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
Christopher E. Mason chm2042@med.cornell.edu

Abstract

Recent studies have revealed that many cellular mRNAs contain the modified base m^{6}A and 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, we detected few changes under most conditions and were unable to detect consistent changes across studies using 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.

Introduction

Methylation at the *N*6 position in adenosine (m^6A) 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^6A 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^6A is lost and gained in response to various physiological changes remains contentious (3,10–15). To assess the evidence for proposed dynamic changes in m^6A, a reliable and reproducible method to detect changes in methylation as distinct from changes in gene expression is necessary.

The first and most widely-used method to enable transcriptome-wide studies of m^6A, MeRIP-seq or m^6A-seq, involves the immunoprecipitation of m^6A-modified RNA fragments followed by peak detection through comparison to background gene coverage (16,17). A second method has since been developed, miCLIP or m^6A-CLIP, which involves crosslinking at the site of antibody binding to induce mutations during reverse transcription for single-nucleotide detection of methylated bases (2,18). MeRIP-seq is still more often used than miCLIP, despite less precise localization of m^6A to peak regions of approximately 50-200 base pairs that can contain multiple DRAC motifs, since it follows a simpler protocol, requires less starting material, and generally produces higher coverage of more transcripts. Antibodies for m^6A can also detect a second base modification, N^6,2′-*O*-dimethyladenosine (m^6A_m), found at a lower abundance than m^6A and located at the 5′ ends of select transcripts (15,18). We thus refer to the base modifications detected through MeRIP-seq collectively as $m^6A_{(m)}$, although most are likely m^6A. As of late 2018, over fifty studies have used MeRIP-seq to detect $m^6A_{(m)}$ in mammalian mRNA (**Supplementary Table 1**).

Although MeRIP-seq can reveal approximate sites of $m^6A_{(m)}$, it cannot be used to quantitatively measure the fraction of transcript copies that are methylated (19). Therefore, studies of m^6A variation in response to stimuli estimate differences at individual loci through changes in peak presence or peak height. Using these approaches, studies have reported changes to m^6A with heat shock, microRNA expression, transcription factor expression, cancer, oxidative stress, human immunodeficiency virus (HIV) infection, Kaposi's sarcoma herpesvirus (KSHV) infection, and Zika virus infection, including hundreds to thousands of changes in enrichment at specific sites (20–29). Statistical approaches to analysis have only recently been published and there have been no comprehensive evaluations of methods to detect changes in m^6A based on MeRIP-seq data (30,31). Thus, while these studies could suggest that m^6A is highly variable in response to diverse stimuli, they have applied inconsistent analysis methods to detect changes in m^6A and often don't control for differences in RNA expression or typical variability in peak heights between replicates. In some cases, these studies have reported m^6A changes based on simple differences in peak count (24,26,27,32). However, others have applied statistical tests or thresholds for differences in immunoprecipitated (IP) over input fraction enrichment and visual analysis of coverage

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

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

Results

Detection of peaks across replicates, experiments, and cell types

The first steps in MeRIP-seq data analysis are to align sequencing reads to the genome or transcriptome of origin and to identify peaks in transcript coverage in the IP fraction relative to the input control. Several methods have been developed for MeRIP-seq peak detection, including exomePeak, MeTPeak, MeTDiff, and bespoke scripts. Another method often used for MeRIP-seq peak detection is MACS2, which was originally designed to detect protein binding sites in DNA from chromatin immunoprecipitation sequencing (ChIP-seq). We compared $m^6A_{(m)}$ peak detection by exomePeak, MeTPeak, MeTDiff, and MACS2 (31,47–49) in seven replicates of MeRIP-seq data obtained from mouse cortices under basal conditions (34), and in 12 replicates of MeRIP-seq data we generated from human liver Huh7 cells. The intersect between all tools tested was high and we saw minimal differences in DRAC motif enrichment (**Supplementary Figure 1a**). Since MACS2 is the most commonly used tool and was previously used to compare MeRIP-seq experimental methods (42), we used MACS2 for the remainder of our analyses.

We next defined the threshold of detection for peaks in MeRIP-seq data in terms of coverage. For $m^6A_{(m)}$ peak detection, a transcript must be sufficiently expressed for enrichment by $m^6A_{(m)}$ antibody and adequate sequencing coverage in both the IP and input fractions. Because previous reports have suggested that $m^6A_{(m)}$ presence does not decrease with expression level (9), we assume that the detection of fewer peaks as read counts at a gene or peak decrease indicates inadequate coverage. To estimate the level of coverage necessary for peak detection, we analyzed the percent of genes with at least one, two, or three peaks relative to mean input transcript coverage in both the mouse cortex and Huh7 cell data (**Figure 1a**). Based on the upper shoulders of the sigmoidal curves as the percent of genes with peaks begins to plateau, we estimate that gene coverage of approximately 10-50X is necessary to avoid missing peaks based on insufficient coverage. Input RNA-seq coverage of peak regions alone supports a similar threshold; few peaks are detected with median input read counts below 10 across replicates (**Supplementary Figure 1b**).

To evaluate the reproducibility of MeRIP-seq data, we next examined the consistency of $m^6A_{(m)}$ peak calling between replicates. Previous studies have reported that peak overlap between replicates is approximately 80% (9,16,50,51). Similarly, we found that between two replicates, $\log_2$ enrichment of IP over input reads at detected peaks showed a Pearson correlation of approximately 0.8 to 0.86 (**Supplementary Figure 1c**, top). A single replicate captured a median of 78% of the peaks found in seven replicates of mouse cortex data and 67% of peaks found in twelve replicates of Huh7 cell data. The number of detected peaks increased log-linearly with the addition of more replicates, such that with three replicates, 84-92% of the peaks found with 7-12 replicates were detected (**Figure 1b**, top). Conversely, the number of peaks in common across replicates decreased as the number of replicates increased, such that while ~80% of peaks were detected in at least two replicates, only ~60% were detected in six replicates for both data sets and ~30% in all twelve replicates of Huh7 cell data (**Figure 1b**, bottom). Detection of peaks in more replicates did not increase DRAC motif enrichment (**Supplementary Figure 1c**, bottom). These results suggest that many $m^6A_{(m)}$ sites may be missed in studies that use one to three replicates, and that increasing replicates could enable detection of more peaks. However, not all peaks correspond to true $m^6A_{(m)}$ sites. A recent reanalysis found that at least one published data set lacked enrichment for the canonical DRAC motif under MeRIP-seq peaks, suggesting a high ratio of false to true positives (3).

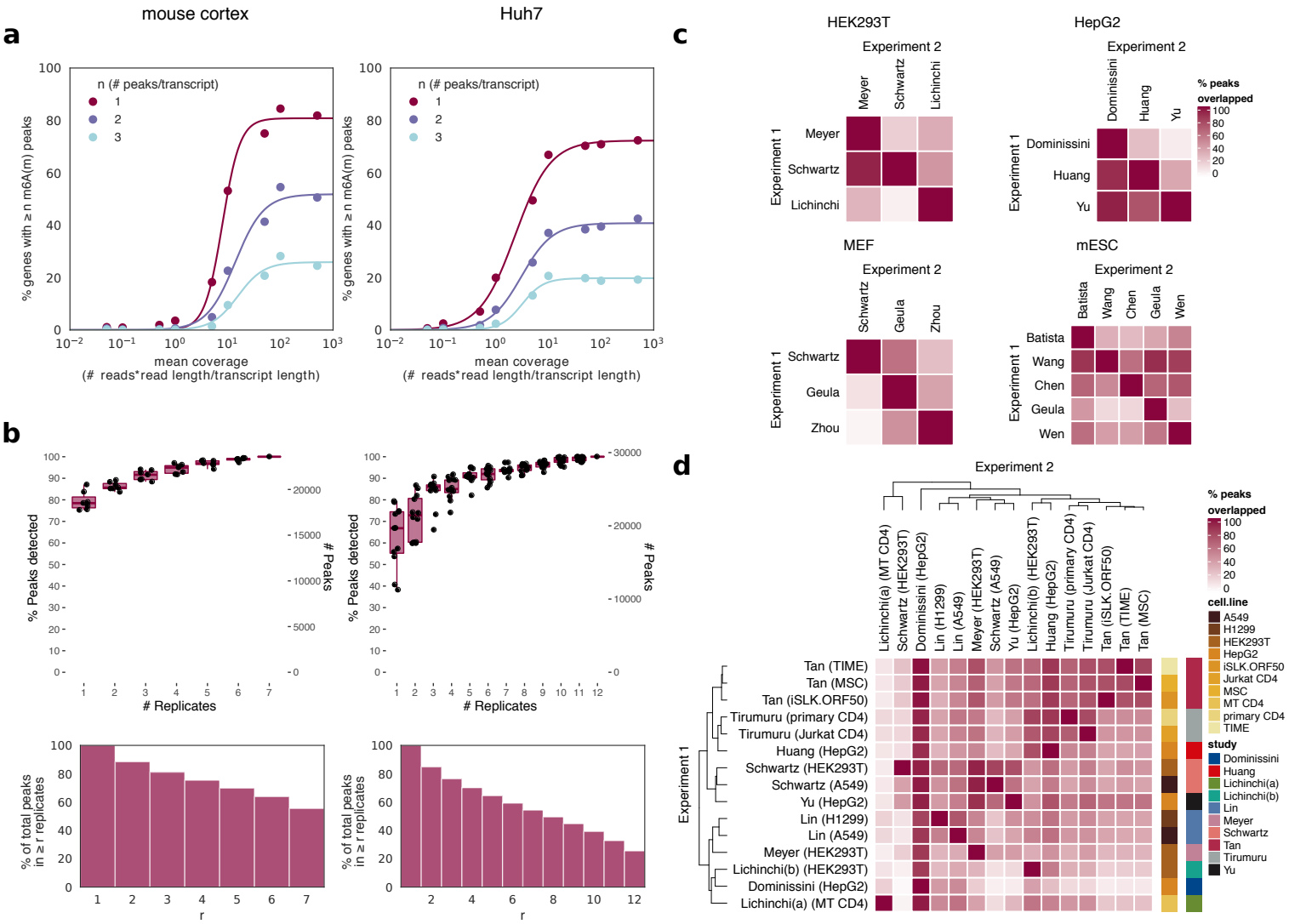


Figure 1: Thresholds and reproducibility of peak detection. a) $m^6A_{(m)}$ site detection in MeRIP-seq data from mouse cortex (left) and human liver cells (Huh7, right) shows saturation of peak detection as transcript coverage approaches 10-50X for replicates at basal conditions, with peaks merged from all replicates. **b)** The total number of peaks captured increases with more replicates, with single replicates capturing a median of 67-79% 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)** 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. **d)** Peak detection across tissue and cell types shows samples from the same study cluster better together than samples from the same tissue. Studies used in (c) and (d) are described in **Supplementary Table 2**.

138 The number of peaks detected across studies varies. Given that coverage affects peak detection,
139 we hypothesized that variation in sequencing depth could contribute to differences in peak count. Zeng et
140 al. (2018) reported that peak count begins to saturate by around 20 million reads by subsampling data

within individual studies(42). However, we found that there is no positive correlation between peak count and input or IP sequencing depth across data sets from different published studies, each of which had 3-81M reads per replicate (input Pearson's R = -0.37, p = 0.015; IP Pearson's R = -0.18, p = 0.24) (**Supplementary Table 2, Supplementary Figure 1d**). This implies that other experimental factors contribute to the variability of peak counts across studies.

We next analyzed the overlap of peaks among studies and found similar inconsistency in peak localization on transcripts. Within four commonly used cell types, the percent of peaks detected in one experiment that were also detected in a second varied among pairs of studies from as low as 2% of peaks to as high as 90%, filtering for transcripts expressed above a mean of 10X input coverage in both (**Figure 1c,** there were insufficient shared transcripts at 50X for most combinations to use that threshold). In fact, peaks showed higher overlap within different cell types from the same study than within the same cell type from different studies, suggesting that MeRIP-seq data is prone to strong batch effects (**Figure 1d**). We were unable to identify a link between peaks called and differences among experimental protocols used (summarized in **Supplementary Table 2**). Overall, most percent overlaps of $m^6A_{(m)}$ peaks fell between ~30% (1st quartile) and ~60% (3rd quartile). These results thus suggest that multiple labs running MeRIP-seq on the same cell type will not detect the same $m^6A_{(m)}$ sites.

Detection of changes in peaks between conditions

Following $m^6A_{(m)}$ peak detection, many studies seek to compare the expression of peaks between two conditions. Looking at plots of IP and input gene coverage under different conditions can help evaluate the evidence for peak changes (33), however, statistical or heuristic methods are necessary to narrow down a list of candidate sites to plot. Several tools used for statistical analysis by the studies in **Supplementary Table 1** or for other types of RNA IP sequencing assays model peak counts using either (a) the Poisson distribution, in which the variance of a measure (here, read counts) is assumed to be equal to the mean (MeTDiff), or (b) the negative binomial distribution, in which a second parameter allows for independent adjustment of mean and variance (QNB and two implementations of a generalized linear model approach using DESeq2 or edgeR, **Table 1**) (30,31,52–54). In the mouse cortex and Huh7 cell data, we found that, similar to RNA-seq data (24,53,55), the variance in read counts under peaks exceeded their mean, indicative of overdispersion (**Supplementary Figure 2a**). The log likelihood (the probability of an observation given a distribution with known parameters) for our sample also fell within the distribution of expected log likelihoods for the negative binomial distribution (bottom) but not the Poisson distribution (top) (**Figure 2a**). Thus, the negative binomial distribution captures the mean-variance relationship in MeRIP-seq data, suggesting that tools that account for overdispersion better model the distribution of read counts at $m^6A_{(m)}$ peaks than tools that do not.

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	

We next defined positive and negative controls to evaluate tool performance for detection of changes in $m^6A_{(m)}$ peaks. Past evaluations of methods have used data sets in which methylation machinery genes or the methyl donor were disrupted compared to baseline conditions as positive controls, and have simulated negative controls by randomly swapping labels in the positive controls (30,31). However, swapping labels for conditions that may feature differences in gene expression in addition to m^6A levels could unrealistically increase variance in read counts within groups. Therefore, we instead used the two data sets from mouse cortex 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 replicates for comparison and divided the Huh7 replicates by lab of incubation, which did not affect sample clustering (**Supplementary Figure 2b**). 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 (summarized in **Supplementary Table 3**).

When we compared the negative and positive controls, we found that the percent of peak changes called below a p-value threshold of 0.05 were similar (**Figure 2b**). With all tools except MeTDiff, a knockout of Mettl3 showed the largest effects on m^6A (56), while fewer significant peaks in other positive controls suggested variable effects of the positive control conditions on $m^6A_{(m)}$, possibly related to knockdown or overexpression efficiency (7,33,57–61). 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 (**Supplementary Figure 2c**) 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 2b**). By contrast, 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 (**Supplementary Figure 2c**), suggesting fewer false positives than MeTDiff.

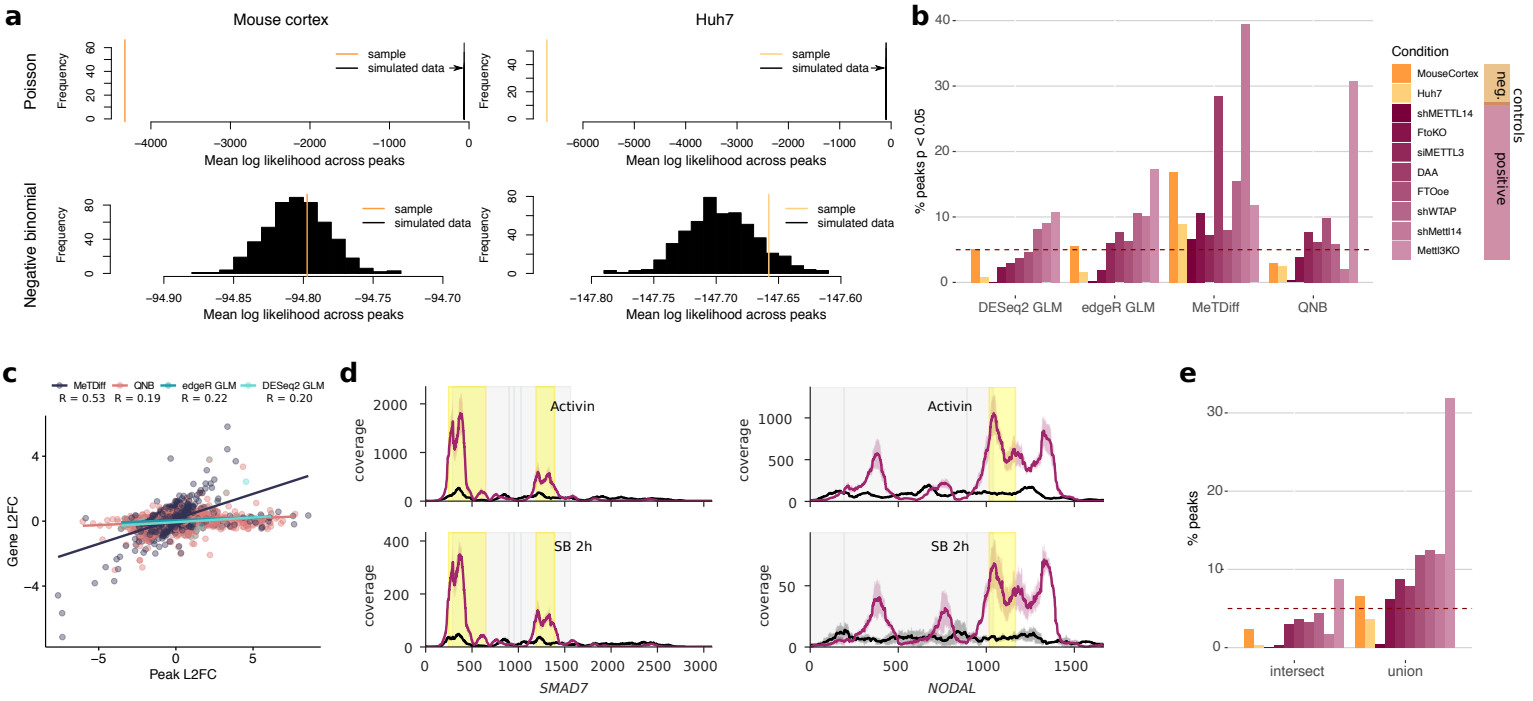


Figure 2: Analysis of methods to detect peak changes disproportional to gene expression changes. a) A comparison of Poisson (above) and negative binomial (below) models for read counts under peaks. The negative binomial mean log likelihood of the sample data fell within the 63rd and 91st percentiles 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 (**Supplementary Table 3**). The line at 5% indicates the expected proportion of sites given a uniform p-value distribution (see **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, 0.20, 0.53, and 0.19 for edgeR, DESeq2, MeTDiff, and QNB, respectively, with p = 7.0E-4, 0.10, 1.1E-71, and 2.9E-21. **d)** Coverage plots showing changes in peak expression are proportional to changes in gene expression for genes identified using MeTDiff as less methylated after two hours of treatment using an activin-NODAL inhibitor, SB431542 (SB), than with incubation in the presence of activin by Bertero et al. (2018). 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).

To ensure significant peaks detected by each of the tools reflected changes in IP enrichment independent of differential gene expression, we measured the correlation between changes in IP read counts at peak sites and changes in input read counts across their encompassing genes. For significant peaks (FDR-adjusted p-value < 0.05) from the positive controls, correlation between $\log_2$ fold change in

peak IP and gene input read counts was low for the GLMs and QNB (Pearson's R = 0.18 to 0.21, 2.9E-21 < p < 0.1) but reached 0.53 (p = 2.3E-74) for MeTDiff (**Figure 2c**). The higher correlation for MeTDiff was driven by peaks with proportional changes in IP and input levels, which suggests that MeTDiff often detects differential expression of methylated genes rather than differential methylation and is therefore of relevance for published studies that have used MeTDiff (22,62). Plotting coverage for genes reported as differentially methylated in one of these studies with the y-axis scaled separately per condition confirmed that changes in m^6A identified by MeTDiff were proportional to changes in gene expression (**Figure 2d**) (22). Given these results, QNB or the GLM implementations are better methods than MeTDiff to detect differential methylation. Taking the intersect of significant peaks for the GLMs and QNB may help determine the most probable sites, while taking the union of predictions provides a less conservative approach to selecting sites for further validation (**Figure 2e**). However, there were still significant peaks for which the difference between peak $\log_2$ fold change and gene $\log_2$ fold change was close to zero, particularly with QNB (**Supplementary Figure 2d**). For the remainder of our analyses, we therefore added a filter to the combined predictions from QNB and the two GLMs for difference in peak and gene $\log_2$ fold change ≥ 1, with an additional filter where noted for peak read counts ≥ 10 across all replicates and conditions to ensure sufficient coverage for peak detection.

Reanalyzing peak changes between conditions

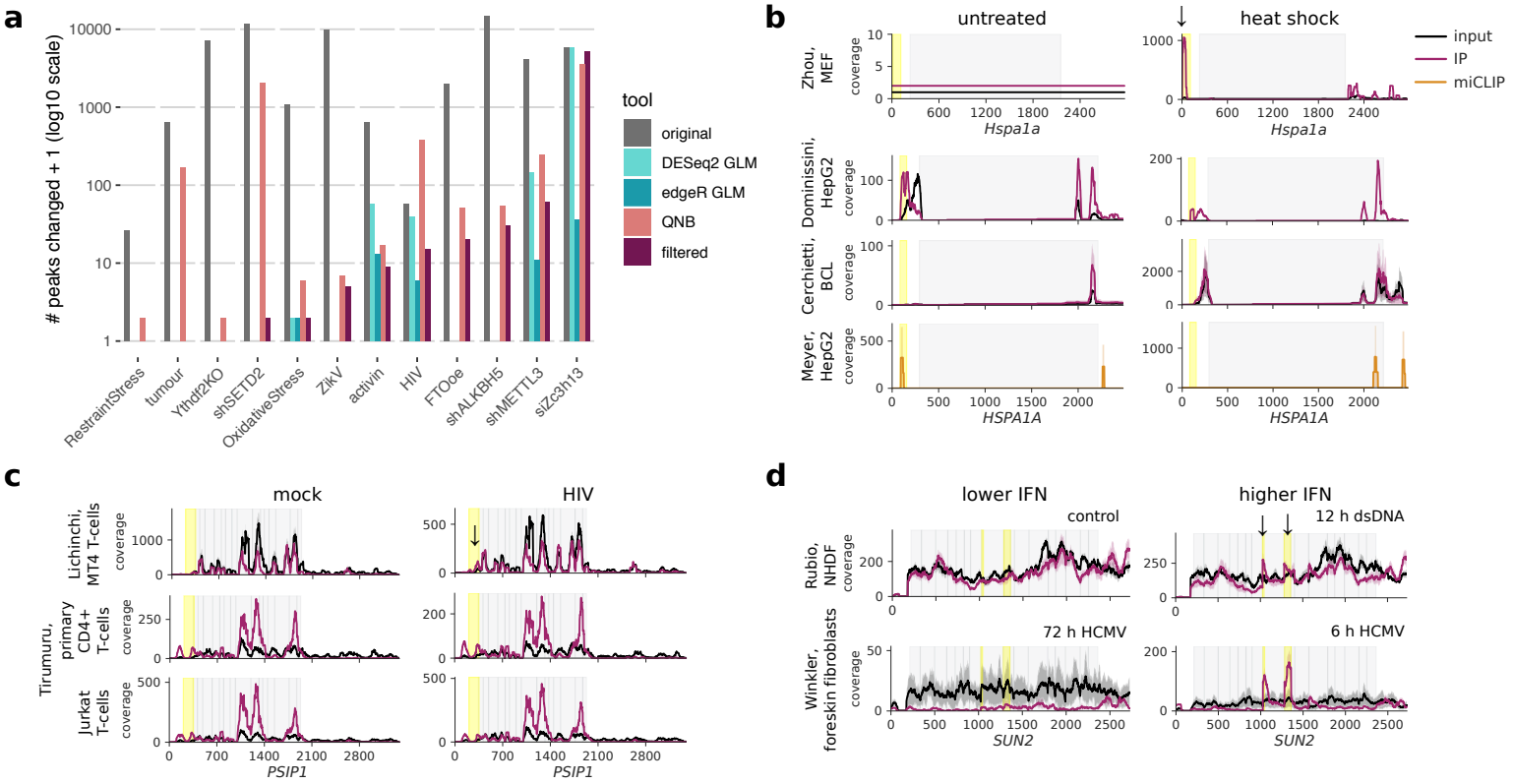
We next estimated the scale of statistically detectable peak changes under various conditions using our approaches and compared these results to previously reported estimates of these changes (**Figure 3a**, **Supplementary Table 4**). We identified fewer peaks as differentially methylated than originally 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 (22–26,34,59,62–66). Notably, knockdown of Zc3h13 did appear to disrupt $m^6A_{(m)}$, suggesting the gene does participate in methylation as recently suggested (64). 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 found 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. Without the thresholds for fold change and expression, the adjusted p-value for enrichment of the KEGG pathway "signaling pathways regulating pluripotency of stem cells" dropped to 0.15 based on three genes, LEFTY2, FZD28, and FGFR3 (**Supplementary Figure 3a**). 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, with 2064/2065 sites predicted by QNB falling below that threshold due to low input coverage in the first and second replicates (**Supplementary Figure 3b-e**) (65). 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 replicate per condition (40–42).

Multiple studies have investigated $m^6A_{(m)}$ in the context of heat shock, HIV infection, KSHV infection, and dsDNA treatment or human cytomegalovirus (HCMV) infection (**Supplementary Table 5**). Since each step in MeRIP-seq analysis risks introducing false negatives, we cannot rule out consistent changes between studies that used similar experimental interventions based on statistical detection alone. Therefore, we plotted coverage for specific genes reported as differentially methylated to evaluate reproducibility across these studies. Zhou, et al. (2015) reported 5′ UTR methylation of *Hspa1a* with heat shock (20). Coverage was too low for untreated controls to determine if *Hspa1a* was newly methylated or newly expressed with heat shock based on our alignment of their data using STAR (67). We were also unable to detect a change in methylation of *HSPA1A* using data from other heat shock studies, including a new data set from a B-cell lymphoma cell line and a published miCLIP data set, although coverage was again low (**Figure 3b**) (4). Lichinchi, et al. (2016) reported that 56 genes showed increased methylation with HIV infection in MT4 T-cells, with enrichment for genes involved in viral gene expression (25). Specific genes, for example *PSIP1*, in which we also detected a peak using MACS2 and see a change in the peak when plotting coverage using the data from Lichinchi et al. (2016), did not show the same changes in data from two other $CD4^+$ cell types, primary $CD4^+$ cells and Jurkat cells (**Figure 3c**) (68). Two other studies both used MeRIP-seq to detect m^6A in *IFNB1* induced through dsDNA treatment or infection by the dsDNA virus HCMV (69,70). The different treatments, time points and use of a fibroblast cell line versus primary foreskin fibroblasts make it difficult to compare $m^6A_{(m)}$ changes between the two experiments. Nevertheless, using QNB and the GLM approaches, we found five peaks in three genes (*AKAP8*, *SUN2*, and *TMEM140*) that showed significant changes both after 12 h of dsDNA treatment compared to untreated controls (69) and after 6 h post-HCMV infection compared to 72 h, when interferon levels have declined (70) (**Figure 3d**). Overall, we were unable to detect the same changes in $m^6A_{(m)}$ across studies of heat shock or HIV and few common changes in the response to dsDNA, but cell line-specific differences in $m^6A_{(m)}$ regulation and differences in experimental protocols could account for some of the variability among these studies.

While we did not have MeRIP-seq data for two studies from exactly the same conditions and cell lines to compare, two studies both used cell lines derived from iSLK to study the effects of KSHV on host m^6A (27,28). Both suggested that KSHV infection could decrease the number of m^6A sites in host transcripts. Hesser et al. (2018) found that lytic KSHV infection decreased the number of peaks on host transcripts by >25%; Tan et al. (2018) suggested a loss of 17-59% of peaks in two different cell types, but that $m^6A_{(m)}$ peak fold enrichment showed better clustering by cell type than by infection status. Neither discussed specific genes that showed differential methylation with lytic infection. For our comparison of $m^6A_{(m)}$ peak changes in these data sets, we identified probable changes in peaks based on statistical

283 significance using QNB or the GLMs with $\log_2$ fold change difference between peaks and genes of ≥1. We
284 detected 80 peak changes in the data from Hesser et al. (2018) and 18 in the data from Tan et al. (2018)
285 but found no peaks that changed in both iSLK data sets with lytic KSHV infection.



286

Figure 3: Changes in peaks between conditions. a) Detected $m^6A_{(m)}$ changes in ten published data sets that measured $m^6A_{(m)}$ peak changes between two conditions **(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. 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.

MeRIP-RT-qPCR validation

Although statistical approaches revealed fewer changes in $m^6A_{(m)}$ with various stimuli than published estimates, and we were unable to confirm changes in $m^6A_{(m)}$ methylation of specific genes across studies of similar conditions, many of the studies we looked at do include additional validation of $m^6A_{(m)}$ changes from MeRIP-seq using MeRIP-RT-qPCR. Recently it was shown that MeRIP-RT-qPCR can capture differences in m^6A:A ratios at specific sites (34), but it is unknown how MeRIP-RT-qPCR is affected by changes in gene expression. To test this, we ran MeRIP-RT-qPCR on in vitro transcribed RNA oligonucleotides that lacked or contained m^6A spiked into total RNA extracted from Huh7 cells (**Supplementary Table 6**). We found that MeRIP-RT-qPCR was able to detect the direction of change in m^6A levels at different spike-in concentrations (**Figure 4a-b**). However, technical variation could also lead to spuriously significant differences between two dilutions of in vitro controls with the same ratio of m^6A:A. For example, a comparison of m^6A enrichment between 30% methylated spike-ins at 0.1 fmol and 1 fmol returned a p-value of 0.004 (unpaired Student's *t*-test).

We next assessed the correlation between m^6A enrichment observed using MeRIP-seq and MeRIP-RT-qPCR using data from our recent work on changes in m^6A in Huh7 cells following infection by different viruses. For those experiments, we again selected peaks that change based on results from QNB and the GLM approaches. We found that the magnitude of changes in common among viruses correlated between MeRIP-seq and MeRIP-RT-qPCR, both across peaks (Pearson's R = 0.57, p = 3.7E-6) and within single peaks across viruses (13 out of 19 peaks showed positive correlations, four of which had p-values < 0.05 with three data points) (**Figure 4c, Supplementary Figure 4**). Given the correlation we found between MeRIP-seq and MeRIP-RT-qPCR, it is unclear why changes in IP over input sequencing reads are undetectable at the peaks reported by Bertero et al. (2018) and Huang et al. (2019) but differences in peaks were successfully validated using MeRIP-RT-qPCR (22,65). While MeRIP-RT-qPCR can be used as an initial method of validation for predicted peak changes, additional methods are necessary to confirm quantitative differences in m^6A levels and to resolve points where the assays do not agree.

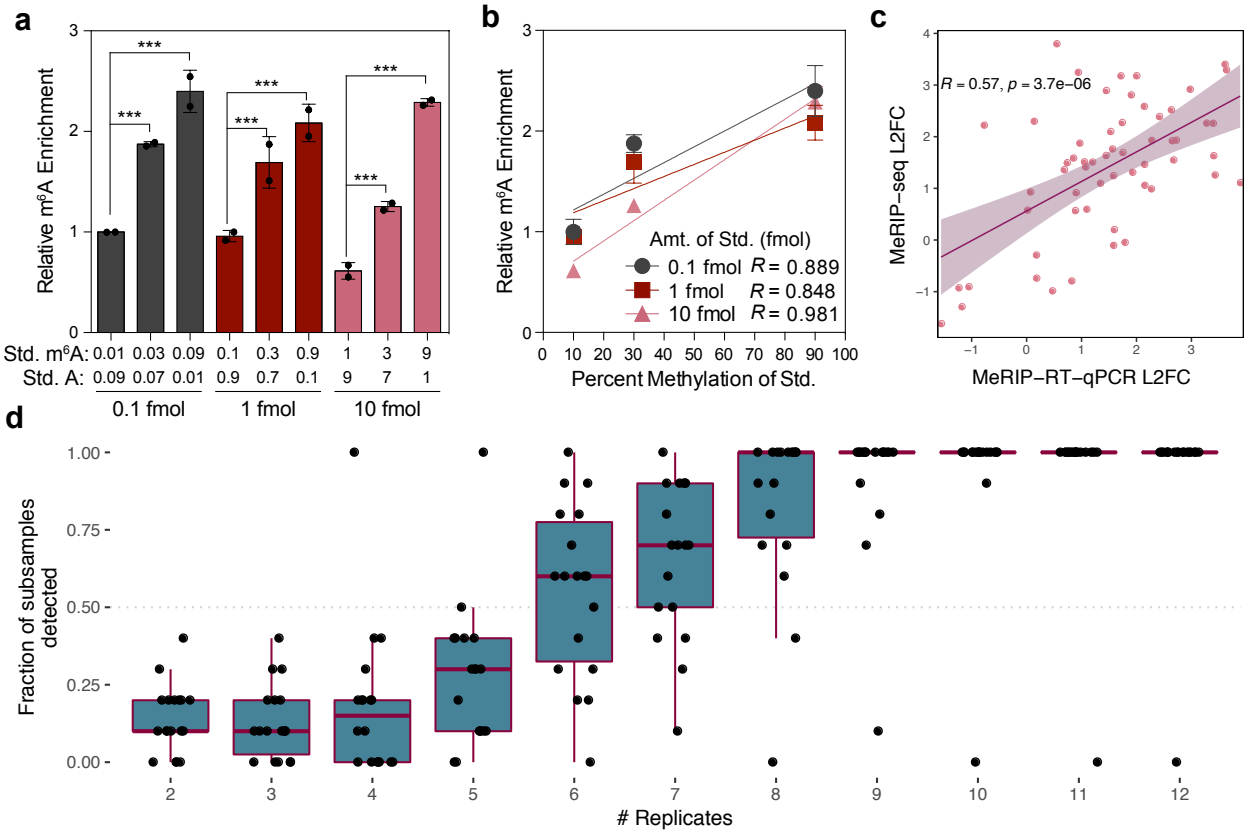


Figure 4: MeRIP-RT-qPCR validation and replicates necessary for the detection of peak changes. a) Relative enrichment of the indicated amounts of an in vitro transcribed standard containing unmodified A or m^6A, as measured by MeRIP-RT-qPCR. Data are shown for two independent replicates of three technical replicates each as IP enrichment over input relative to pulldown of a positive control spike-in, with the 0.1 fmol (0.01 m^6A: 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^6A 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.

We next used our peaks validated using MeRIP-RT-qPCR to estimate the number of replicates necessary for detection of changes with either the GLM or QNB methods. Using a permutation test, we downsampled infected and uninfected replicates and reran statistical detection of changes. We found that approximately 6-9 replicates were necessary for consistent detection (in at least 50% of subsamples) of

most peak changes (**Figure 4d**), suggesting that almost all published MeRIP-seq studies to date are underpowered.

Discussion

In the seven years since MeRIP-/m^{6}A-seq was first published (16,17), many studies have used these methods to examine the function of m^{6}A, its distribution along mRNA transcripts, and how it might be regulated under various conditions. While 35 out of 64 MeRIP- and miCLIP-seq papers we surveyed **(Supplementary Table 1)** refer to m^{6}A as "dynamic", and, by contrast, only two describe the modification as "static", the literature is unclear on what is meant by the word "dynamic". There is mixed evidence as to whether m^{6}A is reversible through demethylation by FTO and ALKBH5 (66,71–73). While m^{6}A does not appear to change over the course of an mRNA's lifetime at steady-state (3), whether it changes in response to a particular stimulus and at what point is less clear. Some studies have suggested that m^{6}A may be modulated through changes in methyltransferase and demethylase expression, producing consistent directions of change across transcripts (8,23,34), through alternative mechanisms involving microRNA, transcription factors, promoters, or histone marks (21,22,62,65,74), or through indeterminate mechanisms (17,20,25–28,50). However, based on our reanalysis of available MeRIP-seq data, there is still only meagre support for widespread changes in m^{6}A independent of changes in the expression of methylation machinery (e.g. increases or decreases in METTL3 expression).

In particular, replication of peaks and changes in peaks across studies is limited. As with other RNA IP-based methods, MeRIP-seq data contains noise, owing to technical and biological variation (75). In fact, while peak overlaps reach ~80% between replicates of the same study, they decrease to a median of 45% between studies, most of which use 2-3 replicates each **(Figure 1)**. Given that the detection of peaks is so variable and that peak heights differ among replicates, it is perhaps not surprising that peak changes have yet to be reproduced between multiple studies of similar conditions. Indeed, variability in MeRIP-seq could also mask differences in m^{6}A regulation among cell types, which have been described in mouse brains (34) and in cell lines exposed to KSHV (28). To distinguish biological and technical variation, it will therefore be particularly important to test if multiple groups using the same cell line and conditions can better reproduce changes in m^{6}A.

Disparities in the methods used to detect changes in m^{6}A$_{(m)}$ peaks also play a role in differing conclusions among studies. Here, we analyzed four statistical methods to detect changes in peaks and found that three of these methods showed uniform or conservatively shifted p-value distributions and were able to identify changes in m^{6}A$_{(m)}$ independent of changes in gene expression. We therefore suggest that these statistical methods, in combination with filters for input levels in both conditions and the difference in $\log_2$ fold change between peaks and genes, can be used to identify candidate m^{6}A$_{(m)}$ sites from MeRIP-seq data for further analysis and validation (**Figure 5)**. Based on our results, we do not recommend MeTDiff for the detection of peak changes as it does not control well for differences in gene expression

353 **(Figure 2)**. Similar to others (33), we found that plotting predicted m^{6}A changes was invaluable with
354 appropriate scaling for gene coverage to reveal changes proportional to gene expression. In addition,
355 plotting the standard deviation in transcript coverage can help assess typical variation in peak height
356 among replicates. We note that both differential methylation of a gene and methylation of a gene that is
357 differentially expressed could be important, but they should not be conflated when considering the role of
358 m^{6}A in transcript regulation.

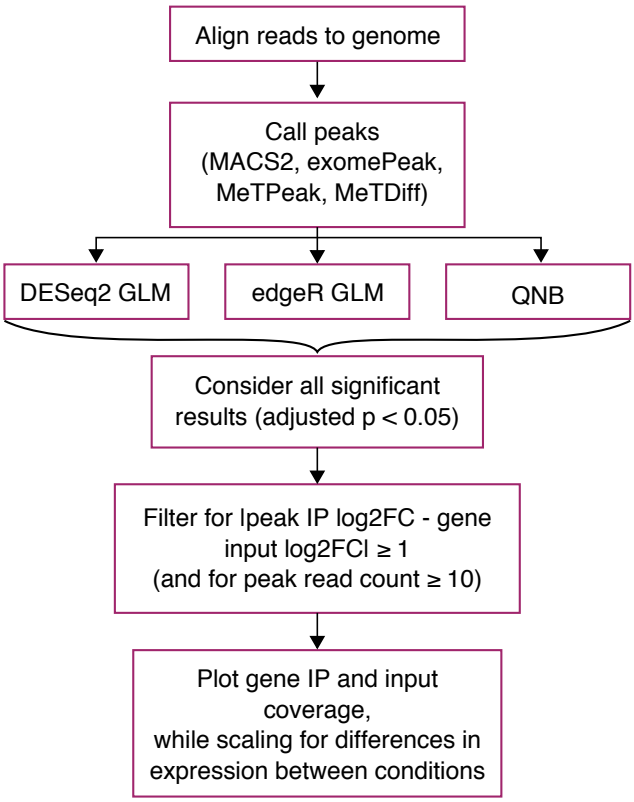


Figure 5: Proposed approach to identify candidates for m^{6}A$_{(m)}$ changes for further validation using MeRIP-seq data.

359

360 The extent to which m^{6}A changes on particular transcripts and whether it changes in binary
361 presence/absence or in degree is unclear. Although we found that MeRIP-RT-qPCR could detect
362 differences in in vitro transcribed RNA methylation and that these changes correlated with differences in
363 MeRIP-seq enrichment, neither MeRIP-seq nor MeRIP-RT-qPCR can reveal the precise fraction of
364 transcript copies modified by m^{6}A. In general, antibody-based methods are subject to biases, including
365 from differences in binding efficiencies based on RNA structure and motif preferences (76). There is an
366 oft-cited but little used antibody-independent method for site-specific quantification of m^{6}A, site-specific
367 cleavage and radioactive-labeling followed by ligation-assisted extraction and thin-layer chromatography
368 (SCARLET) (19). However, methods that can directly detect and quantify m^{6}A over the transcriptome are
369 still needed (e.g. direct RNA sequencing, which has not yet been shown to accurately detect m^{6}A across
370 a cellular transcriptome (46)). A recently developed endoribonuclease-based approach is promising but
371 limited to sites within DRAC motifs ending in ACA, which comprise only a third of known m^{6}A sites

(44,45). Thus, endoribonuclease digestion specific to unmodified strands may enable quantitative analyses of changes in m^6A:A ratios at ACA motif sites (44), but for now, site-specific SCARLET is the only option to biochemically validate proposed changes in m^6A at other motifs.

Conclusions

Our work reveals the limits of MeRIP-seq reproducibility for the detection of $m^6A_{(m)}$ and in particular suggests caution when using MeRIP-seq for the detection of changes in $m^6A_{(m)}$. To increase confidence in predicted changes in $m^6A_{(m)}$, we propose statistical approaches that account for differences in gene expression between conditions and variability among replicates. These methods can be used to gain insight into the regulation and function of $m^6A_{(m)}$ and to predict specific sites for validation before the development of high-throughput alternatives to MeRIP-seq, and similar strategies may be applicable to other types of RNA sequencing assay.

Methods

New MeRIP-seq data

- Huh7 data

Total RNA was extracted from Huh7 cells using Trizol (Thermo-Fisher). mRNA was purified from 200 µg total RNA using the Dynabeads mRNA purification kit (Thermo-Fisher) and concentrated by ethanol precipitation. Purified mRNA was fragmented using the RNA Fragmentation Reagent (Thermo-Fisher) for 15 minutes followed by ethanol precipitation. Then, MeRIP was performed using EpiMark N6-methyladenosine Enrichment kit (NEB). 25 µL Protein G Dynabeads (Thermo-Fisher) per sample were washed three times in MeRIP buffer (150 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40) and incubated with 1 µL anti-m^6A antibody (NEB) for 2 hours at 4°C with rotation. After washing three times, anti-m^6A conjugated beads were incubated with purified mRNA with rotation at 4°C overnight in 300 µL MeRIP buffer with 1 µL RNAse inhibitor (recombinant RNasein; Promega). Beads were then washed twice with 500 µL MeRIP buffer, twice with low salt wash buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40), twice with high salt wash buffer (500 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.1% NP-40), and once again with MeRIP buffer. m^6A-modified RNA was eluted twice in 100 µL MeRIP buffer containing 5mM m^6A salt (Santa Cruz Biotechnology) for 30 minutes at 4°C with rotation and concentrated by ethanol precipitation. RNA-seq libraries were prepared from eluate and the 10% of RNA set aside as input using the TruSeq mRNA library prep kit (Illumina) and checked for fragment length using the Agilent 2100 Bioanalyzer. Single-end 50 base pair reads were sequenced on an Illumina HiSeq 2500.

- Heat shock

Early passage OCI-Ly1 diffuse large B-cell lymphoma cells were grown in Iscove's modified Eagle Medium (IMDM) with 10% fetal bovine serum (FBS). OCI-Ly1 cells were obtained from the Ontario Cancer Institute and regularly tested for *Mycoplasma* contamination by PCR and identified by single

nucleotide polymorphism. Cells were maintained with 1% penicillin/streptomycin in a 37°C, 5% CO_2, humidified incubator. In these growing conditions, heat shocked cells were exposed to 43 °C for 1 hour, followed by 1 hour of recovery at 37°C while control cells were maintained at 37°C. Following treatment, cells were processed at 4°C to obtain total cell lysates. Lysates were immunoprecipitated for $m^6A_{(m)}$ using Synaptic Systems antibody (SYSY 202 003) following the protocol described in Meyer, et al (2012) and sequenced on an Illumina HiSeq 2500 (16).

Read processing

Reads were trimmed using Trimmomatic (77) and aligned to the human genome (hg38) or the mouse genome (mm10), as appropriate, using STAR (67).

Peak detection and comparison

IP over input peaks were called using MACS2 (49). Transcript coverage was estimated using Kallisto (78) with an index construct 31mers, except for the Schwartz et al (2014) data set, where the reads were too short and an alternative index based on 29mers was constructed (33). For **Figure 1b**, the full union of unique peaks was taken and the percent of that set detected in single replicates calculated. Intersects between peaks that overlapped for transcripts with ≥10X mean coverage in both samples were taken using bedtools (79) for **Figure 1c-d**, allowing a generous minimum of 1 overlapping base. Heatmaps for peak overlaps were generated using the ComplexHeatmap package in R (80). MeRIP-seq data sets in **Figure 1d** included those for human cell lines in **Figure 1c**, other data sets from the same studies and any data sets that shared the same cell lines, and other data sets that looked at multiple human cell types. We considered only data sets from baseline conditions in **Figure 1** (untreated cells and knockdown controls).

Poisson and negative binomial fits

Poisson and negative binomial models were fit to input and IP read counts at peaks using maximum likelihood estimation. Simulated read counts were generated with Poisson or negative binomial distributions based on estimated parameters from the sample, with 500 random generations per model. The log likelihood of seeing read counts from the sample and the simulations given the model parameters was then calculated and the mean taken across all peaks.

Peak change detection and generalized linear models

Generalized linear models to detect changes in IP coverage while controlling for differences in input coverage were implemented based on a method previously applied to HITS-CLIP data (54). Full and reduced models were constructed as follows:

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

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

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

Statistical significance was then assessed using a chi-squared test (df=1) for the difference in deviances between the full and reduced models, with the null hypothesis that the interaction term ($\beta_i^{STIM:IP}$) for differential antibody enrichment driven by the experimental intervention is zero. The likelihood ratio test was implemented through DESeq2 (52) and edgeR (53), two programs developed for RNA-seq analysis that differ in how they filter data and in how they estimate dispersions for negative binomial distributions. Generalized linear models implemented through edgeR included a term for the normalized library size of sample j.

QNB was run as suggested for experiments with biological replicates, where each IP and input variable ("ip1", etc.) consisted of a matrix of peak counts for either condition 1 or condition 2:

```
> qnbtest(ip1, ip2, input1, input2, mode="per-condition")
```

We extracted functions from MeTDiff so that we could supply our own peaks and thus control for differences in peak detection among tools. The main post-peak calling function, diff.call.module, was run as follows using the same count matrices as for QNB:

```
> diff.call.module(ip1, input1, ip2, input2)
```

Gene and peak expression changes were estimated as $\log_2$ fold changes from DESeq2 based on differences in input read counts aligned to genes and IP read counts aligned to peaks, respectively, and the change in peak relative to gene enrichment was calculated as the absolute difference in $\log_2$ fold change between those values.

Comparison to published studies

We selected data sets for reanalysis in **Figure 3a** based on the availability of ≥ 2 replicates and a published estimate of the number of m^6A changes between two conditions. The sources for published estimates of m^6A peak changes included in our comparison are listed in **Supplementary Table 4**. Significant (FDR-adjusted $p < 0.05$) peaks were considered for DESeq2, edgeR, and QNB, run as described above. We also considered a filtered set of peaks derived from the union of significant peaks from the three tools with additional filters for location within exons, $\log_2$ fold change between peak IP and

gene inputl ≥ 1, and a minimum peak read count of 10 across replicates and conditions. We used gProfiler to calculate enrichment of functional categories using a hypergeometric test (81).

In **Figure 3b-c**, we selected *Hspa1a/HSPA1A* as our representative gene for heat shock because it was the primary example cited by Zhou et al. (2015) and Meyer et al. (2015) (4,20). For HIV, we selected *PSIP1* because it was among the 56 genes reported by Lichinchi et al. (2016a) (25), it plays a known role in HIV infection, and we detected a peak in the gene using MACS2.

For KSHV, we compared significant results (adjusted $p < 0.05$) from QNB and GLMs (DESeq2 and edgeR), with additional filtering for lpeak IP – gene input $\log_2$ fold changel ≥ 1 (lowering this threshold to 0.5 did not change results), for data from Hesser et al. (2018) (27) in lytic vs. latent iSLK.219 cells and data from Tan et al. (2018) (28) in lytic vs. latent iSLK BAC16 cells. We used the same approach to compare data from Rubio et al. (2018) and Winkler et al. (2019) (69,70) for response to dsDNA. Data sets used for site-specific comparisons are summarized in **Supplementary Table 5**.

Gene coverage was plotted using CovFuzze (https://github.com/al-mcintyre/CovFuzze), which summarizes mean and standard deviation in coverage across available replicates.

Spike-in controls and MeRIP-RT-qPCR

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

IVT_Std_F: TGCCTTTTCTTTCGGTTGCG

IVT_Std_R: CAAACACAAGAAGGCACGGG

GLuc_F: CGACATTCCTGAGATTCCTGG

GLuc_R: TTGAGCAGGTCAGAACACTG

CLuc_F: GCTTCAACATCACCGTCATTG

CLuc_R: CACAGAGGCCAGAGATCATTC

Cell culture and infection (data used for MeRIP-RT-qPCR experiments)

Huh7 cells were grown in DMEM (Mediatech) supplemented with 10% fetal bovine serum (HyClone), 2.5 mM HEPES, and 1X non-essential amino acids (Thermo-Fisher). The identity of the Huh7 cell lines was verified using the Promega GenePrint STR kit (DNA Analysis Facility, Duke University), and cells were verified as mycoplasma free by the LookOut Mycoplasma PCR detection kit (Sigma). Infectious stocks of a cell culture-adapted strain of genotype 2A JFH1 HCV were generated and titered on Huh7.5 cells by focus-forming assay (FFA), as described (82). Dengue virus (DENV2-NGC), West Nile virus (WNV-NY2000), and Zika virus (ZIKV-PRVABC59) viral stocks were generated in C6/36 cells and titered on Vero cells as described (82). All viral infections were performed at a multiplicity of infection of 1 for 48 hours.

Availability of Data and Materials

MeRIP-seq data for the Huh7 negative controls is available in the GEO repository, under accession number GSE130891. MeRIP-seq data for heat shock in B-cell lymphoma is available under accession number GSE130892. Accession numbers for all other data sets reanalyzed in the study are included in Supplementary Tables 1-5. Scripts used for analyses are available at https://github.com/al-mcintyre/merip_reanalysis_scripts.

Competing interests

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

Funding

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

548 **Authors' contributions**

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

553 **Acknowledgements**

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

558 **References**

559 1. Balacco DL, Soller M. The m6A Writer: Rise of a Machine for Growing Tasks. Biochemistry. 2019;

560 2. Ke S, Alemu EA, Mertens C, Gantman EC, Fak JJ, Mele A, et al. A majority of m6A residues are in
561 the last exons, allowing the potential for 3′ UTR regulation. Genes & development. 2015;

562 3. Ke S, Pandya-Jones A, Saito Y, Fak JJ, Vågbø CB, Geula S, et al. m6A mRNA modifications are
563 deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic
564 turnover. Genes & Development. 2017;31(10):990–1006.

565 4. Meyer KD, Patil DP, Zhou J, Zinoviev A, Skabkin MA, Elemento O, et al. 5′ UTR m6A promotes cap-
566 independent translation. Cell. 2015;163(4):999–1010.

567 5. Wang X, Zhao BS, Roundtree IA, Lu Z, Han D, Ma H, et al. N6-methyladenosine modulates
568 messenger RNA translation efficiency. Cell. 2015;161(6):1388–99.

569 6. Lin S, Choe J, Du P, Triboulet R, Gregory RI. The m6A methyltransferase METTL3 promotes
570 translation in human cancer cells. Molecular cell. 2016;62(3):335–45.

571 7. Wang Y, Li Y, Toth JI, Petroski MD, Zhang Z, Zhao JC. N 6-methyladenosine modification
572 destabilizes developmental regulators in embryonic stem cells. Nature cell biology. 2014;16(2):191.

573 8. Molinie B, Wang J, Lim KS, Hillebrand R, Lu Z, Van Wittenberghe N, et al. m 6 A-LAIC-seq reveals
574 the census and complexity of the m 6 A epitranscriptome. Nature methods. 2016;13(8):692.

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

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

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

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

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

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

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

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

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

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

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

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

602 21. Chen T, Hao Y-J, Zhang Y, Li M-M, Wang M, Han W, et al. m6A RNA Methylation Is Regulated by
603 MicroRNAs and Promotes Reprogramming to Pluripotency. Cell Stem Cell. 2015 Mar 5;16(3):289–
604 301.

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

607 23. Liu J, Eckert MA, Harada BT, Liu S-M, Lu Z, Yu K, et al. m 6 A mRNA methylation regulates AKT
608 activity to promote the proliferation and tumorigenicity of endometrial cancer. Nature cell biology.
609 2018;20(9):1074.

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

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

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

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

619 28. Tan B, Liu H, Zhang S, da Silva SR, Zhang L, Meng J, et al. Viral and cellular N 6-methyladenosine
620 and N 6, 2′-O-dimethyladenosine epitranscriptomes in the KSHV life cycle. Nature microbiology.
621 2018;3(1):108.

622 29. Vu LP, Pickering BF, Cheng Y, Zaccara S, Nguyen D, Minuesa G, et al. The N 6-methyladenosine (m
623 6 A)-forming enzyme METTL3 controls myeloid differentiation of normal hematopoietic and
624 leukemia cells. Nature medicine. 2017;23(11):1369.

625 30. Liu L, Zhang S-W, Huang Y, Meng J. QNB: differential RNA methylation analysis for count-based
626 small-sample sequencing data with a quad-negative binomial model. BMC bioinformatics.
627 2017;18(1):387.

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

631 32. Zhao X, Yang Y, Sun B-F, Shi Y, Yang X, Xiao W, et al. FTO-dependent demethylation of N6-
632 methyladenosine regulates mRNA splicing and is required for adipogenesis. Cell research.
633 2014;24(12):1403.

634 33. Schwartz S, Mumbach MR, Jovanovic M, Wang T, Maciag K, Bushkin GG, et al. Perturbation of m6A
635 writers reveals two distinct classes of mRNA methylation at internal and 5′ sites. Cell reports.
636 2014;8(1):284–96.

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

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

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

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

647 38. Geula S, Moshitch-Moshkovitz S, Dominissini D, Mansour AA, Kol N, Salmon-Divon M, et al. m6A
648 mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. Science.
649 2015;347(6225):1002–6.

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

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

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

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

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

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

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

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

666 47. Meng J, Lu Z, Liu H, Zhang L, Zhang S, Chen Y, et al. A protocol for RNA methylation differential
667 analysis with MeRIP-Seq data and exomePeak R/Bioconductor package. Methods. 2014;69(3):274–
668 81.

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

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

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

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

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

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

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

684 55. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y. RNA-seq: an assessment of technical
685 reproducibility and comparison with gene expression arrays. Genome research. 2008;18(9):1509–
686 17.

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

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

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

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

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

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

700 62. Barbieri I, Tzelepis K, Pandolfini L, Shi J, Millán-Zambrano G, Robson SC, et al. Promoter-bound
701 METTL3 maintains myeloid leukaemia by m 6 A-dependent translation control. Nature.
702 2017;552(7683):126.

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

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

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

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

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

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

715 69. Rubio RM, Depledge DP, Bianco C, Thompson L, Mohr I. RNA m6 A modification enzymes shape
716 innate responses to DNA by regulating interferon β. Genes & development. 2018;32(23–24):1472–
717 84.

718 70. Winkler R, Gillis E, Lasman L, Safra M, Geula S, Soyris C, et al. m 6 A modification controls the
719 innate immune response to infection by targeting type I interferons. Nature immunology.
720 2019;20(2):173.

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

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

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

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

730 75. Chakrabarti AM, Haberman N, Praznik A, Luscombe NM, Ule J. Data science issues in studying
731 Protein–RNA interactions with CLIP technologies. Annual Review of Biomedical Data Science.
732 2018;1:235–61.

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

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

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

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

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

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

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

748