CDDpress

www.nature.com/cdd

REVIEW ARTICLE OPEN

Check for updates

Mapping epigenetic modifications by sequencing technologies

Xiufei Chen[1,2], Haiqi Xu[1,2], Xiao Shu[1,2] and Chun-Xiao Song[1,2]✉

The "epigenetics" concept was first described in 1942. Thus far, chemical modifications on histones, DNA, and RNA have emerged as three important building blocks of epigenetic modifications. Many epigenetic modifications have been intensively studied and found to be involved in most essential biological processes as well as human diseases, including cancer. Precisely and quantitatively mapping over 100 [1], 17 [2], and 160 [3] different known types of epigenetic modifications in histone, DNA, and RNA is the key to understanding the role of epigenetic modifications in gene regulation in diverse biological processes. With the rapid development of sequencing technologies, scientists are able to detect specific epigenetic modifications with various quantitative, high-resolution, whole-genome/transcriptome approaches. Here, we summarize recent advances in epigenetic modification sequencing technologies, focusing on major histone, DNA, and RNA modifications in mammalian cells.

Cell Death & Differentiation; https://doi.org/10.1038/s41418-023-01213-1

FACTS

- Traditional, bisulfite sequencing is the gold standard in DNA methylation sequencing but it severely damages DNA [4].
- Recently, EM-Seq and TAPS have been developed to replace bisulfite sequencing [5, 6].
- Very recently, base-resolution and quantitative RNA epitranscriptomic modifications sequencing methods have started to emerge [7–11].
- Third-generation sequencing is promising for epigenetic sequencing; however, many challenges remain to be resolved [12].

OPEN QUESTIONS

- Could further RNA modification sequencing methods be developed with improved efficiency and accuracy?
- Could simultaneous base-resolution sequencing of a wide range of epigenetic modifications be achieved?
- Could third-generation sequencing finally deliver epigenetic data as accurate and cost-effective as next-generation sequencing?
- Could we develop large-scale live cell temporal/spatial epigenetic sequencing?

INTRODUCTION

In 1942, embryologist Conrad Waddington first established the concept of "epigenetics" with the famous "epigenetic landscape" model [13]. However, the explosion of epigenetic studies has only occurred over the last two decades. The term "epigenetic" refers to the alteration of gene expression without change of the DNA sequence, which mainly occurs in the form of a myriad of chemical modifications in histone, DNA, and RNA. Among them, chromatin structure modification studies started in the 1990s. Over 100 distinct modifications have been found in histone, including acetylation (Ac), methylation (Me), phosphorylation (P), ubiquitylation (Ub), SUMOylation (SUMO), ADP ribosylation (ADP), O-GlcNAcylation (O-Glc), and biotinylation (Biotin) (Fig. 1) [1, 14, 15]. In the late 1940s, 5-methylcytosine (5mC) was the first identified DNA chemical modification [16], and to date, over 17 types of DNA chemical modifications have been identified [2]. 5mC is the most predominant and important modification in mammalian DNA, the so-called "fifth base". Its oxidative products, 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), also exist in mammalian DNA (Fig. 1). Compared to DNA, RNA modifications are more diverse, with over 160 types of RNA modifications reported thus far [3]. The most common dynamic RNA modifications include N^6-methyladenosine (m^6A), pseudouridine (Ψ), N^1-methyladenosine (m^1A), N^7-methylguanosine (m^7G), 5-methylcytidine (m^5C) / 5-hydroxymethylcytidine (hm^5C) in mammalian cells, leading to the exciting field of epitranscriptomics [17–20] (Fig. 1). In the past two decades, the dysregulation of epigenetic modifications has been revealed in many studies and serves as a hallmark of cancer: many somatic mutations in human cancers occur in epigenetic regulators [21–23].

The first wave of methods to identify these epigenetic modifications used bulk measurements such as thin layer chromatography (TLC), LC/GC-MS, immunofluorescence, and immunoprecipitation. These methods, however, could not provide sequence information about the modifications [24]. With the rapid development of next generation sequencing (NGS), sequencing based detection methods for epigenetic modifications have been developed at a rapid pace. Earlier methods usually relied on affinity enrichment (i.e. immunoprecipitation, biotin pull-down, etc.). While useful and cost-effective, they only provide limited semi-quantitative and low-resolution (a few hundred base pairs) information about the modification. More recent developments focus on high-resolution (e.g., base-level resolution)

[1]Ludwig Institute for Cancer Research, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK. [2]Target Discovery Institute, Nuffield Department of Medicine, University of Oxford, Oxford OX3 7FZ, UK. ✉email: chunxiao.song@ludwig.ox.ac.uk

Received: 24 March 2023 Revised: 9 August 2023 Accepted: 14 August 2023
Published online: 01 September 2023

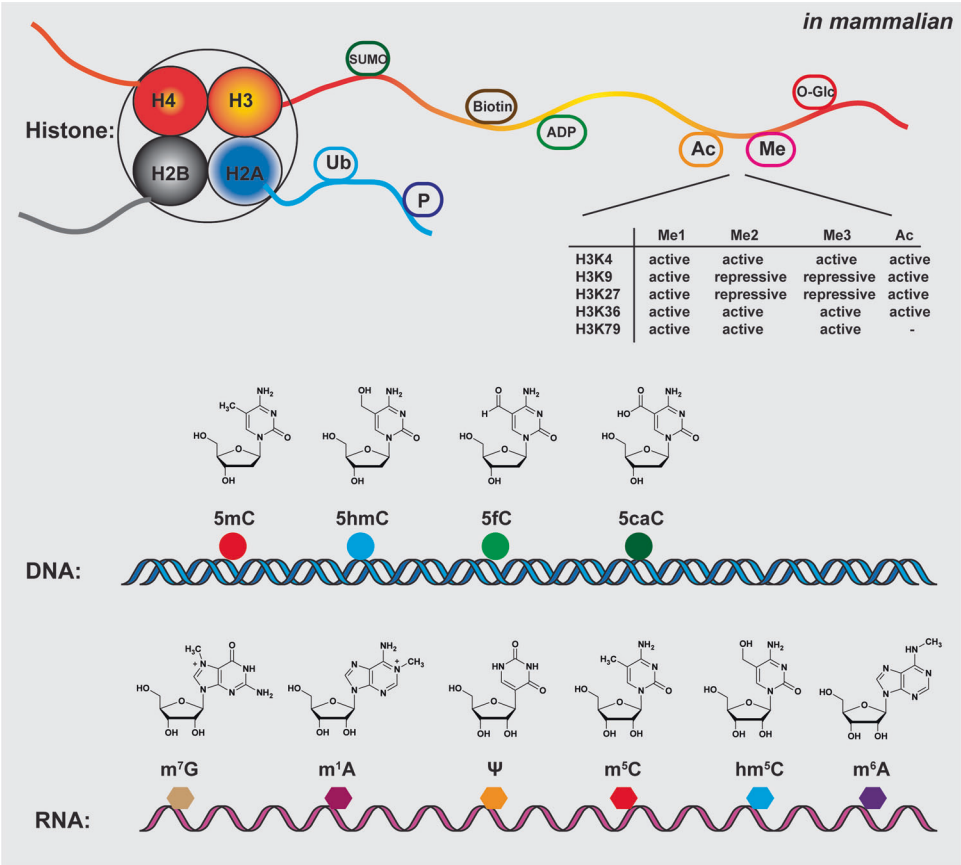

Fig. 1 Epigenetic modifications in Histone, DNA and RNA. Histone Modifications: Nucleosomes are composed of DNA wrapped around the four core histones (H3, H4, H2A, H2B). Post-translational modifications, including acetylation (Ac), methylation (Me), phosphorylation (P), ubiquitylation (Ub), SUMOylation (SUMO), ADP ribosylation (ADP), O-GlcNAcylation (O-Glc), and biotinylation (Biotin), are commonly observed on the N-terminal histone tails. Notably, significant modifications of Histone H3 have been associated with either active or repressive gene expression. Major DNA Modifications in Mammals: DNA modifications include DNA 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC). These modifications possess distinct chemical structures and are crucial in epigenetic regulation. Major RNA modifications in Mammals: RNA modifications include N^6-methyladenosine (m^6A), pseudouridine (Ψ), N^1-methyladenosine (m^1A), N^7-methylguanosine (m^7G), 5-methylcytidine (m^5C) and 5-hydroxymethylcytidine (hm^5C) modifications, and their chemical structures.

and quantitative sequencing methods, which provide a more complete picture of the modification. While NGS has significantly advanced the field, it still has limitations, including short read lengths, biases introduced by amplification steps, and difficulties in accurately resolving repetitive genomic regions. In this regard, the future development of third-generation sequencing emerges as a promising solution [25, 26]. In this review, we summarize the diversity and complexity of epigenetic modifications and their related sequencing methods, focusing on recent advances in quantitative and base-resolution technologies for major DNA (5mC/5hmC/5fC/5caC) and RNA (m^6A/Ψ) modifications in mammalian cells. This review is an update and extension of our previous summary on DNA and RNA modification detection [19].

HISTONE MODIFICATION DETECTION BY SEQUENCING TECHNOLOGIES

Chromatin is composed of DNA and histone proteins with nucleosomes as the basic structural units. Around 147 base pairs of DNA are packaged into an octamer of the four core histones (H3, H4, H2A, H2B) as the nucleosome [27]. Histone modifications usually occur at the unstructured N-terminal of the histones, which could modify chromatin accessibility to chromatin remodeling enzymes and transcription factors, thus regulating gene expression [14, 28]. Some modifications such as H3K27ac and H3K4me1/2/3 are active markers, while others such as H3K27me3 and H3K9me3 are repressive markers in histone H3 tail (Fig. 1) [14, 29–32]. It has been widely established that dysregulation of histone modifications causes a number of diseases, including cancer [30, 33, 34]. It is therefore important to study their location and distribution across the genome. Among many approaches to mapping these modifications, chromatin immunoprecipitation followed by sequencing (ChIP–Seq) is the most classical method to sequence histone modifications in a genome-wide manner. ChIP–Seq is based on formaldehyde/paraformaldehyde-mediated protein-DNA crosslinking, followed by incubation with the specific antibodies to enrich the target histone modification and finally library construction for next-generation sequencing to detect genome-wide histone modification distribution. The earliest formaldehyde-mediated protein-DNA cross-linking ChIP experiment was conducted by Solomon et al. [35] to probe histone H4 and hsp70 DNA interactions in vivo in 1988, whilst the first ChIP-Seq method was established by Barski et al. in 2007, who mapped genome-wide distributions of 20 histone lysine and arginine methylations [36]. Despite this technique being widely used, ChIP–Seq has several limitations, such as the requirement for larger amounts of input DNA, false positive rates induced by cross-linking between DNA and protein, and high background noise due to poor antibody specificity. [37]. Recently, two novel technologies have been established to overcome these limitations. In 2017, Cleavage Under Targets and Release Using Nuclease (CUT&RUN) technology was described, for semi-quantitative detection of protein-DNA interactions in situ at

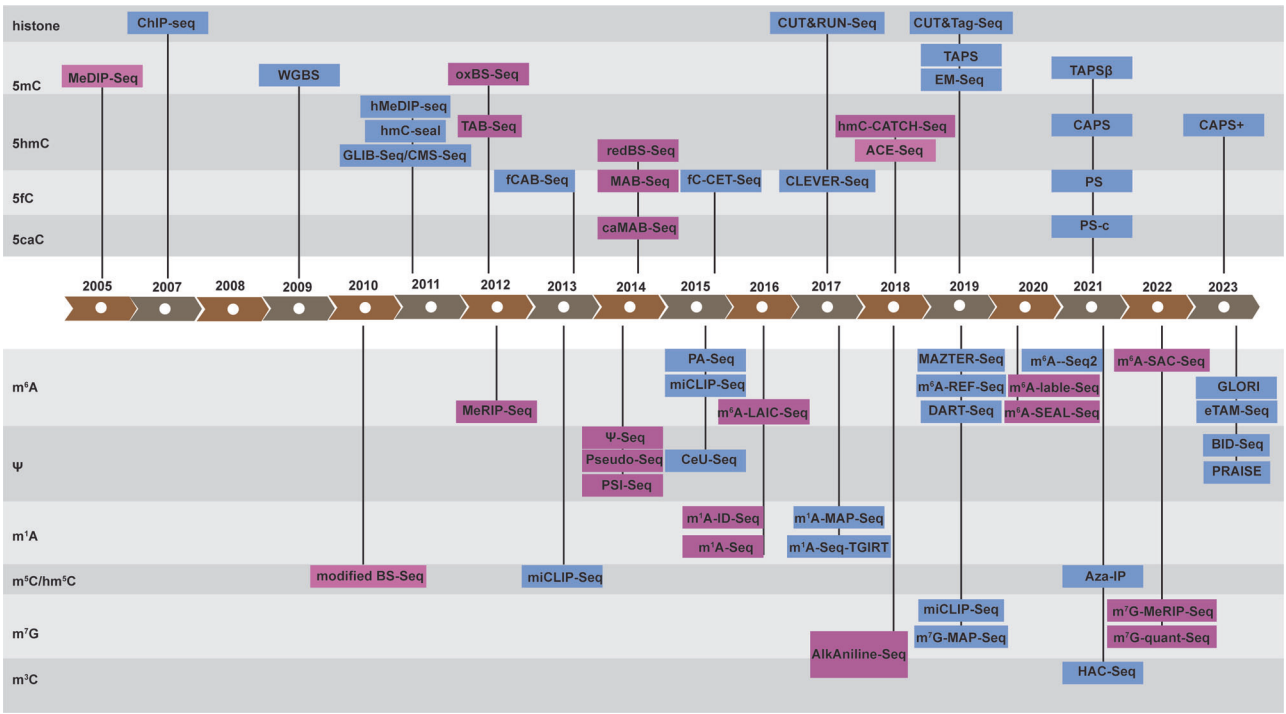

Fig. 2 Timeline landscape of epigenetic modification detection technologies based on next-generation sequencing. A comprehensive overview of the progression of next-generation sequencing techniques developed over time for detecting major Histone, DNA (5mC, 5hmC, 5fC, 5caC) and RNA (m^6A, Ψ, m^1A, m^5C/ hm^5C, m^7G, m^3C) modifications.

~20 bp resolution [38, 39] (Fig. 2). To avoid the epitope masking and false positive binding sites generated by the crosslinking step in ChIP-Seq, CUT&RUN immobilizes the cells on lectin-coated magnetic beads and incubates with specific antibodies and protein A-MNase. Ca^{2+} was then added to initiate the cleavage reaction to release the target protein-DNA complexes for sequencing. In 2019, an improved Cleavage Under Targets and Tagmentation (CUT&Tag) technology was developed, which replaces the MNase digestion with Tn5 tagmentation (Fig. 2). CUT&Tag simplifies the library construction steps [40], making it more feasible for single-cell experiments to probe histone modifications, such as those that are characteristic of active promoters (H3K4me3), enhancers (H3K27ac), gene bodies (H3K36me3) and inactive regions (H3K27me3), as demonstrated in mice brains [41]. However, CUT&RUN and CUT&Tag inherit the disadvantages of antibody-based approaches. Future development could focus on novel antibody-free and enrichment-free approaches to map histone modifications quantitatively at single base resolution.

DNA METHYLATION (5MC) SEQUENCING TECHNOLOGIES

DNA methylation of cytosine (5mC) is the most predominant modification, and accounts for ~5% of all cytosine (C) [42]. 5mC modification on DNA is catalyzed by a family of enzymes known as DNA methyl transferases (DNMTs), predominantly in the symmetrical CpG dinucleotides: 70–80% of CpGs in the mammalian genome are methylated [43]. DNA methylation has been found to relate to X chromosome inactivation, and genomic imprinting, and is dysregulated in most diseases, including cancer [33, 34, 44]. The hypermethylation of tumor suppressor genes or hypomethylation of tumor oncogenes has been commonly identified in many cancers [45–47]. This makes DNA methylation an attractive therapeutic target for cancer treatment. For example, DNA methylation inhibitors, such as 5-azacytidine, 5-aza-2′-deoxycytidine (decitabine, DAC), and cytarabine (Ara-C), have been widely used to inhibit tumor growth in vivo and also applied to treat myelodysplastic syndromes (MDS) and acute myeloid leukemia (AML) patients [48–52].

In 2005, MeDIP-Seq (Fig. 2) was developed to capture differential DNA methylation in normal and transformed human cells [53]. However, as it is based on antibody enrichment, it suffers from the same limitations as ChIP-Seq. In 1992, bisulfite treatment for 5mC detection was first developed by Frommer, et al. [54]. This method utilizes sodium bisulfite, which specifically deaminates unmodified C to Uracil (U) while leaving 5mC and 5hmC intact. During the subsequent PCR and sequencing, U is therefore read as thymine (T). When comparing bisulfite-converted sequences with the reference genome, 5mC and 5hmC can be distinguished from unmodified C. Based on this bisulfite reaction, the base-resolution and quantitative whole-genome bisulfite sequencing (WGBS) (Fig. 2) for 5mC and 5hmC mapping was developed in 2009 [55]. With over 99% conversion of unmodified C to U, WGBS became the powerful and widely used gold standard for mapping 5mC and 5hmC (Fig. 3). However, WGBS has two significant drawbacks. First, bisulfite treatment is a very harsh chemical reaction that will cause severe DNA damage and loss [4]. Second, the converted unmodified cytosine, which accounts for around 95% of all cytosine in the genome, leads to reduced DNA sequence complexity, lower mapping efficiency, and biased genomic coverage after bisulfite treatment. Nevertheless, various improvements have been made, such as post-bisulfite adaptor tagging (PBAT), which salvages fragmented DNA caused by bisulfite treatment to mitigate the bisulfite-induced loss of intact sequencing templates [56]. By adopting PBAT, various single-cell WGBS protocols have been developed [57, 58].

To overcome the limitations of WGBS, two bisulfite-free DNA methylation sequencing methods have recently been developed. The enzymatic methyl-Seq (EM-Seq) (Fig. 2) is an enzymatic deamination method involving three enzymes in two reactions [6]. In the first reaction, Tet methylcytosine dioxygenase 2 (TET2) is used to catalyze the oxidization of 5mC to 5hmC, 5fC, and 5caC. In the same reaction, β-glucosyltransferase (βGT) glucosylates both TET2-derived and genomic 5hmC to form 5-(β-glucosyloxymethyl) cytosine (5gmC). In the second reaction, AID/APOBEC family DNA deaminase

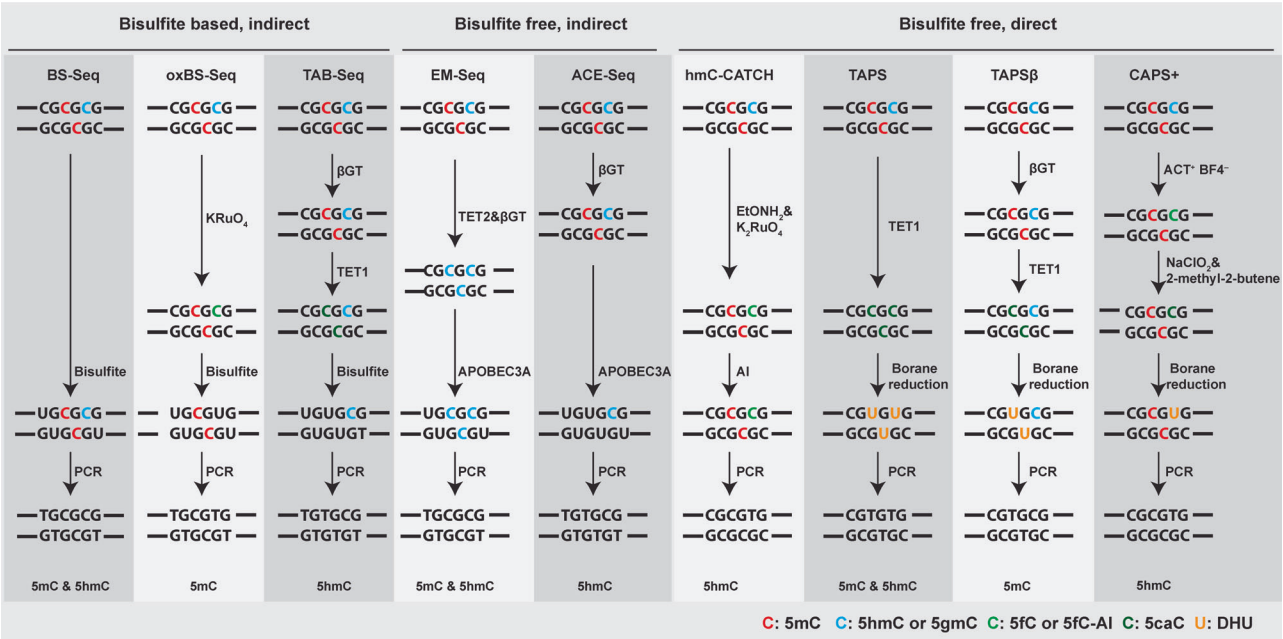

Fig. 3 DNA 5mC and 5hmC modification detection sequencing technologies. Summary of prominent techniques employed in the detection of DNA 5mC and 5hmC epigenetic modifications. Bisulfite-based and indirect methodologies encompass bisulfite sequencing (BS-Seq) for both 5mC and 5hmC, oxidative bisulfite sequencing (oxBS-Seq) for 5mC, and TET-assisted bisulfite sequencing (TAB-Seq) for 5hmC. Bisulfite-free yet indirect methods include enzymatic methyl-Seq (EM-Seq) for both 5mC and 5hmC and APOBEC-coupled epigenetic sequencing (ACE-Seq) for 5hmC. Alternatively, bisulfite-free and direct strategies entail chemical-assisted C-to-T conversion of 5hmC sequencing (hmC-CATCH) for 5hmC, TET-assisted pyridine borane sequencing (TAPS) for both 5mC and 5hmC, TAPSβ for 5mC, chemical-assisted pyridine borane sequencing plus (CAPS+) for 5hmC.

APOBEC3A deaminates unmodified C to U, while 5fC, 5caC, and 5gmC are protected from APOBEC3A deamination (Fig. 3). EM-Seq achieved over 96% protection rate on 5mC with less than 0.6% non-conversion rate (false positive rate) of unmodified cytosine. The mild enzymatic reactions make EM-Seq compatible with input DNA quantities as low as 100 pg. Recently, single-cell EM-Seq has also been established [59]. Similar to WGBS, EM-Seq indirectly maps the 5mC and 5hmC modifications by converting unmodified cytosine in DNA, which leads to a low complexity genome.

Unlike WGBS and EM-Seq, TET-assisted pyridine borane sequencing (TAPS) (Fig. 2) developed by Liu et al. [5] in 2019 is a direct DNA 5mC and 5hmC detection method. It is based on a novel borane reduction chemistry to convert 5caC to dihydrouracil (DHU), which is a mild reaction with little DNA damage compared to the bisulfite reaction. TAPS combines ten-eleven translocation (TET) oxidation of 5mC and 5hmC to 5caC and borane reduction conversion of 5caC to DHU, which subsequently reads as T after PCR amplification (Fig. 3). TAPS achieved a high conversion rate of over 96% on 5mC and a low false positive rate of 0.23% on unmodified cytosine. A key advantage of TAPS is that it induces a C-to-T transition only at modified cytosine (5mC and 5hmC), which only accounts for 5% of all cytosine. Such direct detection preserves the underlying genomic information, enabling TAPS to achieve substantially higher mapping rate and sequencing quality at half the sequencing cost to WGBS [60]. By incorporating βGT to convert 5hmC to 5gmC and to protect 5hmC from TET oxidation and borane reduction, Liu et al. extended TAPS to TAPSβ to enable 5mC-specific sequencing (Fig. 3) [61].

DNA METHYLATION OXIDATIVE PRODUCTS: 5HMC, 5FC, AND 5CAC SEQUENCING TECHNOLOGIES

The removal of 5mC in cells is accomplished by TET family proteins [62, 63] or passively by DNA replication [64, 65]. While TET1/2/3 show tissue specific distributions, they can all convert 5mC to 5hmC, 5fC, and 5caC in three consecutive steps. 5fC and 5caC can be excised and restored to C through thymine DNA glycosylase (TDG) and base excision repair (BER) pathway, resulting in active demethylation [66]. 5hmC was first identified in 2009 [62, 67] and is highly abundant in neurons and related to active genes, such as *Pcp4*, *Neurod2* [68]. 5fC and 5caC have much lower abundances in the mammalian genome and are considered to be demethylation intermediates, however, recent investigations also showed the possibility of these modifications being stable in nature [69, 70]. 5hmC plays an important role in many biological processes such as zygote/embryonic development [71, 72], and cell differentiation [73], and its dysregulation has been shown in tumorigenesis [74, 75], while the 5fC and 5caC's role in cells remain to be delineated.

To map genome-wide 5hmC distribution, various sequencing methods have been developed, such as hMeDIP-Seq [76], hmC-seal [77], CMS-Seq and GLIB-Seq (Fig. 2), all of which are antibody or biotin based affinity enrichment of 5hmC-containing genomic DNA with limited resolution. To map 5hmC quantitatively across the whole genome at the single base resolution, several technologies have been developed, including oxBS-Seq, TAB-Seq [78–81] in 2012, ACE-Seq [82], hmC-CATCH-Seq [83] in 2018, CAPS [61] in 2021, and CAPS+ [84] in 2023 (Fig. 2). Oxidative bisulfite sequencing (oxBS-Seq) utilizes potassium perruthenate ($KRuO_4$) oxidation of 5hmC to 5fC to remove the 5hmC signal from WGBS, thereby only detecting 5mC (Fig. 3). By comparing and subtracting the result of oxBS from WGBS, the base-resolution 5hmC level can be obtained. oxBS-Seq was employed in mouse embryonic stem (ES) cells, leading to the identification of approximately 800 5hmC-modified CpG islands (CGIs), exhibiting an average hydroxymethylation level of 3.3%. Notably, the highly modified CGIs discovered through this method were correlated to intragenic and intergenic CGIs, but not transcription start site (TSS) CGIs [80]. TET-assisted bisulfite sequencing (TAB-Seq), on the other hand, uses TET oxidation and βGT glucosylation to eliminate the 5mC signal from WGBS, thereby detecting 5hmC directly without the need to

compare with WGBS (Fig. 3). Utilizing TAB-Seq on mouse embryonic stem cells (mES), a total of 2 057 636 5hmC sites were identified. These sites are found in close proximity to, but not directly on, transcription factor-binding sites. Moreover, the distribution of these sites demonstrates significant variation across different distal-regulatory elements [79]. Both oxBS-Seq and TAB-Seq are based on bisulfite sequencing and have therefore inherited the disadvantages of WGBS, namely severe loss of DNA integrity and complexity.

Recently, four bisulfite-free base-resolution 5hmC sequencing methods have been developed. In 2018, Chemical-assisted C-to-T conversion of 5hmC sequencing (hmC-CATCH) was developed [83], using potassium ruthenate (K_2RuO_4) oxidation of 5hmC to 5fC, followed by chemical labeling of 5fC by an azido derivative of 1,3-indandione (AI). The labeled adduct is read as T after PCR amplification, making hmC-CATCH a direct sequencing method of 5hmC (Fig. 3). However, hmC-CATCH is not quantitative due to the use of a biotin pull-down enrichment step. Utilizing the hmC-CATCH technique, 607 021 5hmC sites were detected in human embryonic stem cells, and it is also the first time to unveil the base-resolution hydroxymethylome in the cell-free DNA (cfDNA) of both healthy individuals and individuals diagnosed with cancer [83]. The second of these methods, APOBEC-coupled epigenetic sequencing (ACE-Seq), is an enzymatic deamination method for sequencing 5hmC like EM-Seq. It uses βGT to block 5hmC as 5gmC, before APOBEC3A deamination of unmodified C and 5mC to U. ACE-Seq is nondestructive and can achieve a 98.5% protection rate on 5hmC with 0.1% and 0.5% non-conversion rates (false positive rate) on unmodified cytosine and 5mC, respectively (Fig. 3). ACE-Seq identified 798 643 5hmC sites in mES and also provided valuable insights into the distribution of 5hmC in cortical excitatory neurons [82]. Thirdly, in 2021, chemical-assisted pyridine borane sequencing (CAPS) was developed as a sister method for TAPS. CAPS utilizes potassium ruthenate oxidation of 5hmC to 5fC, followed by borane reduction of 5fC to DHU (Fig. 3). It achieved 83.1% 5hmC-to-T conversion, 0.72% and 0.38% false-positive rates on unmodified cytosine and 5mC, respectively. Owing to the direct 5hmC readout, CAPS showed improved sequence complexity, higher base quality, and mapping rate compared to TAB-Seq and ACE-Seq. CAPS detected 1 762 287 5hmC sites in mES [61]. Finally, in 2023, chemical-assisted pyridine borane sequencing plus (CAPS+) [84] was developed as an updated version of CAPS. CAPS+ replaced potassium ruthenate oxidation in CAPS with two milder chemical oxidation reactions: using 4-acetamido-2,2,6,6-tetramethylpiperidine-1-oxoammonium tetra-fluoroborate (ACT^+ $BF4^-$) to oxidize 5hmC to 5fC, then employing sodium chlorite ($NaClO_2$) in the Pinnick oxidation to convert 5fC into 5caC. Borane reduction converts 5caC to DHU, which is similar to the process used in CAPS (Fig. 3). The CAPS+ builds upon the strengths of CAPS, with enhancements in conversion rate (achieving 94.5% for 5hmC) and reductions in false-positive rates (0.15% for 5mC and 0.17% for unmodified cytosine) [84].

Beyond 5hmC, various 5fC and 5caC sequencing methods have also been developed [85]. Several bisulfite-based 5fC and 5caC sequencing methods have been developed utilizing various approaches to modulate the behavior of 5fC and 5caC in bisulfite mediated deamination. This includes 5fC chemically assisted bisulfite sequencing method (fCAB-Seq) [86], reduced bisulfite sequencing (redBS-Seq) [87], M.SssI methylase-assisted bisulfite sequencing (MAB-Seq) and caMAB-Seq [88, 89] (Fig. 2). The investigations presented in these studies uncovered the preferential occurrence of 5fC at poised enhancers, highlighting the crucial involvement of TDG in active DNA demethylation process. Additionally, these studies also detected a notable asymmetry between the strands for both 5fC and 5caC in mES [87, 88]. In 2015, the bisulfite-free cyclization-enabled C-to-T transition of 5fC sequencing (fC-CET-Seq) [90] (Fig. 2) was developed utilizing an azido derivative of 1, 3-indandione to convert 5fC to an adduct that can be read as T following PCR amplification. In 2017, it was further developed into chemical-labeling-enabled C-to-T Conversion Sequencing (CLEVER-Seq) (Fig. 2) for single-cell 5fC sequencing [91]. CLEVER-Seq unraveled the inherent heterogeneity of 5fC in mES. Additionally, 5fC exhibited parental-specific patterns, and its localization on promoters correlated with gene activation throughout the preimplantation development of mice. Based on the borane reduction chemistry, Liu et al. also developed pyridine borane sequencing (PS) and pyridine borane sequencing for carboxylcytosine (PS-c) (Fig. 2) for whole-genome base-resolution sequencing of 5fC and 5caC [61].

OTHER DNA MODIFICATIONS (6MA, 4MC, BASE J)

Beyond 5mC, two other DNA methylation forms have been reported: N^6-methyladenine (6mA) and N^4-methylcytosine (4mC). 6mA is the most prevalent form of methylation in prokaryotes, but its presence in mammalian cells is still in debate [92–96]. Similar to 6mA, 4mC is well known to exist in bacteria but its presence in eukaryotic genomic DNA remains unclear [95]. Beta-D-glucopyranosyloxymethyluracil (base J) is the first hypermodified base found in eukaryotic DNA. It was initially discovered in 1993 and has since been predominantly observed in kinetoplastids [97].

RNA M^6A MODIFICATION SEQUENCING TECHNOLOGIES

m^6A is the most prevalent internal modification in mRNA and has been intensively studied since its first identification in 1974 [98]. m^6A is involved in RNA splicing, translation, stability, translocation, and high-level structure regulation [99, 100], and it has been linked to diverse developmental processes and cancers [18, 100–103]. m^6A is enriched around stop codons and 3′ UTRs (3′ untranslated regions) in mammals. The installation of the modification is catalyzed by m^6A methyltransferase complex proteins (so called “writers”), such as METTL3/14/16, RBM15/15B, ZC3H13, VIRMA, CBLL1, WTAP, and KIAA1429; the m^6A recognition is accomplished by m^6A-binding proteins (“readers”), such as YTHDF1/2/3, YTHDC1/2, IGF2BP1/2/3 and HNRNPA2B1 [17, 100, 104–109]. The demethylases (“erasers”), including FTO [110] and ALKBH5 [111], can remove m^6A modifications.

Traditionally, the most widely used method for mapping m^6A modifications is m^6A-Seq, also called MeRIP-Seq, which is similar to MeDIP-Seq, using antibody enrichment [112, 113]. Other strategies based on the m^6A antibody have also been developed, such as PA-Seq, miCLIP, m^6A-Seq2, m^6A-LAIC-Seq [114–117] (Fig. 2). As these antibody-based enrichment methods are largely influenced by the quality and intrinsic bias of the antibody, antibody-free m^6A sequencing methods have subsequently been developed. MAZTER-Seq [118] and m^6A-REF-Seq [119] (Fig. 2) use a methylation sensitive *E. coli* toxin and RNA endoribonuclease (MazF) which can recognize and cut the ACA motif sequence from 5′ sites. Although these methods can provide single-base resolution mapping of m^6A sites, their preference for ACA sites means that only around 16–25% of all m^6A sites across the whole transcriptome can be mapped. DART-Seq (deamination adjacent to RNA modification targets) [120] (Fig. 2) depends on cytidine deaminase APOBEC1 and m^6A-binding YTH domain fusion protein to induce C-to-U deamination at sites adjacent to m^6A modifications. However, this method requires cellular transfection which limits its application to primary cells and tissue samples. m^6A-SEAL-Seq (FTO-assisted m6A selective chemical labeling method) [121] (Fig. 2) uses FTO to oxidize m^6A into N^6-hydroxymethyladenosin (hm^6A), followed by dithiothreitol (DTT)-mediated thiol-addition reaction to generate N^6-dithiolsitolmethyladenosine (dm^6A): dm^6A can then be labeled with biotin for pull-down and sequencing. m^6A-label-Seq [122] (Fig. 2) is a metabolic labeling method that feeds the cells with a methionine analog, *Se*-allyl-L-selenohomocysteine, which can replace m^6A with N^6-

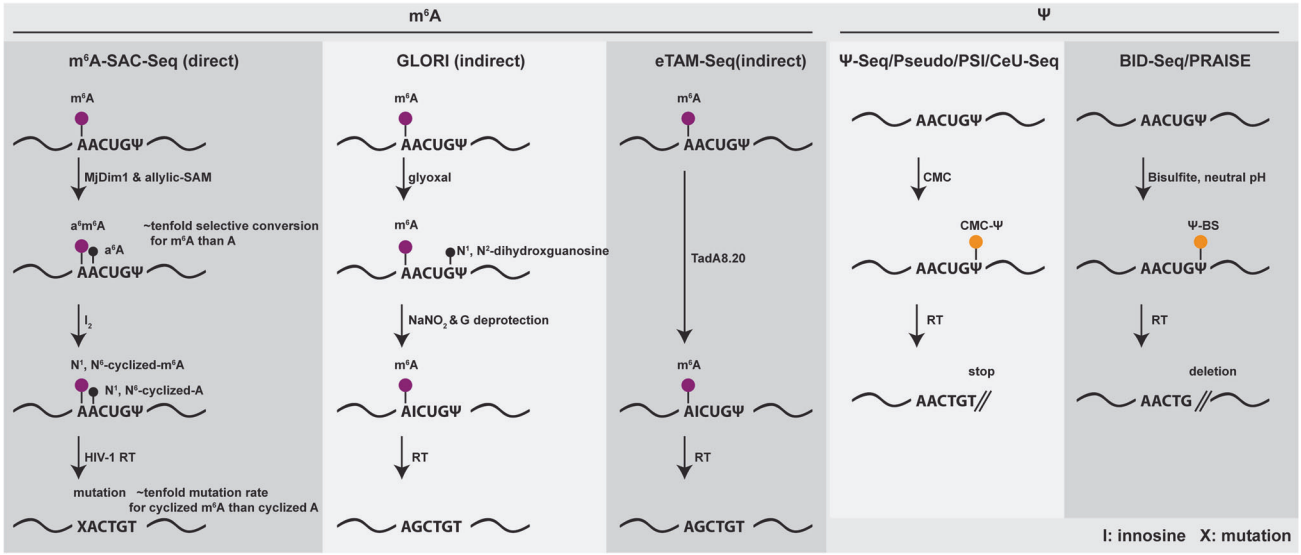

Fig. 4 RNA m^{6}A and Ψ modification detection sequencing technologies. Three methodologies for achieving single-base resolution and stoichiometric m^{6}A sequencing have been developed: m^{6}A-SAC-Seq, Glyoxal and nitrite-mediated deamination of unmethylated adenosines (GLORI), Evolved TadA-assisted N^6-methyladenosine sequencing (eTAM-Seq). In the realm of Ψ sequencing, various approaches are based on CMC-Ψ, such as Ψ-Seq, Pseudo-Seq, pseudouridine site identification sequencing (PSI-Seq), and N_3-CMC-enriched psedouridine sequencing (CeU-Seq). Notably, the most recent advancements in this field have yielded two innovative Ψ-BS based approaches: bisulfite-induced deletion sequencing (BID-Seq) and PRAISE.

allyladenosine (a^{6}A). The a^{6}A positions can be detected using iodination-induced misincorporation during reverse transcription. With this method, the authors demonstrated the detection of 2 479 and 3 808 m^{6}A modification sites in HeLa and HEK293T cells, respectively. However, like DART-seq, m^{6}A-label-Seq can only be applied to in vivo samples.

Very recently, three single-base resolution and stoichiometric m^{6}A sequencing methods have been developed. m^{6}A-SAC-Seq [7] (Fig. 2) is the first enzyme dependent, direct, and quantitative m^{6}A sequencing method. It uses the dimethyltransferase MjDim1 (*Methanocaldococcus jannaschii homolog*) which can convert m^{6}A into a^{6}m^{6}A (N^6-allyl, N^6-methyladnosine) in the presence of allylic-S-adenosyl-methionine (SAM). With the iodination-induced misincorporation used in the m^{6}A-label-Seq, a^{6}m^{6}A can induce mutations during reverse transcription, thereby detecting m^{6}A at base-resolution (Fig. 4). However, MjDim1 exhibits strong sequence biases and poor selectivity of m^{6}A over A, as well as high false positives. To achieve stoichiometric m^{6}A measurement, calibration curves must be generated using spike-in probes with varying modification fractions. Nevertheless, m^{6}A-SAC-Seq captures more than 10 000 m^{6}A sites in HEK293, HeLa and HepG2 cells. This study revealed dynamic m^{6}A modification stoichiometry during human hematopoietic stem and progenitor cells (HSPCs) differentiation.

Glyoxal and nitrite-mediated deamination of unmethylated adenosines (GLORI) [8] (Fig. 2) is the first indirect (similar to bisulfite sequencing for DNA methylation), base-resolution, and quantitative m^{6}A sequencing method. This method is based on nitrous acid mediated adenosine deamination, which had been discovered half a century ago [123, 124]. To overcome nitrous acid's low reaction efficiency towards A but a high rate of G deamination, GLORI uses glyoxal to protect G by forming an N^1, N^2-dihydroxyguanosine adduct (less than 3% G will convert after protection), which surprisingly also acts as a catalyst to boost the A to I (inosine) conversion (98–99%). m^{6}A does not undergo nitrous-acid-mediated deamination, which allows it to be identified as the remaining A sites during sequencing (Fig. 4). GLORI captures 176 642 m^{6}A modification sites in HEK293T cells (at 140x coverage) with around 40% median methylation level [8]. Most m^{6}A sites reside in the canonical DRAC (D = G/A/T, R = A/G) motifs, while one third of m^{6}A sites occur within clusters. As a chemical method, whether GLORI can be applied to low-input samples remains to be tested. Nevertheless, GLORI successfully charted the impact of hypoxia and heat shock conditions on the dynamic modification of m^{6}A, thereby suggesting distinct regulatory mechanisms governing m^{6}A's role in gene expression through its influence on translation efficiency.

Evolved TadA-assisted N^6-methyladenosine sequencing (eTAM-Seq) [9] (Fig. 2) is an enzymatic equivalent of GLORI, which uses a hyperactive transfer RNA adenosine deaminase (TadA) variant TadA8.20 to achieve up to 99% global adenosine-to-inosine deamination (Fig. 4). It identified 80 941 and 65 835 m6A sites in HeLa and mESC, respectively, and can be applied to RNA input quantities as low as 250 pg (10 cells) for loci-specific detection. It also unveiled an inverse relationship existing between m^{6}A modification and the stability of mRNA molecules. However, due to the sensitivity of TadA8.20 to secondary structures, eTAM-seq requires control transcriptomes to eliminate false positives and may be less accurate at lowly methylated sites (<25%).

RNA Ψ MODIFICATION SEQUENCING TECHNOLOGIES

Ψ, also known as the 'fifth nucleotide' in RNA, was first identified in 1951. It is the most abundant modification in total RNA [125] and has been known to exist in tRNA, rRNA, and snRNAs for decades, whilst being recently found in mRNA. Ψ is generated by pseudouridine synthases (PUS) and accounts for ~0.2% of uridine in mammalian mRNA [10, 126, 127]. Ψ's base pairing is similar to uridine, being a structural isomer. Most methods for Ψ mapping rely on the labeling reaction of Ψ by *N*-cyclohexyl-*N*′-(2-morpholinoethyl) carbodiimide methyl-*p*-toluenesulfonate (CMC). The CMC-Ψ adduct can generate a stop signature during reverse transcription. These technologies, including Ψ-Seq [128], Pseudo-Seq [129], pseudouridine site identification sequencing (PSI-Seq) [130], and N_3-CMC-enriched psedouridine sequencing (CeU-Seq) [126], however, lack stoichiometric information of Ψ (Fig. 2).

Recently, bisulfite-induced deletion sequencing (BID-Seq) [10] and PRAISE [11] (Fig. 2) have been developed using a bisulfite-mediated reaction on Ψ [131, 132]. Using reaction conditions at neutral pH, they can efficiently convert Ψ to Ψ-bisulfite (Ψ-BS) adduct without C to U conversion. The Ψ-bisulfite can cause

deletion during reverse transcription, with fully modified Ψ generating about 70–90% detection rate (Fig. 4). With the help of calibration curves from spike-in controls, the BID-Seq quantified over 500 abundant Ψ modification sites in human HeLa, HEK293T, and A549 cells, which are enriched in the coding sequence (CDS) and 3′-UTR. In this study, Ψ modification is shown to be positively correlated with mRNA stability, expression and promotes read-through of stop codons [10]. The PRAISE study, on the other hand, identified 2 209 confident Ψ sites within the HEK293T cell line [11].

OTHER RNA METHYLATION MODIFICATIONS ($M^1A/M^5C/HM^5C/M^3C/M^7G$)

Numerous other modifications exist in RNA, usually in low abundance. As an isomer of m^6A, the established m^1A-ID-Seq [133], m^1A-Seq [134], m^1A-MAP [135], and m^1A-Seq-TGIRT [136] (Fig. 2) can be used to detect m^1A modifications. Similar to DNA methylation, m^5C, and hm^5C have also been found in mRNA. m^5C modification can be detected by modified bisulfite sequencing [137], Aza-IP [138], and miCLIP-Seq [139] (Fig. 2). m^7G is another RNA methylation which can be detected by m^7G-MeRIP-Seq [140], m^7G miCLIP-Seq [141], AlkAniline-Seq [142], m^7G-MAP-Seq [143], and m^7G-quant-Seq [144] (Fig. 2). 3-Methylcytidine (m^3C) can be detected by AlkAniline-Seq [142] and HAC-Seq [145] (Fig. 2).

THIRD GENERATION SEQUENCING FOR DNA AND RNA MODIFICATIONS

PacBio Single-Molecule Real-Time (SMRT) sequencing by Pacific Biosciences [25, 146, 147] and Oxford Nanopore sequencing [148–150] are novel third-generation technologies, which enable long-read sequencing and can directly detect native DNA and RNA modifications [12]. SMRT sequencing uses polymerase kinetics to discriminate different types of DNA modifications whilst nanopore sequencing utilizes the ionic current changes that occur when different nucleotides transit the nanopore channel to discriminate various DNA and RNA modifications. While the Nanopore platform has demonstrated its ability to detect various nucleic acid modifications, such as 5mC, 5hmC in DNA and m^6A, m^5C in RNA [151–157], the SMRT platform exhibits promise in the identification of RNA modification (m^6A) through the analysis of fluorescent signal alterations during reverse transcription. However, its reliability currently remains limited [158, 159]. While they hold great potential for epigenomic and epitranscriptomic sequencing, at present, significant challenges and limitations remain, including high error rate, high cost, and high-sample input requirement. Currently, combining chemical or enzymatic reactions of DNA and RNA modifications with third generation sequencing could provide a solution to enhance their detection abilities [26, 160–163].

CONCLUDING REMARKS

In this review, we have described the various sequencing technologies for histone, DNA, and RNA modifications. A modification is best determined by an absolute quantitative sequencing method at base resolution. DNA methylation was traditionally achieved by bisulfite sequencing. Recent advances in this area resulted in EM-Seq and TAPS overcoming the issues of bisulfite sequencing. Compared to the more mature DNA epigenetic sequencing, recent efforts have started shifting to RNA epitranscriptomic sequencing. Very recently, breakthroughs such as m^6A-SAC-Seq, GLORI, BID-Seq, and eTAM-Seq started to deliver base-resolution and quantitative sequencing of RNA modifications. Their efficiencies and accuracies are still lacking compared to DNA methylation sequencing methods, highlighted by the fact that different methods for the same modification often produce very distinct profiles (e.g., m^6A and Ψ maps of the same cell line obtained by different sequencing methods showed low overlaps) [8–11]. A non-modified negative control (such as in vitro transcribed RNA) could be beneficial to eliminate false positives [8, 9, 11, 164]. We expect rapid development to continue in the field.

Future development of chemical methods and computational analysis in third-generation sequencing technologies will continue to improve their epigenomic and epitranscriptomic sequencing ability. It remains to be seen whether they can achieve the same accuracy and cost-effectiveness as next-generation sequencing in sequencing DNA and RNA modifications. Another potential of third-generation sequencing is for the simultaneous detection of multiple epigenetic or epitranscriptomic modifications in a single molecule [132, 165]. Such information could help reveal the interplay between different modifications.

With the rapid development of temporal/spatial DNA and RNA sequencing in recent years, including techniques such as Slide-Seq [166, 167] and Live-Seq [168], researchers have successfully examined the DNA and RNA molecules in their natural locations and in a dynamic manner. These advanced detection technologies also offer valuable insights into their potential application in uncovering epigenetic modifications. While the ability to investigate biological phenomena in real time and within their spatial context is of utmost significance for comprehensive understanding in the field of biology, it is tempting to envision future large-scale live cell temporal/spatial epigenetic sequencing to further enable biological discoveries.

REFERENCES

1. Zhao Y, Garcia BA. Comprehensive catalog of currently documented histone modifications. Cold Spring Harb Perspect Biol. 2015;7:a025064.
2. Gommers-Ampt JH, Borst P. Hypermodified bases in DNA. FASEB J. 1995;9:1034–42.
3. Boccaletto P, Machnicka MA, Purta E, Piatkowski P, Baginski B, Wirecki TK, et al. MODOMICS: a database of RNA modification pathways. 2017 update. Nucleic Acids Res. 2018;46:D303–D7.
4. Tanaka K, Okamoto A. Degradation of DNA by bisulfite treatment. Bioorg Med Chem Lett. 2007;17:1912–5.
5. Liu Y, Siejka-Zielinska P, Velikova G, Bi Y, Yuan F, Tomkova M, et al. Bisulfite-free direct detection of 5-methylcytosine and 5-hydroxymethylcytosine at base resolution. Nat Biotechnol. 2019;37:424–9.
6. Vaisvila R, Ponnaluri VKC, Sun Z, Langhorst BW, Saleh L, Guan S, et al. Enzymatic methyl sequencing detects DNA methylation at single-base resolution from picograms of DNA. Genome Res. 2021;31:1280–9.
7. Hu L, Liu S, Peng Y, Ge R, Su R, Senevirathne C, et al. m(6)A RNA modifications are measured at single-base resolution across the mammalian transcriptome. Nat Biotechnol. 2022;40:1210–9.
8. Liu C, Sun H, Yi Y, Shen W, Li K, Xiao Y, et al. Absolute quantification of single-base m(6)A methylation in the mammalian transcriptome using GLORI. Nat Biotechnol. 2023;41:355–66.
9. Xiao Y-L, Liu S, Ge R, Wu Y, He C, Chen M, et al. Transcriptome-wide profiling and quantification of N6-methyladenosine by enzyme-assisted adenosine deamination. Nat Biotechnol. 2023;41:993–1003.
10. Dai Q, Zhang LS, Sun HL, Pajdzik K, Yang L, Ye C, et al. Quantitative sequencing using BID-seq uncovers abundant pseudouridines in mammalian mRNA at base resolution. Nat Biotechnol. 2023;41:344–54.
11. Zhang M, Jiang Z, Ma Y, Liu W, Zhuang Y, Lu B, et al. Quantitative profiling of pseudouridylation landscape in the human transcriptome. Nat Chem Biol. 2023.
12. Lucas MC, Novoa EM. Long-read sequencing in the era of epigenomics and epitranscriptomics. Nat Methods. 2023;20:25–9.
13. Deichmann U. Epigenetics: the origins and evolution of a fashionable topic. Dev Biol. 2016;416:249–54.
14. Kouzarides T. Chromatin modifications and their function. Cell. 2007;128:693–705.
15. Zentner GE, Henikoff S. Regulation of nucleosome dynamics by histone modifications. Nat Struct Mol Biol. 2013;20:259–66.
16. Moore LD, Le T, Fan G. DNA methylation and its basic function. Neuropsychopharmacology. 2013;38:23–38.
17. Arzumanian VA, Dolgalev GV, Kurbatov IY, Kiseleva OI, Poverennaya EV Epitranscriptome: review of top 25 most-studied RNA modifications. Int J Mol Sci. 2022;23:13851.
18. Barbieri I, Kouzarides T. Role of RNA modifications in cancer. Nat Rev Cancer. 2020;20:303–22.

19. Zhao LY, Song J, Liu Y, Song CX, Yi C. Mapping the epigenetic modifications of DNA and RNA. Protein Cell. 2020;11:792–808.
20. Fu Y, He C. Nucleic acid modifications with epigenetic significance. Curr Opin Chem Biol. 2012;16:516–24.
21. Consortium ITP-CAoWG. Pan-cancer analysis of whole genomes. Nature. 2020;578:82–93.
22. Ley TJ, Ding L, Walter MJ, McLellan MD, Lamprecht T, Larson DE, et al. DNMT3A mutations in acute myeloid leukemia. N Engl J Med. 2010;363:2424–33.
23. Cancer Genome Atlas Research N, Ley TJ, Miller C, Ding L, Raphael BJ, Mungall AJ, et al. Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med. 2013;368:2059–74.
24. Dai Y, Yuan BF, Feng YQ. Quantification and mapping of DNA modifications. RSC Chem Biol. 2021;2:1096–114.
25. Rhoads A, Au KF. PacBio sequencing and its applications. Genom Proteom Bioinforma. 2015;13:278–89.
26. Searle B, Muller M, Carell T, Kellett A. Third-generation sequencing of epigenetic DNA. Angew Chem Int Ed Engl. 2022;62:e202215704.
27. Ruiz-Carrillo A, Jorcano JL. An octamer of core histones in solution: central role of the H3-H4 tetramer in the self-assembly. Biochemistry. 1979;18:760–8.
28. Bartova E, Krejci J, Harnicarova A, Galiova G, Kozubek S. Histone modifications and nuclear architecture: a review. J Histochem Cytochem. 2008;56:711–21.
29. Zhou VW, Goren A, Bernstein BE. Charting histone modifications and the functional organization of mammalian genomes. Nat Rev Genet. 2011;12:7–18.
30. Zhao Z, Shilatifard A. Epigenetic modifications of histones in cancer. Genome Biol. 2019;20:245.
31. Rosenfeld JA, Wang Z, Schones DE, Zhao K, DeSalle R, Zhang MQ. Determination of enriched histone modifications in non-genic portions of the human genome. BMC Genom. 2009;10:143.
32. Messier TL, Gordon JA, Boyd JR, Tye CE, Browne G, Stein JL, et al. Histone H3 lysine 4 acetylation and methylation dynamics define breast cancer subtypes. Oncotarget. 2016;7:5094–109.
33. Esteller M. Cancer epigenomics: DNA methylomes and histone-modification maps. Nat Rev Genet. 2007;8:286–98.
34. Portela A, Esteller M. Epigenetic modifications and human disease. Nat Biotechnol. 2010;28:1057–68.
35. Solomon MJ, Larsen PL, Varshavsky A. Mapping protein-DNA interactions in vivo with formaldehyde: evidence that histone H4 is retained on a highly transcribed gene. Cell. 1988;53:937–47.
36. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, et al. High-resolution profiling of histone methylations in the human genome. Cell. 2007;129:823–37.
37. Park PJ. ChIP-seq: advantages and challenges of a maturing technology. Nat Rev Genet. 2009;10:669–80.
38. Skene PJ, Henikoff S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. Elife. 2017;6:e21856.
39. Meers MP, Bryson TD, Henikoff JG, Henikoff S. Improved CUT&RUN chromatin profiling tools. Elife. 2019;8:e46314.
40. Kaya-Okur HS, Wu SJ, Codomo CA, Pledger ES, Bryson TD, Henikoff JG, et al. CUT&Tag for efficient epigenomic profiling of small samples and single cells. Nat Commun. 2019;10:1930.
41. Bartosovic M, Kabbe M, Castelo-Branco G. Single-cell CUT&Tag profiles histone modifications and transcription factors in complex tissues. Nat Biotechnol. 2021;39:825–35.
42. Song CX, He C. Potential functional roles of DNA demethylation intermediates. Trends Biochem Sci. 2013;38:480–4.
43. Bird A. DNA methylation patterns and epigenetic memory. Genes Dev. 2002;16:6–21.
44. Reik W, Dean W, Walter J. Epigenetic reprogramming in mammalian development. Science. 2001;293:1089–93.
45. Esteller M. CpG island hypermethylation and tumor suppressor genes: a booming present, a brighter future. Oncogene. 2002;21:5427–40.
46. Feinberg AP, Vogelstein B. Hypomethylation of ras oncogenes in primary human cancers. Biochem Biophys Res Commun. 1983;111:47–54.
47. Van Tongelen A, Loriot A, De Smet C. Oncogenic roles of DNA hypomethylation through the activation of cancer-germline genes. Cancer Lett. 2017;396:130–7.
48. Roulois D, Loo Yau H, Singhania R, Wang Y, Danesh A, Shen SY, et al. DNA-demethylating agents target colorectal cancer cells by inducing viral mimicry by endogenous transcripts. Cell. 2015;162:961–73.
49. Blagitko-Dorfs N, Schlosser P, Greve G, Pfeifer D, Meier R, Baude A, et al. Combination treatment of acute myeloid leukemia cells with DNMT and HDAC inhibitors: predominant synergistic gene downregulation associated with gene body demethylation. Leukemia. 2019;33:945–56.
50. Kelly TK, De Carvalho DD, Jones PA. Epigenetic modifications as therapeutic targets. Nat Biotechnol. 2010;28:1069–78.
51. Salehi B, Selamoglu Z, S Mileski K, Pezzani R, Redaelli M, Cho WC, et al. Liposomal cytarabine as cancer therapy: from chemistry to medicine. Biomolecules. 2019;9:773.
52. Chen X, Zhou W, Song RH, Liu S, Wang S, Chen Y, et al. Tumor suppressor CEBPA interacts with and inhibits DNMT3A activity. Sci Adv. 2022;8:eabl5220.
53. Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, et al. Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. Nat Genet. 2005;37:853–62.
54. Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, et al. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. Proc Natl Acad Sci USA. 1992;89:1827–31.
55. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462:315-22.
56. Miura F, Enomoto Y, Dairiki R, Ito T. Amplification-free whole-genome bisulfite sequencing by post-bisulfite adaptor tagging. Nucleic Acids Res. 2012;40:e136.
57. Clark SJ, Smallwood SA, Lee HJ, Krueger F, Reik W, Kelsey G. Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). Nat Protoc. 2017;12:534–47.
58. Farlik M, Sheffield NC, Nuzzo A, Datlinger P, Schonegger A, Klughammer J, et al. Single-cell DNA methylome sequencing and bioinformatic inference of epigenomic cell-state dynamics. Cell Rep. 2015;10:1386–97.
59. Wang J, Fang Y, Chen W, Zhang C, Chen Z, Xie Z, et al. High coverage of single cell genomes by T7-assisted enzymatic methyl-sequencing. bioRxiv. 2022:2022.02.23.481567.
60. Liu Y, Song CX. TAPS: the development of a direct and base-resolution sequencing method for DNA methylation. ACS Chem Biol. 2022;17:2683–5.
61. Liu Y, Hu Z, Cheng J, Siejka-Zielinska P, Chen J, Inoue M, et al. Subtraction-free and bisulfite-free specific sequencing of 5-methylcytosine and its oxidized derivatives at base resolution. Nat Commun. 2021;12:618.
62. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, et al. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science. 2009;324:930–5.
63. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, et al. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. Science .2011;333:1303–7.
64. Onodera A, Gonzalez-Avalos E, Lio CJ, Georges RO, Bellacosa A, Nakayama T, et al. Roles of TET and TDG in DNA demethylation in proliferating and non-proliferating immune cells. Genome Biol. 2021;22:186.
65. Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, et al. Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. Nucleic Acids Res. 2012;40:4841–9.
66. Kohli RM, Zhang Y. TET enzymes, TDG and the dynamics of DNA demethylation. Nature. 2013;502:472–9.
67. Kriaucionis S, Heintz N. The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. Science. 2009;324:929–30.
68. Mellen M, Ayata P, Dewell S, Kriaucionis S, Heintz N. MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. Cell. 2012;151:1417–30.
69. Bachman M, Uribe-Lewis S, Yang X, Burgess HE, Iurlaro M, Reik W, et al. 5-Formylcytosine can be a stable DNA modification in mammals. Nat Chem Biol. 2015;11:555–7.
70. Inoue A, Shen L, Dai Q, He C, Zhang Y. Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development. Cell Res. 2011;21:1670–6.
71. Inoue A, Zhang Y. Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos. Science. 2011;334:194.
72. Wossidlo M, Nakamura T, Lepikhov K, Marques CJ, Zakhartchenko V, Boiani M, et al. 5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming. Nat Commun. 2011;2:241.
73. Cimmino L, Abdel-Wahab O, Levine RL, Aifantis I. TET family proteins and their role in stem cell differentiation and transformation. Cell Stem Cell. 2011;9:193–204.
74. Ficz G, Gribben JG. Loss of 5-hydroxymethylcytosine in cancer: cause or consequence? Genomics. 2014;104:352–7.
75. Pfeifer GP, Xiong W, Hahn MA, Jin SG. The role of 5-hydroxymethylcytosine in human cancer. Cell Tissue Res. 2014;356:631–41.
76. Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, et al. Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. Nature. 2011;473:398-402.
77. Song CX, Szulwach KE, Fu Y, Dai Q, Yi C, Li X, et al. Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. Nat Biotechnol. 2011;29:68–72.
78. Yu M, Hon GC, Szulwach KE, Song CX, Jin P, Ren B, et al. Tet-assisted bisulfite sequencing of 5-hydroxymethylcytosine. Nat Protoc. 2012;7:2159–70.
79. Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, et al. Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. Cell. 2012;149:1368–80.

80. Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, et al. Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Science. 2012;336:934–7.
81. Booth MJ, Ost TW, Beraldi D, Bell NM, Branco MR, Reik W, et al. Oxidative bisulfite sequencing of 5-methylcytosine and 5-hydroxymethylcytosine. Nat Protoc. 2013;8:1841–51.
82. Schutsky EK, DeNizio JE, Hu P, Liu MY, Nabel CS, Fabyanic EB, et al. Non-destructive, base-resolution sequencing of 5-hydroxymethylcytosine using a DNA deaminase. Nat Biotechnol. 2018;36:1083–1090.
83. Zeng H, He B, Xia B, Bai D, Lu X, Cai J, et al. Bisulfite-free, nanoscale analysis of 5-hydroxymethylcytosine at single base resolution. J Am Chem Soc. 2018;140:13190–4.
84. Xu H, Chen J, Cheng J, Kong L, Chen X, Inoue M, et al. Modular oxidation of cytosine modifications and their application in direct and quantitative sequencing of 5-hydroxymethylcytosine. J Am Chem Soc. 2023;145:7095–100.
85. Wu X, Zhang Y. TET-mediated active DNA demethylation: mechanism, function and beyond. Nat Rev Genet. 2017;18:517–34.
86. Song CX, Szulwach KE, Dai Q, Fu Y, Mao SQ, Lin L, et al. Genome-wide profiling of 5-formylcytosine reveals its roles in epigenetic priming. Cell. 2013;153:678–91.
87. Booth MJ, Marsico G, Bachman M, Beraldi D, Balasubramanian S. Quantitative sequencing of 5-formylcytosine in DNA at single-base resolution. Nat Chem. 2014;6:435–40.
88. Wu H, Wu X, Shen L, Zhang Y. Single-base resolution analysis of active DNA demethylation using methylase-assisted bisulfite sequencing. Nat Biotechnol. 2014;32:1231–40.
89. Wu H, Wu X, Zhang Y. Base-resolution profiling of active DNA demethylation using MAB-seq and caMAB-seq. Nat Protoc. 2016;11:1081–100.
90. Xia B, Han D, Lu X, Sun Z, Zhou A, Yin Q, et al. Bisulfite-free, base-resolution analysis of 5-formylcytosine at the genome scale. Nat Methods. 2015;12:1047–50.
91. Zhu C, Gao Y, Guo H, Xia B, Song J, Wu X, et al. Single-Cell 5-Formylcytosine Landscapes of Mammalian Early Embryos and ESCs at Single-Base Resolution. Cell Stem Cell. 2017;20:720–31.e5.
92. Yao B, Cheng Y, Wang Z, Li Y, Chen L, Huang L, et al. DNA N6-methyladenine is dynamically regulated in the mouse brain following environmental stress. Nat Commun. 2017;8:1122.
93. Li H, Zhang N, Wang Y, Xia S, Zhu Y, Xing C, et al. DNA N6-methyladenine modification in eukaryotic genome. Front Genet. 2022;13:914404.
94. Liu J, Zhu Y, Luo GZ, Wang X, Yue Y, Wang X, et al. Abundant DNA 6mA methylation during early embryogenesis of zebrafish and pig. Nat Commun. 2016;7:13052.
95. O'Brown ZK, Boulias K, Wang J, Wang SY, O'Brown NM, Hao Z, et al. Sources of artifact in measurements of 6mA and 4mC abundance in eukaryotic genomic DNA. BMC Genom. 2019;20:445.
96. Douvlataniotis K, Bensberg M, Lentini A, Gylemo B, Nestor CE. No evidence for DNA N (6)-methyladenine in mammals. Sci Adv. 2020;6:eaay3335.
97. Borst P, Sabatini R. Base J: discovery, biosynthesis, and possible functions. Annu Rev Microbiol. 2008;62:235–51.
98. Desrosiers R, Friderici K, Rottman F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. Proc Natl Acad Sci USA. 1974;71:3971–5.
99. Niu Y, Zhao X, Wu YS, Li MM, Wang XJ, Yang YG. N6-methyl-adenosine (m6A) in RNA: an old modification with a novel epigenetic function. Genom Proteom Bioinforma. 2013;11:8–17.
100. Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Jin Z, et al. The role of m6A modification in the biological functions and diseases. Signal Transduct Target Ther. 2021;6:74.
101. Ma Z, Ji J. N6-methyladenosine (m6A) RNA modification in cancer stem cells. Stem Cells. 2020;38:1511–1519.
102. Feng Q, Wang D, Xue T, Lin C, Gao Y, Sun L, et al. The role of RNA modification in hepatocellular carcinoma. Front Pharm. 2022;13:984453.
103. Sun T, Wu R, Ming L. The role of m6A RNA methylation in cancer. Biomed Pharmacother. 2019;112:108613.
104. He PC, He C. m(6) A RNA methylation: from mechanisms to therapeutic potential. EMBO J. 2021;40:e105977.
105. Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. Nat Chem Biol. 2014;10:93–5.
106. Ping XL, Sun BF, Wang L, Xiao W, Yang X, Wang WJ, et al. Mammalian WTAP is a regulatory subunit of the RNA N6-methyladenosine methyltransferase. Cell Res. 2014;24:177–89.
107. Wen J, Lv R, Ma H, Shen H, He C, Wang J, et al. Zc3h13 regulates nuclear RNA m(6)A methylation and mouse embryonic stem cell self-renewal. Mol Cell. 2018;69:1028–38.e6.
108. Xiao W, Adhikari S, Dahal U, Chen YS, Hao YJ, Sun BF, et al. Nuclear m(6)A reader YTHDC1 regulates mRNA splicing. Mol Cell. 2016;61:507–19.
109. Yue Y, Liu J, Cui X, Cao J, Luo G, Zhang Z, et al. VIRMA mediates preferential m(6) A mRNA methylation in 3'UTR and near stop codon and associates with alternative polyadenylation. Cell Discov. 2018;4:10.
110. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol. 2011;7:885–7.
111. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang CM, Li CJ, et al. ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. Mol Cell. 2013;49:18–29.
112. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature. 2012;485:201–6.
113. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. Cell. 2012;149:1635–46.
114. Chen K, Luo GZ, He C. High-resolution mapping of N(6)-methyladenosine in transcriptome and genome using a photo-crosslinking-assisted strategy. Methods Enzymol. 2015;560:161–85.
115. Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. Nat Methods. 2015;12:767–72.
116. Molinie B, Wang J, Lim KS, Hillebrand R, Lu ZX, Van Wittenberghe N, et al. m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. Nat Methods. 2016;13:692–8.
117. Dierks D, Garcia-Campos MA, Uzonyi A, Safra M, Edelheit S, Rossi A, et al. Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution. Nat Methods. 2021;18:1060–7.
118. Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, et al. Deciphering the "m(6)A Code" via antibody-independent quantitative profiling. Cell. 2019;178:731–47.e16.
119. Zhang Z, Chen LQ, Zhao YL, Yang CG, Roundtree IA, Zhang Z, et al. Single-base mapping of m(6)A by an antibody-independent method. Sci Adv. 2019;5:eaax0250.
120. Meyer KD. DART-seq: an antibody-free method for global m(6)A detection. Nat Methods. 2019;16:1275–80.
121. Wang Y, Xiao Y, Dong S, Yu Q, Jia G. Antibody-free enzyme-assisted chemical approach for detection of N(6)-methyladenosine. Nat Chem Biol. 2020;16:896–903.
122. Shu X, Cao J, Cheng M, Xiang S, Gao M, Li T, et al. A metabolic labeling method detects m(6)A transcriptome-wide at single base resolution. Nat Chem Biol. 2020;16:887–95.
123. Shapiro R, Pohl SH. The reaction of ribonucleosides with nitrous acid. Side-Products Kinet Biochem. 1968;7:448–55.
124. Schuster H, Wilhelm RC. Reaction differences between tobacco mosaic virus and its free ribonucleic acid with nitrous acid. Biochim Biophys Acta. 1963;68:554–60.
125. Shi H, Chai P, Jia R, Fan X. Novel insight into the regulatory roles of diverse RNA modifications: Re-defining the bridge between transcription and translation. Mol Cancer. 2020;19:78.
126. Li X, Zhu P, Ma S, Song J, Bai J, Sun F, et al. Chemical pulldown reveals dynamic pseudouridylation of the mammalian transcriptome. Nat Chem Biol. 2015;11:592–7.
127. Zhang LS, Dai Q, He C. BID-seq: The Quantitative and Base-Resolution Sequencing Method for RNA Pseudouridine. ACS Chem Biol. 2023;18:4-6.
128. Schwartz S, Bernstein DA, Mumbach MR, Jovanovic M, Herbst RH, Leon-Ricardo BX, et al. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. Cell. 2014;159:148–62.
129. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature. 2014;515:143–6.
130. Lovejoy AF, Riordan DP, Brown PO. Transcriptome-wide mapping of pseudouridines: pseudouridine synthases modify specific mRNAs in S. cerevisiae. PLoS One. 2014;9:e110799.
131. Fleming AM, Alenko A, Kitt JP, Orendt AM, Flynn PF, Harris JM, et al. Structural elucidation of bisulfite adducts to pseudouridine that result in deletion signatures during reverse transcription of RNA. J Am Chem Soc. 2019;141:16450–60.
132. Khoddami V, Yerra A, Mosbruger TL, Fleming AM, Burrows CJ, Cairns BR. Transcriptome-wide profiling of multiple RNA modifications simultaneously at single-base resolution. Proc Natl Acad Sci USA. 2019;116:6784–9.
133. Li X, Xiong X, Wang K, Wang L, Shu X, Ma S, et al. Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. Nat Chem Biol. 2016;12:311–6.
134. Dominissini D, Nachtergaele S, Moshitch-Moshkovitz S, Peer E, Kol N, Ben-Haim MS, et al. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. Nature. 2016;530:441–6.
135. Li X, Xiong X, Zhang M, Wang K, Chen Y, Zhou J, et al. Base-resolution mapping reveals distinct m(1)A methylome in nuclear- and mitochondrial-encoded transcripts. Mol Cell. 2017;68:993–1005.e9.

136. Safra M, Sas-Chen A, Nir R, Winkler R, Nachshon A, Bar-Yaacov D, et al. The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. Nature. 2017;551:251–5.
137. Pollex T, Hanna K, Schaefer M Detection of cytosine methylation in RNA using bisulfite sequencing. Cold Spring Harb Protoc. 2010;6.
138. Chen YS, Yang WL, Zhao YL, Yang YG. Dynamic transcriptomic m(5) C and its regulatory role in RNA processing. Wiley Interdiscip Rev RNA. 2021;12:e1639.
139. Hussain S, Sajini AA, Blanco S, Dietmann S, Lombard P, Sugimoto Y, et al. NSun2-mediated cytosine-5 methylation of vault noncoding RNA determines its processing into regulatory small RNAs. Cell Rep. 2013;4:255–61.
140. Li W, Li X, Ma X, Xiao W, Zhang J. Mapping the m1A, m5C, m6A and m7G methylation atlas in zebrafish brain under hypoxic conditions by MeRIP-seq. BMC Genom. 2022;23:105.
141. Malbec L, Zhang T, Chen YS, Zhang Y, Sun BF, Shi BY, et al. Dynamic methylome of internal mRNA N(7)-methylguanosine and its regulatory role in translation. Cell Res. 2019;29:927–41.
142. Marchand V, Ayadi L, Ernst FGM, Hertler J, Bourguignon-Igel V, Galvanin A, et al. AlkAniline-Seq: profiling of m(7) G and m(3) C RNA modifications at single nucleotide resolution. Angew Chem Int Ed Engl. 2018;57:16785–90.
143. Enroth C, Poulsen LD, Iversen S, Kirpekar F, Albrechtsen A, Vinther J. Detection of internal N7-methylguanosine (m7G) RNA modifications by mutational profiling sequencing. Nucleic Acids Res. 2019;47:e126.
144. Zhang LS, Ju CW, Liu C, Wei J, Dai Q, Chen L, et al. m(7)G-quant-seq: quantitative detection of RNA internal N(7)-methylguanosine. ACS Chem Biol. 2022;17:3306–12.
145. Cui J, Liu Q, Sendinc E, Shi Y, Gregory RI. Nucleotide resolution profiling of m3C RNA modification by HAC-seq. Nucleic Acids Res. 2021;49:e27.
146. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, et al. Direct detection of DNA methylation during single-molecule, real-time sequencing. Nat Methods. 2010;7:461–5.
147. Saletore Y, Meyer K, Korlach J, Vilfan ID, Jaffrey S, Mason CE. The birth of the Epitranscriptome: deciphering the function of RNA modifications. Genome Biol. 2012;13:175.
148. Laszlo AH, Derrington IM, Brinkerhoff H, Langford KW, Nova IC, Samson JM, et al. Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. Proc Natl Acad Sci USA. 2013;110:18904–9.
149. Rand AC, Jain M, Eizenga JM, Musselman-Brown A, Olsen HE, Akeson M, et al. Mapping DNA methylation with high-throughput nanopore sequencing. Nat Methods. 2017;14:411–3.
150. Wallace EV, Stoddart D, Heron AJ, Mikhailova E, Maglia G, Donohoe TJ, et al. Identification of epigenetic DNA modifications with a protein nanopore. Chem Commun (Camb). 2010;46:8195–7.
151. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA sequencing on an array of nanopores. Nat Methods. 2018;15:201–6.
152. Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, et al. Accurate detection of m(6)A RNA modifications in native RNA sequences. Nat Commun. 2019;10:4079.
153. Simpson JT, Workman RE, Zuzarte PC, David M, Dursi LJ, Timp W. Detecting DNA cytosine methylation using nanopore sequencing. Nat Methods. 2017;14:407–10.
154. Pratanwanich PN, Yao F, Chen Y, Koh CWQ, Wan YK, Hendra C, et al. Identification of differential RNA modifications from nanopore direct RNA sequencing with xPore. Nat Biotechnol. 2021;39:1394–402.
155. Schadt EE, Banerjee O, Fang G, Feng Z, Wong WH, Zhang X, et al. Modeling kinetic rate variation in third generation DNA sequencing data to detect putative modifications to DNA bases. Genome Res. 2013;23:129–41.
156. Parker MT, Knop K, Sherwood AV, Schurch NJ, Mackinnon K, Gould PD, et al. Nanopore direct RNA sequencing maps the complexity of Arabidopsis mRNA processing and m(6)A modification. Elife. 2020;9:e49658.
157. Begik O, Lucas MC, Pryszcz LP, Ramirez JM, Medina R, Milenkovic I, et al. Quantitative profiling of pseudouridylation dynamics in native RNAs with nanopore sequencing. Nat Biotechnol. 2021;39:1278–91.
158. Huang G, Ding Q, Xie D, Cai Z, Zhao Z. Technical challenges in defining RNA modifications. Semin Cell Dev Biol. 2022;127:155–65.
159. Vilfan ID, Tsai YC, Clark TA, Wegener J, Dai Q, Yi C, et al. Analysis of RNA base modification and structural rearrangement by single-molecule real-time detection of reverse transcription. J Nanobiotechnol. 2013;11:8.
160. Sun Z, Vaisvila R, Hussong LM, Yan B, Baum C, Saleh L, et al. Nondestructive enzymatic deamination enables single-molecule long-read amplicon sequencing for the determination of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. Genome Res. 2021;31:291–300.
161. Sakamoto Y, Zaha S, Nagasawa S, Miyake S, Kojima Y, Suzuki A, et al. Long-read whole-genome methylation patterning using enzymatic base conversion and nanopore sequencing. Nucleic Acids Res. 2021;49:e81.
162. Chen J, Cheng J, Chen X, Inoue M, Liu Y, Song CX. Whole-genome long-read TAPS deciphers DNA methylation patterns at base resolution using PacBio SMRT sequencing technology. Nucleic Acids Res. 2022;50:e104.
163. Liu Y, Cheng J, Siejka-Zielinska P, Weldon C, Roberts H, Lopopolo M, et al. Accurate targeted long-read DNA methylation and hydroxymethylation sequencing with TAPS. Genome Biol. 2020;21:54.
164. Zhang Z, Chen T, Chen HX, Xie YY, Chen LQ, Zhao YL, et al. Systematic calibration of epitranscriptomic maps using a synthetic modification-free RNA library. Nat Methods. 2021;18:1213–22.
165. Fullgrabe J, Gosal WS, Creed P, Liu S, Lumby CK, Morley DJ, et al. Simultaneous sequencing of genetic and epigenetic bases in DNA. Nat Biotechnol. 2023.
166. Zhao T, Chiang ZD, Morriss JW, LaFave LM, Murray EM, Del Priore I, et al. Spatial genomics enables multi-modal study of clonal heterogeneity in tissues. Nature. 2022;601:85–91.
167. Rodriques SG, Stickels RR, Goeva A, Martin CA, Murray E, Vanderburg CR, et al. Slide-seq: a scalable technology for measuring genome-wide expression at high spatial resolution. Science. 2019;363:1463–7.
168. Chen W, Guillaume-Gentil O, Rainer PY, Gabelein CG, Saelens W, Gardeux V, et al. Live-seq enables temporal transcriptomic recording of single cells. Nature. 2022;608:733–40.

ACKNOWLEDGEMENTS

We would also like to acknowledge K. Dunning for editing the manuscript, H.H. and Y.-Q.Y. for discussion on histone modifications.

AUTHOR CONTRIBUTIONS

X.C. and C.-X.S. wrote the manuscript with the help of H.X. and X.S.

FUNDING

Research in our lab is supported by Ludwig Institute for Cancer Research, Cancer Research UK (C63763/A26394), and National Institute for Health Research (NIHR) Oxford Biomedical Research Centre (BRC). H.X. is supported by China Scholarship Council. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, or the Department of Health. We apologize for not being able to cite all the publications related to this topic due to the space constraints of the journal.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to Chun-Xiao Song.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.