Limits in the detection of m^{6}A changes using MeRIP/m^{6}A-seq

Alexa B.R. McIntyre[1,2]*, Nandan S. Gokhale[3], Leandro Cerchietti[4], Samie R. Jaffrey[5], Stacy M. Horner[3,6]*, Christopher E. Mason[1,7,8,9]*

[1]Department of Physiology and Biophysics, Weill Cornell Medicine, New York City, NY 10065

[2]Tri-Institutional Program in Computational Biology and Medicine, New York City, NY 10065

[3]Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710

[4]Division of Hematology and Medical Oncology. Weill Cornell Medicine, New York City, NY 10065

[5]Department of Pharmacology, Weill Cornell Medicine, New York City, NY 10065

[6]Department of Medicine, Duke University Medical Center, Durham, NC 27710

[7]The HRH Prince Alwaleed Bin Talal Bin Abdulaziz Alsaud Institute for Computational Biomedicine, Weill Cornell Medicine, New York, NY 10021

[8]The Feil Family Brain and Mind Research Institute, Weill Cornell Medicine, New York, NY 10065

[9]The WorldQuant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY 10021

*corresponding authors:

Alexa B.R. McIntyre, abm237@cornell.edu

Stacy M. Horner, stacy.horner@duke.edu

Christopher E. Mason chm2042@med.cornell.edu

Abstract

Many cellular mRNAs contain the modified base m^{6}A, and recent studies have suggested that various stimuli can lead to changes in m^{6}A. The most common method to map m^{6}A and to predict changes in m^{6}A between conditions is methylated RNA immunoprecipitation sequencing (MeRIP-seq), through which methylated regions are detected as peaks in transcript coverage from immunoprecipitated RNA relative to input RNA. Here, we generated replicate controls and reanalyzed published MeRIP-seq data to estimate reproducibility across experiments. We found that m^{6}A peak overlap in mRNAs varies from ~30 to 60% between studies, even in the same cell type. We then assessed statistical methods to detect changes in m^{6}A peaks as distinct from changes in gene expression. However, from these published data sets, we detected few changes under most conditions and were unable to detect consistent changes across studies of similar stimuli. Overall, our work identifies limits to MeRIP-seq reproducibility in the detection both of peaks and of peak changes and proposes improved approaches for analysis of peak changes.

Keywords

RNA base modifications, *N*6-methyladenosine, MeRIP-seq, m^{6}A-seq, RNA immunoprecipitation sequencing, statistical methods, bioinformatics

Introduction

Methylation at the *N*6 position in adenosine (m^{6}A) is the most common internal modification in eukaryotic mRNA. A methyltransferase complex composed of METTL3, METTL14, WTAP, VIRMA, and other cofactors catalyzes methylation at DRACH/DRAC motifs, primarily in the last exon (1,2). Most m^{6}A methylation occurs during transcription (3). The modification then affects mRNA metabolism through recognition by RNA-binding proteins that regulate processes including translation and mRNA degradation (4–9). However, whether m^{6}A is lost and gained in response to various cellular changes remains contentious (3,10–15). To assess the evidence for proposed dynamic changes in m^{6}A, a reliable and reproducible method to detect changes in methylation as distinct from changes in gene expression is necessary.

55 The first and most widely-used method to enable transcriptome-wide studies of m^6A, MeRIP-seq
56 or m^6A-seq, involves the immunoprecipitation of m^6A-modified RNA fragments followed by peak detection
57 through comparison to background gene coverage (16,17). A second method was developed in 2015,
58 miCLIP or m^6A-CLIP, which involves crosslinking at the site of antibody binding to induce mutations
59 during reverse transcription for single-nucleotide detection of methylated bases (2,18). MeRIP-seq is still
60 more often used than miCLIP, despite less precise localization of m^6A to peak regions of approximately
61 50-200 base pairs that can contain multiple DRAC motifs, since it follows a simpler protocol, requires less
62 starting material, and generally produces higher coverage of more transcripts. Antibodies for m^6A can
63 also detect a second base modification, N^6,2′-*O*-dimethyladenosine (m^6A_m), found at a lower abundance
64 than m^6A and located at the 5′ ends of select transcripts (15,18). We thus refer to the base modifications
65 detected through MeRIP-seq collectively as $m^6A_{(m)}$, although most are likely m^6A. As of late 2018, over
66 fifty studies used MeRIP-seq to detect $m^6A_{(m)}$ in mammalian mRNA (**Additional File 1: Supplementary**
67 **Table 1**).
68 Although MeRIP-seq can reveal approximate sites of $m^6A_{(m)}$, it cannot be used to quantitatively
69 measure the fraction of transcript copies that are methylated (19). Studies of m^6A variation in response to
70 stimuli instead estimate differences at individual loci through changes in peak presence or peak height.
71 Using these approaches, studies have reported changes to m^6A with heat shock, microRNA expression,
72 transcription factor expression, cancer, oxidative stress, human immunodeficiency virus (HIV) infection,
73 Kaposi's sarcoma herpesvirus (KSHV) infection, and Zika virus infection, including hundreds to thousands
74 of changes in enrichment at specific sites (20–29). Statistical approaches to analysis have only recently
75 been published and there have been no comprehensive evaluations of methods to detect changes in m^6A
76 based on MeRIP-seq data (30,31). Thus, while studies have suggested that m^6A shows widespread
77 changes in response to diverse stimuli, they have applied inconsistent analysis methods to detect
78 changes in m^6A and often don't control for differences in RNA expression between conditions or typical
79 variability in peak heights between replicates. In some cases, these studies have reported m^6A changes
80 based on simple differences in peak count (24,26,27,32). However, others have applied statistical tests or
81 thresholds for differences in immunoprecipitated (IP) over input fraction enrichment and visual analysis of

82 coverage plots, and have reported fewer m^6A changes or suggested that m^6A is a relatively stable mark
83 (33,34). As in RNA-seq, there is noise in MeRIP-seq, and multiple replicates are therefore necessary to
84 estimate variance and statistically identify the effects of experimental intervention (35–37). To date, only
85 one MeRIP-seq study has used more than three replicates per condition (34), while ten have used only
86 one (17,20,32,33,38–43), suggesting that most studies may not have enough power to detect changes in
87 $m^6A_{(m)}$.

88 To re-evaluate the evidence for $m^6A_{(m)}$ changes under various conditions, we first examined the
89 variability in $m^6A_{(m)}$ detection across replicates, cell lines, and experiments using our own negative
90 controls (12 replicates) as well as 24 published MeRIP-seq data sets. We then compared statistical
91 methods to detect differences in IP enrichment using biological negative and positive controls for m^6A
92 changes. We found that these methods are limited by noise, including biological variability from changes
93 in RNA expression and technical variability from immunoprecipitation and sequencing that limits
94 reproducibility across studies. Our results suggest that the scale of statistically detectable $m^6A_{(m)}$ changes
95 in response to various stimuli is orders of magnitude lower than the scale of changes reported in many
96 studies. However, we also found that statistical detection could miss the majority of changed sites when
97 using only 2-3 replicates. We use our results to propose approaches to MeRIP-seq experimental design
98 and analysis to improve reproducibility and more accurately measure differential regulation of $m^6A_{(m)}$ in
99 response to stimuli. These data and analyses emphasize the need for further research and alternative
100 assays, for example recently developed endoribonuclease-based sequencing methods (44,45) or direct
101 RNA nanopore sequencing (46), to resolve the extent to which m^6A changes in response to specific
102 conditions.

103

104 **Results**

105 *Detection of peaks across replicates, experiments, and cell types*

106 The first steps in MeRIP-seq data analysis are to align sequencing reads to the genome or
107 transcriptome of origin and to identify peaks in transcript coverage in the IP fraction relative to the input
108 control. Several methods have been developed for MeRIP-seq peak detection, including exomePeak,

109 MeTPeak, MeTDiff, and bespoke scripts. Another method often used for MeRIP-seq peak detection is
110 MACS2, which was originally designed to detect protein binding sites in DNA from chromatin
111 immunoprecipitation sequencing (ChIP-seq). We compared $m^6A_{(m)}$ peak detection by exomePeak,
112 MeTPeak, MeTDiff, and MACS2 (31,47–49) in seven replicates of MeRIP-seq data obtained from mouse
113 cortices under basal conditions (34), and in 12 replicates of MeRIP-seq data we generated from human
114 liver Huh7 cells (50). The intersect between all tools tested was high, and we saw minimal differences in
115 DRAC motif enrichment, which we use to provide an estimate of tool precision in the absence of true
116 positive m^6A sites (**Additional File 2: Supplementary Figure 1a**). In addition, we assessed the
117 METTL3/METTL14-dependence of specific peaks identified by single tools using MeRIP-RT-qPCR. We
118 found that of these peaks, 4/4 from MACS2, 5/5 from MeTPeak, and 4/5 from MeTDiff showed decreased
119 $m^6A_{(m)}$ enrichment following METTL3/METTL14 depletion, suggesting that these are true m^6A sites. By
120 comparison, only 1/5 of the peaks uniquely called by exomePeak showed statistically significant
121 decreases ($p < 0.05$), although replicate variance was high and 4/5 showed a downward trend
122 (**Additional File 2: Supplementary Figure 1b**). Since MACS2 was the most commonly used tool for
123 peak calling and was previously found to perform well in comparison with a graphical user interface tool
124 and several other peak callers (51), we used MACS2 for the remainder of our analyses. Repeating the
125 analyses shown in Figures 2-4 using the MeTDiff peak caller instead of MACS2 did not affect any of our
126 conclusions (**Additional File 3**).

127 For $m^6A_{(m)}$ peak detection, a transcript must be sufficiently expressed for enrichment by the
128 $m^6A_{(m)}$ antibody and for adequate sequencing coverage in both the IP and input fractions. Previous
129 reports have suggested that $m^6A_{(m)}$ presence does not decrease with lower mRNA expression level, and,
130 if anything, is higher in mRNAs with lower expression as methylated transcripts tend to be less stable
131 (9,38). Peak callers, however, identify fewer peaks in genes at low expression, which we therefore
132 assume reflects inadequate coverage for peak calling. To estimate the level of coverage necessary for
133 peak detection, we analyzed the percent of genes with at least one, two, or three peaks relative to mean
134 input transcript coverage in both the mouse cortex and Huh7 cell data (**Figure 1a**). Based on the upper
135 shoulders of the sigmoidal curves as the percent of genes with peaks begins to plateau, we estimate that

mean gene coverage of approximately 10-50X is necessary to avoid missing peaks based on insufficient coverage. Including a wider array of samples in this analysis likewise showed an increase in the percent of transcripts with ≥1 peak as coverage rose to 10X (**Additional File 2: Supplementary Figure 1c**). Our analysis of the input RNA-seq coverage of peak regions alone again supported a similar threshold; few peaks are detected with median input read counts below 10 across replicates (**Additional File 2: Supplementary Figure 1d**). These thresholds do not mean that peaks in genes with mean coverage < 10X or peaks with fewer than 10 input reads are false positives, but that the likelihood of false negatives rises with lower coverage (**Additional File 2: Supplementary Figure 1e**).

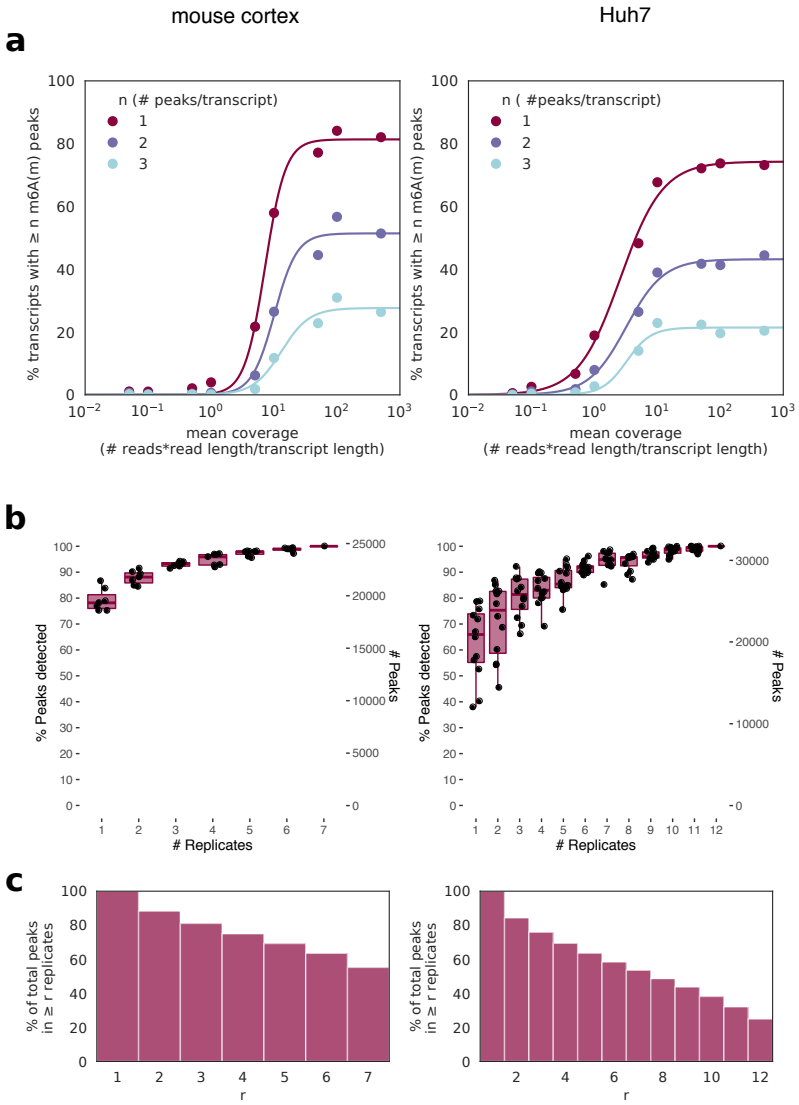


Figure 1: Thresholds for peak detection. a) $m^6A_{(m)}$ site detection in MeRIP-seq data from mouse cortex (left) and human liver cells (Huh7, right) shows saturation of peak detection as transcript coverage approaches 10-50X for replicates at basal conditions, with peaks merged from all replicates. **b)** The total number of peaks captured increases with more replicates, with single replicates capturing a median of 66-78% of total peaks depending on study. Boxes span the 1st to 3rd quartiles of distributions for random subsamples of replicates, with lines indicating the median number of peaks, and whiskers showing the minimum and maximum points within ±1.5x the interquartile distance from the boxes. Jittered points show results for each random subsample (a total of 6 subsamples per replicate number for the mouse cortex data and 12 for the Huh7 data). **c)** The percent of peaks detected in at least r replicates for the same data sets.

145 To evaluate the reproducibility of MeRIP-seq data, we next examined the consistency of $m^6A_{(m)}$
146 peak calling between replicates. Previous studies have reported that peak overlap between replicates is
147 approximately 80% (9,16,52,53). Similarly, we found that between two replicates, $\log_2$ fold enrichment of
148 IP over input reads at detected peaks showed a Pearson correlation of approximately 0.81 to 0.86
149 (**Additional File 2: Supplementary Figure 1f**). A single sample captured a median of 78% of the peaks
150 found in seven replicates of mouse cortex data and 66% of peaks found in twelve replicates of Huh7 cell
151 data. The number of detected peaks increased log-linearly with the addition of more replicates, such that
152 with three replicates, 84-92% of the peaks found with 7-12 replicates were detected (**Figure 1b**).
153 Conversely, the number of peaks in common across replicates decreased as the number of replicates
154 increased, such that while ~80% of peaks were detected in at least two replicates, only ~60% were
155 detected in six replicates for both data sets and ~25% in all twelve replicates of Huh7 cell data (**Figure**
156 **1c**). Detection of peaks in more replicates did not increase DRAC motif enrichment (**Additional File 2:**
157 **Supplementary Figure 1g**). These results suggest that many $m^6A_{(m)}$ sites may be missed in studies that
158 use one to three replicates, and that increasing replicates could enable detection of more peaks.
159 However, not all peaks correspond to true $m^6A_{(m)}$ sites. A recent comparison to data from an
160 endoribonuclease-based method for m^6A detection suggested MeRIP-seq has a false positive rate of
161 ~11%, although this would differ by study and detection threshold (3,54).
162 The number of peaks detected across studies varies. Given that coverage affects peak detection,
163 we hypothesized that variation in sequencing depth could contribute to differences in peak count. Zeng et
164 al. (2018) reported that peak count begins to saturate by around 20 million reads by subsampling data
165 within individual studies (42). However, we found that there is no positive correlation between peak count
166 and input or IP sequencing depth across data sets from different published studies, each of which had 3-
167 81M reads per replicate (input Pearson's R = -0.37, p = 0.015; IP Pearson's R = -0.17, p = 0.28)
168 (**Additional File 1: Supplementary Table 2, Additional File 2: Supplementary Figure 2a-b**). This
169 implies that experimental factors beyond sequencing depth contribute to the variability of peak counts
170 across studies.

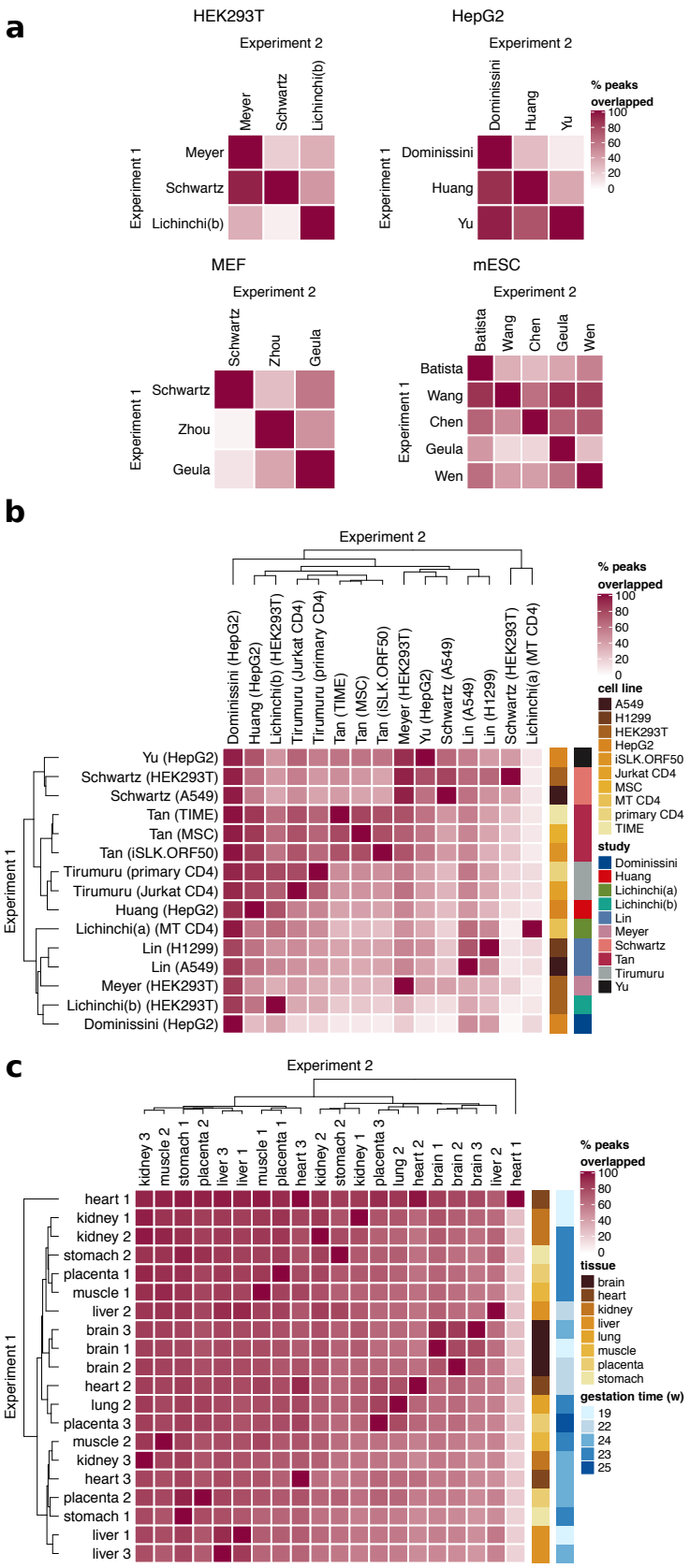


Figure 2: Reproducibility of peak detection. a) Peak detection between studies that used the same cell type shows variable overlap. Overlap was calculated as the percent of peaks detected in Experiment 1 with an overlap of ≥ 1 base pair with peaks from Experiment 2. **b)** Peak detection across tissue and cell types shows samples from the same study cluster better together than samples from the same tissue. Median overlap was 46%. **c)** Peak detection across tissue types for data from the same study (Xiao et al., 2019). Median overlap was 72%. Studies used in (a) and (b) are described in **Additional File 1: Supplementary Table 2**.

171

172 We next analyzed the overlap of peaks among studies and found inconsistency in peak
173 localization on transcripts as well. Within four commonly used cell types, the percent of peaks detected in
174 one experiment that were also detected in a second varied among pairs of studies from as low as 2% of
175 peaks to as high as 90% (median = 45%), after filtering for transcripts expressed above a mean of 10X
176 input coverage in both to ensure sufficient expression for peak detection (**Figure 2a**). In fact, peaks
177 showed higher overlap within different cell types from the same study than within the same cell type from
178 different studies, suggesting that MeRIP-seq data is prone to strong batch effects (**Figure 2b**). While this
179 could be due to differences among experimental protocols used (summarized in **Additional File 1:**
180 **Supplementary Table 2**), we were unable to identify such a link. Overall, most percent overlaps of $m^6A_{(m)}$
181 peaks fell between ~30% (1st quartile) and ~60% (3rd quartile) (**Figure 2b**). With rare exceptions (e.g. that
182 described by Ke et al., 2017 in their Supplementary Figure 8 (3)), most MeRIP-seq data sets do show
183 enrichment of the m^6A motif DRAC. These results indicate, however, that multiple labs running MeRIP-
184 seq on the same cell type will detect different subsets of $m^6A_{(m)}$ sites. Possible contributing factors in the
185 differences among studies include cell state (e.g. different stages of the cell cycle), experimental
186 conditions, and sequencing depth. Despite predictions that tissue or cell type would be a large factor in
187 differences among samples, though, peaks detected in different tissues analyzed in a single experiment
188 showed high overlap and little clustering by tissue type (**Figure 2c**) (55). This suggests that although
189 there is evidence that m^6A levels vary by tissue (19), modified sites are consistent.
190

191 *Detection of changes in peaks between conditions*

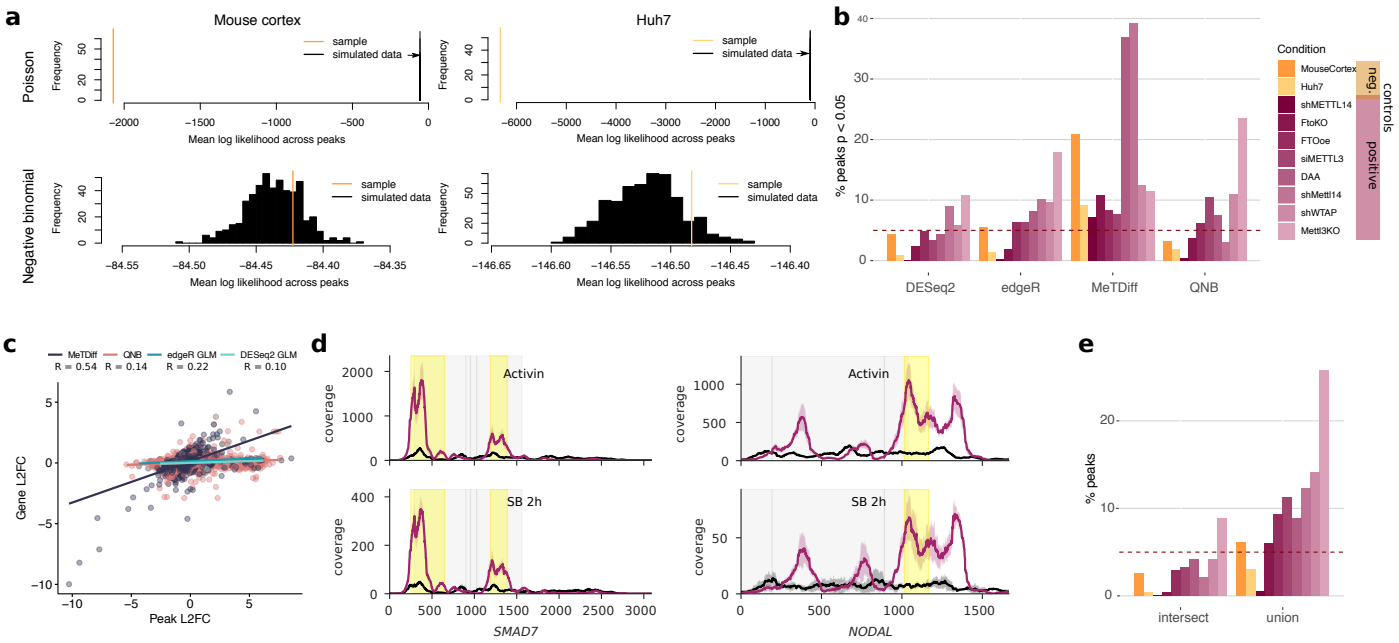
192 Following $m^6A_{(m)}$ peak detection, many studies compare the expression of peaks between two
193 conditions to predict peak changes. While looking at plots of IP and input gene coverage under different
194 conditions can help evaluate the evidence for these changes (33), statistical or heuristic methods are first
195 necessary to narrow down a list of candidate sites to plot. Several tools used for statistical analysis by
196 the studies in **Additional File 1: Supplementary Table 1** or for other types of RNA IP sequencing assays
197 model peak counts using either (a) the Poisson distribution, in which the variance of a measure (here,
198 read counts) is assumed to be equal to the mean (MeTDiff), or (b) the negative binomial distribution, in

199 which a second parameter allows for independent adjustment of mean and variance (QNB and two
200 implementations of a generalized linear model approach using DESeq2 or edgeR, **Table 1**) (30,31,56–
201 58). In the mouse cortex and Huh7 cell data, we found that, similar to RNA-seq data (24,57,59), the
202 variance in read counts under peaks exceeded their mean, indicative of overdispersion (**Additional File**
203 **2: Supplementary Figure 3a**). The log likelihood (the probability of an observation given a distribution
204 with known parameters) for our sample also fell within the distribution of expected log likelihoods for the
205 negative binomial distribution (bottom) but not the Poisson distribution (top) (**Figure 3a**). Thus, the
206 negative binomial distribution captures the mean-variance relationship in MeRIP-seq data, suggesting
207 that tools that account for overdispersion better model the distribution of read counts at $m^6A_{(m)}$ peaks than
208 tools that do not.
209
210 **Table 1:** Statistical methods for the detection of peak changes

Method	**Read count distribution**	**Publication**
MeTDiff	Poisson	Cui et al. (2018)
Quad-negative binomial (QNB)	Negative binomial	Liu et al. (2017)
GLM (DESeq2)	Negative binomial	based on Park et al. (2014) method for HITS-CLIP
GLM (edgeR)	Negative binomial	

211
212 We next defined positive and negative controls to evaluate tool performance for detection of
213 changes in $m^6A_{(m)}$ peaks. Past publications describing new methods to detect $m^6A_{(m)}$ peak changes have
214 used data sets in which methylation machinery genes or the methyl donor were disrupted compared to
215 baseline conditions as positive controls, and have simulated negative controls by randomly swapping
216 labels in the positive controls (30,31). However, swapping labels for conditions that may feature
217 differences in gene expression in addition to differences in m^6A levels could unrealistically increase
218 variance in read counts within groups. Therefore, we instead used the two data sets from mouse cortex
219 and Huh7 cells, which each comprised many replicates at baseline conditions (n=7 and n=12,

respectively), as negative controls. We randomly divided the mouse cortex data into two groups of three to four replicates for comparison and divided the Huh7 replicates by lab of incubation, which did not affect sample clustering (**Additional File 2: Supplementary Figure 3b**). We would expect to see minimal changes in IP enrichment at m^6A peaks between groups for our negative controls, whereas our positive controls, which featured genetic or chemical interference with the m^6A machinery, should show discernible differences in peaks when compared to baseline or wildtype conditions in the same cell lines (summarized in **Additional File 1: Supplementary Table 3**). Indeed, the absolute difference in $\log_2$ fold change between peaks and genes was centered around 0 for the negative controls and showed small shifts that varied in magnitude and direction for the positive controls (**Additional File 2: Supplementary Figure 3c**).



230

Figure 3: Analysis of methods to detect peak changes disproportional to gene expression changes. a) A comparison of Poisson (above) and negative binomial (below) models for read counts under peaks. The negative binomial mean log likelihood of the sample data fell within the 74th and 89st percentiles of 500 simulations for mouse cortex and Huh7 cell data, respectively, while the Poisson model failed to capture the sample distributions. **b)** The percent of sites below an unadjusted p-value threshold of 0.05 for different methods (described in **Table 1**) to detect differential methylation in negative controls between two groups at baseline conditions and positive controls in which methylation processes were disrupted with respect to baseline conditions (**Additional File 1: Supplementary Table 3**). The line at 5% indicates the expected proportion of sites given a uniform p-value distribution (see **Additional File 2: Supplementary Figure 2c**), while colours indicate negative (orange) and positive (purple) control experiments. **c)** The correlation between change in gene expression and change in peak expression between conditions for sites identified as differentially methylated in the eight positive control experiments. Pearson's R = 0.22, 0.10, 0.55, and 0.14 for edgeR, DESeq2, MeTDiff, and QNB, respectively, with p = 0.05, 0.09, 5.8E-87, and 2.4E-11. **d)** Coverage plots showing changes in peak expression are proportional to changes in gene expression for genes identified as differentially methylated by Bertero et al. (2018) using MeTDiff with activin signaling and an activin-NODAL inhibitor, SB431542 (SB). Lines show the mean coverage across three replicates, while shading shows the standard deviation. Peaks detected as significantly changed are highlighted in yellow. Coding sequences are shown in grey. **e)** The intersect and union of peaks with $p < 0.05$ from DESeq2, edgeR, and QNB from (b), coloured as in (b).

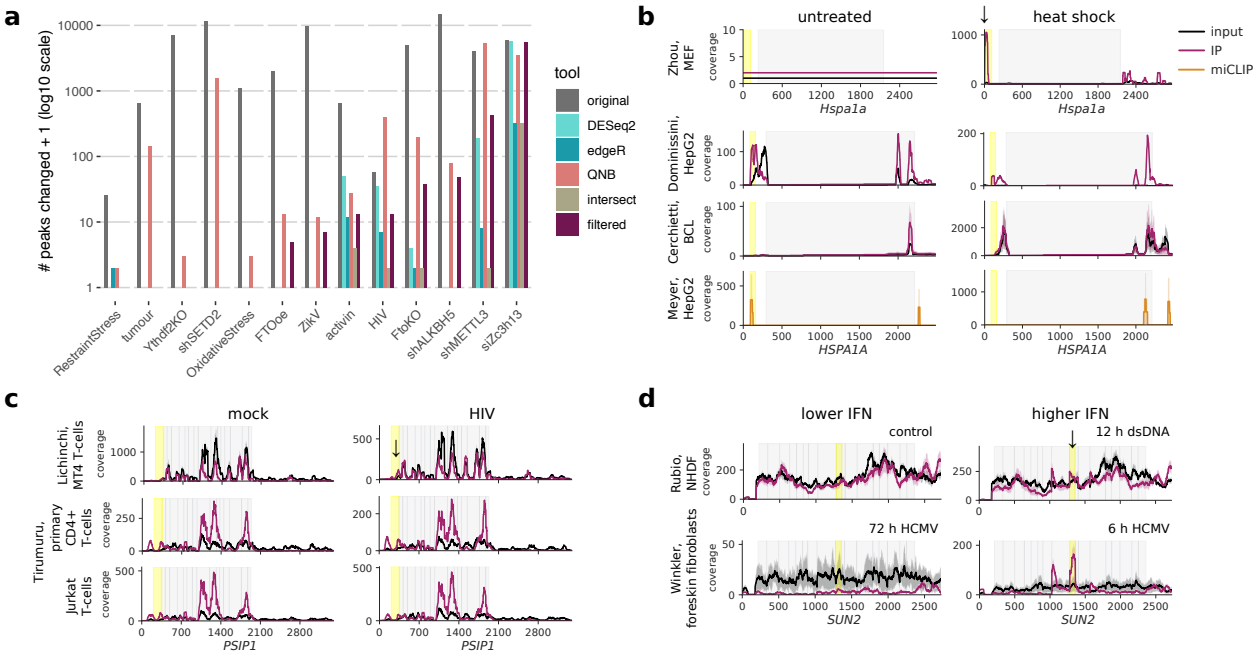
Using statistical methods to detect changes in peak enrichment, we found that the percent of changes called below a p-value threshold of 0.05 were similar in the positive and negative controls (**Figure 3b**). With all tools except MeTDiff, a knockout of Mettl3 showed the largest effects on m^6A (60), while fewer significant peaks in other positive controls suggested variable effects of the positive control conditions on $m^6A_{(m)}$, possibly related to efficiency of the methylation machinery knockdown or overexpression (7,33,61–65). In the absence of true differences between groups, p-value distributions should be uniform for well-calibrated statistical tests, meaning that ~5% of peaks should have p-values < 0.05 for the negative controls. MeTDiff reported an excess number of sites with p-values below 0.05 (**Additional File 2: Supplementary Figure 3d**) and identified a higher percentage of sites as differentially methylated in the mouse cortex negative control data set than in all but two positive controls (**Figure 3b**). On the other hand, the generalized linear models (GLMs) and QNB showed uniform to conservatively shifted p-value distributions, with differences between the mouse cortex and Huh7 data sets (**Additional File 2: Supplementary Figure 3d**), suggesting that these tools detect fewer false positives.

To ensure significant peak changes detected by each of the tools reflected changes in IP enrichment independent of differential gene expression, we measured the correlation between changes in

247 IP read counts at peak sites and changes in input read counts across their encompassing genes. For
248 significant peaks (FDR-adjusted p-value < 0.05) from the positive controls, correlation between $\log_2$ fold
249 change in peak IP and gene input read counts was low for the GLMs and QNB (Pearson's R = 0.10 to
250 0.22) but reached 0.55 (p = 5.8E-87) for MeTDiff (**Figure 3c**). The higher correlation for MeTDiff was
251 driven by peaks with proportional changes in IP and input levels, which suggests that MeTDiff often
252 detects differential expression of methylated genes rather than differential methylation. Therefore,
253 published studies that have used MeTDiff may actually be detecting differential expression and not
254 differential methylation (22,66). Indeed, plotting coverage for genes reported as differentially methylated
255 in one of these studies, with the y-axis scaled separately per condition, confirmed that changes in m^6A
256 identified by MeTDiff were proportional to changes in gene expression (**Figure 3d**) (22). Given these
257 results, QNB or the GLM implementations are better methods than MeTDiff to detect differential
258 methylation. Taking the intersect of significant peaks for the GLMs and QNB may help determine the
259 most probable sites of m^6A changes, while taking the union of predictions provides a less conservative
260 approach to selecting sites for further validation (**Figure 3e**). However, additional filters are needed for
261 robust peak change detection as there were still significant peaks for which the difference between peak
262 $\log_2$ fold change and gene $\log_2$ fold change was close to zero, particularly with QNB (**Additional File 2:**
263 **Supplementary Figure 3e**). For microarray and RNA-seq data, a filter of absolute $\log_2$ fold change > 1
264 has been recommended to reduce false positive rates (67); in the remainder of our analyses, we
265 implemented a similar filter for absolute difference in peak and gene $\log_2$ fold change ≥ 1 to the combined
266 predictions from QNB and the two GLMs, with an additional filter where noted for peak read counts ≥ 10
267 across all replicates and conditions to ensure sufficient coverage for consistent peak detection (as
268 discussed in Figure 1a).
269
270 *Reanalyzing peak changes between conditions*
271 We next estimated the scale of statistically detectable peak changes under various conditions
272 using our approaches and compared these results to previously reported estimates of these changes
273 (**Figure 4a**, **Additional File 1: Supplementary Table 4**). We identified fewer peaks as differentially

methylated than originally reported under most conditions, with zero to hundreds of peaks significantly changed (depending on experiment and method), versus hundreds to over ten thousand described in publications (22–26,34,63,66,68–71). Notably, knockdown of Zc3h13 did appear to disrupt $m^6A_{(m)}$, suggesting the gene does participate in methylation as recently described (69). Another study reported that activin treatment of human pluripotent stem cells led to differential methylation of genes that encode pluripotency factors (22). However, our reanalysis only found a few peak changes that passed our filters for significance, fold change, and expression (minimum input read count across peaks ≥ 10), and no enrichment for pluripotency factors among affected genes. Even when we removed the thresholds for fold change and expression, the adjusted p-value for "signaling pathways regulating pluripotency of stem cells" was still 0.15 and driven by only three genes, LEFTY2, FZD28, and FGFR3 (**Additional File 2: Supplementary Figure 4a**). Interestingly, the minimum read threshold made a particularly dramatic difference in the case of a recent study that looked at the effects of knocking down the histone methyltransferase SETD2 on m^6A in mRNA. For this data, of the 2065 sites predicted by QNB, 2064 fell below the minimum read threshold due to low input coverage in the first and second replicates (**Figure 4a**, **Additional File 2: Supplementary Figure 4b-e**) (70). We could not compare our approach to results reported by Su et al. (2018), who found 6,024 peaks changed with R2HG treatment, Zeng et al. (2018), who found 465-599 peaks changed between tumour samples, or Ma et al. (2018), who found 12,452 peaks were gained and 11,192 lost between P7 and P20 mouse cerebella, as each relied on a single

292 sample per condition, with no replicates (40–42).



293

Figure 4: Changes in peaks between conditions. a) Detected $m^6A_{(m)}$ changes in thirteen published data sets that measured $m^6A_{(m)}$ peak changes between two conditions **(Additional File 1: Supplementary Table 4)**. The number of peaks detected as changed in the original published analyses are compared to the number of peaks with FDR-adjusted p-values < 0.05 in our reanalysis using DESeq2, edgeR, or QNB, and taking the union of results from these three tools with additional filters for $\log_2$ fold difference in peak and gene changes of ≥1 and peak read counts ≥10 across all replicates and conditions ("filtered"). **b)** Gene coverage plots for *Hspa1a* in mouse embryonic fibroblasts (MEFs) and *HSPA1A* in human cells (HepG2 and BCL) before and after heat shock. Input coverage is shown in black and IP coverage in raspberry, with putative m^6A peaks changed highlighted in yellow and marked by arrows. miCLIP coverage for an experiment in HepG2 cells is shown in orange. **c)** Coverage plots for *PSIP1*, which was reported to have a change in 5′ UTR m^6A with HIV infection by Lichinchi et al (2016). **d)** Coverage plots for *SUN2*, in which we detected changes in m^6A with HCMV infection and dsDNA treatment suggesting a possible increase in methylation under higher interferon conditions (after 12 h of dsDNA treatment compared to untreated controls and after 6 h post-HCMV infection compared to 72 h, when interferon levels have declined). Lines in coverage plots (**b-d**) show the mean across all replicates for each experiment, while shading shows the standard deviation. Coding sequences are shown in grey.

294

295 Multiple studies have investigated $m^6A_{(m)}$ in the context of heat shock, HIV infection, KSHV

296 infection, and dsDNA treatment or human cytomegalovirus (HCMV) infection (**Additional File 1:**

297 **Supplementary Table 5**). Since each step in MeRIP-seq analysis risks introducing false negatives, we

298 cannot rule out consistent changes between studies that used similar experimental interventions based

299 on statistical detection alone. Therefore, we started by plotting coverage for specific genes reported as

300 differentially methylated to evaluate reproducibility across these studies. Zhou, et al. (2015) reported 5′
301 UTR methylation of *Hspa1a* with heat shock (20). Coverage was too low for untreated controls to
302 determine if *Hspa1a* was simply newly expressed or was actually newly methylated with heat shock
303 based on our alignment of their data using STAR (72). We were also unable to detect a change in
304 methylation of *HSPA1A* using data from other heat shock studies, including a new data set from a B-cell
305 lymphoma cell line and a published miCLIP data set, although coverage was again low (**Figure 4b**)
306 (4,73). Lichinchi, et al. (2016) reported that 56 genes showed increased methylation with HIV infection in
307 MT4 T-cells, with enrichment for genes involved in viral gene expression (25). Specific genes, for
308 example *PSIP1*, in which we also detected a peak using MACS2 and see a change in the peak when
309 plotting coverage using the data from Lichinchi et al. (2016), did not show the same changes in data from
310 two other $CD4^+$ cell types, primary $CD4^+$ cells and Jurkat cells (**Figure 4c**) (74). Two other studies both
311 used MeRIP-seq to establish the presence of m^6A in *IFNB1* induced through dsDNA treatment or by
312 infection with the dsDNA virus HCMV (75,76). While these studies did not discuss changes in m^6A, we
313 used these data sets to examine the replicability of $m^6A_{(m)}$ changes in response to dsDNA sensing and
314 interferon induction. Although different dsDNA stimuli, time points, and use of a fibroblast cell line versus
315 primary foreskin fibroblasts make it difficult to compare between the two experiments, using QNB and the
316 GLM approaches, we found four peaks in three genes (*AKAP8*, *SUN2*, and *TMEM140*) that showed
317 significant changes with higher interferon (**Figure 4d**). Overall, we were unable to detect the same
318 changes in $m^6A_{(m)}$ across studies of heat shock or HIV, and we detected only a few common changes in
319 the response to dsDNA. However, we do note that cell line-specific differences in $m^6A_{(m)}$ regulation and
320 differences in experimental protocols could account for some of the variability among these studies.
321 We did not have MeRIP-seq data for two studies from exactly the same conditions and cell lines
322 to compare, but two studies both used cell lines derived from iSLK to study the effects of KSHV on host
323 m^6A (27,28). Both suggested that KSHV infection could decrease the number of m^6A sites in host
324 transcripts. Hesser et al. (2018) found that lytic KSHV infection decreased the number of peaks on host
325 transcripts by >25%; Tan et al. (2018) suggested a loss of 17-59% of peaks in two different cell types, but
326 that $m^6A_{(m)}$ peak fold enrichment showed better clustering by cell type than by infection status. Neither of

327 these studies discussed specific genes that showed differential methylation with lytic infection. For our
328 comparison of $m^6A_{(m)}$ peak changes in these data sets, we identified probable changes in peaks based on
329 statistical significance using QNB or the GLMs with $\log_2$ fold change difference between peaks and genes
330 of ≥1. We detected 80 peak changes in the data from Hesser et al. (2018) and 18 in the data from Tan et
331 al. (2018) but found no peaks that changed in both iSLK data sets with lytic KSHV infection. Applying the
332 same statistical approaches, we were likewise unable to detect any shared peak changes between the
333 studies of HIV infection, and there were insufficient replicates to compare heat shock studies
334 (16,20,25,73,74). Thus, in our reanalysis of m^6A changes in response to stimuli, we detected only four
335 statistically reproducible peak changes, all in response to dsDNA.
336 Disparities between experiments were not simply due to significance thresholding or differences
337 in peak detection. Taking the union of peaks called in two experiments for KSHV, HIV, and dsDNA
338 treatment, we found minimal to negative correlations in changes in m^6A enrichment induced by treatment
339 at the same sites, further showing that changes with similar treatments are not reproducible (**Additional**
340 **File 2: Supplementary Figure 4e**).
341
342 *MeRIP-RT-qPCR validation*
343 Although statistical approaches revealed fewer changes in $m^6A_{(m)}$ with various stimuli than
344 published estimates, and we were unable to confirm changes in $m^6A_{(m)}$ methylation of specific genes
345 across studies of similar conditions, many of the studies we looked at do include additional validation of
346 $m^6A_{(m)}$ changes from MeRIP-seq using MeRIP-RT-qPCR. Recently it was shown that MeRIP-RT-qPCR
347 can capture differences in m^6A:A ratios at specific sites (34), but it is unknown how MeRIP-RT-qPCR is
348 affected by changes in gene expression. To test this, we ran MeRIP-RT-qPCR on in vitro transcribed
349 RNA oligonucleotides that lacked or contained m^6A spiked into total RNA extracted from Huh7 cells
350 (**Additional File 1: Supplementary Table 6**). We found that MeRIP-RT-qPCR detected the direction of
351 change in m^6A levels at different concentrations of spike-in RNAs (**Figure 5a-b**). However, technical
352 variation could also lead to spuriously significant differences. For example, a comparison of m^6A

enrichment between two dilutions (0.1 fmol and 10 fmol) of a 30% methylated spike-in mixture returned a p-value of 0.004 (unpaired Student's *t*-test).

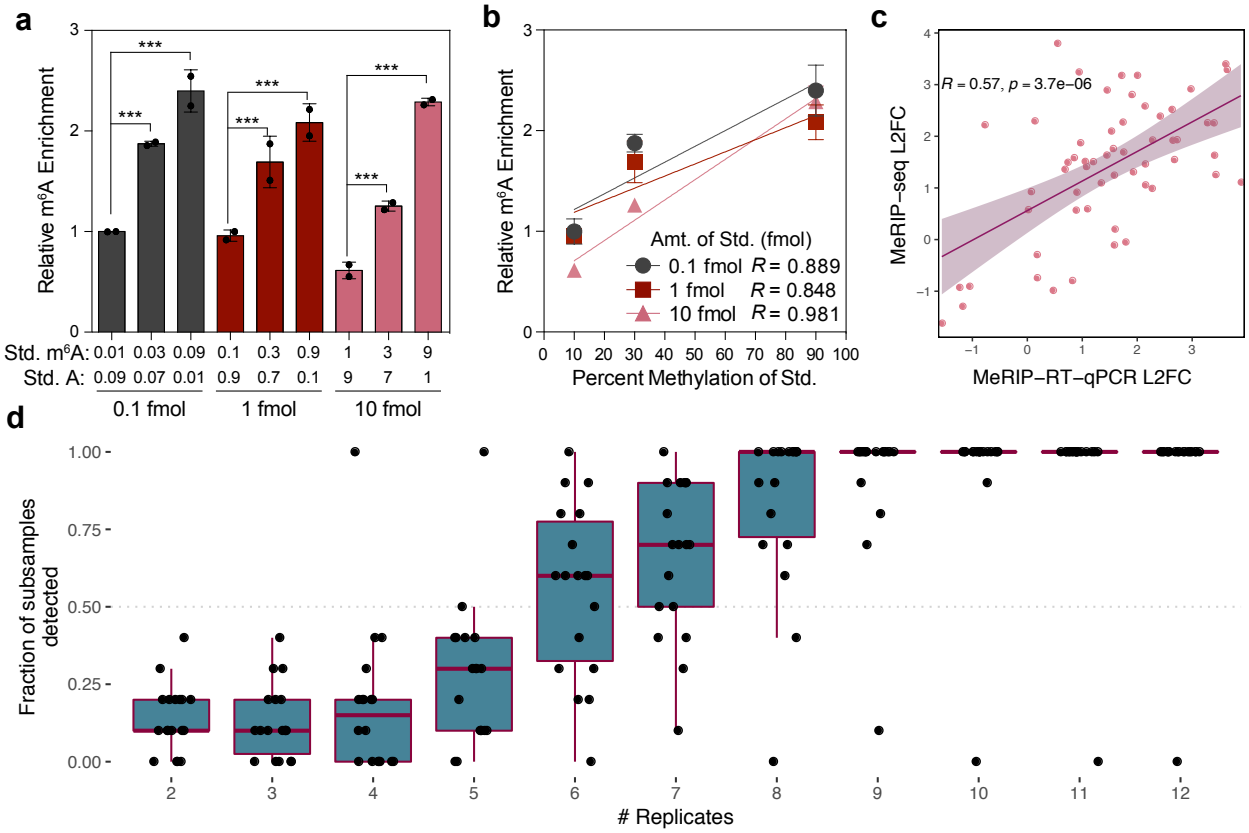


Figure 5: MeRIP-RT-qPCR validation and replicates necessary for the detection of peak changes. a) Relative enrichment of the indicated amounts of an in vitro transcribed standard containing unmodified A or m^{6}A, as measured by MeRIP-RT-qPCR. Data are shown for two independent replicates of three technical replicates each as IP enrichment over input relative to pulldown of a positive control spike-in, with the 0.1 fmol (0.01 m^{6}A: 0.09 A) sample normalized to 1. Bars represent mean ± SEM of two independent replicates. *** $p \leq 0.005$ by unpaired Student's *t*-test. **b)** Linear regression of relative m^{6}A enrichment from (a). Points and error bars mark mean ± SEM of two independent replicates. **c)** Change in MeRIP-RT-qPCR vs. MeRIP-seq enrichment for peaks detected as significantly differentially expressed with infection of Huh7 cells by dengue virus, Zika virus, and hepatitis C virus. **d)** Number of replicates of infected vs. uninfected cells needed to detect the peaks in (c). Replicates were randomly subsampled 10 times to calculate the fraction of subsamples in which peaks were called as significant by the GLMs or QNB. Boxes span the 1st to 3rd quartiles, with medians indicated. Whiskers show the minimum and maximum points within ±1.5x the interquartile distance from the boxes. Results for each subsample of replicates are shown as jittered points.

358 We next assessed the correlation between m^6A enrichment observed using MeRIP-seq and
359 MeRIP-RT-qPCR using data from our recent work that identified 58 peak changes in m^6A in Huh7 cells
360 following infection by four different viruses (50). For those experiments, we again selected peaks that
361 change based on results from the union of QNB and the GLM approaches. We found that the magnitude
362 of changes in common among viruses correlated between MeRIP-seq and MeRIP-RT-qPCR, both across
363 peaks (Pearson's R = 0.57, p = 3.7E-6) and within single peaks across viruses (13 out of 19 peaks
364 showed positive correlations, four of which had p-values < 0.05 with three data points) (**Figure 5c,**
365 **Additional File 2: Supplementary Figure 5**). Given the correlation we found between MeRIP-seq and
366 MeRIP-RT-qPCR, it is unclear why changes in IP over input sequencing reads were undetectable at the
367 peaks reported by Bertero et al. (2018) and Huang et al. (2019) but differences in peaks were
368 successfully validated using MeRIP-RT-qPCR (22,70). Based on these discrepancies, while MeRIP-RT-
369 qPCR can be used as an initial method of validation for predicted peak changes, additional methods are
370 necessary to confirm quantitative differences in m^6A levels and to resolve points where the assays do not
371 agree.
372 We next used our peaks validated with MeRIP-RT-qPCR to estimate the number of replicates
373 necessary for detection of changes with either the GLM or QNB methods. Using a permutation test, we
374 downsampled infected and uninfected replicates and reran statistical detection of changes. We found that
375 approximately 6-9 replicates were necessary for consistent detection (in at least 50% of subsamples) of
376 most peak changes (**Figure 5d**). Schurch et al. (2016) and Conesa et al. (2016) produced similar
377 recommendations for basic RNA-seq studies, finding that 6-12 replicates were necessary to detect most
378 changes in gene expression and that changes of 1.25 were detectable 25% of the time with five
379 replicates, rising to 44% with ten replicates, respectively (36,77). While our findings broadly agree with
380 these recommendations for RNA-seq, they also suggest that almost all published MeRIP-seq studies to
381 date are underpowered.
382
383 **Discussion**

384 In the eight years since MeRIP-/m^{6}A-seq was first published (16,17), many studies have used
385 these methods to examine the function of m^{6}A, its distribution along mRNA transcripts, and how it might
386 be regulated under various conditions. While 35 out of 64 of the MeRIP- and miCLIP-seq papers we
387 surveyed **(Additional File 1: Supplementary Table 1)** refer to m^{6}A as "dynamic", and, by contrast, only
388 two describe the modification as "static", the literature is unclear on what is meant by the word "dynamic".
389 There is mixed evidence as to whether m^{6}A is reversible through demethylation by the proposed
390 demethylases FTO and ALKBH5 (71,78–80). Recent research using an endoribonuclease-based method
391 for m^{6}A detection suggests that ALKBH5 has only a mild suppressive effect on m^{6}A levels and FTO no
392 effect (54). Although m^{6}A does not appear to change over the course of an mRNA's lifetime at steady-
393 state (3), whether it changes in response to a particular stimulus and at what point is less clear. Some
394 studies have suggested that m^{6}A may be modulated through changes in methyltransferase and
395 demethylase expression, producing consistent directions of change across transcripts (8,23,34), through
396 alternative mechanisms involving microRNA, transcription factors, promoters, or histone marks
397 (21,22,66,70,81), or through indeterminate mechanisms (17,20,25–28,52). However, based on our
398 reanalysis of available MeRIP-seq data, there is still only meagre support for widespread changes in m^{6}A
399 across the transcriptome independent of changes in the expression of methylation machinery (e.g.
400 increases or decreases in METTL3 expression).
401 In particular, replication of peaks and changes in peaks across studies is limited. As with other
402 RNA IP-based methods, MeRIP-seq data contains noise, owing to technical and biological variation (82).
403 In fact, while peak overlaps reach ~80% between replicates of the same study, they decrease to a
404 median of 45% between studies, most of which use 2-3 replicates each **(Figure 1)**. Given that the
405 detection of peaks is so variable and that peak heights differ among replicates, it is perhaps not surprising
406 that peak changes have yet to be reproduced between multiple studies of similar conditions. Indeed,
407 variability in MeRIP-seq could also mask differences in m^{6}A regulation among cell types, which have
408 been described in mouse brains (34) and in cell lines exposed to KSHV (28). To distinguish biological and
409 technical variation, it will therefore be particularly important to test if multiple groups using the same cell
410 line and conditions can better reproduce changes in m^{6}A.

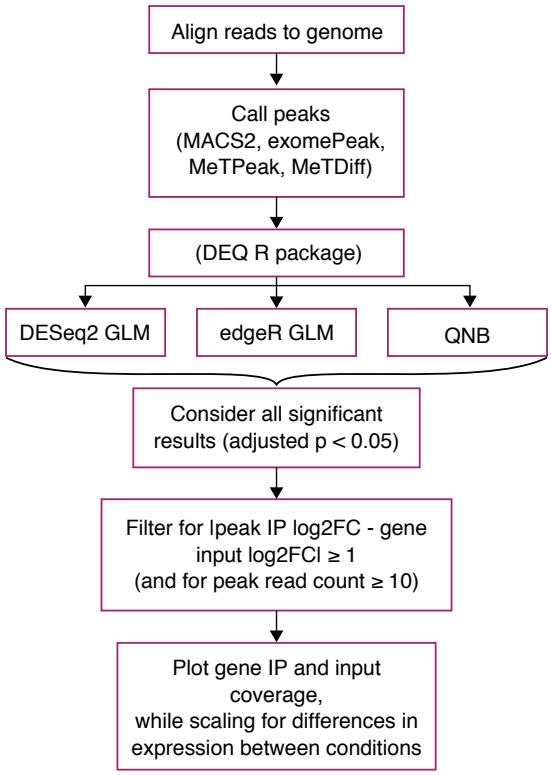


Figure 6: Proposed approach to identify candidates for $m^6A_{(m)}$ changes for further validation using MeRIP-seq data. We suggest predicting changes in $m^6A_{(m)}$ using DESeq2, edgeR, and QNB, and have implemented the DEQ package in R to facilitate this.

411

412 Disparities in the methods used to detect changes in $m^6A_{(m)}$ peaks also play a role in differing
413 conclusions among studies. Here, we analyzed four statistical methods to detect changes in peaks and
414 found that three of these methods showed uniform or conservatively shifted p-value distributions and
415 were able to identify changes in $m^6A_{(m)}$ independent of changes in gene expression. We therefore
416 suggest that these statistical methods, in combination with filters for input levels in both conditions and
417 the difference in $\log_2$ fold change between peaks and genes, can be used to identify candidate $m^6A_{(m)}$
418 sites from MeRIP-seq data for further analysis and validation (**Figure 6**). Based on our results, while
419 MeTDiff works for peak detection, we do not recommend MeTDiff for peak change detection as it does
420 not control well for differences in gene expression **(Figure 3)**. Similar to others (33), we found that plotting
421 predicted m^6A changes was invaluable and that appropriate scaling for gene coverage could reveal
422 changes proportional to gene expression. In addition, plotting the standard deviation in transcript
423 coverage can help assess typical variation in peak height among replicates. We note that both differential
424 methylation of a gene and methylation of a gene that is differentially expressed could be important, but
425 they should not be conflated when considering the role of m^6A in transcript regulation.

426 The extent to which m^{6}A changes on particular transcripts and whether it changes in binary
427 presence/absence or in degree is unclear. MeRIP-RT-qPCR could detect methylation differences in in
428 vitro transcribed RNA. Further, we found that these changes correlated with differences in MeRIP-seq
429 enrichment. However, neither MeRIP-seq nor MeRIP-RT-qPCR can reveal the precise fraction of
430 transcript copies modified by m^{6}A. In general, antibody-based methods are subject to biases, including
431 from differences in binding efficiencies based on RNA structure and motif preferences (83). There is an
432 oft-cited but little-used method for quantification of m^{6}A, site-specific cleavage and radioactive-labeling
433 followed by ligation-assisted extraction and thin-layer chromatography (SCARLET) (19). However, this
434 method can be challenging, works only for highly abundant transcripts, and is impractical for
435 transcriptome-wide analysis. A recently developed endoribonuclease-based, antibody-independent
436 approach for m^{6}A detection is promising in terms of quantification of m^{6}A, but its use is limited to a subset
437 of m^{6}A sites within DRAC motifs ending in ACA (~16% of all sites) (44,45). So far, comparison to this data
438 suggests that antibody-based approaches may underestimate the number of m^{6}A sites (54). Alternative
439 methods to detect m^{6}A based using single-molecule sequencing (including direct RNA sequencing and
440 real-time cDNA synthesis) are under development and may offer ways to detect, quantify, and phase m^{6}A
441 sites, but these have not yet been shown to accurately detect m^{6}A across a cellular transcriptome (84–
442 86). For now, site-specific SCARLET is the only option to biochemically validate proposed changes in
443 m^{6}A at most motifs.
444
445
446 **Conclusions**
447 Our work reveals the limits of MeRIP-seq reproducibility for the detection of m^{6}A$_{(m)}$ and in
448 particular suggests caution when using MeRIP-seq for the detection of changes in m^{6}A$_{(m)}$. To increase
449 confidence in predicted changes in m^{6}A$_{(m)}$, we propose statistical approaches that account for differences
450 in gene expression between conditions and variability among replicates. These methods can be used to
451 gain insight into the regulation and function of m^{6}A$_{(m)}$ and to predict specific sites for validation before the

452 development of high-throughput alternatives to MeRIP-seq, and similar strategies may be applicable to
453 other types of RNA sequencing assay.
454
455 **Methods**
456 *New MeRIP-seq data*
457 *- Huh7 data*
458 Total RNA was extracted from Huh7 cells using Trizol (Thermo-Fisher). mRNA was purified from 200 µg
459 total RNA using the Dynabeads mRNA purification kit (Thermo-Fisher) and concentrated by ethanol
460 precipitation. Purified mRNA was fragmented using the RNA Fragmentation Reagent (Thermo-Fisher) for
461 15 minutes followed by ethanol precipitation. Then, MeRIP was performed using EpiMark N6-
462 methyladenosine Enrichment kit (NEB). 25 µL Protein G Dynabeads (Thermo-Fisher) per sample were
463 washed three times in MeRIP buffer (150 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40) and incubated
464 with 1 µL anti-m^6A antibody (NEB) for 2 hours at 4°C with rotation. After washing three times, anti-m^6A
465 conjugated beads were incubated with purified mRNA with rotation at 4°C overnight in 300 µL MeRIP
466 buffer with 1 µL RNAse inhibitor (recombinant RNasein; Promega). Beads were then washed twice with
467 500 µL MeRIP buffer, twice with low salt wash buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-
468 40), twice with high salt wash buffer (500 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40), and once
469 again with MeRIP buffer. m^6A-modified RNA was eluted twice in 100 µL MeRIP buffer containing 5mM
470 m^6A salt (Santa Cruz Biotechnology) for 30 minutes at 4°C with rotation and concentrated by ethanol
471 precipitation. RNA-seq libraries were prepared from eluate and the 10% of RNA set aside as input using
472 the TruSeq mRNA library prep kit (Illumina) and checked for fragment length using the Agilent 2100
473 Bioanalyzer. Single-end 50 base pair reads were sequenced on an Illumina HiSeq 2500.
474 *- Heat shock*
475 Early passage OCI-Ly1 diffuse large B-cell lymphoma cells were grown in Iscove's modified Eagle
476 Medium (IMDM) with 10% fetal bovine serum (FBS). OCI-Ly1 cells were obtained from the Ontario
477 Cancer Institute and regularly tested for *Mycoplasma* contamination by PCR and identified by single
478 nucleotide polymorphism. Cells were maintained with 1% penicillin/streptomycin in a 37°C, 5% CO_2,

479 humidified incubator. In these growing conditions, heat shocked cells were exposed to 43 °C for 1 hour,
480 followed by 1 hour of recovery at 37°C while control cells were maintained at 37°C. Following treatment,
481 cells were processed at 4°C to obtain total cell lysates. Lysates were immunoprecipitated for $m^6A_{(m)}$ using
482 Synaptic Systems antibody (SYSY 202 003) following the protocol described in Meyer, et al (2012) and
483 sequenced on an Illumina HiSeq 2500 (16).
484
485 *Read processing*
486 Reads were trimmed using Trimmomatic (87) and aligned to the human genome (hg38) or the mouse
487 genome (mm10), as appropriate, using STAR, a splice-aware aligner for RNA-seq data (72). We used the
488 flag "--outFilterMultimapNmax 1" to keep only uniquely aligned reads. Scripts used for alignment are
489 provided with the rest of the analysis scripts at https://github.com/al-mcintyre/merip_reanalysis_scripts.
490
491 *Peak detection and comparison*
492 IP over input peaks were called using MACS2 callpeak using the parameters "--nomodel --extsize 100 (or,
493 if available, the approximate fragment size for a specific experiment to extend reads at their 3' ends to a
494 fixed length) --gsize 100e6 (the approximate size of mouse and human transcriptomes based on gencode
495 annotations)" (49). No filter for coverage was applied at the stage of peak detection. Transcript coverage
496 was estimated using Kallisto (88) with an index construct 31mers, except for the Schwartz et al (2014)
497 data set, where the reads were too short and an alternative index based on 29mers was constructed (33).
498 For **Figure 1b**, the full union of unique peaks was taken and the percent of that set detected in single
499 replicates calculated. Intersects between peaks that overlapped for transcripts with ≥10X mean coverage
500 in both samples were taken using bedtools (89) for **Figure 2**, allowing a generous minimum of 1
501 overlapping base. Heatmaps for peak overlaps were generated using the ComplexHeatmap package in R
502 (90). MeRIP-seq data sets in **Figure 2b** included those for human cell lines in **Figure 2a**, other data sets
503 from the same studies and any data sets that shared the same cell lines, and other data sets that looked
504 at multiple human cell types. We considered only data sets from baseline conditions in **Figure 2**
505 (untreated cells and knockdown controls).

506

507 *Poisson and negative binomial fits*

508 Reads aligned to peaks were counted using featureCounts from the Rsubread package (91). Poisson and
509 negative binomial models were fit to input and IP read counts at peaks using maximum likelihood
510 estimation. Simulated read counts were generated with Poisson or negative binomial distributions based
511 on estimated parameters from the sample, with 500 random generations per model. The log likelihood of
512 seeing read counts from the sample and the simulations given the model parameters was then calculated
513 and the mean taken across all peaks.

514

515 *Peak change detection and generalized linear models*

516 Generalized linear models to detect changes in IP coverage while controlling for differences in input
517 coverage were implemented based on a method previously applied to HITS-CLIP data (58). Full and
518 reduced models were constructed as follows:

519 $$\log \mu_{ij} = \beta_i^0 + \beta_i^{IP} X_j^{IP} + \beta_i^{STIM} X_j^{STIM} + \beta_i^{STIM:IP} X_j^{STIM:IP}$$

520 $$\log \mu_{ij} = \beta_i^0 + \beta_i^{IP} X_j^{IP} + \beta_i^{STIM} X_j^{STIM}$$

521

522 Where μ_{ij} is the expected read count for peak i in sample j, modelled as a negative binomial distribution,
523 $X_j^{IP} = 1$ for IP samples and 0 for input samples, and $X_j^{STIM} = 1$ for samples under the experimental
524 intervention and 0 for control samples.

525

526 Statistical significance was then assessed using a chi-squared test (df=1) for the difference in deviances
527 between the full and reduced models, with the null hypothesis that the interaction term ($\beta_i^{STIM:IP}$) for
528 differential antibody enrichment driven by the experimental intervention is zero. The likelihood ratio test
529 was implemented through DESeq2 (56) and edgeR (57), two programs developed for RNA-seq analysis
530 that differ in how they filter data and in how they estimate dispersions for negative binomial distributions.
531 Generalized linear models implemented through edgeR included a term for the normalized library size of
532 sample j.

533

534 QNB was run as suggested for experiments with biological replicates, where each IP and input variable
535 ("ip1", etc.) consisted of a matrix of peak counts for either condition 1 or condition 2:
536 > qnbtest(ip1, ip2, input1, input2, mode="per-condition")
537
538 We extracted functions from MeTDiff so that we could supply our own peaks and thus control for
539 differences in peak detection among tools. The main post-peak calling function, diff.call.module, was run
540 as follows using the same count matrices as for QNB:
541 > diff.call.module(ip1, input1, ip2, input2)
542
543 Gene and peak expression changes were estimated as $\log_2$ fold changes from DESeq2 based on
544 differences in input read counts aligned to genes and IP read counts aligned to peaks, respectively, and
545 the change in peak relative to gene enrichment was calculated as the absolute difference in $\log_2$ fold
546 change between those values.
547
548 *Comparison to published studies*
549 The sources for published estimates of m^6A peak changes included in our comparison are listed in
550 **Additional File 1: Supplementary Table 4**. Significant (FDR-adjusted $p < 0.05$) peaks were considered
551 for DESeq2, edgeR, and QNB, run as described above. We also considered a filtered set of peaks
552 derived from the union of significant peaks from the three tools with additional filters for location within
553 exons, $|\log_2$ fold change between peak IP and gene input$| \geq 1$, and a minimum peak read count of 10
554 across replicates and conditions. We used gProfiler to calculate enrichment of functional categories (92).
555
556 In **Figure 4b-c**, we selected *Hspa1a/HSPA1A* as our representative gene for heat shock because it was
557 the primary example cited by Zhou et al. (2015) and Meyer et al. (2015) (4,20). For HIV, we selected
558 *PSIP1* because it was among the 56 genes reported by Lichinchi et al. (2016a) (25), it plays a known role
559 in HIV infection, and we detected a peak in the gene using MACS2.

560

561 For KSHV, we compared significant results (adjusted $p < 0.05$) from QNB and GLMs (DESeq2 and
562 edgeR), with additional filtering for |peak IP – gene input $\log_2$ fold change| ≥ 1 (lowering this threshold to
563 0.5 did not change results), for data from Hesser et al. (2018) (27) in lytic vs. latent iSLK.219 cells and
564 data from Tan et al. (2018) (28) in lytic vs. latent iSLK BAC16 cells. We used the same approach to
565 compare data from Rubio et al. (2018) and Winkler et al. (2019) (75,76) for response to dsDNA. Data sets
566 used for site-specific comparisons are summarized in **Additional File 1: Supplementary Table 5**.
567

568 Gene coverage was plotted using CovFuzze (https://github.com/al-mcintyre/CovFuzze), which
569 summarizes mean and standard deviation in coverage across available replicates (93). Pearson's
570 correlations were taken for **Supplementary Figure 4** for peaks expressed above a minimum input peak
571 read count of 10 across replicates and conditions.
572

573 *Spike-in controls and MeRIP-RT-qPCR*
574 In vitro transcribed (IVT) controls were provided by the Jaffrey Lab and consisted of 1001 base long RNA
575 sequences with three adenines in GAC motifs (**Additional File 1: Supplementary Table 6**) either fully
576 methylated or unmethylated. m^6A and A controls were mixed in various ratios (1:9, 3:7, and 9:1) that
577 approximate the variation in m^6A levels detected by SCARLET (m^6A levels at specific sites have been
578 reported to vary from 6-80% of transcripts (19)). Modified and unmodified standards were mixed at the
579 indicated ratios to yield a final quantity of 0.1 fmol, 1 fmol, and 10 fmol. Mixed RNA standards were added
580 to 30 µg total RNA from Huh7 cells, along with 0.1 fmol of positive (m^6A-modified *Gaussia* luciferase
581 RNA, "GLuc") and negative control (unmodified *Cypridina* luciferase, "CLuc") spike-in RNA provided with
582 the N6-methyladenosine Enrichment kit (EpiMark). Following MeRIP as described above, cDNA was
583 synthesized from eluate and input samples using the iScript cDNA synthesis kit (Bio-Rad), and RT-qPCR
584 was performed on a QuantStudio Flex 6 instrument. Data was analyzed as a percent of input of the spike-
585 in RNA in each condition relative to that of the provided positive control spike-in.

586 For MeRIP-RT-qPCR to test peak callers, Huh7 cells plated in 6-well plates were transfected with
587 siRNAs against METTL3 and METTL14 (Qiagen; SI04317096 and SI00459942) or non-targeting control
588 siRNa (SI03650318) using Lipofectamine RNAiMax (Thermo Fisher) twice, 24 hours apart. 48 hours
589 following the second round of siRNA transfection, cells were harvested in TRIzol reagent and total RNA
590 was extracted. 30 µg total RNA was fragmented for 3 mins at 75°C, concentrated by ethanol precipitation,
591 and MeRIP-RT-qPCR was performed as described above. Primers used for RT-qPCR are provided in
592 **Additional File 1: Supplementary Table 7** and siRNA sequences in **Additional File 1: Supplementary**
593 **Table 8.**
594
595 *Cell culture and infection (data used for MeRIP-RT-qPCR comparisons)*
596 Huh7 cells were grown in DMEM (Mediatech) supplemented with 10% fetal bovine serum (HyClone), 2.5
597 mM HEPES, and 1X non-essential amino acids (Thermo-Fisher). The identity of the Huh7 cell lines was
598 verified using the Promega GenePrint STR kit (DNA Analysis Facility, Duke University), and cells were
599 verified as mycoplasma free by the LookOut Mycoplasma PCR detection kit (Sigma). Infectious stocks of
600 a cell culture-adapted strain of genotype 2A JFH1 HCV were generated and titered on Huh7.5 cells by
601 focus-forming assay (FFA), as described (94). Dengue virus (DENV2-NGC), West Nile virus (WNV-
602 NY2000), and Zika virus (ZIKV-PRVABC59) viral stocks were generated in C6/36 cells and titered on
603 Vero cells as described (94). All viral infections were performed at a multiplicity of infection of 1 for 48
604 hours.
605
606 **Declarations**

607 **Ethics approval and consent to participate**

608 Not applicable

609 **Consent for publication**

610 Not applicable

611 **Availability of Data and Materials**

612 MeRIP-seq data for the Huh7 negative controls is available in the GEO repository, under accession
613 number GSE130891. MeRIP-seq data for heat shock in B-cell lymphoma is available under accession
614 number GSE130892. Accession numbers for all other data sets reanalyzed in the study are included in
615 Additional File 1: Supplementary Tables 1-5. Scripts used for analysis are available at
616 https://github.com/al-mcintyre/merip_reanalysis_scripts and a pipeline implementing generalized linear
617 models through DESeq2 and edgeR, as well as QNB, is provided at https://github.com/al-mcintyre/deq.

618 **Competing interests**

619 C.E.M. is a cofounder and board member for Biotia and Onegevity Health, as well as an advisor or
620 compensated speaker for Abbvie, Acuamark Diagnostics, ArcBio, BioRad, DNA Genotek, Genialis,
621 Genpro, Karius, Illumina, New England Biolabs, QIAGEN, Whole Biome, and Zymo Research.

622 **Funding**

623 We are grateful for funding from the Starr Cancer Consortium (I9-A9-071), the Bert L and N. Kuggie Vallee
624 Foundation, the WorldQuant Foundation, The Pershing Square Sohn Cancer Research Alliance, NASA
625 (NNX14AH50G), the National Institutes of Health (R01AI125416, R21AI129851, R01MH117406), the
626 Leukemia and Lymphoma Society (LLS) grants (LLS 9238-16, Mak, LLS-MCL-982, Chen-Kiang), the
627 Burroughs Wellcome Fund (S.M.H.), the National Science and Engineering Research Council of Canada
628 (A.B.R.M. PGS-D funding), and the American Heart Association (N.S.G. Pre-doctoral Fellowship,
629 17PRE33670017).

630 **Authors' contributions**

A.B.R.M. and C.E.M. conceived the study. A.B.R.M. developed and ran the analyses and wrote the manuscript with N.S.G. and S.M.H. S.R.J. provided in vitro controls for MeRIP-RT-qPCR. N.S.G. prepared MeRIP-seq libraries for the Huh7 controls and ran MeRIP-RT-qPCR tests. L.C. contributed additional heat shock data. All authors read and edited the manuscript.

Acknowledgements

We would like to thank Christina Leslie and Jonathan Victor for statistical advice, Aashiq Mirza for the MeRIP-RT-qPCR controls, and Helen Lazear for WNV infection. We would also like to thank the Epigenomics Core at Weill Cornell for preparing MeRIP-seq libraries, the Scientific Computing Unit (SCU), and New England Biolabs for donating anti-m^6A antibodies.

References

1. Balacco DL, Soller M. The m6A Writer: Rise of a Machine for Growing Tasks. Biochemistry. 2019;
2. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, et al. A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. Genes & development. 2015;
3. Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbø CB, Geula S, et al. m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. Genes & Development. 2017;31(10):990–1006.
4. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5′ UTR m6A promotes cap-independent translation. Cell. 2015;163(4):999–1010.
5. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N6-methyladenosine modulates messenger RNA translation efficiency. Cell. 2015;161(6):1388–99.
6. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m6A methyltransferase METTL3 promotes translation in human cancer cells. Molecular cell. 2016;62(3):335–45.
7. Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. N 6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nature cell biology. 2014;16(2):191.
8. Molinie B, Wang J, Lim KS, Hillebrand R, Lu Z, Van Wittenberghe N, et al. m 6 A-LAIC-seq reveals the census and complexity of the m 6 A epitranscriptome. Nature methods. 2016;13(8):692.

658 9. Yoon K-J, Ringeling FR, Vissers C, Jacob F, Pokrass M, Jimenez-Cyrus D, et al. Temporal control of
659 mammalian cortical neurogenesis by m6A methylation. Cell. 2017;171(4):877–89.

660 10. Meyer KD, Jaffrey SR. Rethinking m6A readers, writers, and erasers. Annual review of cell and
661 developmental biology. 2017;33:319–42.

662 11. Roundtree IA, Evans ME, Pan T, He C. Dynamic RNA modifications in gene expression regulation.
663 Cell. 2017;169(7):1187–200.

664 12. Rosa-Mercado NA, Withers JB, Steitz JA. Settling the m6A debate: methylation of mature mRNA is
665 not dynamic but accelerates turnover. Genes & development. 2017;31(10):957–8.

666 13. Darnell RB, Ke S, Darnell JE. Pre-mRNA processing includes N6 methylation of adenosine residues
667 that are retained in mRNA exons and the fallacy of “RNA epigenetics.” RNA. 2018;24(3):262–7.

668 14. Zhao BS, Nachtergaele S, Roundtree IA, He C. Our views of dynamic N6-methyladenosine RNA
669 methylation. RNA. 2018;24(3):268–72.

670 15. Mauer J, Jaffrey SR. FTO, m6Am, and the hypothesis of reversible epitranscriptomic mRNA
671 modifications. FEBS letters. 2018;592(12):2012–22.

672 16. Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. Comprehensive analysis of
673 mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell. 2012;149(7):1635–
674 46.

675 17. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al.
676 Topology of the human and mouse m 6 A RNA methylomes revealed by m 6 A-seq. Nature.
677 2012;485(7397):201.

678 18. Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. Single-nucleotide-
679 resolution mapping of m6A and m6Am throughout the transcriptome. Nature methods.
680 2015;12(8):767.

681 19. Liu N, Parisien M, Dai Q, Zheng G, He C, Pan T. Probing N6-methyladenosine RNA modification
682 status at single nucleotide resolution in mRNA and long noncoding RNA. Rna. 2013;

683 20. Zhou J, Wan J, Gao X, Zhang X, Jaffrey SR, Qian S-B. Dynamic m 6 A mRNA methylation directs
684 translational control of heat shock response. Nature. 2015;526(7574):591.

685 21. Chen T, Hao Y-J, Zhang Y, Li M-M, Wang M, Han W, et al. m6A RNA Methylation Is Regulated by
686 MicroRNAs and Promotes Reprogramming to Pluripotency. Cell Stem Cell. 2015 Mar 5;16(3):289–
687 301.

688 22. Bertero A, Brown S, Madrigal P, Osnato A, Ortmann D, Yiangou L, et al. The SMAD2/3 interactome
689 reveals that TGFβ controls m 6 A mRNA methylation in pluripotency. Nature. 2018;555(7695):256.

690 23. Liu J, Eckert MA, Harada BT, Liu S-M, Lu Z, Yu K, et al. m 6 A mRNA methylation regulates AKT
691 activity to promote the proliferation and tumorigenicity of endometrial cancer. Nature cell biology.
692 2018;20(9):1074.

693 24. Anders M, Chelysheva I, Goebel I, Trenkner T, Zhou J, Mao Y, et al. Dynamic m6A methylation
694 facilitates mRNA triaging to stress granules. Life science alliance. 2018;1(4):e201800113.

695 25. Lichinchi G, Gao S, Saletore Y, Gonzalez GM, Bansal V, Wang Y, et al. Dynamics of the human and
696 viral m6A RNA methylomes during HIV-1 infection of T cells. Nature Microbiology. 2016;1.

697 26. Lichinchi G, Zhao BS, Wu Y, Lu Z, Qin Y, He C, et al. Dynamics of human and viral RNA methylation
698 during Zika virus infection. Cell host & microbe. 2016;20(5):666–73.

699 27. Hesser CR, Karijolich J, Dominissini D, He C, Glaunsinger BA. N6-methyladenosine modification and
700 the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's
701 sarcoma-associated herpesvirus infection. PLoS pathogens. 2018;14(4):e1006995.

702 28. Tan B, Liu H, Zhang S, da Silva SR, Zhang L, Meng J, et al. Viral and cellular N 6-methyladenosine
703 and N 6, 2′-O-dimethyladenosine epitranscriptomes in the KSHV life cycle. Nature microbiology.
704 2018;3(1):108.

705 29. Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, et al. The N 6-methyladenosine (m
706 6 A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and
707 leukemia cells. Nature medicine. 2017;23(11):1369.

708 30. Liu L, Zhang S-W, Huang Y, Meng J. QNB: differential RNA methylation analysis for count-based
709 small-sample sequencing data with a quad-negative binomial model. BMC bioinformatics.
710 2017;18(1):387.

711 31. Cui X, Zhang L, Meng J, Rao MK, Chen Y, Huang Y. MeTDiff: a novel differential RNA methylation
712 analysis for MeRIP-seq data. IEEE/ACM Transactions on Computational Biology and Bioinformatics
713 (TCBB). 2018;15(2):526–34.

714 32. Zhao X, Yang Y, Sun B-F, Shi Y, Yang X, Xiao W, et al. FTO-dependent demethylation of N6-
715 methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell research.
716 2014;24(12):1403.

717 33. Schwartz S, Mumbach MR, Jovanovic M, Wang T, Maciag K, Bushkin GG, et al. Perturbation of m6A
718 writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell reports.
719 2014;8(1):284–96.

720 34. Engel M, Eggert C, Kaplick PM, Eder M, Röh S, Tietze L, et al. The role of m6A/m-RNA methylation
721 in stress response regulation. Neuron. 2018;99(2):389–403.

722 35. Zhou J, White KP, Liu Y. RNA-seq differential expression studies: more sequence or more
723 replication? Bioinformatics. 2013 Dec 4;30(3):301–4.

724 36. Schurch NJ, Schofield P, Gierliński M, Cole C, Sherstnev A, Singh V, et al. How many biological
725 replicates are needed in an RNA-seq experiment and which differential expression tool should you
726 use? Rna. 2016;22(6):839–51.

727 37. Su Z, Łabaj PP, Li S, Thierry-Mieg J, Thierry-Mieg D, Shi W, et al. A comprehensive assessment of
728 RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control
729 Consortium. Nature biotechnology. 2014;32(9):903.

730 38. Geula S, Moshitch-Moshkovitz S, Dominissini D, Mansour AA, Kol N, Salmon-Divon M, et al. m6A
731 mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science.
732 2015;347(6225):1002–6.

733 39. Cui Q, Shi H, Ye P, Li L, Qu Q, Sun G, et al. m6A RNA methylation regulates the self-renewal and
734 tumorigenesis of glioblastoma stem cells. Cell reports. 2017;18(11):2622–34.

735 40. Ma C, Chang M, Lv H, Zhang Z-W, Zhang W, He X, et al. RNA m 6 A methylation participates in
736 regulation of postnatal development of the mouse cerebellum. Genome biology. 2018;19(1):68.

737 41. Su R, Dong L, Li C, Nachtergaele S, Wunderlich M, Qing Y, et al. R-2HG exhibits anti-tumor activity
738 by targeting FTO/m6A/MYC/CEBPA signaling. Cell. 2018;172(1–2):90–105.

739 42. Zeng Y, Wang S, Gao S, Soares F, Ahmed M, Guo H, et al. Refined RIP-seq protocol for
740 epitranscriptome analysis with low input materials. PLoS biology. 2018;16(9):e2006092.

741 43. Zhou J, Wan J, Shu XE, Mao Y, Liu X-M, Yuan X, et al. N6-methyladenosine guides mRNA
742 alternative translation during integrated stress response. Molecular cell. 2018;69(4):636–47.

743 44. Garcia-Campos MA, Edelheit S, Toth U, Shachar R, Nir R, Lasman L, et al. Deciphering the'm6A
744 code'via quantitative profiling of m6A at single-nucleotide resolution. BioRxiv. 2019;571679.

745 45. Zhang Z, Chen L-Q, Zhao Y-L, Yang C-G, Roundtree IA, Zhang Z, et al. Single-base mapping of m6A
746 by an antibody-independent method. bioRxiv. 2019;575555.

747 46. Liu H, Begik O, Lucas MC, Mason CE, Schwartz S, Mattick JS, et al. Accurate detection of m6A RNA
748 modifications in native RNA sequences. bioRxiv. 2019;525741.

749 47. Meng J, Lu Z, Liu H, Zhang L, Zhang S, Chen Y, et al. A protocol for RNA methylation differential
750 analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. Methods.
751 2014;69(3):274–81.

752 48. Cui X, Meng J, Zhang S, Chen Y, Huang Y. A novel algorithm for calling mRNA m 6 A peaks by
753 modeling biological variances in MeRIP-seq data. Bioinformatics. 2016;32(12):i378–85.

754 49. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. Model-based analysis of
755 ChIP-Seq (MACS). Genome biology. 2008;9(9):R137.

756 50. Gokhale NS, McIntyre ABR, Mattocks MD, Holley CL, Lazear HM, Mason CE, et al. Altered m6A
757 modification of specific cellular transcripts affects Flaviviridae infection. bioRxiv. 2019 Jan 1;670984.

758 51. Antanaviciute A, Baquero-Perez B, Watson CM, Harrison SM, Lascelles C, Crinnion L, et al.
759 m6aViewer: software for the detection, analysis, and visualization of N6-methyladenosine peaks
760 from m6A-seq/ME-RIP sequencing data. RNA. 2017;23(10):1493–501.

761 52. He S, Wang H, Liu R, He M, Che T, Jin L, et al. mRNA N6-methyladenosine methylation of postnatal
762 liver development in pig. PloS one. 2017;12(3):e0173421.

763 53. Tao X, Chen J, Jiang Y, Wei Y, Chen Y, Xu H, et al. Transcriptome-wide N 6-methyladenosine
764 methylome profiling of porcine muscle and adipose tissues reveals a potential mechanism for
765 transcriptional regulation and differential methylation pattern. BMC genomics. 2017;18(1):336.

766 54. Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, et al. Deciphering the "m6A
767 Code" via Antibody-Independent Quantitative Profiling. Cell. 2019 Jul 25;178(3):731-747.e16.

768 55. Xiao S, Cao S, Huang Q, Xia L, Deng M, Yang M, et al. The RNA N6-methyladenosine modification
769 landscape of human fetal tissues. Nature Cell Biology. 2019 May 1;21(5):651–61.

770 56. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data
771 with DESeq2. Genome biology. 2014;15(12):550.

772 57. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression
773 analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.

774 58. Park S-M, Deering RP, Lu Y, Tivnan P, Lianoglou S, Al-Shahrour F, et al. Musashi-2 controls cell
775 fate, lineage bias, and TGF-β signaling in HSCs. Journal of Experimental Medicine.
776 2014;211(1):71–87.

777 59. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical
778 reproducibility and comparison with gene expression arrays. Genome research. 2008;18(9):1509–
779 17.

780 60. Batista PJ, Molinie B, Wang J, Qu K, Zhang J, Li L, et al. m6A RNA modification controls cell fate
781 transition in mammalian embryonic stem cells. Cell stem cell. 2014;15(6):707–19.

782 61. Fustin J-M, Doi M, Yamaguchi Y, Hida H, Nishimura S, Yoshida M, et al. RNA-methylation-dependent
783 RNA processing controls the speed of the circadian clock. Cell. 2013;155(4):793–806.

784 62. Hess ME, Hess S, Meyer KD, Verhagen LA, Koch L, Brönneke HS, et al. The fat mass and obesity
785 associated gene (Fto) regulates activity of the dopaminergic midbrain circuitry. Nature
786 neuroscience. 2013;16(8):1042.

787 63. Li Z, Weng H, Su R, Weng X, Zuo Z, Li C, et al. FTO plays an oncogenic role in acute myeloid
788 leukemia as a N6-methyladenosine RNA demethylase. Cancer cell. 2017;31(1):127–41.

789 64. Huang H, Weng H, Sun W, Qin X, Shi H, Wu H, et al. Recognition of RNA N 6-methyladenosine by
790 IGF2BP proteins enhances mRNA stability and translation. Nature cell biology. 2018;20(3):285.

791 65. Zhong X, Yu J, Frazier K, Weng X, Li Y, Cham CM, et al. Circadian Clock Regulation of Hepatic Lipid
792 Metabolism by Modulation of m6A mRNA Methylation. Cell reports. 2018;25(7):1816–28.

793 66. Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millán-Zambrano G, Robson SC, et al. Promoter-bound
794 METTL3 maintains myeloid leukaemia by m 6 A-dependent translation control. Nature.
795 2017;552(7683):126.

796 67. SEQC/MAQC-III Consortium, Su Z, Łabaj PP, Li S, Thierry-Mieg J, Thierry-Mieg D, et al. A
797 comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the
798 Sequencing Quality Control Consortium. Nature Biotechnology. 2014 Aug 24;32:903.

799 68. Li M, Zhao X, Wang W, Shi H, Pan Q, Lu Z, et al. Ythdf2-mediated m 6 A mRNA clearance
800 modulates neural development in mice. Genome biology. 2018;19(1):69.

801 69. Wen J, Lv R, Ma H, Shen H, He C, Wang J, et al. Zc3h13 regulates nuclear RNA m 6 A methylation
802 and mouse embryonic stem cell self-renewal. Molecular cell. 2018;69(6):1028–38.

803 70. Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, et al. Histone H3 trimethylation at lysine 36
804 guides m 6 A RNA modification co-transcriptionally. Nature. 2019;1.

805 71. Zheng G, Dahl JA, Niu Y, Fedorcsak P, Huang C-M, Li CJ, et al. ALKBH5 is a mammalian RNA
806 demethylase that impacts RNA metabolism and mouse fertility. Molecular cell. 2013;49(1):18–29.

807 72. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-
808 seq aligner. Bioinformatics. 2013;29(1):15–21.

809 73. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, et al.
810 Topology of the human and mouse m 6 A RNA methylomes revealed by m 6 A-seq. Nature.
811 2012;485(7397):201.

812 74. Tirumuru N, Zhao BS, Lu W, Lu Z, He C, Wu L. N6-methyladenosine of HIV-1 RNA regulates viral
813 infection and HIV-1 Gag protein expression. Elife. 2016;5:e15528.

814 75. Rubio RM, Depledge DP, Bianco C, Thompson L, Mohr I. RNA m6 A modification enzymes shape
815 innate responses to DNA by regulating interferon β. Genes & development. 2018;32(23–24):1472–
816 84.

817 76. Winkler R, Gillis E, Lasman L, Safra M, Geula S, Soyris C, et al. m 6 A modification controls the
818 innate immune response to infection by targeting type I interferons. Nature immunology.
819 2019;20(2):173.

820 77. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of
821 best practices for RNA-seq data analysis. Genome Biology. 2016 Jan 26;17(1):13.

822 78. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, Yang Y, et al. N6-methyladenosine in nuclear RNA is a major
823 substrate of the obesity-associated FTO. Nature chemical biology. 2011;7(12):885.

824 79. Wei J, Liu F, Lu Z, Fei Q, Ai Y, He PC, et al. Differential m6A, m6Am, and m1A demethylation
825 mediated by FTO in the cell nucleus and cytoplasm. Molecular cell. 2018;71(6):973–85.

826 80. Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, et al. Reversible methylation of m 6 A m in
827 the 5′ cap controls mRNA stability. Nature. 2017;541(7637):371.

828 81. Aguilo F, Zhang F, Sancho A, Fidalgo M, Di Cecilia S, Vashisht A, et al. Coordination of m6A mRNA
829 methylation and gene transcription by ZFP217 regulates pluripotency and reprogramming. Cell stem
830 cell. 2015;17(6):689–704.

831 82. Chakrabarti AM, Haberman N, Praznik A, Luscombe NM, Ule J. Data science issues in studying
832 Protein–RNA interactions with CLIP technologies. Annual Review of Biomedical Data Science.
833 2018;1:235–61.

834 83. Liu B, Merriman DK, Choi SH, Schumacher MA, Plangger R, Kreutz C, et al. A potentially abundant
835 junctional RNA motif stabilized by m6A and Mg2+. Nature Communications. 2018 Jul 17;9(1):2761.

836 84. Saletore Y, Meyer K, Korlach J, Vilfan ID, Jaffrey S, Mason CE. The birth of the Epitranscriptome:
837 deciphering the function of RNA modifications. Genome Biology. 2012;13(10):175.

838 85. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, et al. Highly parallel direct RNA
839 sequencing on an array of nanopores. Nature Methods. 2018 Jan 15;15:201.

840 86. Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, et al. Accurate detection of m6A RNA
841 modifications in native RNA sequences. Nature Communications. 2019 Sep 9;10(1):4079.

842 87. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
843 Bioinformatics. 2014;30(15):2114–20.

88. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. Nature biotechnology. 2016;34(5):525.

89. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 2010;26(6):841–2.

90. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics. 2016;32(18):2847–9.

91. Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Research. 2019 Feb 20;47(8):e47–e47.

92. Reimand J, Arak T, Adler P, Kolberg L, Reisberg S, Peterson H, et al. g: Profiler—a web server for functional interpretation of gene lists (2016 update). Nucleic acids research. 2016;44(W1):W83–9.

93. Imam H, Khan M, Gokhale NS, McIntyre AB, Kim G-W, Jang JY, et al. N6-methyladenosine modification of hepatitis B virus RNA differentially regulates the viral life cycle. Proceedings of the National Academy of Sciences. 2018;115(35):8829–34.

94. Gokhale NS, McIntyre ABR, McFadden MJ, Roder AE, Kennedy EM, Gandara JA, et al. N6-methyladenosine in Flaviviridae viral RNA genomes regulates infection. Cell host & microbe. 2016 Nov;

Additional file 1: Supplementary Tables. (xls)

Additional file 2: Supplementary Figures. (pdf)

Additional file 3: Results with MeTDiff peak caller. (pdf)

Additional file 4: Peaks and statistical test results for experiments in Figures 3-4. (zip)

Additional file 5: Response to reviewers. (txt)