Direct RNA sequencing reveals m^{6}A modifications on adenovirus RNA are necessary for efficient splicing

Alexander M. Price[1], Katharina E. Hayer[2], Alexa B.R. McIntyre[3,4,a], Nandan S. Gokhale[5,b], Ashley N. Della Fera[1,c], Christopher E. Mason[3,7,8,9], Stacy M. Horner[5,6], Angus C. Wilson[10], Daniel P. Depledge[11,*], and Matthew D. Weitzman[1,12,*]

[1]Division of Protective Immunity and Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia, Philadelphia, PA

[2]Department of Biomedical and Health Informatics, The Children's Hospital of Philadelphia, Philadelphia, PA

[3]Department of Physiology and Biophysics, Weill Cornell Medicine, New York, NY

[4]Tri-Institutional Program in Computational Biology and Medicine, New York, NY

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

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

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

[8]The World Quant Initiative for Quantitative Prediction, Weill Cornell Medicine, New York, NY

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

[10]Department of Microbiology, New York University School of Medicine, New York, NY

[11]Department of Medicine, New York University School of Medicine, New York, NY

[12]Department of Pathology and Laboratory Medicine, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA

Current address: [a]Department of Molecular Life Sciences, University of Zurich; [b]Department of Immunology, University of Washington, Seattle, WA; [c]Biological Sciences Graduate Group, University of Maryland, College Park, MD

*Co-corresponding authors

Contact Information:

daniel.depledge@nyulangone.org; weitzmanm@email.chop.edu

Abstract

Adenovirus is a nuclear replicating DNA virus reliant on host RNA processing machinery. Processing and metabolism of cellular RNAs can be regulated by METTL3, which catalyzes the addition of *N6*-methyladenosine (m^6A) to mRNAs. While m^6A-modified adenoviral RNAs have been previously detected, the location and function of this mark within the infectious cycle is unknown. Since the complex adenovirus transcriptome includes overlapping spliced units that would impede accurate m^6A mapping using short-read sequencing, we profiled m^6A within the adenovirus transcriptome using a combination of meRIP-seq and direct RNA long-read sequencing to yield both nucleotide and transcript-resolved m^6A detection. Although both early and late viral transcripts contain m^6A, depletion of m^6A writer METTL3 specifically impacts viral late transcripts by reducing their splicing efficiency. These data showcase a new technique for m^6A discovery within individual transcripts at nucleotide resolution, and highlight the role of m^6A in regulating splicing of a viral pathogen.

Introduction

Adenovirus is a nuclear-replicating DNA virus with a linear double-stranded genome that is dependent on the host cell machinery for productive infection[1]. To maximize coding capacity of the 36 kilobase genome, adenovirus employs a tightly regulated gene transcription process. Early genes and subsequent late genes are produced from both strands of DNA using cellular RNA polymerase II and the spliceosomal machinery to generate capped, spliced, and polyadenylated messenger RNAs (mRNA). Besides the four canonical ribose nucleosides, adenoviral RNA is also known to contain RNA modifications[2,3]. RNA modifications comprise a family of over one hundred chemical modifications of nucleic acid that can play important roles in both RNA biogenesis and function[4,5]. In eukaryotic messenger RNA, *N6*-methyladenosine (m^6A) is the most prevalent RNA modification besides the 7-methylguanosine cap[6]. This mark has been implicated in regulating multiple processes of RNA maturation, including splicing, polyadenylation, export, translation, and ultimately decay[7–12]. Current understanding suggests that m^6A is added to messenger RNAs co-transcriptionally in the nucleus by recruitment of a writer complex composed of METTL3, METTL14, WTAP, and other accessory proteins to RNA polymerase II[13–15]. These modified mRNAs are bound by reader proteins such as the YTH family (YTHDC1-2, YTHDF1-3)[16,17], various hnRNPs[18,19], and the IGF2B[20] family of RNA binding proteins, which affect various downstream fates of the modified mRNAs. This mark is reversible through the action of erasers, and FTO and ALKBH5 have been proposed to demethylate m^6A[8,21,22].

Shortly after the discovery of m^6A in cellular RNAs, RNAs from several diverse viruses such as adenovirus, Rous sarcoma virus, simian virus 40, herpes simplex virus, and influenza A virus were also shown to contain m^6A in internal regions[2,3,23–26]. In particular, adenovirus serotype 2 was shown to contain m^6A at sites away from the cap[2]. These marks were added to nuclear pre-mRNA and retained in the fully processed cytoplasmic RNA[3]. However, no subsequent studies have elucidated the functions of m^6A modification in adenovirus. With the advent of high-throughput m^6A sequencing methods, interest in viral RNA modifications has been rekindled. RNA viruses such as HIV, influenza A, picornavirus, and various *Flaviviridae* including Zika, dengue, and hepatitis C virus are influenced both positively and negatively by m^6A added via METTL3, and many of these viral RNAs are bound by cytoplasmic YTHDF proteins[27–34]. In hepatitis B virus, m^6A at the same site can both stimulate reverse transcription as well as reduce mRNA stability[35]. For DNA viruses such as SV40 and KSHV, deposition of m^6A on viral RNA transcripts can enhance viral replication[36–39]. Interestingly, multiple labs have published conflicting functions for m^6A within the same viral transcript of KSHV, which suggests

cell-type-specific roles[39]. Of note, recent work using human cytomegalovirus also implicates m^{6}A in controlling aspects of the interferon response, thereby indirectly regulating viral infection[40,41]. Since adenovirus is reliant on cellular polymerases and mechanisms to generate and process its viral RNAs, adenovirus infection provides an excellent opportunity to study the consequences of co-transcriptional m^{6}A addition.

Until recently, sequencing methods to map m^{6}A have relied on antibody-based immunoprecipitations to enrich for methylated RNA within a relatively large nucleotide window (methylated RNA immunoprecipitation sequencing, meRIP-seq or m^{6}A-seq)[42,43]. These techniques are indirect, because antibody-precipitated RNA has to be converted to cDNA before sequencing. Although other RNA modifications can be located due to mutations or truncations resulting from reverse transcription[44–46], these events are not generated in the case of m^{6}A due to efficient base pairing with thymine and uracil. Several techniques have circumvented some of these limitations, such as photo-crosslinking assisted m^{6}A sequencing (PA-m^{6}A-Seq)[47], m^{6}A individual nucleotide resolution crosslinking and immunoprecipitation (miCLIP)[48,49], and RNA digestion via m^{6}A sensitive RNase (MAZTER-seq)[50]. In general, these methods are labor intensive, and require either specialized chemical addition to cell culture, large amounts of input material, or higher unique read counts than meRIP-seq[49]. Furthermore, the antibodies used to precipitate m^{6}A may themselves have sequence or structure biases, and cannot distinguish between m^{6}A and the similar modifications m^{6}A$_m$[22,51]. To this end, the ability to sequence native RNA molecules directly using nanopore arrays provides a new approach to locate RNA modifications. While detecting modified DNA nucleotides is possible using both PacBio and Oxford Nanopore Technologies platforms[52,53], detection of RNA modifications has proven much more challenging. Recently, two groups have shown detection of m^{6}A using nanopores in yeast total RNA and in human cell lines[54,55]. In addition to detecting RNA modifications directly, production of long reads by these platforms provides distinct advantages in the study of gene-dense viral genomes, which encode complex and often overlapping sets of transcripts[56]. To date, the ability to use direct RNA sequencing to map full-length transcripts and their RNA modifications unambiguously has not been realized.

In this study, we found that adenovirus infection does not alter expression of m^{6}A-interacting enzymes but instead concentrates these host proteins at sites of nascent viral RNA synthesis. While meRIP-Seq was able to identify numerous methylated regions on both early and late kinetic classes of viral mRNA, the complex splicing structure and overlapping nature of the adenovirus transcriptome precluded unambiguous transcript assignments and m^{6}A localization by this method alone. To overcome this limitation, we developed a method to predict

sites of m^{6}A modification at single-base resolution within full-length RNA by direct RNA sequencing and used this technique to predict m^{6}A specific to transcript isoforms. While we found that both viral early and late genes are marked by m^{6}A, expression of viral late RNAs in particular decreased dramatically with loss of the cellular m^{6}A writer METTL3. This late gene-biased effect was primarily mediated by decreased RNA splicing efficiency in the absence of METTL3, and could be extended to all of the multiply spliced adenovirus late RNAs. Overall, these results highlight a new technological advancement in long-read RNA sequencing, and reveal that m^{6}A influences the splicing and expression from a viral pathogen.

Results

Nuclear m^{6}A-interacting factors concentrate at viral RNA during adenovirus infection

While it is known that adenovirus RNA transcripts contain m^{6}A, the impact of infection on cellular components involved in writing, reading, or erasing m^{6}A is unknown. The majority of m^{6}A on messenger RNA is installed co-transcriptionally in the nucleus, where adenovirus replicates, by a writer complex composed minimally of METTL3, METTL14, and WTAP[13–15]. To examine whether m^{6}A-interacting enzymes are altered during adenovirus infection, we performed immunoblot analysis over a time-course of infection with adenovirus serotype 5 (Ad5) in A549 lung adenocarcinoma cells (**Figure 1a**). Over the course of infection, levels of the assayed writers (METTL3, METTL14, and WTAP) and readers (YTHDC1, YTHDF1, and YTHDF2) remain unchanged. There was a modest increase in levels of both purported erasers, FTO and ALKBH5, including the appearance of a faster migrating band detected with the ALKBH5 antibody.

Adenovirus is known to recruit specific cellular factors to viral replication centers or mislocalize antiviral cellular factors[57]. To determine if localization of m^{6}A-interacting factors were similarly altered, we performed indirect immunofluorescence microscopy to localize cellular proteins, as well including an antibody against the viral DNA binding protein (DBP) to demarcate viral replication centers (**Figure 1b**). When comparing mock-infected A549 cells to cells infected with Ad5 for 18 hours, we observed that METTL3, METTL14, WTAP, and YTHDC1 relocalized from their diffuse nuclear pattern into ring-like structures surrounding the sites of viral DNA replication marked by DBP. These structures have been previously characterized as sites of viral RNA transcription[58], and was consistent with staining for phospho-serine 2 on the RNA Pol II C-terminal domain, a marker of actively transcribing polymerase (**Figure 1c**). The localization of cytoplasmic readers (YTHDF1, YTHDF2) and demethylases (ALKBH5 and FTO) was mostly unchanged. These data highlight that while adenovirus does not significantly alter expression

levels of known m^{6}A writing enzymes, these nuclear proteins are concentrated at sites of viral RNA synthesis during infection.

Adenovirus transcripts contain METTL3-dependent m^{6}A modifications

While it is known that adenovirus RNAs contain m^{6}A, it is not known exactly where these marks are located or whether adenovirus infection affects m^{6}A localization within host transcripts[2,3]. To address these questions, we performed meRIP-seq on poly(A)-selected RNA from A549 cells that were mock-infected or infected with Ad5 for 24 hours. Strand-specific sequencing was performed on fragmented RNA immunoprecipitated (IP) with an anti-m^{6}A antibody, as well as on total input RNA. We used the MACS2 algorithm to call peaks in IP over input reads for both viral and cellular transcriptomes across three biological replicates (**Figure 2a**). HOMER motif analysis revealed the characteristic DRACH signature (Where D=A/G/U, R=A/G, and H=A/U/C) as the highest ranked motif in cellular m^{6}A peaks in both mock and infected conditions, indicating that the immunoprecipitation was successful and that use of the canonical m^{6}A motif was unperturbed by infection (**Figure 2b**). Furthermore, the general location of m^{6}A peaks in cellular transcripts was unchanged during infection and showed the characteristic stop codon/3'UTR bias in mRNA metagene plots (**Figure 2c**). We next focused on m^{6}A addition to viral transcripts and identified 19 peaks in the forward transcripts and 6 in the reverse transcripts (**Figure 2a**). While these peaks covered every viral kinetic class and transcriptional unit produced by the virus, the MACS2-generated peak areas were very broad, and it was impossible to identify which of many overlapping viral transcripts were m^{6}A methylated.

We next asked whether viral RNAs were m^{6}A-methylated by cellular enzyme METTL3. To achieve this, METTL3 was knocked down by siRNA prior to Ad5 infection of A549 cells and meRIP followed by qRT-PCR of total RNA was performed at 24 hours post-infection (**Figure 2d, 2e**). Our results demonstrate that the amount of m^{6}A-marked RNA available for immunoprecipitation of both the early viral transcript DBP, late viral RNAs generated from the Major Late Promoter (MLP), as well as the positive control cellular transcript *MALAT1*, were all reduced after knockdown of METTL3. These data indicate that viral early and late transcriptional units contain m^{6}A, and that this modification is added by the cellular writer complex that includes METTL3. However, the exact location of the m^{6}A mark could not be assessed by this approach.

A statistical framework predicts sites of m^{6}A methylation using direct RNA sequencing

To address shortcomings of short read-based m^{6}A-sequencing platforms, we sought a technique that would provide both single nucleotide resolution as well as long read length to allow for unambiguous assignment of m^{6}A sites to specific adenovirus mature transcripts. Nanopore sequencing has been used to call DNA modifications directly using differences between measured and expected current values as nucleotides travel through the pore[53]. It is established that signal deviations during direct RNA sequencing can result from the presence of one or more base modifications, and that this leads to an increase in base-call error rate around the modified base[54,59]. The likelihood of a given nucleotide being modified (i.e. carrying an m^{6}A mark) can be assessed by a 2x5 contingency table to examine the distribution of base-calls between two datasets (e.g. m^{6}A positive and m^{6}A negative) at a given genome position, as was recently suggested[60]. Here, a G-test on the distribution of A, C, G, U, and indels provides a score and p-value that requires subsequent (Bonferroni) correction for multiple testing (**Supp Figure 1a**). Expanded methodology and detailed statistical testing are described in the methods.

To produce m^{6}A positive and negative datasets, we generated METTL3 knockout A549 cells using CRISPR-Cas9 and a strategy that included regulated expression of a nuclease-insensitive transgene (**Figure 2f**). We performed direct RNA sequencing using two biological replicates each of RNA collected from parental wild-type A549 cells (WT) and METTL3 knockout A549 cells (M3KO), each infected with Ad5 for 24 hours (**Figure 3a**). Since the datasets are unlinked, each METTL3 KO dataset was compared to each WT parental dataset, yielding four distinct comparisons (**Supp Figure 1b**). Each comparison yielded between 335-452 candidate sites with significant G-test statistics. To account for only the positions at which the base error rate was greater in the WT (m^{6}A positive) dataset, we filtered for a one-fold or greater increase in the ratio of mismatch:match base-calls compared to the M3KO (m^{6}A negative) dataset. This reduced the number of putative sites to 191-217 (**Supp Figure 1b**). Reasoning that multiple sites proximal to a single m^{6}A modification could show significant differences (adjusted $p<0.01$ in error rate), we next calculated the distance between each candidate site and its nearest neighbor candidate in a strand-specific manner (**Supp Figure 1c**). We determined that the majority of candidate sites had at least one neighboring m^{6}A candidate site within five nucleotides, so we implemented an additional filtering strategy to mask all but one candidate site within five nucleotides of a given AC motif (**Figure 3b**), retaining only a single candidate site with the highest G-test statistic. We subsequently plotted the distance (number of nucleotides) from the identified base to the nearest upstream or downstream AC motif (the minimal possible m^{6}A motif). Most significant candidate sites filtered for mismatch:match rates

(94.2-99%) were located within five nucleotides of an AC motif (**Figure 3c**), the maximum distance at which a modified base is thought to affect basecall error[59]. Masking reduced this fraction slightly (92.7-98.9%, **Figure 3c**) and yielded 89-111 distinct predicted m^{6}A sites, of which 53 were conserved across all four dataset comparisons (**Figure 3d,** highlighted by dark blue line in **Figure 2a**). The majority (83.1%) of these 53 sites mapped directly to adenines in AC motifs, with the remainder mapping within four nucleotides (**Figure 3c**). When random non-candidate nucleotides were selected from our dataset we found that the distance to the nearest AC motif was much greater than for m^{6}A-candidate sites, indicating that this was not due to random chance (**Supp Figure 1d**). We subsequently extracted the ten base sequence centered on a candidate m^{6}A site and generated a sequence logo (DRAC, **Figure 3e**) that closely matched the m^{6}A DRACH logo that we confirmed for m^{6}A modifications in the human genome. Concordance between the predicted m^{6}A sites and the previously established motif supports the validity of this unbiased mapping approach.

Exome versus isoform-level m^{6}A analysis

After identifying putative m^{6}A-modified bases within the viral RNA exome using direct RNA-seq, we extended our approach to transcript isoform-level analysis. Here, we aligned our nanopore sequencing reads against 75 distinct transcripts derived from the recently re-annotated adenovirus Ad5 genome and observed that the aligner (MiniMap2[61]) produced secondary alignments for many reads and supplementary alignments for a smaller subset of reads (**Supp Figure 1e**). Secondary alignments indicate that a region of a given sequence read aligns with high confidence against two or more distinct transcripts while supplementary alignments indicate potentially chimeric reads where two segments of the same read align to separate overlapping transcripts (**Figure 4a**). Given the possibility that overlapping transcript isoforms may share the same NNACN motifs but still undergo differential methylation, we reasoned that reads with multiple alignments could reduce sensitivity of detection. We thus retained only reads that produced unique alignments with mapping qualities >0 and had no insertions greater than 20 nucleotides, with the latter requirement intended to exclude reads from incompletely spliced RNAs. While this approach discarded 63-69% of our sequence reads (**Supp Figure 1e**), the specificity achieved by unambiguous isoform-level assignment of sequence reads allowed us to identify 747 putative m^{6}A-modified bases across 47 transcripts. This translated to 352 exome-level sites, of which 204 were conserved across all pairwise comparisons (**Figure 4b,** highlighted by magenta lines in **Figure 2a**). Isoform-level analysis recapitulated the majority of exome-level sites identified above (47/53), while identifying an

additional 50 m^6A sites (**Figure 4c**). This represents a four-fold increase when compared to the equivalent exome-level analysis, demonstrating the greater sensitivity of the isoform-level approach. Furthermore, this strategy allowed us to detect transcripts that had unique m^6A sites, even compared to overlapping transcripts that shared the same potential DRACH motifs (**Figure 4d**). Overall, these data highlight a novel technique that reveals m^6A marks at both single nucleotide and isoform-specific levels, greatly improving our ability to map m^6A modifications across complex viral transcriptomes.

Loss of METTL3 and m^6A methylation differentially impacts late viral gene expression

Since both viral early and late genes are marked by METTL3-dependent m^6A, we asked what role this mark might play in the viral infectious cycle. As reported by others, siRNA-mediated knockdown of METTL3 results in a concomitant loss of METTL14, and vice versa[14]. Upon knockdown of either METTL3 or METTL14, we observed reductions of viral late proteins Hexon, Penton, and Fiber. In contrast, the viral early protein DBP was largely unaffected (**Figure 5a**), even though the DBP transcript is also marked by m^6A (**Figure 2a**). When we assessed viral genome amplification by qPCR, we detected only a minimal change after knockdown of METTL3 or METTL14 (**Figure 5b**), indicating m^6A is not required for early stage adenovirus infection. This contrasts with robust decreases in the number of infectious particles in lysates derived from infected cells knocked down for either METTL3 or METTL14, as measured by plaque assay (**Figure 5c**). Consistent with these findings, when we used reverse transcription coupled with quantitative PCR (qRT-PCR) to measure total viral mRNA levels over a time course of infection (**Figure 5d**), we observed that early viral genes were only modestly (less than two-fold or not significantly) decreased by METTL3 knockdown, whereas viral late genes were significantly reduced (sometimes greater than 10-fold). To corroborate these findings from siRNA-mediated knockdown, we also assayed specific stages of the viral infectious cycle in METTL3 KO A549 cells. Two independently generated CRISPR-mediated knockout cells were infected with Ad5 for 24 hours and assayed for viral early and late genes by immunoblotting and qRT-PCR. Consistent with siRNA findings, accumulation of DBP transcript and protein was not affected, but viral late gene products were significantly reduced in METTL3 KO cells (**Supp Figure 2a,b**). Finally, we blocked methylation by treating cells with the small molecule inhibitor 3-Deazadenosine (DAA) at the time of infection. DAA is an S-Adenosylmethionine synthesis inhibitor that has preferential effects on m^6A deposition when used at low concentrations[62]. Although treatment with DAA reduced early gene protein expression and DNA replication more than METTL3 knockdown or knockout, the most robust

effects observed were reductions in the amount of viral late gene transcripts, protein production, and plaque forming units (**Supp Figure 2c-f**). Overall, these data highlight a preferential effect of m^6A addition at late stages of the viral infectious cycle, with viral early gene transcription and genome replication largely unaffected while late RNAs, late proteins, and infectious progeny production are greatly reduced in the absence of m^6A.

Two groups have recently highlighted that in primary cells m^6A destabilizes the *IFNB1* transcript, such that loss of METTL3 results in increased production of IFNβ[40,41]. When these cells were infected with viruses, including by adenovirus serotype 4, they restricted viral infectivity through indirect priming of the interferon pathway. To determine if loss of METTL3 similarly increased *IFNB1* and innate immune-related transcripts in our experiments, we used siRNA to knock down METTL3 for 48 hours prior to infection with Ad5 either in the presence or absence of ruxolitinub, a JAK/STAT inhibitor that blocks signaling downstream of type-I interferons such as IFNβ (**Figure 6a**). As a positive control for interferon activation and the efficacy of ruxolitinub, uninfected cells were transfected with poly(I:C) as a surrogate viral RNA agonist. Our data demonstrate that in A549 cells, neither type-I interferon (*IFNB1*) or type-III interferon (*IFNL1*) transcripts are induced by Ad5 infection or METTL3 depletion, whereas poly(I:C) transfection induces these transcripts several hundred-fold. Furthermore, expression of interferon-stimulated genes such as *MX1* and *OAS1* was blocked downstream of poly(I:C) transfection by JAK inhibition but not induced by viral infection or METTL3 knockdown. Additionally, blocking the interferon pathway with JAK inhibition did not alter the specific decrease in viral late gene transcripts with METTL3 depletion previously described (**Figure 6b**). Overall these data suggest a specific role for m^6A in altering accumulation of late adenoviral RNA transcripts at late stages of the adenoviral infectious cycle, independent of innate immune activation, in the cell types we used.

Cytoplasmic readers and m^6A erasers do not affect the adenoviral infectious cycle

Since we found that loss of the m^6A writer METTL3 resulted in decreases in viral late gene transcripts and proteins, we next examined whether loss of m^6A erasers led to increases in viral protein production. We used siRNA to knockdown either ALKBH5 or FTO in A549 cells and performed a time-course infection with Ad5. There was no defect in either viral early proteins or late proteins with either ALKBH5 or FTO knockdown (**Supp Figure 3a**). Furthermore, when assaying viral DNA replication by qPCR or infectious particle production by plaque forming units after knockdown of the two erasers, we saw less than 3-fold or non-significant changes (**Supp Figure 3b,c**). Using siRNA in A549 cells, we were able to knock

down the cytoplasmic m^{6}A reader proteins YTHDF1-3 efficiently (**Supp Figure 3d**) but did not observe changes in viral protein production (**Supp Figure 3e**). These results highlight that whatever amount of m^{6}A deposition is necessary for adenoviral function, removal of ALKBH5 and FTO does not increase m^{6}A beyond necessary levels.

Splicing efficiency of late viral RNA is mediated by m^{6}A

We investigated various mechanisms to explain the differential impact of m^{6}A on early and late viral RNAs. To rule out the role of transcription in masking effects of this post-transcriptional modification, we examined early and late promoter activity by profiling nascent RNA transcription with a 4-thiouridine (4sU) labeling approach[63]. After 24 hour infection of A549 cells with Ad5 following knockdown of METTL3, nascent RNA was labeled with 4sU for the last ten minutes of infection. Total RNA was harvested, nascent RNA conjugated with biotin, and precipitated for analysis by qRT-PCR with primers to viral transcripts. We examined unspliced regions of viral RNA near the early E1A and E4 regions, as well as within the tripartite leader of the Major Late Promoter (MLP) that precedes nearly all adenovirus late genes. While relative transcription of the E1A promoter was increased two-fold after METTL3 knockdown, a separate early promoter (E4), and the viral late promoter itself showed no significant change (**Figure 7a**). Furthermore, by comparing 4sU-labeled nascent RNA to total RNA, we observed no significant change in the RNA decay rate of spliced viral early and late genes after METTL3 knockdown (**Figure 7b**). These data imply that adenoviral infections reach the late stage of the viral infectious cycle in cells lacking METTL3, yet still accumulate fewer late viral RNAs.

Besides transcription from the MLP, which is dependent on viral DNA replication, the other major difference between early and late viral transcripts is the sheer number of alternative splicing and polyadenylation events that must take place to generate mature mRNAs. While most early transcripts contain only one splice junction and no more than two potential poly(A) sites, the major late transcriptional unit has over 25 splice acceptors and 5 major sites of polyadenylation, leading to over 20 functional open reading frames. Mounting evidence suggests that m^{6}A and the nuclear reader YTHDC1 can regulate mRNA splicing[7,37]. We used qRT-PCR to measure spliced and unspliced viral RNAs and determined the splicing efficiency of a particular transcript as a ratio of these two products (**Figure 7c**). We examined the m^{6}A-marked transcripts E1A (early gene) and Fiber (late gene) as representative of their kinetic classes. Of note, Fiber is the only late gene that can be incorporated into this type of qPCR assay, since no other intron amplifying primer set within the late gene transcriptional unit is specific to a single gene. Upon knockdown of METTL3 or WTAP, the splicing efficiency of the

E1A gene did not change, whereas the splicing efficiency of Fiber significantly decreased (**Figure 7d**). The loss of Fiber splicing efficiency was phenocopied by METTL3 knockout (**Figure 7e**), as well as by knockdown of the nuclear m^{6}A reader protein YTHDC1 (**Figure 7f**). This decreased splicing efficiency appears to explain the total RNA decrease in viral late genes as well as the decrease in protein expression, since both METTL3 and WTAP knockdown lead to significant losses in viral late gene proteins (**Figure 7g**). A similar result was observed with knockdown of YTHDC1, although to a much lesser degree.

To assay the role of m^{6}A *in cis* on viral transcript splicing we turned to a transgene splicing assay in which we can ablate m^{6}A sites (**Figure 7h**). In this system, we generated a construct with an exogenous plasmid-based promoter to drive expression of an RNA containing the splice donor exon of the viral late tripartite leader, as much intervening Fiber intron as possible without including additional viral genes or splice sites, and the 5' region of Fiber that encompasses the meRIP-seq peak. This viral cassette was fused to a *Renilla* luciferase gene in which all m^{6}A DRACH motifs were silently mutated[34]. These constructs allow transgene expression with wild-type viral context (WT Fiber), or with all 15 DRACH sites present in both exonic and intronic regions silently mutated to ablate deposition of METTL3-dependent m^{6}A (m^{6}A Mut Fiber, mutations shown below in **Figure 7h**). Potential intronic sites were also ablated, since both meRIP-seq and direct RNA sequencing were performed on polyadenylated mRNA and would not have detected potential m^{6}A sites within introns. HeLa cells were depleted of METTL3 with siRNA for 48 hours before the respective transgene plasmids were transfected into uninfected cells and incubated for a further 24 hours. Using this system, splicing efficiency is read out with qRT-PCR using the primers for endogenous Fiber splicing. Upon knockdown of METTL3, the WT Fiber construct showed a decrease in splicing efficiency (**Figure 7i**), similar to that observed during viral infection (**Figure 7d**). Importantly, m^{6}A Mut Fiber showed a similar decrease in splicing efficiency when compared to the WT Fiber construct during control siRNA knockdown, and had no further decrease upon METTL3 knockdown (**Figure 7i**). Overall, these data suggest m^{6}A positively regulates the splicing reaction of viral late transcripts.

METTL3 knockdown globally dysregulates adenoviral late RNA processing

To determine if m^{6}A positively regulates the splicing of all late transcripts, we returned to two orthologous sequencing technologies, short-read sequencing and long-read direct RNA sequencing. Due to the limitations of short-read mapping to viral genomes with overlapping transcripts, standard expression algorithms such as mapped fragments per kilobase transcript per million base pairs (FPKM) cannot be used to quantify these viral transcripts accurately.

Instead, we focused exclusively on the short reads that contained splice junctions, allowing unambiguous assignment to one viral transcript. Since our RNA-seq library preparation was strand-specific, this allowed us to map reads accurately to their strand of origin. Furthermore, we were able to address intron retention of the late transcriptional unit by counting only unspliced reads overlapping the major splice donor in the third exon of the tripartite leader (intron retention, IR). This strategy is schematized for top-strand viral genes (**Figure 8a**). When we applied this technique to adenovirus-infected A549 cells after METTL3 knockdown, we observed that the prevalence of early gene splice junctions was largely unchanged (**Figure 8b**). Of these, E1B-19K appears to be an outlier, since the prevalence of this particular splice junction increases as the virus moves into the late stage, and a potential ORF of this transcript encodes for protein IX, an alternate promoter-driven adenovirus late gene. When we performed this analysis on late gene splice sites, we observed that every late gene originating from the major late promoter was significantly decreased, and that intron retention in this transcriptional unit was increased, consistent with our previous results. The only exception we detected was the first possible alternative splice junction encoding L1-52K, which was unchanged.

While short reads targeting splice junctions provided high depth quantitative results, we turned to long-read sequencing with unambiguous transcript mapping as an orthogonal technique to validate these results. We counted full-length RNA transcripts after infection of METTL3 knockdown A549 cells normalized to the read depth of the experiment (**Figure 8c**). Confirming the short-read sequencing data, we saw that early transcripts were essentially unaffected by loss of METTL3, while the overall abundance of every viral late transcript, except L1-52K, was decreased. Overall, these data highlight that METTL3 promotes the expression and splicing efficiency of the alternatively spliced viral late transcriptional unit.

Discussion

In this work we have identified sites of m^6A methylation in both cellular and viral transcripts during infection with adenovirus. Using a conventional antibody-based approach to detect m^6A, we discovered that both viral early and late kinetic classes of mRNA were modified. To identify m^6A within overlapping viral transcripts, we devised a technique to predict METTL3-dependent methylation sites in full-length RNAs using direct RNA sequencing. This technique identified specific m^6A sites at nucleotide resolution within m^6A-enriched regions, and revealed adenovirus RNA isoform-dependent sites of methylation. Ultimately, we found that m^6A addition in the alternatively spliced viral late transcripts was important for mediating the efficiency of

splicing and accumulation of these messages leading to productive infection. Taken together, this work demonstrates m^{6}A-mediated control over the splicing of viral transcripts.

We also defined a new method for isoform-specific m^{6}A detection within complex transcriptomes. Compared to conventional antibody-based approaches, direct detection of m^{6}A modification in RNA provides many advantages. Focusing on adenovirus RNAs, we show that our technique is highly reproducible between biological replicates with a very low false-positive rate (here defined as candidates that did not map within five nucleotides of an AC dinucleotide). In fact, after candidate site masking, greater than 80% of all sites mapped directly to the adenine within an AC dinucleotide, with the rest often mapping within four nucleotides. This is consistent with the fact that nanopores read RNA in a five nucleotide window[59]. While we demonstrate this method by assaying for METTL3-dependent m^{6}A modifications, theoretically this technique could work for other RNA modifications where the writing enzyme is known, such as methyl-5-cytosine and pseudouridine[44,45]. Compared to other single-nucleotide resolution approaches to m^{6}A detection[48–50], direct RNA sequencing avoids the known biases associated with antibody targeting[51] and cDNA synthesis[42,49] and, unlike endoribonuclease cutting[50], is not specific to a small subset of methylated motifs. Currently, our technique works best with samples of high read-depth, and is thus ideally suited to the study of viruses that take over the RNA processing of their host cell. However, with improvements in nanopore sequencing technologies such as the PromethION which can yield much higher read depth, this technique should be scalable to the study of cellular transcripts as well.

The lack of long-read assays to map m^{6}A methylation has hampered the identification of transcript isoform-specific modifications. While our method still does not enable single molecule resolution, the ability to aggregate similar isoforms in bulk using the long reads produced through direct RNA sequencing allows for isoform-level prediction of m^{6}A sites. Indeed, we demonstrate that isoform-level identification was more sensitive for the detection of potential m^{6}A sites in adenovirus mRNA than the aggregate exome-level data by as much as four-fold. While isoform-level alignment and filtering rely on quality annotations, this technique may need additional customization based on transcript structures in the target of choice. Furthermore, we have only shown the efficacy of this design on polyadenylated messenger RNA. Further studies focusing on non-adenylated or nascent RNAs should prove highly informative for defining the role of m^{6}A-mediated regulation within introns and structural RNAs.

While the influence of m^{6}A on splicing has been widely reported in both viral and cellular contexts[7,8,37,64], this finding has not been without controversy[65]. One study in particular showed that while m^{6}A can be detected in nascent pre-mRNA, these locations do not change upon RNA

maturation and nuclear export[15]. Furthermore, knockout of METTL3 in mouse embryonic stem cells led to no detectable change in cassette exon splicing[15]. In general, our findings agree with this as we do not see the global loss of any particular viral splicing event in the absence of METTL3, only a decrease in total amount of spliced transcripts. However, we did find that m^6A modifications present in adenoviral late transcripts increase the splicing efficiency of these RNAs. This result is consistent with the results of others that have found that the presence of m^6A near splice sites correlates with the rate and efficiency of splicing when studied in a time-resolved fashion in nascent RNA[66]. Furthermore, it is known that unspliced cellular RNAs resident in the nucleus can be targeted for decay by the nuclear exosome[67,68]. Therefore, our hypothesis is that efficient splicing requires m^6A and leads to the accumulation of mature mRNAs that would otherwise be destroyed.

The m^6A modification has also been implicated in the regulation of alternative polyadenylation[47,64]. While the majority of the adenovirus early transcripts that were insensitive to METTL3 knockdown contain only one cleavage and polyadenylation site, the adenovirus late transcriptional unit contains five poly(A) sites, and usage must be regulated to be able to splice the downstream isoforms. It is intriguing to note that transcripts generated from the L1 region upstream of the first polyadenylation site were the viral late transcripts least affected by METTL3 knockdown. While our quantitative PCR reaction to test the splicing efficiency of the L5 region cannot fully rule out the possibility of fewer transcripts reaching this region due to altered poly(A) cleavage, our splicing reporter transgene assay that does not contain upstream poly(A) sites argues against this hypothesis. Splicing and polyadenylation are thought to happen co-transcriptionally[69], and any effect that m^6A might have on coordinating these processes remains an interesting topic for future study.

In summary, we discovered that while both early and late adenoviral genes are marked by m^6A, we only detected a loss of expression within late genes after depletion of METTL3. This was because m^6A increased the efficiency of splicing within adenovirus late transcriptional units, where a plethora of potential splice sites leads to an abundance of alternative splice isoforms. Our data show that even when rates of transcription and RNA decay remain constant, increased splicing efficiency can lead to increased transcript abundance, presumably via protection against nuclear decay mechanisms that target unspliced transcripts. This work highlights how RNA modifications can regulate distinct stages of viral gene expression.

Materials and Methods

Cell Culture

All cell lines were obtained from American Type Culture Collection (ATCC) and cultured at 37 °C and 5% CO_2. All cell lines tested negative for mycoplasma infection and were routinely tested afterwards using the LookOut Mycoplasma PCR Detection Kit (Sigma-Aldrich). A549 cells (ATCC CCL-185) were maintained in Ham's F-12K medium (Gibco, 21127-022) supplemented with 10% v/v FBS (VWR, 89510-186) and 1% v/v Pen/Strep (100 U/ml of penicillin, 100 µg/ml of streptomycin, Gibco, 15140-122). HeLa cells (ATCC CCL-2) and HEK293 cells (ATCC CRL-1573) were grown in DMEM (Corning, 10-013-CV) were grown in DMEM supplemented with 10% v/v FBS and 1% v/v Pen/Strep.

Viral Infection

Adenovirus serotype 5 (Ad5) was originally purchased from ATCC. All viruses were expanded on HEK293 cells, purified using two sequential rounds of ultracentrifugation in CsCl gradients, and stored in 40% v/v glycerol at -20 °C (short term) or -80 °C (long term). Viral stock titer was determined on HEK293 cells by plaque assay, and all subsequent infections were performed at a multiplicity of infection (MOI) of 20 PFU/cell. Cells were infected at 80-90% confluent monolayers by incubation with diluted virus in a minimal volume of low serum (2%) F-12K for two hours. After infection viral inoculum was removed by vacuum and full serum growth media was replaced for the duration of the experiment.

Plasmids, siRNA, and transfections

Fiber-Transgene constructs were created from the previously generated m^6A-null psiCheck2 reporter plasmid[34], where both Firefly and Renilla luciferase genes have all m^6A-DRACH motifs ablated by silent mutations. The original 3'UTR reporter was removed from the end of *Renilla* luciferase using *XhoI* and *NotI* restriction enzymes and this site was healed by the insertion of a recombinant multiple cloning site containing *XhoI*, *AgeI*, and *NotI* using two annealed DNA oligos with complementary sticky ends (Obtained from IDT). From this reporter construct lacking significant 3'UTR, the chimeric intron downstream of the SV40 promoter and upstream of *Renilla* luciferase was excised using *StuI* and *NheI*. Using compatible cut sites, a DNA fragment was cloned into this site encoding the third exon of the Ad5 late gene unit (nucleotides 9,644 to 9,733), all intervening intron that did not contain other Ad5 genes or splice sites (nucleotides 9,733 to 10,631 and 30,867 to 31,042), and the first 482 nucleotides of the

Fiber exon and ORF (31,042 to 31,524) such that the Fiber ORF would continue in frame to the *Renilla* luciferase ORF (WT Fiber). Alternatively, the same fragment but with all 15 DRACH motifs silently mutated was used to generate m^6A Mut Fiber. Both fragments were excised from plasmids created as a PriorityGENE service from GeneWiz from provided sequences. All DNA oligos can be found in **Table 1**. DNA transfections were performed using the standard protocol for Lipofectamine2000 (Invitrogen).

The following siRNA pools were obtained from Dharmacon: non-targeting control (D-001206-13-05), METTL3 (M-005170-01-0005), METTL14 (M-014169-00-0005), WTAP (M-017323-02-0005), YTHDC1 (M-015332-01-0005), FTO (M-004159-01-0005), and ALKBH5 (M-004281-01-0005). The following siRNAs were obtained from Qiagen: YTHDF1 (SI00764715), YTHDF2 (SI04174534), YTHDF3 (SI04205761), FTO (SI04177530), and ALKBH5 (SI04138869). All siRNA transfections were performed using the standard protocol for Lipofectamine RNAiMAX (Invitrogen). Poly(I:C) was provided pre-complexed with transfection reagent (Invivogen tlrl-piclv), and was reconstituted fresh with molecular grade water before added to cells at a concentration of 500 ng/mL.

METTL3 Knockout

Since METTL3 has previously been reported to be an essential gene, a rescue cell line was pre-constructed that contained a CRISPR-insensitive METTL3 transgene under a tetracycline-inducible promoter. In brief, the 2X-Flag Tagged METTL3 transgene[32] was cloned from pEFTak into the BB726[70] entry vector and the NGG Cas9 PAM site at nucleotide position 117 (G to C) was silently mutated via the Stratagene QuickChange Site-Directed Mutagenesis protocol (primer sequence in **Table 1**). This plasmid was then co-transfected into A549 HiLo cells with a transiently expressed Cre Recombinase before being selected by puromycin[70]. This allows integration of METTL3 into a conserved tetracycline-regulated genomic locus, and expression was tested both by immunoblot and immunofluorescence for selection efficiency.

GFP-Cas9 expressing plasmid pX330[71] was constructed to contain the single guide RNA (**Table 1**) from the GECKO CRISPR library. The A549-METTL3-HiLo cell line was then induced to express Cas9-insensitive METTL3 for one day before transfected with the single pX330 plasmid containing both GFP-Cas9 and METTL3 sgRNA. 24 hours post transfection, GFP expressing cells were sorted as single cells by fluorescence-activated cell sorting (FACS) into 96-well plates for clonal expansion. Tetracycline was refreshed biweekly to maintain transgene expression for the entire outgrowth until resulting cell lines reached large enough numbers to be viably frozen. Afterwards, tetracycline was withdrawn and cells were cultured for a further two

weeks before the presence of endogenous METTL3 was assayed for by both immunoblot and immunofluorescence. Two clones were picked that displayed undetectable levels of endogenous METTL3 at the single cell level, and Cas9-mediated lesions were detected by Sanger sequencing of a PCR amplicon derived from genomic DNA (PureLink Genomic DNA kit, Invitrogen).

meRIP-Seq and meRIP-qPCR

A549 cells in 10 cm plates were lysed directly in TRIzol (Thermo Fisher) and RNA was extracted and DNaseI treated (Qiagen). Poly(A) RNA was selected with the Poly(A)Purist MAG Kit before being fractionated (Thermo Fisher) with RNA Fragmentation Reagents (Thermo Fisher) and ethanol precipitated. For meRIP-qPCR the following protocol was identical but RNA was not poly(A) purified or fragmented. MeRIP was performed using the EpiMark N6-methyladenosine Enrichment Kit (NEB) following manufacturer's instructions. The following modifications were made following the previously published protocol[32]. Immunoprecipitations were washed with low and high salt wash buffers and RNA was eluted from the beads with 5 mM m^6A salt (Santa Cruz Biotechnology).

For meRIP-Seq RNA-seq libraries were prepared from both eluate and 10% input mRNA using the TruSeq mRNA library prep kit (Illumina), subjected to quality control (MultiQC), and sequenced on the HiSeq4000 instrument. Both the IP and input samples were mapped to the GRCh37/hg19 genome assembly and the Ad5 genome using the RNA-seq aligner GSNAP[72] (version 2019-09-12) or STAR[73]. We called peaks using MACS2[74] with the following flags: "-q 0.05 -B --call-summits --keep-dup auto --nomodel --extsize 150". HOMER[75] was then used to identify motifs enriched in the identified m^6A-peaks. Metagene analysis and visualization was done using deepTools2[76].

Antibodies, Immunoblotting, and Immunofluorescence

The following primary antibodies were used for cellular proteins: METTL3 (Novus Biologicals H00056339, WB: 1:400, IF: 1:50), METTL14 (Sigma-Aldrich HPA038002, WB: 1:5000, IF: 1:100), WTAP (Proteintech 60188, WB: 1:400, IF: 1:100), YTHDC1 (Abcam ab122340, WB: 1:2000, IF: 1:100), YTHDF1 (Abcam ab99080, WB: 1:500, IF: 1:100), YTHDF2 (Proteintech 24744-1-AP, WB: 1;1000, IF: 1:100), YTHDF3 (Sigma-Aldrich SAB2102735, WB: 1:500), FTO (Abcam ab92821, WB: 1:300, IF: 1:100), ALKBH5 (Sigma-Aldrich SAB1407587, WB: 1:250, IF: 1:100), β-Actin (Sigma-Aldrich A5441-100UL, WB 1:5000), GAPDH (GeneTex 41577, WB:1:20,000), and RNA Pol II p-Ser2 (Abcam ab5095, IF: 1:400). Primary anti-Flag tag

antibody was obtained from Sigma-Aldrich (F7425-.2MG, WB 1:2000). Primary antibodies against viral proteins were obtained from: rabbit polyclonal against adenovirus Hexon, Penton, and Fiber (Gift from J. Wilson, WB 1:10,000), mouse anti-DBP (Gift from A. Levine, Clone: B6-8, WB 1:1000, IF 1:400), polyclonal rabbit anti-DBP (Gift from A. Levine, IF: 1:40,000), mouse anti-E1A (BD 554155, WB: 1:500), and mouse anti-E1B55K (Gift from A. Levine, Clone: 58K2A6, WB 1:500).

For western immunoblotting protein samples were prepared by directly lysing cells in lithium dodecyl sulfate (LDS) loading buffer (NuPage) supplemented with 1% beta-mercaptoethanol (BME) and boiled at 95 °C for 10 min. Equal volumes of protein lysate were separated SDS-PAGE in MOPS buffer (Invitrogen) before being transferred onto a nitrocellulose membrane (Millipore) at 35 V for 90 minutes in 20% methanol solution. Membranes were stained with Ponceau to confirm equal loading and blocked in 5% w/v non-fat milk in TBST supplemented with 0.05% w/v sodium azide. Membranes were incubated with primary antibodies in milk overnight, washed for three times in TBST, incubated with HRP-conjugated secondary (Jackson Laboratories) for 1 h and washed an additional three times in TBST. Proteins were visualized with Pierce ECL or Femto Western Blotting Substrate (Thermo Scientific) and detected using a Syngene G-Box. Images were processed and assembled in Adobe Photoshop and Illustrator CS6.

For immunofluorescence A549 cells were grown on glass coverslips in 24-well plates, mock-infected or infected with Ad5 for 18 hours, washed twice with PBS and then fixed in 4% w/v Paraformaldehyde for 15 minutes. Cells were permeabilized with 0.5% v/v Triton-X in PBS for 10 mins, and blocked in 3% w/v BSA in PBS (+ 0.05% w/v sodium azide) for one hour. Primary antibody dilutions were added to coverslips in 3% w/v BSA in PBS (+ 0.05% w/v sodium azide) for 1 h, washed with 3% BSA in PBS three times, followed by secondary antibodies and 4,6-diamidino-2-phenylindole (DAPI) for one hour. Secondary antibodies were Alexa-Fluor 488, 555, or 568 in anti-mouse or anti-rabbit (Invitrogen). Coverslips were mounted onto glass slides using ProLong Gold Antifade Reagent (Cell Signaling Technologies). Immunofluorescence was visualized using a Zeiss LSM 710 Confocal microscope (Cell and Developmental Microscopy Core at UPenn) and ZEN 2011 software. Images were processed in FIJI and assembled in Adobe Photoshop and Illustrator CS6.

Inhibitors and Small Molecules

3-Deazadenosine was purchased from Sigma-Aldrich (D8296) and resuspended in Dimethyl Sulfoxide (DMSO) before being added to cells at the time of infection to a final

concentration of 25 μM. Ruxolitinub was purchased from Sigma-Aldrich (G-6185) and resuspended in DMSO before being added to cells immediately prior to infection with Ad5 or transfection with poly(I:C) at a final concentration of 4 μM.

RNA Isolation and qPCR

Total RNA was isolated from cells by either TRIzol extraction (Thermo Fisher) or RNeasy Micro kit (Qiagen), following manufacturer protocols. RNA was treated with RNase-free DNase I (Qiagen), either on-column or after ethanol precipitation. RNA was converted to complementary DNA (cDNA) using 1 µg of input RNA in the High Capcity RNA-to-cDNA kit (Thermo Fisher). Quantitative PCR was performed using the standard protocol for SYBR Green reagents (Thermo Fisher) in a QuantStudio 7 Flex Real-Time PCR System (Applied Biosystems). All primers were used at 10 µM and sequences can be found in **Table 1**. All values were normalized by the ΔΔCt method by normalizing first to internal controls such as HPRT1 and GAPDH.

Viral DNA Replication qPCR

Infected cells were harvested at the indicated time points post infection by trypsin and total DNA was harvested using the PureLink Genomic DNA kit (Invitrogen). DNA quantity was assessed by qPCR and SYBR green reagents using primers for genomic regions of Ad5 and normalized to cellular tubulin (See **Table 1** for primer sequence). Entry of the viral genome was assessed at the 4 hour infection time point, and all subsequent values were normalized to this by the ΔΔCt method.

Plaque Assays

Infected cells seeded in 12-well plates were harvested by scraping at the indicated time points and lysed by three cycles of freeze-thawing in liquid nitrogen. Cell debris was removed from lysates by centrifugation at max speed (21,130 g), 4°C, 5 min. Lysates were serially diluted in DMEM supplemented with 2% v/v FBS and 1% v/v Pen/Strep to infect a confluent monolayer of HEK293 cells seeded in 12-well plates. After incubation for 2 h at 37 °C, the infection media was removed, and cells were overlaid with DMEM containing 0.45% w/v SeaPlaque agarose (Lonza) in addition to 2% v/v FBS and 1% v/v Pen/Strep. Plaques were stained using 1% w/v crystal violet in 50% v/v ethanol between 6 to 7 days post-infection.

Metabolic Labeling of RNA for Transcription and Half-life determination

To assess relative RNA transcription rate and RNA half-life, cells were treated with 1 mM 4-thiouridine (4sU; Sigma T4509) for exactly 10 minutes. Infection was stopped and RNA harvested using 1 mL TRIzol (Thermo Fisher), following manufacturer's instructions. A fraction of the total RNA was reserved as input, and the remaining 4sU-labeled nascent RNA was biotinylated using MTSEA-Biotin-XX (Biotium, 90066) as previously described[63,77]. Nascent RNA was separated from unlabeled RNA using MyOne C1 Streptavidin Dynabeads (Thermo Fisher Scientific, 65-001), biotin was removed from nascent RNA using 100 mM dithiothreitol (DTT), and RNA was isopropanol precipitated. One μg of total RNA (T) and an equivalent volume of nascent RNA (N) were converted to cDNA and qPCR was performed as described above. Relative transcription rates were determined by the ΔΔCt method to compare nascent transcript levels between control and siRNA treated cells normalized to nascent GAPDH RNA. RNA half-life was determined using the previously described formula $t0.5 = -t \times [\ln(2)/DR]$ where t is the 4sU labeling time (0.16667 hours) and DR is the decay rate defined as Nascent/Total RNA[78]. Half-lives were normalized to the half-life of GAPDH set at 8 hours as previously determined[79].

RNA-sequencing

Total RNA from three biological replicates of Control knockdown or three biological replicates of METTL3-knockdown A549 cells infected with Ad5 for 24 hours were sent to Genewiz for preparation into strand-specific RNA-Seq libraries. Libraries were then run spread over three lanes of an Illumina HiSeq 2500 using a 150bp paired-end protocol. Raw reads were mapped to the GRCh37/hg19 genome assembly and the Ad5 genome using the RNA-seq aligner GSNAP[72] (version 2019-09-12). The algorithm was given known human gene models provided by GENCODE (release_27_hg19) to achieve higher mapping accuracy. We used R package ggplot2 for visualization. Downstream analysis and visualization was done using deepTools2[76]. Splice junctions were visualized in Integrative Genomics Viewer using the junctions output from GSNAP.

Nanopore Direct RNA Sequencing

Direct RNA sequencing libraries were generated from 680 – 1,800 ng of poly(A) RNA, isolated using the Dynabeads™ mRNA Purification Kit (Invitrogen, 61006). Isolated poly(A) RNA was subsequently spiked with 0.3 µl of a synthetic Enolase 2 (ENO2) calibration RNA (Oxford Nanopore Technologies Ltd.) and prepared for sequencing as described previously[56,59] Sequencing was performed on a MinION MkIb using R9.4.1 (rev D) flow cells (Oxford Nanopore Technologies Ltd.) for 18 – 23 hours (one library per flowcell) and yielded between 1,000,000 –

2,000,000 reads per dataset. Raw fast5 datasets were then basecalled using Guppy v3.2.2 (-f FLO-MIN106 -k SQK-RNA002) with only reads in the pass folder used for subsequent analyses. Sequence reads were aligned against the adenovirus Ad5 genome, using MiniMap2 (-ax splice -k14 -uf --secondary=no), a splice aware aligner[61], with subsequent parsing through SAMtools[80] and BEDtools[81]. Here sequence reads were filtered to retain only primary alignments (SAM FLAG 0 (top strand) or 16 (bottom strand)). Coverage plots (Figure 2 and 4) were generated using the R packages Gviz[82] and GenomicRanges[83]. For isoform-level analysis, sequence reads were aligned against an adenovirus Ad5 transcriptome, derived from the latest annotation, using MiniMap2. Subsequent parsing with SAMtools was used to filter out sequence reads that could not be unambiguously assigned to a single transcript. Here, only primary alignments (SAM flag 0) with mapping qualities (MapQ) > 0 were retained.

Nucleotide Resolution Analysis of m^6A Sites using Direct RNA Sequencing

For both exome and isoform level analyses, the aligned sequence datasets were individually parsed in a stranded manner using BAM Readcount (https://github.com/genome/bam-readcount) to yield base-call distributions (i.e. the number of A, C, G, U, and indels) at each position in the alignment. For each comparison (e.g. M3KO vs WT), a G-test was performed on the base-call distributions at each position in each dataset with subsequent Bonferroni correction for multiple testing. Positions with a read coverage of < 100x in one or both datasets being compared were excluded from this analysis. To further filter the dataset following the G-test, we calculated the ratio of match:mismatch fractions at each position. Candidate sites were those that gave a statistically significant G-test result (bonferroni adjusted $p < 0.01$) and had a one-fold or higher reduction in the match:mismatch fraction in the WT dataset compared to the METTL3 knockout dataset. Subsequent masking was performed by collapsing all sites located within five nucleotides of the same NNACN motif into a single representative, chosen by picking the site with highest G-test score. Where sites were located within ten nucleotides of each other, but further than five nucleotides from an NNACN motif, collapsing was performed by selecting the site with the highest G-test score.

Acknowledgments

We extend special thanks to Ian Mohr (New York University School of Medicine) for his many insightful comments and for support of DPD in part through National Institutes of Health (NIH) grants R01-AI073898 and R01-GM056927. We also thank members of the Weitzman and Mohr/Wilson Labs for insightful discussions and input. We are grateful to A. Berk, P. Branton, G. Ketner, A. Levine, and D. Ornelles for generous gifts of reagents. We thank our colleagues Kristen Lynch, Fange Liu, and Peter Choi for helpful suggestions and careful reading of the manuscript. We thank the UPenn Cell and Developmental Biology Microscopy Core for imaging assistance. This work was supported through NIH grants R21-AI130618 and R21-AI147163 (ACW), R01-AI125416 and R21-AI129851 (SMH and CEM), and R01-AI145266, R01-AI121321, and R01-CA097093 (MDW). SMH was supported by the Burroughs Wellcome Fund (17PRE336700170). NSG was supported by an American Heart Association Predoctoral Fellowship. ABRM was supported by the Natural Sciences and Engineering Research Council of Canada. Additional support came from the NCI T32 Training Grant in Tumor Virology T32-CA115299 (AMP) and Individual National Research Service Award F32-AI138432 (AMP).

Data Availability

Basecalled fast5 (Nanopore) and fastq (Illumina) datasets generated as part of this study can be downloaded from the European Nucleotide Archive (ENA) under the following study accession: PRJEB35652. The authors declare that all other data supporting the findings of this study are available within the article and its Supplementary Information files, or are available from the authors upon request.

Author Contributions

A.M.P., D.P.D., and M.D.W. conceived of the project and designed the experiments; C.E.M., A.C.W. and S.M.H. provided additional input into study design; A.M.P., N.S.G., A.N.D.F, and D.P.D. performed the experiments; D.P.D. designed the nanopore m^6A detection approach; K.E.H., A.B.R.M., and D.P.D. performed computational analyses; A.M.P. and D.P.D analyzed all additional data; A.M.P., D.P.D., and M.D.W. wrote the manuscript; All authors read, edited and approved the final paper.

Figures

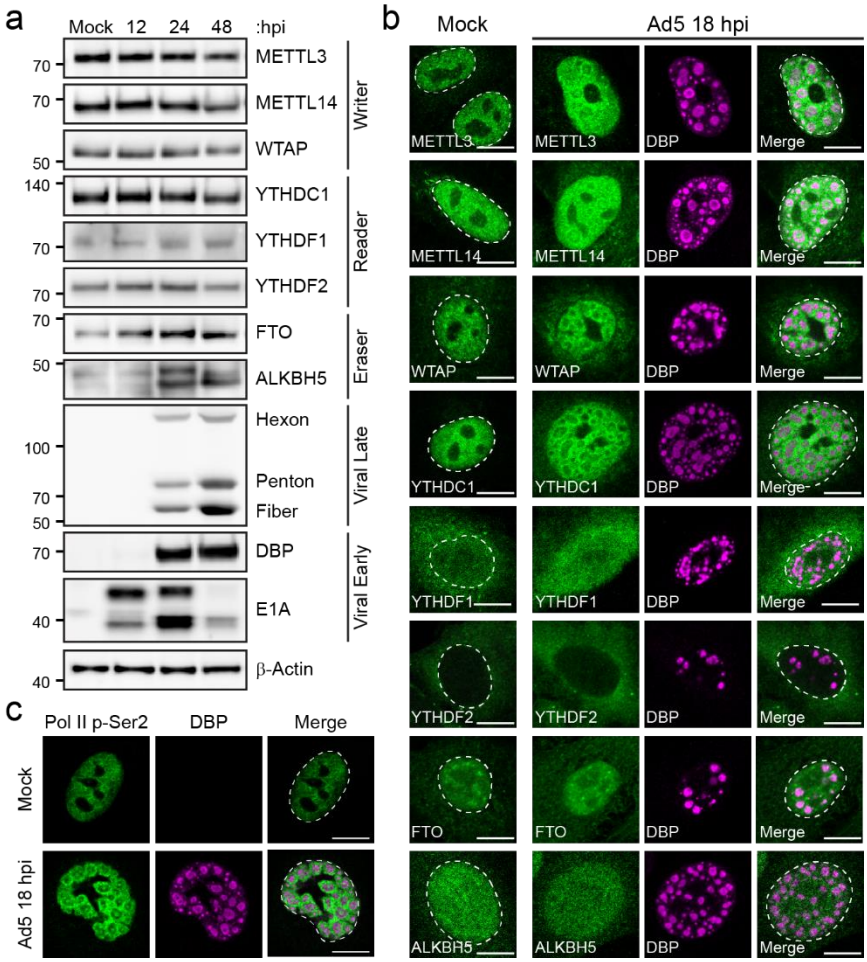


Figure 1. Nuclear m^{6}A-interacting factors concentrate at sites of nascent viral RNA synthesis. (a) Abundance of cellular m^{6}A proteins is unchanged during infection. Immunoblot showing abundance of m^{6}A writers, readers, and putative erasers over a time-course of adenovirus infection. Viral early (E1A and DBP) and late (Hexon, Penton, and Fiber) proteins demonstrate representative kinetic classes. β-Actin is the loading control. Kilodalton size markers shown on the left. **(b)** Confocal microscopy of m^{6}A-interacting proteins (green) in mock-infected or Ad5-infected A549 cells 18 hours post-infection (hpi). DBP (magenta) is the viral DNA binding protein that marks sites of nuclear viral replication centers. The nuclear periphery is shown by a dotted white line as assessed by DAPI staining. Scale bar = 10 μm. **(c)** Confocal microscopy showing the pattern of actively transcribing RNA Polymerase II phosphorylated on serine 2 of CTD (Pol II p-Ser2, green) in mock-infected cells or relative to DBP (magenta) in infected cells. Scale bar = 10 μm.

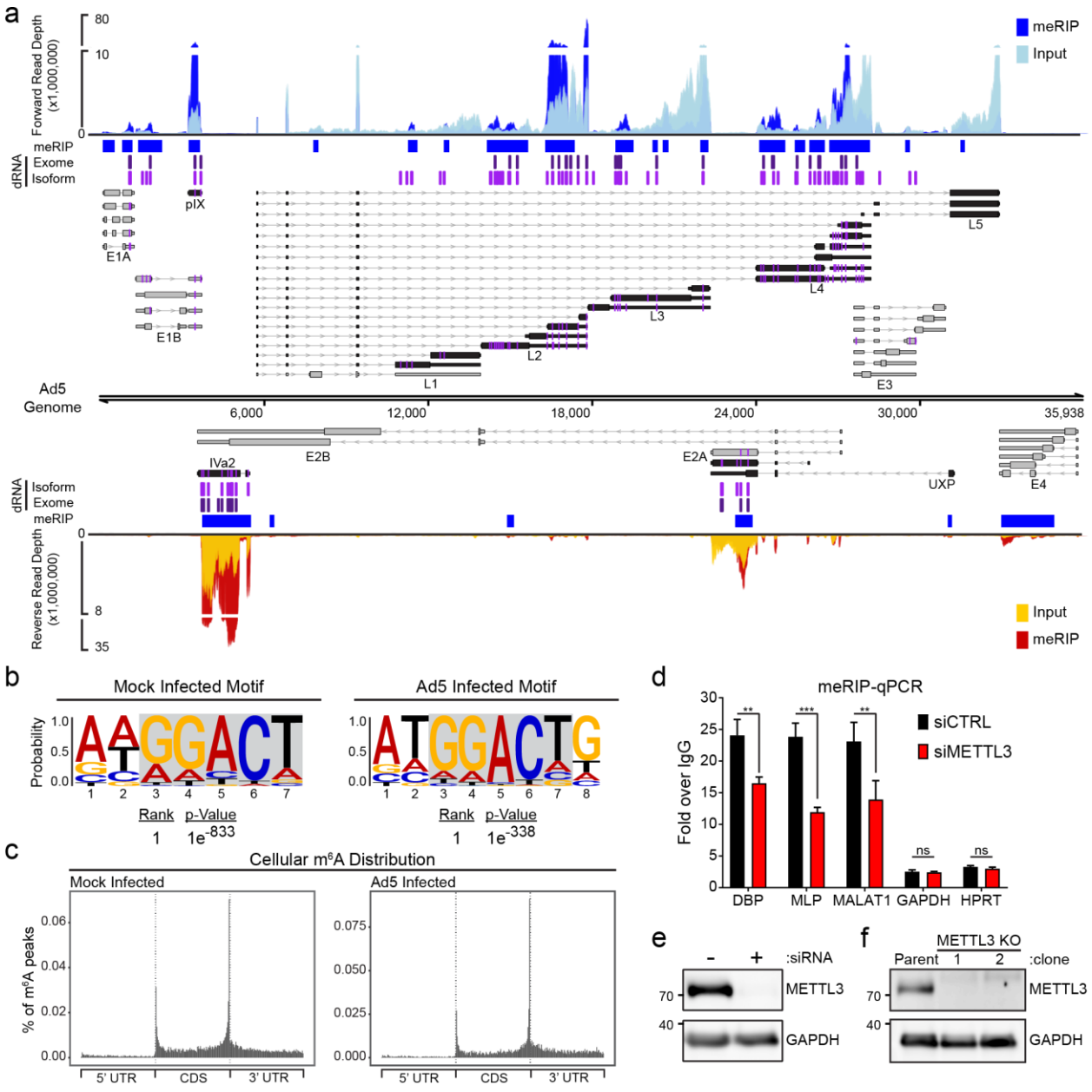


Figure 2. Transcript-specific analysis reveals adenovirus RNAs contain METTL3-dependent m^{6}A modifications. (a) The viral transcriptome is schematized with forward facing transcripts above the genome and reverse transcripts below. Viral gene kinetic classes are color-coded to denote early (grey) or late (black) genes. Lines with arrows denote introns, thin bars are untranslated exonic regions, and thick bars represent open reading frames. The names of each viral transcriptional unit are shown below the transcript cluster. meRIP-Seq was performed in triplicate on Ad5-infected A549 cells at 24 hpi. Representative meRIP data (blue/red) and total input RNA (light blue/yellow) sequence coverage is plotted against the adenovirus genome. Peaks containing increased meRIP-seq signal over input were called with MACS2 and denoted by blue boxes. Using direct RNA (dRNA) sequencing, full-length RNAs

were sequenced from A549 parental cells or METTL3 knockout cells infected with adenovirus for 24 hours. Specific m^6A sites were predicted by comparing the nucleotide error rate of dRNA sequence data from WT to KO cells. Indicated in purple vertical lines are individual adenosines predicted to be modified by m^6A that reach statistical significance when applied to all RNA that maps to a single nucleotide of the Ad genome (dRNA Exome). All Ad5-mapping transcripts were binned into unique full-length reads spanning entire transcript isoform and the same m^6A prediction was applied on a transcript-by-transcript basis. Magenta vertical lines indicate predicted m^6A residues found on the transcriptome level (dRNA Isoform). Additionally, the position of m^6A present in each viral transcript is highlighted in magenta directly on the transcript schemes. **(b)** HOMER nucleotide motif analysis of MACS2 called peaks in cellular meRIP-seq data from Mock or Ad5-infected samples. **(c)** Metagene analysis of m^6A-peak distribution across cellular mRNA molecules containing 5' and 3' untranslated regions (UTR) and coding sequence (CDS) in Mock or Ad5-infected samples. **(d)** meRIP-qRT-PCR was performed on total RNA isolated from Ad5-infected control or METTL3 knockdown A549 cells 24 hours post-infection. **(e)** Immunoblot showing knockdown efficiency of METTL3 in A549 cells. **(f)** Representative immunoblot showing two clones generated from Cas9-mediated knockout of METTL3 in A549 cells. For all assays, significance was determined by Student's T-test, ** $p \leq 0.01$, *** $p \leq 0.001$, ns=not significant. Experiments are representative of three biological replicates and error bars represent standard deviation.

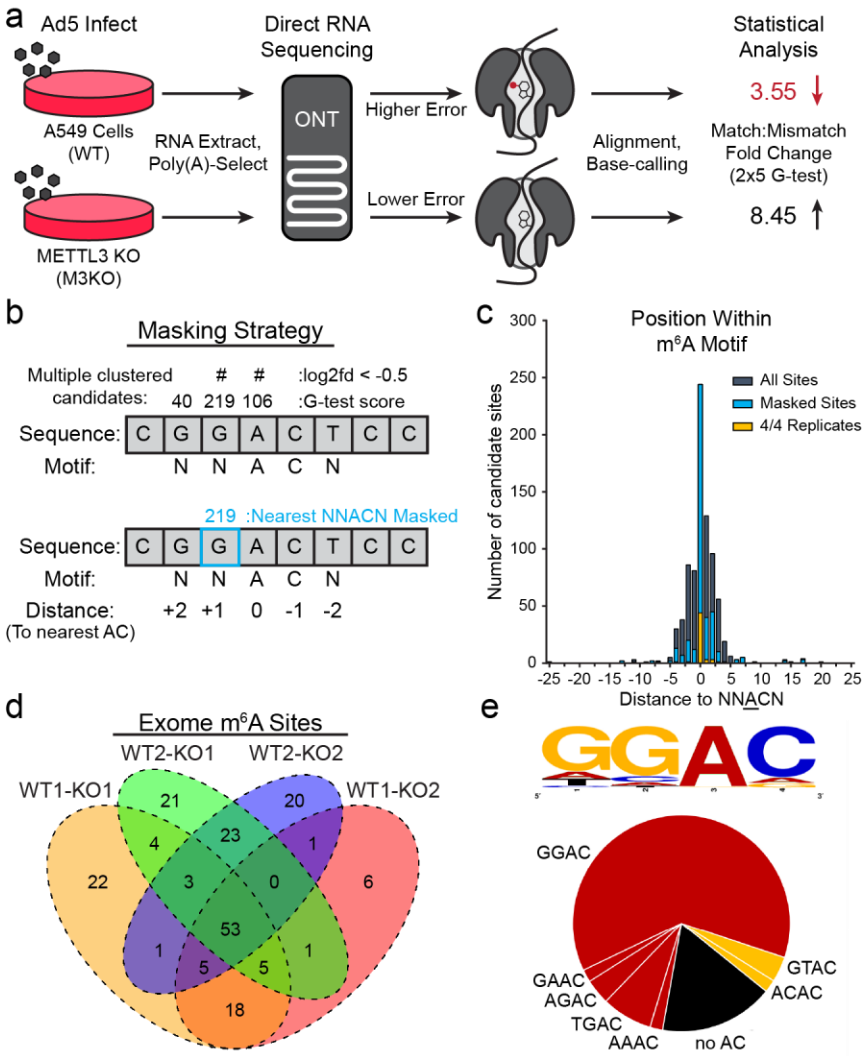


Figure 3. A statistical framework for m^{6}A detection using direct RNA sequencing. (a) Schematic diagram of proposed strategy to detect m^{6}A sites using direct RNA sequencing. RNAs generated in WT cells (METTL3 positive) will contain m^{6}A modifications, while RNA from METTL3 KO cells (M3KO) will not. As modified RNA passes through the pore there will be a higher error rate during base-calling compared to unmodified RNA. When compared to the reference transcriptome, the aggregate fold change in the Match:Mismatch ratio will be lower in at nucleotides containing m^{6}A. **(b)** Proposed strategy for masking neighboring candidates. Here, three sites within five nucleotides of an AC produce significant G-test scores. All but the candidate giving the highest G-test statistic are masked. **(c)** For each significant candidate site with a one-fold or greater difference in the match:mismatch ratio, the distance to the nearest AC motif was calculated and plotted (grey). This was repeated after masking neighboring candidates (blue) and for the 53 genome-level sites identified across all four comparisons (gold). **(d)** Comparisons between WT and M3KO (KO) datasets yields 184 putative m^{6}A modified bases (post-masking) of which 53 are consistently detected across all four comparisons. **(e)** The consensus four-nucleotide motif (DRAC) for putative m^{6}A modified bases in the Ad5 transcriptome is predominantly comprised of five common DRACH motifs.

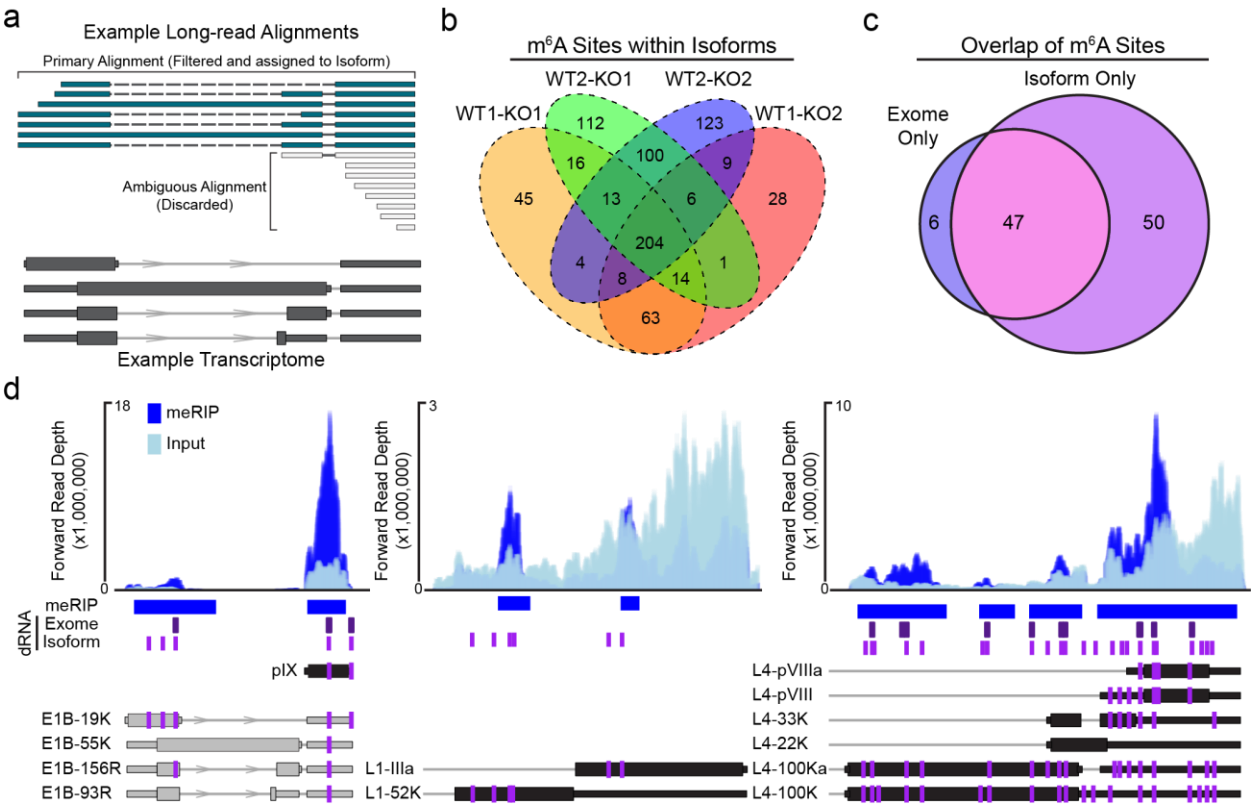
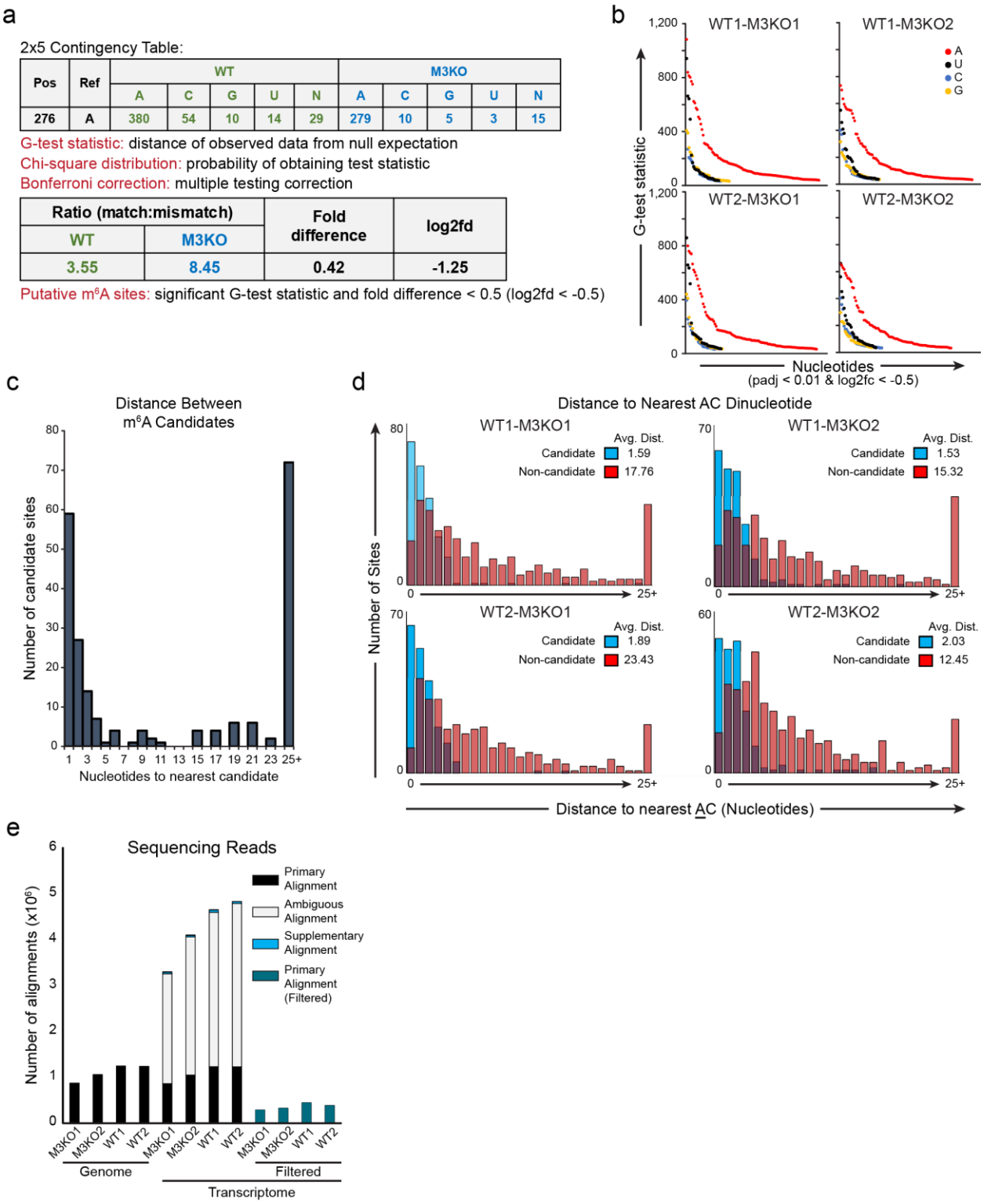


Figure 4. Exome versus isoform-level m^{6}A analysis. (a) Example of read filtering. Overlapping RNA isoforms are shown in dark grey. Sequence reads that map unambiguously (teal) are filtered and retained for m^{6}A analysis. Sequence reads with multiple primary alignments (grey) are discarded if no single alignment is considered superior. **(b)** Isoform-level comparisons between WT and M3KO datasets yields 747 candidates (post-masking) of which 204 are consistently detected across all four comparisons. **(c)** Overlap of all m^{6}A sites detected in 4/4 exome-mapped replicates and 4/4 isoform-mapped replicates. Isoform only candidates were collapsed to unique genomic loci. **(d)** Representative plots showing isoform-specific m^{6}A predictions on adenovirus transcripts. MACS2 peaks and meRIP-seq signal (blue) is shown on top. Direct RNA (dRNA) predicted m^{6}A sites at the exome-level (purple) or isoform-level (magenta) are shown below meRIP peaks as vertical lines. The positioning of m^{6}A within individual viral transcripts from Figure 2 are shown as magenta lines. Transcript names are shown to the left of each transcript.



Supp Figure 1. Direct RNA m^{6}A prediction clusters on AC dinucleotides. (a) Example 2x5 contingency table and statistical tests used to predict m^{6}A sites at nucleotide resolution by comparing the error rate between RNA generated in WT (METTL3-positive) versus METTL3 Knockout (M3KO) cells. **(b)** Distributions of the significant G-test statistic for each nucleotide

with additional filtering for a one-fold or greater difference in the match:mismatch ratio between the m^6A positive (WT) and m^6A negative (M3KO) datasets. All G-test statistic p-values are corrected for multiple comparisons using Bonferroni correction. **(c)** The distance (in nucleotides) between nearest neighbor candidates indicates that multiple candidate sites (significant G-test statistics) lie proximal to many putative m^6A modified bases. **(d)** Distance to the nearest A nucleotide in an AC dinucleotide within the Ad5 genome for candidate m^6A sites (blue) versus 200 randomly selected non-candidate sites (red). Average distance for the four distinct comparisons are shown in the legend. **(e)** Alignment of sequence reads to the adenovirus genome results in a single primary alignment per read (SAM flag 0). By comparison, aligning against the transcriptome can yield multiple primary alignments (termed secondary alignments, SAM flag 256, grey) and/or supplementary alignments (SAM flag 2048, blue). Subsequent filtering to retain only reads that unambiguously align to a single transcript (SAM flag 0, mapping quality >0, teal) results in approximately 70% of the sequence datasets being discarded.

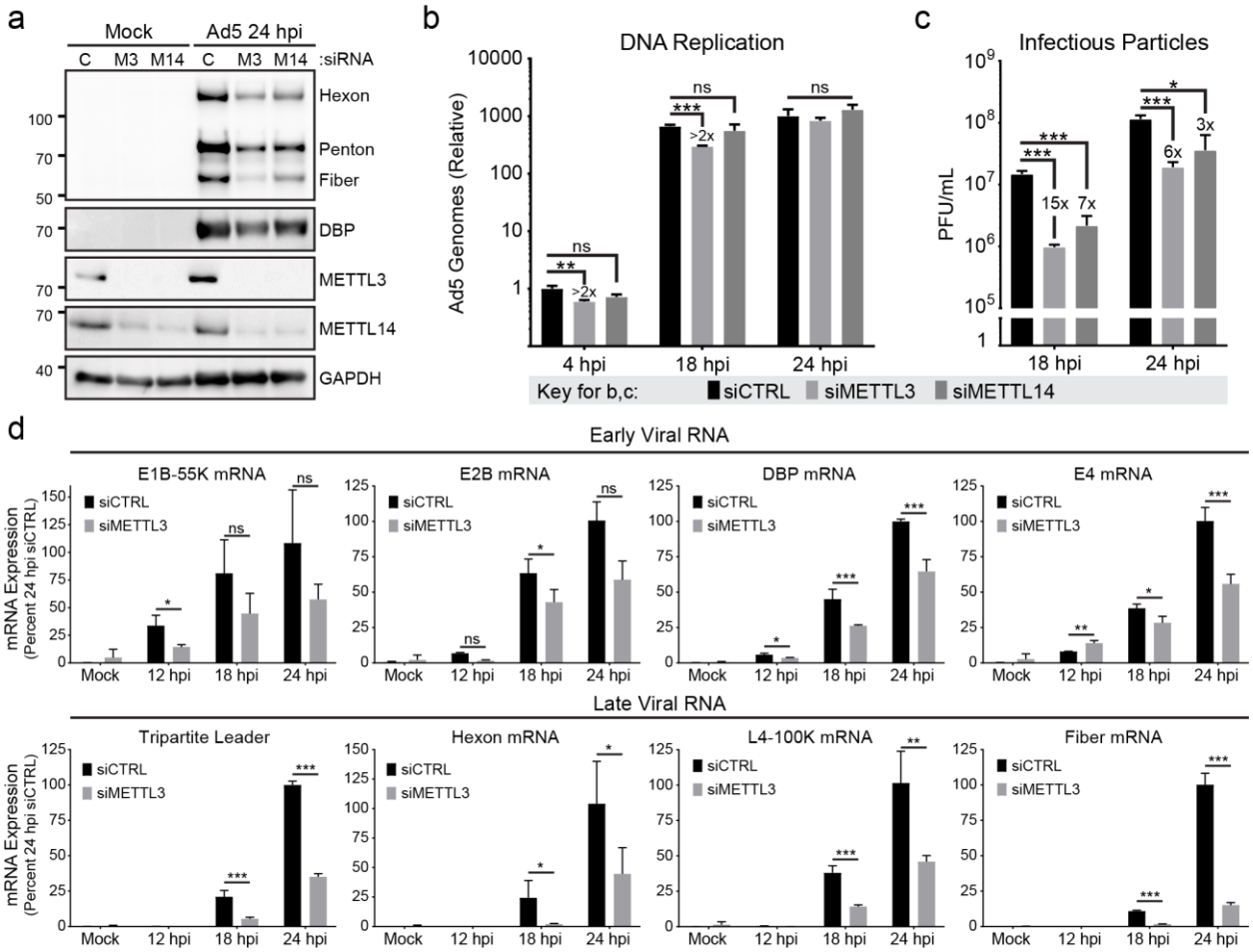
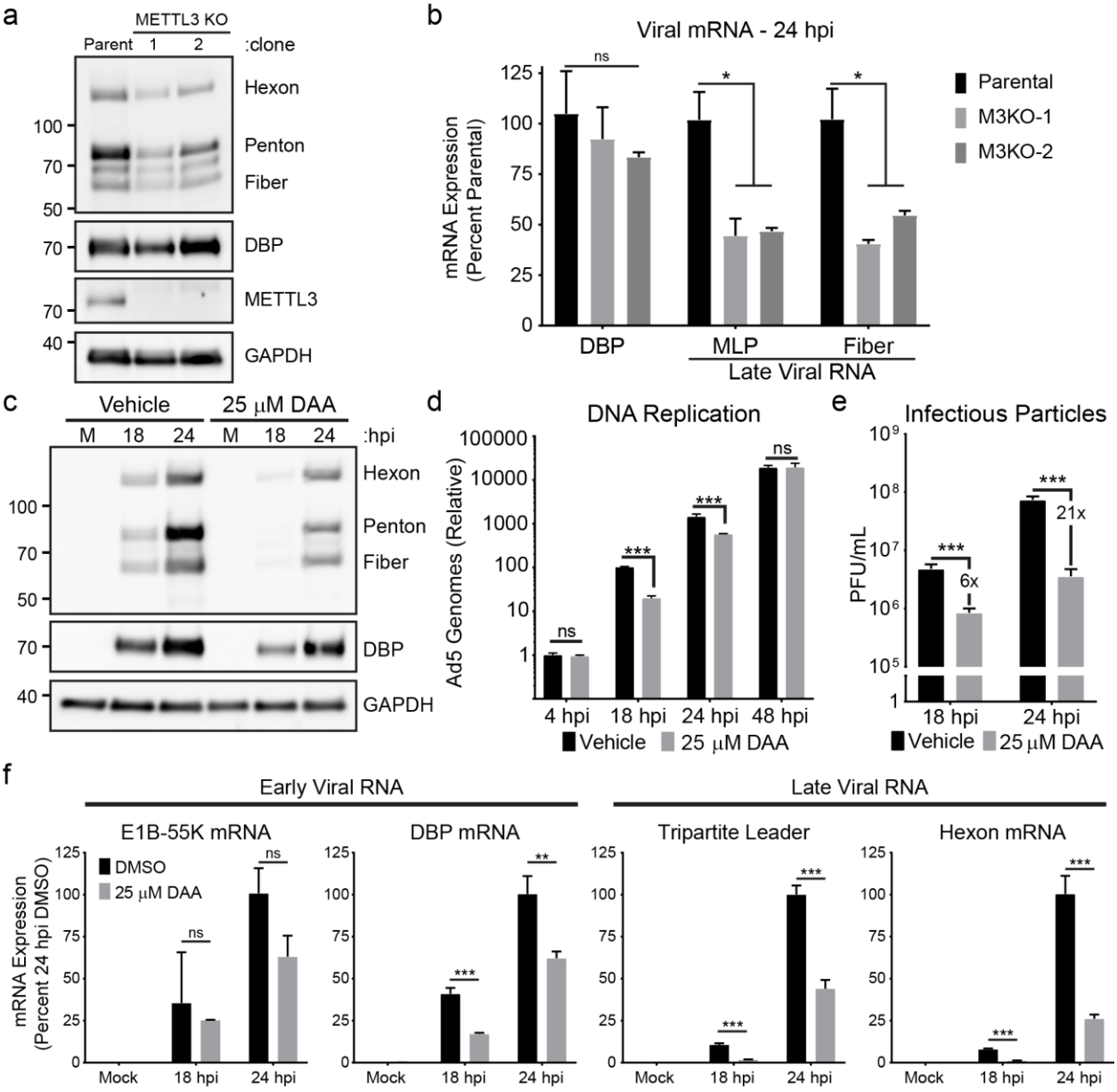


Figure 5. Loss of METTL3 differentially impacts late viral gene expression. (a) A549 cells were transfected with control siRNA (C), or siRNA targeting METTL3 (M3) or METTL14 (M14) for 48 hours before Ad5 infection. Viral late proteins were analyzed at 24 hpi with polyclonal antibody against structural proteins (Hexon, Penton, and Fiber) or early protein DBP. **(b)** Viral DNA replication was measured in biological triplicate samples with qPCR at different times after infecting A549 cells depleted of METTL3 or METTL14 by siRNA. Fold-change in genome copy was normalized to the amount of input DNA at 4 hpi in control siRNA treated cells. **(c)** Infectious particle production was measured at different times by plaque assay after infecting A549 cells depleted of METTL3 or METTL14 by siRNA. **(d)** Viral RNA expression was measured by qRT-PCR in a time-course of infection after control or METTL3 knockdown for 48 hours. Viral early transcripts are shown on the top, while late stage transcripts are shown on the bottom. For all assays significance was determined by Student's T-test, * p≤0.05, ** p≤0.01, *** p≤0.001, ns=not significant. Experiments are representative of three biological replicates and error bars represent standard deviation.



Supp Figure 2. METTL3 knockdown is phenocopied by methylation inhibition or METTL3 knockout. (a) Two independent clones of Cas9-generated METTL3 knockout cells (M3KO) or the parental A549 cells were infected with Ad5 for 24 hours and analyzed for early protein expression (DBP) or late proteins (Hexon, Penton, and Fiber) by immunoblot. **(b)** Viral early and late RNA expression was measured by qRT-PCR at 24 hpi with Ad5. Samples include three cell culture triplicates of A549 Parental or M3KO clones 1 and 2. **(c)** Viral early and late protein expression was analyzed by immunoblot over a time-course of infection in cells treated with 25 μM m^6A-inhibitor 3-Deazaadenosine (DAA) or vehicle control. **(d)** Viral DNA replication was measured with qPCR in biological triplicate samples at different times after infecting A549 cells and concurrently treating with vehicle or 25 μM DAA. Fold-change in genome copy was normalized to the amount of input DNA at 4 hpi in vehicle treated cells. **(c)** Infectious particle production was measured by plaque assay at different times after infecting A549 cells and

concurrently treating with vehicle or 25 μM DAA. **(f)** Viral RNA expression was measured by qRT-PCR in a time-course of infection after concurrently treating with vehicle or 25 μM DAA. Viral early transcripts are shown on the left, while late stage transcripts are shown on the right. Data include three biological replicates and significance was determined by Student's T-test, * p≤0.05, ** p≤0.01, *** p≤0.001, ns=not significant. Experiments are representative of three biological replicates and error bars represent standard deviation.

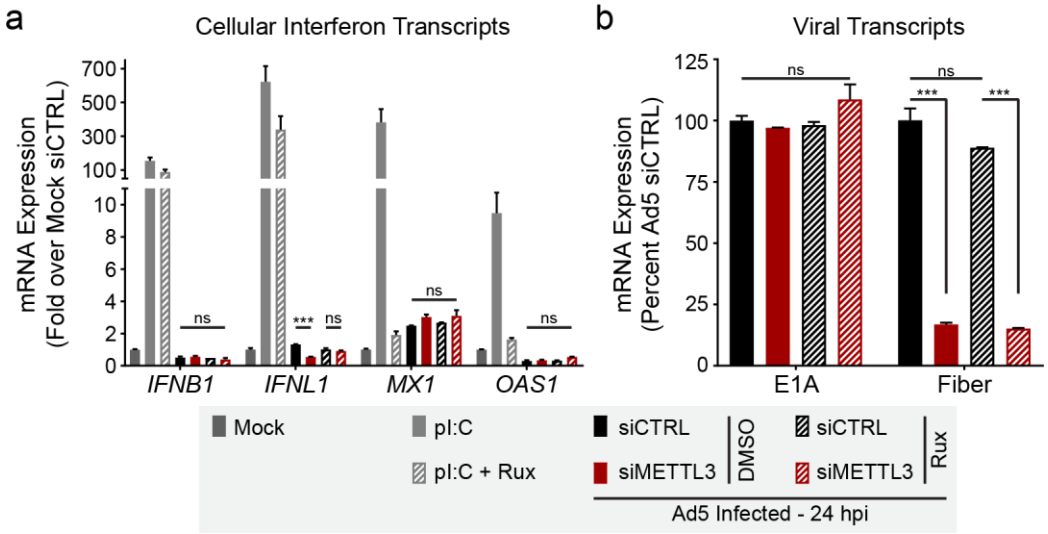
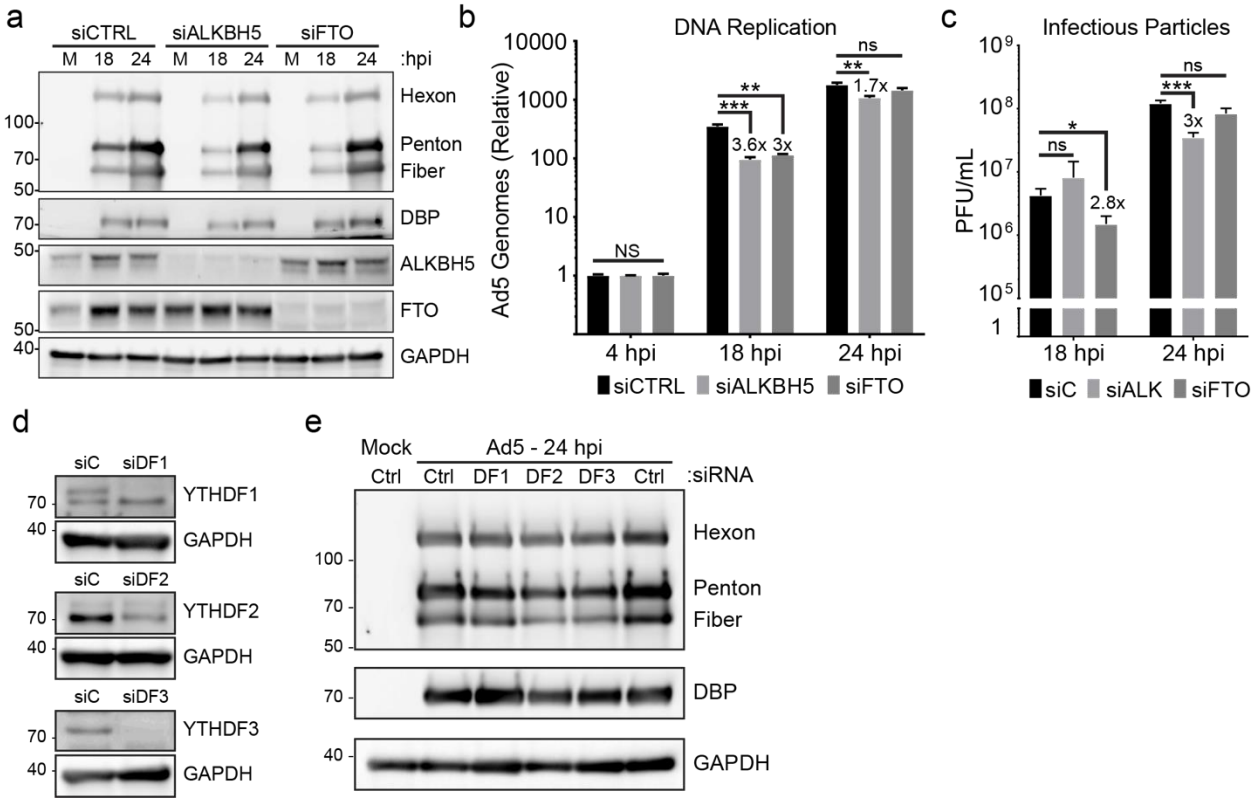


Figure 6. Interferon gene expression is not induced by METTL3 knockdown or adenovirus infection. (a) A549 cells were mock treated (dark grey bars) or transfected with the viral RNA agonist poly(I:C) for 24 hours (light grey bars). Concurrently, A549 cells were depleted of METTL3 with siRNA (red bars) or control treated with siRNA (black bars) before being infected with Ad5 for 24 hours. All sample were either treated with 4 μM Ruxolitinub (hashed bars) or vehicle control (solid bars) for the duration of infection or poly(I:C) transduction. Interferon RNAs were measured by qRT-PCR and displayed as fold-change over mock treated A549 cells. **(b)** Viral transcripts were quantified by qRT-PCR to assay expression of a viral early gene (E1A) or a viral late gene (Fiber). For all assays, significance was determined by Student's T-test, *** p≤0.001, ns=not significant.



Supp Figure 3. Cytoplasmic readers and m^{6}A-erasers have minimal effects on the adenoviral infectious cycle. (a) A549 cells were transfected with control siRNA (CTRL), or siRNA targeting ALKBH5 or FTO for 48 hours before Ad5 infection. Viral late proteins were analyzed over a time-course of infection with polyclonal antibody against structural genes (Hexon, Penton, and Fiber) or the early protein DBP. **(b)** Viral DNA replication was measured in biological triplicate samples with qPCR over a time-course of infection in A549 cells depleted of ALKBH5 or FTO by siRNA. Fold-change in genome copy was normalized to the amount of input DNA at 4 hpi in control siRNA treated cells. **(c)** Infectious particle production was measured by plaque assay at different times post-infection after infecting A549 cells depleted of ALKBH5 or FTO by siRNA. Data includes three biological replicates. **(d)** siRNA-mediated knockdown of YTHDF1, YTHDF2, or YTHDF3 in A549 cells. **(e)** A549 cells were transfected with control siRNA (Ctrl), or siRNA targeting YTHDF1, -2, or -3 (DF1, DF2, DF3) for 48 hours prior to Ad5 infection. Viral early and late proteins were analyzed by immunoblot with the indicated antibodies. For all assays, significance was determined by Student's T-test, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, ns=not significant.

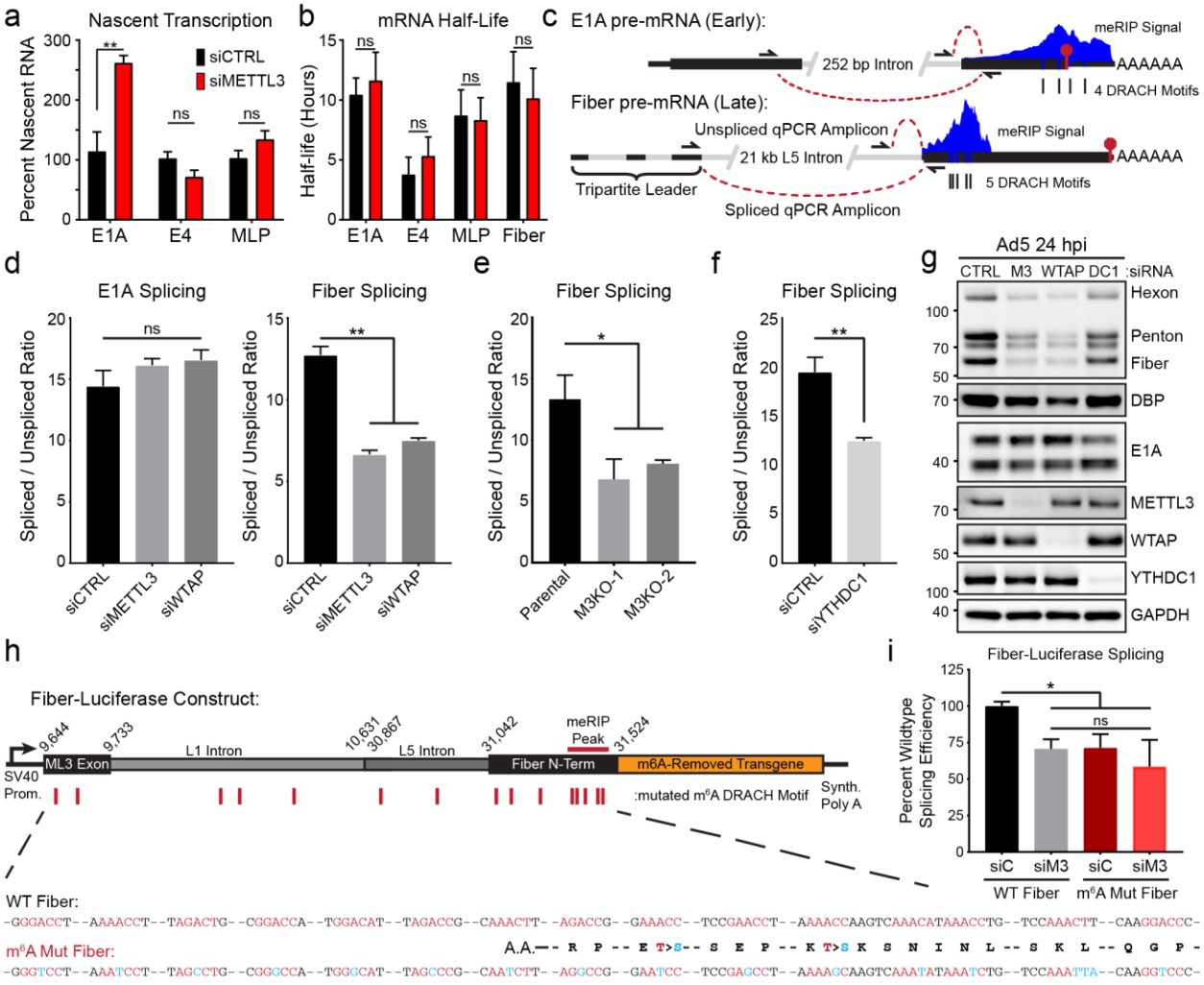


Figure 7. Splicing efficiency of late viral RNA is mediated by m^{6}A. (a) Nascent transcription was analyzed by labeling RNA with 1 mM 4-thiouridine (4sU) for exactly ten minutes at 24 hpi for infections of A549 cells transfected with control (siCTRL) or siRNA-mediated knockdown of METTL3 (siMETTL3). Nascent 4sU-labeled RNA was extracted for use in qRT-PCR for analysis of relative transcription rates of two viral early genes (E1A and E4) and the tripartite leader (MLP) found in all Ad5 late transcripts. Samples include four biological replicates. **(b)** mRNA half-life was approximated using 4sU-labeled nascent RNA and total input RNA levels for viral transcripts normalized to GAPDH. Samples include four biological replicates. **(c)** Schematic of the early transcript E1A and the late transcript Fiber that both contain m^{6}A sites. Three primers allow for the distinction between spliced and unspliced PCR products that can be analyzed by qRT-PCR. **(d)** Splice efficiency as defined by the relative ratio of spliced to unspliced transcripts of E1A and Fiber were analyzed by qRT-PCR in A549 cells infected with Ad5 for 24 hours after depletion of METTL3 or WTAP. Data represents three biological experiments. **(e)** Fiber splice efficiency was analyzed in Parental A549 cells or two independent METTL3 KO cell lines. **(f)** Splice efficiency of Fiber was analyzed after siRNA-mediated depletion of the nuclear m^{6}A reader YTHDC1. **(g)** Immunoblot showing viral late proteins Hexon, Penton, and Fiber as well as viral early proteins DBP and E1A. A549 cells were depleted of METTL3 (M3), WTAP,

YTHDC1 (DC1) or control siRNA (siC) for 48 hours prior to infection with Ad5 for 24 hours. **(h)** Schematic design for a luciferase construct that expresses the third adenovirus tripartite leader to Fiber splice site with intervening L1 and L5 adenoviral intron. The 5' fragment of Fiber that contains the m^6A signal peak was fused in-frame to a *Renilla* luciferase transgene where all m^6A DRACH motifs have been ablated by silent mutation. A matching construct was generated with all 15 potential m^6A DRACH motifs ablated by silent or synonymous mutation (m^6A Mut Fiber). **(i)** HeLa cells were control transfected (siC) or depleted of METTL3 (siM3) by siRNA for 48 hours before transfection with either WT Fiber or m^6A Mut Fiber plasmid. At 24 hours after the second transfection, splicing efficiency of the transgene was assayed using Fiber-specific primers. Data represent four biological replicates. For all assays significance was determined by Student's T-test, * $p \leq 0.05$, ** $p \leq 0.01$, ns=not significant. Error bars represent standard deviation.

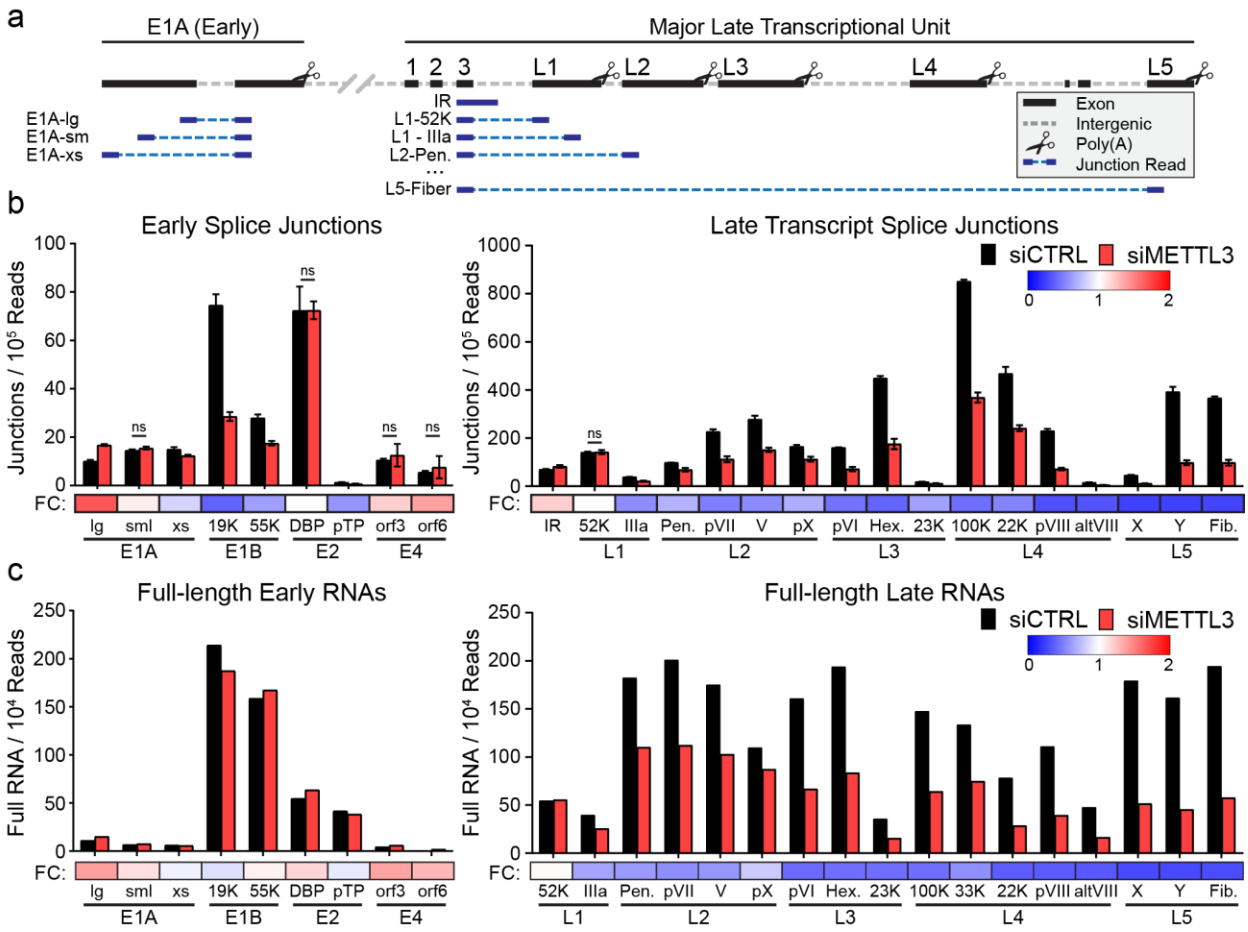


Figure 8. METTL3 knockdown globally dysregulates adenoviral late RNA processing. (a) Schematic showing how junction-containing splice reads generated by Illumina sequencing can be used to predict specific transcript abundances when genes overlap. Short reads (dark blue) aligning specifically to viral exons (black) were filtered for the presence of a splice junction (dashed light blue line) that was only present in one viral transcript. **(b)** Splice junction containing reads indicative of adenovirus transcripts present in Illumina RNA-Seq data generated after infection of A549 cells where METTL3 was depleted by siRNA. Splice junction read-depth was normalized to the total amount of reads mapping to both human and viral RNA per library. Early viral transcripts are shown on the left, MLP-derived late transcripts shown on the right. Fold-change (FC) between control siRNA and siMETTL3 for each transcript was plotted as a heatmap below the bar chart. Data depict three biological replicates with error bars showing standard deviation. For all assays significance was determined by Student's T-test, everything un-labeled was significant with $p \leq 0.05$. Otherwise specific comparisons were labeled ns (not significant). **(c)** Independently derived RNA was sequenced by ONT after METTL3 knockdown and adenovirus infection to yield full-length RNA sequences indicative of the labeled early and late viral transcripts. Fold-change (FC) between control siRNA and siMETTL3 for each transcript was plotted as a heatmap below the bar chart.

References

1. Arnold J. Berk. Adenoviridae. in *Fields Virology* (eds. David M. Knipe & Peter M. Howley) vol. 2 1704–1731 (Wolters Kluwer Health/Lippincott Williams & Wilkins, 2013).
2. Sommer, S. *et al.* The methylation of adenovirus-specific nuclear and cytoplasmic RNA. *Nucleic Acids Res* **3**, 749–65 (1976).
3. Chen-Kiang, S., Nevins, J. R. & Darnell, J. E. N-6-methyl-adenosine in adenovirus type 2 nuclear RNA is conserved in the formation of messenger RNA. *J Mol Biol* **135**, 733–52 (1979).
4. Zhao, B. S., Roundtree, I. A. & He, C. Post-transcriptional gene regulation by mRNA modifications. *Nat Rev Mol Cell Biol* **18**, 31–42 (2017).
5. Williams, G. D., Gokhale, N. S. & Horner, S. M. Regulation of Viral Infection by the RNA Modification N6-Methyladenosine. *Annu. Rev. Virol.* **6**, 235–253 (2019).
6. Yue, Y., Liu, J. & He, C. RNA N6-methyladenosine methylation in post-transcriptional gene expression regulation. *Genes Dev* **29**, 1343–55 (2015).
7. Xiao, W. *et al.* Nuclear m(6)A Reader YTHDC1 Regulates mRNA Splicing. *Mol Cell* **61**, 507–19 (2016).
8. Bartosovic, M. *et al.* N6-methyladenosine demethylase FTO targets pre-mRNAs and regulates alternative splicing and 3'-end processing. *Nucleic Acids Res* **45**, 11356–11370 (2017).
9. Roundtree, I. A. *et al.* YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* **6**, (2017).
10. Wang, X. *et al.* N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**, 117–20 (2014).
11. Lin, S., Choe, J., Du, P., Triboulet, R. & Gregory, R. I. The m(6)A Methyltransferase METTL3 Promotes Translation in Human Cancer Cells. *Mol Cell* **62**, 335–45 (2016).

12. Wang, X. *et al.* N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**, 1388–99 (2015).
13. Liu, J. *et al.* A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**, 93–5 (2014).
14. Schwartz, S. *et al.* Perturbation of m6A writers reveals two distinct classes of mRNA methylation at internal and 5' sites. *Cell Rep* **8**, 284–96 (2014).
15. Ke, S. *et al.* m(6)A 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).
16. Meyer, K. D. & Jaffrey, S. R. Rethinking m(6)A Readers, Writers, and Erasers. *Annu Rev Cell Dev Biol* **33**, 319–342 (2017).
17. Hazra, D., Chapat, C. & Graille, M. m6A mRNA Destiny: Chained to the rhYTHm by the YTH-Containing Proteins. *Genes* **10**, 49 (2019).
18. Liu, N. *et al.* N(6)-methyladenosine-dependent RNA structural switches regulate RNA-protein interactions. *Nature* **518**, 560–564 (2015).
19. Zhou, K. I. *et al.* Regulation of Co-transcriptional Pre-mRNA Splicing by m6A through the Low-Complexity Protein hnRNPG. *Mol. Cell* **76**, 70-81.e9 (2019).
20. Huang, H. *et al.* Recognition of RNA N(6)-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol* **20**, 285–295 (2018).
21. Zheng, G. *et al.* ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* **49**, 18–29 (2013).
22. Mauer, J. *et al.* Reversible methylation of m 6 A m in the 5′ cap controls mRNA stability. *Nature* **541**, 371–375 (2017).
23. Kane, S. E. & Beemon, K. Precise localization of m6A in Rous sarcoma virus RNA reveals clustering of methylation sites: implications for RNA processing. *Mol. Cell. Biol.* **5**, 2298–2306 (1985).

24. Lavi, S. & Shatkin, A. J. Methylated simian virus 40-specific RNA from nuclei and cytoplasm of infected BSC-1 cells. *Proc Natl Acad Sci U A* **72**, 2012–6 (1975).
25. Moss, B., Gershowitz, A., Stringer, J. R., Holland, L. E. & Wagner, E. K. 5'-Terminal and internal methylated nucleosides in herpes simplex virus type 1 mRNA. *J Virol* **23**, 234–9 (1977).
26. Krug, R. M., Morgan, M. A. & Shatkin, A. J. Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *J Virol* **20**, 45–53 (1976).
27. Kennedy, E. M. *et al.* Posttranscriptional m(6)A Editing of HIV-1 mRNAs Enhances Viral Gene Expression. *Cell Host Microbe* **19**, 675–85 (2016).
28. Lichinchi, G. *et al.* Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of T cells. *Nat Microbiol* **1**, 16011 (2016).
29. Tirumuru, N. *et al.* N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein expression. *Elife* **5**, (2016).
30. Courtney, D. G. *et al.* Epitranscriptomic Enhancement of Influenza A Virus Gene Expression and Replication. *Cell Host Microbe* **22**, 377-386 e5 (2017).
31. Hao, H. *et al.* N6-methyladenosine modification and METTL3 modulate enterovirus 71 replication. *Nucleic Acids Res.* **47**, 362–374 (2019).
32. Gokhale, N. S. *et al.* N6-Methyladenosine in Flaviviridae Viral RNA Genomes Regulates Infection. *Cell Host Microbe* **20**, 654–665 (2016).
33. Lichinchi, G. *et al.* Dynamics of Human and Viral RNA Methylation during Zika Virus Infection. *Cell Host Microbe* **20**, 666–673 (2016).
34. Gokhale, N. S. *et al.* Altered m6A Modification of Specific Cellular Transcripts Affects Flaviviridae Infection. *Mol. Cell* (2019) doi:10.1016/j.molcel.2019.11.007.
35. Imam, H. *et al.* N6-methyladenosine modification of hepatitis B virus RNA differentially regulates the viral life cycle. *Proc. Natl. Acad. Sci. U. S. A.* **115**, 8829–8834 (2018).

36. Tsai, K., Courtney, D. G. & Cullen, B. R. Addition of m6A to SV40 late mRNAs enhances viral structural gene expression and replication. *PLOS Pathog.* **14**, e1006919 (2018).

37. Ye, F., Chen, E. R. & Nilsen, T. W. Kaposi's Sarcoma-Associated Herpesvirus Utilizes and Manipulates RNA N6-Adenosine Methylation to Promote Lytic Replication. *J Virol* (2017) doi:10.1128/JVI.00466-17.

38. Tan, B. *et al.* Viral and cellular N 6 -methyladenosine and N 6 ,2′- O -dimethyladenosine epitranscriptomes in the KSHV life cycle. *Nat. Microbiol.* **3**, 108–120 (2018).

39. Hesser, C. R., Karijolich, J., Dominissini, D., He, C. & Glaunsinger, B. A. N6-methyladenosine modification and the YTHDF2 reader protein play cell type specific roles in lytic viral gene expression during Kaposi's sarcoma-associated herpesvirus infection. *PLOS Pathog.* **14**, e1006995 (2018).

40. Rubio, R. M., Depledge, D. P., Bianco, C., Thompson, L. & Mohr, I. RNA m6 A modification enzymes shape innate responses to DNA by regulating interferon β. *Genes Dev.* **32**, 1472–1484 (2018).

41. Winkler, R. *et al.* m 6 A modification controls the innate immune response to infection by targeting type I interferons. *Nat. Immunol.* **20**, 173–182 (2019).

42. Meyer, K. D. *et al.* Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**, 1635–46 (2012).

43. Li, X., Xiong, X. & Yi, C. Epitranscriptome sequencing technologies: decoding RNA modifications. *Nat. Methods* **14**, 23–31 (2017).

44. Carlile, T. M. *et al.* Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature* **515**, 143–146 (2014).

45. Squires, J. E. *et al.* Widespread occurrence of 5-methylcytosine in human coding and non-coding RNA. *Nucleic Acids Res.* **40**, 5023–5033 (2012).

46. Li, X. *et al.* Base-Resolution Mapping Reveals Distinct m1A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol. Cell* **68**, 993-1005.e9 (2017).

47. Chen, K. *et al.* High-Resolution N6-Methyladenosine (m6A) Map Using Photo-Crosslinking-Assisted m6A Sequencing. *Angew. Chem. Int. Ed.* **54**, 1587–1590 (2015).

48. Ke, S. *et al.* A majority of m6A residues are in the last exons, allowing the potential for 3′ UTR regulation. *Genes Dev.* **29**, 2037–2053 (2015).

49. Linder, B. *et al.* Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat. Methods* **12**, 767–772 (2015).

50. Garcia-Campos, M. A. *et al.* Deciphering the “m6A Code” via Antibody-Independent Quantitative Profiling. *Cell* **178**, 731-747.e16 (2019).

51. Zeng, Y. *et al.* Refined RIP-seq protocol for epitranscriptome analysis with low input materials. *PLOS Biol.* **16**, e2006092 (2018).

52. Flusberg, B. A. *et al.* Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* **7**, 461–465 (2010).

53. McIntyre, A. B. R. *et al.* Single-molecule sequencing detection of N 6-methyladenine in microbial reference materials. *Nat. Commun.* **10**, 1–11 (2019).

54. Liu, H. *et al.* Accurate detection of m 6 A RNA modifications in native RNA sequences. *Nat. Commun.* **10**, 1–9 (2019).

55. 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* rna.072785.119 (2019) doi:10.1261/rna.072785.119.

56. Depledge, D. P. *et al.* Direct RNA sequencing on nanopore arrays redefines the transcriptional complexity of a viral pathogen. *Nat. Commun.* **10**, 1–13 (2019).

57. Reyes, E. D. *et al.* Identifying Host Factors Associated with DNA Replicated During Virus Infection. *Mol Cell Proteomics* **16**, 2079–2097 (2017).

58. Pombo, A., Ferreira, J., Bridge, E. & Carmo-Fonseca, M. Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells. *EMBO J* **13**, 5075–85 (1994).

59. Garalde, D. R. *et al.* Highly parallel direct RNA sequencing on an array of nanopores. *Nat. Methods* **15**, 201–206 (2018).

60. Parker, M. T. *et al.* Nanopore direct RNA sequencing maps an Arabidopsis N6 methyladenosine epitranscriptome. *bioRxiv* 706002 (2019) doi:10.1101/706002.

61. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

62. Fustin, J.-M. *et al.* RNA-Methylation-Dependent RNA Processing Controls the Speed of the Circadian Clock. *Cell* **155**, 793–806 (2013).

63. Russo, J., Heck, A. M., Wilusz, J. & Wilusz, C. J. Metabolic labeling and recovery of nascent RNA to accurately quantify mRNA stability. *Methods San Diego Calif* **120**, 39–48 (2017).

64. Kasowitz, S. D. *et al.* Nuclear m6A reader YTHDC1 regulates alternative polyadenylation and splicing during mouse oocyte development. *PLOS Genet.* **14**, e1007412 (2018).

65. Darnell, R. B., Ke, S. & Darnell, J. E. Pre-mRNA processing includes N6 methylation of adenosine residues that are retained in mRNA exons and the fallacy of 'RNA epigenetics'. *RNA N. Y. N* **24**, 262–267 (2018).

66. Louloupi, A., Ntini, E., Conrad, T. & Ørom, U. A. V. Transient N-6-Methyladenosine Transcriptome Sequencing Reveals a Regulatory Role of m6A in Splicing Efficiency. *Cell Rep.* **23**, 3429–3437 (2018).

67. Doma, M. K. & Parker, R. RNA Quality Control in Eukaryotes. *Cell* **131**, 660–668 (2007).

68. Bresson, S. M., Hunter, O. V., Hunter, A. C. & Conrad, N. K. Canonical Poly(A) Polymerase Activity Promotes the Decay of a Wide Variety of Mammalian Nuclear RNAs. *PLOS Genet.* **11**, e1005610 (2015).

69. Misra, A. & Green, M. R. From polyadenylation to splicing: Dual role for mRNA 3' end formation factors. *RNA Biol.* **13**, 259–264 (2015).

70. Khandelia, P., Yap, K. & Makeyev, E. V. Streamlined platform for short hairpin RNA interference and transgenesis in cultured mammalian cells. *Proc Natl Acad Sci U A* **108**, 12799–804 (2011).

71. Cong, L. *et al.* Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–23 (2013).

72. Wu, T. D., Reeder, J., Lawrence, M., Becker, G. & Brauer, M. J. GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality. *Methods Mol. Biol. Clifton NJ* **1418**, 283–334 (2016).

73. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinforma. Oxf. Engl.* **29**, 15–21 (2013).

74. Zhang, Y. *et al.* Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

75. 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).

76. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160-165 (2016).

77. Poling, B. C., Price, A. M., Luftig, M. A. & Cullen, B. R. The Epstein-Barr virus miR-BHRF1 microRNAs regulate viral gene expression in cis. *Virology* **512**, 113–123 (2017).

78. Price, A. M., Messinger, J. E. & Luftig, M. A. c-Myc Represses Transcription of Epstein-Barr Virus Latent Membrane Protein 1 Early after Primary B Cell Infection. *J. Virol.* **92**, (2018).

79. Dolken, L. *et al.* High-resolution gene expression profiling for simultaneous kinetic parameter analysis of RNA synthesis and decay. *RNA* **14**, 1959–72 (2008).

80. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25**, 2078–2079 (2009).

81. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

82. Hahne, F. & Ivanek, R. Visualizing Genomic Data Using Gviz and Bioconductor. in *Statistical Genomics: Methods and Protocols* (eds. Mathé, E. & Davis, S.) 335–351 (Springer, 2016). doi:10.1007/978-1-4939-3578-9_16.

83. Lawrence, M. *et al.* Software for Computing and Annotating Genomic Ranges. *PLOS Comput. Biol.* **9**, e1003118 (2013).