Long-read RNA sequencing reveals allele-specific N^6-methyladenosine modifications

Dayea Park and Can Cenik

Genome Res. 2025 35: 999-1011 originally published online October 29, 2024
Access the most recent version at doi:10.1101/gr.279270.124

Method

Long-read RNA sequencing reveals allele-specific N^6-methyladenosine modifications

Dayea Park and Can Cenik

Department of Molecular Biosciences, University of Texas at Austin, Austin, Texas 78712, USA

Long-read sequencing technology enables highly accurate detection of allele-specific RNA expression, providing insights into the effects of genetic variation on splicing and RNA abundance. Furthermore, the ability to directly sequence RNA enables the detection of RNA modifications in tandem with ascertaining the allelic origin of each molecule. Here, we leverage these advantages to determine allele-biased patterns of N^6-methyladenosine (m^6A) modifications in native mRNA. We used human and mouse cells with known genetic variants to assign the allelic origin of each mRNA molecule combined with a supervised machine learning model to detect read-level m^6A modification ratios. Our analyses reveal the importance of sequences adjacent to the DRACH motif in determining m^6A deposition, in addition to allelic differences that directly alter the motif. Moreover, we discover allele-specific m^6A modification events with no genetic variants in close proximity to the differentially modified nucleotide, demonstrating the unique advantage of using long-reads and surpassing the capabilities of antibody-based short-read approaches. This technological advance will further our understanding of the role of genetics in determining mRNA modifications.

[Supplemental material is available for this article.]

Allele-specific expression (ASE) refers to the differences in gene expression from two alleles of the same gene. Such an imbalance in expression can contribute to phenotypic variation and the pathophysiology of diseases (Gicquel et al. 2005; de la Chapelle 2009; Castel et al. 2015, 2020; Fan et al. 2020). In mammalian development, a predominant form of ASE, genomic imprinting, plays a critical role as only one allele is expressed. Allele-specific DNA methylation and chromatin composition are two well-established epigenetic systems that control imprinted gene expression (Fournier et al. 2002; Singh et al. 2010; Prendergast et al. 2012).

ASE can reflect differential rates of transcription, mRNA stability, or alternative splicing between the alleles due to genetic variation (Nembaware et al. 2008; Pai et al. 2012; Amoah et al. 2021). That is, local genetic variants can influence transcriptional or posttranscriptional processes to modulate the mRNA abundance of each allele (Robles-Espinoza et al. 2021). While the significance of allele-specific RNA expression is well-acknowledged, allele-specific RNA modification remains underexplored.

N^6-methyladenosine (m^6A), the most prevalent RNA modification of mRNAs, has been suggested to impact diverse mechanisms to regulate gene expression (Lee et al. 2020; He and He 2021). Various interactions with the methyltransferase complex or m^6A reader proteins impact several steps of mRNA metabolism, including splicing, export, translation, recruitment of RNA-binding proteins, and stability (Lin and Gregory 2014; Akhtar et al. 2021; Jiang et al. 2021; Wang et al. 2022).

Transcriptome-wide patterns of m^6A RNA modifications have typically been studied using short-read sequencing coupled with either antibody-dependent methods such as methylated RNA immunoprecipitation sequencing (MeRIP-seq) (Meyer et al. 2012) or enzymatic/chemical approaches (Garcia-Campos et al. 2019; Meyer 2019a; Song et al. 2021). Among these methods, MeRIP-seq remains the most popular choice despite its limitations leading to elevated false-positive rates, attributable to nonspecific antibody binding (Helm et al. 2019; McIntyre et al. 2020; Zhang et al. 2021). Furthermore, all short-read sequencing strategies to detect m^6A are inherently limited to aggregate measurements and are incapable of quantification at an individual molecular level.

In contrast, Oxford Nanopore Technologies (ONT) RNA sequencing enables direct detection of RNA modifications such as m^6A with single-molecule resolution (Garalde et al. 2018). The electric signal recorded by the ONT sequencing platform is altered by the presence of RNA modifications (Garalde et al. 2018; Workman et al. 2018; Dassi et al. 2020). Subsequently, machine learning methods can use the electronic current signal intensity to identify potential m^6A sites from long-read data (Hendra et al. 2022).

The ability to directly detect m^6A modifications on the ONT RNA sequencing platform provides a unique opportunity to combine these advantages with the ability of long-read sequencing to facilitate ASE analysis. Long-read sequencing improves upon the fundamental limitations of short-read sequencing for allele-specific analysis by detecting an increased number of single-nucleotide polymorphisms (SNPs) on a read, enabling its precise allelic assignment (Cho et al. 2014). This feature has been leveraged to characterize the genetic effects of rare and common variants in the transcriptome (Glinos et al. 2022). Furthermore, long-read sequencing enables comprehensive analysis of splicing (Tilgner et al. 2015, 2018; Joglekar et al. 2021) which has fundamental importance for determining mRNA modifications due to their dependence on splicing patterns and transcript architecture (Cenik et al. 2017; Yang et al. 2022; He et al. 2023).

Here, we introduce a novel approach harnessing ONT direct RNA sequencing (ONT DRS) to surmount the persistent constraints of m^6A detection methods for allele-specific analyses.

Corresponding author: ccenik@austin.utexas.edu
Article published online before print. Article, supplemental material, and publication date are at https://www.genome.org/cgi/doi/10.1101/gr.279270.124.

Our findings establish long-read sequencing of RNA as a robust solution for allele-specific m^{6}A modification (ASM) analysis.

Results

ONT DRS enables the identification of allele-specific m^{6}A modifications in hybrid mESCs

We leveraged DRS to simultaneously determine the allelic origin of each molecule along with its m^{6}A modification status. To achieve high accuracy of allelic assignment of individual molecules, we used mouse embryonic stem cells (mESCs) that were derived from a cross of two highly genetically diverse mouse strains (C57BL/6J × CAST/EiJ, B6 × CAST) (Balasooriya and Spector 2022). Using DRS, we generated two replicates of 2.3 and 2.2 million reads from these hybrid mESCs.

To assess our ability to accurately detect m^{6}A modifications, we generated mESC clones where methyltransferase-like 3 (*Mettl3*), the major methyltransferase for m^{6}A modifications, is knocked out (Bokar et al. 1994; Liu et al. 2014; Supplemental Fig. 1A; Supplemental Table 1; Methods). As expected, in wild-type mESCs ~6% of the adenines within the context of a DRACH motif had a high probability (>0.85) of being modified, compared to only 0.2% in those with *Mettl3* knockout (Supplemental Fig. 1B). Among the sites displaying a high-probability modification ratio, the levels of modification ratios were consistently higher in wild-type compared to *Mettl3* knockout cells (median modification ratio 0.629 and 0.512, respectively). Moreover, the modified adenines were predominantly clustered near the 3′ end of coding sequences, which is consistent with the expected pattern of m^{6}A RNA modifications (Fig. 1A,B; Supplemental Fig. 1C,D; Meyer et al. 2012).

Using SNPs, we assigned 1,110,260 (replicate 1) and 837,011 (replicate 2) long-reads to their allelic origins across more than 13,000 transcripts (Fig. 1C; Methods). Of the detected transcripts, 8657 transcripts had at least 10 reads in both replicates. In allele-specific analyses, a common challenge is reference allele bias which is the tendency for reads that match the reference genome to align with a higher probability than reads containing the alternate allele, potentially skewing variant detection and analysis (Castel et al. 2015). To minimize this bias, we employed an N-masked transcriptome reference. This approach led to a mean CAST allele ratio across all transcripts of 0.505 as opposed to 0.485 when using an unmasked reference (Methods). These assignments were based on 135,380 (replicate 1) and 134,585 (replicate 2) informative positions on the long-reads that overlapped known genetic variation between the strains (210,004 total SNPs).

As an orthogonal approach to determine ASE, we used Illumina short-read sequencing. We found that RNA expression levels from the two methods were significantly correlated (Supplemental Fig. 2A) (Spearman's correlations [rho], 0.792–0.816 across replicates). Moreover, gene-level allele-specific RNA expression was moderately concordant between the two approaches (Supplemental Fig. 2B) (weighted rho = 0.61, Methods). Although short-read sequencing produced nearly eight times more aligned reads, long-read sequencing identified 2.3 times as many SNPs, demonstrating that the greater number of informative positions in long-read data enhances allelic assignment accuracy and gene-specific ASE reproducibility (rho = 0.63 and 0.51 for long-read and

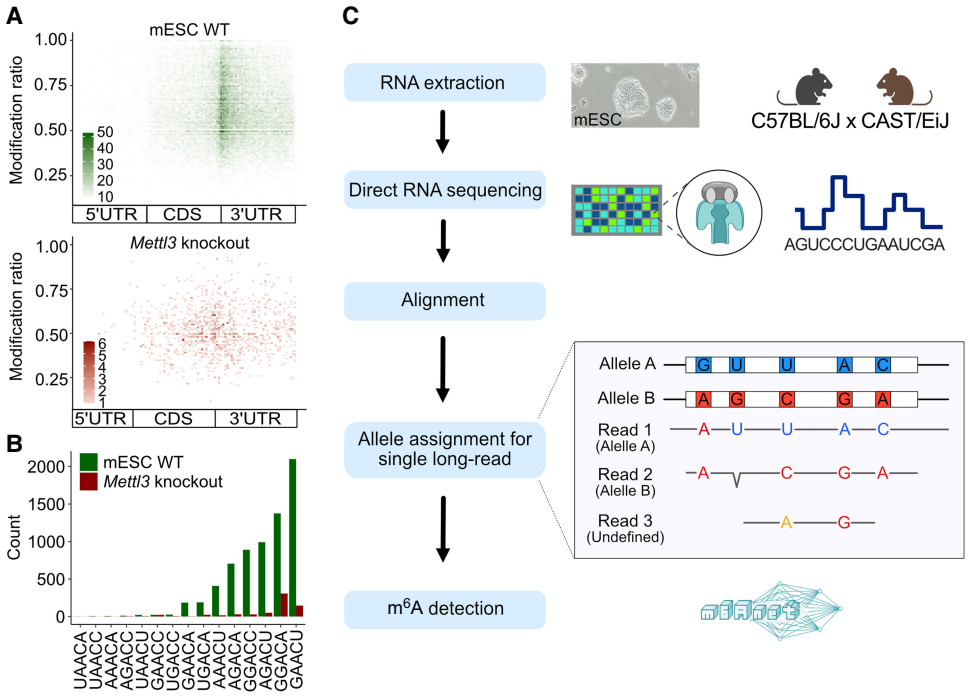

Figure 1. Allelic read assignment and m^{6}A modification analysis using ONT DRS in hybrid mESCs. (*A*) m^{6}A modification ratio and locations detected from m^{6}Anet using all reads (*top*, green, WT; *bottom*, red, *Mettl3* knockout). The relative m^{6}A locations within the transcript body were determined. It presents modification ratios after high-probability selection (>0.85). The color darkness represents the counts of the ratio on the position. (*B*) Comparison of the frequencies of instances of DRACH motif sequences (green, WT; red, *Mettl3* knockout). (*C*) Schematic overview of the strategy used for allelic long-read assignment for ASM analysis. Total RNA from hybrid mESC (C57BL/6J × CAST/EiJ) underwent DRS. To avoid reference bias, we used an N-masked transcriptome for alignment. Reads were then allocated to each allele. This process classified reads into Allele A (B6), Allele B (CAST), and undefined categories, enabling m^{6}A detection within each group individually.

short-read sequencing, respectively) (Supplemental Fig. 2C,D). Taken together, these measures of quality control underscore the high precision in allelic assignment.

We then employed a supervised machine learning approach (Liu et al. 2019; Hendra et al. 2022) to quantify m^{6}A RNA modifications for the reads attributed to each allele. This process revealed an equivalence in the number of reads and m^{6}A sites among alleles, indicating the allelic impartiality of our approach. Specifically, we observed similar numbers of reads (564,944 and 444,514 for B6; 557,787 and 443,131 for CAST) and potential modification sites (114,457 and 105,190 for B6; 112,947 and 105,117 for CAST) for each allele (Fig. 2A; Supplemental Table 2). This result indicates minimal to no allelic bias in the assignment of reads and identification of modification sites.

The modification ratios of the candidate m^{6}A sites were highly correlated between replicates and demonstrated an even higher correlation within the same allele. Specifically, rho within the same allele were 0.82 and 0.83 for the modification ratios of the B6 and CAST alleles, respectively. Conversely, correlations between different alleles were slightly lower, with the values of 0.77 (B6 replicate 1, CAST replicate 2) and 0.75 (CAST replicate 1 and B6 replicate 2). In contrast, modification ratios from *Mettl3* knockout cells exhibited significantly lower correlations, falling below 0.46 (Fig. 2B; Supplemental Fig. 3).

To identify ASMs, we established a selection criterion centered on a site probability aggregated from all reads. Therefore, we focused on m^{6}A sites that demonstrated a high probability of modification (>0.85) across reads. In mESC wild-type, an average of ~7% of these candidate sites met our selection criteria. Notably, these m^{6}A modification sites were predominantly localized at the junction between the coding region and the 3′ UTR, showing high modification ratios (with median ratios of 0.621 for B6 and 0.627 for CAST) (Fig. 2A,C; Supplemental Table 3). In contrast, *Mettl3* knockout had only 0.8% allelic sites exhibiting high probabilities of m^{6}A modification. Moreover, these sites demonstrated a wider dispersion across various transcript regions (Fig. 2A,D). Overall, these observations affirm the ability of our methodology in detecting allelic m^{6}A modifications subsequent to the assignment of reads to alleles.

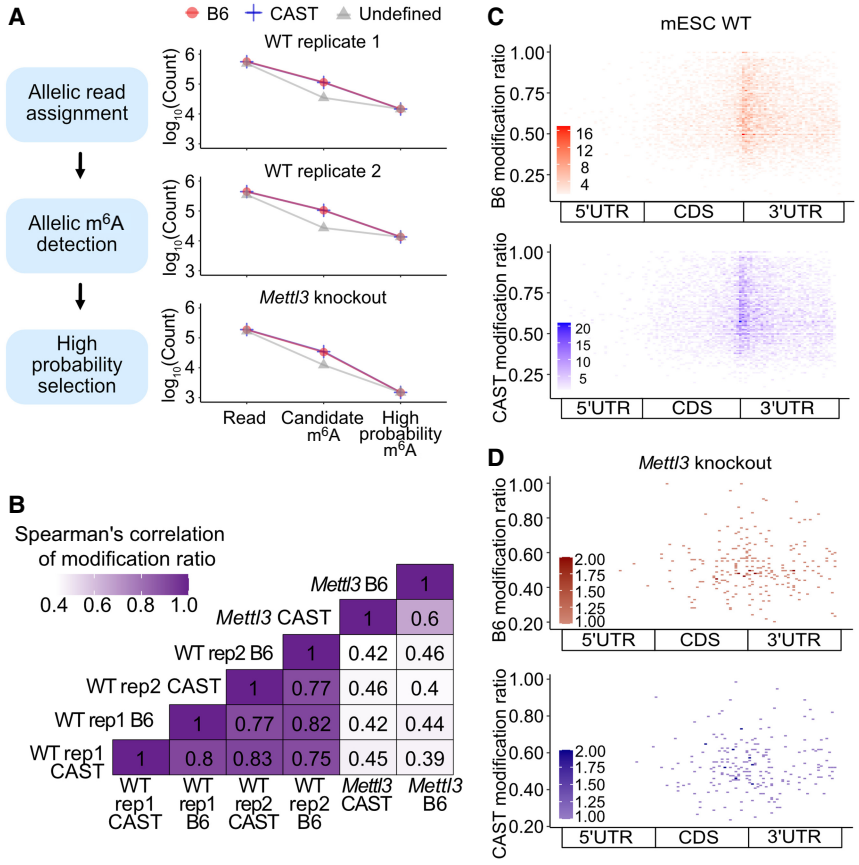

Figure 2. Comparative analysis of allelic modifications in wild-type (WT) and *Mettl3* knockout mESCs. (*A*) Allelic impartiality while allelic read assignment and m^{6}A detection. The *left* panels outline the schematic of the data procedural steps for ASMs analysis. The *right* panels display the counts for allelic reads, candidate m^{6}A sites, and high-probability m^{6}A sites selected through our criteria (red circles, B6; blue plusses, CAST; and gray triangles, undefined group). The *left top* two plots show the counts from mESC WT replicates and the *bottom* plot exhibits the numbers from mESC *Mettl3* knockout. (*B*) Spearman's correlation of modification ratio between alleles from WT and *Mettl3* knockout cells (rep1: mESC WT replicate 1; rep2: mESC WT replicate 2; *Mettl3*: *Mettl3* knockout). (*C*,*D*) The distribution of sites with high probability of m^{6}A modification (prob > 0.85) is displayed in a metagene plot by calculating the relative positions of these sites within gene regions. The color scale represents the number of m^{6}A sites with the given modification ratio inferred from reads assigned to either of the two alleles in mESC WT (*C*) or *Mettl3* knockout (*D*) cells (red, B6 allele; blue, CAST allele).

Detection of sites with significant allele-specific m^{6}A modification

The capability to accurately assign each RNA molecule to its allelic origin while concurrently identifying RNA modifications allows for the investigation of positions within mRNAs that exhibit differential modifications between alleles. While numerous statistical methods have been developed to identify allele-specific differences in gene expression phenotypes (DeVeale et al. 2012), these studies underscore the challenges inherent in this analysis, including a propensity for false positives when employing simple binomial tests to assess deviations from expected expression levels across the two alleles (Mohammadi et al. 2017; Zitovsky and Love 2019).

To address these challenges, we implemented a conservative strategy that leverages bootstrap sampling to quantify uncertainty in modification ratio estimates (Methods). This method enabled us to pinpoint mRNA positions showing significant ASM (Fig. 3A). Among detected 14,609 and 13,542 candidate m^{6}A sites in the replicate experiments, we identified 57 ASM sites (false discovery rate [FDR] < 0.1) with an average modification difference between the two alleles of 0.32.

In allele-specific analysis, previous research revealed that events with larger effect sizes are more likely to be reproducible and biologically relevant (Mohammadi et al. 2017; Castel et al. 2020). Therefore, we repeated our statistical analyses using an effect size

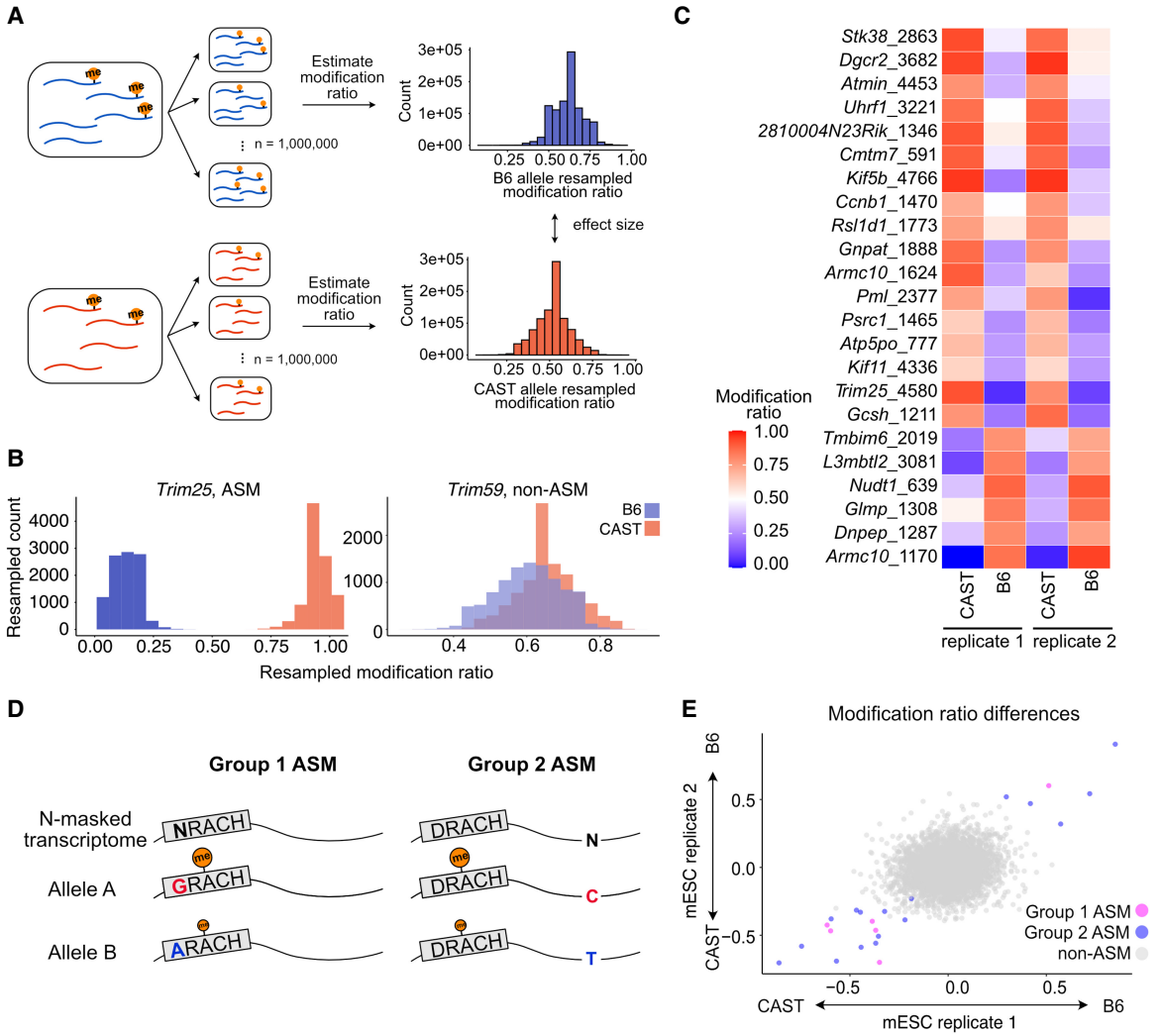

Figure 3. Identification and classification of ASM sites. (*A*) Schematic of statistical procedure for ASM detection (Methods). Reads overlapping the site under consideration were resampled, and the modification ratio was estimated in each bootstrap sample. A statistically significant ASM site was defined as an adjusted harmonic *P*-value (FDR < 0.1; Methods). (*B*) An ASM site within *Trim25*, exhibits distinct modification ratio samples. Conversely, a non-ASM site within *Trim59* displays substantial overlap in modification ratios between bootstrap sampling distributions. (*C*) Modification ratios of each allele across mESC wild-type replicates. The *y*-axis displays the name of the gene and m^{6}A position in the transcript. (*D*,*E*) ASM sites were classified into two groups. Group 1 is defined by genetic variants within the DRACH motif, and Group 2 is characterized by variants adjacent to or distal from the DRACH motif (*D*). The modification differences of the defined ASM were represented by color according to their classification (Group 1 in magenta, Group 2 in blue, and non-ASM in gray). Each axis is the modification ratio, where negative values denote CAST allele bias and positive values indicate B6 allele bias in m^{6}A modification (*E*).

threshold of 0.1 corresponding to the inferred modification ratio difference between the two alleles (Methods). This analysis uncovered 23 sites across 22 genes indicating ASM. Notably, at these ASM sites, the distributions of the resampled modification ratios from the two alleles were consistently distinct and had large effect sizes, with a mean modification ratio difference of 0.48 (Fig. 3B; Supplemental Fig. 5A; Supplemental Table 4). We focused our detailed analyses on this subset of ASM sites.

One inherent limitation of ONT sequencing lies in its limited sequencing depth. Consistently, transcripts with statistically significant ASM sites had significantly greater RNA expression levels than those without, underscoring the dependence of ASM detection on transcript abundance (Supplemental Fig. 4).

The genes with ASM sites are distributed across chromosomes without any discernible location preferences (Supplemental Fig. 5B), and are associated with a wide range of functions (Fig. 3C). A particularly notable finding was the identification of two distinct sites of ASM on the *Armc10* transcript, which encodes a protein involved in mitochondrial dynamics (Serrat et al. 2014; Chen et al. 2019). Moreover, our analysis identified six B6-biased ASM sites with a higher modification ratio on the B6 allele and 17 CAST-biased ASM sites with a higher ratio on the CAST allele. While the majority of ASM sites were located on 3′ UTRs, one B6-biased ASM (on *Dnpep*) and two CAST-biased ASMs (on *Gnpat*, and *Pml*) were found in coding regions, near the stop codons (Supplemental Fig. 5C).

Genome sequencing was used to identify genetic differences between the mouse inbred lines (Tsang et al. 2005; Adams et al. 2015); however, potential genotyping errors from these could result in erroneous ASM calls. Hence, we verified the genomic

DNA sequences near the m^{6}A modification sites using Sanger sequencing (Supplemental Table 5; Methods). In six selected ASM sites, we confirmed annotated SNPs (*Nudt1*, D site), and the absence of unannotated genetic variants. These results indicate that the detected modifications are genuinely posttranscriptional and do not reflect genotyping errors. Taken together, these findings highlight a key strength of our approach based on the ONT DRS technique, which enables the detection of m^{6}A modifications at the individual molecule level, rather than relying on aggregate measurements.

Genetic variants influence allele-specific m^{6}A modification patterns

We hypothesized that local genetic differences could influence methylation efficacy, leading to differential m^{6}A deposition. Accordingly, we categorized ASM-biased sites into two groups based on the proximity of the nearest genetic variation to the canonical m^{6}A methylation motif (DRACH) (Fig. 3D,E). Of the 23 sites identified, six had genetic variants located within the DRACH motif itself (Group 1), with three variants at the D position, one at R, and two at H (Fig. 4A). In total, 41 m^{6}A sites had SNPs in D, R, or H positions, with 6 of these (14%) classified as ASM. These results suggest that SNPs within the DRACH motif are, as expected, more likely to lead to ASM. Furthermore, specific instances of the DRACH motifs are more likely to lead to modified adenines (Fig. 4B). In agreement with expectation, alleles for Group 1 ASM sites that exhibited higher modification ratios were more likely to match instances of the DRACH motif with higher propensity for modification (sole exception site on *Pml*).

Analysis of the remaining 17 sites (Group 2) revealed that five possessed genetic differences adjacent to the DRACH motif. Specifically, we found six SNPs near five m^{6}A sites: *L3mbtl2* (D – 1 and H + 1); *Trim25* and *Dnpep* (D – 1); *Atmin* and *Tmbim6* (H + 2) (Fig. 4C). Notably, all instances of SNPs at the D – 1 position included a U at the D – 2 position (UNUGACU). In this context, a cytosine at the D – 1 site correlated with higher m^{6}A levels (*Dnpep*, 0.768; *Trim25*, 0.855; *L3mbtl2*, 0.789) compared to an adenine or a guanine at the D – 1 site (adenine on *Dnpep*, 0.324 and *Trim25*, 0.073; guanine on *L3mbtl2*, 0.157) (Fig. 4D). This finding highlights the significant influence of nucleotides adjacent to the DRACH motif on m^{6}A deposition, contingent upon their specific genetic context.

Among remaining Group 2 ASM sites, eight had a SNP within 100 bp of the modified adenine (*Stk38*, −64; *2810004N23Rik*, −51; *Cmtm7*, −12 bp; *Glmp*, +10 bp; *Rsl1d1*, +34 bp; *Gcsh*, +42 bp; *Kif11*, +59; *Armc10* at 1170 position, +99). Despite the limitations imposed by the read length, short-read-based m^{6}A detection methods are theoretically capable of detecting SNPs within 50–100 bp of the methylated site (Dominissini et al. 2012; Chen et al. 2015). However, our method also identified four ASM sites that had no SNPs within this range hence highlighting the unique strength of long-read sequencing for ASM detection.

We also noticed that Group 2 ASMs were highly enriched for the UGACU motif sequence over the most commonly observed GAACU instance of the DRACH among m^{6}A-modified sites (Fig. 4E, *P*-value 0.0084; Methods). This finding suggests that ASM may be more prevalent for specific motif sequences distinct from those typically seen in m^{6}A-modified sites. In summary, we uncovered differential m^{6}A modification of alleles that may depend on genetic differences that are proximal to the DRACH motif as well as ASM sites which have no nearby genetic differences.

Validation of ASM sites through orthogonal approaches

We next assessed the robustness of ASM detection by visualizing read pileups and conducting an orthogonal experimental method. Computational approaches to detect m^{6}A modifications from DRS have been developed to leverage the increased propensity of base-calling errors around modified bases (Liu et al. 2019). We visualized sequencing reads that overlap ASM sites, enabling us to verify the expected enrichment of base-calling errors around sites with a higher modification ratio. The phenomenon was observed consistently across both replicates, characterized by a correspondence between base-calling errors and modification ratios (Supplemental Fig. 6).

An orthogonal experimental approach that can potentially detect transcript regions with ASM is MeRIP-seq (Cao et al. 2023). MeRIP-seq relies on antibodies to differentiate modified loci and uses short-read sequencing; thus, this strategy lacks single-molecule and single-nucleotide resolution. Nonetheless, we reasoned that some ASM sites would overlap MeRIP-seq peaks and provide additional experimental support for allelic bias.

Among the 23 ASM sites, we detected 19 in the MeRIP-seq. Only four out of 19 sites, which contain SNPs within or nearby the DRACH motif had sufficient coverage in our MeRIP-seq data (Supplemental Table 6; Methods). The allele bias ratio measured from MeRIP-seq in these four sites demonstrated consistency with the allele bias detected by our approach (Supplemental Fig. 7). For example, *Atp5po* (Group 1 ASM) displayed allelic bias consistent with expectation in all three MeRIP-seq replicates (Fig. 4F). Another Group 2 ASM site, *Gcsh* (SNP at position 1253, 42 bp downstream from methylation site), exhibited the same allele-bias pattern in both long-read sequencing and MeRIP-seq data (Fig. 4G). In short, while MeRIP-seq cannot capture all ASM sites detected by the long-read approach due to inherent limitations, we observed consistent allele bias in m^{6}A patterns at four sites with sufficient read coverage.

The reliance of MeRIP-seq on short-read sequencing can lead to errors in allelic assignment, primarily due to dependence on a limited number of SNPs, which increases susceptibility to reference allele bias, genotyping errors, and systematic biases in library preparation. To assess potential genotyping errors, we examined 37 SNP sites within ASM genes using Sanger sequencing. Of these, 33 sites showed the expected genetic variants with strong peak signals, however, four sites (*Atp5po*, 776; *Psrc1*, 977; *Trim25*, 5007 and 5041) displayed nucleotides from only one allele suggesting potential genotyping errors or limitations in our Sanger sequencing experiments. Importantly, these findings further underscore the challenges of accurate allelic detection especially for the short-read sequencing approach that relies on one or few SNPs (Supplemental Table 7).

Applicability of ONT DRS to detect ASM sites in human cells

The analytical and empirical workflow we developed to detect ASM sites is broadly applicable to any cell type with known genetic information. Given that systematic replication is essential to validate new approaches (Piccolo and Frampton 2016), we next replicated ASM detection using a lymphoblastoid cell line derived from a human with a well-characterized genome. Specifically, we analyzed five replicates of ONT DRS data generated using the NA12878 cell line (Hansen 2016; Workman et al. 2018), assigned reads to their allelic origin, and quantified m^{6}A modifications for each group of reads (Fig. 5A; Supplemental Table 8).

Unlike hybrid mESCs, a typical human harbors many fewer heterozygous SNPs per transcript (Rozowsky et al. 2011; Workman et al. 2018). In our long-read sequencing analysis, among 21,569 mouse transcripts, 16,242 contain at least one heterozygous SNP in hybrid mESCs, while only 8889 human transcripts contain such sites in NA12878. The transcriptome overall

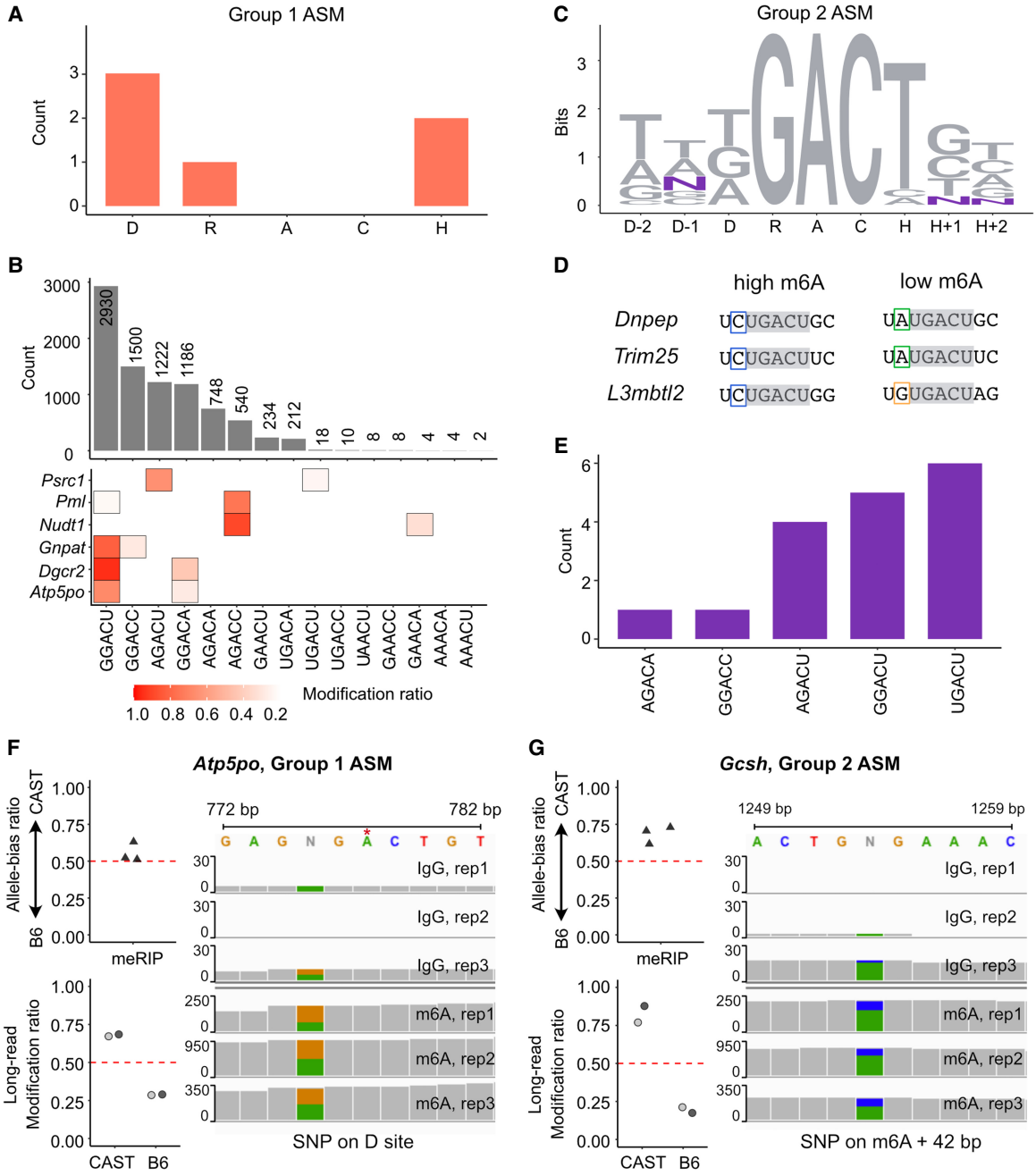

Figure 4. Characterization of ASM sites and orthogonal detection with MeRIP-seq. (*A*) SNP distribution in Group 1 ASM. (*B*) Motif frequencies and modification ratios of motif sequences. The *top* bar plot illustrates motif sequence frequencies in all m^{6}A instances, while the *bottom* heatmap indicates modification ratios. The first row presents the modification ratio of all instances and the following six rows represent the modification ratio on each motif sequence differentiated by SNPs from two alleles of Group 1 ASM sites. (*C*) Information content of the extended DRACH motif in 17 Group 2 ASM sites shown in DNA sequence. The D – 1 site has three SNPs, while the H + 1 and H – 2 sites each have one SNP. (*D*) Extended motif sequences where the D – 1 site possesses SNPs. The gray box represents the DRACH motif, in which all three genes share the same sequence (UGACU) followed by U on the D – 2 site. (*E*) Motif prevalence in Group 2 ASM. The UGACU motifs are predominantly observed, contrasting with the common m^{6}A motifs, which are typically represented by GGACU. (*F*,*G*) Orthogonal detection of ASM through MeRIP-seq and long-read sequencing. On the *top* panel, points illustrate the allele-bias m^{6}A ratio (proportion of reads from CAST allele) derived from three replicates of MeRIP-seq analysis. The Integrative Genomics Viewer browser displays MeRIP-seq reads on SNPs adjacent to m^{6}A sites, which correspond to the MeRIP-seq allele-bias ratio. The points in the *bottom* panel indicate the modification ratio of each allele from long-read sequencing, with gray color pairs representing data from two replicates. Two examples from *Atp5po*, Group 1 ASM (*F*), and *Gcsh*, Group 2 ASM (*G*) are shown.

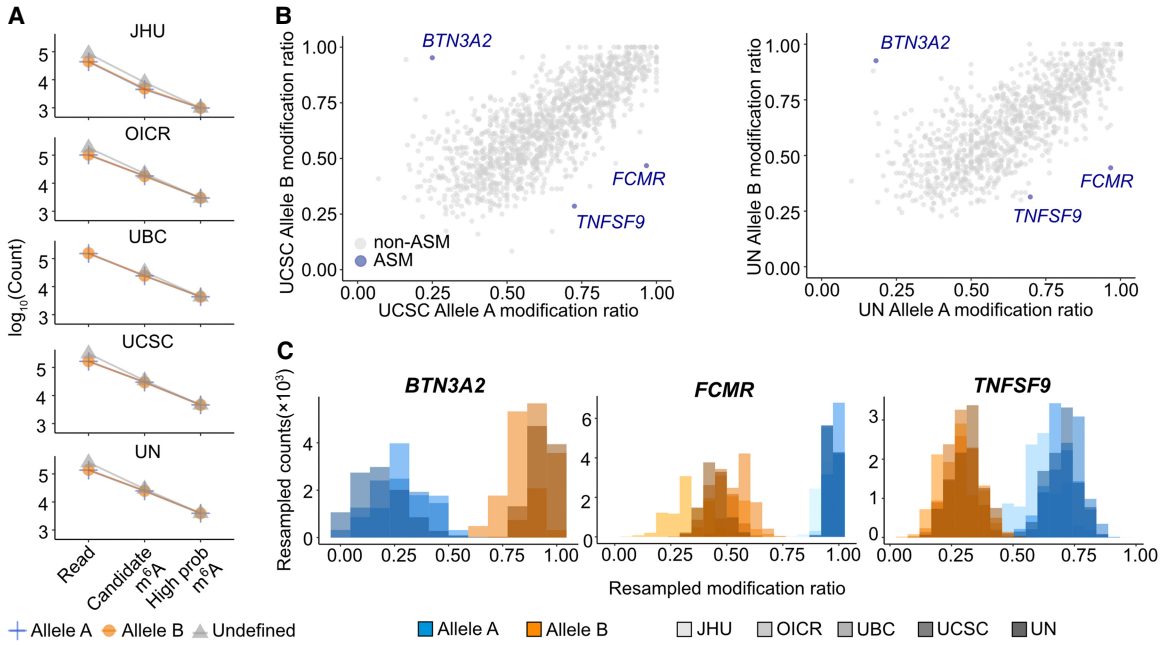

Figure 5. Reproducibility of ASM profiling procedure in human cells. (*A*) Number of detected candidate m^{6}A modification sites among five replicates (blue, Allele A; orange, Allele B; gray, undefined read classification). (*B*) Modification ratios from each allele, including three ASM sites (blue) and non-ASM sites (gray) in UCSC (*left*) and UN (*right*) data sets, the highest depth data sets among five replicates. The *x*-axis represents the modification ratio of Allele A reads, while the *y*-axis represents the modification ratio of Allele B reads. (*C*) Resampled modification ratios from bootstrapping. Each color represents an allele (blue for Allele A, orange for Allele B), and the gray gradient indicates each replicate.

harbors nearly 10 times fewer heterozygous SNPs in NA12878 compared to hybrid mESCs (210,004 in mESCs; 27,269 in NA12878) (Supplemental Fig. 8). Hence, the percentage of long-reads that can be assigned to their allelic origin with high confidence is reduced (Supplemental Tables 2, 8).

Despite having lower depth of sequencing and fewer informative SNPs per transcript, we identified three ASM sites with reproducible and large effect sizes. These three sites were found on the *BTN3A2* (FDR = 0.006), *FCMR* (FDR = 0.006), and *TNFSF9* (FDR = 0.47) transcripts (Fig. 5B; Supplemental Figs. 9, 10). All three events were observed within the 3′ UTR and exhibited large differences in modification ratio between the two alleles (Fig. 5C) (mean difference in modification ratio 0.657, *BTN3A2*; 0.500, *FCMR*; and 0.390, *TNFSF9*). *BTN3A2* plays a crucial role in T cell activation and proliferation (Kabelitz and Déchanet-Merville 2015; Vantourout et al. 2018), *FCMR*, which encodes the IgM Fc receptor, is vital for B cell activation and survival (Wang et al. 2016), and *TNFSF9*, a member of the TNF superfamily, enhances T cell responses by interacting with CD137 on activated T lymphocytes (Wang et al. 2016; Hashimoto 2021). The applicability of our ONT DRS method for ASM detection in human cells supports the wide-ranging utility of our approach in any system with known genetic information.

RNA abundance is higher for the allele with higher m^{6}A modification ratio

Allele-specific differences in m^{6}A modification provide a powerful platform to assess their functional impact on expression dynamics as the genetic background, environmental factors, and sample preparation are identical for the two alleles. Hence, we generated matched RNA-seq and ribosome profiling data in hybrid mESCs and leveraged existing measurements for the NA12878 cells (Methods; Cenik et al. 2015). This data enabled us to determine the relative RNA expression and ribosome occupancy on each allele and correlate these with their m^{6}A modification status.

In hybrid mESCs, transcripts harboring ASM sites demonstrated statistically significant RNA expression bias toward the allele with higher m^{6}A modification. This pattern was consistent across both long-read and short-read sequencing methods (Fig. 6A; Binomial test *P*-value 0.004 and 0.011, respectively). Specifically, the mean proportion of RNA reads from the CAST allele for genes exhibiting CAST-biased ASM were 0.557 and 0.558 for long-read and short-read sequencing. Similarly, genes with B6-biased ASM had higher mean proportion of RNA reads from the B6 allele (0.460 and 0.398, respectively). These observations suggest that ASM is associated with ASE in the same allelic direction.

In NA12878 cells, the association of ASM and allele-specific RNA expression was similarly evident. *BTN3A2*, possessing Allele B biased methylation site, demonstrated a high proportion of RNA reads from Allele B (mean allelic ratios 0.851 and 0.704 for long-read and short-read sequencing). Similarly, *TNFSF9* and *FCMR*, with Allele A biased methylation sites, showed slightly elevated proportions of RNA reads from the Allele A (Supplemental Fig. 11A,B). These findings further support an association with ASM and allele-specific RNA expression (Fig. 6B).

Recent studies have postulated the role of m^{6}A modification in regulating translation (Mao et al. 2019; Meyer 2019b; Jain et al. 2023). In particular, we analyzed allele-specific ribosome occupancy on genes with ASM in hybrid mESCs. We did not detect a significant correlation between allele-specific ribosome occupancy and ASM (Fig. 6A; *P*-value 0.83; Supplemental Fig. 11C). Collectively, our results indicate that alleles with higher m^{6}A modification ratios are associated with increased RNA abundance but similar ribosome occupancy.

Discussion

In this study, we used ONT DRS as a new method to detect allele-specific m^{6}A RNA modifications in both human and mouse cells. Notably, the long-reads generated by the ONT approach revealed ASM sites with no nearby genetic differences, suggesting that m^{6}A modification on a given site may depend on factors beyond the local sequence context. These ASM may potentially be governed from long-range interactions that modulate mRNA secondary structure, differences in allele-specific interactions with RNA-binding proteins, or the local chromatin context of each allele (Berlivet et al. 2019; Huang et al. 2019; Deng et al. 2022).

The deposition of m^{6}A modification on mRNA is dependent on the presence of a specific motif (DRACH) surrounding the adenosine that is modified (Linder et al. 2015). Our analysis of ASM sites revealed that the nucleotide identity of the positions that surround this canonical motif may also influence m^{6}A deposition in particular contexts. Specifically, we found that alleles containing cytosine at the D – 1 site followed by uracil at the D – 2 site exhibit higher methylation levels (Supplemental Fig. 6).

A major strength of our approach over short-read-based methods is its ability to cover many more informative SNPs to assign reads to their allelic origin (Supplemental Fig. 12) (maximum SNP count per read, 12 with short-read; 78 with long-read in mESC). Hence, long-read technology has higher confidence in allelic assignment (Glinos et al. 2022; Wu et al. 2023). In contrast to short-read methods, which rely heavily on single SNPs within a read, our approach minimizes the impact of errors in genetic variant annotations by leveraging an increased number of variants. Furthermore, in samples with less genetic variation, long-reads increase the chance of linking genetic variants that may be far away from the site of interest which would not be detectable by short-read-based approaches.

A recent study leveraged previously generated MeRIP-seq data and claimed to detect numerous ASM sites (Cao et al. 2023). Their approach involved calculating *P*-values from Fisher's exact test on tables of reads per kilobase of transcript per million mapped reads for each allele from the input control and immunoprecipitation. They interpreted the resulting *P*-values as evidence of ASM. However, this method is fundamentally flawed. Fisher's exact test is specifically designed for categorical data and applying it to continuous data in this context is inappropriate. This misuse of the test raises serious concerns about the validity of their conclusions.

Furthermore, MeRIP-seq suffers from the additional limitations of antibody-based enrichment. Antibody-based approaches introduce specificity artifacts which result in variability in the number and location of peaks detected across experiments (Helm et al. 2019). Similarly, the immunoprecipitation step creates variable yields, limiting quantitative measurements among experiments (McIntyre et al. 2020). Therefore, the large number of sites reported by Cao et al. are inflated with a large number of false positives. In our study, we focused on large effect differences using a bootstrap resampling strategy and minimum effect size threshold to reduce statistical artifacts as previously recommended (Mohammadi et al. 2017; Castel et al. 2020). Consequently, the number of sites described here likely reflects the extent of allele-biased methylation more accurately.

To address the limitations of antibody-based detection of m^{6}A modifications, recent work developed enzymatic and chemical approaches (Meyer 2019a; Song et al. 2021). However, the applicability of the enzymatic approach is currently restricted to a subset of m^{6}A sites within DRACH motifs ending in ACA, constituting ~16% of total sites. While these advances are promising, they will likely be limited for allele-specific analysis due to the use of short-read sequencing (Garcia-Campos et al. 2019).

Allelic imbalances in m^{6}A modification ratios between transcripts can potentially lead to allele-specific RNA expression and translation based on their impact on mRNA stability, transcription, and translation efficiency (Mauer et al. 2017; Min et al. 2018; Cesaro et al. 2023). Specifically, m^{6}A reader proteins such as YTHDC1 and YTHDC2, which interact with m^{6}A sites on 3′ UTRs, enhance mRNA stability (Wang et al. 2014; Lee et al. 2021). Our study revealed a positive relationship between ASM and allele-specific RNA expression. A potential mechanism explaining this association is the allele-specific association with m^{6}A reader proteins that subsequently stabilize m^{6}A-enriched mRNAs.

In contrast, we did not observe an association between ASM and allele-specific ribosome occupancy. Given that ribosome occupancy is a composite

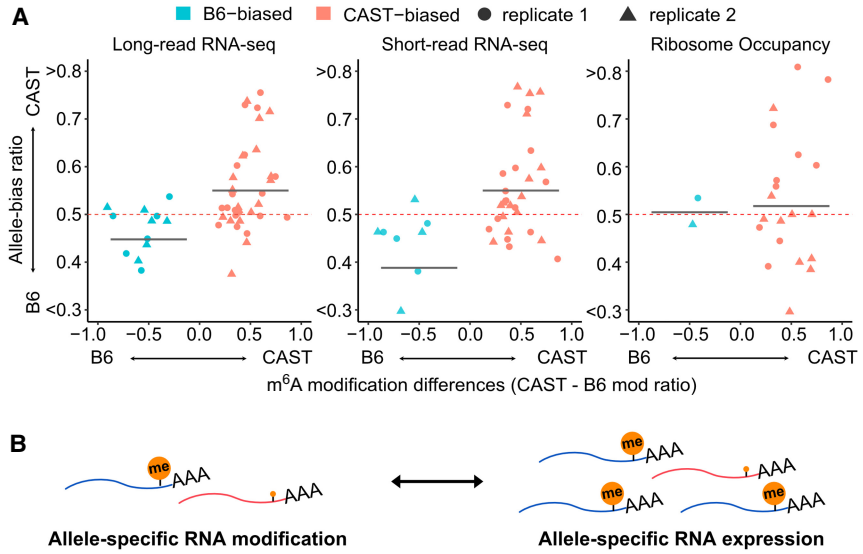

Figure 6. Effects of ASM on allele-specific RNA expression and ribosome occupancy. (*A*) Allele bias ratio of genes containing ASM sites (sky blue for B6 biased; pink for CAST biased). Different shapes represent the replicates (circle for replicate 1, triangle for replicate 2) The *y*-axis displays the allele bias ratio obtained from long-read (*left*), short-read (*middle*) sequencing, and ribosome profiling (*right*). The *x*-axis shows the difference in m^{6}A modification ratios between the two alleles (CAST–B6). The red dashed horizontal bar indicates an allele-bias ratio 0.5, an allele-bias cutoff point. The gray bar represents the mean allele bias ratio for genes with B6 or CAST-biased ASM sites. (*B*) Model for regulation of ASM and ASE. The two-sided arrow in the model reflects the possibility of bidirectional influence, rather than suggesting unidirectional causality. Differential transcription driven by allele-specific epigenetic contexts may regulate ASM, indicating that both phenomena could share a common underlying mechanism, without one necessarily being the direct cause of the other.

measurement of RNA expression and translation efficiency (Liu et al. 2024; Zheng et al. 2024), this observation may indicate that alleles with higher modification ratios are less efficiently translated despite having higher steady-state RNA abundance. Such a mechanism would be in agreement with a previously proposed model of coupling between cotranscriptional m^{6}A deposition and translation (Slobodin et al. 2017).

Our method has several important limitations. First, the supervised machine learning framework we adopted is predicated on the assumption that modifications occur exclusively within DRACH motifs. Consequently, our analysis does not account for genetic variations that alter the motif into sequences not matching the DRACH pattern, which are presumed to result in methylation loss. Second, the limitation in the number of reads generated by ONT DRS constrains our method's ability to detect ASM sites in lowly expressed transcripts. Hence, ASM sites identified in this study occur in genes within the top 30th percentile of RNA expression (Supplemental Fig. 4).

In summary, we present a novel method for identifying ASM using ONT DRS. Our analyses emphasize the benefits of long-read sequencing and direct detection of RNA modifications for ASM analysis. Future ASM studies are likely to extend the catalog of allelic variants that influence RNA modifications, and characterize the mechanisms leading to ASM and its functional consequences on gene expression.

Methods

Cell culture

The C57BL/6J-CAST/EiJ F1 Hybrid mESCs were generously provided by Dr. David Spector. Cells were cultured in 2i medium as followed by the original cell culture method (Balasooriya and Spector 2022).

Generation of *Mettl3* knockout mESCs

Mettl3 knockout cells were generated by introducing Cas9/sgRNA ribonucleoprotein (RNP) complexes into mESCs via nucleofection (Kirton et al. 2013). The sgRNA was synthesized by Synthego (Supplemental Table 2). To form the RNP, 300 pmol of Cas9 protein (NEB M0386M) and 600 pmol of sgRNA were incubated in Cas9 Buffer (150 mM KCl, 1 mM $MgCl_2$, 10% v/v Glycerol, 20 mM HEPES–KOH [pH 7.5]) at room temperature for 30 min. Subsequently, 65 µL of 4D-Nucleofector X Solution was added to the RNP solution. Nucleofection was performed using the optimized protocol recommended by the manufacturer (SF Cell Line 4D-NucleofectorTM X Kit L). A cell pellet was collected from 2×10^6 cells, resuspended in the RNP solution, and transferred into a 100 µL Nucleocuvette Vessel. Electroporation was carried out using the 4D-Nucleofector X Unit (Lonza) with the FF120 program. Postnucleofection, cells were equilibrated at room temperature for 8 min, then transferred to a gelatin-coated culture dish containing prewarmed 2i media. The cells were allowed to recover at 37°C for 72 h, followed by the isolation of single clones using serial dilution. The genomic DNA was isolated from cells grown from single clones and mutations were confirmed using the primers listed in Supplemental Table 2.

ONT DRS

mESCs were collected from two different numbers of passages on separate days, and considered as two biological replicates. The cells were lysed in TRIzol reagent (Zymo Research R2050) and RNA was extracted according to the manufacturer's instructions (Zymo Direct-zol RNA Kits R2061). Five micrograms of total RNA without poly(A) RNA isolation was used for DRS (Viscardi and Arribere 2022). The library was generated using the Oxford Nanopore Direct RNA Sequencing Kit (Nanopore SQK-RNA002) following the manufacturer's protocol. The RNA sequencing from each RNA replicate was performed on four MinION MkIb with R9.4 flow cells (Oxford Nanopore Technologies) with a 24 h runtime for each run.

We used published ONT DRS data from the human cell line NA12878 downloaded from GitHub (https://github.com/nanopore-wgs-consortium/NA12878) (Hansen 2016; Workman et al. 2018). Raw FAST5 files were used directly for analysis.

m^{6}A detection from ONT DRS

Guppy v. 6.3.2 (Wick et al. 2019) (quality score cutoff = 7) was used for base-calling from FAST5 files, and error rates were analyzed with Pomoxis v0.3.15 (Buttler and Drown 2023). Reads were aligned to the transcriptome with minimap 2.1 "-ax map-ont" option (Li 2018). To reduce alignment biases, we used a transcriptome reference in which SNPs were masked with Ns as previously described (Ozadam et al. 2023). A mouse VCF was downloaded from the Mouse Genome Project (https://www.mousegenomes.org/), and the NA12878 VCF file was obtained from https://hgdownload.soe.ucsc.edu/gbdb/hg38/platinumGenomes/.

To identify m^{6}A modifications, we used Nanopolish v0.11.3 (Loman et al. 2015) to generate an index with the "--scale-events" and "--signal-index" options, aligning events to the N-masked transcriptome reference. Detection of m^{6}A RNA modifications was conducted using m6Anet v-2.0.0 and a pretrained model (Hct116_RNA002) (Hendra et al. 2022). A minimum of 20 reads per site was required to call modification sites.

Assignment of reads to their allelic origin

To assign aligned reads to their allelic origin, we identified the positions on each read that correspond to an SNP, adjusting for any deletions and insertions in the read with respect to the reference. The number of heterozygous SNPs in mESC and NA12878 transcriptomes was 210,004, and 27,269, respectively (Supplemental Fig. 8). Consequently, we required at least three SNPs per read for mESCs, and one SNP for NA12878. Then, we calculated the number of matches to each allele and defined a read-level allele-bias ratio:

$$\text{Allele A bias ratio} = \frac{\text{SNP count from Allele A in a read}}{\text{total number of detected SNPs in a read}}$$

We assigned each read into one of three groups based on this ratio: Allele A (bias ratio exceeding 0.7), Allele B (bias ratio <0.3), and undefined (bias ratio between 0.3 and 0.7). The process was implemented in a Python script that is provided on GitHub: allele_assignment.py. Subsequently, the three groups of reads underwent processing through m6Anet separately to predict m^{6}A probability and modification ratios as described above.

Identification of allele-specific m^{6}A modifications

We first selected m^{6}A sites with a high probability of modification (prob > 0.85) using all reads. When the SNPs on the motif convert DRACH motifs to non-DRACH motifs, we exclude them from the analysis because non-DRACH motifs are by definition assumed to be unmethylated (Hendra et al. 2022). In mESCs, 29 SNP overlapping the motif led to different instances of the DRACH sequences among a total of 178 and 145 sites, respectively (Supplemental Fig.

13). In NA12878, the corresponding numbers were 1 in 5 (JHU), 1 in 11 (OICR), 1 in 21 (UBC), 2 in 24 (UCSC), and 2 in 22 (UN).

If both alleles had more than 20 reads, the modification ratios were used directly as calculated by m6Anet (63% of total). However, when one of the two alleles has a read count <20, we recalculated the modification ratio leveraging modification information from all reads, without distinguishing the two alleles. Without loss of generality, let us assume that allele A had fewer than 20 reads assigned and hence was not considered by m6Anet. In this case, we first calculated its read count by following:

$$\text{Allele A read count} = \text{Total read counts} - \text{Allele B read counts} - \text{Undefined read counts}$$

If at least 10 reads were assigned to Allele A, we retained this site for further analysis and recalculated the modification ratio of Allele A using the following formula:

$$\text{Allele A mod ratio} = \frac{(\text{Total read count} \times \text{mod ratio from total reads}) - (\text{Allele B read count} \times \text{Allele B mod ratio})}{\text{Allele A read count}}$$

See Supplemental Figure 14 for a schematic description of this procedure.

We identified statistically significant ASM sites using a bootstrapping-based statistical test. First, for each allele, methylated read counts were derived by multiplying modification ratios with total read numbers. We then sampled the number of methylated reads for each allele with replacement and calculated the difference between the modification ratios using the resampled read counts (McLachlan and Rathnayake 2014; Banjanovic and Osborne 2016). This resampling procedure was repeated 1,000,000 times and a one-sided *P*-value was calculated by using an effect size threshold (T) of 0 or 0.1 as follows:

$$P \text{ value} = 1 - \frac{|\{b: (\text{Mod Ratio}_b^A - \text{Mod Ratio}_b^B) > (T)\}|}{1,000,000}$$

where, Mod Ratio_b^A represents the bootstrap resampling value for the allele with the higher observed m^{6}A modification ratio. An aggregate *P*-value was calculated by combining the *P*-values from each replicate using the harmonic mean method (Wilson 2019). FDR was calculated by the Benjamini and Hochberg method (Yoav Benjamini 1995). Finally, statistically significant ASM sites were defined if adjusted harmonic mean *P*-values (FDR) were below 0.1. For instance, with an effect size threshold of 0.1 (T= 0.1), if none of the randomizations exceed this cutoff, it would suggest that the modification ratios of the two alleles from resampled reads are highly similar. Consequently, the probability of this site being an ASM would be very low, corresponding to a *P*-value of 1.

To assess the significance of UGACU being the most common DRACH variant among Group 2 ASM sites, we randomly resampled 17 motifs 10,000 times. For the resampling, we used the observed frequency of each of the 15 instances of the DRACH motif among sites with a high probability of modification. In these random samples, only 84 instances had UGACU as the most frequent motif hence corresponding to a *P*-value of 0.0084.

mESC MeRIP-seq experiments and analyses

MeRIP-seq libraries were prepared with EpiNext CUT&RUN RNA m^{6}A-Seq Kit (EpiGentek). The three replicates of mESCs were collected from different numbers of passages on separate dates. The total RNA was extracted with Direct-zol RNA Purification Kits (Zymo Research R2050). Seven micrograms of total RNA was subjected to immunoprecipitation with an m^{6}A antibody (P9016, EpiGentek, 1:100 dilution) and digested with cleavage enzyme on beads. The beads were then washed three times with a wash buffer and protein digestion buffer, and RNA was eluted in 13 µL of the elution buffer. The sequencing libraries were generated using the Diagenode small RNA sequencing kit (Diagenode C05030001). The libraries were sequenced on a NovaSeq 6000 system (Illumina).

Adaptor sequences were trimmed from raw reads with cutadapt v4.7 (Martin 2011) using the following parameters: -a `AAAAAAAAAACAAAAAAAAAA` -G `^TTTTTTTTTGTTTTTTTTTT` -A `AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT` -n 2 --overlap=4 --trimmed-only --maximum-length=150 --minimum-length=31 --quality-cutoff=28. Trimmed reads were aligned to the N-masked mouse transcriptome with STAR v2.7.10b (Dobin et al. 2013). Reads with low mapping quality were discarded (mapping quality <10) and indexed with SAMtools v1.15.1 (Bonfield et al. 2021; Danecek et al. 2021).

To compute the allele bias ratio, we counted the number of allelic reads that harbor at least one SNP within 100 bp of the ASM sites. Out of 23 ASM sites, four did not have genetic differences within 100 bp of the methylated position and 15 had fewer than 40 reads across the three replicates (Supplemental Table 6). Allele bias for the remaining four sites was calculated as

$$\text{Allele bias ratio} = \frac{\text{CAST allele read count}}{\text{B6 allele read count} + \text{CAST allele read count}}$$

mESC RNA-seq and ribosome profiling library preparation

Five million mESCs were scraped and transferred to 1.5 mL tubes in iced lysis buffer (20 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 100 µg/mL Cycloheximide, 1% Triton-X). The lysate was clarified by centrifugation at 1300*g* for 10 min at 4°C. Ten percent of the clarified lysate by volume was reserved for RNA extraction using Direct-zol RNA Purification Kits (Zymo Research R2050). The RNA-seq libraries were prepared using the NEBNext Ultra II RNA Library Prep Kit.

Seven microliters of RNase I (Invitrogen AM2294) was added to the remaining lysate for ribosome profiling (following REF with minor modifications) (Rao et al. 2021). Digestion was carried out for 1 h at 4°C, and stopped with ribonucleoside vanadyl complex (NEB S1402S) at a final concentration of 20 mM. Digested lysates were layered on a sucrose cushion (20 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 34% sucrose, 1 mM DTT) and the ribosomes were pelleted by centrifugation in a SW 41 Ti rotor (Beckman Coulter) at 38,000 rpm for 2.5 h at 4°C. RNA was isolated with the RNeasy Mini RNA Kit (Qiagen 74104) and size-selected by running 5 µg of each sample on a 15% polyacrylamide TBE-UREA gel. The 21–34 nt RNA fragments were excised and extracted by crushing the gel fragment in the RNA extraction buffer (300 mM sodium acetate pH 5.5, 1 mM EDTA, 0.25% SDS) followed by an overnight incubation at room temperature. The sample was passed through a Spin X filter (Corning 8160) and the flowthrough was ethanol precipitated in the presence of 5 mM $MgCl_2$ and 1 µL GlycoBlue (Invitrogen AM9516). The RNA pellet was resuspended in 10 µL of RNase-free water and immediately processed for library preparation using D-Plex Small RNA-seq kit (Diagenode C05030001) with slight modifications. The dephosphorylation reaction was supplemented with 0.5 µL T4 PNK (NEB M0201S), and the reaction was incubated for 25 min. Subsequently, the complementary DNA (cDNA) was amplified for 12 PCR cycles. We used AMPure XP bead cleanup (1.8×), followed by size selection using 3% agarose, dye-free gel cassettes with internal standards (Sage Science BDQ3010) on the BluePippin platform. Sequencing was performed on a NovaSeq 6000 platform.

Read processing of RNA-seq and ribosome profiling

For mESC, RNA-seq and Ribo-seq data were processed using RiboFlow v0.0.1 (Ozadam et al. 2020). For the Ribo-seq library, Unique Molecular Identifier (UMI) sequences were isolated employing the following parameters: "umi_tools extract -p '^(?P<umi_1>.{12})(?P<discard_1>.{4}).+$' --extract-method=regex". Subsequently, reads underwent clipping with the parameters "-a AAAAAAAAAACAAAAAAAAAA --overlap=4 --trimmed-only". Trimmed reads were then filtered by alignment to mouse rRNA and tRNA sequences with Bowtie 2 v 7.3.0 (Langmead and Salzberg 2012), and utilizing unaligned reads for subsequent alignment to the N-masked transcriptome. Following transcriptome alignment, reads (mapping quality >2) were preserved and deduplicated utilizing UMI-tools (Smith et al. 2017) directional adjacency method with the parameter "--read-length".

In mESC RNA-seq analysis, we clipped read adaptors using cutadapt, v1.18 (Martin 2011) "-a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT -O 8 -m 20 --cores=8". The reads were aligned to the N-masked transcriptome, and the read count for RNA-seq and Ribo-seq were obtained from .Ribo files with RiboR (Ozadam et al. 2020). Quality control metrics were inspected using RiboR and RiboGraph (Chacko et al. 2024).

For the NA12878 sample, we analyzed RNA-seq and Ribo-seq data (NCBI Gene Expression Omnibus [GEO; https://www.ncbi.nlm.nih.gov/geo/] under accession number GSE65912) based on the study by Cenik et al. (2015). We trimmed read NEB adaptors using the same option that we used for mESC RNA-seq, then filtered by aligning to human rRNA and tRNA sequences with Bowtie 2 v7.3.0. Unaligned reads were subsequently mapped to the N-masked human transcriptome.

Allele-specific RNA expression and ribosome occupancy analyses

Utilizing the aligned BAM files obtained from RNA-seq and Ribo-seq, ASE counts were acquired using GATK v3.8.1 ASEReadCounter (McKenna et al. 2010). The fraction of reads corresponding to the two alleles was calculated for all loci. After normalization by count per million reads, ASE scores were computed by dividing the read count from a certain allele to the sum of the read counts from both alleles (Castel et al. 2015; Liu et al. 2018).

To compare the allele-specific RNA expression and ribosome occupancy ratio in genes which have ASM, we obtained an allele bias ratio to the same allele (e.g., Allele A) which showed ASM (e.g., Allele A bias methylation).

$$\text{Allele A bias ratio} = \frac{\text{Allele A read count}}{\text{Allele A read count} + \text{Allele B read count}}$$

To quantify the relationship between the allele-bias ratio from long-read sequencing and short-read sequencing, we calculate the weighted rho using long-read sequencing read counts as weights. The correlation was calculated by using the R package "boot", v1.3-30 (Davison and Hinkley 1997).

Genomic DNA extraction and PCR for genetic variant verification

Genomic DNA from mESCs was extracted using the Quick-DNA Miniprep Plus Kit (Zymo D4068). The target regions were amplified by PCR using Q5 High-Fidelity DNA Polymerase (NEB M0491S) using 17 primer pairs (Supplemental Table 9). The resulting PCR products were purified using the NucleoSpin Gel and PCR Clean-up Kit (Takara 740609.250) and sequenced by Sanger sequencing (ACGT, Inc. DNA sequencing service).

Data access

All mESC short-read sequencing data sets and the base-called DRS data generated in this study have been submitted to the NCBI Sequence Read Archive (SRA; https://www.ncbi.nlm.nih.gov/sra) under accession number SRP486746. The ONT DRS data are available on Zenodo (https://zenodo.org) under the following record numbers: mESC replicate 1 (10815502, 13255832, and 13256383), mESC replicate 2 (13257639, 13259594, 13273847, 13275906, 13278114, and 13277067), and *Mettl3* knockout cells (13257082). All custom scripts used to perform bioinformatics analyses are available at GitHub (https://github.com/DayeaPark/Allele-specific-m6A-modification) and as Supplemental Code.

Competing interest statement

The authors declare no competing interests.

Acknowledgments

We thank Dr. David Spector for kindly providing hybrid mESCs. We appreciate the insightful comments provided by Dr. Ian Hoskins on the manuscript. This work was supported by the National Institutes of Health grants (HD110096, GM150667), as well as a Welch Foundation grant (F-2027-20230405) (C.C.). Figures were generated using BioRender.com under a publication license (JO26GB978U). All original text in this paper was authored by the researchers. Additionally, we acknowledge the assistance of a Large Language Model (OpenAI, ChatGPT v3.5) for suggesting edits aimed at improving clarity and grammar.

Author contributions: D.P.: Conceptualization, methodology, investigation, formal analysis, visualization, and writing—original draft; C.C.: funding acquisition, project administration, and supervision; C.C. and D.P.: editing and review.

References

Adams DJ, Doran AG, Lilue J, Keane TM. 2015. The Mouse Genomes Project: a repository of inbred laboratory mouse strain genomes. *Mamm Genome* **26:** 403–412. doi:10.1007/s00335-015-9579-6

Akhtar J, Lugoboni M, Junion G. 2021. m^{6}A RNA modification in transcription regulation. *Transcription* **12:** 266–276. doi:10.1080/21541264.2022.2057177

Amoah K, Hsiao Y-HE, Bahn JH, Sun Y, Burghard C, Tan BX, Yang E-W, Xiao X. 2021. Allele-specific alternative splicing and its functional genetic variants in human tissues. *Genome Res* **31:** 359–371. doi:10.1101/gr.265637.120

Balasooriya GI, Spector DL. 2022. Allele-specific differential regulation of monoallelically expressed autosomal genes in the cardiac lineage. *Nat Commun* **13:** 5984. doi:10.1038/s41467-022-33722-x

Banjanovic ES, Osborne JW. 2016. Confidence intervals for effect sizes: applying bootstrap resampling. *Pract Assess Res Eval* **21:** 5. doi:10.7275/dz3r-8n08

Berlivet S, Scutenaire J, Deragon J-M, Bousquet-Antonelli C. 2019. Readers of the m^{6}A epitranscriptomic code. *Biochim Biophys Acta Gene Regul Mech* **1862:** 329–342. doi:10.1016/j.bbagrm.2018.12.008

Bokar JA, Rath-Shambaugh ME, Ludwiczak R, Narayan P, Rottman F. 1994. Characterization and partial purification of mRNA N^6-adenosine methyltransferase from HeLa cell nuclei. Internal mRNA methylation requires a multisubunit complex. *J Biol Chem* **269:** 17697–17704. doi:10.1016/S0021-9258(17)32497-3

Bonfield JK, Marshall J, Danecek P, Li H, Ohan V, Whitwham A, Keane T, Davies RM. 2021. HTSlib: C library for reading/writing high-throughput sequencing data. *Gigascience* **10:** giab007. doi:10.1093/gigascience/giab007

Buttler J, Drown DM. 2023. Accuracy and completeness of long read metagenomic assemblies. *Microorganisms* **11:** 96. doi:10.3390/microorganisms11010096

Cao S, Zhu H, Cui J, Liu S, Li Y, Shi J, Mo J, Wang Z, Wang H, Hu J, et al. 2023. Allele-specific RNA N^6-methyladenosine modifications reveal

functional genetic variants in human tissues. *Genome Res* **33:** 1369–1380. doi:10.1101/gr.277704.123
Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T. 2015. Tools and best practices for data processing in allelic expression analysis. *Genome Biol* **16:** 195. doi:10.1186/s13059-015-0762-6
Castel SE, Aguet F, Mohammadi P, GTEx Consortium, Ardlie KG, Lappalainen T. 2020. A vast resource of allelic expression data spanning human tissues. *Genome Biol* **21:** 234. doi:10.1186/s13059-020-02122-z
Cenik C, Cenik ES, Byeon GW, Grubert F, Candille SI, Spacek D, Alsallakh B, Tilgner H, Araya CL, Tang H, et al. 2015. Integrative analysis of RNA, translation, and protein levels reveals distinct regulatory variation across humans. *Genome Res* **25:** 1610–1621. doi:10.1101/gr.193342.115
Cenik C, Chua HN, Singh G, Akef A, Snyder MP, Palazzo AF, Moore MJ, Roth FP. 2017. A common class of transcripts with 5′-intron depletion, distinct early coding sequence features, and N^1-methyladenosine modification. *RNA* **23:** 270–283. doi:10.1261/rna.059105.116
Cesaro B, Tarullo M, Fatica A. 2023. Regulation of gene expression by m^6A_m RNA modification. *Int J Mol Sci* **24:** 2277. doi:10.3390/ijms24032277
Chacko J, Ozadam H, Cenik C. 2024. RiboGraph: an interactive visualization system for ribosome profiling data at read length resolution. *Bioinformatics* **40:** btae369. doi:10.1093/bioinformatics/btae369
Chen K, Luo G-Z, He C. 2015. High-resolution mapping of N^6-methyladenosine in transcriptome and genome using a photo-crosslinking-assisted strategy. *Methods Enzymol* **560:** 161–185. doi:10.1016/bs.mie.2015.03.012
Chen Z, Lei C, Wang C, Li N, Srivastava M, Tang M, Zhang H, Choi JM, Jung SY, Qin J, et al. 2019. Global phosphoproteomic analysis reveals ARMC10 as an AMPK substrate that regulates mitochondrial dynamics. *Nat Commun* **10:** 104. doi:10.1038/s41467-018-08004-0
Cho H, Davis J, Li X, Smith KS, Battle A, Montgomery SB. 2014. High-resolution transcriptome analysis with long-read RNA sequencing. *PLoS One* **9:** e108095. doi:10.1371/journal.pone.0108095
Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools. *Gigascience* **10:** giab008. doi:10.1093/gigascience/giab008
Dassi E, Baranov PV, Pelizzola M. 2020. Editorial: Computational epitranscriptomics: bioinformatic approaches for the analysis of RNA modifications. *Front Genet* **11:** 630360. doi:10.3389/fgene.2020.630360
Davison AC, Hinkley DV. 1997. *Bootstrap methods and their application*. Cambridge University Press, New York.
de la Chapelle A. 2009. Genetic predisposition to human disease: allele-specific expression and low-penetrance regulatory loci. *Oncogene* **28:** 3345–3348. doi:10.1038/onc.2009.194
Deng S, Zhang J, Su J, Zuo Z, Zeng L, Liu K, Zheng Y, Huang X, Bai R, Zhuang L, et al. 2022. RNA m^6A regulates transcription via DNA demethylation and chromatin accessibility. *Nat Genet* **54:** 1427–1437. doi:10.1038/s41588-022-01173-1
DeVeale B, van der Kooy D, Babak T. 2012. Critical evaluation of imprinted gene expression by RNA-seq: a new perspective. *PLoS Genet* **8:** e1002600. doi:10.1371/journal.pgen.1002600
Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29:** 15–21. doi:10.1093/bioinformatics/bts635
Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, et al. 2012. Topology of the human and mouse m^6A RNA methylomes revealed by m^6A-seq. *Nature* **485:** 201–206. doi:10.1038/nature11112
Fan J, Hu J, Xue C, Zhang H, Susztak K, Reilly MP, Xiao R, Li M. 2020. ASEP: gene-based detection of allele-specific expression across individuals in a population by RNA sequencing. *PLoS Genet* **16:** e1008786. doi:10.1371/journal.pgen.1008786
Fournier C, Goto Y, Ballestar E, Delaval K, Hever AM, Esteller M, Feil R. 2002. Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J* **21:** 6560–6570. doi:10.1093/emboj/cdf655
Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A, et al. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* **15:** 201–206. doi:10.1038/nmeth.4577
Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, Winkler R, Nir R, Lasman L, Brandis A, et al. 2019. Deciphering the "m^6A code" via antibody-independent quantitative profiling. *Cell* **178:** 731–747.e16. doi:10.1016/j.cell.2019.06.013
Gicquel C, Rossignol S, Cabrol S, Houang M, Steunou V, Barbu V, Danton F, Thibaud N, Le Merrer M, Burglen L, et al. 2005. Epimutation of the telomeric imprinting center region on chromosome 11p15 in Silver-Russell syndrome. *Nat Genet* **37:** 1003–1007. doi:10.1038/ng1629
Glinos DA, Garborcauskas G, Hoffman P, Ehsan N, Jiang L, Gokden A, Dai X, Aguet F, Brown KL, Garimella K, et al. 2022. Transcriptome variation in human tissues revealed by long-read sequencing. *Nature* **608:** 353–359. doi:10.1038/s41586-022-05035-y
Hansen NF. 2016. Variant calling from next generation sequence data. In *Statistical genomics: methods and protocols* (eds. Mathé E, Davis S), pp. 209–224. Springer, New York.
Hashimoto K. 2021. CD137 as an attractive T cell co-stimulatory target in the TNFRSF for immuno-oncology drug development. *Cancers (Basel)* **13:** 2288. doi:10.3390/cancers13102288
He PC, He C. 2021. m^6a RNA methylation: from mechanisms to therapeutic potential. *EMBO J* **40:** e105977. doi:10.15252/embj.2020105977
He PC, Wei J, Dou X, Harada BT, Zhang Z, Ge R, Liu C, Zhang L-S, Yu X, Wang S, et al. 2023. Exon architecture controls mRNA m^6A suppression and gene expression. *Science* **379:** 677–682. doi:10.1126/science.abj9090
Helm M, Lyko F, Motorin Y. 2019. Limited antibody specificity compromises epitranscriptomic analyses. *Nat Commun* **10:** 5669. doi:10.1038/s41467-019-13684-3
Hendra C, Pratanwanich PN, Wan YK, Goh WSS, Thiery A, Göke J. 2022. Detection of m^6A from direct RNA sequencing using a multiple instance learning framework. *Nat Methods* **19:** 1590–1598. doi:10.1038/s41592-022-01666-1
Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, Chen Z, Deng X, Xiao G, Auer F, et al. 2019. Histone H3 trimethylation at lysine 36 guides m^6A RNA modification co-transcriptionally. *Nature* **567:** 414–419. doi:10.1038/s41586-019-1016-7
Jain S, Koziej L, Poulis P, Kaczmarczyk I, Gaik M, Rawski M, Ranjan N, Glatt S, Rodnina MV. 2023. Modulation of translational decoding by m^6A modification of mRNA. *Nat Commun* **14:** 4784. doi:10.1038/s41467-023-40422-7
Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Jin Z, Yang C, Chen Y. 2021. The role of m^6A modification in the biological functions and diseases. *Signal Transduct Target Ther* **6:** 74. doi:10.1038/s41392-020-00450-x
Joglekar A, Prjibelski A, Mahfouz A, Collier P, Lin S, Schlusche AK, Marrocco J, Williams SR, Haase B, Hayes A, et al. 2021. A spatially resolved brain region- and cell type-specific isoform atlas of the postnatal mouse brain. *Nat Commun* **12:** 463. doi:10.1038/s41467-020-20343-5
Kabelitz D, Déchanet-Merville J. 2015. Editorial: "Recent advances in γ/δ T cell biology: new ligands, new functions, and new translational perspectives". *Front Immunol* **6:** 371. doi:10.3389/fimmu.2015.00371
Kirton HM, Pettinger L, Gamper N. 2013. Transient overexpression of genes in neurons using nucleofection. In *Ion channels: methods and protocols* (ed. Gamper N), pp. 55–64. Humana Press, Totowa, NJ.
Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9:** 357–359. doi:10.1038/nmeth.1923
Lee Y, Choe J, Park OH, Kim YK. 2020. Molecular mechanisms driving mRNA degradation by m^6A modification. *Trends Genet* **36:** 177–188. doi:10.1016/j.tig.2019.12.007
Lee J-H, Wang R, Xiong F, Krakowiak J, Liao Z, Nguyen PT, Moroz-Omori EV, Shao J, Zhu X, Bolt MJ, et al. 2021. Enhancer RNA m6A methylation facilitates transcriptional condensate formation and gene activation. *Mol Cell* **81:** 3368–3385.e9. doi:10.1016/j.molcel.2021.07.024
Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34:** 3094–3100. doi:10.1093/bioinformatics/bty191
Lin S, Gregory RI. 2014. Methyltransferases modulate RNA stability in embryonic stem cells. *Nat Cell Biol* **16:** 129–131. doi:10.1038/ncb2914
Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat Methods* **12:** 767–772. doi:10.1038/nmeth.3453
Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, et al. 2014. A METTL3–METTL14 complex mediates mammalian nuclear RNA N^6-adenosine methylation. *Nat Chem Biol* **10:** 93–95. doi:10.1038/nchembio.1432
Liu Z, Dong X, Li Y. 2018. A genome-wide study of allele-specific expression in colorectal cancer. *Front Genet* **9:** 570. doi:10.3389/fgene.2018.00570
Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, Schwartz S, Mattick JS, Smith MA, Novoa EM. 2019. Accurate detection of m^6A RNA modifications in native RNA sequences. *Nat Commun* **10:** 4079. doi:10.1038/s41467-019-11713-9
Liu Y, Hoskins I, Geng M, Zhao Q, Chacko J, Qi K, Persyn L, Wang J, Zheng D, Zhong Y, et al. 2024. Translation efficiency covariation across cell types is a conserved organizing principle of mammalian transcriptomes. bioRxiv doi:10.1101/2024.08.11.607360
Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled de novo using only nanopore sequencing data. *Nat Methods* **12:** 733–735. doi:10.1038/nmeth.3444
Mao Y, Dong L, Liu X-M, Guo J, Ma H, Shen B, Qian S-B. 2019. m^6A in mRNA coding regions promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun* **10:** 5332. doi:10.1038/s41467-019-13317-9

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J* **17:** 10. doi:10.14806/ej.17.1.200

Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur J-J, Chen Q, et al. 2017. Reversible methylation of m^6A_m in the 5′ cap controls mRNA stability. *Nature* **541:** 371–375. doi:10.1038/nature21022

McIntyre ABR, Gokhale NS, Cerchietti L, Jaffrey SR, Horner SM, Mason CE. 2020. Limits in the detection of m^6A changes using MeRIP/m^6A-seq. *Sci Rep* **10:** 6590. doi:10.1038/s41598-020-63355-3

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* **20:** 1297–1303. doi:10.1101/gr.107524.110

McLachlan GJ, Rathnayake S. 2014. On the number of components in a Gaussian mixture model. *Wiley Interdiscip Rev Data Min Knowl Discov* **4:** 341–355. doi:10.1002/widm.1135

Meyer KD. 2019a. DART-seq: an antibody-free method for global m^6A detection. *Nat Methods* **16:** 1275–1280. doi:10.1038/s41592-019-0570-0

Meyer KD. 2019b. m^6A-mediated translation regulation. *Biochim Biophys Acta Gene Regul Mech* **1862:** 301–309. doi:10.1016/j.bbagrm.2018.10.006

Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. *Cell* **149:** 1635–1646. doi:10.1016/j.cell.2012.05.003

Min K-W, Zealy RW, Davila S, Fomin M, Cummings JC, Makowsky D, Mcdowell CH, Thigpen H, Hafner M, Kwon S-H, et al. 2018. Profiling of m6A RNA modifications identified an age-associated regulation of *AGO2* mRNA stability. *Aging Cell* **17:** e12753. doi:10.1111/acel.12753

Mohammadi P, Castel SE, Brown AA, Lappalainen T. 2017. Quantifying the regulatory effect size of *cis*-acting genetic variation using allelic fold change. *Genome Res* **27:** 1872–1884. doi:10.1101/gr.216747.116

Nembaware V, Lupindo B, Schouest K, Spillane C, Scheffler K, Seoighe C. 2008. Genome-wide survey of allele-specific splicing in humans. *BMC Genomics* **9:** 265. doi:10.1186/1471-2164-9-265

Ozadam H, Geng M, Cenik C. 2020. Riboflow, RiboR and RiboPy: an ecosystem for analyzing ribosome profiling data at read length resolution. *Bioinformatics* **36:** 2929–2931. doi:10.1093/bioinformatics/btaa028

Ozadam H, Tonn T, Han CM, Segura A, Hoskins I, Rao S, Ghatpande V, Tran D, Catoe D, Salit M, et al. 2023. Single-cell quantification of ribosome occupancy in early mouse development. *Nature* **618:** 1057–1064. doi:10.1038/s41586-023-06228-9

Pai AA, Cain CE, Mizrahi-Man O, De Leon S, Lewellen N, Veyrieras J-B, Degner JF, Gaffney DJ, Pickrell JK, Stephens M, et al. 2012. The contribution of RNA decay quantitative trait loci to inter-individual variation in steady-state gene expression levels. *PLoS Genet* **8:** e1003000. doi:10.1371/journal.pgen.1003000

Piccolo SR, Frampton MB. 2016. Tools and techniques for computational reproducibility. *Gigascience* **5:** 30. doi:10.1186/s13742-016-0135-4

Prendergast JGD, Tong P, Hay DC, Farrington SM, Semple CAM. 2012. A genome-wide screen in human embryonic stem cells reveals novel sites of allele-specific histone modification associated with known disease loci. *Epigenetics Chromatin* **5:** 6. doi:10.1186/1756-8935-5-6

Rao S, Hoskins I, Tonn T, Garcia PD, Ozadam H, Sarinay Cenik E, Cenik C. 2021. Genes with 5′ terminal oligopyrimidine tracts preferentially escape global suppression of translation by the SARS-CoV-2 Nsp1 protein. *RNA* **27:** 1025–1045. doi:10.1261/rna.078661.120

Robles-Espinoza CD, Mohammadi P, Bonilla X, Gutierrez-Arcelus M. 2021. Allele-specific expression: applications in cancer and technical considerations. *Curr Opin Genet Dev* **66:** 10–19. doi:10.1016/j.gde.2020.10.007

Rozowsky J, Abyzov A, Wang J, Alves P, Raha D, Harmanci A, Leng J, Bjornson R, Kong Y, Kitabayashi N, et al. 2011. AlleleSeq: analysis of allele-specific expression and binding in a network framework. *Mol Syst Biol* **7:** 522. doi:10.1038/msb.2011.54

Serrat R, Mirra S, Figueiro-Silva J, Navas-Pérez E, Quevedo M, López-Doménech G, Podlesniy P, Ulloa F, Garcia-Fernàndez J, Trullas R, et al. 2014. The Armc10/SVH gene: genome context, regulation of mitochondrial dynamics and protection against Aβ-induced mitochondrial fragmentation. *Cell Death Dis* **5:** e1163. doi:10.1038/cddis.2014.121

Singh P, Cho J, Tsai SY, Rivas GE, Larson GP, Szabó PE. 2010. Coordinated allele-specific histone acetylation at the differentially methylated regions of imprinted genes. *Nucleic Acids Res* **38:** 7974–7990. doi:10.1093/nar/gkq680

Slobodin B, Han R, Calderone V, Vrielink JAFO, Loayza-Puch F, Elkon R, Agami R. 2017. Transcription impacts the efficiency of mRNA translation via co-transcriptional N^6-adenosine methylation. *Cell* **169:** 326–337.e12. doi:10.1016/j.cell.2017.03.031

Smith T, Heger A, Sudbery I. 2017. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res* **27:** 491–499. doi:10.1101/gr.209601.116

Song H, Song J, Cheng M, Zheng M, Wang T, Tian S, Flavell RA, Zhu S, Li H-B, Ding C, et al. 2021. METTL3-mediated m^6A RNA methylation promotes the anti-tumour immunity of natural killer cells. *Nat Commun* **12:** 5522. doi:10.1038/s41467-021-25803-0

Tilgner H, Jahanbani F, Blauwkamp T, Moshrefi A, Jaeger E, Chen F, Harel I, Bustamante CD, Rasmussen M, Snyder MP. 2015. Comprehensive transcriptome analysis using synthetic long-read sequencing reveals molecular co-association of distant splicing events. *Nat Biotechnol* **33:** 736–742. doi:10.1038/nbt.3242

Tilgner H, Jahanbani F, Gupta I, Collier P, Wei E, Rasmussen M, Snyder M. 2018. Microfluidic isoform sequencing shows widespread splicing coordination in the human transcriptome. *Genome Res* **28:** 231–242. doi:10.1101/gr.230516.117

Tsang S, Sun Z, Luke B, Stewart C, Lum N, Gregory M, Wu X, Subleski M, Jenkins NA, Copeland NG, et al. 2005. A comprehensive SNP-based genetic analysis of inbred mouse strains. *Mamm Genome* **16:** 476–480. doi:10.1007/s00335-005-0001-7

Vantourout P, Laing A, Woodward MJ, Zlatareva I, Apolonia L, Jones AW, Snijders AP, Malim MH, Hayday AC. 2018. Heteromeric interactions regulate butyrophilin (BTN) and BTN-like molecules governing γδ T cell biology. *Proc Natl Acad Sci* **115:** 1039–1044. doi:10.1073/pnas.1701237115

Viscardi MJ, Arribere JA. 2022. Poly(a) selection introduces bias and undue noise in direct RNA-sequencing. *BMC Genomics* **23:** 530. doi:10.1186/s12864-022-08762-8

Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, et al. 2014. N^6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505:** 117–120. doi:10.1038/nature12730

Wang H, Coligan JE, Morse HC III. 2016. Emerging functions of natural IgM and its Fc receptor FCMR in immune homeostasis. *Front Immunol* **7:** 99. doi:10.3389/fimmu.2016.00099

Wang S, Lv W, Li T, Zhang S, Wang H, Li X, Wang L, Ma D, Zang Y, Shen J, et al. 2022. Dynamic regulation and functions of mRNA m^6A modification. *Cancer Cell Int* **22:** 48. doi:10.1186/s12935-022-02452-x

Wick RR, Judd LM, Holt KE. 2019. Performance of neural network basecalling tools for Oxford Nanopore sequencing. *Genome Biol* **20:** 129. doi:10.1186/s13059-019-1727-y

Wilson DJ. 2019. The harmonic mean *p*-value for combining dependent tests. *Proc Natl Acad Sci* **116:** 1195–1200. doi:10.1073/pnas.1814092116

Workman RE, Tang AD, Tang PS, Jain M, Timp W. 2018. Nanopore native RNA sequencing of a human poly(A) transcriptome. https://www.researchgate.net/publication/328855598_Nanopore_native_RNA_sequencing_of_a_human_polyA_transcriptome (Accessed December 13, 2023).

Wu J, Hu W, Li S. 2023. Long-read transcriptome sequencing reveals allele-specific variants at high resolution. *Trends Genet* **39:** 31–33. doi:10.1016/j.tig.2022.09.001

Yang X, Triboulet R, Liu Q, Sendinc E, Gregory RI. 2022. Exon junction complex shapes the m^6A epitranscriptome. *Nat Commun* **13:** 7904. doi:10.1038/s41467-022-35643-1

Yoav Benjamini YH. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc* **57:** 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x

Zhang Z, Chen T, Chen H-X, Xie Y-Y, Chen L-Q, Zhao Y-L, Liu B-D, Jin L, Zhang W, Liu C, et al. 2021. Systematic calibration of epitranscriptomic maps using a synthetic modification-free RNA library. *Nat Methods* **18:** 1213–1222. doi:10.1038/s41592-021-01280-7

Zheng D, Wang J, Persyn L, Liu Y, Montoya FU, Cenik C, Agarwal V. 2024. Predicting the translation efficiency of messenger RNA in mammalian cells. bioRxiv doi:10.1101/2024.08.11.607362

Zitovsky JP, Love MI. 2019. Fast effect size shrinkage software for beta-binomial models of allelic imbalance. *F1000Res* **8:** 2024. doi:10.12688/f1000research.20916.1

Received March 14, 2024; accepted in revised form October 23, 2024.