Fundamental Research 3 (2023) 657–664

Contents lists available at ScienceDirect

Fundamental Research

journal homepage: http://www.keaipublishing.com/en/journals/fundamental-research/

Article

a^{6}A-seq: N^6-allyladenosine-based cellular messenger RNA metabolic labelling and sequencing

Xiao Shu[a], Chenyang Huang[a], Tengwei Li[a], Jie Cao[a,b], Jianzhao Liu[a,b,*]

[a] MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Department of Polymer Science and Engineering, Zhejiang University, Zheda Road 38, Hangzhou 310027, China

[b] Life Sciences Institute, Zhejiang University, 866 Yuhangtang Rd, Hangzhou 310058, China

ARTICLE INFO

Article history:
Received 11 November 2022
Received in revised form 4 January 2023
Accepted 19 April 2023
Available online 13 May 2023

Keywords:
RNA sequencing
Adenosine analogue
Metabolic labelling
Base mutation
Nascent RNA
RNA dynamics
RNA tag

ABSTRACT

The integration of RNA metabolic labelling by nucleoside analogues with high-throughput RNA sequencing has been harnessed to study RNA dynamics. The immunoprecipitation purification or chemical pulldown technique is generally required to enrich the analogue-labelled RNAs. Here we developed an a^{6}A-seq method, which takes advantage of N^6-allyladenosine (a^{6}A) metabolic labelling on cellular mRNAs and profiles them in an immunoprecipitation-free and mutation-based manner. a^{6}A plays a role as a chemical sequencing tag in that the iodination of a^{6}A in mRNAs results in 1,N^6-cyclized adenosine (cyc-A), which induces base misincorporation during RNA reverse transcription, thus making a^{6}A-labelled mRNAs detectable by sequencing. A nucleic acid melting assay was utilized to investigate why cyc-A prefers to be paired with guanine. a^{6}A-seq was utilized to study cellular gene expression changes under a methionine-free stress condition. Compared with regular RNA-seq, a^{6}A-seq could more sensitively detect the change of mRNA production over a time scale. The experiment of a^{6}A-containing mRNA immunoprecipitation followed by qPCR successfully validated the high-throughput a^{6}A-seq data. Together, our results show a^{6}A-seq is an effective tool to study RNA dynamics.

1. Introduction

High throughput RNA-seq has been widely used to profile global cellular gene expression and understand the cellular biological processes [1], however, it represents a steady state of messenger RNA (mRNA) transcript abundance. To follow dynamic change of cellular mRNA expression, mRNA labelling and sequencing over time is necessary [2]. In this regard, how to label cellular mRNAs and identify them in a simple and high throughput manner remains critical.

The nucleoside analogues [3,4], such as 4-thiouridine (4sU) [5,6], 5-ethynyluridine (5-EU) [7], 5-bromouridine (5-BrU) [8], 6-thioguanosine (6sG) [9], N^6-propargyladenosine (p^{6}A) [10–12] and 2′-azidocytidine (2′-AzC) [13], have been used to label cellular mRNAs and study mRNA dynamics [14]. They are able to be metabolically incorporated into the newly transcribed RNAs by nucleoside salvage pathways [13]. From the structure point of view, these analogues can serve as functional tags because they are either immunoprecipitable or post-modifiable for subsequent chemical pulldowns. For instance, specific antibodies have been used to immunoprecipitate 5-BrU-labelled mRNA from total mRNA [8]. Similarly, 4sU and 6sG in RNA can be modified by a 2-pyridylthio-activated disulfide of biotin (HPDP-biotin) [15,16] or methylthiosulfonate-activated biotin (MTS-biotin) [17], and then be enriched by streptavidin. The 5-EU-, p^{6}A- or 2′-AzC-labelled RNAs can be modified with an azide-fluorophore or azide-biotin probe by bioorthogonal azide-alkyne click reaction and then be quantified by a fluorescence or chemical pulldown assay [18–20].

Recently, RNA nucleoside analogues for chemical sequencing have brought considerable attention due to enrichment-free steps in the process of RNA sequencing library construction. The nucleoside analogues are chemically treated to induce base misincorporation in the RNA reverse transcription (RT), and the mutation signals in RNA-seq are counted to indicate the abundance of labelled RNA transcripts. For example, chemical treatments of 4sU by thiol-linked alkylation [21], oxidative-nucleophilic-aromatic substitution [22], or thiol-ene addition [23] enable U-to-C mutations [24–26] in the RNA-seq. However, these methods were only limited to 4sU analogue, and they either utilized harsh chemical treatments or lacked the *in vitro* validation data to prove the high-throughput sequencing results. Taken together, nucleoside analogues for chemical sequencing of RNA are very rare, and the addition to this toolkit will be highly beneficial to the field. In our previous work, an adenosine analogue, N^6-allyladenosine (a^{6}A) was synthesized [27], and under mild iodination, it could be transformed into 1,N^6-cyclized

* Corresponding author.
E-mail address: liujz@zju.edu.cn (J. Liu).

https://doi.org/10.1016/j.fmre.2023.04.010

adenosine (cyc-A) to induce mismatch during RT [28,29]. Although a^6A was able to metabolically label cellular mRNA [30], an a^6A-based sequencing technique has not been developed yet for studying RNA dynamics.

In this work, we developed an a^6A-seq method, which takes advantage of a^6A metabolic labelling on cellular mRNAs and profiles them in an immunoprecipitation-free and mutation-based manner. First, a nucleic acid melting assay was performed to explain the preferred base pairing of cyc-A with guanine. Then, a^6A-seq was utilized to study the change of the cellular gene expression profile under a methionine-free stress condition. Compared with regular RNA-seq, a^6A-seq could more sensitively detect the level change of newly produced mRNAs. Finally, the experiment of a^6A-containing mRNA immunoprecipitation followed by quantitative real-time PCR (qPCR) was successfully used to validate the high-throughput a^6A-seq data. Together, our results show a^6A-seq is an effective tool to study RNA dynamics.

2. Material and methods

2.1. Materials

All chemicals and reagents were used as purchased without further purification. m^6ATP (N^6-methyladenosine-5′-Triphosphate, catalogue no. N-1013) was purchased from TriLink BioTechnologies. MEGAscript T7 kit (Invitrogen, catalogue no. AM1334), RevertAid enzyme (Thermo, catalogue no. EP0441), 5x RT reaction buffer (Thermo Scientific, catalogue no. EP0441), DMEM/high-glucose medium (HyClone, catalogue no. SH30243.01), FBS–fetal bovine serum (Gibco, catalogue no. 10270), Penicillin–Streptomycin (HyClone, catalogue no. SV30010) and Roswell Park Memorial Institute (RPMI) 1640 medium (no methionine) (Gibco, catalogue no. A1451701) were purchased from Thermo Fisher Scientific. iTaq™ Universal SYBR® Green Supermix (Bio-Rad, catalogue no. 1725124) was purchased from Bio-Rad Laboratories, Inc. Reverse transcriptase recombinant HIV (catalogue no. LS05003) was purchased from Worthington Biochemical Corporation. N^6-isopentenyladenosine antibody (catalogue no. AS09 415) was purchased from Agrisera.

2.2. Measurement of melting temperatures of RNA/cDNA hybrids

91-nt and 89-nt RNA probes (Tables S1, S2) were synthesized by *in vitro* transcription using ATP, m^6ATP, a^6ATP [10] respectively with a MEGAscript T7 kit. The transcribed products were digested by TURBO DNase at 37 °C for 15 min to remove the template DNA and then purified with RNA clean and concentrator. For cyc-A RNA probe, a^6A-RNA probe (10 μg) was dissolved in 52 μL water, and 8 μL 0.125 M I_2 (dissolved in 0.25 M KI solution) was added. The mixture was incubated at 37 °C for 0.5 h, then approximately 8 μL 0.2 M $Na_2S_2O_3$ was added until the mixture became colourless. Followed by adding 12 μL 0.1 M Na_2CO_3 (pH 9.5), the mixture was incubated at 37 °C for another 0.5 h. After purified with RNA clean and concentrator, cyc-A-RNA product was obtained. Next, 3 μg RNA probe was mixed with 4.5 μg corresponding cDNA (Tables S1, S2) in 20 μL RNase-free water. The mixture was heated at 95 °C for 5 min to denature, and cooled to 25 °C with a rate of 0.5 °C/s for hybridization. Afterwards, 20 μL SYBR Green Supermix was added, and the mixtures were divided into 10 μL portions for determination of melting temperatures. The melting curve was measured by quantitative real-time PCR (Bio-Rad CFX96 Connect) with temperature control from 65 °C to 95 °C at a rate of 0.5 °C/5 s.

2.3. RNA metabolic a^6A labelling and its quantification by mass-spectroscopy

HeLa and HEK293T cells were cultured in DMEM/high-glucose medium supplemented with 10% FBS and 1% 100 × pen-strep at 37 °C with 5% CO_2. The medium was changed to methionine-deficient medium supplemented with 10% FBS, 1% 100 × pen-strep and 1 mM L-cysteine to deplete the intracellular methionine. In the meantime, a^6A was added at indicated concentrations. For the control samples, 1 mM methionine was added. After culturing for 24 h, the total RNAs of HeLa cells and HEK293T cells were isolated using Trizol Reagent, and the mRNA was purified from the total RNA by GenElute™ mRNA Miniprep Kit (Sigam, MRN10-1KT). For each sample, around 300–500 ng RNA was digested by using 1 U nuclease P1 (Wako) in 30 μL reaction containing 20 mM NH_4OAc at 42 °C for 2 h. Next, 1 μL Alkaline Phosphatase (1 U, Sigma-Aldrich) and 3 μL 1 M NH_4HCO_3 (Sigma–Aldrich) was added and the reaction was incubated at 37 °C for another 2 h. Samples were then filtered by 0.22 μm filter (Millipore) and diluted to 80 μL. 10 μL of the sample was injected into LC–MS/MS. The levels of m^6A and a^6A in total RNA and mRNA were measured by reverse phase ultra-high performance liquid chromatography on a C18 column coupled with online mass spectrometry detection using Waters TQ MS triple-quadrupole LC spectrometer in positive electrospray ionization mode. The a^6A and m^6A nucleosides were quantified by using the nucleoside to base ion mass transitions of 308 to 176, and 282 to 150, respectively.

2.4. Cell viability assay

Five thousand HeLa or HEK293T cells were seeded per well in the 96-well plate on the day before the experiment. After onset of the experiment, cells were cultured with medium containing the indicated concentration of a^6A. The conditions that a^6A-containing culture medium was exchanged every 6 h or 12 h were also tested. Cell viability experiments were performed by CellTiter-Glo Luminescent Cell Viability Assay (Promega) following the instructions of the manufacturer. Luminescent signals were measured on SynergyNEO2 (BioTek).

2.5. Library construction and high-throughput sequencing analysis

Library preparations were conducted following the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina E7760 manual. Instead of using the RT enzyme mix from the NEB kit, 20 Unites HIV (Reverse transcriptase recombinant HIV) and 5x RT reaction buffer were added for first strand synthesis Reaction. The RT reaction was performed at 25 °C for 10 min, 37 °C for 60 min, and 70 °C for 15 min. Next, standard protocol was followed for second strand cDNA synthesis and other continuing steps to complete the library construction. The library quality was checked using an Agilent 2100 Bioanalyzer and sequenced on an Illumina Hiseq X10 system with paired-end 2 × 150 bp read length. The sequencing reads data was first subject to fastqc (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) for quality control, and adapters were trimmed using fastp (https://github.com/OpenGene/fastp) [31]. The minimum quality threshold was set to 20, and the maximum amplicon end distance was 146 nt. For alignment, reads data was mapped to hg38 as the reference genome using hisat2 (http://daehwankimlab.github.io/hisat2/) [32]. The uniquely aligned sam files (flag: 83/163, 99/147) were sorted and reordered into bam files with samtools (http://www.htslib.org/) [33]. To confirm the mutation pattern in high-throughput sequencing, mutation sites were identified by samtools mpileup and annotated to hg38_refGene. To quantify the total reads, all reads in cyc samples were used for reads counting. To quantify the mutation reads with A-to-C/T, we used sam2tsv (http://lindenb.github.io/jvarkit/Sam2Tsv.html) to search each read with mutation sites for reads counting. In order to eliminate the background signals, all mutation sites that occurred in input samples were ignored in cyc samples. Afterwards, total reads and mutation reads were annotated to hg38_refGene by gene name. The fold changes of different genes ('Met−' samples versus 'Met+' samples) were calculated according to reads count after normalization by sequencing depth. Statistical significance was set to P-value <0.05. Gene ontology (GO) analysis was performed by Metascape (http://metascape.org/gp/index.html#/main/step1).

2.6. a^6A-RIP and qPCR assay

The metabolically a^6A-labelled mRNA (10 μg, isolated from HeLa cells) was fragmented into 100–300 nt length using 10 × fragmentation buffer (1 μL 1 M $ZnCl_2$, 1 μL Tris–HCl pH = 7.0, 8 μL RNase-free water) at 70 °C for 7 min. The fragmented RNAs were subjected to immunoprecipitation as follows: 5 μg RNA fragments were mixed with 80 μL 5 × IP buffer (0.5 mL 1 M Tris–HCl pH = 7.4, 1.5 mL 5 M NaCl, and 0.5 mL 10% vol/vol Igepal CA-630 in total volume of 10 mL), 2 μL RNase inhibitor, and 1 μL of stock antibody (anti-N^6-isopentenyladenosine, Agrisera) in a total volume of 400 μL, and incubated with head-over-tail rotation for 4 h at 4 °C. Meanwhile, 40 μL Protein A beads (Invitrogen, washed with 200 μL 1 × IP buffer 3 times) were mixed with 80 μL 5 × IP buffer, 10 μL 20 mg/μL BSA (MP Biomedicals) in total volume of 400 μL and incubated with head-over-tail rotation for 2 h at 4 °C. After washed with 200 μL 1× IP buffer for 3 times, the Protein A beads were mixed with the RNA fragment-antibody mixture described above and incubated with head-over-tail rotation for 2 h at 4 °C. Finally, the Protein A beads on magnetic stand were washed with 800 μL 1 × IP buffer 2 times and were eluted by 100 μL elution buffer (90 μL 5 × IP buffer, 3 μL RNase inhibitor, and 40 μL 75 mM a^6ATP in total volume of 450 μL) with head-over-tail rotation for 3 h at 4 °C. The IP products were purified by isopropanol precipitation, dissolved in RNase-free water, and stored at −80 °C. For qPCR assay, GAPDH was chosen as control and primer sequences are shown in Table S3. Through RT with RevertAid enzyme and qPCR with SYBR Green Supermix, the enrichment folds were calculated using the $2^{-\Delta\Delta CT}$ method, where CT is the cycle number to reach the detection threshold. For example, in order to calculate enrichment folds of *MAT2A*, $CT_{IP-MAT2A}$ and $CT_{Input-MAT2A}$ represented the cycle number of *MAT2A* to reach detection threshold in IP and Input samples, respectively. When $\Delta CT_1 = CT_{IP-MAT2A} - CT_{IP\text{-}GAPDH}$, $\Delta CT_2 = CT_{Input-MAT2A} - CT_{Input\text{-}GAPDH}$, and $\Delta\Delta CT(MAT2A) = \Delta CT_1 - \Delta CT_2$, the enrichment fold of *MAT2A* is equal to $2^{-\Delta\Delta CT(MAT2A)}$. The same strategy was used to calculate enrichment of other selected genes.

3. Results and discussion

3.1. Preferred cyc-A:G base pairing in RNA/DNA hybrid

In our previous study, the a^6A-derived cyc-A in RNA probe has been found to mismatch with guanine (G) and adenine (A) in a descending order during reverse transcription, with cyc-A:G pairing dominant (Fig. 1a) [28]. From the structure point of view, the hydrogen bonding sites for canonical Watson–Crick base pairing in cyc-A are impaired [34]. We proposed the potential mode of hydrogen bond binding of cyc-A with G as shown in Fig. 1b, however, there is no further evidence to prove the preferred cyc-A:G base pairing. We first designed a nucleic acid melting assay to validate the cyc-A pairing mode. Three 91-nt RNA probes containing 13 unmodified versus modified adenosines at the same sequence locations, named A-RNA, m^6A-RNA, and a^6A-RNA (Table S1), were synthesized by *in vitro* transcription (IVT), and then hybridized with their complementary DNA (cDNA) for measurement of melting temperature (T_m). Since T_m indicates the capacity of base pairing, impaired base pairing will generally lead to decreased T_m. Obviously, a decreasing order of T_m was found for hybrids of A-RNA/cDNA (84.1 °C), m^6A-RNA/cDNA (81.5 °C), and a^6A-RNA/cDNA (77.1 °C) (Fig. S1), suggesting that chemical modifications significantly weaken the canonical A:T (thymine) base pairing. We further converted the a^6A-RNA into cyc-A-RNA probe by iodination and tested its cDNA hybrid T_m, but no reliable T_m was detected, indicating the occurrence of even worse base pairing between cyc-A and T.

Given the above results, we next prepared an 89-nt RNA probe (Table S2) with only 3 adenosines in order to investigate the cyc-A base pairing behaviour with an expectation to observe more efficient cDNA hybridization and reliable T_m. Besides, to investigate the base pairing of modified A with different bases, we designed another three kinds of cDNAs by replacing the corresponding T in the cDNA sequence with A, C and G, named A-, C-, and G-cDNA, respectively (Table S2). The T_m of A-RNA, m^6A-RNA, a^6A-RNA hybridized with T-cDNA were 83.3 °C, 82.9 °C and 82.0 °C (Fig. 1c), well consistent with the result of the above-mentioned 91-nt RNA probes. Furthermore, the hybrids of A, m^6A and a^6A-RNAs with A-, C-, and G-cDNAs exhibited significantly lower T_m than their T-cDNAs due to the weakened hydrogen bonding interactions (Figs. 1c and S2). Both 91-nt and 89-nt RNA probes indicated that the T_m of RNA decreased approximately 0.2 °C and 0.5 °C per m^6A and a^6A site respectively. By looking into the a^6A-derived cyc-A-RNA probe, we observed a decreasing order of T_m's in hybrids with G-, A- and C-cDNAs, supporting that cyc-A mainly pairs with G in the RT-induced misincorporation process (Figs. 1c and S2). These data provide excellent evidence to explain the major mutation pattern of cyc-A to C/T in the cDNA sequencing data.

3.2. Development of high-throughput a^6A-seq in mammalian cells

We next moved to develop high-throughput a^6A-seq based on RNA a^6A metabolic labelling followed by RNA sequencing. First, HEK293T cells were labelled with 1 mM a^6A for different time within a period of 4 h, and the a^6A labelling level in total RNA were characterized by UHPLC-QQQ-MS/MS (ultra-high-performance liquid chromatography coupled with triple quadrupole mass spectrometry) (Fig. S3). The a^6A/A ratio significantly increased with labelling time, suggesting the occurrence of RNA metabolic labelling event. In order to screen the optimal conditions for a^6A labelling, the effect of a^6A concentration on cellular cytotoxicity was studied in both HEK293T and HeLa cells. We treated cells with different a^6A concentrations for 24 h of incubation and then tested the cell viability. The results showed that the cytotoxicity gradually increased with a^6A concentration (Fig. S4), and the suitable concentration range was 100 μM–1 mM, which gave the acceptable cytotoxicity (>70% viability). It has been reported that the exchange of medium with labelling molecules every time interval would improve the labelling ratio [21]. However, our data showed that the level of mRNA a^6A only slightly increased by exchanging a^6A-containing medium every 6 h or 12 h (Fig. S5a), and that the toxicity was increased. It was clearly seen that the a^6A labelling levels on mRNAs increased significantly with a^6A concentrations (Fig. S5b). To achieve both low cytotoxicity and high labelling yield, we finally chose the condition of 1 mM a^6A and 24 h of incubation time to perform a^6A-seq in HeLa cells. Over the time period of 24 h, the a^6A/A level increased to around 2.6‰, approaching the native cellular m^6A level (Fig. 2a). Meanwhile, the a^6A/A level in mRNA increased much faster than that in total RNA, suggesting that a^6A preferred to be incorporated into mRNA.

In a^6A-seq, the metabolically labelled mRNA was isolated, fragmented, iodinated, and subjected to standard RNA-seq library construction. The samples with and without iodination treatment were named cyc and input, respectively. In order to acquire the mutation pattern of cyc-A, we used two reverse transcriptases in the RT step to construct sequencing libraries and compared their information on mutation sites at transcriptome-wide level. The one is HIV RT enzyme which has been reported to induce mutation of cyc-A in RNA, and the other is the commercial RT enzyme from NEBNext Ultra II Directional RNA Library Prep Kit for Illumina. After alignment of high-throughput a^6A-seq sequencing reads, mutation sites and certain base mutation ratios could be calculated. Compared with HIV-input samples, HIV-cyc samples showed significant mutations of A-to-C and A-to-T, but they showed no significance for other types of base mutations (Figs. 2b and S6) [35]. The NEB RT enzyme did not demonstrate significant differences amongst mutation sites of A to other bases between input and cyc samples (Figs. 2b and S6), it was possibly because that NEB RT enzyme could not read through the cyc-A base. For HIV enzyme, the mutation ratios of A-to-C and A-to-T reached 0.75‰ and 0.45‰ (A-to-C/A-to-T, 5/3), respectively, under

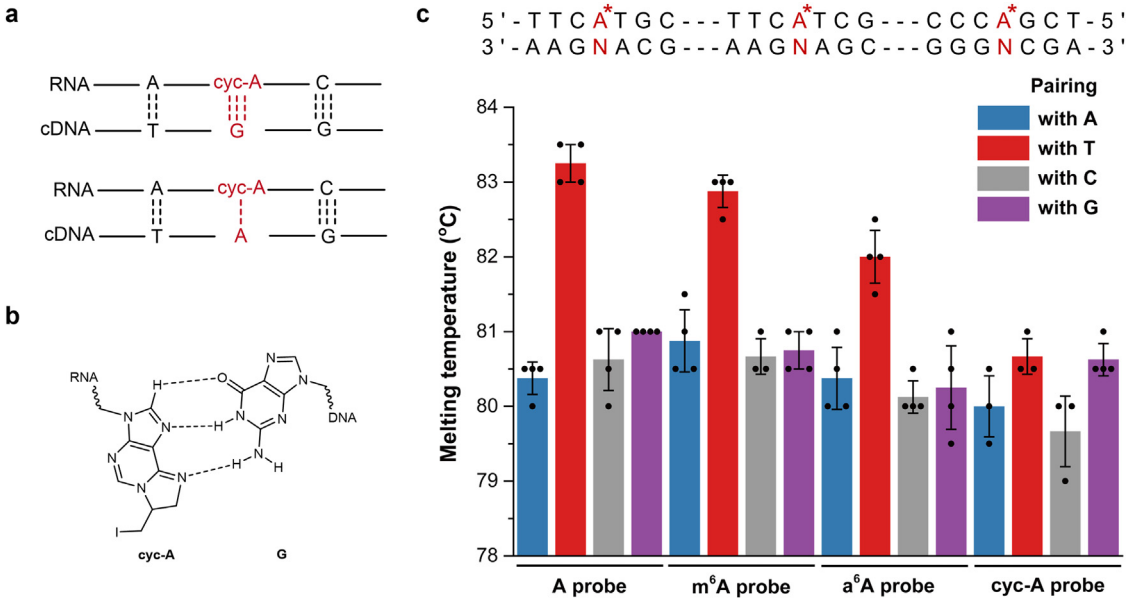


Fig. 1. Study of base pairing behaviour of different adenosine analogue-containing RNA probes with their complementary DNA (cDNA) by a melting assay. (a) Schematic illustrations of RNA/cDNA hybrids having base pairing of cyc-A with G or A, (b) a predicted structure showing hydrogen bonding of cyc-A with G, (c) melting temperatures of various RNA/cDNA hybrids. The RNA probes differ in adenosine modification designated as A, m^6A, a^6A and cyc-A. The melting temperatures for each modification paired with A, T, C or G were tested. Mean values ± s.e.m. are shown; $n = 4$ biologically independent samples.

the mRNA a^6A labelling level of 2.6‰ (Fig. 2a, b). Together, the HIV RT system was suitable for a^6A-seq.

Next, we annotated all sequencing reads and counted the numbers of A-to-C/T mutation sites in each gene transcriptome-wide. The high correlation between A-to-C and A-to-T mutations in cyc samples (Pearson's $R = 0.99$, Figs S7, S8) made it feasible to identify the cyc-A (viz. a^6A) sites from total transcripts. By removing the background signals from input samples, we found that the obtained effective mutation sites from different samples were highly correlated (Pearson's $R \geq 0.98$). The ratio of A-to-C to A-to-T sites was around 1.6 (Fig. 2c, d), which was almost the same as mutation ratio (5/3) calculated from total mRNA transcripts (Fig. 2b). Therefore, it is reliable to apply A-to-C/T mutation to distinguish the newly produced a^6A-lablled mRNAs from total transcripts (Fig. 2e).

3.3. Application of a^6A-seq to study the change of gene expression profile under stressed condition

We next explored the feasibility of a^6A-seq to profile the gene expression in various cellular contexts. It was anticipated that a^6A-seq could not only provide the profile of global gene expression derived from total mRNA reads, but also generate the information of metabolically produced a^6A-labelled mRNAs from mutation reads. We chose a methionine-free condition as outlined in Fig. 3 to test how stress affects mRNA production over a time period. Both the total reads and mutation reads from a^6A-seq were analysed in order to see dynamic fine change in gene expression profile. Meanwhile, the IP of a^6A-containing mRNA and qPCR assay (Fig. 3) were used to validate the targets identified from mutation reads in a^6A-seq.

Methionine is the precursor of *S*-adenosyl methionine (SAM), which is the methyl donor of various cellular methyltransferases [36–39], therefore we thought the deficiency of methionine would affect the cellular methylation processes. The N^6-methyladenosione (m^6A) modification is a star epitranscriptomic mark and plays significant biological roles in regulating gene expressions [40–43]. In this regard, decreased supply of methionine will lead to the reduction of the m^6A formation. Indeed, after depletion of methionine in culturing medium of HeLa cells for 16 h, a significant decrease of m^6A level by 37.5% in mRNA was observed (Fig. 4a, Fig. S9), while the m^6A change in total RNA was not significant possibly due to the large contribution of m^6A modification in ribosomal RNA with a longer turnover time than mRNA.

The mutation analysis from a^6A-seq data showed that the mutation ratios of A-to-C/T in methionine-deficient samples (Met−) and methionine-sufficient samples (Met+) were almost identical (Fig. 4b). After annotation of high-throughput sequencing reads, a standard RNA-seq protocol was used to calculate the reads of mRNA transcripts. Moreover, to identify the a^6A-labelled transcripts, sequencing reads with mutation sites of A-to-C/T were selected by removing background signals from input samples (Fig. S10). By normalization according to sequencing depth, the fold changes of expressed genes in Met− versus Met+ were calculated by total reads and mutation reads, respectively (Fig. 4c, d). It is clearly seen that the mutation reads can demonstrate the fine change of certain mRNA productions over the labelling period, which cannot be identified from the total reads analysis. The information from the mutation reads indicates the difference in mRNA transcription production over this specific stress period, while the total reads represent the addition of both existing and newly generated transcripts. In the analysis of total reads, 2476 and 131 transcripts exhibited increased and decreased expressions, respectively (Fig. 4c). However, 425 up-regulated and 377 down-regulated genes were identified from mutation reads (Fig. 4d).

We next selected targets and performed an independent assay to validate the a^6A-seq result. It was found that *MAT2A* showed up-regulated expression from the analysis of both total reads and mutation reads (Fig. 5a), and was thus selected as the common target [44]. Special attention was paid to the genes showing an opposite trend of fold changes in between total and mutation reads analysis, and four genes including *NR_027295, COMMD1, KLHDC7B* and *ZFAND2B* were chosen (Fig. 5a). The combination of IP of specific a^6A-conatining mRNAs [30] and qPCR is an ideal tool to validate the targets from the a^6A-seq. The qPCR data of the a^6A IP product will reflect the amount of a^6A-labelled mRNA transcripts. The Met− and Met+ mRNA samples were subjected to fragmentation and a^6A-spcific IP, then the products and their input samples were characterized by qPCR assay. *GAPDH* was chosen as a control for the qPCR assay due to its stable expression. The fold changes of *MAT2A, NR_027295, COMMD1, KLHDC7B* and *ZFAND2B* between input and IP

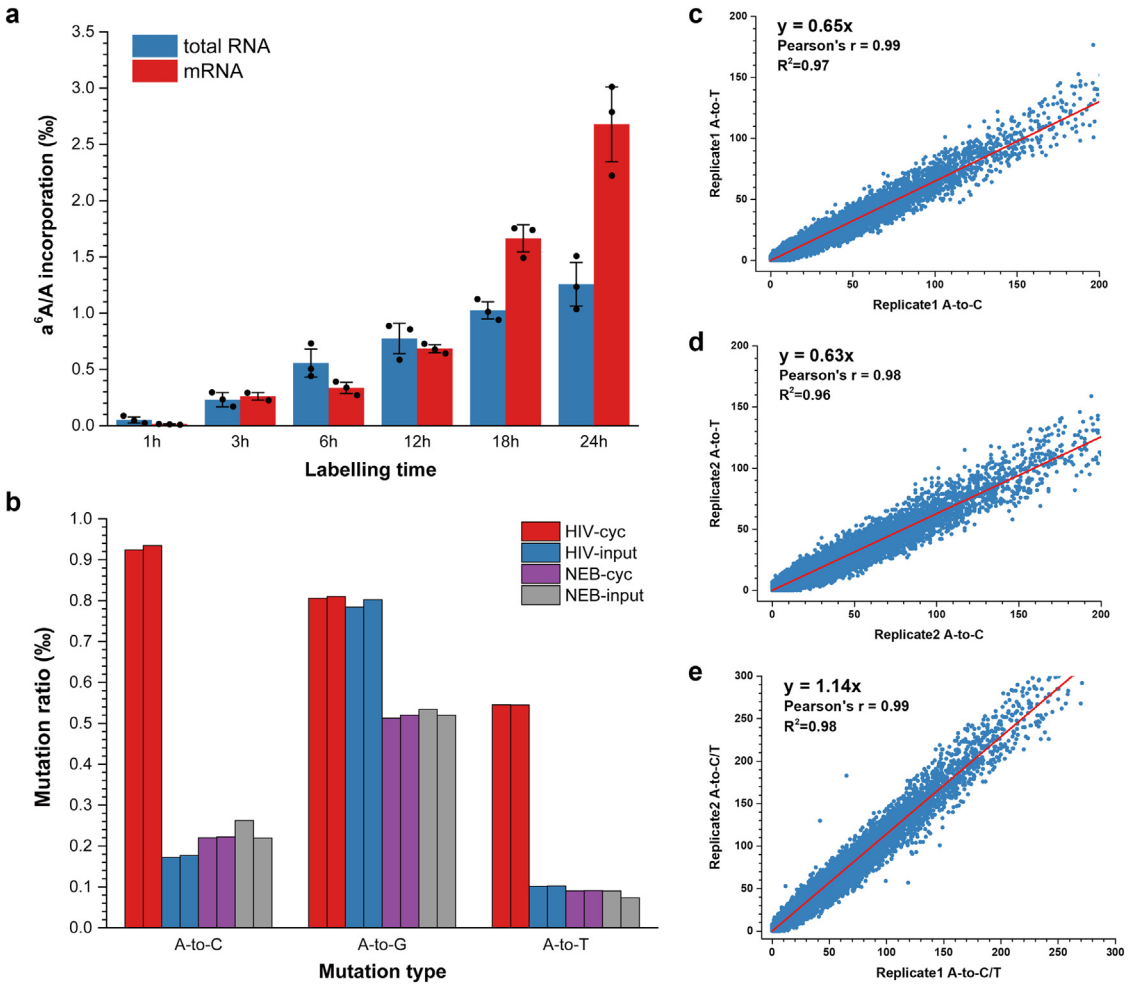


Fig. 2. The metabolic incorporation ratio of a^{6}A/A in RNA and analysis of a^{6}A-seq in HeLa cells. (a) The a^{6}A incorporation ratio in total RNA and mRNA of HeLa cells. Mean values ± s.e.m. are shown; n = 3 biologically independent samples. (b) The calculated A-to-C/G/T mutation ratios from mRNA a^{6}A-seq. In the mRNA RT step during library construction, HIV enzyme and RT enzyme from NEB RNA Library Prep Kit were used, respectively. HIV-/NEB-input: a^{6}A-labelled mRNA samples and RT by HIV/NEB; HIV-/NEB-cyc: a^{6}A-labelled samples after chemical cyclization treatment and RT by HIV/NEB. n = 2 biologically independent samples. The mutation rates of other bases are provided in Fig. S10. (c–e) The correlation plots of transcriptome-wide mRNA mutation sites from a^{6}A-seq. The A-to-C/T means that the mutation sites of A-to-C or A-to-T were counted.

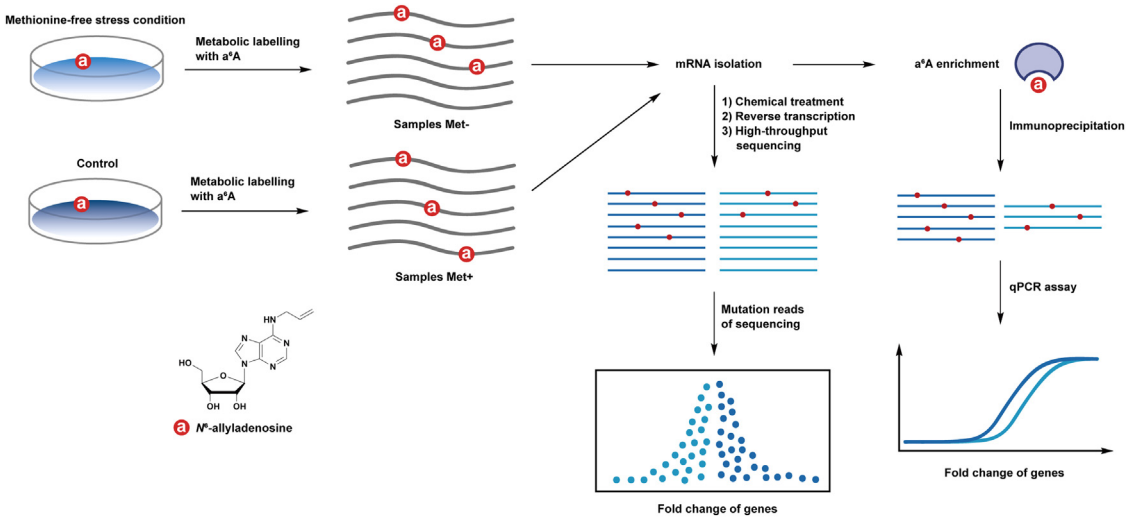


Fig. 3. A schematic illustration to study the changes of gene expression profiles under methionine-free condition by a^{6}A-seq. The mutation reads can be used to identify the produced a^{6}A-containing transcripts over a time scale. The combination of a^{6}A-specific antibody immunoprecipitation and qPCR assay can be applied to validate the high-throughput a^{6}A-seq data.

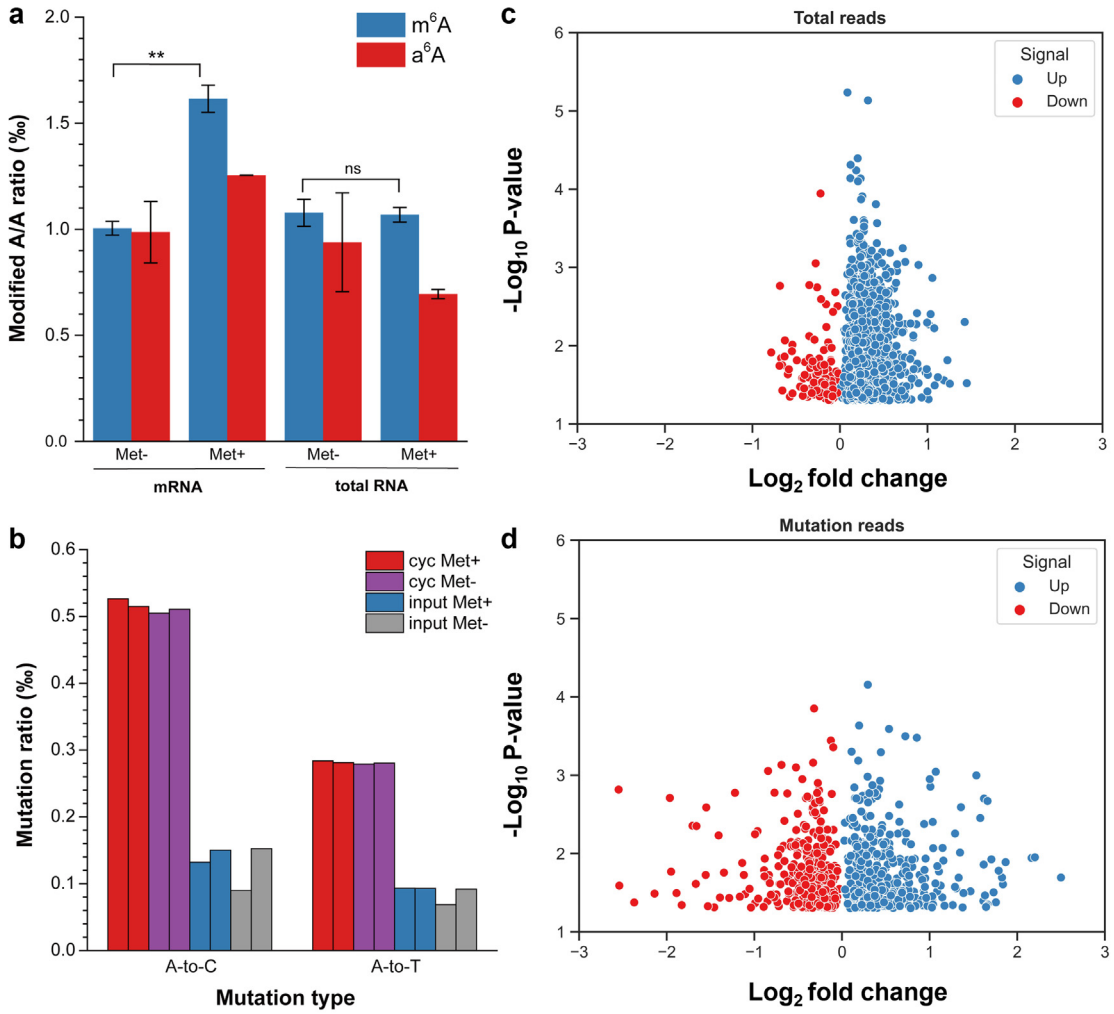


Fig. 4. The application of a^{6}A-seq to study cellular gene expression change under the methionine-deficient (Met−) versus methionine-sufficient (Met+) conditions. (a) The measured a^{6}A and m^{6}A levels of mRNA and total RNA from HeLa cells labelled with a^{6}A for 16 h under conditions of Met− versus Met+. $^{**}P < 0.01$. ns, non-significant by *t*-test (two-tailed). (b) The mutation ratios of A-to-C and A-to-T in a^{6}A-seq data of different samples. Input: a^{6}A-labelled samples; cyc: a^{6}A-labelled samples after chemical cyclization treatment. $n = 2$ biologically independent samples. (c, d) The up- and down-regulated mRNAs identified from total reads and mutation reads of cyc samples in the form of $\log_2$ fold change (Met−)/(Met+), respectively. $P < 0.05$.

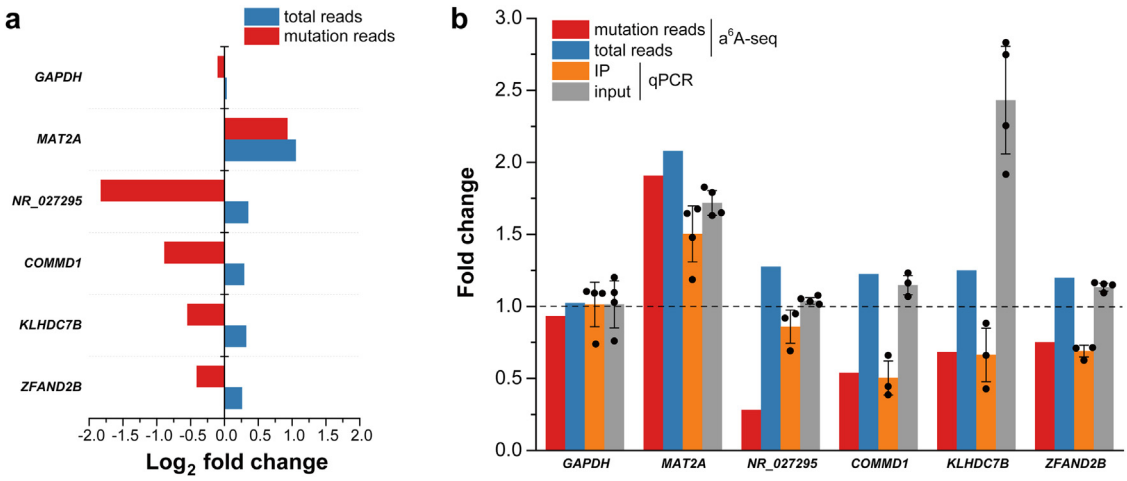


Fig. 5. Validation of a^{6}A-seq result by antibody immunoprecipitation (IP) and qPCR assay. (a) Selected examples of up- and down-regulated genes identified from total reads and mutation reads in HeLa cells under conditions of Met− versus Met+. $P < 0.05$. (b) The qPCR results of selected genes are shown in panel A in both input mRNA and a^{6}A-specific antibody IP mRNA samples. The associated a^{6}A-seq data is also provided for comparison. $n = 4$ biologically independent samples.

samples were highly consistent with those from high-throughput a^{6}A-seq (Fig. 5b). In order to confirm the gene expression changes calculated from a^{6}A-seq, we compared the data in between a^{6}A-seq and normal RNA-seq in the absence of a^{6}A labelling [30]. The overlap of up- and down-regulated genes between a^{6}A-seq and normal RNA-seq was more than 60% (Fig. S11). It should be noticed that *MAT2A, NR_027295, KLHDC7B* and *ZFAND2B* were included in the overlapped transcripts. The normal RNA-seq showed the same trend of gene expressions with that calculated from total reads by a^{6}A-seq.

Finally, gene ontology (GO) analysis was performed to understand the potentially influenced biological processes (BP) by differentially expressed genes in methionine-deficient samples versus control (Figs S12, S13). The BP enrichment analysis from both mutation reads and total reads showed some identical pathways, such as mitochondrion organization and RNA splicing. However, there were a quite number of differently enriched pathways, especially in the group of down-regulated genes (Fig. S13), indicating that mutation reads in a^{6}A-seq could capture the decrease of mRNA production more efficiently.

4. Conclusion

In summary, we report an a^{6}A-seq technique which is built on metabolic labelling of cellular mRNAs with nucleoside a^{6}A and enables the detection of labelled mRNAs in an immunoprecipitation-free and mutation-based manner. a^{6}A-seq can not only measure global mRNA expression change but also quantify the RNA production over a certain time period under stimulating conditions, and thus functions as an enhanced version of regular RNA-seq. Obviously, a^{6}A plays a role of chemical sequencing tag, and its A-to-C/T mutation pattern in a^{6}A-seq was solidified by a melting assay on cyc-A-RNA/cDNA hybrids. It should be noticed that there are many methods to detect RNA modification sites based on mutation during reverse transcription. To help exclude false positive mutation signals from background, our T_m-based qPCR assay can be used to validate the mismatches of modified nucleosides on RNA. The T_m of RNA/DNA hybrids provides a standard of mutation species for screening reverse transcriptases which can induce misincorporation on cDNA. In our previous work about 1,N^6-cyclized adenosine, most RT enzymes could not introduce mutation sites, but the T_m provides evidence of mismatches, which led us to find a suitable reverse transcriptase. Therefore, T_m-based qPCR assay is a fast way to identify mutation-capable modified nucleosides. Finally, combining the advantages of low-throughput and high-throughput sequencing, we provided an effective and orthogonal a^{6}A-RNA IP followed by qPCR assay, which was developed to validate the a^{6}A-seq results. We believe that the application of multiple orthogonal methods will greatly improve the reliability of sequencing results, especially for RNA-seq.

Inspired by the success of a^{6}A-seq, it is highly rewarding to develop new RNA chemical sequencing tags for biorthogonal labelling on different kinds of cells or temporal labelling on the same cell type. Until now, only a^{6}A and 4sU have been reported for applications of RNA labelling and chemical sequencing. There is a lot more room to explore cytosine- and guanosine-based chemical sequencing tags with the hope of achieving specific labelling on a certain kind of RNA. Because each type of analogue has its own mutation pattern, cellular or intercellular temporal events can be identified in the order. The work along this line will contribute to the elucidation of sophisticated RNA dynamics.

Declaration of competing interest

The authors declare that they have no conflicts of interest in this work.

Acknowledgments

We thank the support from the National Key R&D Program of China (2022YFA1103702 and 2017YFA0506800), the National Natural Science Foundation of China (22022702 and 21977087), the Fundamental Research Funds for the Central Universities, and MOE Key Laboratory of Macromolecular Synthesis and Functionalization, Zhejiang University (2022MSF04).

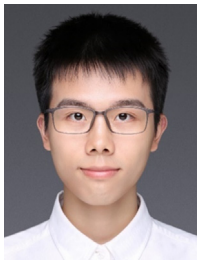
Supplementary materials

The high-throughput sequencing data reported in this paper have been deposited in the Gene Expression Omnibus database at https://www.ncbi.nlm.nih.gov/geo/ (accession no. GSE192818). Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.fmre.2023.04.010.

References

[1] B. Schwanhausser, D. Busse, N. Li, et al., Global quantification of mammalian gene expression control, Nature 473 (7347) (2011) 337–342.
[2] R. Stark, M. Grzelak, J. Hadfield, RNA sequencing: the teenage years, Nat. Rev. Genet. 20 (11) (2019) 631–656.
[3] H. Tani, N. Akimitsu, Genome-wide technology for determining RNA stability in mammalian cells historical perspective and recent advantages based on modified nucleotide labeling, RNA Biol. 9 (10) (2012) 1233–1238.
[4] J.M. Holstein, A. Rentmeister, Current covalent modification methods for detecting RNA in fixed and living cells, Methods 98 (2016) 18–25.
[5] W.T. Melvin, H.B. Milne, A.A. Slater, et al., Incorporation of 6-thioguanosine and 4-thiouridine into RNA - application to isolation of newly synthesized RNA by affinity chromatography, Eur. J. Biochem. 92 (2) (1978) 373–379.
[6] M. Rabani, J.Z. Levin, L. Fan, et al., Metabolic labeling of RNA uncovers principles of RNA production and degradation dynamics in mammalian cells, Nat. Biotechnol. 29 (5) (2011) 436–U237.
[7] C.Y. Jao, A. Salic, Exploring RNA transcription and turnover in vivo by using click chemistry, Proc. Natl. Acad. Sci. U.S.A. 105 (41) (2008) 15779–15784.
[8] H. Tani, R. Mizutani, K.A. Salam, et al., Genome-wide determination of RNA stability reveals hundreds of short-lived noncoding transcripts in mammals, Genome Res. 22 (5) (2012) 947–956.
[9] L. Kiefer, J.A. Schofield, M.D. Simon, Expanding the nucleoside recoding toolkit: revealing RNA population dynamics with 6-thioguanosine, J. Am. Chem. Soc. 140 (44) (2018) 14567–14570.
[10] M. Grammel, P. Luong, K. Orth, et al., A chemical reporter for protein AMPylation, J. Am. Chem. Soc. 133 (43) (2011) 17103–17105.
[11] M. Grammel, H. Hang, N.K. Conrad, Chemical reporters for monitoring RNA synthesis and poly(A) tail dynamics, ChemBioChem 13 (8) (2012) 1112–1115.
[12] X.Y. Gao, X. Shu, Y.N. Song, et al., Visualization and quantification of cellular RNA production and degradation using a combined fluorescence and mass spectrometry characterization assay, Chem. Commun. 55 (57) (2019) 8321–8324.
[13] D.Y. Wang, Y. Zhang, R.E. Kleiner, Cell- and polymerase-selective metabolic labeling of cellular RNA with 2 '-azidocytidine, J. Am. Chem. Soc. 142 (34) (2020) 14417–14421.
[14] R.E. Kleiner, Interrogating the transcriptome with metabolically incorporated ribonucleosides, Mol. Omics 17 (6) (2021) 833–841.
[15] M.D. Cleary, C.D. Meiering, E. Jan, et al., Biosynthetic labeling of RNA with uracil phosphoribosyltransferase allows cell-specific microarray analysis of mRNA synthesis and decay, Nat. Biotechnol. 23 (2) (2005) 232–237.
[16] M.R. Miller, K.J. Robinson, M.D. Cleary, et al., TU-tagging: cell type-specific RNA isolation from intact complex tissues, Nat. Methods 6 (6) (2009) 439–U457.
[17] E.E. Duffy, M. Rutenberg-Schoenberg, C.D. Stark, et al., Tracking distinct RNA populations using efficient and reversible covalent chemistry, Mol. Cell 59 (5) (2015) 858–866.
[18] S. Nainar, S. Beasley, M. Fazio, et al., Metabolic incorporation of azide functionality into cellular RNA, ChemBioChem 17 (22) (2016) 2149–2152.
[19] Y. Zhang, R.E. Kleiner, A metabolic engineering approach to incorporate modified pyrimidine nucleosides into cellular RNA, J. Am. Chem. Soc. 141 (8) (2019) 3347–3351.
[20] N. Klocker, F.P. Weissenboeck, A. Rentmeister, Covalent labeling of nucleic acids, Chem. Soc. Rev. 49 (23) (2020) 8749–8773.
[21] V.A. Herzog, B. Reichholf, T. Neumann, et al., Thiol-linked alkylation of RNA to assess expression dynamics, Nat. Methods 14 (12) (2017) 1198–1204.
[22] J.A. Schofield, E.E. Duffy, L. Kiefer, et al., TimeLapse-seq: adding a temporal dimension to RNA sequencing through nucleoside recoding, Nat. Methods 15 (3) (2018) 221–225.
[23] Y.Q. Chen, F. Wu, Z.G. Chen, et al., Acrylonitrile-mediated nascent RNA sequencing for transcriptome-wide profiling of cellular RNA dynamics, Adv. Sci. 7 (8) (2020) 1900997.

[24] C. Riml, T. Amort, D. Rieder, et al., Osmium-mediated transformation of 4-thiouridine to cytidine as key to study RNA dynamics by sequencing, Angew. Chem. Int. Edit. 56 (43) (2017) 13479–13483.
[25] C. Gasser, I. Delazer, E. Neuner, et al., Thioguanosine conversion enables mRNA-lifetime evaluation by RNA sequencing using double metabolic labeling (TUC-seq DUAL), Angew. Chem. Int. Edit. 59 (17) (2020) 6881–6886.
[26] F. Erhard, A.E. Saliba, A. Lusser, et al., Time-resolved single-cell RNA-seq using metabolic RNA labelling, Nat. Rev. Methods Primers 2 (2022) 77.
[27] R. Ottria, S. Casati, E. Baldoli, et al., N^6-alkyladenosines: synthesis and evaluation of in vitro anticancer activity, Biorg. Med. Chem. 18 (23) (2010) 8396–8402.
[28] X. Shu, Q. Dai, T. Wu, et al., N^6-allyladenosine: a new small molecule for RNA labeling identified by mutation assay, J. Am. Chem. Soc. 139 (48) (2017) 17213–17216.
[29] D. Bartee, S.T. Gamage, C.N. Link, et al., Arrow pushing in RNA modification sequencing, Chem. Soc. Rev. 50 (17) (2021) 9482–9502.
[30] X. Shu, J. Cao, M.H. Cheng, et al., A metabolic labeling method detects m^{6}A transcriptome-wide at single base resolution, Nat. Chem. Biol. 16 (8) (2020) 887–895.
[31] S.F. Chen, Y.Q. Zhou, Y.R. Chen, et al., Fastp: an ultra-fast all-in-one FASTQ preprocessor, Bioinformatics 34 (17) (2018) 884–890.
[32] D. Kim, J.M. Paggi, C. Park, et al., Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype, Nat. Biotechnol. 37 (8) (2019) 907–915.
[33] H. Li, B. Handsaker, A. Wysoker, et al., The sequence alignment/map format and SAMtools, Bioinformatics 25 (16) (2009) 2078–2079.
[34] A. Calabretta, C.J. Leumann, Base pairing and miscoding properties of 1,N^6-ethenoadenine- and 3,N^4-ethenocytosine-containing RNA oligonucleotides, Biochemistry 52 (11) (2013) 1990–1997.
[35] H.Q. Zhou, S. Rauch, Q. Dai, et al., Evolution of a reverse transcriptase to map N^1-methyladenosine in human messenger RNA, Nat. Methods 16 (12) (2019) 1281–1288.
[36] C. Dalhoff, G. Lukinavicius, S. Klimasauskas, et al., Direct transfer of extended groups from synthetic cofactors by DNA methyltransferases, Nat. Chem. Biol. 2 (1) (2006) 31–32.
[37] R. Wang, K. Islam, Y. Liu, et al., Profiling genome-wide chromatin methylation with engineered posttranslation apparatus within living cells, J. Am. Chem. Soc. 135 (3) (2013) 1048–1056.
[38] K. Hartstock, B.S. Nilges, A. Ovcharenko, et al., Enzymatic or in vivo installation of propargyl groups in combination with click chemistry for the enrichment and detection of methyltransferase target sites in RNA, Angew. Chem. Int. Edit. 57 (21) (2018) 6342–6346.
[39] S. Mikutis, M.X. Gu, E. Sendinc, et al., meCLICK-Seq, a substrate-hijacking and RNA degradation strategy for the study of RNA methylation, ACS Cent. Sci. 6 (12) (2020) 2196–2208.
[40] Y. Fu, D. Dominissini, G. Rechavi, et al., Gene expression regulation mediated through reversible m^{6}A RNA methylation, Nat. Rev. Genet. 15 (5) (2014) 293–306.
[41] I.A. Roundtree, M.E. Evans, T. Pan, et al., Dynamic RNA modifications in gene expression regulation, Cell 169 (7) (2017) 1187–1200.
[42] Y. Yang, P.J. Hsu, Y.S. Chen, et al., Dynamic transcriptomic m^{6}A decoration: writers, erasers, readers and functions in RNA metabolism, Cell Res. 28 (6) (2018) 616–624.
[43] H.L. Shi, J.B. Wei, C. He, Where, when, and how: context-dependent functions of RNA methylation writers, readers, and erasers, Mol. Cell 74 (4) (2019) 640–650.
[44] C.L. Quinlan, S.E. Kaiser, B. Bolanos, et al., Targeting S-adenosylmethionine biosynthesis with a novel allosteric inhibitor of Mat2A, Nat. Chem. Biol. 13 (7) (2017) 785–792.



Xiao Shu received his Ph.D. degree from Zhejiang University under the mentorship of Professor Jianzhao Liu in 2022. Now, he is currently working as a postdoctoral research scientist at University of Oxford. His research interests focus on developing new technologies to detect RNA and DNA modifications. Dr. Shu has developed a new and direct detection method for cellular transcriptome-wide RNA m^{6}A at single base resolution, which was published in *Nature Chemical Biology* and *Journal of the American Chemical Society* and highlighted by *Nature Review Genetics*.

Jianzhao Liu is a tenured full professor at Zhejiang University. His current research focuses on nucleic acid chemistry and biology. He got bachelor's degree from Zhejiang University and PhD from The Hong Kong University of Science and Technology under mentorship of Professor Ben Zhong Tang. After postdoc training at The University of Chicago with Professor Chuan He, he became tenure-tracked principal investigator at Zhejiang University in 2015 and got tenured in 2020. He has published over 100 papers and received NSFC fund for excellent young scientists. He won the 2021 and 2022 Highly Cited Researcher in cross-field ranked by Clarivate.