Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution

Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution

Dierks, David; Garcia-Campos, Miguel Angel; Uzonyi, Anna

Dierks, Garcia-Campos, M. A., Uzonyi, A., Safra, M., Edelheit, S., Rossi, A., Sideri, T., Varier, R. A., Brandis, A., Stelzer, Y., van Werven, F., Scherz-Shouval, R., & Schwartz, S. (2021). Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution. Nature Methods, 18(9), 1060–1067. https://doi.org/10.1038/s41592-021-01242-z

Published Version: https://doi.org/10.1038/s41592-021-01242-z

Multiplexed profiling facilitates robust m6A quantification at site, gene and sample resolution

David Dierks[1,5], Miguel Angel Garcia-Campos[1], Anna Uzonyi[1,4], Modi Safra[1], Sarit Edelheit[1], Alice Rossi[2], Theodora Sideri[2], Radhika A. Varier[2], Alexander Brandis[3], Yonatan Stelzer[4], Folkert van Werven[2], Ruth Scherz-Shouval[5] & Schraga Schwartz[1,6,*]

Affililations

[1] Department of Molecular Genetics, Weizmann Institute of Science, 7610001 Rehovot, Israel

[2] Cell Fate and Gene Regulation Laboratory, The Francis Crick Institute, 1 Midland Road, London, NW1 1AT, UK.

[3] Life Sciences Core Facilities, Weizmann Institute of Science, 7610001 Rehovot, Israel

[4] Department of Molecular Cell Biology, Weizmann Institute of Science, 7610001 Rehovot, Israel

[5] Department of Biomolecular Sciences, Weizmann Institute of Science, 7610001 Rehovot, Israel

[6] Lead Contact

* Correspondence: schwartz@weizmann.ac.il

Abstract

N6-methyladenosine (m6A) is the most prevalent modification of mRNA in mammals. To interrogate its functions and dynamics, there is a critical need to quantify m6A at three levels: *site*, *gene*, and *sample*. Current approaches address these needs in a limited manner. Here we develop m6A-seq2, relying on multiplexed m6A immunoprecipitation of barcoded pooled samples. m6A-seq2 allows a major increase in throughput, while dramatically reducing technical variability, requirements of input material, cost, and labor. m6A-seq2 is furthermore uniquely capable of providing sample-level relative quantitations of m6A, serving as an important, orthogonal alternative to mass spectrometry-based approaches. Finally, we develop a computational approach for gene-level quantitation of m6A. We demonstrate that using this metric, roughly 30% of the variability in RNA half-life in mouse embryonic stem cells can be explained, establishing m6A as a major driver of RNA stability. m6A-seq2 thus provides an experimental and analytic framework for dissecting m6A-based regulation at three critical resolutions.

Introduction

The most abundant modification in mammalian mRNA is N6-methyladenosine (m6A). Initially discovered nearly five decades ago [1,2], research into this modification was catalyzed with the advent of transcriptome-wide detection methodology, namely m6A-seq/m6A-meRIP [3,4], relying on immunoprecipitation of m6A-containing RNA fragments, followed by high-throughput sequencing. m6A has been implicated in most steps of mRNA fate, and the disruption of the m6A-methylation machinery has critical manifestations in diverse systems [5–13]. Yet, despite intensive research, important aspects pertaining to m6A regulation and functions remain intensely debated [14–17], to a large extent due to our limited technical toolkit for interrogating m6A dynamics and functions at a transcriptome-wide scale. Nearly all genomic techniques to date have focused on the identification of methylated *sites*. However, connecting m6A with functions and phenotypes requires an ability to quantify m6A at the levels of *genes* and *samples*. Specifically, the functional outcomes of m6A (e.g. impact on mRNA stability, translation or localization) require a metric quantifying relative abundance of m6A per *gene*. Conversely, quantitation of the abundance of m6A across *samples* is indispensable towards interrogating its regulation, dynamics and responses to genetic perturbations. These two aspects are currently both addressed in a very limited manner. For gene-level m6A estimates, most studies, with a single notable exception [18], have converged on employing a binary metric (m6A methylated or unmethylated) or the number of m6A sites gene methylation quantification. This approach has resulted in significant relationships between m6A and different functional readouts, albeit typically of small effect sizes. It remains unclear whether these small effect-sizes are due to biological or technical reasons, in which case an improved metric could reveal stronger associations. In parallel, we currently lack genomic methodologies for *sample* m6A estimates, and typically rely on immunoassays or liquid chromatography followed by mass-spectrometry (HPLC-MS), which cannot provide information at the site/gene levels and is susceptible to non-mRNA contaminants (e.g. tRNA, rRNA)[19].

Finally, our ability to quantitatively measure m6A also at the site level is limited. The dynamic range of m6A changes in human samples is likely to be low [20,21]. Hence, identification of changes in m6A level of specific sites between conditions hinges on

the absence of strong *technical* sources of noise. The need to immunoprecipitate each sample separately in m6A-seq leads to considerable levels of technical noise and batch effect giving rise to erroneous calling of differentially methylated sites. All the more so when studies are conducted with a low number of replicates[22], which is often the case given expenses and input requirements associated with current protocols. This underscores the need for an improved methodology, reducing the extent of technical variability between samples, and facilitating robust experimental designs.

Here we establish an experimental and analytic framework for dissecting m6A-based regulations at all three relevant resolutions. Specifically, we develop an experimental approach, m6A-seq2, allowing to robustly and cost-efficiently interrogate m6A at the site-, sample- and gene- level resolution. We anticipate that the combined experimental and analytical framework developed here will be of high utility for functional and mechanistic interrogation of this widespread modification.

Results

The main principle introduced in m6A-seq2 is that a single m6A-IP is performed on pooled RNA samples, instead of on a single sample at a time (Fig. 1a, upper part). This is achieved by ligating barcoded RNA adapters to fragmented RNA originating from up to 24 distinct samples, subsequent pooling, followed by a single anti-m6A immunoprecipitation conducted on the pooled samples. These steps are followed by pooled cDNA generation, ligation of an adapter to the cDNA, library amplification and sequencing. The assignment of each read to its original sample is achieved on the basis of the barcode sequence.

This protocol has several advantages over traditional m6A-seq: **First**, all m6A-IPs are conducted in a single tube, which could reduce technical variability **(Fig. 1a)**. **Second**, given the competition between different samples over a fixed amount of antibody, this protocol could potentially allow comparing *global* m6A levels across different samples. **Third**, the pooling of *n* samples together *prior* to the IP step allows reducing the amount of starting material by a factor of n, in the case of 24 samples

from ~1.2 µg per sample to ~50 ng. **Fourth**, this protocol significantly reduces the costs of library preparation, to roughly 14% of the cost in standard m6A-seq **(based on 24 samples, Table S2)**. **Fifth**, this protocol increases the scale at which m6A can be interrogated by a factor of n, permitting more robust experimental designs (e.g. more replicates, controls) (**Fig. 1a**).

m6A-seq2 reduces technical variability in m6A quantification

To directly assess whether m6A-seq2 reduces technical variability, we isolated RNA from an Ndt80Δ/Δ yeast strain that is genetically synchronized at meiotic prophase). Of note, yeast has very low levels of m6A under vegetative growth conditions, but induces m6A during meiosis, peaking at the prophase stage [27–29]. M6A sites under these conditions have been well characterized and validated using orthogonal techniques [20,27]. We then conducted two sets of experiments: (1) we applied the standard m6A-seq protocol to 3 technical replicates beginning with 1.2 µg/sample, and (2) we applied m6A-seq2 to 12 technical replicates of 100 ng/sample.

Despite the order of magnitude difference in RNA starting material, the sequenced samples show similar complexity **(Supplementary Fig. 1a)**, similar enrichment of methylated targets **(Supplementary Fig. 1b)**, similar enrichment towards the 3' end **(Supplementary Fig. 1c),** and similar enrichment of the m6A consensus motif in a *de novo* peak calling analysis **(Supplementary Fig. 1d)**. We furthermore found that all barcodes gave rise to more or less similar numbers of reads per library **(Supplementary Fig. 1e)** and that the distribution of the 3'-most base in the library (undergoing the ligation with the adapter) was uniform across all adapters (**Supplementary Fig. 1f**), ruling out sequence biases due to ligation of distinct RNA adapters [30,31].

Figure 1

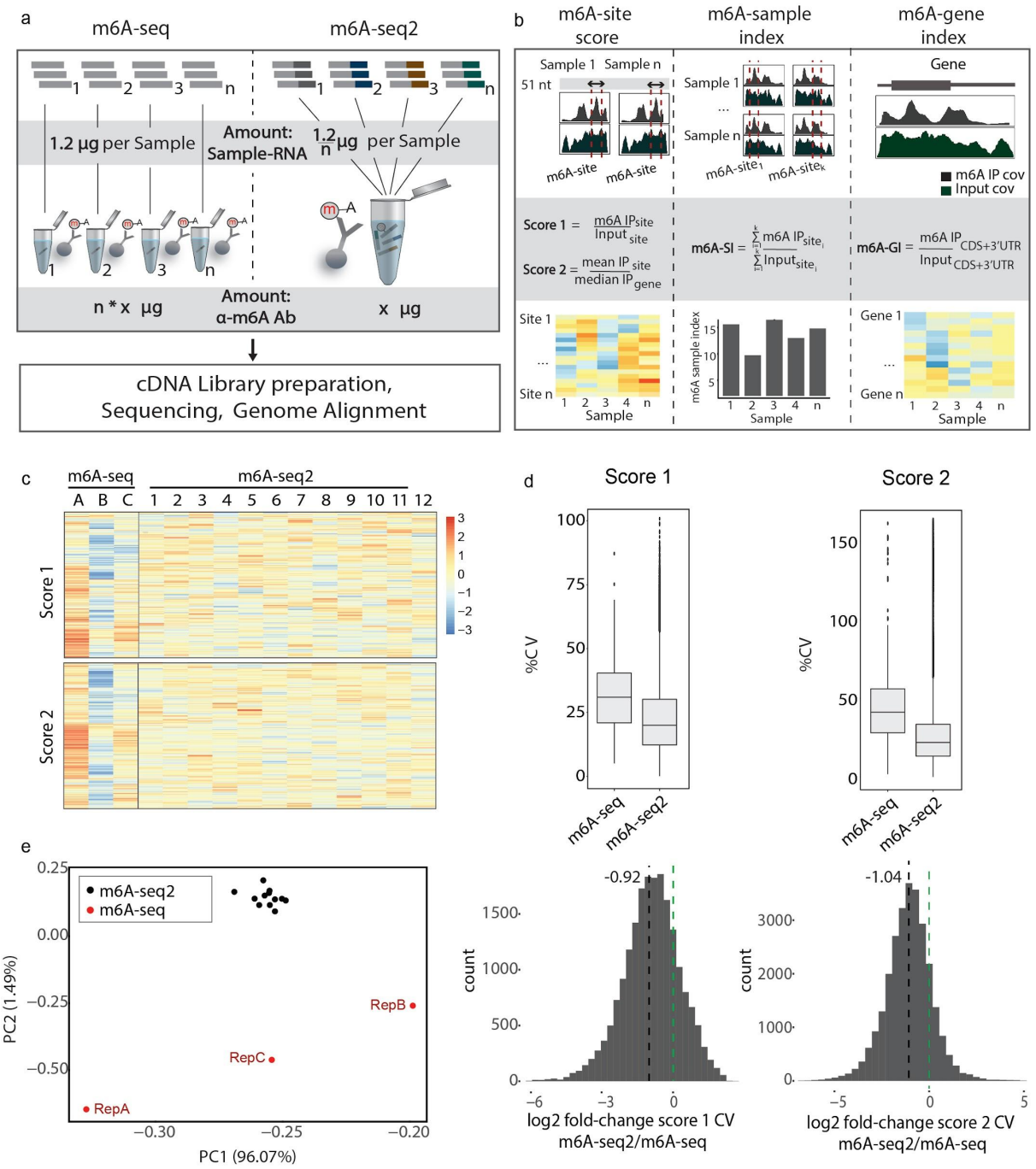


Figure 1. m6A-seq2 reduces batch-induced variability between replicates. **a)** Schematic comparison of m6A-seq versus m6A-seq2 experimental approach. **b)** Schematic overview of the m6A quantification approaches on three resolutions. **c)** m6A-site score comparison between technical replicates generated via m6A-seq (n=3) or m6A-seq2 (n=12). (Top) Heatmap of score1 for 489 m6A-sites. (Bottom) score2 for 577 m6A-sites passing the coverage thresholds for high-confidence m6A site estimation (see methods). **d) Top**: Comparison of the percentage coefficient of variation (%CV) of score1 (left) and score2 (right) estimates across the triplicates of samples measured via m6A-seq, and randomly selected triplicates (from the 12 replicates) measured via m6A-seq2. Boxplot: Center is the median, lower and upper hinges depict the first and third quartile and the whiskers stretch to 1.5 times the inter-quartile range from the corresponding hinge. **Bottom**: Histogram of the log2 fold-change of m6A-seq2 to m6A-seq triplicate

coefficient of variation (CV) of the m6A-site score1 & 2. Mean log2 fold-change CV indicated with black dashed line. **e)** Principal component analysis of Score1 of 486 m6A - sites

We next quantified m6A intensities across a set of 486 previously detected high-confidence m6A sites [20,27], using two previously-established metrics [21]: (1) an input-normalized score ('score 1'), in which the sequence coverage of the IP sample is normalized by the corresponding value in the input sample, and (2) a background-normalized score ('score 2') in which the coverage at the site in the IP sample is normalized by median IP coverage levels of the gene **(Fig. 1b)**. For both sets of quantitations, m6A-seq2 results showed substantially reduced variability **(Fig. 1c & 1d)**, with roughly 50% reduction in the coefficient of variation (CV) **(Fig. 1d, lower panel)**. This reduced variability is captured also in a principal component analysis of m6A-site score1 **(Fig. 1e)** and score2 **(Supplementary Fig. 1g)**. Such reduced variability should both increase the statistical power to detect truly differentially modified sites, and reduce the number of sites erroneously detected as being differentially modified between two conditions. The latter aspect is explored in a bootstrapping-based scheme, revealing that m6A-seq2 is associated with a dramatic reduction in erroneously detected differentially modified sites (**Supplementary Fig. 1h**). Erroneous calling of differentially modified sites has been the source of numerous misinterpretations in the field[22], and can thus be considerably mitigated by m6A-seq2.

To further confirm these results in a mammalian transcriptome, we isolated mRNA from mouse embryonic fibroblasts (mEFs), and split it into 10 aliquots. Five of these were subjected to traditional m6A-seq, and the others were subjected to m6A-seq2. We confirmed that also in mammalian contexts, the complexity of the m6A-seq2 libraries was similar compared to those obtained using the traditional m6A-seq protocol **(Supplementary Fig. 2a)**. m6A-seq2 allowed similar *de novo* identification of m6A sequence motifs (**Supplementary Fig. 2b**). Also metagene profiles showed a comparable enrichment around the stop codon (**Supplementary Fig. 2c**). m6A-seq2 also resulted in a significant reduction of technical variability and therefore to decreased liability of falsely calling a site as 'differentially modified' **(Supplementary Fig. 2d-g)**.

m6A-seq2 enables estimation of global m6A levels across samples

The standard m6A-seq protocol is limited in its ability to estimate the m6A level of a sample. If sample A contains twice the amount of m6A than sample B, the *relative* extent to which m6A is enriched in two separate m6A-IPs will not necessarily reflect this two-fold enrichment. In contrast, in m6A-seq2 samples A and B 'compete' over a fixed amount of antibody, in which case sample A should be immunoprecipitated twice as efficiently as B. We thus explored whether m6A-seq2 could allow estimating an 'm6A index' of the relative amounts of m6A in each sample. We first explored this question in the context of yeast meiosis. We applied m6A-seq2 to 12 samples, comprising 5 time points across meiosis sampled from two biological replicates, in addition to an Ndt80Δ/Δ positive control sample (PC) and an Ime4Δ/Δ/Ndt80Δ/Δ negative control (NC) lacking m6A [27,28]. Indeed, a crude readout of the number of reads of the m6A-IP divided by the corresponding number in the input sample, mirrored the expected accumulation of m6A during meiosis **(Fig. 2a)**. Of note, in a standard m6A-seq experiment, the raw number of reads is entirely technical and determined by the user-defined depth of sequencing. In contrast, in m6A-seq2, samples harboring higher levels of m6A will outcompete samples harboring lower m6A levels in their ability to bind the anti-m6A antibody, whereas all samples will be equally represented in the input samples, and hence the ratio is informative. Rather than quantify raw read-numbers, we next sought to define a more specific 'm6A sample index' (m6A-SI), quantifying the coverage over a predefined set of m6A sites in the IP sample, divided by the corresponding coverage in the input sample **(Fig. 2b)**. The m6A-SI displayed an increased dynamic range along the meiosis time course, as well as reduced variability in comparison to the read number fold-change, and captured the anticipated accumulation of m6A during meiosis. The m6A-SI strongly correlated with the HPLC-MS-based m6A level **(Fig. 2c & 2d)**.

For analysis at the site level, relatively high sequencing depth is required; However, the m6A-SI could, in principle, rely on much more shallow sequencing depth, given that it relies on the aggregation of reads from across hundreds of different sites [33]. To directly explore this, we subsampled sequencing reads from the meiosis time course at varying depths and assessed the correlation between the subsampled m6A-SIs per sample and their counterparts based on full coverage. We found that

subsampling down to 10% of the original read number didn't impact the m6A-SI **(Fig. 2e)**. At this depth, there are on average about 132,000 IP and input reads per sample, dramatically lower than is typically acquired for site-level analysis. m6A-seq2 thus not only reduces the library preparation cost but also requires an order of magnitude reduced sequencing depth for sample-level quantifications.

The scale at which m6A-seq2 can be applied also allowed us to perform a genetic screen in yeast for genes potentially involved in RNA methylation. We selected 16 genes involved in RNA metabolism **(Methods & Material Table 2)**, which we either fully deleted or inducibly depleted via auxin-inducible degrons. We applied m6A-seq2 to a total of 41 samples, isolated from meiotic prophase, either following induction of the degron (when applicable) or in the absence of such induction, in 4 batches. We observed varying levels of methylation **(Supplementary Fig. 2a)** that correlated strongly (Spearman ρ=0.71) with the expression of a 'meiotic prophase' signature set of genes **(Fig. 2f),** suggesting that the efficiency with which the different strains progressed throughout meiosis varied considerably between the strains. This strong correlation suggests that the observed differences in m6A levels primarily reflect differences in the kinetics and/or efficiency of sporulation between the different strains which *indirectly* impact m6A levels. Our ability to observe such a clear correlation across dozens of samples underscores the scalability of m6A-seq2, the quality of its output, and its utility for genetic screens.

Figure 2

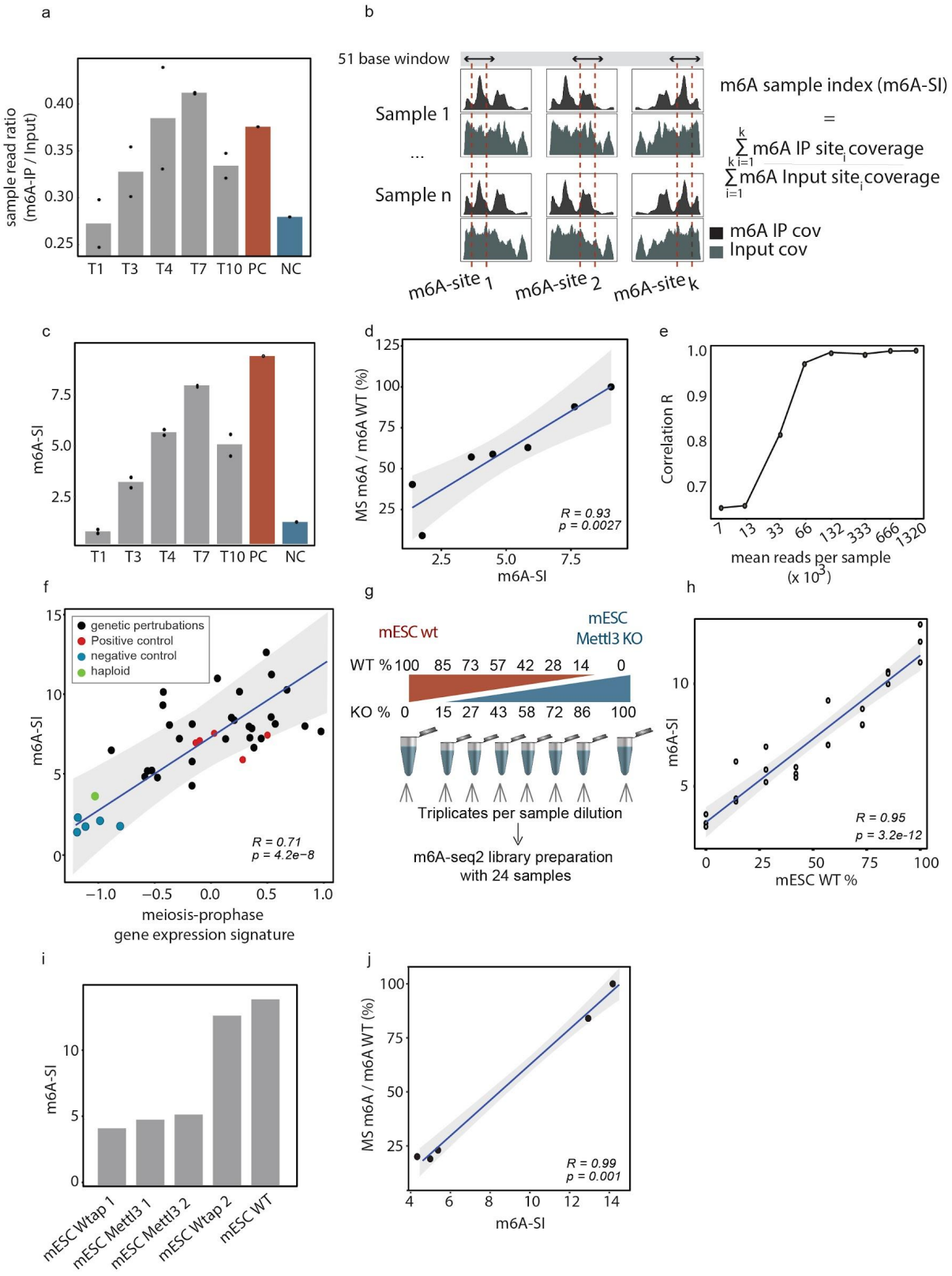


Figure 2. Development of an m6A sample index. **a)** Mean ratio of read-numbers (the number of reads in m6A-IP divided by the corresponding number in input) of a sample (time point in h) from the m6A-seq2 experiment of the yeast meiosis time course, barplot represents the mean of biological replicates n = 2 **b)** Scheme of m6A sample

index (m6A-SI) calculation. **c)** m6A-SI for meiosis time course samples. Biological duplicates for each time point (h) as well as a positive and negative control. As in a) , barplot represents the mean of biological replicates n = 2 **d)** m6A-SI across a meiosis time course plotted against mass-spec derived m6A levels [20]. Annotated Pearson's R coefficient and p-value and 95 % confidence interval **e)** Correlation between m6A-SIs using all available reads in comparison to counterparts estimated on the basis of down-sampling of read-number. The x-axis shows the average depth per sample. **f)** m6A-SIs plotted against a gene-expression derived index on the basis of a set of genes induced at meiosis prophase (see Methods). Annotated Pearson's R coefficient and p-value and 95 % confidence interval **g)** Experimental design of the 24 sample m6A-seq2 mESC experiment. Indicated are the 8 samples with varying amounts (rounded %) of Mettl3 KO (green) and poly-A selected RNA from WT mES cells (red). These sample dilutions were used for 3 technical replicates each (8 samples times 3 technical replicates) and applied to m6A-seq2 **h)** Scatterplot of the m6A-SI of 24 mESC samples and the relative amount of mESC WT derived RNA compared to Mettl3 KO. Annotated Pearson's R coefficient and p-value and 95 % confidence interval **i)** m6A-SIs in mESCs for WT cells and clones with partial knockouts of METTL3 and WTAP. **j)** Relative mass-spec m6A abundances plotted against m6A-SIs for the WT and methylation-writer perturbed mESCs. Annotated Pearson's R coefficient and p-value and 95 % confidence interval

To explore the applicability of sample-level quantitation in mammalian contexts, we generated 8 mixes of mESC WT and METTL3 KO mRNA at ratios spanning the range from 0% WT (0 m6A) to 100% WT (corresponding to highest levels), with triplicate measurements obtained for each sample dilution (**Fig. 2g**). m6A-seq2 was applied to these 24 samples, and subjected to shallow sequencing. At the peak level, an analysis of a subset of annotated high-confidence m6A-sites showed that the peak intensities at these sites were enriched proportionally to the amount of mESC WT mRNA **(Supplementary Fig. 3b-c)**. At the sample level, we observed an excellent agreement between the m6A-SI and between the relative ratios at which the WT and KO samples had been mixed (**Fig. 2h**). Next, we utilized m6A-seq2, to validate the elimination of m6A within mESC clones in which we had targeted m6A writers (Mettl3, Wtap) with a CRISPR approach. From each perturbation we chose two clones which appeared to lack the targeted exon based on Sanger sequencing. We conducted m6A-seq2 to all clones and in parallel analyzed the poly(A) mRNA of all samples with HPLC-MS. Both techniques indicated that none of the samples showed a full loss of methylation. While the extent of m6A depletion varied from one clone to another, it was highly correlated based on the two measurements **(Pearson R=0.99, P=0.001, Fig. 2i-j).** These data thus support the extension of m6A indexes to mammalian samples and validate a strong agreement with orthogonal, HPLC-MS

derived data, even at a shallow sequencing depth. These data further demonstrate the difficulties in achieving a full knockout of the methylation machinery components, consistent with previous attempts that only led to a partial reduction of methylation [12].

Development of an m6A gene index

Estimation of the overall m6A levels per gene is essential for connecting m6A to the wide range of functional readouts that are acquired at the gene level (e.g. stability, localization, translation). The conceptual framework introduced by m6A-seq2 of aggregating reads from across entire samples to estimate *sample* methylation levels motivated us to explore whether aggregation of reads from entire *genes* could provide a better estimate for gene-level methylation than approaches used to date, which have relied primarily on counting of peaks within genes [9,17,21,34–38]. Accordingly, we developed an m6A gene index **(m6A-GI)**, defined as the ratio of the number of reads overlapping a gene in m6A-IP divided by the corresponding value in the input sample **(Fig. 3a)**.

To calibrate and assess the validity of the m6A-GI, we relied on the *functional* propensity of m6A to enhance mRNA turnover. We utilized previously published measurements of mRNA stability in WT or METTL3 KO mESCs **(Material & Methods Table 1)**, based on which we calculated for each gene a fold-change (FC) half-life (WT/KO). Correlation of FC-half-life with the *number* of m6A sites recapitulated a low, negative correlation (R=-0.17), thus explaining only ~2.9% of the FC in half-life between WT and METTL3 KO cells **(Fig. 3c)**. In stark contrast, the m6A-GI established a dramatically stronger correlation (R=-0.55) **(Fig. 3c-d)**, thus explaining ~30% of the variability in mRNA stability. Moreover, even when correlating the m6A-GI directly against mRNA half-life we observed a similar negative correlation in mESC WT samples (R=-0.54) **(Fig. 3e).** Critically, the associations of the m6A-GI with FC-half-life and with half-life are both abolished in METTL3 KO mESCs **(Fig. 3d-e, Supplementary Figure 4a)**. The m6A-GI thus indicates that rather than being a minor contributor to stability, m6A in mESCs is a major force shaping mRNA half life.

To further assess the performance of the m6A-GI, we obtained m6A-LAIC-seq derived estimates of m6A abundances. m6A-LAIC-seq, which is based on subjecting samples to an m6A-IP without prior fragmentation of the RNA [18], is the only approach employed to date to derive gene-level estimates of m6A. We observed a strong correlation (R=0.84) between LAIC-seq-derived estimates of m6A in H1 hES cells and the m6A-GI derived from A549 cells m6A-seq data **(Fig. 3b-c)**.

Figure 3

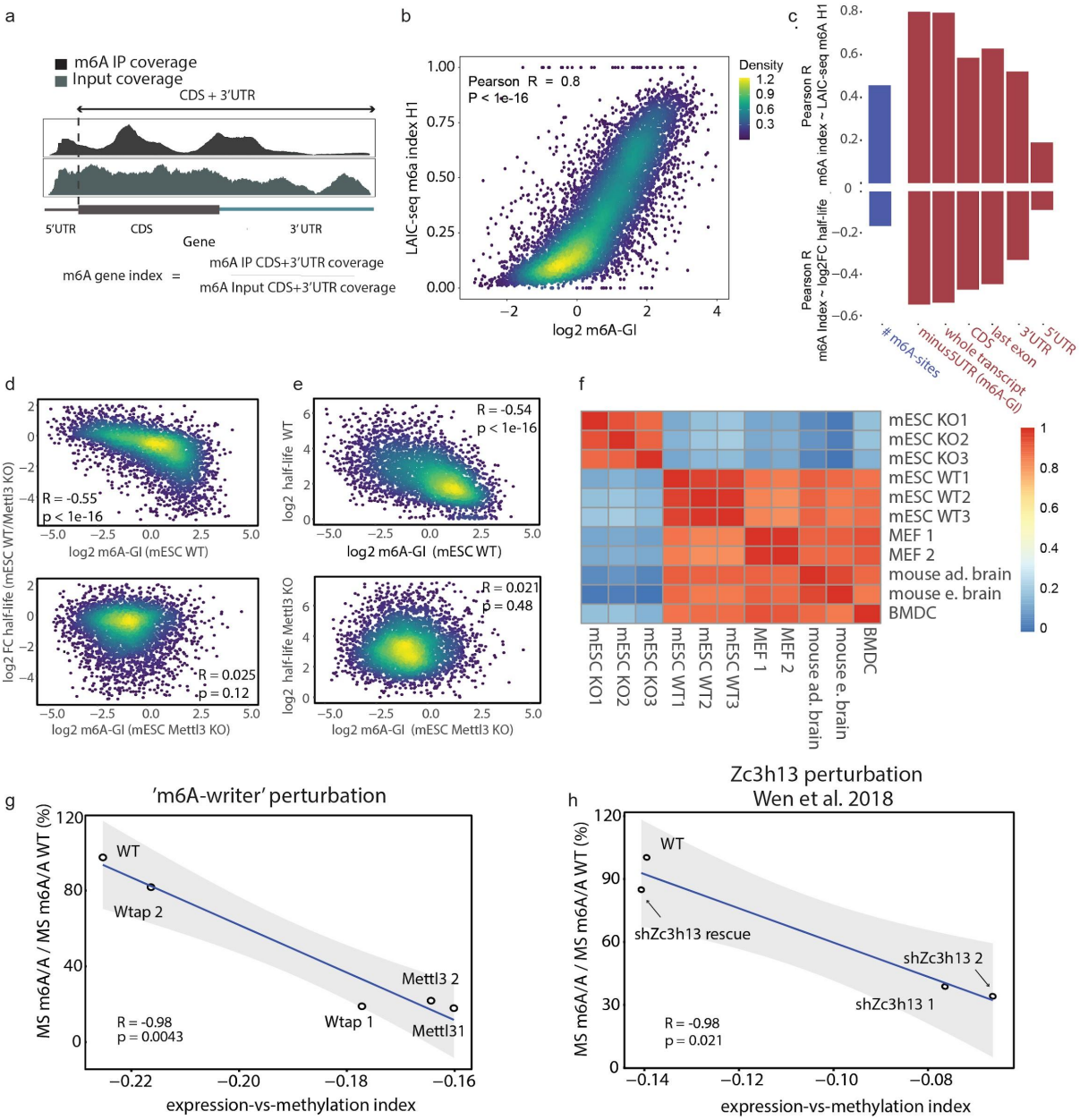


Figure 3. Development of an m6A-gene index. **a)** Scheme of the m6A-gene index (m6A-GI) **b)** Scatter plot of the LAIC-seq based and reported m6A-index of H1 hES cells with the log2 transformed m6A-GI derived from a published A549 m6A-seq dataset. Pearson's R and p-Value are annotated[21] **c)** Barplot of Pearson correlation

coefficients (*R*) (Y-axis) between tested m6A-indexes based on various transcript regions of A549 m6A-seq dataset and the LAIC-seq m6A-gene estimates of H1 hES cells (upper panel). Pearson correlation coefficients (*R*) (Y-axis) between measured m6A-indexes of various transcript regions based on mESC m6A-seq dataset and Mettl3-dependent half-life estimates (log2 fold-change WT/ Mettl3 KO half-lives), (lower panel) (**Supplementary Table 3**). (1) Calculation of read-number enrichment (IP versus input) in diverse regions within the gene **(red)** and The number of annotated high-confidence m6A-sites (Methods Table 1) (**blue**). **d)** m6A-GI in WT (top) and Mettl3 KO (bottom) mESC cells plotted against the L2FC of mRNA half-life between WT and Mettl3 KO. **e)** m6A-GI plotted against mRNA half-life in WT mESC **(top)** and KO mESCs **(bottom)**. Spearman R labeled. **f)** Pairwise Spearman correlation matrix of m6A-GIs calculated based on different m6A-seq datasets in different cells/tissues (**Supplementary Table 9**). **g)** HPLC-MS abundances of m6A across WT and m6A-writer perturbed mESC clones plotted against the 'expression-vs-methylation index', defined as the Spearman correlation between WT m6A-GI and normalized expression (TPM) (**see methods**). Annotated Pearson's R coefficient and p-value and 95 % confidence interval **h)** analysis as in **g,** conducted on published HPLC-MS determined m6A levels and RNA-Seq of mESC with perturbations of m6A-'writer' gene Zc3h13 [45]

We further explored whether the above-observed correlations with FC-half-life or LAIC-seq could be increased by summarizing coverage only across different regions within the gene (5' UTR, CDS, 3' UTR, last exon) or by focusing exclusively on regions displaying enrichment. We found, restricting the index to specific regions led to poorer correlations with both metrics, with the exception of the 5' UTR region, whose elimination resulted in slightly improved correlation with gene stability **(Fig. 3c)**. This is consistent with the notion that the enrichment of this region is not due to m6A but to m6Am [3,21,39], which is not thought to be associated with transcript destabilization [40–44].

Finally, we explored the extent to which m6A-GI varies between different cells and tissue-types. We found that m6A-GIs from widely different tissues (mESCs, embryonic fibroblasts, embryonic and adult brains, dendritic cells[21] correlated strongly with each other, with a mean pairwise correlation of 0.89 **(Fig. 3f)**. This suggests that across mouse tissues and cell types, m6A profiles at the gene level are to a large extent constitutively maintained, consistent with studies conducted at the m6A site level [21,22,38].

Inference of m6A sample levels directly from gene expression

The strong contribution of m6A to mRNA stability suggested the possibility of directly inferring m6A levels from steady-state levels of gene expression. This would be highly advantageous, given that the latter measurements are publicly available at large scales. To explore this possibility, we took the RNA-seq based normalized expression levels of the METTL3/WTAP perturbed mESC clones and computed the correlation of their steady-state gene expression levels to the m6A-GIs based on a reference m6A-seq dataset in WT mESCs. As expected, the magnitude of the negative correlation decreased in the perturbed samples, consistent with a reduced extent of m6A-dependent mRNA degradation. Importantly, we observed an excellent agreement between the magnitude of these correlations and the mass-spectrometry derived estimates of m6A levels in these samples **(Fig. 3g**). Given the high similarity in m6A-GIs across different cell types/tissues (**Fig. 3f**), we found that using m6A-GIs derived from mouse embryonic fibroblasts, dendritic cells or mouse adult brain tissue performed well in quantitatively classifying the extent of impairment of the methylation machinery (**Supplementary Figure Fig. 4b**). We could further extend and reproduce these analyses with a previous dataset of measurements in mESCs in which m6A-writer ZC3H13 was perturbed and rescued (**Fig. 3h**). These results demonstrate the potential of inferring methylation levels directly from measurements of RNA expression. Nonetheless, it should be noted that the absolute *magnitude* of this correlation is low (R is within a range of -0.1 to -0.3) which may hence limit the utility of this approach, in particular within contexts in which m6A levels are modulated within a narrower range than in the context of the genetic perturbations conducted here.

Discussion

The development of m6A-seq revolutionized the technical toolkit for interrogating regions modified by m6A [3,4,46–48]. Technical efforts in the field over the years have primarily concentrated on improving the *resolution* of m6A detection, either by introducing a cross-linking step following the IP [49,50] , or using antibody-independent methodologies [20,23–26] including nanopore-based approaches [51–53]. However, none of these approaches were geared towards the quantification of m6A levels across *genes* or across *samples*, two critical resolutions whose proper quantification may -

under many settings - be even more important than the identification of modified sites. Moreover, the vast majority of these methods have not been established to provide a robust quantitative readout even at the resolution of individual sites/peaks, and the ones that have (e.g. MAZTER-seq, [20]) can only do so for a low fraction of sites. These limitations have constrained interrogation of m6A dynamics and functions across biological responses and trajectories [22].

m6A-seq2 combines multiple advantages. At the peak level, it substantially reduces technical variability. This is critical, given that the inherent variability in m6A-seq has been pinpointed as one of the key factors giving rise to misinterpretations concerning the dynamics of this modification and its response to different genetic perturbations [22,54]. A related advantage of m6A-seq2 is that it reduces cost, input-material and labor by a factor proportional to the number of samples, while simultaneously increasing the scale. This feature is particularly important, given that lack of replicates (2-3 in the majority of studies) and insufficient statistical power was found to be among the major determinants giving rise to flawed inference of differential methylation [22,54].

M6A-seq2 further allows establishing sample-level indexes of m6A levels, establishing a critical alternative for HPLC-MS based approaches. Such an orthogonal readout is of critical importance, given that HPLS-MS is subject to both technical variability and is sensitive to 'contaminants' originating from non-poly(A) fractions, which have limited the ability to robustly detect RNA modifications [19,55,56]. Nonetheless, the m6A-SI has limitations. First, sample-level quantifications are inherently relative. We anticipate that in the future it may become possible to obtain absolute quantifications by employing spike-in samples with known m6A concentrations. Second, an implicit assumption made in the calculation of the m6A-SI is that differences in methylation between samples are - to a first approximation - *global*, impacting all sites to a similar extent. The m6A-SI will be less sensitive to capturing changes impacting only a fraction of sites. Third, m6A-SI provides more 'weight' to peaks originating in more highly expressed genes, given that they contribute a greater number of reads. It might be beneficial to utilize variants of the m6A-SI incorporating a step of normalizing m6A intensities by gene

expression, for applications requiring capturing changes in m6A levels at only a subset of lowly-expressed sites.

We further develop a conceptually-similar approach for *gene-level* quantifications of m6A. Application of this approach to mESCs explains ~30% of variability in mRNA stability, a dramatic increase with respect to previously employed peak-counting approaches, which suggests that m6A is a key determinant of RNA stability in mESCs. The ability of the m6A-GI to capture m6A levels suggests that mRNA may be methylated at low levels in a pervasive and diffuse manner, which can be best captured by aggregating the signal from the m6A-IP along the entire length of the mRNA transcript. This conclusion draws support from a recent study, which concluded that the sensitivity of antibody-based approaches is quite limited and that m6A is likely to be more widespread [20].

Of note, the sequencing coverage required to accurately quantify these three dimensions differs, with peaks requiring the greatest sequencing depth, followed by genes, followed by samples.

It is important to note that m6A immunoprecipitations are subject to technical biases due to antibody promiscuity and sequence-specific enrichment [27], which may impact m6A quantifications at the site or at the gene level. In the case of comparing the m6A-GI (or site index) for the *same* gene (or site) across distinct samples, this effect should be negligible. Indeed, an examination of the m6A-GI upon mixing WT and METTL3 KO cells at varying levels, revealed a striking linear increase in m6A-GI as a function of the contribution of mRNA from the WT cells **(Supplementary Fig. 5a-c)**. Comparison of m6A-GI across *different* genes is more prone to biases, given that the antibody may selectively enrich some genes more than others. Indeed, a set of five synthetic oligonucleotides, which were spiked into distinct samples at varying stoichiometries, demonstrated a strong correlation between the *relative* enrichment levels and the methylation stoichiometry, but demonstrated variability in the *absolute* enrichment levels **(Supplementary Fig. 5d)**. Nonetheless, the strong agreement of the m6A-GI, both with mRNA stability data and calibrated m6A-LAIC-seq measurements, suggest that the m6A-GI has important advantages in comparison to previously available approaches.

Collectively, m6A-seq2 expands the scale and resolutions at which m6A can be interrogated, while simultaneously decreasing the associated input requirements, technical variability costs and labor. We anticipate that m6A-seq2 will be widely adopted to interrogate the distribution, functions, and dynamics of m6A.

Materials & Methods

Table 1

	Source
Antibodies	
Rabbit anti-m6A antibody	Synaptic Systems (Cat# 202003; RRID: AB_2279214)
Deposited data	
m6A-seq: 841-strain, 3 technical replicates	GSE ##
m6A-seq2: 841-strain , 12 technical replicates	GSE ##
m6A-seq2: Yeast Meiosis time course	GSE ##
m6A-seq2: Yeast Genetic perturbation dataset	GSE ##
m6A-seq2: mESC m6A-'writer' perturbation clones	GSE ##
m6A-seq2: MEF, 5 technical replicates	GSE ##
m6A-seq: MEF, 5 technical replicates	GSE ##
m6A-seq2: mESC m6A-'writer' perturbation clones	GSE ##
Published datasets	
Actinomycin D mRNA half-life data: mESC WT & Mettl3 KO	9
m6A-seq data: mESC data WT & Mettl3 KO	20
RNA-seq data: mESC Zc3h13 perturbation	45
m6A-seq data: MEF shGFP, mouse adult brain & mouse dendritic cells	21
m6A-seq data: A549 shGFP	21
LAIC-seq: H1 ES cell m6A index	18
Experimental Models	
S. cerevisiae WT (SAy821) MAT a/α lys2/lys2 ho::LYS2/ho::LYS2	28

S. cerevisiae ndt80Δ/Δ (SAy841) MAT a/α lys2/lys2 ho::LYS2/ho::LYS2 ndt80::LEU2/ndt80::LEU2	28
S. cerevisiae Ime4Δ/Δ (SAy966) MAT a/α lys2/lys2 ho::LYS2/ho::LYS2 ndt80::LEU2/ndt80::LEU2 ime4::HIS3/ime4::HIS3	28
Murine Embryonic Stem Cells (mESCs) WT (male)	57
Murine Embryonic Stem Cells (mESCs) WT (male)	20
Murine Embryonic Fibroblasts (MEFs) WT	58
Software and Algorithms	
STAR v2.5.3a	59
RSEM v1.3.3	60
Samtools v1.3.1	61
bedtools v2.26.0	62
preseq v2.0.1	63
HOMER v4.9.1	64
Cutadapt v2.10	65
Reference genomes & gene annotations	
sk1 gene annotation table	**Supplementary Table 4**
mm9 gene annotation table	**Supplementary Table 5**
sk1 m6A site annotation annotation	**Supplementary Table 6**
mm9 high-confidence m6A site annotation	**Supplementary Table 7**
hg19 gene annotation table	**Supplementary Table 9**
hg19 high-confidence m6A site annotation	**Supplementary Table 10**

Yeast meiosis time course

All yeast strains used in this work were derived from the sporulation proficient SK1 strain background. S. cerevisiae cells (WT, PC: Ndt80Δ/Δ, NC: Ndt80Δ/Δ & Ime4Δ/Δ) were grown in YPD (2% dextrose) at 30C. To induce synchronous meiotic entry, cells were grown for 24 h in 1% yeast extract, 2% peptone, 4% dextrose at 30C, diluted in BYTA (1% yeast extract, 2% tryptone, 1% potassium acetate, 50 mM potassium phthalate) to OD600 = 0.2 and grown for another 16 h at 30C, 200 rpm. Cells were then washed once with water and re-suspended in SPO (0.3% potassium acetate) at OD600 = 2.0 and incubated at 30C at 190 rpm. Cells were isolated from SPO at the indicated times and collected by 2 min centrifugation at 3000 g. Pellets were snap-frozen and stored at -80 for RNA extraction.

Genetic perturbation yeast-strain construction

Gene deletion strains were generated by a one-step promoter replacement protocol, and genetic crosses [66]. For generating depletion alleles, we used the auxin-induced degron allele (AID) and a plasmid expressing Oryza sativa osTIR1 ligase from a CUP1 promoter (pCUP1-TIR1) (gift from Elçin Ünal) [67]. All gene deletions and AID alleles were crossed into the ndt80-deletion background, which ensures that cells do not exit meiotic prophase [68].

Table 2. SK1 genetic perturbation candidate genes affiliated role in the regulation of the yeast transcriptome

Gene	Involvement
NDT80	meiosis [68]
RPB3	transcription [69]
RAT1	transcription, mRNA degradation [70,71]
NRD1	transcription, mRNA degradation [72,73]
LSM1	transcription, mRNA degradation [74]
CAF40	transcription, mRNA decapping [75]
NOT5	transcription, mRNA degradation [76]
NOT3	transcription, mRNA decapping [77]

RRP6	mRNA degradation [78]
DCP2	mRNA decapping[79]
XRN1	mRNA degradation [80]
POP2	mRNA degradation [81]
PAT1	mRNA degradation [82], mRNA decapping [83]
PAN3	polyadenylation [84,85]
PHO92	m6A metabolism [86]
GIS2	mRNA translation [87]

Growth and medium conditions

To induce depletion, AID tagged allele cells expressing pCUP-TIR1 were grown to BYTA as described above. Four hours after shifting to SPO, 50 μM CuSO4 was added to induce TIR1 expression from the CUP1 promoter and 500 μM indole-3-acetic acid (IAA) to induce degradation. Samples were harvested after 2 hours (thus 6 hours in SPO).

Cell line & cell culture

MEFs cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Biological industries, 01-052-1A) supplemented with 10% fetal bovine serum (FBS).

Cloning of gRNA for targeting of methylation writers in mESCs

Single guide RNAs were annealed and cloned into a px330 vector according to Zhang lab standard cloning protocol [88]. gRNA pair sequences used to remove exon 4 of Mettl3 or exon 5 of Wtap can be found in (**Supplementary Table 1**)

RNA preparation

Yeast total RNA samples were prepared by MasterPure Yeast RNA extraction kit (Lucigen, MPY03100). For mESCs cells, total RNA was extracted using Nucleozol (Macherey-Nagel, 740404.200).

Mouse embryonic stem cells - m6A writer perturbation

Commonly used V6.5 [57] mouse embryonic stem cell line was kept in DMEM (Gibco) supplemented with penicillin-streptomycin, 1 mM L-glutamin, 1% non-essential amino acids, 20% high-grade fetal bovine serum, beta-mercaptoethanol and 10 µg recombinant leukemia inhibiting factor (LIF). Cells were kept on tissue culture plates covered by 0.2% gelatin in co-culture with in-house generated, radiation-inactivated mouse embryonic fibroblasts (MEF).

CRISPR-Cas9 vector was transfected into the mESCs with TransIT-X2 reagent (Mirus). Cells were FACS-sorted three days post-transfection, and cells that obtained both targeting gRNAs were seeded on 0.2% gelatin without MEF in low density. Single colonies were picked under stereomicroscope a week later. Successful targeting was confirmed via Sanger sequencing of genomic DNA of selected clones.

HPLC-MS for determination of m6A/A

The analysis was performed following the procedure described [20], with the difference that, 500 ng of double selected polyA RNA fractions were used for digestion.m6A-seq2 The proposed protocol builds on the previously reported protocol [27,89].

Synthesis of in-vitro transcribed RNA

Two synthetic RNA fragments (IS1,IS2,IS3,IS4,IS5, **Supplementary Table 1**), each comprising a 102 nt long sequence with a single ACA in the center were in vitro transcribed from dsDNA templates containing a T7 promoter, either in the presence of ATP or N6-methyl-ATP, using MaxiScriptT7 kit (Invitrogen, AM1320). Purified m6A-containing products were serially diluted in non-m6A containing products.

m6A-seq2 library preparation

RNA ligation and pooling: Starting material for the m6A-seq2 protocol were 1.2/n µg per sample (with n samples) of double polyA selected RNA (Dynabeads® mRNA DIRECT™). RNA fragmentation to ~150 nt fragment size was performed according to the protocol with the Invitrogen™ RNA Fragmentation Reagents kit (Invitrogen™). After every step, apart from those in which the sample was eluted in H2O, a cleanup with Dynabeads® MyOne™ Silane was performed according to the protocol. For the subsequent DNAse and dephosphorylation treatment, each sample was incubated in T4 PNK (NEB T4 Polynucleotide Kinase), TURBO™ DNase and FastAP (Thermo Scientific™ FastAP Thermosensitive Alkaline Phosphatase) for 30 min in 37c in 5x FNKBuffer (1:1:2 ratio of T4PNK-buffer : FastAP-buffer : H2O). 3' RNA barcode adapter ligation was performed with 100 pmol of RNA ILL adapter (**Supplementary Table 1**) and 36 U T4 RNA ligase (NEB) for 1.5 h at room temperature (23c) for each sample. Following the 3' ligation of barcoded RNA adapters, all samples were pooled for the multiplexed m6A-Immunoprecipitation (m6A-IP). 10% of the sample-pool was taken as an Input-RNA sample.

Multiplexed m6A-IP: 25µl of Dynabeads protein G beads (Invitrogen) were washed twice in 100 µl IPP buffer and for 30 min hybridized with 3.5 µl anti-m6A antibody (Synaptic Systems). The sample pool, after heat denaturation (2 min 70°C, stored in ice afterward), was incubated with the anti-m6A AB-Protein G beads at 4°C for 2 hours. The RNA-AB-Protein G beads were then washed twice in 200 µl of IPP buffer, twice in low-salt IPP buffer (50 mM NaCl, 0.1% NP-40, 10 mM Tris-HCl, pH 7.5), and twice in high-salt IPP buffer (500 mM NaCl, 0.1% NP-40, 10 mM Tris-HCl, pH 7.5). RNA was eluted in 30 µl RLT (Qiagen). were washed in 100 µl RLT, resuspended in 30 µl RLT, and added to the eluted RNA. 60 µl of 100% ethanol was added to the mixture, the mixture attached to the magnet, and the supernatant discarded. After two washes in 100 µl of 70% ethanol, the RNA was eluted from the beads in 100 µl of IPP for the second round of the m6A-IP cycle with a final elution from the beads in 13.5 µl H_2O for cDNA synthesis with the rTd RT primer (**Supplementary Table 1**) and SuperScript™ III Reverse Transcriptase. Alkaline RNA hydrolysis in 1 M NaOH in 70c for 12 min. Illumina 5'adapter ligation was performed with 50 pmol 5iLL-22 DNA adapter with 45U T4 RNA Ligase 1 for 3 h at room temperature. PCR

enrichment was performed with KAPA HiFi PCR Kit with the universal forward primer and the reverse primer containing the barcode. Amplified libraries were cleaned with AMPure XP beads (Agencourt, A63881), quantified using Qubit (Life Technologies) and the distribution of library size was determined using TapeStation (Agilent Technologies).

m6A-Seq

For a better comparison of batch-induced variability, the m6A-seq protocol was performed according to the protocol above, but with 1.2 µg of double poly-A selected RNA per sample (~1% yield of the total RNA) and without sample-pooling. The steps prior to pooling were adjusted to the increased amount of input RNA. 250 ul Dynabeads® (according to the protocol) was used for the poly-A selection.

m6A-seq2 data analysis

Read alignment

Paired-end reads of m6A-seq2 libraries were demultiplexed into individual samples, using an in-house python script which distributes reads into a sample FASTQ file according to the barcoded-sequence in Read2 (Pos 4-10). The genomic paired-end alignment was performed with STAR v.2.5.3a [59] to the SK1 reference genome [27] with default parameters but limiting the intron length to 500 nt (--alignIntronMax 500' parameter). Mouse-derived samples were aligned to the mm9 reference genome with default parameters ('--alignIntronMax 1000000' parameter). Unique aligned reads were extracted for further use. For the technical replicates comparison, one million paired-end alignments were subsampled for Input and m6A-IP per sample using Samtools v1.3.1 [61]. Paired-end alignment mRNA gene coverage (whole insert coverage) was calculated using the bam2ReadEnds.R script (which calculates paired-end read coverage for all reads whose alignment patterns are consistent with the gene annotation, excluding reads from intergenic or intronic regions)[20] with a sk1 gene annotation table for yeast, and mm9 gene annotation table for mouse (**Supplementary Table 4 & 5**).

Merging technical replicates

For the m6A gene- and site quantification in the 24-sample mESC m6A-seq2 experiment, the FASTQ files of the 3 technical replicates for each sample (0%, 14%, …, 100% mESC WT RNA) were merged (concatenated) to increase the coverage.

m6A site scores

The m6A site coordinates are based on a reference table with a single-nucleotide resolution of confirmed m6A sites (**Supplementary Table 6 & 7**). Coverage comparison was performed in a 51 base window centered around the annotated m6A-site, defined below.

m6A site score 1

The m6A site score 1 calculation was performed as a ratio of the sum of the m6A-IP coverage per base in a 51 bp window centered around the annotated m6A-site divided by the corresponding value in the input sample. To ensure adequate coverage, we demanded a sum of position coverage > 250 for both Input and m6A-IP coverage, corresponding to ~5 reads/base on average.

m6A site score 2

The median normalized m6A-site score (adapted from [27]) calculation was based on the enrichment of the sum m6A-IP coverage of 50-nt window centered around the annotated m6A-site over the median m6A-IP coverage (sum of 10 base window) of the whole annotated transcript.

m6A sample index - m6A-SI

The m6A-SI for a transcriptome-wide m6A quantification estimate was defined as the ratio of the sum of coverage (over a set of predefined peaks) in the m6A-IP dataset divided by the corresponding number in the input dataset. Note that for this score, no normalization for library size was performed. For the 4 batches of genetic perturbation in yeast-strains, the scores calculated in a batch were multiplied by a correction factor, which was the ratio of the sum of m6A-IP reads to the sum of m6A-Input reads of the whole batch. This was performed to make the scores comparable between the batches.

$$\frac{\sum_{i=1}^{k} \text{m6A IP site}_i \text{coverage}}{\sum_{i=1}^{k} \text{m6A Input site}_i \text{coverage}}$$

m6A gene index - m6A-GI

The calculation of the m6A-GI was performed by calculating the ratio in the number of reads aligned to the entire gene excluding the annotated 5' UTR in the m6A-IP sample and normalizing by the corresponding number in the input. The various region-based m6A indexes were calculated in the same way, with the coordinates adjusted to the targeted and annotated regions (5'UTR, 3'UTR, CDS, last exon, gene). For the last exon calculation, single-exon genes were treated as the last exon. Calculated region-based scores are listed in **Supplementary Table 3**.

m6A peak detection

*D*e novo m6A-seq peak detection was performed with adaptations according to the published protocol [21]. The peak detection is based on a 'window-score' for each nucleotide position centered in a rolling window of 100 nt. The window-score is the ratio of the m6A-IP FC mean coverage of the window divided by the median m6A-IP coverage of the gene, divided in turn by the corresponding number in the input sample. The thresholds for defining an 'm6A-peak' were a window score > 2 and a median Input coverage of a given gene > 10. Finally, neighboring positions with a sufficient window-score were merged.

Motif enrichment analysis

For all the defined 'm6A peaks' of a sample, de *novo* motif enrichment and P values of consensus motifs were generated by HOMER (**Material & Methods Table 1**) with a one-sided binomial test.

Metagene analysis

Metagene coverage density was calculated by aggregating the normalized coverage densities of every detected gene. Specifically, each gene was binned into 100 intervals and the sum of coverage for each of the intervals was normalized by the

sum of coverage per gene. Then the coverage density for each interval of all the genes were aggregated and normalized by the sum of coverage over all intervals.
The metagene motif distribution analysis was performed by using the transcriptomic coordinate of the closest DR**A**CH motif to the *de novo* detected m6A peak summit (max. normalized m6A-IP coverage of the peak region). Further the relative localization of the motif coordinate in a transcript region (5'UTR, CDS, 3'UTR) was calculated. Finally we defined the relative lengths of the three regions by scaling all genes that contained at least one detected m6A peak and using the mean of all the relative lengths of the 3 transcript regions.

Yeast Meiosis Signature

The yeast meiosis gene signature was defined as the median scaled expression of genes, which exhibit an expression profile of specific upregulation during yeast meiosis time course at T4 (h). Detection of the 67 genes was based on a Pearson R > 0.5 of expression profiles sampled from a yeast meiosis time course (RSEM calculated TPM) against a vector with a maximum (1) at T4 position (residual time point positions had 0)

RNA-seq data analysis

Each dataset was first trimmed with cutadapt (**Material & Methods Table 1**). The subsequent genome alignment (sk1 and mm9) and expression estimation (TPM) was performed using rsem-calculate-expression (rsem/1.3.3, [60]).

mRNA half-life estimation

mRNA decay-rate estimation was performed by linear modeling of the log-transformed normalized expression of the Actinomycin D time course samples against the respective time points. Decay rates that were positive or ones emerging from non-significant fits of the linear model ($p>0.05$) were removed.

m6A inference from expression levels

To infer the m6A-signature from normalized mRNA expression levels, the ‘expression-vs-methylation index’ was calculated. The ‘expression-vs-methylation index’ was defined as the Spearman ρ calculated between the normalized gene expression estimates (TPM) of the sample and the m6A-GIs of a reference m6A-seq experiment. Genes were filtered exhibiting < -0.5 log2 fold-change of Mettl3- (m6A-) dependent half-life (WT/Mettl3 KO).

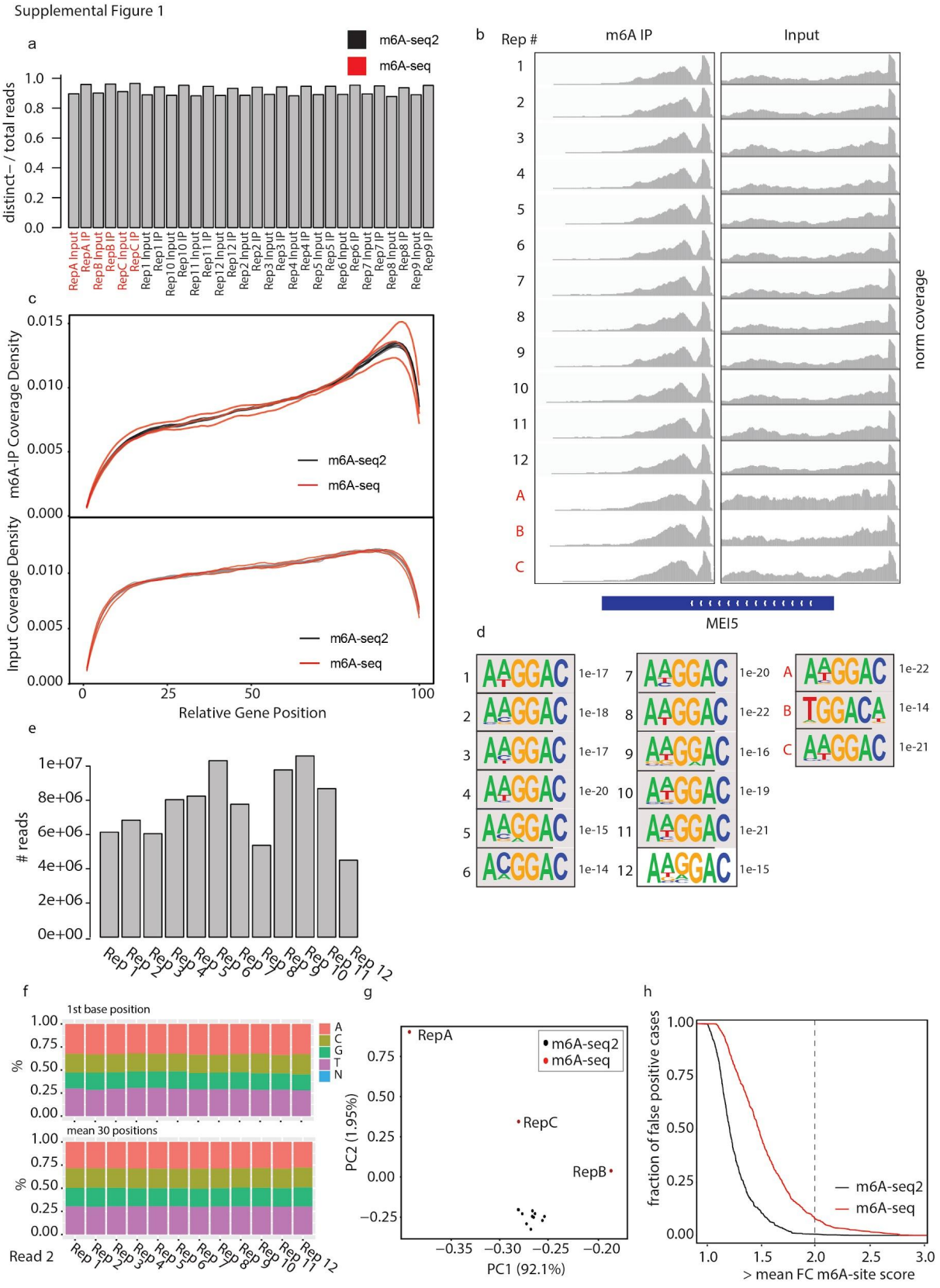
Acknowledgments

We would like to thank Prof. Elçin Ünal for giving us the plasmid expressing Oryza sativa osTIR1 ligase from a CUP1 promoter (pCUP1-TIR1).

Competing interests

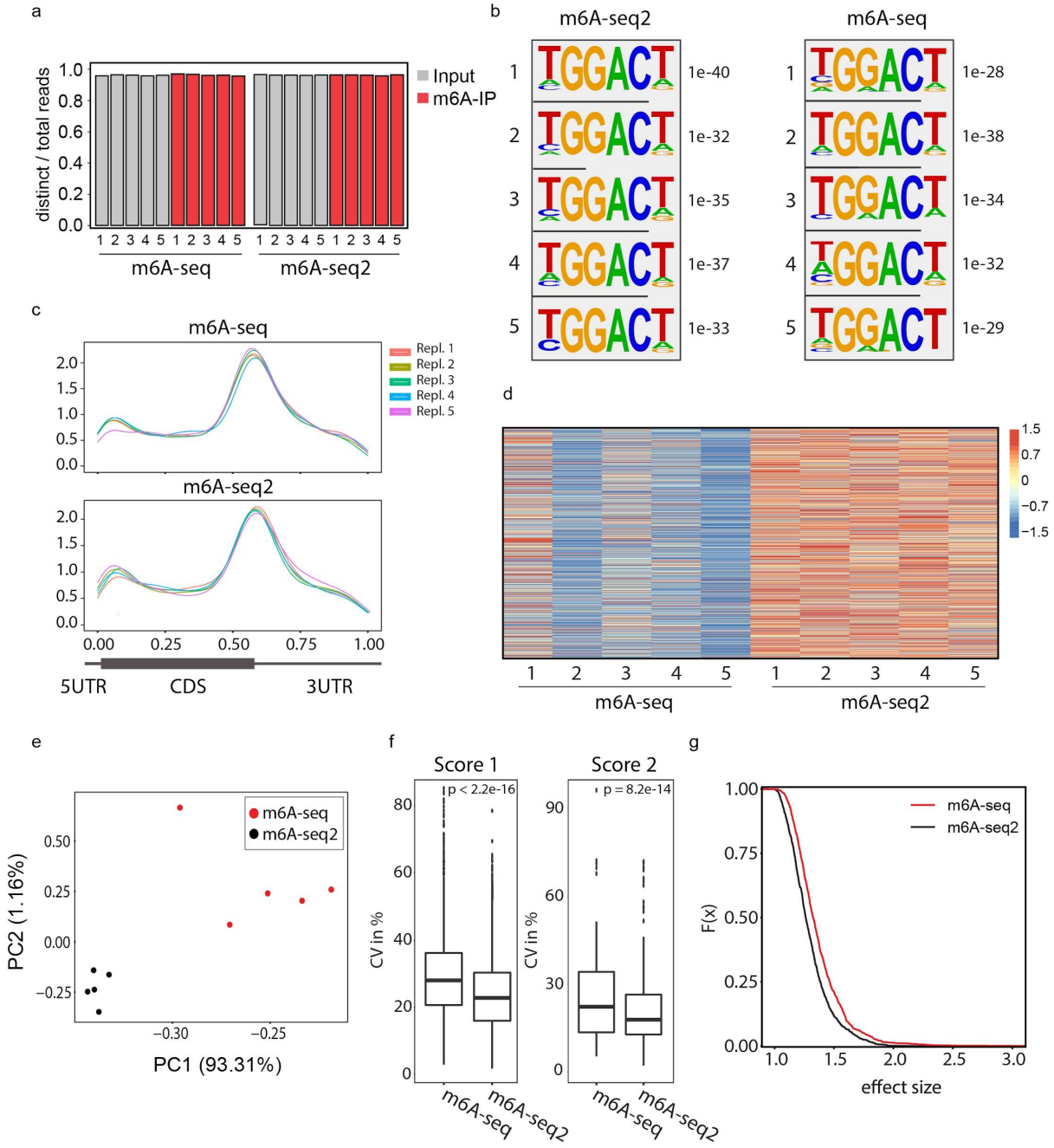
The authors declare no competing interests.

Supplementary Figures



Supplementary Figure 1. m6A-seq2 reduces batch-induced variability while maintaining m6A-enrichment capacities **a)** Sequencing complexity metric based on the ratio of distinct- to total read numbers for 500,000 sampled reads (**Methods Table 1**) **b)** m6A-IP and Input normalized coverage profile of MEI5 for the 12 m6A-seq2 replicates and the 3 m6A-seq replicates **c)** Metagene coverage densities of m6A-IP and Input for the 12 m6A-seq2 replicates (red) and 3 m6A-seq replicates (black). **d)** Motif analysis showing the most significant HOMER **(Material & Methods Table 1)** *de novo* motif analysis, based on detected peaks (**see Methods)** of the 12 m6A-seq2 replicates (red) and 3 m6A-seq replicates (black) **e)** m6A-seq2 number of reads of the 12 technical replicates. **f)** Identity of nucleotide of the first position of read 2 (top), which resembles the 3' terminus of a sequenced RNA fragment. Stacked barplot of the mean nucleotide identity of the first 30 positions (bottom). **g)** Principal component analysis of m6A-site score 2 of 486 m6A sites. **h** Cumulative plot showing the fraction of false-positive detected 'differential-methylation' cases and the effect-size as the absolute fold-change of the mean score1 between the sampled groups. Conducted in a bootstrapping approach to generate 8 measurements for each m6A site. Random assignment into 2 groups, and used a combination of a statistical test (t-test) and m6A-site score fold changes.

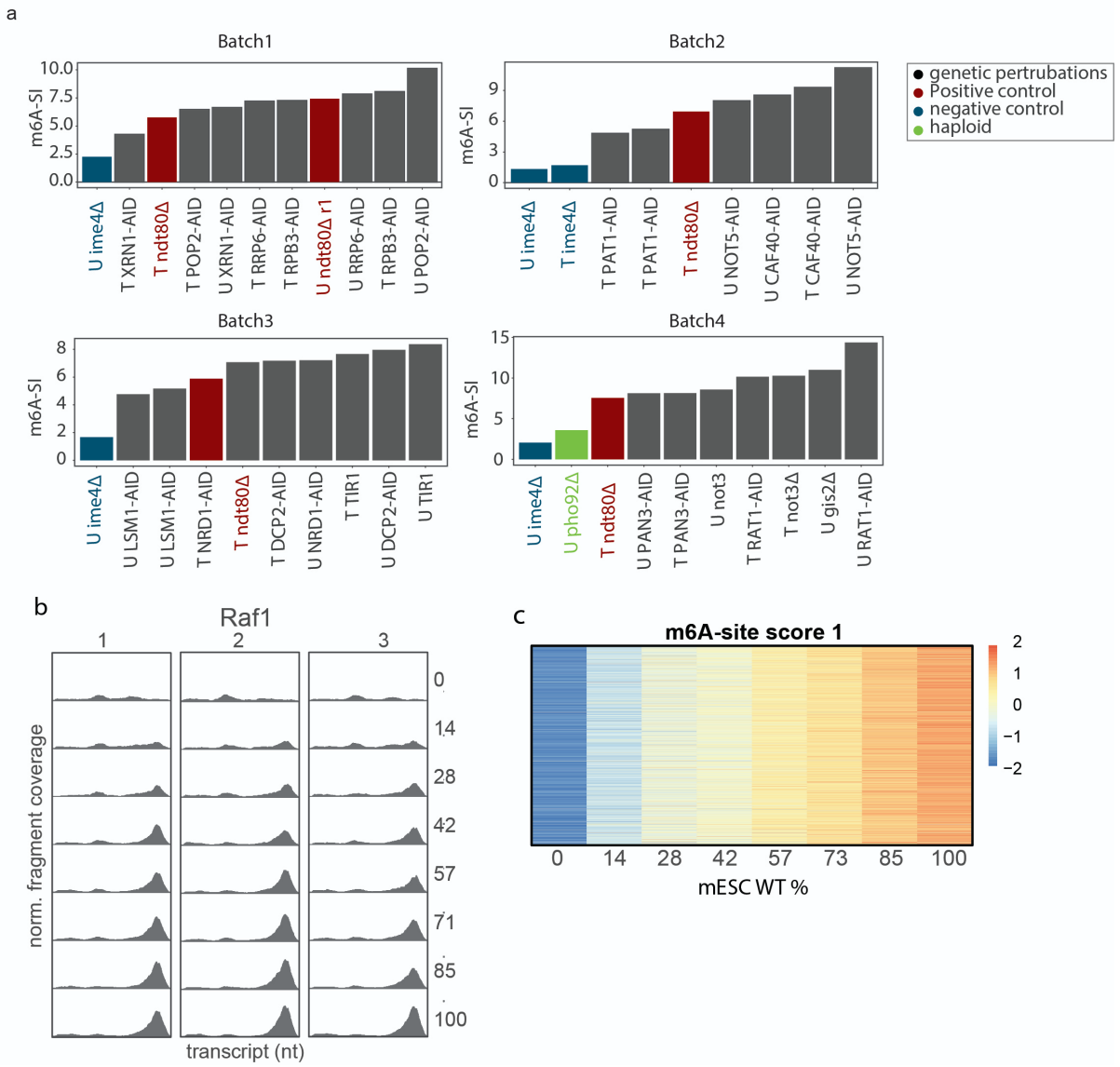
Supplementary Figure 2



Supplementary Figure 2. m6A-seq2 reduces batch-induced variability in technical replicates of mouse embryonic fibroblasts. **a)** Library complexity metric as the ratio of distinct- to total read numbers based on one million reads of the m6A-seq2 and m6A-seq MEF dataset (**preseq, see Methods Table 1**) **b)** Motif analysis showing the most significant HOMER **(Material & Methods Table 1)** de *novo* motif analysis, based on detected peaks (**see Methods**) of the m6A-seq2 technical replicates (left) and m6A-seq technical replicates (right). **c)** Metagene m6A-peak distribution (**see Methods**) of the m6A-seq technical replicates (top) and m6A-seq2 replicates

(bottom). **d)** m6A-site score 1 comparison between technical replicates generated via m6A-seq or m6A-seq2. score2 for 3416 m6A-sites passing the coverage thresholds for high-confidence m6A site estimation (**see methods**). **e)** Principal component analysis of m6A-site score 1 of 3416 high-confidence m6A sites. **f)** Comparison of the percentage coefficient of variation (%CV) of score1 (left) and score2 (right) estimates across the technical replicates measured via m6A-seq and m6A-seq2. Wilcoxon test p value annotated **g)** Cumulative plot showing the fraction of false-positive detected 'differential-methylation' cases and the effect-size as the absolute fold-change of the mean score between the sampled groups.

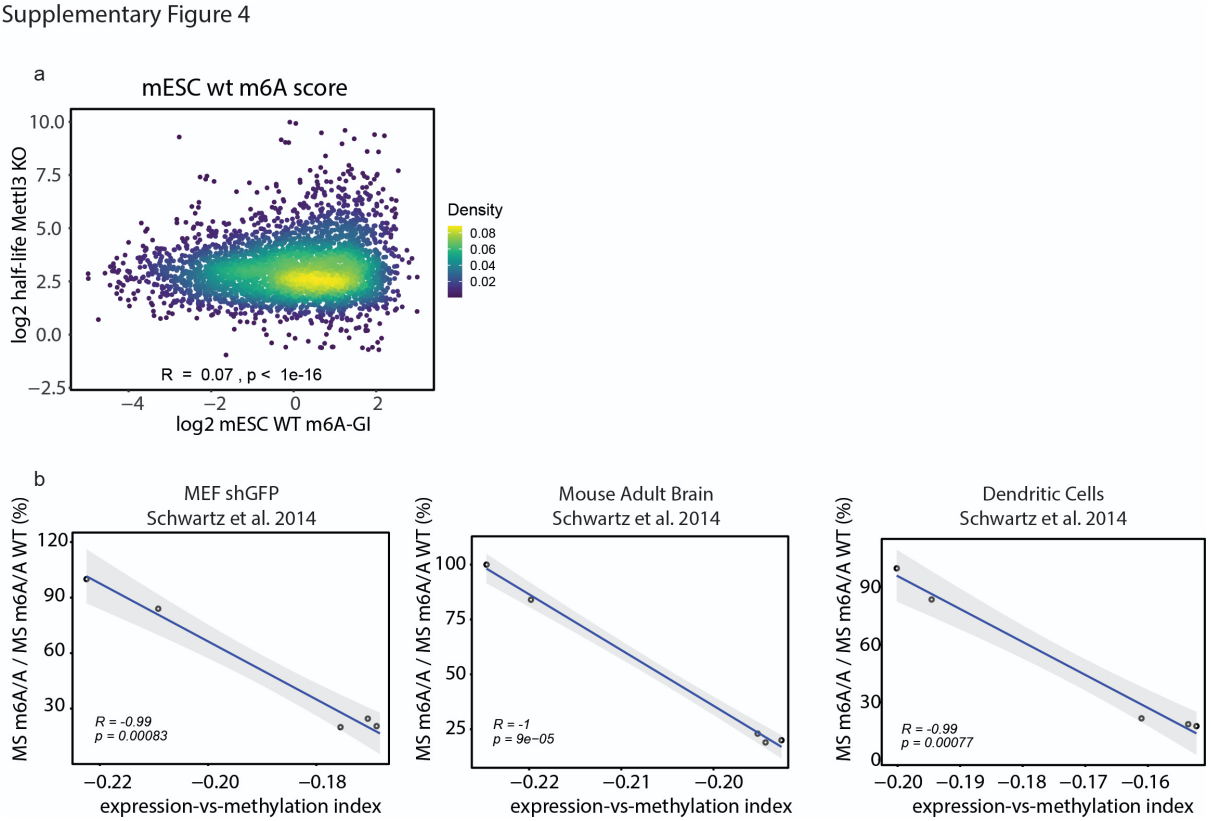
Supplementary Figure 3



Supplementary Figure 3. m6A-methylome profiling upon genetic perturbations during yeast meiosis **a)** m6A-SI based on m6A-Seq2 measurements on individual genetic perturbations strains with ndt80Δ/Δ background. Barplot for each of the m6A-seq2 batches. In total 3 gene deletion strains (*pho92*Δ/Δ, *not3*Δ/Δ, *gis2*Δ/Δ), 12 genes with an auxin-inducible degrons (AID), 5 positive controls (only ndt80Δ/Δ, red) and 5 negative controls (ime4Δ/Δ & ndt80Δ/Δ, blue) (noted in legend) and AID-control (TIR1). U: untreated samples and T: treatment with CU and IAA (**see Methods**). **b)** Raf1 m6A-IP fragment coverage normalized by sample library size and Raf1 expression level (RPM) based on the input sample for all 24 samples (% amount of WT mESC RNA per triplicate is annotated at the right). The Y-axis is fixed to an identical range across all samples. **c)**

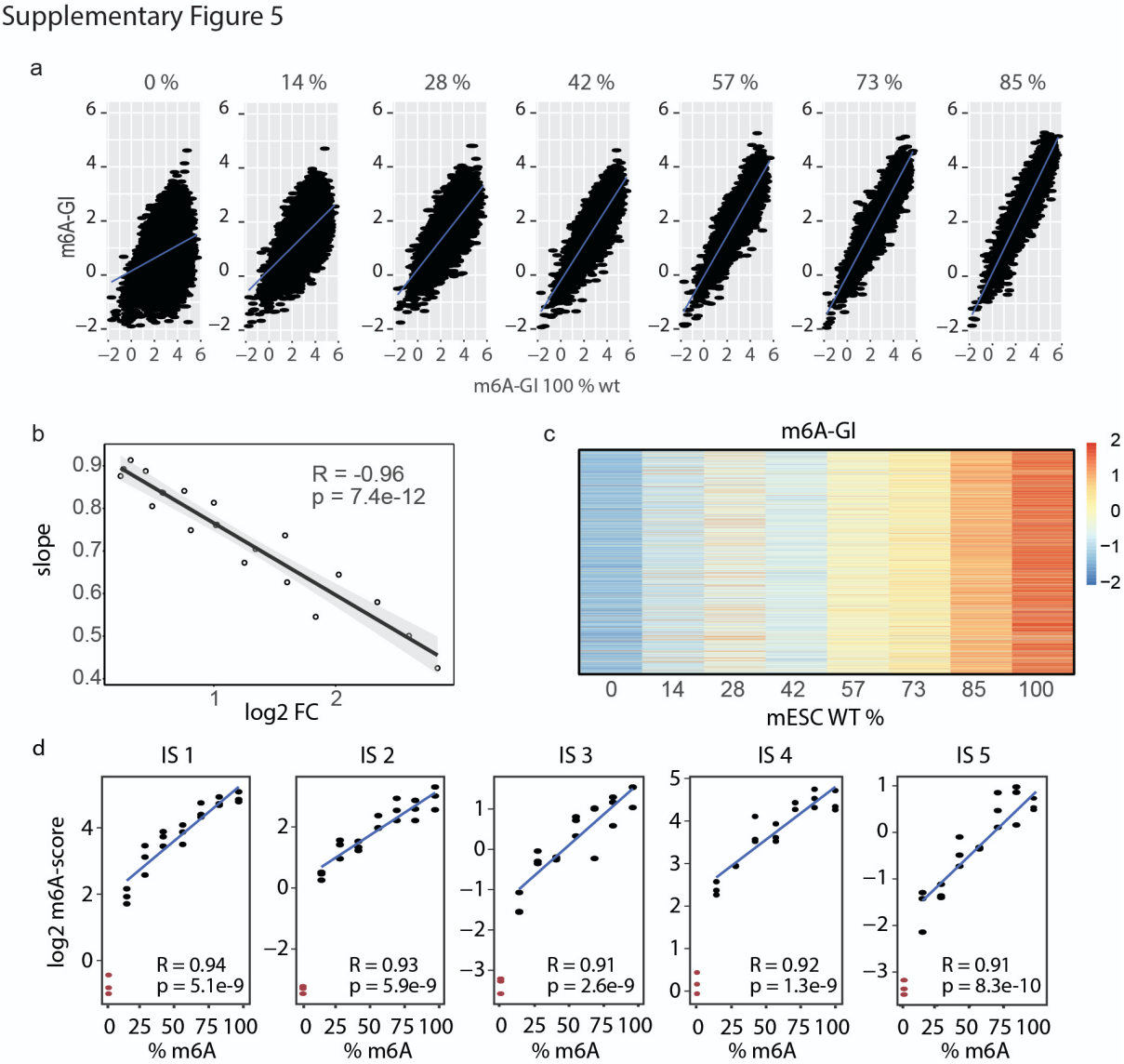
Heatmap of log2 transformed m6A-site score 1 of 5379 annotated high-confidence m6A-sites (scaled by row) of the samples (merged triplicates).

Supplementary Figure 4



Supplementary Figure 4. m6A-GI infers with half-life and steady-state expression **a)** m6A-GI of the mESC WT cells plotted against mRNA half-life in Mettl3-KO mESC **b)** HPLC-MS abundances of m6A across WT and m6A-writer perturbed mESC clones plotted against the 'expression-vs-methylation index', defined as the spearman correlation between normalized expression (TPM, see Methods) and m6A-GIs calculated based on 3 previously published m6A-seq datasets [21] .

Supplementary Figure 5



Supplementary Figure 5 m6A-gene index correlates with the increase of m6A **a)** m6A-gene index (m6A-GI) of 100 % WT sample (merged triplicates) plotted against the other sample concentrations (merged triplicates) with a linear model fit (blue line). **b)** Scatterplot of the slope of the linear modeling fit of all unique combinations of m6A-GIs derived from different % WT mESC (as in a)), plotted against the log2 transformed fold-change (FC) of the corresponding unique combination out of the 7 samples (14 % WT to 100 % WT). **c)** Heatmap of the log2 transformed m6A-GIs (scaled by gene) for 6572 genes of the different samples (merged triplicates). **d)** Scatter plots of the m6A-index of five distinct in-vitro transcribed spikes (IS1-5, **see Supplementary Table 1**) with a defined amount of m6A ranging from 0% m6A (in red) over 7 increasing concentrations to 100% (in black). Linear-modeling and Pearson's R determined for all data points with m6A

(black) annotated. Spikes with 0% methylation are plotted in red, and were not taken into account for the linear model.

References

1. Perry, R. P. & Kelley, D. E. Existence of methylated messenger RNA in mouse L cells. *Cell* **1**, 37–42 (1974).
2. Desrosiers, R., Friderici, K. & Rottman, F. Identification of methylated nucleosides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3971–3975 (1974).
3. Dominissini, D. *et al.* Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201–206 (2012).
4. Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–1646 (2012).
5. Wang, X. *et al.* N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–120 (2014).
6. Wang, X. *et al.* N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**, 1388–1399 (2015).
7. Roundtree, I. A. *et al.* YTHDC1 mediates nuclear export of N6-methyladenosine methylated mRNAs. *Elife* **6**, (2017).
8. Shi, H. *et al.* YTHDF3 facilitates translation and decay of N6-methyladenosine-modified RNA. *Cell Res.* **27**, 315–328 (2017).
9. Ke, S. *et al.* m6A mRNA modifications are deposited in nascent pre-mRNA and are not required for splicing but do specify cytoplasmic turnover. *Genes Dev.* **31**, 990–1006 (2017).
10. Li, A. *et al.* Cytoplasmic m6A reader YTHDF3 promotes mRNA translation. *Cell Res.* **27**, 444–447 (2017).
11. Liu, N. *et al.* N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
12. Batista, P. J. *et al.* m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. *Cell Stem Cell* **15**, 707–719 (2014).

13. Cui, Q. *et al.* m6A RNA Methylation Regulates the Self-Renewal and Tumorigenesis of Glioblastoma Stem Cells. *Cell Rep.* **18**, 2622–2634 (2017).
14. Rosa-Mercado, N. A., Withers, J. B. & Steitz, J. A. Settling the m6A debate: methylation of mature mRNA is not dynamic but accelerates turnover. *Genes Dev.* **31**, 957–958 (2017).
15. Darnell, R. B., Ke, S. & Darnell, J. E., Jr. Pre-mRNA processing includes N6 methylation of adenosine residues that are retained in mRNA exons and the fallacy of 'RNA epigenetics'. *RNA* **24**, 262–267 (2018).
16. Zhao, B. S., Nachtergaele, S., Roundtree, I. A. & He, C. Our views of dynamic N6-methyladenosine RNA methylation. *RNA* **24**, 268–272 (2018).
17. Zaccara, S. & Jaffrey, S. R. A Unified Model for the Function of YTHDF Proteins in Regulating m6A-Modified mRNA. *Cell* **181**, 1582–1595.e18 (2020).
18. Molinie, B. *et al.* m(6)A-LAIC-seq reveals the census and complexity of the m(6)A epitranscriptome. *Nat. Methods* **13**, 692–698 (2016).
19. Legrand, C. *et al.* Statistically robust methylation calling for whole-transcriptome bisulfite sequencing reveals distinct methylation patterns for mouse RNAs. *Genome Res.* **27**, 1589–1596 (2017).
20. Garcia-Campos, M. A. *et al.* Deciphering the 'm6A Code' via Antibody-Independent Quantitative Profiling. *Cell* **178**, 731–747.e16 (2019).
21. Schwartz, S. *et al.* Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep.* **8**, 284–296 (2014).
22. McIntyre, A. B. R. *et al.* Limits in the detection of m6A changes using MeRIP/m6A-seq. *Sci. Rep.* **10**, 6590 (2020).
23. Meyer, K. D. DART-seq: an antibody-free method for global m6A detection. *Nature Methods* vol. 16 1275–1280 (2019).
24. Wang, Y., Xiao, Y., Dong, S., Yu, Q. & Jia, G. Antibody-free enzyme-assisted chemical approach for detection of N6-methyladenosine. *Nature Chemical Biology* vol. 16 896–903 (2020).

25. Shu, X. *et al.* A metabolic labeling method detects m6A transcriptome-wide at single base resolution. *Nature Chemical Biology* vol. 16 887–895 (2020).
26. Zhang, Z. *et al.* Single-base mapping of m6A by an antibody-independent method. *Sci Adv* **5**, eaax0250 (2019).
27. Schwartz, S. *et al.* High-resolution mapping reveals a conserved, widespread, dynamic mRNA methylation program in yeast meiosis. *Cell* **155**, 1409–1421 (2013).
28. Agarwala, S. D., Blitzblau, H. G., Hochwagen, A. & Fink, G. R. RNA methylation by the MIS complex regulates a cell fate decision in yeast. *PLoS Genet.* **8**, e1002732 (2012).
29. Clancy, M. J., Shambaugh, M. E., Timpte, C. S. & Bokar, J. A. Induction of sporulation in Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA: a potential mechanism for the activity of the IME4 gene. *Nucleic Acids Res.* **30**, 4509–4518 (2002).
30. Zhuang, F., Fuchs, R. T., Sun, Z., Zheng, Y. & Robb, G. B. Structural bias in T4 RNA ligase-mediated 3'-adapter ligation. *Nucleic Acids Res.* **40**, e54 (2012).
31. Fuchs, R. T., Sun, Z., Zhuang, F. & Robb, G. B. Bias in ligation-based small RNA sequencing library construction is determined by adaptor and RNA structure. *PLoS One* **10**, e0126049 (2015).
32. Shishkin, A. A. *et al.* Simultaneous generation of many RNA-seq libraries in a single reaction. *Nat. Methods* **12**, 323–325 (2015).
33. Roth, S. H., Levanon, E. Y. & Eisenberg, E. Genome-wide quantification of ADAR adenosine-to-inosine RNA editing activity. *Nat. Methods* **16**, 1131–1138 (2019).
34. Geula, S. *et al.* m6A mRNA methylation facilitates resolution of naïve pluripotency toward differentiation. *Science* **347**, 1002–1006 (2015).
35. Ries, R. J. *et al.* m6A enhances the phase separation potential of mRNA. *Nature* vol. 571 424–428 (2019).
36. Anders, M. *et al.* Dynamic m6A methylation facilitates mRNA triaging to stress granules. *Life Sci Alliance* **1**, e201800113 (2018).
37. Lasman, L., Krupalnik, V., Geula, S., Zerbib, M. & Viukov, S. Context-dependent

functional compensation between Ythdf m6A readers. *bioRxiv* (2020).

38. Gokhale, N. S. *et al.* Altered m6A Modification of Specific Cellular Transcripts Affects Flaviviridae Infection. *Mol. Cell* **77**, 542–555.e8 (2020).
39. Mauer, J. *et al.* Reversible methylation of m6Am in the 5′ cap controls mRNA stability. *Nature* **541**, 371–375 (2016).
40. Akichika, S. *et al.* Cap-specific terminal N6-methylation of RNA by an RNA polymerase II-associated methyltransferase. *Science* **363**, (2019).
41. Boulias, K. *et al.* Identification of the m6Am Methyltransferase PCIF1 Reveals the Location and Functions of m6Am in the Transcriptome. *Mol. Cell* **75**, 631–643.e8 (2019).
42. Sun, H., Zhang, M., Li, K., Bai, D. & Yi, C. Cap-specific, terminal N6-methylation by a mammalian m6Am methyltransferase. *Cell Res.* **29**, 80–82 (2019).
43. Sendinc, E. *et al.* PCIF1 Catalyzes m6Am mRNA Methylation to Regulate Gene Expression. *Mol. Cell* **75**, 620–630.e9 (2019).
44. Cowling, V. H. CAPAM: The mRNA Cap Adenosine N6-Methyltransferase. *Trends Biochem. Sci.* **44**, 183–185 (2019).
45. Wen, J. *et al.* Zc3h13 Regulates Nuclear RNA m6A Methylation and Mouse Embryonic Stem Cell Self-Renewal. *Mol. Cell* **69**, 1028–1038.e6 (2018).
46. Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat. Rev. Mol. Cell Biol.* **18**, 31–42 (2017).
47. Zaccara, S., Ries, R. J. & Jaffrey, S. R. Reading, writing and erasing mRNA methylation. *Nat. Rev. Mol. Cell Biol.* **20**, 608–624 (2019).
48. Schwartz, S. Cracking the epitranscriptome. *RNA* **22**, 169–174 (2016).
49. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* **12**, 767–772 (2015).
50. Grozhik, A. V., Linder, B., Olarerin-George, A. O. & Jaffrey, S. R. Mapping m6A at Individual-Nucleotide Resolution Using Crosslinking and Immunoprecipitation (miCLIP). *Methods Mol. Biol.* **1562**, 55–78 (2017).

51. Liu, H. *et al.* Accurate detection of m6A RNA modifications in native RNA sequences. *Nat. Commun.* **10**, 4079 (2019).
52. Drexler, H. L., Choquet, K. & Churchman, L. S. Splicing Kinetics and Coordination Revealed by Direct Nascent RNA Sequencing through Nanopores. *Mol. Cell* **77**, 985–998.e8 (2020).
53. Lorenz, D. A., Sathe, S., Einstein, J. M. & Yeo, G. W. Direct RNA sequencing enables m6A detection in endogenous transcript isoforms at base-specific resolution. *RNA* **26**, 19–28 (2020).
54. Zhang, Z. *et al.* RADAR: differential analysis of MeRIP-seq data with a random effect model. *Genome Biol.* **20**, 294 (2019).
55. Sas-Chen, A. *et al.* Dynamic RNA acetylation revealed by quantitative cross-evolutionary mapping. *Nature* (2020) doi:10.1038/s41586-020-2418-2.
56. Schwartz, S. m1A within cytoplasmic mRNAs at single nucleotide resolution: a reconciled transcriptome-wide map. *RNA* **24**, 1427–1436 (2018).
57. Stelzer, Y., Shivalila, C. S., Soldner, F., Markoulaki, S. & Jaenisch, R. Tracing dynamic changes of DNA methylation at single-cell resolution. *Cell* **163**, 218–229 (2015).
58. Levi-Galibov, O. *et al.* Heat Shock Factor 1-dependent extracellular matrix remodeling mediates the transition from chronic intestinal inflammation to colon cancer. *Nat. Commun.* **11**, 6245 (2020).
59. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
60. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
61. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
62. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
63. Daley, T. & Smith, A. D. Predicting the molecular complexity of sequencing libraries. *Nat.*

Methods **10**, 325–327 (2013).

64. Heinz, S. *et al.* Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).
65. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).
66. Longtine, M. S. *et al.* Additional modules for versatile and economical PCR-based gene deletion and modification in Saccharomyces cerevisiae. *Yeast* **14**, 953–961 (1998).
67. Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T. & Kanemaki, M. An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nat. Methods* **6**, 917–922 (2009).
68. Chu, S. & Herskowitz, I. Gametogenesis in yeast is regulated by a transcriptional cascade dependent on Ndt80. *Mol. Cell* **1**, 685–696 (1998).
69. Kolodziej, P. & Young, R. A. RNA polymerase II subunit RPB3 is an essential component of the mRNA transcription apparatus. *Mol. Cell. Biol.* **9**, 5387–5394 (1989).
70. Kim, M. *et al.* The yeast Rat1 exonuclease promotes transcription termination by RNA polymerase II. *Nature* **432**, 517–522 (2004).
71. Luo, W., Johnson, A. W. & Bentley, D. L. The role of Rat1 in coupling mRNA 3'-end processing to transcription termination: implications for a unified allosteric-torpedo model. *Genes Dev.* **20**, 954–965 (2006).
72. Jamonnak, N. *et al.* Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing. *RNA* **17**, 2011–2025 (2011).
73. Honorine, R., Mosrin-Huaman, C., Hervouet-Coste, N., Libri, D. & Rahmouni, A. R. Nuclear mRNA quality control in yeast is mediated by Nrd1 co-transcriptional recruitment, as revealed by the targeting of Rho-induced aberrant transcripts. *Nucleic Acids Res.* **39**, 2809–2820 (2011).
74. Herrero, A. B. & Moreno, S. Lsm1 promotes genomic stability by controlling histone

mRNA decay. *EMBO J.* **30**, 2008–2018 (2011).

75. Miller, J. E. & Reese, J. C. Ccr4-Not complex: the control freak of eukaryotic cells. *Crit. Rev. Biochem. Mol. Biol.* **47**, 315–333 (2012).
76. Villanyi, Z. *et al.* The Not5 subunit of the ccr4-not complex connects transcription and translation. *PLoS Genet.* **10**, e1004569 (2014).
77. Alhusaini, N. & Coller, J. The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. *RNA* **22**, 709–721 (2016).
78. Carneiro, T. *et al.* Depletion of the yeast nuclear exosome subunit Rrp6 results in accumulation of polyadenylated RNAs in a discrete domain within the nucleolus. *Mol. Cell. Biol.* **27**, 4157–4165 (2007).
79. Dunckley, T. & Parker, R. The DCP2 protein is required for mRNA decapping in Saccharomyces cerevisiae and contains a functional MutT motif. *EMBO J.* **18**, 5411–5422 (1999).
80. Parker, R. RNA degradation in Saccharomyces cerevisae. *Genetics* **191**, 671–702 (2012).
81. Daugeron, M. C., Mauxion, F. & Séraphin, B. The yeast POP2 gene encodes a nuclease involved in mRNA deadenylation. *Nucleic Acids Res.* **29**, 2448–2455 (2001).
82. Chowdhury, A., Kalurupalle, S. & Tharun, S. Pat1 contributes to the RNA binding activity of the Lsm1-7-Pat1 complex. *RNA* **20**, 1465–1475 (2014).
83. Charenton, C. *et al.* A unique surface on Pat1 C-terminal domain directly interacts with Dcp2 decapping enzyme and Xrn1 5'-3' mRNA exonuclease in yeast. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E9493–E9501 (2017).
84. Brown, C. E. & Sachs, A. B. Poly(A) tail length control in Saccharomyces cerevisiae occurs by message-specific deadenylation. *Mol. Cell. Biol.* **18**, 6548–6559 (1998).
85. Wolf, J. & Passmore, L. A. mRNA deadenylation by Pan2-Pan3. *Biochem. Soc. Trans.* **42**, 184–187 (2014).
86. Bodi, Z., Bottley, A., Archer, N., May, S. T. & Fray, R. G. Yeast m6A Methylated mRNAs Are Enriched on Translating Ribosomes during Meiosis, and under Rapamycin

Treatment. *PLoS One* **10**, e0132090 (2015).

87. Rojas, M. *et al.* Yeast Gis2 and its human ortholog CNBP are novel components of stress-induced RNP granules. *PLoS One* **7**, e52824 (2012).
88. Ran, F. A. *et al.* Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
89. Safra, M. *et al.* The m1A landscape on cytosolic and mitochondrial mRNA at single-base resolution. *Nature* **551**, 251–255 (2017).