A novel platform of RNA 2′-O-methylation high-throughput and site-specific quantification tools revealed its broad distribution

Yao Tang[1], Yifan Wu[1], Ruilin Xu[1], Xiangwen Gu[1], Yan Wu[1], Jian-Qun Chen[1]*, Qiang Wang[1]*, Qihan Chen[1, 2]*

1 The State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University, Nanjing, Jiangsu 210023, China.

2 Medical School of Nanjing University, Nanjing, Jiangsu 210093, China.

To whom correspondence should be addressed. E-mail: chenqihan@nju.edu.cn, wangq@nju.edu.cn, chenjq@nju.edu.cn.

Abstract

Ribose 2′-O-methylation is involved in critical biological processes, but its biological functions and significance remain largely unknown due to the lack of accurate and efficient identification tools. To overcome this gap, we established an accurate high-throughput single-base Nm site screening method, NJU-seq (Nm-Judged-Universally sequencing), and a site-specific 2′-O-methylation quantification tool, Nm-VAQ (RNase H-based Nm Validation and Absolute Quantification). In this study, we identified several new Nm sites on HeLa rRNA and thousands of new Nm sites on mRNA of human and mouse cell lines, and 13 of them were selected and further validated. The 2′-O-methylation ratios of all validated Nm sites on mRNA were lower than 20%, which was out of the detection abilities of previous methods. Our study suggested that the function and regulation of Nm sites are more complicated than previous studies revealed and worth further mechanistic investigation.

25 **Introduction**

26 More than 100 RNA modifications have been identified, constituting the epi-
27 transcriptome code and affecting RNA functions[1]. Among these RNA modifications,
28 2′-O-methylation (otherwise known as Nm), whereby ribose is methylated at the 2′
29 hydroxyl position without limitation of the base identity, is a widespread RNA
30 modification phenomenon, including multiple RNA types and species[1-5]. Furthermore,
31 several sporadic published studies have demonstrated the importance of 2′-O-
32 methylation in multiple biological processes as an essential signaling regulator. For
33 example, Nm stabilizes the RNA structure by maintaining the ribose C3′-endo
34 conformation, resisting nucleophilic attack, and avoiding 3′-5′ degradation triggered by
35 uridylation[6-8]. In addition, a single 2′-O-methylation modification as a molecular tag
36 distinguishes between internal and external RNAs, generates abnormal alternative
37 splicing, and inhibits translation[5, 9-15]. However, despite its potential wide-ranging
38 cellular activities, an in-depth exploration of 2′-O-methylation biological functions is
39 still limited due to the lack of screening tools to detect Nm sites.

40 Three high-throughput Nm detection methods have been developed in the past few
41 years: 2′OMe-seq[16], RiboMethSeq[7], and Nm-Seq[17]. These methods can reliably detect

42 rRNA Nm modification for an abundant amount of rRNA and a high modification ratio.

43 In addition, LC-MS and RTL-P were also used to validate the rRNA site-specific 2′-O-

44 methylation status[18, 19]. However, the small amount, complex sequence context, and

45 low methylation ratio of mRNAs lead to barriers to the practical application of those

46 methods in mRNAs. To address this problem, we established an accurate high-

47 throughput single-base Nm site screening method NJU-seq and a site-specific 2′-O-

48 methylation quantification tool Nm-VAQ. With the help of these two methods, we

49 efficiently revealed the broad distribution of Nm sites on mRNAs and observed that the

50 2′-O-methylation ratios of all validated Nm sites on mRNA were lower than 20%.

51

52 **Results**

53 **Principle and Application of NJU-seq**

54 The NJU-seq method relied on helicase and hydrolase activities of RNase R from
55 *Mycoplasma genitalium* (MgR) on ssRNA from the 3′ to 5′ direction, which produced
56 the digested products end at every Nm+1 site[13]. A synthesized 22 nt ssRNA substrate
57 carrying one mixed 2′-O-methylation site NNNmNN (N=A/G/C/U) was incubated with
58 MgR to evaluate its sensitivity to Nm sites (Fig. 1a, Supplementary Fig. 1a, b). The 18
59 nt ssRNA products accumulated with increasing MgR concentration, which confirmed
60 the versatility of the enzyme activity. In addition, we also performed a reaction of MgR
61 with the synthesized 40 nt ssRNA substrate specifically carrying only the Am site at
62 position 28 to demonstrate the exact breakpoint of MgR activity, and the products were
63 TA cloned and sequenced (Fig. 1b). Importantly, we discovered that all ssRNA
64 products ended at the Am+1 position (20/20), supporting the detection of Nm sites by
65 the accumulated 3′ end of MgR reaction products. To further confirm its specificity on
66 Nm, an ssRNA substrate with m^6A and m^5C sites downstream of the Am site was
67 incubated with MgR, demonstrating clear products ending at the Nm+1 site but nothing
68 at other modification sites (Supplementary Fig. 1c). Finally, we tested an ssRNA
69 substrate with a long stem and found that it could also be digested to the Nm+1 site by
70 MgR, which proved that MgR could work to RNA with a strong secondary structure
71 (Supplementary Fig 1d, e).

72 Taking all the above information together, we combined MgR with NGS to develop
73 a new high-throughput Nm screening method named NJU-seq (Nm-Judged-
74 Universally sequencing). As detailed in the Methods, there were five main steps in
75 NJU-seq (Fig. 1c). A sample prepared with the same process without MgR treatment
76 was used as a negative control (NC) group to filter the false-positive outcomes caused
77 by strong RNA secondary structure. First, we tested NJU-seq on total RNA from HeLa
78 cells and analyzed the rRNA results. The sequencing data were aligned to a reference
79 sequence, and reads' 3′ ends were counted (Fig. 1d). Comparing rRNA alignments after
80 MgR digestion to NC, 3′ end enrichments correspond to potential Nm sites. For
81 example, between positions 157 to 169 in 18S rRNA, we found that the reads ending at
82 positions 160 and 167 significantly accumulated, indicating two possible Am sites at
83 positions 159 and 166 (Fig. 1e). Besides, sites' read distributions with eight other RNA
84 modifications, including common m^6A and pseudouridine (ψ), were browsed (Fig. 1f,
85 g, Supplementary Fig. 2), confirming that read 3′ end accumulation only occurred at
86 sites 1 nt downstream of the Nm site.

87 To observe the specificity and sensitivity of NJU-seq, we knocked down known
88 rRNA 2′-O-methyltransferase fibrillarin (*FBL*) by siRNA in the HeLa cell line. *FBL*
89 mRNA expression was significantly decreased, as confirmed by qPCR (Fig. 1h). After
90 normalization to total rRNA-aligned reads, fragments ending 1 nt downstream reported
91 Nm sites in the *FBL*-KD group were significantly fewer than those in the untreated
92 group (Fig. 1i, j). Such results supported that NJU-seq can reflect the copy number of
93 fragments with 2′-O-methylation.
94 We scored each site with the difference of reads end 1 nt downstream divided by
95 reads end just at the target site between MgR treatment and NC group (Supplementary
96 Fig. 3). The score was mainly correlated with the copy number of fragments with 2′-O-
97 methylation modification at the Nm site but not the methylation ratio, such as 2′OMe-
98 seq or RiboMethSeq. Candidate Nm sites were further judged by whether reads ending
99 1 nt downstream of the MgR treatment group were significantly more abundant than
100 the NC group to exclude false-positive results by strong RNA secondary structure. The
101 details of the scoring and filtering steps are described in Methods.
102 **HeLa rRNA Nm sites identification by NJU-seq**

103 We screened Nm sites on the HeLa cell line rRNA from three independent sample
104 repeats. The rRNA scores were consistent among the three sample repeats. Since the
105 Nm sites on rRNA of the HeLa cell line were well identified in several previous studies,
106 we compared our results with reported Nm sites to estimate the accuracy of the results
107 generated by the NJU-seq platform (Fig. 2a). We selected all sites with a score higher
108 than 30 and a fold change greater than 3 (Supplementary Fig. 3) and acquired 54, 39,
109 and 1 potential Nm sites in 28S, 18S, and 5.8S rRNA. 46 of 54, 32 of 38, and 1 of 1
110 sites were reported to be 2′-O-methylated in previous studies[14] (Fig. 2a). To validate
111 the novel discovered Nm sites, the retrotranscription-based method RTL-P was
112 applied[19]. Five previously reported and nine discovered Nm sites demonstrated
113 significant retrotranscription efficiency changes compared with unmodified sites (Fig.
114 2b). However, due to the high density of Nm sites around the target Nm site to be tested,
115 some of the RT fold changes may not be contributed only by the target site (Fig. 2c).

116 **Establish Nm-VAQ as a site-specific validation and quantification tool of Nm**
117 **status**

118 RTL-P is a semi-quantification method that can only validate sites with high 2′-O-
119 methylation and would report high false positive or negative results under certain
120 situations[19]. Hence, we established the Nm-VAQ (an RNase H-based Nm Validation
121 and Absolute Quantification tool) to validate candidate sites more accurately. Briefly,
122 Nm-VAQ was based on selective cleavage of RNase H and the guidance of a DNA-
123 RNA hybrid probe to target the cleavage site. With several tests of multiple designs of
124 hybrid probes, we anchored the RNase H cleavage site to the target site with a probe
125 containing 4 DNA and 13 RNA, as shown (Fig. 2d, Supplementary Fig. 4). Combined
126 with qPCR, Nm-VAQ quantified the mixed substrates with 2′-O-methylation gradients
127 accurately as low as 0.78% (Fig. 2e, Supplementary Fig. 4), while RTL-P could only
128 detect Nm sites with modifications higher than 40%[19] and presented significantly
129 different results under different conditions (Supplementary Fig. 5). In addition, the tests
130 of Nm-VAQ with 50% 2′-O-methylated substrates at 6.25 fmol, 0.625 fmol, and 0.0625
131 fmol demonstrated highly consistent results, which proved its evaluation capability in
132 low-amount RNA studies (Fig. 2f). Although nine novel sites were validated as Nm
133 sites with similar retrotranscription efficiency changes, Nm-VAQ further acquired an

134 accurate 2′-O-methylation status of 0.3%-74% (Fig. 2g). Besides, the methylation
135 levels of several selected Nm sites between the HeLa negative control and *FBL*-KD
136 groups decreased by 8-20%, as quantified by Nm-VAQ (Fig. 2h). To date, Nm-VAQ
137 has allowed the study of Nm to the absolute quantification stage.

138 **Mapping of Nm sites on mRNA of human cell lines**

139 We next investigated Nm sites on mRNAs from three cell lines, HeLa, HEK293T,
140 and A549. With rRNA elimination, the proportion of mRNA in the substrate was
141 substantially increased before MgR treatment. mRNA Nm sites shared with all three
142 independent samples of each cell line were selected for the following analysis and
143 statistics, as detailed in Methods. The same standard selected the potential Nm sites on
144 mRNA as rRNA, which acquired 4074/2059/2367 sites in HeLa, HEK293T, and A549
145 cell lines (Fig. 3a). We first selected 5 Nm sites on HeLa mRNA as targets for validation
146 by RTL-P. To our surprise, only three of them showed positive results compared with
147 *GAPDH* as a negative control (Supplementary Fig. 6). On the other hand, the results of
148 Nm-VAQ confirmed that their methylation ratios were in the range of 0.4-16.8%, which
149 was out of the validation capability of RTL-P (Fig. 3b). Since the mRNA 2′-O-
150 methylation status of most sites may be much lower than rRNA and RTL-P was limited
151 by false positive, false negative, semiquantitative and non-single nucleotide resolution
152 defects, Nm-VAQ was more appropriate in the following study than RTL-P (Fig. 3b).

153 Three Nm sites shared by three cell lines were selected for further validation:
154 *BCL6* (3, 187721452, Gm), *PLXNB2* (22, 50290219, Gm), and *TBX2* (17, 61401930,
155 Cm). Nm-VAQ confirmed their methylation and evaluated 2′-O-methylation ratios as
156 0.4-18% (Fig. 3b). One site shared by HeLa and HEK293T cells, *ZNFX1* (20, 49271223,
157 Um), was validated with a 1-2% methylation ratio. Two Nm sites only present in HeLa,
158 *ENOX2* (X, 130703233, Am) and *CLEC2B* (12, 9869269, Am), were confirmed to be
159 methylated, and 2′-O-methylation ratios of 2-7% were evaluated (Fig. 3b). All validated
160 sites demonstrated clear 2’-O-methylation status but were much lower than Nm sites
161 on rRNA.

162 To identify the overall distribution of 2′-O-methylation, we analyzed the sequence
163 composition around the Nm site. We found that Nm modification occurred in a distinct
164 pattern on different ribonucleotides but was similar among the three cell lines (Fig. 3c).
165 Although A on mRNA may act as a branch point in the intron part and regulate
166 alternative splicing by its 2′-O-methylation status[11], more 2′-O-methylations were
167 modified on C and G than A and U. We found high GC content around the Nm sites in
168 the three cell lines, especially G, which might indicate modification bias
169 (Supplementary Fig. 7a). When we further calculated Nm-1 and Nm+1 ribonucleotides
170 with Nm together as a tri-ribonucleotide unit, the results showed an apparent uneven
171 distribution with a clear enrichment of UGmU among all three cell lines (Fig. 3c).

172 **2′-O-methylation is involved in mRNA alternative splicing events and translation**

173 To identify the potential functions of 2′-O-methylation modifications, we analyzed
174 Nm′s distribution patterns in different functional mRNA regions. By exploring the 2′-
175 O-methylation distribution pattern on mRNA transcripts, we found that in the three cell
176 lines, Nm numbers were CDS>3′UTR>5′UTR, which demonstrated the highest density
177 in the 5′UTR considering the length (Fig. 3d). To further investigate the Nm site
178 distribution, we focused on the start and stop codons (Supplementary Fig. 7b, c). 3,479
179 HeLa Nm sites located on mRNAs that could cover -50 to +500 nt around the start
180 codon were selected to count the distribution in 50nt-wide windows. Similarly, 3,074
181 HeLa Nm sites were selected to evaluate the stop codon's -500 to +500 nt regions.
182 Interestingly, both areas demonstrated apparent downward trends of Nm sites from 5′
183 to 3′ direction. Such a pattern was also seen in the other two cell lines.

184 Such results implied that there were relatively stable patterns of Nm distribution
185 in the functional regions of mRNA. We compared the translation efficiency of genes
186 with and without 2′-O-methylation (Fig. 3e). The translation efficiency was reflected
187 by protein abundance/gene expression. Significantly more genes with Nm site(s)

188 showed higher translational efficiency, which strongly suggests a correlation between
189 Nm sites and translation, confirming the vital role of 2′-O-methylation in translational
190 regulation. Previous studies elucidated that pre-mRNA splicing of *ACT1* was affected
191 by Am modification, so we explored whether Nm sites were extensively involved in
192 mRNA splicing[11]. Interestingly, there were more Nm sites closer to the 5′ exon
193 boundary than those closer to the 3′ exon boundary, significant in the 0-100 window
194 from the boundary (Fig. 3f, Supplementary Fig. 7d). Such discoveries positively
195 indicated the involvement of 2′-O-methylation in mRNA splicing. To further
196 demonstrate the effects on splicing, we assigned mRNA Nm sites as constitutively
197 spliced exons (CSEs) or alternatively spliced exons (ASEs)[10, 20]. The total length of
198 ASE and CSE with modifications was counted, and the density of Nm sites was
199 obtained. We found that more Nm sites were significantly located in the CSE than in
200 the ASE when considering their different lengths (Fig. 3g, Supplementary Fig. 7e).
201 Such results suggested the importance of 2′-O-methylation in the occurrence and
202 regulation of splicing during mRNA maturation.
203

204 **Discussion**

205 The 2′-O-methylation modification of RNA ribose affects diverse biological processes.
206 Thus, high-throughput identification for comprehensive and accurate profiling of Nm
207 sites is vital for further research on cellular regulation and functions of 2′-O-
208 methylation. Three high-throughput Nm detection methods have been developed in the
209 past few years: 2′OMe-seq[16], RiboMethSeq[7, 21, 22], and Nm-Seq[17, 23]. Based on this
210 principle, 2′OMe-seq and RiboMethSeq were able to identify Nm sites with abundant
211 amounts and high methylation ratios, such as Nm sites on rRNA. Based on our results,
212 the methylation ratio of all 13 tested mRNA Nm sites was lower than 20%, which
213 indicated that most Nm sites on mRNA were unable to be detected by 2′OMe-seq and
214 RiboMethSeq. On the other hand, the Nm-seq method demonstrated improved mRNA
215 Nm site detection direct targeting to mRNA molecules with 2′-O-methylation[17, 23],
216 similar to our new method NJU-seq. However, two main problems remained and were
217 solved by NJU-seq: 1) Nm-seq based on multiple cycles of OED+RNA purification to
218 enrich the Nm signal, which resulted in a massive loss of RNA content and took at least
219 16 hours. For NJU-seq, enrichment of the Nm+1 site can be accomplished with only 30

min MgR digestion for only one reaction with much less RNA sample. 2) Ligation bias of fragments. After Nm-seq treatment, the 3′ end of the RNA was 2′-O-methylated, resulting in a much lower ligation efficiency than the unmodified end in library construction[24]. Thus, some Nm sites were insufficient to obtain more sequencing reads than surrounding non-modified sites. For NJU-seq, the RNA fragment ended at the Nm+1 site without 2′-O-methylation, which has no bias in the ligation step.

In this study, 14 sites on rRNA and 13 sites on mRNA were selected, and most of them were further confirmed to be 2′-O-methylated, which demonstrated high specificity of detecting Nm sites. The method can also be used to study 2′-O-methylation on other RNAs, such as tRNA. For example, Cm_{32} $tRNA^{Gln}$ (UUG) turned out to be consistent with sites reported in previous studies[25], and Gm_{35} $tRNA^{Asn}$ (GmUU) was a new Nm anticodon (Supplementary Fig. 8). NJU-seq achieved high consistency in mRNA duplicate samples. These results indicated the stable broad-spectrum applicability of NJU-seq in detecting Nm modification in various RNA molecules with high complexity and low content.

235 The site-specific Nm validation and quantification method Nm-VAQ was another
236 new tool established in this study to fulfill the Nm quantification gap in the study of 2′-
237 O-Methylation. Several previously developed low-throughput Nm detection methods,
238 including LC-MS, RTL-P, and DNA polymerase, have various defects[18, 19, 26]. LC-MS
239 is labor-intensive and difficult for mRNA Nm detection due to the requirement for a
240 large number of RNA molecules[18]. Both RTL-P and DNA polymerase relied on the
241 blocking of Nm on reverse transcription, which can be called RT-based methods[19, 26].
242 As the schematic illustrates, any Nm site between the amplification products will
243 generate a methylation signal, and thus, these two detections are non-site-specific
244 methods (Fig. 2c). Meanwhile, although both RTL-P and DNA polymerase methods
245 could acquire linear results correlated with the methylation ratio with synthetic RNA,
246 the result varied with different amounts of target RNA. As we observed in the
247 evaluation of 18S 1391Cm, the results were significantly different in three RNA
248 amounts, making it difficult to accurately evaluate the 2′-O-methylation ratio of the
249 actual target (Supplementary Fig. 5b). In addition, the original study of RTL-P
250 mentioned false positive and negative results, which may be caused by RNA secondary

251 structures that can occur on several rRNA sites[19]. Nm-VAQ demonstrated apparent
252 advantages compared to those methods, as listed in Table 1. Nm-VAQ anchored the
253 cleavage position to the target site directed by the chimera and discriminated 2′-O-
254 methylated RNA from unmethylated RNA molecules by RNase H. This method
255 simultaneously acquired the absolute amount of accurate 2′-O-methylation of the target
256 site. In addition, Nm-VAQ showed its capability to consistently evaluate targets with
257 low amounts or low methylation ratios, which demonstrated its capability to study the
258 status of specific Nm sites and their dynamic changes.

259 In this study, thousands of new Nm sites were uncovered on mRNAs from humans
260 and mice, installing multiple areas (UTR/CDS, CSE/ASE), revealing its broad
261 distribution. Unlike Nm sites on rRNA, all 13 selected and tested Nm sites on mRNA
262 were 2′-O-methylated at less than 20%. This ratio was similar to m^6A, which may be
263 enough to regulate the function of mRNA[27]. Since NJU-seq and Nm-VAQ studied a
264 group of cells from cell lines or tissues, it was unclear whether all cells demonstrated
265 such low modification or some cells were highly methylated while others were not,
266 which required other methods to explore further. A similar distribution pattern of 2′-O-

methylation was shared in different cell lines. The non-random distribution of Nm sites and the occurrence of 2′-O-methylation were tightly regulated in critical functional regions, suggesting that Nm might be an essential regulatory means. Thus, 2′-O-methylation modifications played an essential role in mRNA metabolism. In contrast to m^1A with enrichment in the start codon or m^6A with accumulation around the stop codon[28-30], the Nm distribution pattern was more complicated than these two RNA modifications and may possess broader biological functions.

In the future, with the help of NJU-seq and Nm-VAQ, further studies about the function, modification/demodification/regulation of 2′-O-methylation could be extended. On the other hand, 2′-O-methylation demonstrated a critical function in environmental response and great potential in disease study, diagnosis, and drug discovery.

280 **Materials and Methods**

281 **Cloning, expression, and purification.** The RNase R sequence from *Mycoplasma*
282 *genitalium* (accession number: WP_009885662.1) was synthesized by GenScript
283 Biotech Co. and inserted into the pET28a vector (+) with digestion and ligation at the
284 NdeI and BamHI sites, followed by confirmation by Sanger sequencing. The plasmid
285 was transferred into *E. Coli*. BL21 (DE3) strain. Cells were grown in 1 L LB media
286 under 37 °C. IPTG (0.3 mM) was added to every 1 L medium in a shaker incubator to
287 induce protein expression at 16 °C for 18 h. Cell pellets were harvested and lysed by
288 French press with lysis buffer (50 mM Tris-HCl at pH 8.0, 250 mM NaCl, 0.5 mM
289 TCEP), followed by centrifuging at 10000 rpm for 1 h. MgR protein was purified by
290 passing through a column Nickle resin (GenScript Biotech Co.) and washed by using 3
291 column volumes of wash buffer (50 mM Tris-HCl at pH 8.0, 250 mM NaCl, 20 mM
292 imidazole, 0.5 mM TCEP), then eluted with elution buffer (50 mM Tris-HCl at pH 8.0,
293 250 mM NaCl, 100 mM imidazole, 0.5 mM TCEP). The elution was concentrated and
294 aliquoted into multiple small tubes and stored at -80 °C for further experiments. Protein

295 was confirmed by SDS–PAGE gel, and its concentration was determined by standard

296 BSA protein on the same gel.

297 **In vitro assay of RNA hydrolysis**. Nm-modified ssRNA substrate was synthesized by

298 GenScript Biotech Co. (5′6-Fam-UAACC UAUGA AGNNN UNmNNC UC-3′).

299 Briefly, various purified proteins were reacted with 600 ng ssRNA substrate in reaction

300 buffer (20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.01 mM $ZnCl_2$) at 37 °C for 30 mins.

301 The reaction was terminated by heating to 85 °C for 10 mins. The products were

302 separated by 20% urea-PAGE, visualized with a ChemiDoc XRS+\UnUniversal

303 alHoodII gel imaging system (Bio-Rad) or Tanon 3500, and quantified with ImageJ

304 software. A series of reactions with 400 ng ssRNA were designed to test the range of

305 pH and reaction duration. Subsequently, multiple ssRNA substrates were incubated

306 with 18 μg MgR in a 10 μl reaction (20 mM Tris-HCl pH 8.0, 100 mM KCl, 0.01 mM

307 $ZnCl_2$) at 37 °C for 30 mins to test MgR properties. The ssRNA sequences were as

308 follows: ssRNA substrate with multiple modifications (5′6-Fam-AUACC GCAGC

309 UAGAA UGAmGC Um^5CAGA UGm^6AAA-3′) and ssRNA substrate with hairpin

310 structure (5′6-Fam-UAGCU CUGGA AAUUC UAGAmG CUAAU ACAAA-3′).

311 Another Nm-modified ssRNA substrate was synthesized by GenScript Biotech Co. (5′-
312 AGCCG CCUGG AUACC GCAGC UAGAA UGAmGC UGAGA UGAAA-3′) to test
313 the consistency of hydrolysis. After reacting with MgR, reaction products were purified
314 by an RNA Clean & Concentrator-5 kit (Zymo Research), followed by 5′
315 phosphorylation (New England BioLabs), repurification, and adaptor ligation (New
316 England BioLabs). The final products were TA cloned and analyzed by Sanger
317 sequencing.

318 **Cell culture and RNA extraction.** HeLa cells, HEK293T cells, and A549 cells, were
319 purchased from the Shanghai Institute of Cell Biology, Chinese Academy of Sciences
320 (Shanghai, China). All cells were grown in Dulbecco's modified Eagle's medium
321 (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 units/ml penicillin,
322 and 100 μg/ml streptomycin at 37 °C with 5% CO_2.

323 Total RNA from cells was extracted by the Qiazol RNA extraction kit (QIAGEN).
324 Extracted total RNA was confirmed by Nanodrop (Thermo Fisher) and 1% agarose gel
325 electrophoresis (Sangon Biotech).

326 **Knockdown of FBL in HeLa cell line.** For siRNA experiments, siRNA duplexes were

327 used for fibrillarin silencing: 5′-GUCUU CAUUU GUCGA GGAA-AdTdT-3′ (sense

328 sequence). Control siRNA does not target any human sequence (RiboBio). HeLa cells

329 were transfected using LIPOFECTAMINE RNAIMAX reagent (Invitrogen) according

330 to the manufacturer's instructions and collected after 48 h. *FBL* mRNA expression was

331 quantified by a HiScript II 1st Strand cDNA Synthesis Kit (+gDNA wiper) and ChamQ

332 Universal SYBR qPCR Master Mix (Vazyme Biotech Co., Ltd.) according to

333 instructions.

334 **rRNA elimination.** Except for HeLa, HEK293T, and A549 cells, which used 15 μg

335 total RNA, all the other samples were processed with 10 μg. The RNA was diluted to

336 29 μl, and 1 μl rRNA probe (H/M/R) was added (Vazyme Biotech Co., Ltd.). RNA was

337 hybridized with the probe and heated to 95 °C for 2 min, followed by gradual cooling

338 to 22 °C at 0.1 °C/s and maintenance for 5 min. Next, the rRNA was eliminated by

339 adding 1 μl RNase H (New England Biolabs), 4 μl 10X RNase H buffer, 4 μl DEPC

340 water and incubating at 37 °C for 30 min. Subsequently, the hybridization and RNase

341 H reaction were repeated 5/3 times to eliminate the rRNA fragment as much as possible.

342 Finally, the probe was digested by 2 units of DNaseI (New England Biolabs) at 37 °C

343 for 30 min. According to the manufacturer's protocol, the reaction mixture was purified

344 using VAHTS RNA Clean Beads (Vazyme Biotech Co., Ltd.).

345 **NJU-seq.** RNA was fragmented by an RNA fragmentation reagents kit (Thermo Fisher)

346 at 95 °C for 10 mins and then purified by 10 μl 3-M NaOAc (pH 5.2), 24 μl 5-μg/μl

347 glycogen, and 372 μl 96% cold ethanol. For facilitating MgR binding of fragments with

348 complex secondary structures after solubilization in water, every 1 μg RNA solution

349 was ligated with 3 μl 25-μM poly(A)-ssRNA adaptor (pAGCUA AAAAA AAAAA

350 Ap, synthesized by GenScript Biotech Co.) at 16 °C overnight. The products were then

351 reacted with 30 μg MgR in a 60 μl final volume with 20 mM Tris-Cl (pH 8.0), 100 mM

352 KCl, and 0.01 mM $ZnCl_2$ at 37 °C for 30 min. Next, the products were purified by an

353 RNA Clean & Concentration-5 kit (Zymo Research), followed by 5′ phosphorylation

354 (39 μl RNA, 5 μl (10X T4 PNK Reaction Buffer), 5 μl ATP (10 mM), 1 μl T4 PNK (3′

355 phosphatase minus, New England BioLabs)). Finally, we constructed the NGS library

356 using the products from the last step following the NEBNext Small RNA Library Prep

357 Set for Illumina (New England BioLabs) protocol. Fragments smaller than 50 nt were

358 extracted from a 6% PAGE gel and sent for NGS using Illumina X-Ten. The exact

359 process prepared the NC sample (negative control group) without MgR protein.

360 **NGS data analysis and statistics.** Nm site detection and various statistical analyses

361 were performed using homemade Perl and R scripts and open-source bioinformatics

362 tools. Adaptors were filtered from raw reads using Cutadapt (Version 2.10). Reads were

363 mapped the reference sequences by Bowtie 2[31] (Version 2.3.4). rRNA reference

364 sequences are shown, and the transcript reference sequences were from GENCODE

365 release 34 and M25[32]. Only protein-coding transcripts with 'BASIC' tags and without

366 'cds_start_NF' or 'cds_end_NF' tags were retained. Perl scripts handled alignment

367 results, including filtering low-quality or multiple-hit alignments and calculating each

368 site's reads' end count. The likelihood of 2′-O-methylation modification at each site was

369 judged based on the following scoring system.

370 **Scoring 2′-O-methylation modification.** As shown in Supplementary Fig. 3, we

371 counted the stacking position of the 3′ end of the reads for the treatment group with

372 MgR digestion (TR) and the control group without digestion (NC). First, normalization

373 correction was performed according to the total aligned reads. Then, scoring was

374 performed for each site on the rRNA and transcriptome reference based on the number

375 of reads ending at the current position (N) and the latter position (N+1). The NJU-seq

376 score algorithm was as follows.

377 $$Score = \frac{3'Count(TR_{N+1})}{3'Count(TR_N)} - \frac{3'Count(NC_{N+1})}{3'Count(NC_N)}.$$

378 The average of the three replicates' scores was taken and ranked to complete

379 identification for modifications, and here, we took 30 as a threshold. To prevent some

380 points with a similar high N+1 enrichment in NC from being misclassified, we also

381 required $3'Count(TR_{N+1}) > 3 \times mean(3'Count(NC_{N+1}))$. At this point, we

382 finished processing the NJU-seq data and acquiring the Nm sites for subsequent

383 analysis.

384 **Statistics on genomic features.** We mainly used Perl and R scripts to perform tri-

385 ribonucleotide unit statistics, presentation of distribution positions, and alternative

386 splicing region statistics on all Nm sites.

387 Sequence probability logo plots for Nm sites were drawn by the R package

388 *ggseqlogo*[33]. The length-normalized mRNA transcript was obtained by averaging all

389 Nm-located transcripts' 5′UTR/CDS/3′UTR length of each sample. Equal length bins

390 of three regions were taken for Nm site counting in each standard transcript. We
391 counted the absolute distance for the Nm site position to the start and stop codon. ASE
392 and CSE were obtained from the statistics of all alternative splicing events provided in
393 the SpliceSeq database[20].

394 **Protein abundance.** The HeLa protein abundance data used were downloaded from
395 the Protein Abundances Across Organisms (PaxDb, Experiment ID 182) [34, 35]. We used
396 transcriptome expression data from the ENCODE[36] portal with the following identifiers:
397 ENCFF796REI and ENCFF846THO, and only kept genes with TPM > 0.2. The 6,881
398 genes with both protein abundance and mRNA expression were used for subsequent
399 analysis. Among them, the number of Nm-located genes was 1,397.

400 **Quantitation of RNA Nm status by RTL-P method**. Two sets of primers were
401 designed for the target site, with the F_U forward primer located upstream of the Nm site
402 and the F_D forward primer located downstream of the Nm site. For rRNA Nm detection,
403 the high dNTP concentration reaction mixture consisted of 5× RT Buffer 4 μl, M-MLV
404 (H-) Reverse Transcriptase (Vazyme Biotech Co., Ltd., 200 U/μL) 1 μl, RNase inhibitor
405 (40 U/μL) 1 μl, RNA 100 ng, RT primer (10 μM) 1 μl, RNase-free H_2O 5 μl, dNTPs (1

406 mM each) 1 μl. The low dNTP concentration reaction mixture was replaced by dNTPs

407 (2 μM each). Then, the mixtures were incubated at 45 °C for 1 hour and 85 °C for 2

408 min. The cDNA was diluted 100-fold and subsequently subjected to qPCR, with both

409 F_U and F_D products amplified for each cDNA. For mRNA Nm detection, 1 μg of total

410 RNA was used for further RTL-P tests. RT efficiency and RT fold change were

411 calculated according to the following strategy. RT efficiency=template amount

412 measured by F_U and R/template amount measured by F_D and R. RT fold change=RT

413 efficiency with low dNTPs/RT efficiency with high dNTPs[17, 19].

414 **The assay of RNase H cleavage**. GenScript Biotech Co synthesized RNA

415 oligonucleotides and chimera probes. Briefly, 25 pmol RNA oligonucleotides were

416 mixed with 150 pmol chimera probe, heated to 95 °C for 2 min, cooled to 22 °C at

417 0.1 °C/s, and maintained for 5 min. Next, the hybrid was reacted with 1 μl RNase H

418 (New England BioLabs) at 37 °C for 30 min and then heated to 90 °C for 10 min to

419 terminate the reaction. The cleavage products were added to RNA Dye (New England

420 BioLabs), analyzed by 20% UERA-PAGE, and visualized by a ChemiDoc

421 XRS+\UnUniversal alHoodII gel imaging system (Bio-Rad).

422 **Nm-VAQ assay of RNA 2′-O-methylation ratio.** Synthetic RNA oligonucleotides
423 with/without Nm were obtained from GenScript Biotech Co., and Sangon Biotech Co.
424 synthesized qRT–PCR primers. The oligonucleotides with/without Nm were mixed to
425 obtain gradient 2′-O-methylation ratio substrates. Then, $6.25*10^{-2}$ pmol substrate was
426 added to 6.25 pmol D(4)R(13) chimera probe at 95 °C for 2 min, followed by cooling
427 to 22 °C at 0.1 °C/s and 22 °C for 5 min. The products were divided into 2 parts for the
428 following reaction. One mixture contained 5 μl previous product, 1 μl RNase H (New
429 England BioLabs), 1 μl 10X RNase H Reaction Buffer, and 3 μl RNase-free H_2O.
430 RNase H storage buffer was substituted for RNase H to form the other mixture to serve
431 as a blank control, followed by 30 min at 37 °C and 10 min at 90 °C. The products were
432 diluted 50-fold and then used to obtain cDNA by a HiScript II 1st Strand cDNA
433 Synthesis Kit (Vazyme Biotech Co., Ltd.). cDNA was diluted 100-fold, and then qPCR
434 was conducted in a 20 μl reaction mixture containing 10 μl ChamQ Universal SYBR
435 qPCR Master Mix (Vazyme Biotech Co., Ltd.), 0.4 μl F primer, 0.4 μl R primer, 2 μl
436 cDNA, and 7.2 μl H_2O, following the protocol: 30 s at 95 °C, then 40 cycles of 95 °C

437 for 10 s and 60 °C for 30 s. Each cDNA was analyzed in 3 replicates. ΔCT (cycle

438 threshold) = Ct value of RNase H reaction- Ct value of control.

439 **Quantitation of RNA Nm status by Nm-VAQ method**. All primers were obtained

440 from Sangon Biotech Co. For rRNA Nm detection, 100 ng total RNA and 6.25 pmol

441 chimera probes (molar ratio of rRNA molecules to chimera probes was ~ 1:200) were

442 hybridized, followed by RNase H cleavage, reverse transcription and qPCR (cDNA

443 was diluted 100-fold) according to Nm-VAQ protocol. ΔCT was used calculate the

444 ratio of modification. In addition, some treatments were adjusted to perform mRNA

445 Nm detection. 2 μg of total RNA was used for hybrids, and 4 μl of cDNA was used for

446 the qPCR assay.

447

448 **REFERENCES AND NOTES**

449 1. Motorin, Y. & Helm, M. RNA nucleotide methylation. *Wiley Interdiscip Rev*
450 *RNA* **2**, 611-631 (2011).
451 2. Jockel, S. et al. The 2'-O-methylation status of a single guanosine controls
452 transfer RNA-mediated Toll-like receptor 7 activation or inhibition. *J Exp Med*
453 **209**, 235-241 (2012).
454 3. Deryusheva, S., Choleza, M., Barbarossa, A., Gall, J.G. & Bordonne, R. Post-
455 transcriptional modification of spliceosomal RNAs is normal in SMN-deficient
456 cells. *RNA* **18**, 31-36 (2012).
457 4. Tycowski, K.T., Smith, C.M., Shu, M.D. & Steitz, J.A. A small nucleolar RNA
458 requirement for site-specific ribose methylation of rRNA in Xenopus. *Proc Natl*
459 *Acad Sci U S A* **93**, 14480-14485 (1996).
460 5. Zust, R. et al. Ribose 2'-O-methylation provides a molecular signature for the
461 distinction of self and non-self mRNA dependent on the RNA sensor Mda5. *Nat*
462 *Immunol* **12**, 137-143 (2011).
463 6. Kawai, G. et al. Conformational rigidity of specific pyrimidine residues in
464 tRNA arises from posttranscriptional modifications that enhance steric
465 interaction between the base and the 2'-hydroxyl group. *Biochemistry* **31**, 1040-
466 1046 (1992).
467 7. Krogh, N. et al. Profiling of 2'-O-Me in human rRNA reveals a subset of
468 fractionally modified positions and provides evidence for ribosome
469 heterogeneity. *Nucleic Acids Res* **44**, 7884-7895 (2016).
470 8. Ji, L. & Chen, X. Regulation of small RNA stability: methylation and beyond.
471 *Cell Res* **22**, 624-636 (2012).
472 9. Ma, Y. et al. mRNA cap methylation influences pathogenesis of vesicular
473 stomatitis virus in vivo. *J Virol* **88**, 2913-2926 (2014).
474 10. Ge, J., Liu, H. & Yu, Y.T. Regulation of pre-mRNA splicing in Xenopus oocytes
475 by targeted 2'-O-methylation. *RNA* **16**, 1078-1085 (2010).
476 11. Zhao, X. & Yu, Y.T. Targeted pre-mRNA modification for gene silencing and
477 regulation. *Nat Methods* **5**, 95-100 (2008).
478 12. Elliott, B.A. et al. Modification of messenger RNA by 2'-O-methylation
479 regulates gene expression in vivo. *Nat Commun* **10**, 3401 (2019).

480 13. Lalonde, M.S. et al. Exoribonuclease R in Mycoplasma genitalium can carry
481 out both RNA processing and degradative functions and is sensitive to RNA
482 ribose methylation. *RNA* **13**, 1957-1968 (2007).
483 14. Ayadi, L., Galvanin, A., Pichot, F., Marchand, V. & Motorin, Y. RNA ribose
484 methylation (2'-O-methylation): Occurrence, biosynthesis and biological
485 functions. *Biochim Biophys Acta Gene Regul Mech* **1862**, 253-269 (2019).
486 15. Daffis, S. et al. 2'-O methylation of the viral mRNA cap evades host restriction
487 by IFIT family members. *Nature* **468**, 452-456 (2010).
488 16. Incarnato, D. et al. High-throughput single-base resolution mapping of RNA 2-
489 O-methylated residues. *Nucleic Acids Res* **45**, 1433-1441 (2017).
490 17. Corrigendum: Nm-seq maps 2'-O-methylation sites in human mRNA with base
491 precision. *Nat Methods* **15**, 226-227 (2018).
492 18. Douthwaite, S. & Kirpekar, F. Identifying modifications in RNA by MALDI
493 mass spectrometry. *Methods Enzymol* **425**, 3-20 (2007).
494 19. Dong, Z.W. et al. RTL-P: a sensitive approach for detecting sites of 2'-O-
495 methylation in RNA molecules. *Nucleic Acids Res* **40**, e157 (2012).
496 20. Ryan, M.C., Cleland, J., Kim, R., Wong, W.C. & Weinstein, J.N. SpliceSeq: a
497 resource for analysis and visualization of RNA-Seq data on alternative splicing
498 and its functional impacts. *Bioinformatics* **28**, 2385-2387 (2012).
499 21. Erales, J. et al. Evidence for rRNA 2'-O-methylation plasticity: Control of
500 intrinsic translational capabilities of human ribosomes. *Proc Natl Acad Sci U S*
501 *A* **114**, 12934-12939 (2017).
502 22. Birkedal, U. et al. Profiling of ribose methylations in RNA by high-throughput
503 sequencing. *Angew Chem Int Ed Engl* **54**, 451-455 (2015).
504 23. Dai, Q. et al. Nm-seq maps 2'-O-methylation sites in human mRNA with base
505 precision. *Nat Methods* **14**, 695-698 (2017).
506 24. Saikia, M. et al. A systematic, ligation-based approach to study RNA
507 modifications. *RNA* **12**, 2025-2033 (2006).
508 25. Marchand, V. et al. Next-Generation Sequencing-Based RiboMethSeq Protocol
509 for Analysis of tRNA 2'-O-Methylation. *Biomolecules* **7** (2017).
510 26. Aschenbrenner, J. & Marx, A. Direct and site-specific quantification of RNA 2'-
511 O-methylation by PCR with an engineered DNA polymerase. *Nucleic Acids Res*
512 **44**, 3495-3502 (2016).

27. Liu, N. et al. Probing N6-methyladenosine RNA modification status at single nucleotide resolution in mRNA and long noncoding RNA. *RNA* **19**, 1848-1856 (2013).
28. Dominissini, D. et al. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature* **530**, 441-446 (2016).
29. Meyer, K.D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635-1646 (2012).
30. Li, X. et al. Transcriptome-wide mapping reveals reversible and dynamic N(1)-methyladenosine methylome. *Nat Chem Biol* **12**, 311-316 (2016).
31. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**, 357-359 (2012).
32. Frankish, A. et al. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res* **47**, D766-D773 (2019).
33. Wagih, O. ggseqlogo: a versatile R package for drawing sequence logos. *Bioinformatics* **33**, 3645-3647 (2017).
34. Wang, M., Herrmann, C.J., Simonovic, M., Szklarczyk, D. & von Mering, C. Version 4.0 of PaxDb: Protein abundance data, integrated across model organisms, tissues, and cell-lines. *Proteomics* **15**, 3163-3168 (2015).
35. Geiger, T., Wehner, A., Schaab, C., Cox, J. & Mann, M. Comparative proteomic analysis of eleven common cell lines reveals ubiquitous but varying expression of most proteins. *Mol Cell Proteomics* **11**, M111 014050 (2012).
36. Davis, C.A. et al. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic Acids Res* **46**, D794-D801 (2018).

Acknowledgments

We are grateful to the High-Performance Computing Center of Nanjing University for performing the numerical calculations in this paper on its blade cluster system. In addition, we acknowledge all members of Dr. Chen Jian-Qun's lab and Dr. Li Shanqing

Dr. Liang naixin for comments and discussions. We also gratefully acknowledge Dr. Chan Robin, Dr. Zhang Hui, and Dr. Zhou Li for editing the paper.

Funding: This work was supported by the Youth Program of the National Natural Science Foundation of China (31801065) and Fundamental Research Funds for the Central Universities (021414380507).

Author contributions: QC, QW, and JC designed this research; YT, RX, XG, and YW performed this research; QC, QW, YT, and YW wrote the manuscript; YW and QW analyzed the data.

Competing interests: No competing interests declared.

Table 1 Comparisons of LC-MS, RTL-P, DNA polymerase, and Nm-VAQ methods.

Method	LC-MS	RTL-P	DNA polymerase	Nm-VAQ
Main principle	Purification of RNA fragments containing modification sites and specific enzymatic hydrolysis to separated and analyzed modification signal	2′-O-methylation induced reverse transcription stop at low dNTP concentrations	Am or Cm in 1nt downstream of the primer hindered the reverse transcription reaction	2′-O-methylation inhibited the cleavage of RNase H
Region/site specific	Yes	No	No	Yes
Effective target	Am/Gm/Cm/Um	Am/Gm/Cm/Um	Am /Cm	Am/Gm/Cm/Um
Methylation ratio	Yes	No	Yes	Yes
absolute quantification of methylation copy number	No	No	No	Yes
Potential false result	No	Yes	Unknown	No
Other	Large amounts of high purity RNA, which is very laborious, are not suitable for detecting modifications of low abundance RNA	The methylation ratio varied significantly with the different RNA expression		Precise quantification of Nm ratio and methylation copy numbers at unknown RNA concentrations

Figure 1

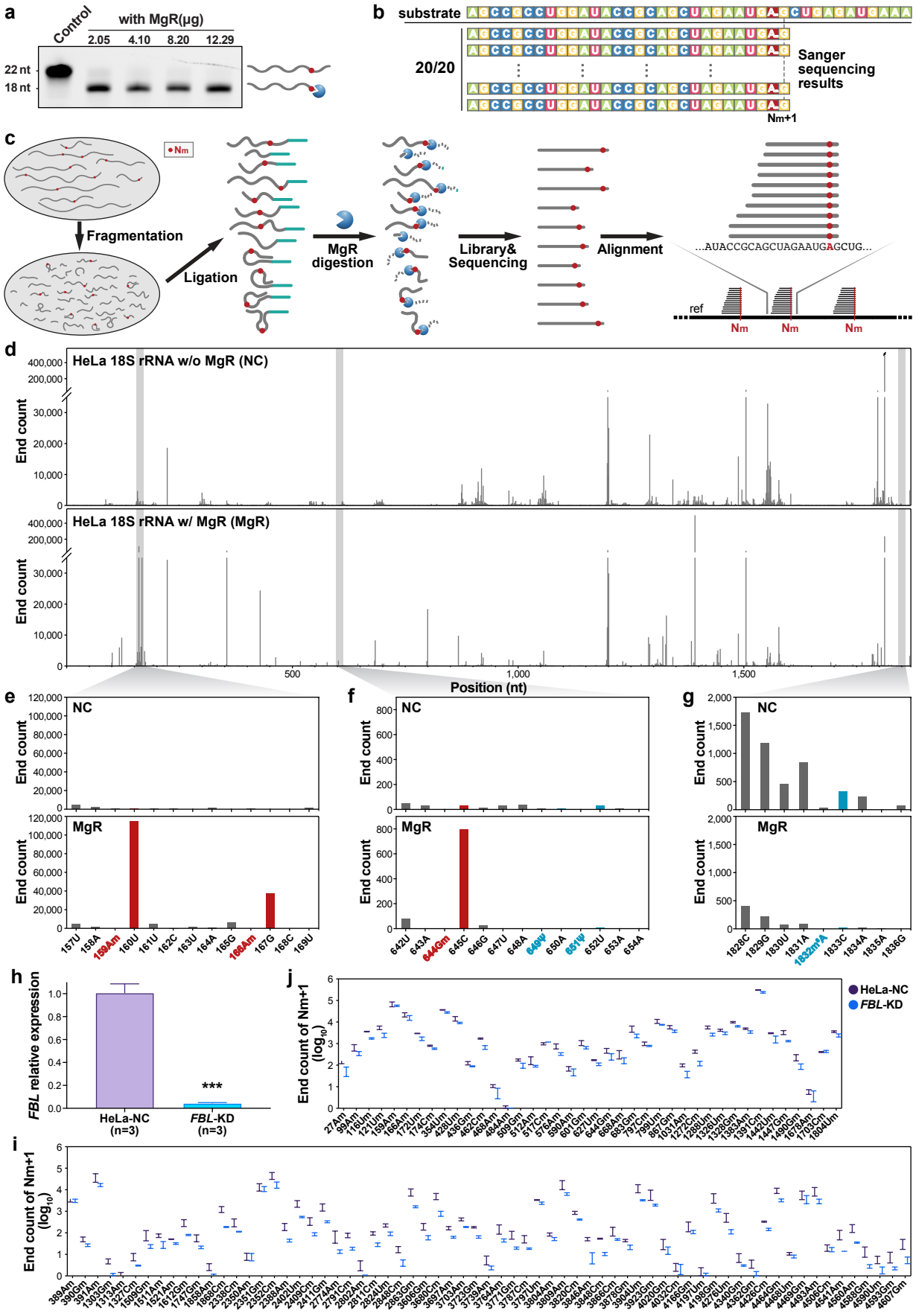

a
Control
with MgR(μg)
2.05
4.10
8.20
12.29
22 nt
18 nt
b
substrate
20/20
Sanger sequencing results
Nm+1
c
Nm
Fragmentation
Ligation
MgR digestion
Library& Sequencing
Alignment
...AUACCGCAGCUAGAAUGAGCUG...
ref
Nm
d
HeLa 18S rRNA w/o MgR (NC)
HeLa 18S rRNA w/ MgR (MgR)
End count
Position (nt)
e
NC
MgR
f
g
h
FBL relative expression
HeLa-NC (n=3)
FBL-KD (n=3)
***
j
End count of Nm+1 (log10)
HeLa-NC
FBL-KD
i

Figure 1 | Establish a novel high-throughput method, NJU-seq, for RNA 2′-O-methylation. a, In vitro ssRNA hydrolysis assay. The ssRNA substrate (5′-6-Fam-UAACCUAUGA AGNNNUNmNNC UC-3′, N represents the mix of A/G/C/U) was reacted with varying amounts of MgR. The left electrophoresis showed digestion products of MgR, and a schematic was shown on the right. **b**, TA clone MgR hydrolyzed ssRNA. The ssRNA substrate (5′-AGCCGCCUGG AUACCGCAGC UAGAAUGAmGC UGAGAUGAAA-3′) was reacted with MgR protein, followed by TA cloning into *E. coli*. Twenty insert sequences from independent clones were sequenced by Sanger sequencing. All products ended 1 nt downstream of the Am site. **c**, Schematic workflow of NJU-seq. 1) total RNA fragmentation; 2) 3′ tail ligation; 3) RNase R from Mycoplasma genitalium (MgR) digestion; 4) NGS library preparation and high-throughput sequencing; 5) alignment and analysis. **d,** Plots showing read 3′ end counts on HeLa 18S rRNA with (MgR, bottom) and without (NC, top) MgR digestion. **e-g,** Plots showing read 3′ end counts of regions 157-169, 642-654, and 1828-1836 on 18S rRNA, without and with MgR digestion. Sites marked with red were previously reported with 2′-O-methylation, and blue were previously reported as m^6A or pseudouridine. **h,** Expression of *FBL* mRNA. *GAPDH* was used as an internal reference gene. *FBL*-KD mRNA expression was calculated by normalization to the HeLa-NC group. Error bars describe SEM (n = 3). ***P<0.001 (unpaired t test). **i**, **j,** The read end count of sites 1 nt downstream previously reported Nm in 18S (j) and 28S rRNA (i) of HeLa-NC and *FBL*-KD samples by NJU-seq (n = 3). The end count was normalized according to total aligned reads.

Figure 2

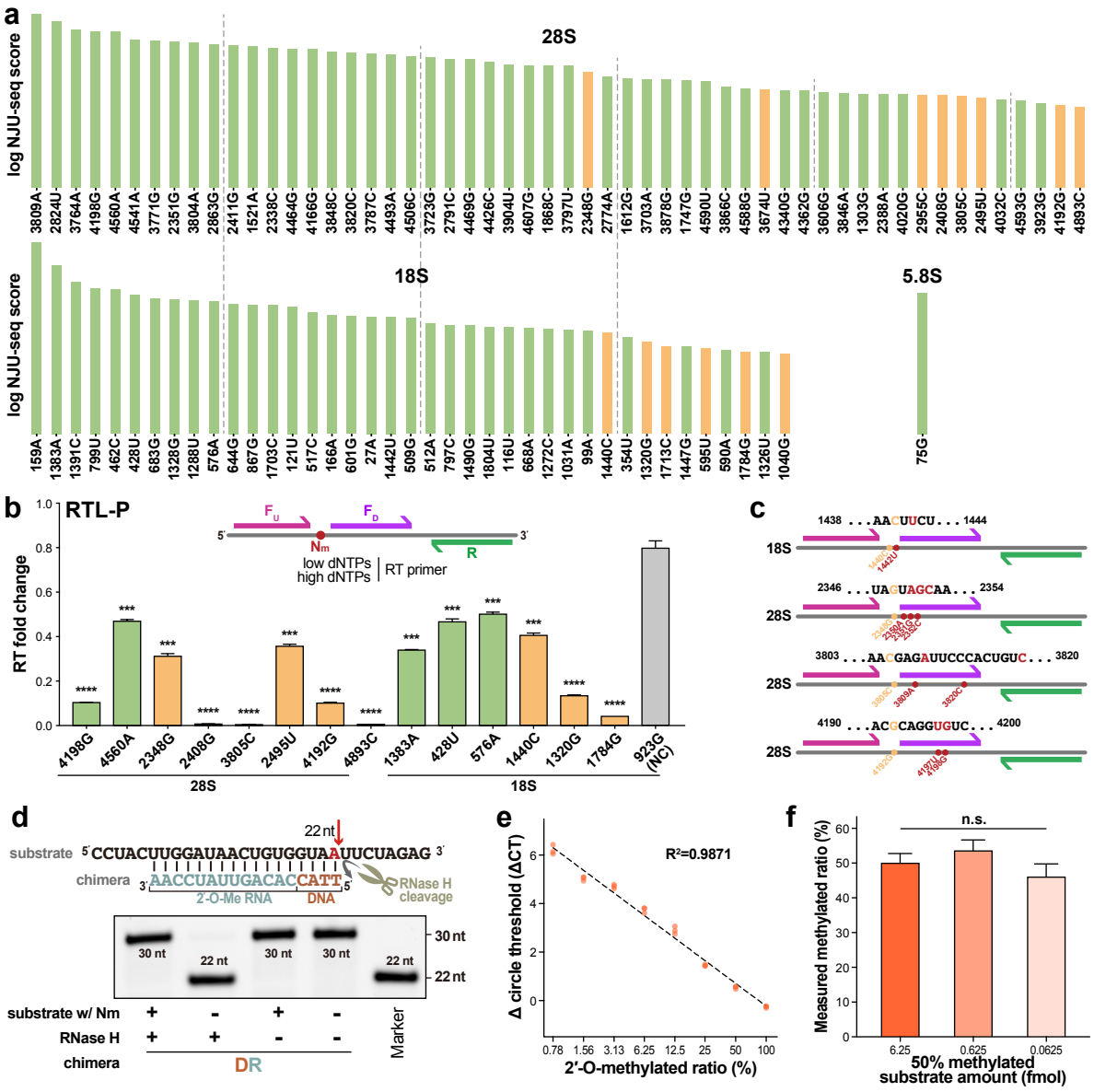

a
log NJU-seq score
28S
18S
5.8S
b
RTL-P
RT fold change
low dNTPs
high dNTPs
RT primer
c
d
substrate
chimera
22 nt
RNase H cleavage
2'-O-Me RNA
DNA
substrate w/ Nm
RNase H
chimera
DR
Marker
e
R²=0.9871
Δ circle threshold (ΔCT)
2'-O-methylated ratio (%)
f
n.s.
Measured methylated ratio (%)
50% methylated substrate amount (fmol)
g
Nm-VAQ
2'-O-methylated ratio (%)
h
HeLa-NC
FBL-KD

Figure 2 | Detection and validation of HeLa rRNA Nm sites by high-throughput and low-throughput tools. a, Nm sites detected from HeLa cell line's 28S, 18S and 5.8S rRNA by NJU-seq. Sites with scores > 30 and fold change of N+1 end counts > 3 were presented. Green indicated an Nm site in previous studies, and orange did not. **b,** Some Nm sites were validated by RTL-P. RT fold change was calculated as Methods detailed and negatively correlated with the modification ratio. Non-modification-reported 18S 923G serves as a negative control. All results were compared with 18S 923G using unpaired t test (***P<0.001, ****P<0.0001). **c,** RTL-P primer designs for detections of four Nm sites. F_U primers (rose) were located upstream of the target site, and F_D primers (purple) were located downstream. The target sites were shown as orange sites, and the surrounding Nm sites were marked red. **d-f**, The construction of Nm-VAQ. **d**, RNase H-dependent site-specific cleavage and qRT–PCR were combined to form Nm-VAQ. Scheme of hybrids of RNA substrates (up sequences) and chimera probes (down sequences) were shown. The Nm site in substrate was marked red, and the arrows indicated the cleavage sites. The RNA of the chimera was 2′-O-methylated, and the DNA of the chimera was marked reddish-brown. The RNase H reaction products were presented by electrophoresis (bottom). **e,** Correlation of the 2′-O-methylated ratio of substrate to ΔCT (cycle threshold) for the Nm-VAQ assay. A $6.25*10^{-2}$ pmol substrate with and without Nm was mixed to obtain a known Nm ratio, and ΔCT was obtained from the RNase H treated and control samples. **f,** Measurement of the Nm ratio of three different substrate concentrations with an Nm ratio of 50%. Nm ratio was calculated with ΔCT. n.s., not significant, by ordinary one-way ANOVA.

g, The Nm ratio of some NJU-seq-detected sites on HeLa rRNA by Nm-VAQ. **h,** Changes in Nm ratio between the *FBL*-KD group and HeLa-NC by Nm-VAQ. Data of NC cells was from Fig. **2g**. *P<0.05, **P<0.01, ****P<0.0001. Error bars in **b**, **e**-**h** described SEM for three technical replicates.

Figure 3

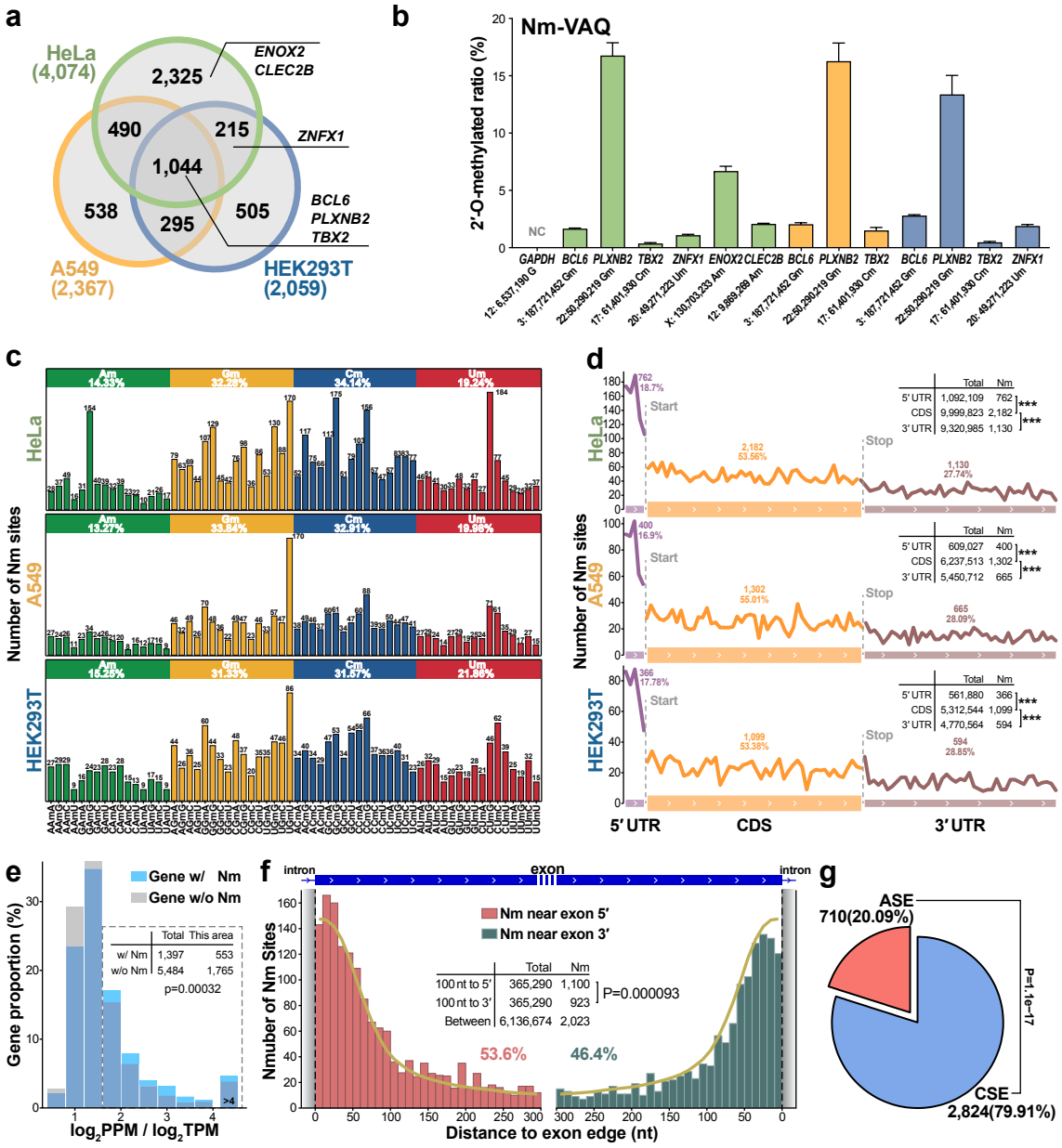
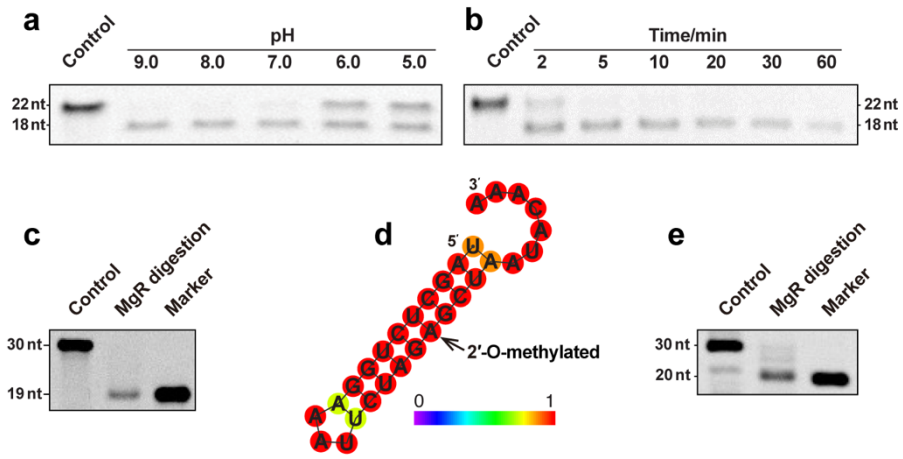
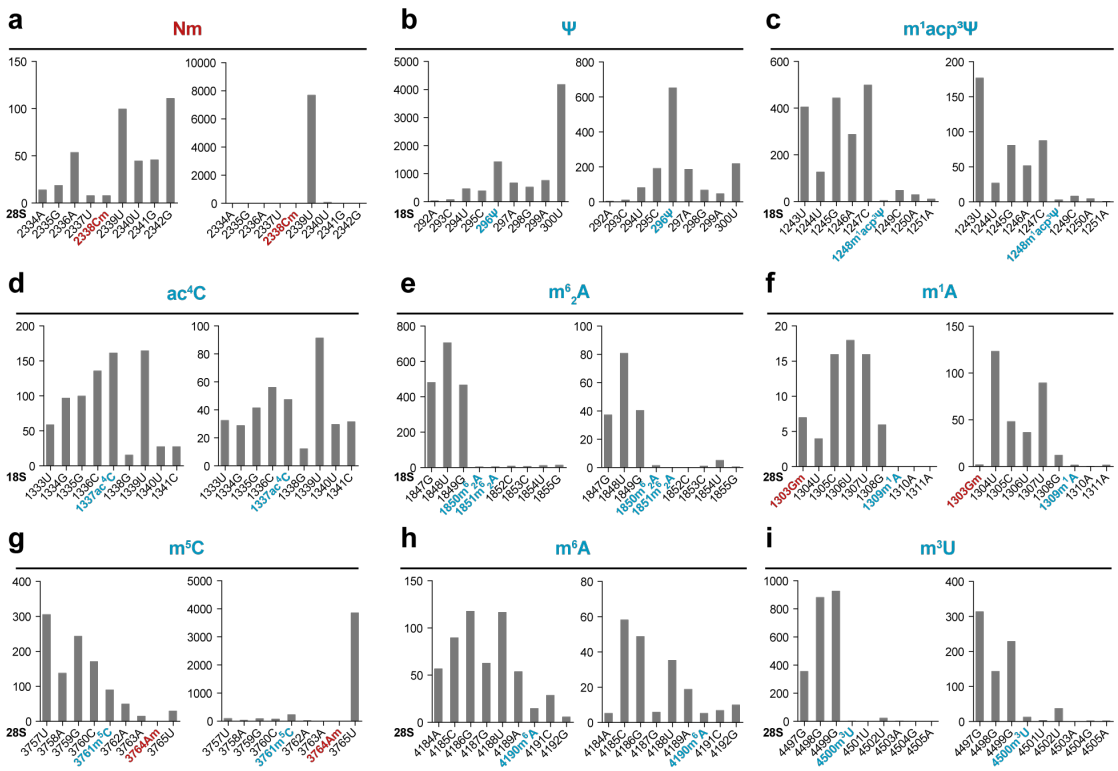

a
HeLa (4,074)
A549 (2,367)
HEK293T (2,059)
2,325
490
215
1,044
538
295
505
ENOX2
CLEC2B
ZNFX1
BCL6
PLXNB2
TBX2
b
Nm-VAQ
2'-O-methylated ratio (%)
NC
GAPDH BCL6 PLXNB2 TBX2 ZNFX1 ENOX2 CLEC2B BCL6 PLXNB2 TBX2 BCL6 PLXNB2 TBX2 ZNFX1
c
Number of Nm sites
HeLa
A549
HEK293T
Am 14.33%
Gm 32.28%
Cm 34.14%
Um 19.24%
Am 13.27%
Gm 33.84%
Cm 32.41%
Um 19.98%
Am 15.25%
Gm 31.33%
Cm 31.57%
Um 21.85%
d
Number of Nm sites
HeLa
A549
HEK293T
Start
Stop
762 18.7%
2,182 53.56%
1,130 27.74%
400 16.9%
1,302 55.01%
665 28.09%
366 17.78%
1,099 53.38%
594 28.85%
Total Nm
5' UTR 1,092,109 762
CDS 9,999,823 2,182
3' UTR 9,320,985 1,130
5' UTR 609,027 400
CDS 6,237,513 1,302
3' UTR 5,450,712 665
5' UTR 561,880 366
CDS 5,312,544 1,099
3' UTR 4,770,564 594
5' UTR
CDS
3' UTR
e
Gene proportion (%)
Gene w/ Nm
Gene w/o Nm
Total This area
w/ Nm 1,397 553
w/o Nm 5,484 1,765
p=0.00032
log2PPM / log2TPM
f
intron
exon
intron
Nm near exon 5'
Nm near exon 3'
Nmuber of Nm Sites
Total Nm
100 nt to 5' 365,290 1,100
100 nt to 3' 365,290 923
Between 6,136,674 2,023
P=0.000093
53.6%
46.4%
Distance to exon edge (nt)
g
ASE 710(20.09%)
CSE 2,824(79.91%)
P=1.1e-17

Figure 3 | mRNA Nm sites in three human cell lines. a, Overlap of mRNA Nm sites among the three cell lines by NJU-seq. **b,** The methylation status in mRNA Nm sites detected by Nm-VAQ. The Nm sites of HeLa, HEK293T, and A549 cells were presented as green, blue, and orange sites, respectively. *GAPDH* (12:6,537,190, G) was not identified or previously reported as an Nm site and thus served as a negative control. **c**, Methylated nucleotide frequency distribution of HeLa mRNA. Nm-1, Nm and Nm+1 sites were considered tri-ribonucleotide units. **d,** Distribution of HeLa Nm sites on a length-normalized mRNA transcript (see Methods). ***P < 0.001 (Pearson's chi-square test). **e**, Translation efficiency of genes with and without Nm site(s) in HeLa. The ratio of protein abundance and RNA expression (log2PPM/log2TPM) represented translation efficiency. *P < 0.05 (Pearson's Chi-square test). **f**, Distribution of Nm sites near the exon 5′ (left) and exon 3′ (right) boundaries. All Nm sites were counted and divided into these two parts. The yellow dash represents the simulation of an even distribution. **g**, Distribution of HeLa Nm sites in alternatively spliced exon (ASE) and constitutively spliced exon (CSE) regions was shown. ***P < 0.001 (Pearson's chi-square test).

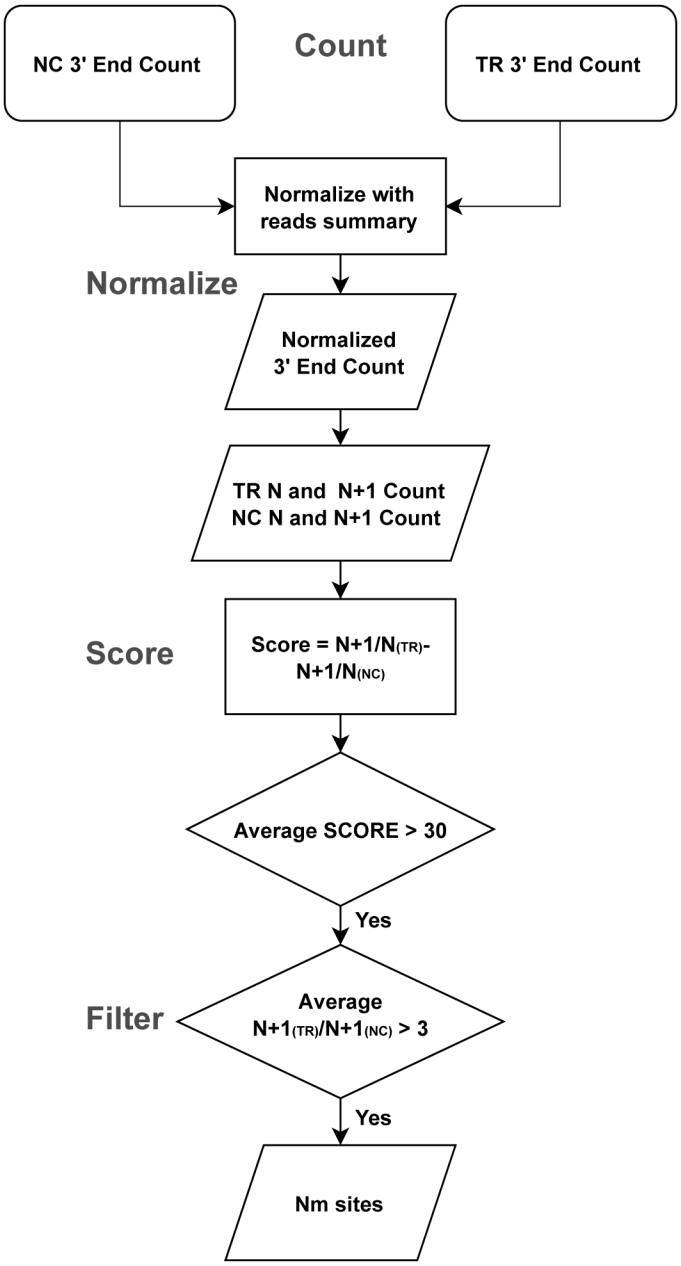


Supplementary Figure 1 | MgR recognizes 2′-O-methylation modifications. a, b, Exploration of MgR hydrolysis conditions in vitro. The ssRNA substrate (5′-6-Fam-UAACC UAUGA AGNNN UNmNNC UC-3′, N represents the mix of A/G/C/U) was reacted with a series of different conditions as shown. **a**, pH ranges from 5.0 to 9.0; **b,** Reaction time ranges from 0-60 min; **c**, Recognition specificity of MgR for Nm modifications. MgR digested the ssRNA substrate (5′-6-Fam-AUACC GCAGC UAGAA UGAmGC Um^{5}CAGA UGm^{6}AAA). The marker band responded to the Nm+1 site. **d**, Testing of MgR to hydrolyze complex RNA structure. RNAFold predicted the substrate (5′-6-Fam-UAGCUCUGGA AAUUCUAGAmG CUAAUACAAA-3′) structure. **e**, Products of the substrate reacted by MgR were shown.

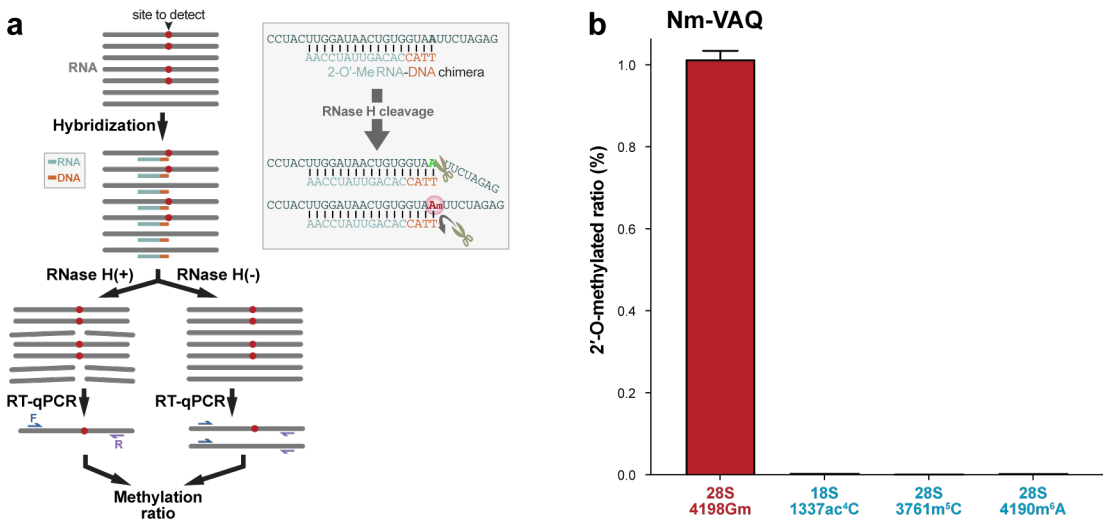


Supplementary Figure 2 | Specificity of NJU-seq for 2′-O-methylation. The left panels illustrated NGS fragments from the NC group (without MgR), and the MgR group was shown on the right. Red and blue marked the presence of different RNA modifications in previous reports. **a,** Position 2334 to 2342 of HeLa 28S rRNA with 2338Cm. **b-i,** other kinds of RNA modifications on HeLa rRNA. **b,** Position 292 to 300 of HeLa 18S rRNA with 296Ψ. **c,** Position 1243 to 1251 of HeLa 18S rRNA with 1248m^{1}acp^{3}Ψ. **d,** Position 1333 to 1341 of HeLa 18S rRNA with 1337ac^{4}C. **e,** Position 1847 to 1855 of HeLa 18S rRNA with 1850m$^6{}_2$A and 1851m$^6{}_2$A. **f,** Position 1303 to 1311 of HeLa 28S rRNA with 1309m^{1}A. **g,** Position 3757 to 3765 of HeLa 28S rRNA

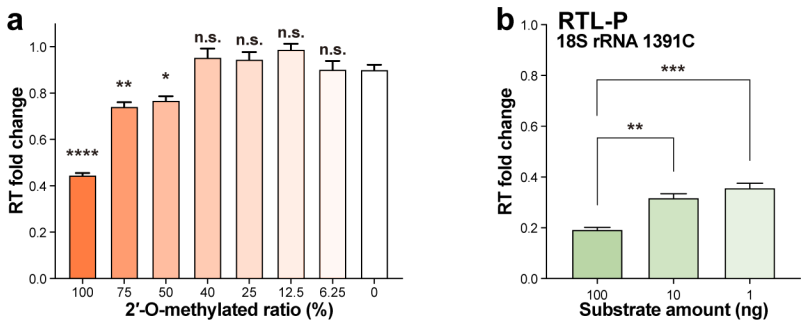
with 3761m^{5}C. **h,** Position 4186 to 4194 of HeLa 28S rRNA with 4190m^{6}A. **i,** Position 4497 to 4505 of HeLa 28S rRNA with 4500m^{3}U.



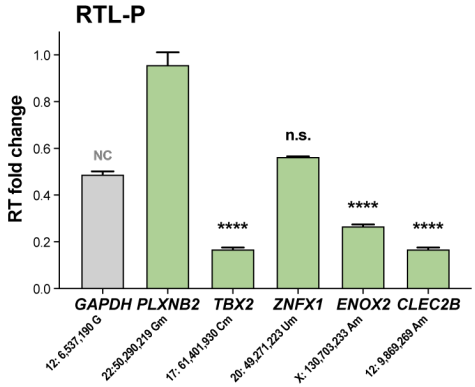
Supplementary Figure 3 | Scoring and filtering process for NJU-seq. According to this flow, Nm sites would be detected from the reads aligned to the reference sequence.



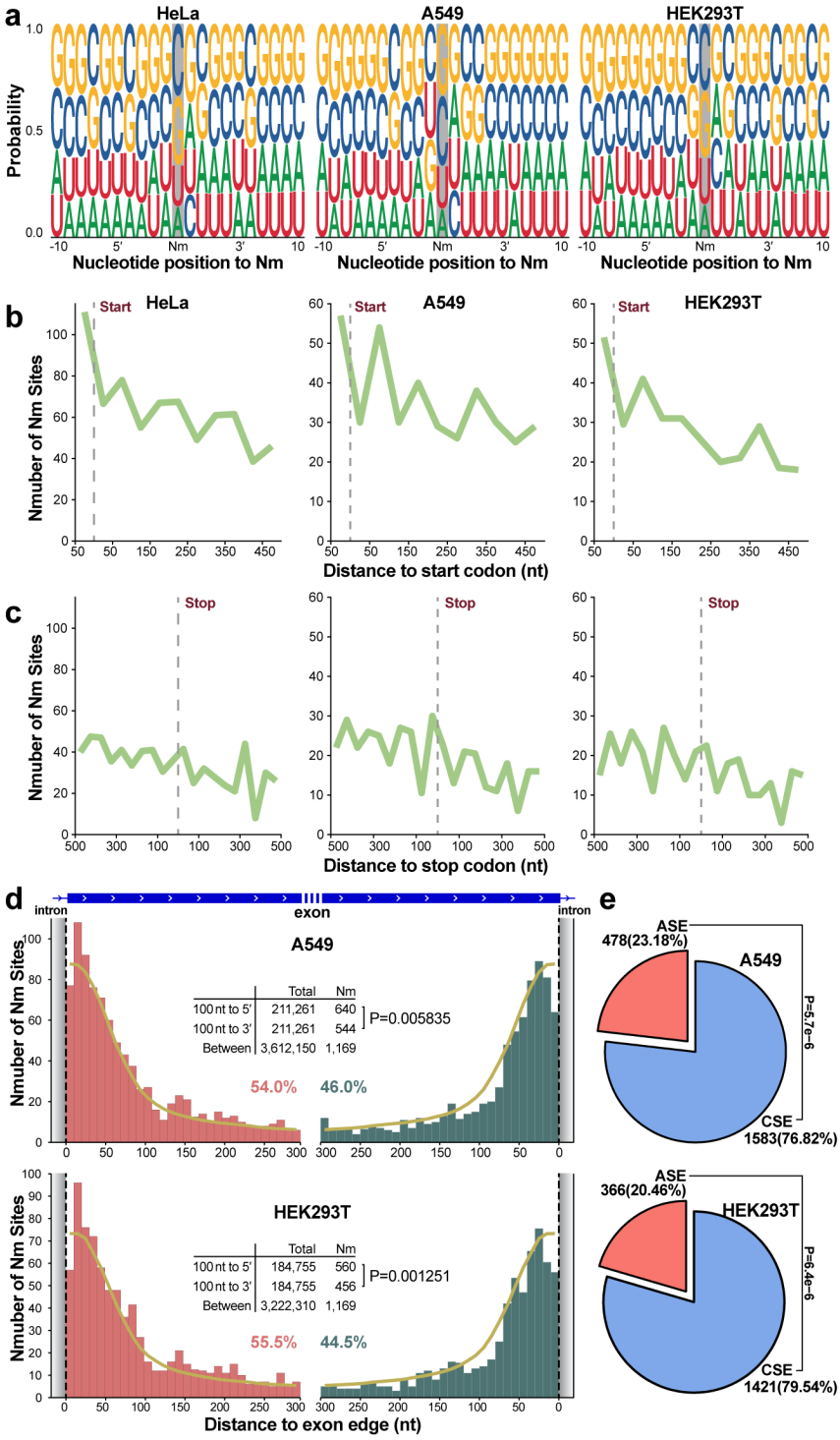
Supplementary Figure 4 | Schematic workflow of Nm-VAQ. a, 1) the hybrid of RNA and chimera probes. In the substrate, the red site indicated the target site was 2′-O-methylated, and the green site indicated the unmethylated site. In chimera probe, the reddish-brown region showed the DNA, and the light green region showed the 2′-O-methylated RNA; 2) with/without RNase H cleavage; 3) RT-qPCR. Methylation ratio was from ΔCT (cycle threshold) of RNase H cleavage and control sample. The box represented a real example. **b**, Detection of Nm-VAQ on other common RNA modifications. 28S 4198, 4190, 3761, and 18S 1337 were reported as Gm, m^6A, m^5C, and ac^4C in previous studies. 28S 4198Cm′ data was from **Fig. 2g**.



Supplementary Figure 5 | RTL-P detected known 2′-O-methylation modification ratio. **a**, $6.25*10^{-2}$ pmol synthesized substrate with and without Nm were mixed to obtain a known Nm ratio. Seeing Methods for calculating RT fold change, negative with Nm ratio. All results were compared with a 0% Nm ratio (negative control). *P ≤ 0.05, **P ≤ 0.01, ****P ≤ 0.0001, n.s., not significant, one-way ANOVA. **b**, The RT fold change of 18S 1391Cm site among three total RNA amounts detected by RTL-P. **P ≤ 0.01, ***P ≤ 0.001 by ordinary one-way ANOVA. Error bars in supplementary Fig. 6a, b describe SEM for three technical replicates.

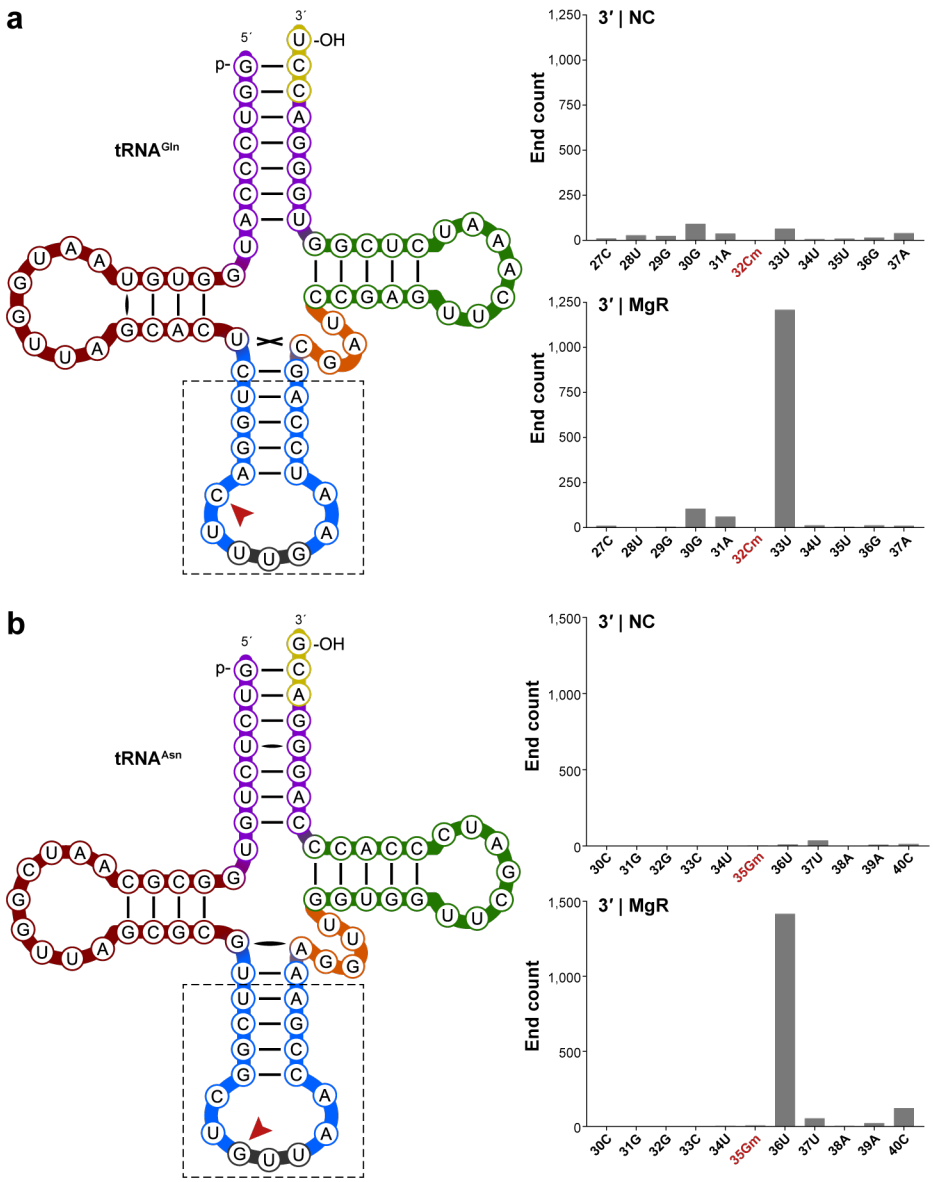


Supplementary Figure 6 | HeLa mRNA Nm status detected by RTL-P. These five sites were predicted as Nm sites by NJU-seq. *GAPDH* (12:6,537,190, G) was used as a negative control. Error bars describe SEM for three technical replicates.



Supplementary Figure 7 | mRNA Nm site distribution in HeLa, A549 and HEK293T cells. a, Sequence probability logo for mRNA Nm sites. Nm sites were marked in gray background. **b**, **c**, Distribution of three cell lines' Nm sites in windows

with absolute length centered on the start (**b**) and stop (**c**) codon. **d**, Distribution of Nm sites near exon 5′ (left) and exon 3′ (right) boundaries. **e**, Distribution of A549 and HEK293T Nm sites in ASE and CSE regions was shown. ***$P < 0.001$ (Pearson's chi-square test).



Supplementary Figure 8 | Identification of transfer RNA (tRNA) Nm sites using NJU-seq. (a, b) Plot of read 3′ end counts at 32Cm of tRNA-Gln-UUG (a) and 35Gm of tRNA-Asn-GmUU (b) without and with MgR digestion.