to quantify 6mA, with the detection limit being as low as 1.4 amol.[110] Our group developed a sensitive 5mC detection method based on capillary liquid chromatography using a hyper-cross-linked polymer monolithic column carrying phenyl and quaternary ammonium groups.[111] The LOD of 5mC was 0.014 pmol, which is comparable to that obtained by MS analysis.

The CE method has been used to address many interesting questions, including the examination of DNA methylation in cancer cell lines,[112] the investigation of the association between the DNA methylation and clinicopathological data from leukemia patients,[113] and the evaluation of the *in vivo* activity of mammalian DNA methyltransferases.[114] In the above-described studies, the typical sample amounts required were in the range of 0.1–10 μg of genomic DNA. Although the CE method provides higher separation efficiency than the LC method, the separation reproducibility can be affected by slight variations.

### 2.6 Thin layer chromatography

Thin layer chromatography (TLC) is based on the differences in net charge, polarity, and hydrophobicity between analytes.[115] The genomic DNA is enzymatically hydrolyzed to nucleosides and then labeled with [$^{32}$P]ATP by T4 polynucleotide kinase, followed by separation on cellulose TLC plates. The relative intensity of the spots can be determined using a phosphorimager. Many DNA modifications, such as 5hmC[116,117] and 5caC,[8] were initially discovered in the genomic DNA of mammalian cells by TLC and then further confirmed by HRMS.

Compared to one-dimensional (1D) chromatography, two-dimensional (2D) separation can effectively improve the separation capability. 2D-TLC has been widely utilized to detect various DNA modifications such as 5mC,[118] base J,[119] and OG.[120] It is simple to identify and quantify $^{32}$P-labeled modified nucleotides down to the femtomole level. For the TLC-based detection of DNA modifications, there is no need for sophisticated instrumentation. However, the analytical procedure with radioactive labeling by TLC is relatively time-consuming.

### 2.7 Immunostaining-based detection

Immunostaining has been widely used to evaluate the global levels of various DNA modifications. Specific antibodies can recognize 5mC,[121] 5hmC,[122] 5fC,[123] 5caC,[123] 6mA,[124] base J[125] and OG[126] for visualization. Although this method is simple, it has the disadvantages of potential low antibody specificity and bias towards DNA modifications. Immunostaining methods generally provide semi-quantitative information.

Enzyme-linked immunosorbent assay (ELISA) coupled with electrochemiluminescence (ECL) detection was used to detect 5mC[127] and OG.[128] The developed method employed an anti-5mC antibody labeled with acetylcholinesterase, which can convert acetylthiocholine (substrate) to thiocholine (product). The thiocholine was accumulated on a gold electrode surface, making it possible to observe bright and distinctive ECL by applying a potential to the gold electrode. The detection limit of 5mC was 0.18 pmol.

### 2.8 Fluorescence labeling-based detection

Fluorescence labeling is a well-established strategy in the detection of DNA modifications. High detection sensitivity can be achieved by selecting suitable fluorescent reagents.[129,130] For example, 2-bromoacetophenone was used to selectively label cytosine moieties from the DNA of salmon testes and nucleic acids from plants, human blood, and earthworms, followed by the detection of these derivatives using LC with a spectrofluorimetric detector.[131] The LOD of 5mC can reach 16.6 fmol. In general, the fluorescent moiety can be efficiently attached to aldehyde-containing modifications (such as 5fC and 5fU) by amine hydrazine or hydroxylamine through aldol-type condensation reactions,[132–138] or by the direct condensation of aminobenzaldehyde with cyano reagents.[139,140] In addition, by the chemical conversion of hydroxymethyl group to an aldehyde group, the quantitative detection of 5hmC in different mouse tissues has been achieved using a similar strategy.[141] The linear correlation between the concentration of labeled modifications and the fluorescence intensity allows us to quantitatively determine the levels of DNA modifications. It should be noted that some developed methods have been successfully used to analyze the DNA modifications of biological samples;[132,136,138,139,142] however, some methods were established using synthesized DNA carrying modifications.[133–135,137,140]

### 2.9 Electrochemical detection

Electrochemical biosensing strategies can offer fast and sensitive detection towards DNA modifications.[143,144] Based on the decrease in charge density between the C·G base pairs caused by DNA methylation, a simple electrochemical strategy was reported for the continuous monitoring of the dynamic DNA methylation process.[145] Although most of the studies can only detect the overall level of 5mC in DNA, the analysis of single DNA methylation sites also has been successfully accomplished.[146]

Signal amplification is a prevalent strategy to increase the detection sensitivity of electrochemical biosensors. The signal amplification technologies based on enzymatic amplification[147] or redox cycles[148] were reported to realize the sensitive quantification of 5mC and 5hmC in genomic DNA. Additionally, analytical approaches for signal amplification through carbon nanocomposite materials,[149] new nano-microspheres,[150] gold nanoparticles,[151] and other novel nanomaterials[152] were developed for the analysis of DNA modifications of various biological samples. Using these methods, the detection limit of 5hmC was as low as 9.06 fM,[148] while the detection limit of OG was 1 pM.[151]

## 3 Methods for mapping DNA modifications

Methods for the overall detection of DNA modifications require the digestion or release of modifications from DNA before

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

measurements, which will not provide information of the positions of modifications within the DNA sequence. Rapidly improving technologies, especially the next-generation DNA sequencing, revolutionize the study of genome-wide mapping of DNA modifications, allowing us to study the site-specific dynamic changes of epigenetic modifications.[153] A variety of methods have been established for the genome-wide mapping of DNA modifications, which is essential and critical to reveal the biological functions of DNA modifications (Table 2).

### 3.1 Immunoprecipitation enrichment-sequencing

A number of immunoprecipitation-based enrichment strategies combined with sequencing have been developed for mapping DNA modifications in genomes.[49] These methods typically include the enrichment of DNA fragments that contain the modified base, sequencing of the enriched fragments, and aligning of the sequence reads to the reference genomes. The position of a modification in genomic DNA can be approximately mapped by fragmenting the DNA into short pieces, and the peak analysis could improve the mapping resolution for modifications.

The immunoprecipitation-mediated enrichment of DNA fragments using antibodies or specific affinity binding proteins has been demonstrated to be very useful for mapping DNA modifications. The immunoprecipitation of methylated DNA is commonly achieved by the isolation of methylated DNA fragments using antibodies for 5mC (MeDIP-seq)[154] or the isolation of DNA fragments containing methylated CpG dinucleotides using methyl-CpG binding domain (MBD) proteins (MBD-seq).[155,156] A secondary antibody captures the methylated DNA-antibody/protein complex (MethylCap-seq, MIRA-seq),[157,158] which can then be further analyzed using PCR, array-based methods, or sequencing technologies. Similarly, affinity-based enrichment of 5hmC,[159] 5fC,[160] 6mA,[161,162] base J,[163] and OG,[164] followed by sequencing, were also developed to map these modifications in genomes of various species.

Although immunoprecipitation-mediated sequencing is straightforward, it does not offer base-resolution information of DNA modifications. In addition, the results of independent DIP-seq studies often show considerable variations between the profiles of the identical genomes and between the profiles obtained by alternative methods, which might be attributed to the low specificity of antibodies and the intrinsic affinity of IgG for short unmodified DNA repeats.[165]

### 3.2 Chemical conversion-sequencing

Specific conversion of modified nucleosides or unmodified nucleosides by chemical treatments can lead to the subsequent characteristic sequencing events for modifications, which therefore can be employed for mapping modified nucleosides in DNA. The selective and efficient chemical conversion is the critical issue in this method.

**3.2.1 Bisulfite sequencing.** Bisulfite sequencing (BS-seq) is the gold standard for the genome-wide mapping of 5mC at single-base resolution.[166] In BS-seq, cytosine, 5fC, and 5caC undergo deamination by bisulfite treatment and are read as thymines, while both 5mC and 5hmC are resistant to deamination by bisulfite treatment and are read as cytosines.[167,168] Therefore, traditional bisulfite sequencing cannot distinguish 5fC or 5caC from cytosine, nor can it differentiate 5hmC from 5mC.

Some methods have been developed by combining chemical labeling and bisulfite treatment to achieve the genome-wide mapping of 5hmC, 5fC, and 5caC (Fig. 2A). Oxidative bisulfite sequencing (oxBS-seq)[167] and TET-assisted bisulfite sequencing (TAB-Seq)[168] methods were established to map 5hmC in genomic DNA. The oxBS-seq approach is based on the specific oxidation of 5hmC by $KRuO_4$ to produce 5fC, which can be converted to uracil under bisulfite treatment.[167] With the oxBS-seq approach, 5hmC is read as thymine, but 5mC is read as cytosine in DNA. In the TAB-seq approach, β-GT specifically transfers a glycosyl group to 5hmC to produce β-glucosyl-5-hydroxymethylcytosine (5gmC), which resists the oxidation by TET proteins.[168] But TET proteins oxidize 5mC to 5fC and 5caC, both of which are read as thymines by bisulfite sequencing. The remaining cytosine signals come from the glycosylated 5hmC, offering a strategy for the single-base resolution mapping of 5hmC. A mirror bisulfite sequencing for 5hmC detection at a single CpG site with glycosylation by β-GT, methylation by M.SssI methylase and bisulfite conversion also achieved the site-specific analysis of 5hmC at the CpG sites of DNA.[169]

Reduced bisulfite sequencing (redBS-seq)[170] was developed to map 5fC based on the selective reduction of 5fC to 5hmC by sodium borohydride combined with bisulfite sequencing. 5fC sites could be detected by comparing the bisulfite-treated reduced and non-reduced DNA. In addition, the fCAB-seq approach[171] used *O*-ethylhydroxylamine ($EtONH_2$) to convert 5fC to oxime, which resists deamination by bisulfite treatment. Comparing the results of fCAB-seq (where 5fC is read as C) with those of BS-seq (where 5fC is read as T) enables the mapping of 5fC at single-base resolution. For 5caC mapping analysis, a chemical modification-assisted bisulfite sequencing (CAB-seq) method was established.[172] The carboxyl group of 5caC reacts with the primary amine group, thereby protecting the labeled 5caC from deamination by bisulfite treatment. Based on CAB-Seq, a DNA immunoprecipitation-coupled CAB-Seq (DIP-CAB-Seq) was further developed and successfully used in the mapping of 5fC and 5caC in the genomes of mouse embryonic stem cells.[173] In addition, the combination of bisulfite treatment with the use of M.SssI methyltransferase could achieve the base-resolution mapping of 5fC and 5caC in DNA (MAB-seq and caMAB-seq).[174,175]

In addition to the 5-substituted cytosine modifications, 4mC was also found to be present in the DNA of thermophilic bacteria and many bacterial mesophiles.[176] Traditional bisulfite sequencing cannot distinguish 4mC from 5mC because both 4mC and 5mC resist deamination and will then be read as cytosines. A 4mC-Tet-assisted-bisulfite-sequencing (4mC-TAB-seq) approach was established for the genome-wide mapping of 4mC.[177] This 4mC-TAB-seq approach can accurately identify the 4mC sites without interference from 5mC. By coupling the oxidation of all 5mC to 5caC mediated by excessive TET proteins under optimized

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 2 Summary of methods for mapping DNA modifications

| Modifications | Methods | Treatment | Resolution | Principles of sequencing | Ref. |
|---|---|---|---|---|---|
| 5mC, 5hmC | BS-seq | Bisulfite treatment | Single-base resolution | C, 5fC and 5caC undergo deamination and all of them are read as T, while 5mC and 5hmC are resistant to deamination and are read as C. | 166 |
| | TAPS | Oxidation of 5mC, 5hmC and 5fC by TET proteins; reduction of 5caC by pyridine borane | | TET oxidizes 5mC, 5hmC and 5fC to 5caC, then pyridine borane reduces 5caC to DHU, which is read as T in PCR and thereby realize the C-to-T conversion of 5mC and 5hmC. | 182 |
| | EM-seq | Oxidation of 5mC by TET2 and glycosylation of 5hmC by β-GT; deamination of C by APOBEC3A | | TET2 oxidation and β-GT glycosylation protect 5mC, 5hmC and 5fC from APOBEC3A deamination; only C is converted to U, which is read as T in sequencing. 5mC, 5hmC and 5fC are read as C. | 204 |
| 5mC | TAPSβ | Glycosylation of 5hmC by β-GT; TET oxidation and pyridine borane reduction of 5mC | Single-base resolution | β-GT glycosylation protects 5hmC from TET oxidation and pyridine borane reduction; only 5mC is converted to DHU, which is read as T in sequencing. | 182 |
| 5hmC | oxBS-seq | Oxidation of 5hmC by $KRuO_4$ | Single-base resolution | $KRuO_4$ oxidizes 5hmC to 5fC, which is read as T in BS-seq, while 5mC is read as C. | 167 |
| | TAB-seq | Glycosylation of 5hmC by β-GT and oxidation of 5mC by TET proteins | | Glycosylation of 5hmC to 5gmC by β-GT; TET oxidizes 5mC to 5fC and 5caC, both of which are read as T in BS-seq, while the original 5hmC is read as C. | 168 |
| | hmC-CATCH | Labeling of 5fC by $EtONH_2$; oxidation of 5hmC by $K_2RuO_4$ and labeling by 1,3-indandione (AI) | | Protection of 5fC by $EtONH_2$; $K_2RuO_4$ oxidizes 5hmC to 5fC followed by AI labeling, which induces a C-to-T transition in sequencing. | 178 |
| | CAM-Seq | $KRuO_4$ oxidation and azi-BP labeling of 5hmC | | Protection of 5fC by hydroxylamine; $KRuO_4$ oxidizes 5hmC to 5fC followed by azi-BP labeling, which induces C-to-T conversion in sequencing. | 180 |
| | CAPS | Oxidation of 5hmC by $KRuO_4$ and reduction by pyridine borane | | $KRuO_4$ oxidizes 5hmC to 5fC, then is reduced to DHU by pyridine borane. DHU is read as T in sequencing. | 182 |
| | hmC-seq | Oxidation of 5hmC by peroxotungstate | | Peroxotungstate converts 5hmC to trihydroxylated thymine ($^{th}$T), leading to a C-to-T transition in polymerase extension. | 181 |
| | AMD-seq, ACE-seq | Glycosylation of 5hmC by β-GT; deamination of C by APOBEC3A | | β-GT glycosylation protects 5hmC from APOBEC3A deamination and 5gmC is read as C; C and 5mC are read as T. | 202 and 203 |
| | hmC-seq | Glycosylation of 5hmC by β-GT and precipitation by JBP1 | Genome-wide | β-GT converts 5hmC to 5gmC that can be pulled down by J-binding protein 1 coupled to magnetic beads. | 195 |
| | nano-hmC-Seal | Glycosylation of 5hmC by β-GT | | Labeling of 5hmC to 6-$N_3$-β-glucosyl-5hmC by β-GT, then a biotin tag is installed onto the azido group for pull down using click chemistry. | 197 |
| 5fC | redBS-seq | Reduction of 5fC by $NaBH_4$ | Single-base resolution | $NaBH_4$ reduces 5fC to 5hmC, which is converted to CMS by bisulfite treatment. The 5fC site is identified by comparing the output of redBS-seq (where 5fC is read as C) with that of BS-seq (where 5fC is read as T). | 170 |
| | fCAB-seq | Conversion of 5fC to oxime by $EtONH_2$ | | $EtONH_2$ converts 5fC to oxime, which resists deamination by bisulfite treatment. The 5fC site is identified by comparing the output of fCAB-seq (where 5fC is read as C) with that of BS-seq (where 5fC is read as T). | 171 |
| | fC-CET | Labeling of 5fC by 1,3-indandione (AI) | | Labeling of 5fC by AI enables a subsequent C-to-T transition in PCR. | 191 |
| | CLEVER-seq | Labeling of 5fC by malononitrile | | Labeling of 5fC by malononitrile induces a C-to-T conversion in sequencing. | 192 |
| | fC-seq | Labeling of 5fC by 2-(5-chlorobenzo[*d*]thiazol-2-yl)acetonitrile (CBAN) or azi-BP | | CBAN or azi-BP reacts with 5fC to generate an intramolecular cyclization nucleobase, leading to a C-to-T conversion in polymerase extension. | 142 and 193 |
| | fC-seq | Labeling of 5fC by *O*-(biotinylcarbazoylmethyl) hydroxylamine (ARP) | Genome-wide | Labeling of 5fC by ARP to form a biotinylated 5fC, which can be enriched by streptavidin-coated magnetic beads and then sequenced. | 160 |
| 5caC | CAB-seq | Labeling of 5caC by EDC-catalyzed xylene-based primary amine | Single-base resolution | Labeling of 5caC with xylene-based primary amine, which protects 5caC from deamination. The labeled 5caC is read as C in BS-seq. | 172 |
| | caMAB-seq | Reduction of 5fC by $NaBH_4$; methylation of C by M.SssI enzyme | | 5fC is reduced to 5hmC by $NaBH_4$; methylation of C by M.SssI. 5caC is sequenced as T in BS-seq, whereas C, 5mC, 5hmC and 5fC are read as C. | 174 |
| 5fC, 5caC | MAB-seq | Methylation of C by M.SssI enzyme | Single-base resolution | Methylation of CpG by M.SssI protects unmodified C from bisulfite conversion to U. 5fC/5caC is read as T in sequencing. | 175 |
| 4mC | 4mC-TAB-seq | Oxidation of 5mC by TET proteins | Single-base resolution | Oxidation of 5mC to 5caC by TET proteins followed by bisulfite treatment; 5mC is read as thymine, while 4mC is read as cytosine in sequencing. | 177 |

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

Table 2 (*continued*)

| Modifications | Methods | Treatment | Resolution | Principles of sequencing | Ref. |
|---|---|---|---|---|---|
| 5hmU | hmU-seq | Oxidation of 5hmU by $KRuO_4$ | Single-base resolution | Oxidation of 5hmU to 5fU by $KRuO_4$ induces a T-to-C base transition in polymerase extension. | 179 |
| | hmU-seq | Glycosylation of 5hmU by base J glucosyltransferase (J-GT) | Genome-wide | Labeling of 5hmU with $N_3$-glucose by J-GT followed by adding biotin tag and enrichment with streptavidin-coupled beads. | 199 |
| 5fU | fU-seq | Labeling of 5fU by azi-BIAN | Genome-wide | Azi-BIAN labeling of 5fU enables pull down of 5fU-containing DNA fragments for sequencing. | 188 |
| OG | OG-seq | Oxidation of OG by $K_2IrBr_6$ | Genome-wide | $K_2IrBr_6$ oxidizes OG to a covalent adduct of a primary-amine-terminated biotin, allowing for enrichment and sequencing. | 189 |
| 6mA | 6mA-seq | Deamination of adenine by sodium nitrite ($NaNO_2$) | Single-base resolution | $NaNO_2$ deaminates unmethylated adenines to hypoxanthine bases, which are read as guanine by polymerase and reverse transcriptase. 6mA site resists deamination and is read as adenine. | 194 |

bisulfite treatment conditions, 5caC is read as T. At the same time, 4mC is partially deaminated by bisulfite treatment, and about half of the 4mC sites are read as C. With this 4mC-TAB-seq approach, three 4mC-containing motifs were identified in the genomes of *Caldicellulosiruptor kristjanssonii*.

**3.2.2 Bisulfite-free sequencing.** The treatment condition by bisulfite is harsh, which could cause the degradation of the majority of input DNA. Thus, mild chemical treatment could provide an advantage in processing a small amount of DNA samples. In addition, bisulfite treatment converts abundant unmodified cytosines to thymines rather than non-canonical modifications, which reduces the sequence complexity of template DNA. Thus, BS-seq often involves low hybridization and mapping selectivity in sequencing. In this respect, bisulfite-free sequencing methods were developed for mapping DNA modifications (Fig. 2B).

The hmC-CATCH approach combines the $EtONH_2$ protection of endogenous 5fC with the selective oxidation of 5hmC to 5fC using $K_2RuO_4$.[178] 1,3-Indandione (AI) was used for the selective labeling of the converted 5fC, which would lead to a C-to-T transition in the subsequent PCR amplification and sequencing.[178] Similarly, the oxidation of 5hmU to form 5fU by $KRuO_4$ was used for the mapping of 5hmU.[179] In this approach, the chemical conversion of 5hmU to 5fU induced a T-to-C base transition in the subsequent polymerase extension, which was utilized to map 5hmU at single-base resolution. $KRuO_4$ was also used to transform 5hmC to 5fC followed by azi-BP labeling. The C-to-T transition was induced because the labeled 5fC lost the exocyclic 4-amino group, enabling the mapping of 5hmC in DNA.[180] Similarly, peroxotungstate can convert 5hmC to trihydroxylated thymine (thT), leading to a C-to-T transition during the polymerase extension and single-base resolution mapping of 5hmC.[181]

TET-assisted pyridine borane sequencing (TAPS)[182,183] includes the TET-mediated oxidation of 5mC, 5hmC and 5fC to 5caC and the pyridine borane reduction of 5caC to dihydrouracil (DHU). DHU was read as thymine in PCR, thereby realizing the C-to-T conversion of 5mC and 5hmC, while unmodified cytosines still were read as C. Thus, 5mC and 5hmC can be differentiated from cytosines, but 5mC and 5hmC cannot be differentiated from each other by TAPS. In this respect, 5hmC was glycosylated using β-GT, and the formed 5gmC resisted the TET oxidation and pyridine borane reduction. Thus, 5hmC was read as C, while 5mC was read as T, realizing the differentiation of 5mC and 5hmC (TAPSβ assay).[182]

Iodine shows high selectivity in the cleavage of PT modifications in DNA.[184] Thus, iodine-induced cleavage quantitative real-time PCR (IC-qPCR) has been developed to evaluate the frequency of PT modifications at a given site in bacterial DNA.[185] The high sensitivity of the IC-qPCR method with a LOQ of 5 copies or 10 copies per reaction suggested that the IC-qPCR method was suitable for quantifying PT modifications at low frequencies. In addition, the PT-IC-seq method was based on the iodine-induced selective cleavage at the PT sites, and a high-throughput sequencing was developed to map the PT modifications in genomic DNA.[186,187] In PT-IC-seq, iodine treatment could selectively induce DNA cleavage at the modified phosphodiester linkage, which produces the characteristic signatures. The PT sites can then be obtained from the cleavage sites that are examined by sequencing analysis. The mapping results for PT modifications indicated a remarkable target selection by the PT-modification proteins, and PT modifications could function in epigenetic control.

### 3.3 Chemical labeling-sequencing

**3.3.1 Chemical labeling-assisted enrichment.** In addition to the immunoprecipitation-mediated sequencing, chemical labeling-based strategies have been developed for the efficient enrichment of DNA fragments with modifications followed by sequencing. The critical point of chemical pull-down is the efficiency and selectivity of the chemical labeling reaction.

The reagent of (2-benzimidazolyl)-acetonitrile (azi-BIAN) was used to selectively label 5fU followed by the enrichment and genome-wide mapping of 5fU in human and mouse tissues (Fig. 3A).[188] The mapping results indicated that most of the 5fU sites mainly exist in the intergenic regions and introns. Similarly, aldehyde-reactive reagents of *O*-(biotinylcarbazoylmethyl) hydroxylamine (ARP)[160] and amine-terminal biotin[189] could achieve the selective labeling of 5fC and OG in DNA, respectively, which can be further enriched by streptavidin (Fig. 3B and C).

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

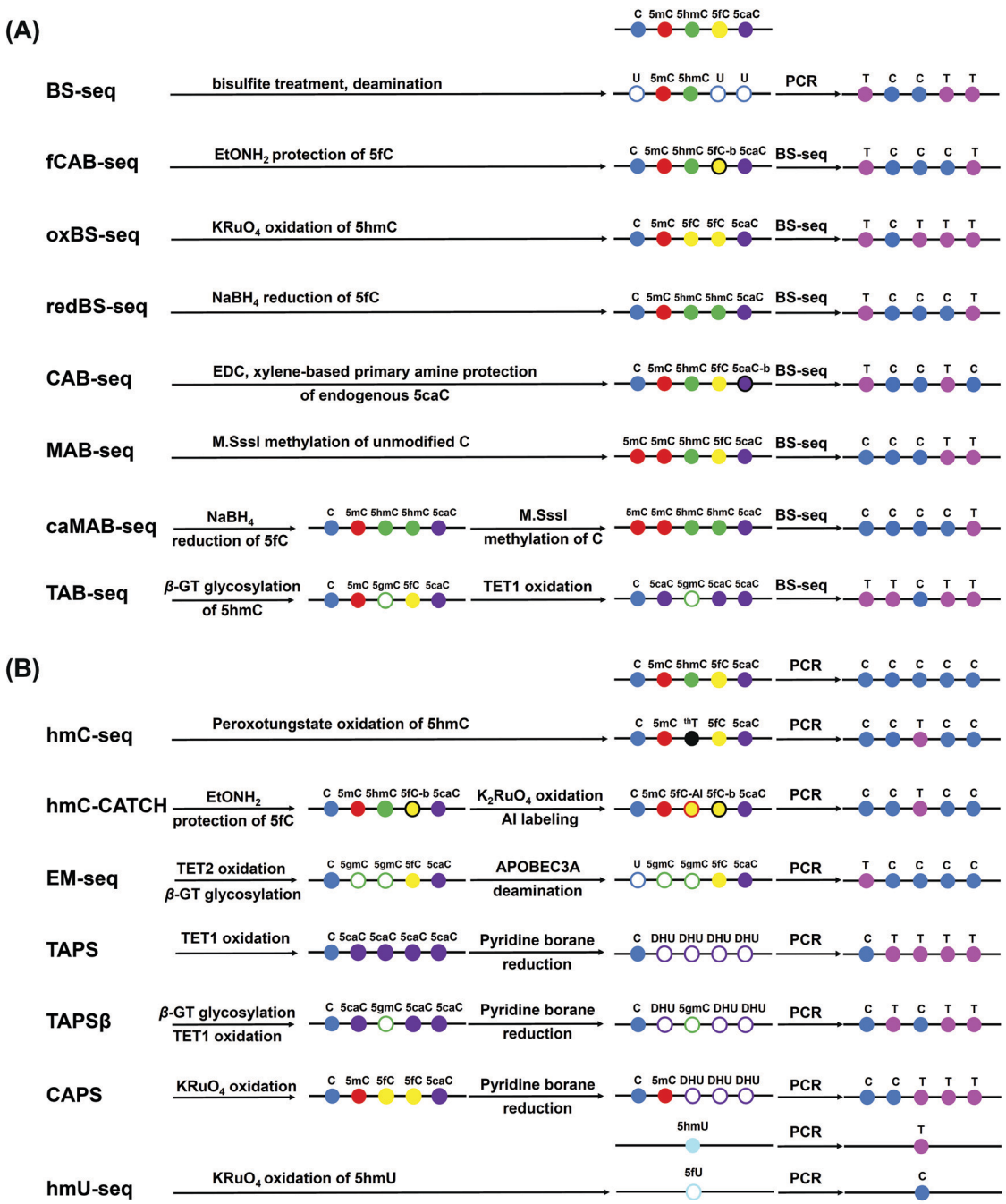



Fig. 2 Schematic illustration of the chemical conversion-sequencing methods for mapping DNA modifications. (A) Bisulfite sequencing. (B) Bisulfite-free sequencing.

The ARP labeling-assisted enrichment and sequencing of 5fC showed that 5fC mainly occurred in the CpG islands (CGIs),[160] supporting the role of 5fC in the maintenance of hypomethylation at CGIs in the ES cells. However, abasic sites as off-target aldehydes could also be labeled by ARP,[190] which may lead to the inaccurate mapping of 5fC. As for the OG-seq approach,[189] a mild one-electron oxidant ($K_2IrBr_6$) was used to selectively oxidize OG to form a covalent adduct of primary-amine terminated biotin (Fig. 3C). The resolution of OG-seq was 0.15 kb. The genome-wide mapping through OG-seq demonstrated that OG was mainly enriched in the promoter and UTR regulatory regions flanking protein-coding sequences.

**3.3.2 Chemical labeling-mediated base transition.** Chemical labeling of DNA modifications may affect their base pairing property, which can be utilized to distinguish DNA modifications from other nucleobases during sequencing. Yi's group developed an approach termed fC-CET sequencing for the genome-wide mapping of 5fC based on the 1,3-indandione (AI)

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

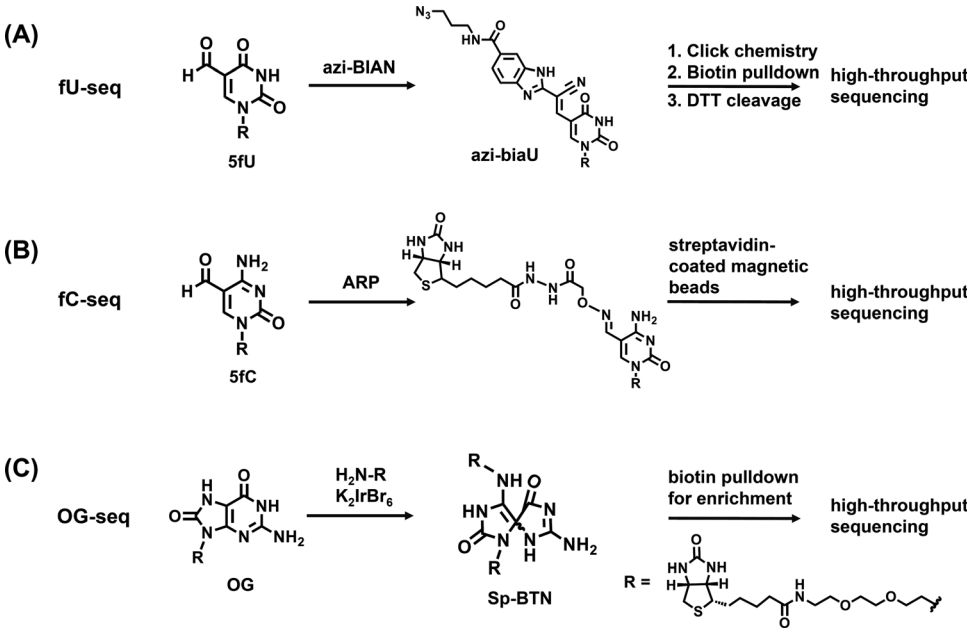



Fig. 3 Reaction of the chemical labeling-assisted enrichment sequencing of fU-seq (A), fC-seq (B) and OG-seq (C) for mapping DNA modifications.

selective labeling of 5fC (Fig. 4A).[191] The AI-labeled 5fC enabled the subsequent C-to-T transition during PCR. However, AI is not an ideal reagent due to its poor solubility. In a recent approach of CLEVER-seq, malononitrile (M) was employed for 5fC labeling (Fig. 4B).[192] Similar to 5fC-AI, 5fC-M enables a C-to-T conversion during sequencing. In this approach, malononitrile showed good solubility, high reactivity, and good biocompatibility. In addition, 2-(5-chlorobenzo[*d*]thiazol-2-yl)acetonitrile (CBAN)[193] and its functional derivative of azi-BP[142] were also demonstrated to be suitable reagents for 5fC labeling (Fig. 4C and D).

A single-nucleotide resolution mapping analysis of 6mA in DNA and RNA has been recently conducted using nitrite labeling-mediated sequencing (Fig. 4E).[194] In this study, treatment of DNA or RNA by sodium nitrite under acidic conditions led to the chemoselective deamination of unmethylated adenines, but not the 6mA. The deamination of adenines resulted in hypoxanthine bases, which were read as guanine during replication and reverse transcription, while 6mA resisted deamination and was read as adenine. The approach coupled with high-throughput sequencing enabled the identification of 6mA sites in DNA within various sequence contexts.

### 3.4 Enzyme-mediated labeling-sequencing

Enzyme-mediated labeling methods have been developed to achieve the efficient enrichment of DNA fragments that carry certain modifications. A typical example is the selective glycosylation of 5hmC by β-GT to produce 5gmC (Fig. 5A).[195] In this respect, an engineered glucose moiety containing an azido group can be labeled to the hydroxyl group of 5hmC (Fig. 5B).[196,197] The azido group on the glucose was then utilized to attach a dibenzocyclooctyne-PEG3-biotin through bioorthogonal chemistry for subsequent affinity enrichment and sequencing. Moreover, the combination of the glycosylation of 5hmC by β-GT, oxidation by sodium periodate, and biotinylation by an aldehyde-reactive probe (GLIB approach) enabled the enrichment and sequencing of 5hmC (Fig. 5C).[198] In addition to β-GT, base J glucosyltransferase (J-GT) was also used to label and enrich 5hmU in genomic DNA (Fig. 5D).[199] J-GT can selectively glycosylate 5hmU to produce glucosylhydroxymethyluracil (base J), which can be pulled down by an antibody against base J or J-binding protein 1. The enriched DNA was then subjected to sequencing analysis.

Deciphering the biological impact of OG requires the single-base resolution sequencing of OG in genomes. Recently, the click-code-seq strategy has been established to enable the base-resolution mapping of OG by coupling the use of repair enzymes with a click DNA ligation reaction to insert a biocompatible locator code.[200] In this approach, DNA was treated with the base excision repair proteins, formamidopyrimidine DNA glycosylase and apurinic/apyrimidinic endonuclease (APE1), to remove OG and yield a gap. The gap was filled with a synthetic *O*-3′-propargyl-modified nucleotide, which was labeled with a code sequence suitable for sequencing the location of the original OG site. With this approach, thousands of OG sites were discovered with distinct patterns related to transcription, chromatin architecture, and chemical oxidation potential.

### 3.5 Deamination-mediated sequencing

Some natural enzymes can selectively recognize specific nucleosides and convert them into other nucleosides, which can be utilized to map modifications in genomic DNA. For example, APOBEC3A (apolipoprotein B mRNA-editing catalytic polypeptide-like 3A) can efficiently deaminate cytosine, 5mC, and 5hmC, but shows no observable deamination activity toward glycosylated 5hmC.[201] Using this unique property of APOBEC3A, we and others reported the APOBEC3A-mediated deamination sequencing

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

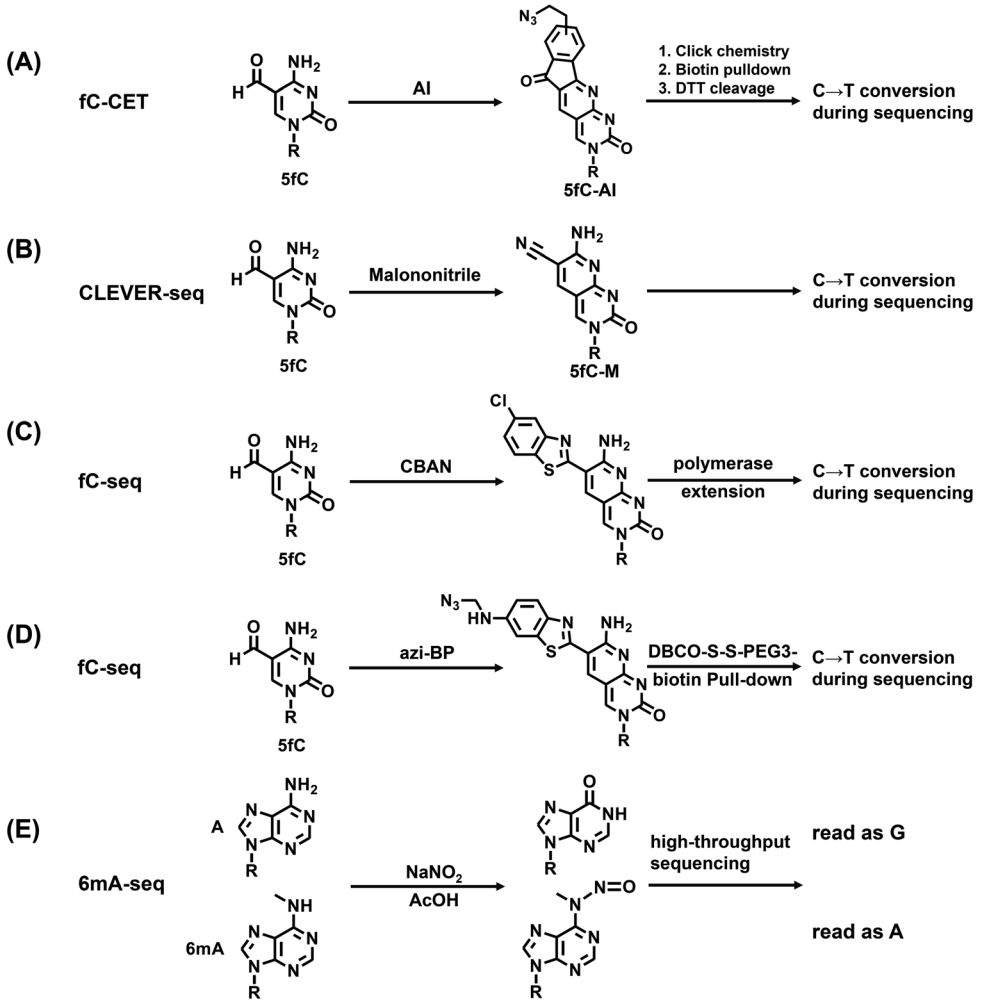



Fig. 4 Reaction of the chemical labeling-mediated base transition sequencing of fC-CET (A), CLEVER-seq (B), fC-seq (C-D) and 6mA-seq (E) for mapping DNA modifications.

(AMD-seq)[202] and APOBEC-coupled epigenetic sequencing (ACE-seq)[203] for the mapping of 5hmC in DNA at single-base resolution. In these approaches, the original C and 5mC in DNA were deaminated by APOBEC3A to form U and T, respectively, both of which were read as T during sequencing, while the glycosylated 5hmC was resistant to deamination and read as C during sequencing. Therefore, the remaining C in the sequence context came from the original 5hmC, which offered the single-base resolution analysis of 5hmC in DNA. These methods based on the unique property of DNA deaminase enzymes can effectively distinguish the modified state of cytosine, and enabled the mapping of 5hmC from 2 ng of input genomic DNA.[203,204]

### 3.6 Polymerase-mediated sequencing

We recently developed an approach for the single-base resolution mapping of OG in DNA using the differential coding properties of *Tth* and *Bsu* DNA polymerases.[205] The *anti* and *syn* conformations of OG led to the dual coding property, *i.e.*, the *anti*-conformation of OG engages in Watson–Crick base pairs with cytosine, while the *syn*-conformation of OG allows stable Hoogsteen base pairs with adenine.[206] We found that *Bsu* DNA polymerase incorporated adenine opposite OG while *Tth* DNA polymerase incorporated cytosine opposite OG. The comparison of the primer extension products by *Bsu* and *Tth* polymerases followed by sequencing provided single-base resolution site information of OG in DNA. The analytical strategy was simple, and the analysis of sequencing data was relatively easy to perform.

### 3.7 Nuclease digestion-mediated sequencing

The nuclease digestion-mediated sequencing method is based on the fact that certain enzymes can specifically recognize specific DNA modifications. Therefore, by combining with various detection methods (such as gel electrophoresis, PCR, and sequencing), the resulting pattern can provide readouts of DNA modifications.

Some restriction enzymes are inhibited by 5mC in the sequence context CpG, so the cutting patterns by such enzymes can provide a readout of DNA methylation. The frequently used methylation-sensitive restriction enzymes for the DNA methylation studies are *Hpa*II and *Sma*I, because they each have an isoschizomer (*Msp*I for *Hpa*II) or neoschizomer (*Xma*I for *Sma*I)

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

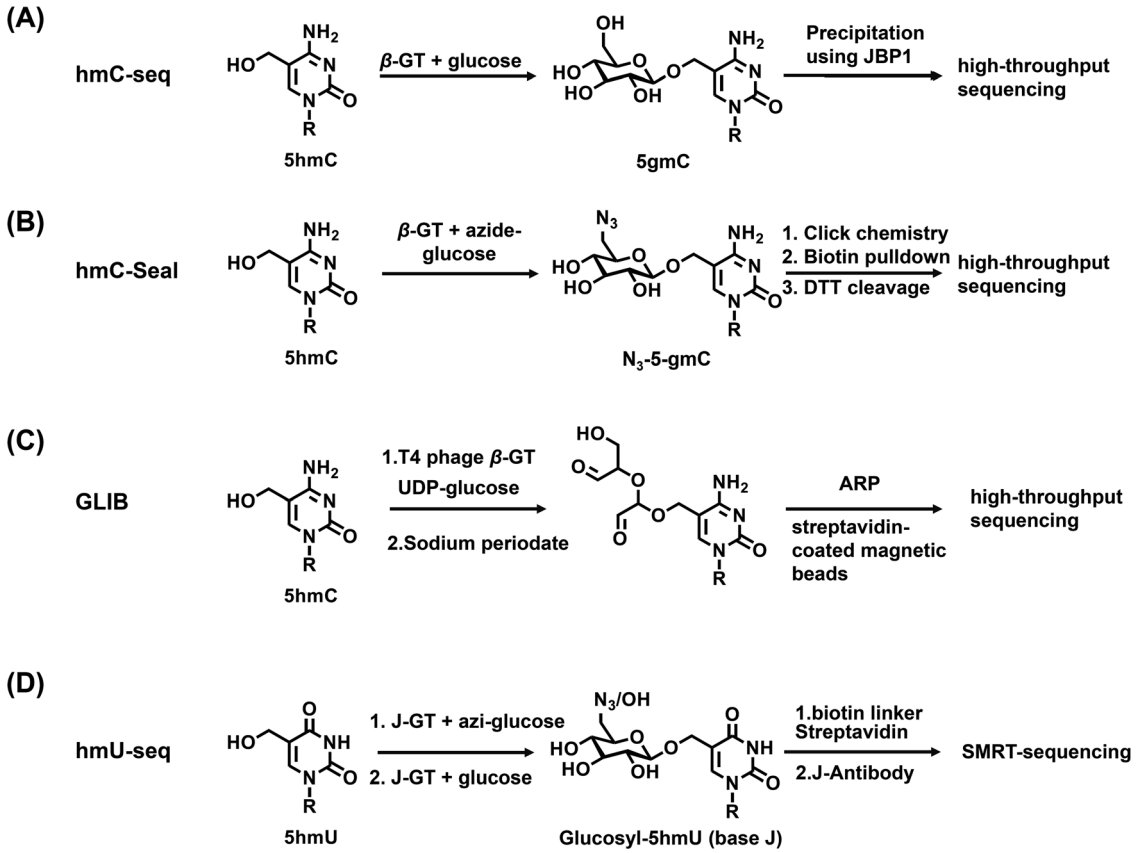



Fig. 5 Reaction of the enzyme-mediated labeling-sequencing of hmC-seq (A), hmC-seal (B), GLIB (C) and hmU-seq (D) for mapping DNA modifications.

that is not inhibited by CpG methylation.[207] Similarly, the methods of DNA-modification-dependent restriction endonuclease *Aba*SI coupled with sequencing (Aba-seq[208] and scAba-seq[209]) were developed to map 5hmC, and the assay of restriction endonuclease *PvuRts*1I coupled with sequencing (Pvu-Seal-seq) was created to map 5hmC and 5fC.[210]

DpnI-assisted $N^6$-methyladenine sequencing (DA-6mA-seq) used DpnI to cleave methylated adenine sites (5′-G6mATC-3′) in duplex DNA.[211] The cutting pattern was then employed to realize the base-resolution mapping of 6mA in sequencing. 6mA cross-linking exonuclease sequencing (6mACE-seq) utilized 6mA-specific antibodies cross-linked to protect the 6mA-DNA fragments from subsequent exonuclease treatment.[212,213] In addition, the 6mA-CLIP-exo combines the immunoprecipitation enrichment, photo-crosslinking, and exonuclease digestion to map the 6mA sites.[14] However, 6mA can only be detected at the sites recognized by these enzymes, and the 6mA site information may be lost if 6mA exists in other motifs.[214]

### 3.8 Nanopore sequencing

The third-generation sequencing technology of nanopore sequencing has become a promising approach in DNA sequencing.[215–219] In nanopore sequencing, different nucleotides passing through nanopores generate different electric currents, which are measured and designated to the corresponding nucleotides or modified nucleotides. Nanopore sequencing technology has been employed to successfully distinguish 5mC from cytosine of human DNA,[220] and 5mC from 5hmC in synthesized DNA strands,[221] thus leading to the mapping of these modifications at single-base resolution. Besides, the 5hmC base was chemically labeled by thiolation in a synthesized single-stranded DNA under bisulfite treatment conditions and detected using nanopore sequencing.[222] Nanopore-based sequencing was also successfully established to distinguish the five known C5-cytosine bases using synthesized DNA.[223] In addition, the α-hemolysin (α-HL) nanopore was employed to map OG in ssDNA. In this work, OG was coupled with 1,12-dodecanediamine and then incubated with cucurbit[7]uril to form a host–guest complex-modified DNA hybrid. Translocation of this DNA hybrid generates current signatures reflective of the presence of OG in DNA.[224]

### 3.9 Single-molecule real-time sequencing

In addition to nanopore sequencing, single-molecule real-time (SMRT) sequencing is another promising third-generation sequencing technology.[225] SMRT sequencing is performed by continuously observing the incorporation of a fluorescently labeled nucleotide with DNA polymerase. On the basis of the differences in the duration of nucleotide incorporation between the unmodified bases and the modified bases, SMRT sequencing

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

can detect the positions and types of base modifications. SMRT sequencing has been applied to the direct detection of 5mC,[226] 6mA,[227–229] 5hmC,[230] and PT.[186] The detection of DNA modifications depends on the magnitude of the effect of the modification on polymerase kinetics. Because the methyl group of 5mC does not directly contribute to base pairing, 5mC shows a subtler impact on nucleotide incorporation than 6mA.[231] Therefore, the detection of 5mC requires relatively high-coverage sequencing or oxidation of 5mC to 5caC using TET1 for signal enhancement.[232]

The emergence of third-generation sequencing technology provides complementary methods for mapping DNA modifications in genomes. Nevertheless, a recent report showed that this method usually overestimates the number of 6mA modifications measured in eukaryotic genomes.[233] SMRT sequencing is also expensive, which impedes its extensive application to larger eukaryotic genomes.

## 4 Conclusions

The existing modifications in genomes broaden our views of the structures and functions of DNA. Elucidation of the biological roles of the modifications in genomes is critical for epigenetic-orientated biological sciences. We highlight the advances in the methods for the quantification and mapping of DNA modifications across prokaryotes and eukaryotes. The development of sensitive analytical methods and techniques in future studies will promote the discovery of new modifications and expand the list of natural modifications in genomes. The knowledge of DNA modifications is still growing, and deciphering the functions of DNA modifications remains an ongoing challenge. Developing appropriate and straightforward sequencing methods for mapping modifications in genomes and elucidating their dynamic changes will facilitate the study of their functions in human diseases, potentially leading to new strategies for therapies. Moreover, the determination of DNA modifications is also valuable in clinical diagnostics and forensics.

Although cells in living organisms have the same DNA sequence context, they can function differently. Single-cell epigenetic studies will provide insights into how heterogeneous DNA modifications may affect the transcriptional output and consequently their functions. The study of DNA modifications in individual cells will offer a new understanding of cell fate decisions, cell development, and disease progression. Mapping DNA modifications in a single cell is particularly valuable in embryos, which may uncover the mystery of epigenetic regulation in embryonic development. Current assays for single-cell DNA modifications frequently suffer from high false-discovery rates, low mapping rates and low detection sensitivity. There is still an urgent need for developing new methods to effectively decipher single-cell DNA modifications, especially for rare DNA modifications.

## Conflicts of interest

The authors declare no competing financial interest.

## Acknowledgements

The work is supported by the National Natural Science Foundation of China (22074110, 21635006, 21721005), and the Fundamental Research Funds for the Central Universities (2042021kf0212).

## References

1 Y. Zhao, X. Zuo, Q. Li, F. Chen, Y. R. Chen, J. Deng, D. Han, C. Hao, F. Huang, Y. Huang, G. Ke, H. Kuang, F. Li, J. Li, M. Li, N. Li, Z. Lin, D. Liu, J. Liu, L. Liu, X. Liu, C. Lu, F. Luo, X. Mao, J. Sun, B. Tang, F. Wang, J. Wang, L. Wang, S. Wang, L. Wu, Z. S. Wu, F. Xia, C. Xu, Y. Yang, B. F. Yuan, Q. Yuan, C. Zhang, Z. Zhu, C. Yang, X. B. Zhang, H. Yang, W. Tan and C. Fan, *Sci. China: Chem.*, 2021, **64**, 171–203.
2 B. Mark and J. F. McGouran, *Nat. Rev. Chem.*, 2018, **2**, 332–348.
3 T. Liu, C. J. Ma, B. F. Yuan and Y. Q. Feng, *Sci. China: Chem.*, 2018, **61**, 381–392.
4 M. V. C. Greenberg and D. Bourc'his, *Nat. Rev. Mol. Cell Biol.*, 2019, **20**, 590–607.
5 C. Rausch, F. D. Hastert and M. C. Cardoso, *J. Mol. Biol.*, 2020, **432**, 1731–1746.
6 C. Luo, P. Hajkova and J. R. Ecker, *Science*, 2018, **361**, 1336–1340.
7 X. Wu and Y. Zhang, *Nat. Rev. Genet.*, 2017, **18**, 517–534.
8 Y. F. He, B. Z. Li, Z. Li, P. Liu, Y. Wang, Q. Tang, J. Ding, Y. Jia, Z. Chen, L. Li, Y. Sun, X. Li, Q. Dai, C. X. Song, K. Zhang, C. He and G. L. Xu, *Science*, 2011, **333**, 1303–1307.
9 M. K. Bilyard, S. Becker and S. Balasubramanian, *Curr. Opin. Chem. Biol.*, 2020, **57**, 1–7.
10 A. J. Sood, C. Viner and M. M. Hoffman, *J. Cheminform*, 2019, **11**, 30.
11 J. Xiong, B. F. Yuan and Y. Q. Feng, *Chem. Res. Toxicol.*, 2019, **32**, 808–819.
12 H. Heyn and M. Esteller, *Cell*, 2015, **161**, 710–713.
13 G. Z. Luo and C. He, *Nat. Struct. Mol. Biol.*, 2017, **24**, 503–506.
14 Y. Fu, G. Z. Luo, K. Chen, X. Deng, M. Yu, D. Han, Z. Hao, J. Liu, X. Lu, L. C. Dore, X. Weng, Q. Ji, L. Mets and C. He, *Cell*, 2015, **161**, 879–892.
15 E. L. Greer, M. A. Blanco, L. Gu, E. Sendinc, J. Liu, D. Aristizabal-Corrales, C. H. Hsu, L. Aravind, C. He and Y. Shi, *Cell*, 2015, **161**, 868–878.
16 G. Zhang, H. Huang, D. Liu, Y. Cheng, X. Liu, W. Zhang, R. Yin, D. Zhang, P. Zhang, J. Liu, C. Li, B. Liu, Y. Luo, Y. Zhu, N. Zhang, S. He, C. He, H. Wang and D. Chen, *Cell*, 2015, **161**, 893–906.
17 M. J. Koziol, C. R. Bradshaw, G. E. Allen, A. S. Costa, C. Frezza and J. B. Gurdon, *Nat. Struct. Mol. Biol.*, 2016, **23**, 24–30.
18 J. Liu, Y. Zhu, G. Z. Luo, X. Wang, Y. Yue, X. Wang, X. Zong, K. Chen, H. Yin, Y. Fu, D. Han, Y. Wang, D. Chen and C. He, *Nat. Commun.*, 2016, **7**, 13052.

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

19 T. P. Wu, T. Wang, M. G. Seetin, Y. Lai, S. Zhu, K. Lin, Y. Liu, S. D. Byrum, S. G. Mackintosh, M. Zhong, A. Tackett, G. Wang, L. S. Hon, G. Fang, J. A. Swenberg and A. Z. Xiao, *Nature*, 2016, **532**, 329–333.
20 C. L. Xiao, S. Zhu, M. He, D. Chen, Q. Zhang, Y. Chen, G. Yu, J. Liu, S. Q. Xie, F. Luo, Z. Liang, D. P. Wang, X. C. Bo, X. F. Gu, K. Wang and G. R. Yan, *Mol. Cell*, 2018, **71**, 306–318.
21 W. Huang, J. Xiong, Y. Yang, S. M. Liu, B. F. Yuan and Y. Q. Feng, *RSC Adv.*, 2015, **5**, 64046–64054.
22 Z. Hao, T. Wu, X. Cui, P. Zhu, C. Tan, X. Dou, K. W. Hsu, Y. T. Lin, P. H. Peng, L. S. Zhang, Y. Gao, L. Hu, H. L. Sun, A. Zhu, J. Liu, K. J. Wu and C. He, *Mol. Cell*, 2020, **78**, 382–395 e388.
23 M. U. Musheev, A. Baumgartner, L. Krebs and C. Niehrs, *Nat. Chem. Biol.*, 2020, **16**, 630–634.
24 X. Liu, W. Lai, Y. Li, S. Chen, B. Liu, N. Zhang, J. Mo, C. Lyu, J. Zheng, Y. R. Du, G. Jiang, G. L. Xu and H. Wang, *Cell Res.*, 2021, **31**, 94–97.
25 F. Kawasaki, D. Beraldi, R. E. Hardisty, G. R. McInroy, P. van Delft and S. Balasubramanian, *Genome Biol.*, 2017, **18**, 23.
26 R. G. Kallen, M. Simon and J. Marmur, *J. Mol. Biol.*, 1962, **5**, 248–250.
27 P. M. Rae, *Science*, 1976, **194**, 1062–1064.
28 T. Pfaffeneder, F. Spada, M. Wagner, C. Brandmayr, S. K. Laube, D. Eisen, M. Truss, J. Steinbacher, B. Hackner, O. Kotljarova, D. Schuermann, S. Michalakis, O. Kosmatchev, S. Schiesser, B. Steigenberger, N. Raddaoui, G. Kashiwazaki, U. Muller, C. G. Spruijt, M. Vermeulen, H. Leonhardt, P. Schar, M. Muller and T. Carell, *Nat. Chem. Biol.*, 2014, **10**, 574–581.
29 P. Borst and R. Sabatini, *Annu. Rev. Microbiol.*, 2008, **62**, 235–251.
30 W. Bullard, J. Lopes da Rosa-Spiegler, S. Liu, Y. Wang and R. Sabatini, *J. Biol. Chem.*, 2014, **289**, 20273–20282.
31 J. H. Xue, G. D. Chen, F. Hao, H. Chen, Z. Fang, F. F. Chen, B. Pang, Q. L. Yang, X. Wei, Q. Q. Fan, C. Xin, J. Zhao, X. Deng, B. A. Wang, X. J. Zhang, Y. Chu, H. Tang, H. Yin, W. Ma, L. Chen, J. Ding, E. Weinhold, R. M. Kohli, W. Liu, Z. J. Zhu, K. Huang, H. Tang and G. L. Xu, *Nature*, 2019, **569**, 581–585.
32 M. M. Greenberg, *Acc. Chem. Res.*, 2012, **45**, 588–597.
33 S. A. Roberts and D. A. Gordenin, *Nat. Rev. Cancer*, 2014, **14**, 786–800.
34 B. F. Yuan, Y. Jiang and Y. Wang, *Chem. Res. Toxicol.*, 2010, **23**, 11–19.
35 A. Tubbs and A. Nussenzweig, *Cell*, 2017, **168**, 644–656.
36 B. F. Yuan, *Chem. Res. Toxicol.*, 2020, **33**, 695–708.
37 A. M. Fleming and C. J. Burrows, *DNA Repair*, 2017, **56**, 75–83.
38 L. Pan, B. Zhu, W. Hao, X. Zeng, S. A. Vlahopoulos, T. K. Hazra, M. L. Hegde, Z. Radak, A. Bacsi, A. R. Brasier, X. Ba and I. Boldogh, *J. Biol. Chem.*, 2016, **291**, 25553–25566.
39 A. M. Fleming, Y. Ding and C. J. Burrows, *Proc. Natl. Acad. Sci. U. S. A.*, 2017, **114**, 2604–2609.
40 X. Wu, B. Cao, P. Aquino, T. P. Chiu, C. Chen, S. Jiang, Z. Deng, S. Chen, R. Rohs, L. Wang, J. E. Galagan and P. C. Dedon, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 14322–14330.
41 T. Tong, S. Chen, L. Wang, Y. Tang, J. Y. Ryu, S. Jiang, X. Wu, C. Chen, J. Luo, Z. Deng, Z. Li, S. Y. Lee and S. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E2988–E2996.
42 L. Wang, S. Jiang, Z. Deng, P. C. Dedon and S. Chen, *FEMS Microbiol. Rev.*, 2019, **43**, 109–122.
43 J. J. Thiaville, S. M. Kellner, Y. Yuan, G. Hutinet, P. C. Thiaville, W. Jumpathong, S. Mohapatra, C. Brochier-Armanet, A. V. Letarov, R. Hillebrand, C. K. Malik, C. J. Rizzo, P. C. Dedon and V. de Crecy-Lagard, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, E1452–E1459.
44 G. Hutinet, W. Kot, L. Cui, R. Hillebrand, S. Balamkundu, S. Gnanakalai, R. Neelakandan, A. B. Carstens, C. Fa Lui, D. Tremblay, D. Jacobs-Sera, M. Sassanfar, Y. J. Lee, P. Weigele, S. Moineau, G. F. Hatfull, P. C. Dedon, L. H. Hansen and V. de Crecy-Lagard, *Nat. Commun.*, 2019, **10**, 5442.
45 M. A. Dawson, *Science*, 2017, **355**, 1147–1152.
46 S. A. Haws, C. M. Leech and J. M. Denu, *Trends Biochem. Sci.*, 2020, **45**, 731–747.
47 U. Sharma and O. J. Rando, *Cell Metab.*, 2017, **25**, 544–558.
48 L. Y. Zhao, J. Song, Y. Liu, C. X. Song and C. Yi, *Protein Cell*, 2020, **11**, 792–808.
49 H. Zeng, B. He and C. Yi, *ChemBioChem*, 2019, **20**, 1898–1905.
50 A. Hofer, Z. J. Liu and S. Balasubramanian, *J. Am. Chem. Soc.*, 2019, **141**, 6420–6429.
51 K. M. Armstrong, E. N. Bermingham, S. A. Bassett, B. P. Treloar, N. C. Roy and M. P. Barnett, *Biotechnol. J.*, 2011, **6**, 113–117.
52 W. Y. Lai, J. Z. Mo, J. F. Yin, C. Lyu and H. L. Wang, *TrAC, Trends Anal. Chem.*, 2019, **110**, 173–182.
53 S. Liu and Y. Wang, *Chem. Soc. Rev.*, 2015, **44**, 7829–7854.
54 X. J. You, T. Liu, C. J. Ma, C. B. Qi, Y. Tong, X. Zhao, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2019, **91**, 10477–10483.
55 C. B. Qi, H. P. Jiang, J. Xiong, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2019, **30**, 553–557.
56 B. Chen, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2019, **91**, 743–756.
57 Y. Dai, C. B. Qi, Y. Feng, Q. Y. Cheng, F. L. Liu, M. Y. Cheng, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2021, **93**, 6938–6946.
58 F. L. Liu, T. T. Ye, J. H. Ding, X. M. Yin, X. K. Yang, W. H. Huang, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2021, **93**, 6848–6856.
59 Q. Y. Li, B. F. Yuan and Y. Q. Feng, *Chem. Lett.*, 2018, **47**, 1453–1459.
60 M. D. Lan, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2019, **30**, 1–6.
61 W. Huang, C. B. Qi, S. W. Lv, M. Xie, Y. Q. Feng, W. H. Huang and B. F. Yuan, *Anal. Chem.*, 2016, **88**, 1378–1384.


© 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

62 L. Wang, S. Chen, K. L. Vergin, S. J. Giovannoni, S. W. Chan, M. S. DeMott, K. Taghizadeh, O. X. Cordero, M. Cutler, S. Timberlake, E. J. Alm, M. F. Polz, J. Pinhassi, Z. Deng and P. C. Dedon, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 2963–2968.
63 Y. Wu, Y. Tang, X. Dong, Y. Y. Zheng, P. Haruehanroengra, S. Mao, Q. Lin and J. Sheng, *ACS Chem. Biol.*, 2020, **15**, 1301–1305.
64 K. Boulias and E. L. Greer, *Methods Mol. Biol.*, 2021, **2198**, 79–90.
65 R. Yin, S. Q. Mao, B. Zhao, Z. Chong, Y. Yang, C. Zhao, D. Zhang, H. Huang, J. Gao, Z. Li, Y. Jiao, C. Li, S. Liu, D. Wu, W. Gu, Y. G. Yang, G. L. Xu and H. Wang, *J. Am. Chem. Soc.*, 2013, **135**, 10396–10403.
66 S. Liu, J. Wang, Y. Su, C. Guerrero, Y. Zeng, D. Mitra, P. J. Brooks, D. E. Fisher, H. Song and Y. Wang, *Nucleic Acids Res.*, 2013, **41**, 6421–6429.
67 R. Yin, J. Mo, M. Lu and H. Wang, *Anal. Chem.*, 2015, **87**, 1846–1852.
68 J. Wang, B. F. Yuan, C. Guerrero, R. Bahde, S. Gupta and Y. Wang, *Anal. Chem.*, 2011, **83**, 2201–2209.
69 Q. Xie, T. P. Wu, R. C. Gimple, Z. Li, B. C. Prager, Q. Wu, Y. Yu, P. Wang, Y. Wang, D. U. Gorkin, C. Zhang, A. V. Dowiak, K. Lin, C. Zeng, Y. Sui, L. J. Y. Kim, T. E. Miller, L. Jiang, C. H. Lee, Z. Huang, X. Fang, K. Zhai, S. C. Mack, M. Sander, S. Bao, A. E. Kerstetter-Fogle, A. E. Sloan, A. Z. Xiao and J. N. Rich, *Cell*, 2018, **175**, 1228–1243.
70 J. Xiong, T. T. Ye, C. J. Ma, Q. Y. Cheng, B. F. Yuan and Y. Q. Feng, *Nucleic Acids Res.*, 2019, **47**, 1268–1277.
71 S. Kellner, M. S. DeMott, C. P. Cheng, B. S. Russell, B. Cao, D. You and P. C. Dedon, *Nat. Chem. Biol.*, 2017, **13**, 888–894.
72 T. Pfaffeneder, B. Hackner, M. Truss, M. Munzel, M. Muller, C. A. Deiml, C. Hagemeier and T. Carell, *Angew. Chem., Int. Ed.*, 2011, **50**, 7008–7012.
73 S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300–1303.
74 M. L. Chen, F. Shen, W. Huang, J. H. Qi, Y. Wang, Y. Q. Feng, S. M. Liu and B. F. Yuan, *Clin. Chem.*, 2013, **59**, 824–832.
75 K. Chen, J. Zhang, Z. Guo, Q. Ma, Z. Xu, Y. Zhou, Z. Xu, Z. Li, Y. Liu, X. Ye, X. Li, B. Yuan, Y. Ke, C. He, L. Zhou, J. Liu and W. Ci, *Cell Res.*, 2016, **26**, 103–118.
76 L. I. Kroeze, B. A. van der Reijden and J. H. Jansen, *Biochim. Biophys. Acta*, 2015, **1855**, 144–154.
77 G. Ficz and J. G. Gribben, *Genomics*, 2014, **104**, 352–357.
78 Y. F. Zhang, C. B. Qi, B. F. Yuan and Y. Q. Feng, *Anal. Chim. Acta*, 2019, **1081**, 103–111.
79 S. C. Wilschefski and M. R. Baxter, *Clin. Biochem. Rev.*, 2019, **40**, 115–133.
80 K. Wrobel, J. A. Landero Figueroa, S. Zaina, G. Lund and K. Wrobel, *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2010, **878**, 609–614.
81 I. Y. Kim, S. H. Suh, I. K. Lee and R. R. Wolfe, *Exp. Mol. Med.*, 2016, **48**, e203.
82 A. C. Newman, C. F. Labuschagne, K. H. Vousden and O. D. K. Maddocks, in *Cancer Metabolism: Methods and Protocols*, ed. M. Haznadar, Springer New York, New York, NY, 2019, pp. 55–67.
83 K. Iwan, R. Rahimoff, A. Kirchner, F. Spada, A. S. Schröder, O. Kosmatchev, S. Ferizaj, J. Steinbacher, E. Parsa, M. Müller and T. Carell, *Nat. Chem. Biol.*, 2018, **14**, 72–78.
84 B. Liu, X. Liu, W. Lai and H. Wang, *Anal. Chem.*, 2017, **89**, 6202–6209.
85 Y. Feng, N. B. Xie, W. B. Tao, J. H. Ding, X. J. You, C. J. Ma, X. Zhang, C. Yi, X. Zhou, B. F. Yuan and Y. Q. Feng, *CCS Chem.*, 2020, **2**, 994–1008.
86 S. Schiesser, B. Hackner, T. Pfaffeneder, M. Muller, C. Hagemeier, M. Truss and T. Carell, *Angew. Chem., Int. Ed.*, 2012, **51**, 6516–6520.
87 M. Zoccali, P. Q. Tranchida and L. Mondello, *TrAC, Trends Anal. Chem.*, 2019, **118**, 444–452.
88 F. Rossella, E. Polledri, V. Bollati, A. Baccarelli and S. Fustinoni, *Rapid Commun. Mass Spectrom.*, 2009, **23**, 2637–2646.
89 Y. Tang, X. D. Gao, Y. Wang, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2012, **84**, 7249–7255.
90 I. D. Podmore, H. R. Griffiths, K. E. Herbert, N. Mistry, P. Mistry and J. Lunec, *Nature*, 1998, **392**, 559.
91 B. L. Qi, P. Liu, Q. Y. Wang, W. J. Cai, B. F. Yuan and Y. Q. Feng, *TrAC, Trends Anal. Chem.*, 2014, **59**, 121–132.
92 F. L. Liu, C. B. Qi, Q. Y. Cheng, J. H. Ding, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2020, **92**, 2301–2309.
93 Y. Feng, C. J. Ma, J. H. Ding, C. B. Qi, X. J. Xu, B. F. Yuan and Y. Q. Feng, *Anal. Chim. Acta*, 2020, **1098**, 56–65.
94 Q. Y. Cheng, J. Xiong, C. J. Ma, Y. Dai, J. H. Ding, F. L. Liu, B. F. Yuan and Y. Q. Feng, *Chem. Sci.*, 2020, **11**, 1878–1891.
95 B. F. Yuan, Q. F. Zhu, N. Guo, S. J. Zheng, Y. L. Wang, J. Wang, J. Xu, S. J. Liu, K. He, T. Hu, Y. W. Zheng, F. Q. Xu and Y. Q. Feng, *Anal. Chem.*, 2018, **90**, 3512–3520.
96 Q. Y. Cheng, J. Xiong, F. Wang, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2018, **29**, 115–118.
97 H. Zeng, C. B. Qi, T. Liu, H. M. Xiao, Q. Y. Cheng, H. P. Jiang, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2017, **89**, 4153–4160.
98 Y. Tang, S. J. Zheng, C. B. Qi, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2015, **87**, 3445–3452.
99 W. Huang, M. D. Lan, C. B. Qi, S. J. Zheng, S. Z. Wei, B. F. Yuan and Y. Q. Feng, *Chem. Sci.*, 2016, **7**, 5495–5502.
100 M. Guo, X. Li, L. Zhang, D. Liu, W. Du, D. Yin, N. Lyu, G. Zhao, C. Guo and D. Tang, *Oncotarget*, 2017, **8**, 91248–91257.
101 J. Xiong, X. Liu, Q. Y. Cheng, S. Xiao, L. X. Xia, B. F. Yuan and Y. Q. Feng, *ACS Chem. Biol.*, 2017, **12**, 1636–1643.
102 Y. Tang, J.-M. Chu, W. Huang, J. Xiong, X.-W. Xing, X. Zhou, Y.-Q. Feng and B.-F. Yuan, *Anal. Chem.*, 2013, **85**, 6129–6135.
103 H. Hong and Y. Wang, *Anal. Chem.*, 2007, **79**, 322–326.
104 Y. Tang, J. Xiong, H. P. Jiang, S. J. Zheng, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2014, **86**, 7764–7772.
105 H. P. Jiang, T. Liu, N. Guo, L. Yu, B. F. Yuan and Y. Q. Feng, *Anal. Chim. Acta*, 2017, **981**, 1–10.

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

106 Y. Yu, F. Yuan, X. H. Zhang, M. Z. Zhao, Y. L. Zhou and X. X. Zhang, *Anal. Chem.*, 2019, **91**, 13047–13053.
107 H. Y. Zhang, J. Xiong, B. L. Qi, Y. Q. Feng and B. F. Yuan, *Chem. Commun.*, 2016, **52**, 737–740.
108 H. P. Jiang, J. Xiong, F. L. Liu, C. J. Ma, X. L. Tang, B. F. Yuan and Y. Q. Feng, *Chem. Sci.*, 2018, **9**, 4160–4167.
109 F. Yuan, X. H. Zhang, J. Nie, H. X. Chen, Y. L. Zhou and X. X. Zhang, *Chem. Commun.*, 2016, **52**, 2698–2700.
110 A. M. Krais, M. G. Cornelius and H. H. Schmeiser, *Electrophoresis*, 2010, **31**, 3548–3551.
111 M. L. Chen, Y. L. Liu, X. W. Xing, X. Zhou, Y. Q. Feng and B. F. Yuan, *Chem. – Eur. J.*, 2013, **19**, 1035–1041.
112 M. F. Paz, M. F. Fraga, S. Avila, M. Guo, M. Pollan, J. G. Herman and M. Esteller, *Cancer Res.*, 2003, **63**, 1114–1121.
113 F. Lyko, D. Stach, A. Brenner, S. Stilgenbauer, H. Dohner, M. Wirtz, M. Wiessler and O. J. Schmitz, *Electrophoresis*, 2004, **25**, 1530–1535.
114 C. Mund, T. Musch, M. Strodicke, B. Assmann, E. Li and F. Lyko, *Biochem. J.*, 2004, **378**, 763–768.
115 Z. Bodi and R. G. Fray, *Methods Mol. Biol.*, 2017, **1562**, 79–87.
116 M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu, L. Aravind and A. Rao, *Science*, 2009, **324**, 930–935.
117 S. Kriaucionis and N. Heintz, *Science*, 2009, **324**, 929–930.
118 H. Gowher, O. Leismann and A. Jeltsch, *EMBO J.*, 2000, **19**, 6918–6923.
119 D. Dooijes, I. Chaves, R. Kieft, A. Dirks-Mulder, W. Martin and P. Borst, *Nucleic Acids Res.*, 2000, **28**, 3017–3021.
120 I. Maatouk, N. Bouaïcha, M. J. Plessis and F. Périn, *Toxicol. Mech. Methods*, 2006, **16**, 313–322.
121 K. L. Thu, L. A. Pikor, J. Y. Kennett, C. E. Alvarez and W. L. Lam, *J. Cell. Physiol.*, 2010, **222**, 522–531.
122 A. Inoue and Y. Zhang, *Science*, 2011, **334**, 194.
123 A. Inoue, L. Shen, Q. Dai, C. He and Y. Zhang, *Cell Res.*, 2011, **21**, 1670–1676.
124 L. Chen, M. Zhang and M. Guo, *Discovery Med.*, 2020, **29**, 85–90.
125 F. van Leeuwen, M. C. Taylor, A. Mondragon, H. Moreau, W. Gibson, R. Kieft and P. Borst, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 2366–2371.
126 N. Machella, F. Regoli and R. M. Santella, *Aquat. Toxicol.*, 2005, **71**, 335–343.
127 R. Kurita, K. Arai, K. Nakamoto, D. Kato and O. Niwa, *Anal. Chem.*, 2012, **84**, 1799–1803.
128 J. Sun, X. D. Lou, H. D. Wang, A. Sollazzo, M. Harms-Ringdahl, S. Skog, E. He and S. Haghdoost, *Int. J. Diabetes Dev. Countries*, 2015, **35**, 368–373.
129 Y. Hori, N. Otomura, A. Nishida, M. Nishiura, M. Umeno, I. Suetake and K. Kikuchi, *J. Am. Chem. Soc.*, 2018, **140**, 1686–1690.
130 C. X. Song, J. Diao, A. T. Brunger and S. R. Quake, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 4338–4343.
131 A. Lopez Torres, E. Yanez Barrientos, K. Wrobel and K. Wrobel, *Anal. Chem.*, 2011, **83**, 7999–8005.
132 P. Guo, S. Yan, J. Hu, X. Xing, C. Wang, X. Xu, X. Qiu, W. Ma, C. Lu, X. Weng and X. Zhou, *Org. Lett.*, 2013, **15**, 3266–3269.
133 J. Hu, X. Xing, X. Xu, F. Wu, P. Guo, S. Yan, Z. Xu, J. Xu, X. Weng and X. Zhou, *Chem. – Eur. J.*, 2013, **19**, 5836–5840.
134 L. Xu, Y. C. Chen, J. Chong, A. Fin, L. S. McCoy, J. Xu, C. Zhang and D. Wang, *Angew. Chem., Int. Ed.*, 2014, **53**, 11223–11227.
135 L. Xu, Y. C. Chen, S. Nakajima, J. Chong, L. Wang, L. Lan, C. Zhang and D. Wang, *Chem. Sci.*, 2014, **5**, 567–574.
136 Y. Wang, C. Liu, W. Yang, G. Zou, X. Zhang, F. Wu, S. Yu, X. Luo and X. Zhou, *Chem. Commun.*, 2018, **54**, 1497–1500.
137 B. Samanta, J. Seikowski and C. Hobartner, *Angew. Chem., Int. Ed.*, 2016, **55**, 1912–1916.
138 C. X. Liu, Y. Q. Chen, Y. F. Wang, F. Wu, X. Zhang, W. Yang, J. Q. Wang, Y. Chen, Z. Y. He, G. R. Zou, S. R. Wang and X. Zhou, *Nano Res.*, 2017, **10**, 2449–2458.
139 C. Liu, Y. Wang, X. Zhang, F. Wu, W. Yang, G. Zou, Q. Yao, J. Wang, Y. Chen, S. Wang and X. Zhou, *Chem. Sci.*, 2017, **8**, 4505–4510.
140 Q. Zhou, K. Li, L. L. Li, K. K. Yu, H. Zhang, L. Shi, H. Chen and X. Q. Yu, *Anal. Chem.*, 2019, **91**, 9366–9370.
141 T. Shahal, O. Green, U. Hananel, Y. Michaeli, D. Shabat and Y. Ebenstein, *Methods Appl. Fluoresc.*, 2016, **4**, 044003.
142 Y. Wang, C. Liu, X. Zhang, W. Yang, F. Wu, G. Zou, X. Weng and X. Zhou, *Chem. Sci.*, 2018, **9**, 3723–3728.
143 Y. Yang, G. Yang, H. Chen, H. Zhang, J. J. Feng and C. Cai, *Analyst*, 2018, **143**, 2051–2056.
144 X. Chen, J. Huang, S. Zhang, F. Mo, S. Su, Y. Li, L. Fang, J. Deng, H. Huang, Z. Luo and J. Zheng, *ACS Appl. Mater. Interfaces*, 2019, **11**, 3745–3752.
145 L. Zhang, X. Xiao, Y. Xu, D. Chen, J. Chen, Y. Ma, Z. Dai and X. Zou, *Biosens. Bioelectron.*, 2018, **100**, 184–191.
146 J. Huang, S. Zhang, F. Mo, S. Su, X. Chen, Y. Li, L. Fang, H. Huang, J. Deng, H. Liu, X. Yang and J. Zheng, *Biosens. Bioelectron.*, 2019, **127**, 155–160.
147 N. Beaujean, J. Salvaing, N. A. A. Hadi and S. Pennings, *Methods Mol. Biol.*, 2018, **1708**, 59–80.
148 L. Cui, J. Hu, M. Wang, C. C. Li and C. Y. Zhang, *Anal. Chem.*, 2019, **91**, 1232–1236.
149 Y. Meng, W. Bai, Y. Zhang, H. Sun and Y. Li, *Talanta*, 2020, **210**, 120597.
150 H. Y. Chen, J. R. Wei, J. X. Pan, W. Zhang, F. Q. Dang, Z. Q. Zhang and J. Zhang, *Biosens. Bioelectron.*, 2017, **91**, 328–333.
151 J. Fan, Y. Liu, E. Xu, Y. Zhang, W. Wei, L. Yin, Y. Pu and S. Liu, *Anal. Chim. Acta*, 2016, **946**, 48–55.
152 Y. Zhang, Y. Li, Y. Wei, H. Sun and H. Wang, *Talanta*, 2017, **170**, 546–551.
153 C. B. Qi, J. H. Ding, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2019, **30**, 1618–1626.
154 M. Weber, J. J. Davies, D. Wittig, E. J. Oakeley, M. Haase, W. L. Lam and D. Schubeler, *Nat. Genet.*, 2005, **37**, 853–862.
155 T. Rauch and G. P. Pfeifer, *Lab. Invest.*, 2005, **85**, 1172–1180.

 © 2021 The Author(s). Published by the Royal Society of Chemistry

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

156 D. Serre, B. H. Lee and A. H. Ting, *Nucleic Acids Res.*, 2010, **38**, 391–399.
157 M. Jung, S. Kadam, W. Xiong, T. A. Rauch, S. G. Jin and G. P. Pfeifer, *Epigenomics*, 2015, **7**, 695–706.
158 O. Bogdanovic, A. Fernandez-Minan, J. J. Tena, E. de la Calle-Mustienes and J. L. Gomez-Skarmeta, *Methods*, 2013, **62**, 207–215.
159 J. P. Thomson and R. R. Meehan, *Methods Mol. Biol.*, 2018, **1708**, 679–696.
160 E. A. Raiber, D. Beraldi, G. Ficz, H. E. Burgess, M. R. Branco, P. Murat, D. Oxley, M. J. Booth, W. Reik and S. Balasubramanian, *Genome Biol.*, 2012, **13**, R69.
161 B. Yao, Y. Li, Z. Wang, L. Chen, M. Poidevin, C. Zhang, L. Lin, F. Wang, H. Bao, B. Jiao, J. Lim, Y. Cheng, L. Huang, B. L. Phillips, T. Xu, R. Duan, K. H. Moberg, H. Wu and P. Jin, *Mol. Cell*, 2018, **71**, 848–857 e846.
162 X. Liu, W. Lai, N. Zhang and H. Wang, *Anal. Chem.*, 2018, **90**, 5546–5551.
163 L. J. Cliffe, T. N. Siegel, M. Marshall, G. A. Cross and R. Sabatini, *Nucleic Acids Res.*, 2010, **38**, 3923–3935.
164 M. Yoshihara, L. Jiang, S. Akatsuka, M. Suyama and S. Toyokuni, *DNA Res.*, 2014, **21**, 603–612.
165 A. Lentini, C. Lagerwall, S. Vikingsson, H. K. Mjoseng, K. Douvlataniotis, H. Vogt, H. Green, R. R. Meehan, M. Benson and C. E. Nestor, *Nat. Methods*, 2018, **15**, 499–504.
166 V. V. Ashapkin, L. I. Kutueva and B. F. Vanyushin, *Methods Mol. Biol.*, 2020, **2138**, 297–312.
167 M. J. Booth, M. R. Branco, G. Ficz, D. Oxley, F. Krueger, W. Reik and S. Balasubramanian, *Science*, 2012, **336**, 934–937.
168 M. Yu, G. C. Hon, K. E. Szulwach, C. X. Song, L. Zhang, A. Kim, X. Li, Q. Dai, Y. Shen, B. Park, J. H. Min, P. Jin, B. Ren and C. He, *Cell*, 2012, **149**, 1368–1380.
169 D. Tan, T. H. Chung, X. Sun and X. Y. Jia, *Anal. Chem.*, 2018, **90**, 13200–13206.
170 M. J. Booth, G. Marsico, M. Bachman, D. Beraldi and S. Balasubramanian, *Nat. Chem.*, 2014, **6**, 435–440.
171 C. X. Song, K. E. Szulwach, Q. Dai, Y. Fu, S. Q. Mao, L. Lin, C. Street, Y. Li, M. Poidevin, H. Wu, J. Gao, P. Liu, L. Li, G. L. Xu, P. Jin and C. He, *Cell*, 2013, **153**, 678–691.
172 X. Lu, C. X. Song, K. Szulwach, Z. Wang, P. Weidenbacher, P. Jin and C. He, *J. Am. Chem. Soc.*, 2013, **135**, 9315–9317.
173 X. Lu, D. Han, B. S. Zhao, C. X. Song, L. S. Zhang, L. C. Dore and C. He, *Cell Res.*, 2015, **25**, 386–389.
174 H. Wu, X. Wu and Y. Zhang, *Nat. Protoc.*, 2016, **11**, 1081–1100.
175 F. Neri, D. Incarnato, A. Krepelova, C. Parlato and S. Oliviero, *Nat. Protoc.*, 2016, **11**, 1191–1205.
176 M. Ehrlich, G. G. Wilson, K. C. Kuo and C. W. Gehrke, *J. Bacteriol.*, 1987, **169**, 939–943.
177 M. Yu, L. Ji, D. A. Neumann, D. H. Chung, J. Groom, J. Westpheling, C. He and R. J. Schmitz, *Nucleic Acids Res.*, 2015, **43**, e148.
178 H. Zeng, B. He, B. Xia, D. Bai, X. Lu, J. Cai, L. Chen, A. Zhou, C. Zhu, H. Meng, Y. Gao, H. Guo, C. He, Q. Dai and C. Yi, *J. Am. Chem. Soc.*, 2018, **140**, 13190–13194.
179 F. Kawasaki, S. Martinez Cuesta, D. Beraldi, A. Mahtey, R. E. Hardisty, M. Carrington and S. Balasubramanian, *Angew. Chem., Int. Ed.*, 2018, **57**, 9694–9696.
180 Y. Wang, X. Zhang, F. Wu, Z. Chen and X. Zhou, *Chem. Sci.*, 2019, **10**, 447–452.
181 G. Hayashi, K. Koyama, H. Shiota, A. Kamio, T. Umeda, G. Nagae, H. Aburatani and A. Okamoto, *J. Am. Chem. Soc.*, 2016, **138**, 14178–14181.
182 Y. Liu, P. Siejka-Zielinska, G. Velikova, Y. Bi, F. Yuan, M. Tomkova, C. Bai, L. Chen, B. Schuster-Bockler and C. X. Song, *Nat. Biotechnol.*, 2019, **37**, 424–429.
183 Y. Liu, J. Cheng, P. Siejka-Zielinska, C. Weldon, H. Roberts, M. Lopopolo, A. Magri, V. D'Arienzo, J. M. Harris, J. A. McKeating and C. X. Song, *Genome Biol.*, 2020, **21**, 54.
184 G. Gish and F. Eckstein, *Science*, 1988, **240**, 1520–1522.
185 Y. Chen, T. Zheng, J. Li, J. Cui, Z. Deng, D. You and L. Yang, *Sci. Rep.*, 2019, **9**, 7485.
186 B. Cao, C. Chen, M. S. DeMott, Q. Cheng, T. A. Clark, X. Xiong, X. Zheng, V. Butty, S. S. Levine, G. Yuan, M. Boitano, K. Luong, Y. Song, X. Zhou, Z. Deng, S. W. Turner, J. Korlach, D. You, L. Wang, S. Chen and P. C. Dedon, *Nat. Commun.*, 2014, **5**, 3951.
187 J. Li, Y. Chen, T. Zheng, L. Kong, S. Zhu, Y. Sun, Z. Deng, L. Yang and D. You, *PLoS Genet.*, 2019, **15**, e1008026.
188 Y. Wang, C. Liu, F. Wu, X. Zhang, S. Liu, Z. Chen, W. Zeng, W. Yang, X. Zhang, Y. Zhou, X. Weng, Z. Wu and X. Zhou, *iScience*, 2018, **9**, 423–432.
189 Y. Ding, A. M. Fleming and C. J. Burrows, *J. Am. Chem. Soc.*, 2017, **139**, 2569–2572.
190 H. Ide, K. Akamatsu, Y. Kimura, K. Michiue, K. Makino, A. Asaeda, Y. Takamori and K. Kubo, *Biochemistry*, 1993, **32**, 8276–8283.
191 B. Xia, D. Han, X. Lu, Z. Sun, A. Zhou, Q. Yin, H. Zeng, M. Liu, X. Jiang, W. Xie, C. He and C. Yi, *Nat. Methods*, 2015, **12**, 1047–1050.
192 C. Zhu, Y. Gao, H. Guo, B. Xia, J. Song, X. Wu, H. Zeng, K. Kee, F. Tang and C. Yi, *Cell Stem Cell*, 2017, **20**, 720–731 e725.
193 C. Liu, Y. Wang, W. Yang, F. Wu, W. Zeng, Z. Chen, J. Huang, G. Zou, X. Zhang, S. Wang, X. Weng, Z. Wu, Y. Zhou and X. Zhou, *Chem. Sci.*, 2017, **8**, 7443–7447.
194 Y. Mahdavi-Amiri, K. Chung Kim Chung and R. Hili, *Chem. Sci.*, 2021, **12**, 606–612.
195 A. B. Robertson, J. A. Dahl, C. B. Vagbo, P. Tripathi, H. E. Krokan and A. Klungland, *Nucleic Acids Res.*, 2011, **39**, e55.
196 C. X. Song, K. E. Szulwach, Y. Fu, Q. Dai, C. Yi, X. Li, Y. Li, C. H. Chen, W. Zhang, X. Jian, J. Wang, L. Zhang, T. J. Looney, B. Zhang, L. A. Godley, L. M. Hicks, B. T. Lahn, P. Jin and C. He, *Nat. Biotechnol.*, 2011, **29**, 68–72.
197 D. Han, X. Lu, A. H. Shih, J. Nie, Q. You, M. M. Xu, A. M. Melnick, R. L. Levine and C. He, *Mol. Cell*, 2016, **63**, 711–719.
198 W. A. Pastor, U. J. Pape, Y. Huang, H. R. Henderson, R. Lister, M. Ko, E. M. McLoughlin, Y. Brudno,

Open Access Article. Published on 21 May 2021. Downloaded on 10/10/2024 11:50:56 AM.
This article is licensed under a Creative Commons Attribution-NonCommercial 3.0 Unported Licence.

S. Mahapatra, P. Kapranov, M. Tahiliani, G. Q. Daley, X. S. Liu, J. R. Ecker, P. M. Milos, S. Agarwal and A. Rao, *Nature*, 2011, **473**, 394–397.
199 W. Bullard, R. Kieft and R. Sabatini, *Biol. Methods Protoc.*, 2017, **2**, bpw006.
200 J. Wu, M. McKeague and S. J. Sturla, *J. Am. Chem. Soc.*, 2018, **140**, 9783–9787.
201 E. K. Schutsky, C. S. Nabel, A. K. F. Davis, J. E. DeNizio and R. M. Kohli, *Nucleic Acids Res.*, 2017, **45**, 7655–7665.
202 Q. Y. Li, N. B. Xie, J. Xiong, B. F. Yuan and Y. Q. Feng, *Anal. Chem.*, 2018, **90**, 14622–14628.
203 E. K. Schutsky, J. E. DeNizio, P. Hu, M. Y. Liu, C. S. Nabel, E. B. Fabyanic, Y. Hwang, F. D. Bushman, H. Wu and R. M. Kohli, *Nat. Biotechnol.*, 2018, **36**, 1083–1090.
204 Z. Sun, R. Vaisvila, L. M. Hussong, B. Yan, C. Baum, L. Saleh, M. Samaranayake, S. Guan, N. Dai, I. R. Corrêa, Jr., S. Pradhan, T. B. Davis, T. C. Evans, Jr. and L. M. Ettwiller, *Genome Res.*, 2021, **31**, 291–300.
205 F. Tang, S. Liu, Q. Y. Li, J. Yuan, L. Li, Y. Wang, B. F. Yuan and Y. Q. Feng, *Chem. Sci.*, 2019, **10**, 4272–4281.
206 K. N. Kirouac and H. Ling, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 3210–3215.
207 M. J. Pajares, C. Palanca-Ballester, R. Urtasun, E. Alemany-Cosme, A. Lahoz and J. Sandoval, *Methods*, 2021, **187**, 3–12.
208 Z. Sun, J. Terragni, J. G. Borgaro, Y. Liu, L. Yu, S. Guan, H. Wang, D. Sun, X. Cheng, Z. Zhu, S. Pradhan and Y. Zheng, *Cell Rep.*, 2013, **3**, 567–576.
209 D. Mooijman, S. S. Dey, J. C. Boisset, N. Crosetto and A. van Oudenaarden, *Nat. Biotechnol.*, 2016, **34**, 852–856.
210 Z. Sun, N. Dai, J. G. Borgaro, A. Quimby, D. Sun, I. R. Correa, Jr., Y. Zheng, Z. Zhu and S. Guan, *Mol. Cell*, 2015, **57**, 750–761.
211 G. Z. Luo, F. Wang, X. Weng, K. Chen, Z. Hao, M. Yu, X. Deng, J. Liu and C. He, *Nat. Commun.*, 2016, **7**, 11301.
212 W. S. S. Goh, *Methods Mol. Biol.*, 2021, **2198**, 369–377.
213 C. W. Q. Koh, Y. T. Goh, J. D. W. Toh, S. P. Neo, S. B. Ng, J. Gunaratne, Y. G. Gao, S. R. Quake, W. F. Burkholder and W. S. S. Goh, *Nucleic Acids Res.*, 2018, **46**, 11659–11670.
214 G. Z. Luo, M. A. Blanco, E. L. Greer, C. He and Y. Shi, *Nat. Rev. Mol. Cell Biol.*, 2015, **16**, 705–710.
215 J. J. Kasianowicz, E. Brandin, D. Branton and D. W. Deamer, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 13770–13773.
216 A. Meller, L. Nivon, E. Brandin, J. Golovchenko and D. Branton, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 1079–1084.
217 D. W. Deamer and D. Branton, *Acc. Chem. Res.*, 2002, **35**, 817–825.
218 D. Branton, D. W. Deamer, A. Marziali, H. Bayley, S. A. Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S. B. Jovanovich, P. S. Krstic, S. Lindsay, X. S. Ling, C. H. Mastrangelo, A. Meller, J. S. Oliver, Y. V. Pershin, J. M. Ramsey, R. Riehn, G. V. Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin and J. A. Schloss, *Nat. Biotechnol.*, 2008, **26**, 1146–1153.
219 J. Shendure, S. Balasubramanian, G. M. Church, W. Gilbert, J. Rogers, J. A. Schloss and R. H. Waterston, *Nature*, 2017, **550**, 345–353.
220 J. T. Simpson, R. E. Workman, P. C. Zuzarte, M. David, L. J. Dursi and W. Timp, *Nat. Methods*, 2017, **14**, 407–410.
221 A. H. Laszlo, I. M. Derrington, H. Brinkerhoff, K. W. Langford, I. C. Nova, J. M. Samson, J. J. Bartlett, M. Pavlenok and J. H. Gundlach, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 18904–18909.
222 W. W. Li, L. Gong and H. Bayley, *Angew. Chem., Int. Ed.*, 2013, **52**, 4350–4355.
223 Z. L. Wescoe, J. Schreiber and M. Akeson, *J. Am. Chem. Soc.*, 2014, **136**, 16582–16587.
224 L. Liu, Y. Li, T. Li, J. Xie, C. Chen, Q. Liu, S. Zhang and H. C. Wu, *Anal. Chem.*, 2016, **88**, 1073–1077.
225 S. Ardui, A. Ameur, J. R. Vermeesch and M. S. Hestand, *Nucleic Acids Res.*, 2018, **46**, 2159–2168.
226 B. A. Flusberg, D. R. Webster, J. H. Lee, K. J. Travers, E. C. Olivares, T. A. Clark, J. Korlach and S. W. Turner, *Nat. Methods*, 2010, **7**, 461–465.
227 Z. Liang, L. Shen, X. Cui, S. Bao, Y. Geng, G. Yu, F. Liang, S. Xie, T. Lu, X. Gu and H. Yu, *Dev. Cell*, 2018, **45**, 406–416.
228 C. Zhou, C. Wang, H. Liu, Q. Zhou, Q. Liu, Y. Guo, T. Peng, J. Song, J. Zhang, L. Chen, Y. Zhao, Z. Zeng and D. X. Zhou, *Nat. Plants*, 2018, **4**, 554–563.
229 G. Fang, D. Munera, D. I. Friedman, A. Mandlik, M. C. Chao, O. Banerjee, Z. Feng, B. Losic, M. C. Mahajan, O. J. Jabado, G. Deikus, T. A. Clark, K. Luong, I. A. Murray, B. M. Davis, A. Keren-Paz, A. Chess, R. J. Roberts, J. Korlach, S. W. Turner, V. Kumar, M. K. Waldor and E. E. Schadt, *Nat. Biotechnol.*, 2012, **30**, 1232–1239.
230 C. X. Song, T. A. Clark, X. Y. Lu, A. Kislyuk, Q. Dai, S. W. Turner, C. He and J. Korlach, *Nat. Methods*, 2012, **9**, 75–77.
231 T. A. Clark, I. A. Murray, R. D. Morgan, A. O. Kislyuk, K. E. Spittle, M. Boitano, A. Fomenkov, R. J. Roberts and J. Korlach, *Nucleic Acids Res.*, 2012, **40**, e29.
232 T. A. Clark, X. Lu, K. Luong, Q. Dai, M. Boitano, S. W. Turner, C. He and J. Korlach, *BMC Biol.*, 2013, **11**, 4.
233 Z. K. O'Brown, K. Boulias, J. Wang, S. Y. Wang, N. M. O'Brown, Z. Hao, H. Shibuya, P. E. Fady, Y. Shi, C. He, S. G. Megason, T. Liu and E. L. Greer, *BMC Genomics*, 2019, **20**, 445.


© 2021 The Author(s). Published by the Royal Society of Chemistry