TITLE

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

RUNNING TITLE

Allele-Specific N^6-methyladenosine profiling

AUTHORS

Dayea Park[1], Can Cenik[1]*

AFFILIATIONS

[1] Department of Molecular Biosciences, University of Texas at Austin, Austin, TX 78712, USA

* Corresponding author: ccenik@austin.utexas.edu

ABSTRACT

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 promises 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 (m6A) modifications in native mRNA. We utilized human and mouse cells with known genetic variants to assign allelic origin of each mRNA molecule combined with a supervised machine learning model to detect read-level m6A modification ratios. Our analyses revealed the importance of sequences adjacent to the DRACH-motif in determining m6A deposition, in addition to allelic differences that directly alter the motif.

Moreover, we discovered allele-specific m6A modification (ASM) 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 advancement promises to advance our understanding of the role of genetics in determining mRNA modifications.

Keywords: LRS Special Issue, N^6-Methyladenosine, Allele-specific expression

INTRODUCTION

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 (Castel et al. 2015, 2020; Fan et al. 2020; de la Chapelle 2009; Gicquel et al. 2005). 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 (Singh et al. 2010; Prendergast et al. 2012; Fournier et al. 2002). Particularly in development, DNA methylation regulates allele-specific expression, and coordinates X-chromosome inactivation in females for dosage compensation (Morcos et al. 2011; Benton et al. 2019). Furthermore, H3K27me3 marks in the early embryo mediate imprinted mono-allelic expression and persist from oocyte development through the blastocyst stage (Santini et al. 2021; Inoue 2023; Sergeeva et al. 2023).

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

N^6-Methyladenosine (m6A), the most prevalent RNA modification of mRNAs, has been suggested to impact diverse mechanisms to regulate gene expression (He and He 2021; Lee et al. 2020). Various interactions with the methyltransferase complex or m6A reader proteins impact several steps of mRNA metabolism, including splicing, export, translation, recruitment of RNA binding

60 proteins, and stability (Wang et al. 2022; Jiang et al. 2021; Lin and Gregory 2014; Akhtar et al.
61 2021).
62
63 Transcriptome-wide patterns of m6A RNA modifications have typically been studied using short-
64 read sequencing coupled with either antibody-dependent methods such as MeRIP-seq (Meyer et
65 al. 2012) or enzymatic/chemical approaches (Garcia-Campos et al. 2019; Meyer 2019a; Song et
66 al. 2021). Among these methods, MeRIP-seq remains the most popular choice despite its
67 limitations leading to elevated false-positive rates, attributable to nonspecific antibody binding
68 (Helm et al. 2019; McIntyre et al. 2020; Zhang et al. 2021). Furthermore, all short-read sequencing
69 strategies to detect m6A are inherently limited to aggregate measurements and are incapable of
70 quantification at an individual molecular level.
71
72 In contrast, Oxford Nanopore Technology (ONT) RNA sequencing enables direct detection of
73 RNA modifications such as m6A with single molecule resolution (Garalde et al. 2018). The electric
74 signal recorded by the ONT sequencing platform was shown to be altered by the presence of
75 RNA modifications (Workman et al. 2018; Garalde et al. 2018; Pelizzola et al. 2021).
76 Subsequently, machine learning methods can utilize the electronic current signal intensity to
77 identify potential m6A sites from long-read data (Hendra et al. 2022).
78
79 The ability to directly detect m6A modifications on the ONT RNA sequencing platform provides a
80 unique opportunity to combine these advantages with the ability of long-read sequencing to
81 facilitate ASE analysis. Long-read sequencing improves upon the fundamental limitations of short-
82 read sequencing for allele-specific analysis by detecting an increased number of single nucleotide
83 polymorphisms (SNPs) on a read, enabling its precise allelic assignment (Cho et al. 2014). This
84 feature has been leveraged to characterize the genetic effects of rare and common variants in
85 the transcriptome (Glinos et al. 2022). Furthermore, long-read sequencing enables

86 comprehensive analysis of splicing (Tilgner et al. 2015, 2018; Joglekar et al. 2021) which has
87 fundamental importance for determining mRNA modifications due to their dependence on splicing
88 patterns and transcript architecture (Yang et al. 2022; He et al. 2023; Cenik et al. 2017).
89
90 Here we introduce a novel approach harnessing ONT direct RNA sequencing to surmount the
91 persistent constraints of m6A detection methods for allele-specific analyses. Our findings
92 establish long-read sequencing of RNA as a robust solution for allele-specific m6A modification
93 (ASM) analysis.
94
95 **RESULTS**
96
97 **ONT direct RNA sequencing enables the identification of allele-specific m6A modifications**
98 **in hybrid mESCs**
99 We leveraged ONT direct RNA sequencing to simultaneously determine the allelic origin of each
100 molecule along with its m6A modification status. To achieve high accuracy of allelic assignment
101 of individual molecules, we decided to use mouse embryonic stem cells (mESCs) that were
102 derived from a cross of two highly genetically diverse mouse strains (C57BL/6J x CAST/EiJ, B6
103 x CAST) (Balasooriya and Spector 2022). Using ONT direct RNA sequencing, we generated two
104 replicates of 2.3 and 2.2 million reads from these hybrid mESCs.
105
106 To assess our ability to accurately detect m6A modifications, we generated mESC clones where
107 methyltransferase-like 3 (*Mettl3*), the major methyltransferase for m6A modifications, is knocked
108 out (Bokar et al. 1994; Liu et al. 2013) (Supplemental Fig. 1A, Supplemental Table 1; Methods).
109 As expected, in wild-type mESCs ~6% of the adenines within the context of a DRACH motif had
110 a high-probability (>0.85) of being modified, compared to only 0.2% in those with *Mettl3* knockout
111 (Supplemental Fig. 1B). Among the sites displaying a high-probability modification ratio, the levels

112 of modification ratios were consistently higher in wild-type compared to *Mettl3* knockout cells
113 (Median modification ratio 0.629 and 0.512, respectively). Moreover, the modified adenines were
114 predominantly clustered near the 3' end of coding sequences, which is consistent with the
115 expected pattern of m6A RNA modifications (Meyer et al. 2012; Zhang et al. 2019) (Fig. 1A-B;
116 Supplemental Fig. 1C-D).
117
118 Using single nucleotide polymorphisms (SNPs), we assigned 1,110,260 (replicate 1) and 837,011
119 (replicate 2) long-reads to their allelic origins across more than 13,000 transcripts (Fig. 1C,
120 Methods). Of the detected transcripts, 8,657 transcripts had at least ten reads in both replicates.
121 In allele-specific analyses, a common challenge is reference allele bias which is the tendency for
122 reads that match the reference genome to align with a higher probability than reads containing
123 the alternate allele, potentially skewing variant detection and analysis. To minimize this bias
124 (Castel et al. 2015), we employed an N-masked transcriptome reference. This approach led to a
125 mean CAST allele ratio across all transcripts of 0.505 as opposed to 0.485 when using an
126 unmasked reference (Methods). These assignments were based on 135,380 (replicate 1) and
127 134,585 (replicate 2) informative positions on the long-reads that overlapped known genetic
128 variation between the strains (210,004 total SNPs).
129
130 As an orthogonal approach to determine ASE, we used Illumina short-read sequencing. We found
131 that RNA expression levels from the two methods were significantly correlated (Supplemental Fig.
132 2A; Spearman Correlations, 0.792-0.816 across replicates). Moreover, gene-level allele-specific
133 RNA expression was moderately concordant between the two approaches (Supplemental Fig.
134 2B; weighted Spearman correlation coefficient 0.61, Methods). Although short-read sequencing
135 produced nearly eight times more aligned reads, long-read sequencing identified 2.3 times as
136 many SNPs, demonstrating that the greater number of informative positions in long-read data
137 enhances allelic assignment accuracy and gene-specific ASE reproducibility (Spearman

138 correlation coefficient 0.63 and 0.51 for long-read and short-read sequencing, respectively;
139 Supplemental Fig. 2C-D). Taken together, these measures of quality control underscore the high
140 precision in allelic assignment achieved through our methodology.
141
142 Following allelic read assignment, we employed a supervised machine learning approach (Hendra
143 et al. 2022; Liu et al. 2019) to quantify m6A RNA modifications for the reads attributed to each
144 allele. This process revealed an equivalence in the number of reads and m6A sites among alleles,
145 indicating the allelic impartiality of our approach. Specifically, we observed similar numbers of
146 reads (564,944 and 444,514 for B6; 557,787 and 443,131 for CAST) and potential modification
147 sites (114,457 and 105,190 for B6; 112,947 and 105,117 for CAST) for each allele (Fig 2A,
148 Supplemental Table 2). This result indicates minimal to no allelic bias in the assignment of reads
149 and identification of modification sites.
150
151 The modification ratios of the candidate m6A sites were highly correlated between replicates and
152 demonstrated an even higher correlation within the same allele. Specifically, Spearman
153 correlations within the same allele were 0.82 and 0.83 for the modification ratios of the B6 and
154 CAST allele, respectively. Conversely, correlations between different alleles were slightly lower,
155 with values of 0.77 (B6 replicate 1, CAST replicate 2) and 0.75 (CAST replicate 1 and B6 replicate
156 2). In contrast, modification ratios from *Mettl3* knockout cells exhibited significantly lower
157 correlations, falling below 0.46 (Fig. 2B, Supplemental Fig. 3).
158
159 To identify allele-specific m6A modifications, we established a selection criterion centered on a
160 site probability aggregated from all reads. Therefore, we focused on m6A sites that demonstrated
161 a high probability of modification (>0.85) across reads. In mESC wild-type, an average of ~7% of
162 these candidate sites met our selection criteria. Notably, these m6A modification sites were
163 predominantly localized at the junction between the coding region and the 3' UTR, showing high

164 modification ratios (with median ratios of 0.621 for B6 and 0.627 for CAST; Fig. 2A, C). In contrast,
165 *Mettl3* knockout had only 0.8% allelic sites exhibiting high probabilities of m6A modification.
166 Moreover, these sites demonstrated a wider dispersion across various transcript regions (Fig. 2A,
167 D). Overall, these observations affirm the ability of our methodology in detecting allelic m6A
168 modifications subsequent to the assignment of reads to alleles.
169
170 **Detection of sites with significant allele-specific m6A modification**
171 The capability to accurately assign each RNA molecule to its allelic origin while concurrently
172 identifying RNA modifications allows for the investigation of positions within mRNAs that exhibit
173 differential modifications between alleles. While numerous statistical methods have been
174 developed to identify allele-specific differences in gene expression phenotypes (DeVeale et al.
175 2012), these studies underscore the challenges inherent in this analysis, including a propensity
176 for false positives when employing simple binomial tests to assess deviations from expected
177 expression levels across the two alleles (Zitovsky and Love 2019; Mohammadi et al. 2017).
178
179 To address these challenges, we implemented a conservative strategy that leverages bootstrap
180 sampling to quantify uncertainty in modification ratio estimates (Methods). This method enabled
181 us to pinpoint mRNA positions showing significant allele-specific m6A modification (ASM) (Fig.
182 3A). Among detected 14,609 and 13,542 candidate m6A sites in the replicate experiments, we
183 identified 57 ASM sites (FDR<0.1) with an average modification difference between the two alleles
184 of 0.32.
185
186 In allele-specific analysis, previous research revealed that events with larger effect sizes are more
187 likely to be reproducible and biologically relevant (Castel et al. 2020; Mohammadi et al. 2017).
188 Therefore, we repeated our statistical analyses using an effect size threshold of 0.1 corresponding
189 to the inferred modification ratio difference between the two alleles (Methods). This analysis

190 uncovered 23 sites across 22 genes indicating ASM. Notably, at these ASM sites, the distributions
191 of the resampled modification ratios from the two alleles were consistently distinct and had large
192 effect sizes, with a mean modification ratio difference of 0.48 (Fig. 3B, Supplemental Fig. 5A,
193 Supplemental Table 3). We focused our detailed analyses on this subset of ASM sites.
194
195 One inherent limitation of Oxford Nanopore Technologies (ONT) sequencing lies in its limited
196 sequencing depth, which constrains our ability to detect ASM in transcripts with low expression
197 levels. Consistently, transcripts with statistically significant ASM sites had significantly greater
198 RNA expression levels than those without, underscoring the dependence of ASM detection on
199 transcript abundance (Supplemental Fig. 4).
200
201 The genes with ASM sites are distributed across chromosomes without any discernible location
202 preferences (Supplemental Fig. 5B), and are associated with a wide range of functions (Fig. 3C).
203 A particularly notable finding was the identification of two distinct sites of ASM on the *Armc10*
204 transcript, which encodes a protein involved in mitochondrial dynamics (Chen et al. 2019; Serrat
205 et al. 2014).
206
207 Moreover, our analysis identified six B6-biased ASM sites with a higher modification ratio on the
208 B6 allele and 17 CAST-biased ASM sites with a higher ratio on the CAST allele. While the majority
209 of ASM sites were located on 3' UTRs, one B6-biased ASM (on *Dnpep*) and two CAST-biased
210 ASMs (on *Gnpat*, and *Pml*) were found in coding regions, near the stop codons (Supplemental
211 Fig. 5C).
212
213 Genome sequencing was used to identify genetic differences between the C57BL/6J (B6) and
214 CAST/EiJ (CAST) inbred lines (Tsang et al. 2005; David J. Adams, Anthony G. Doran, Jingtao
215 Lilue & Thomas M. Keane 2015), however, potential genotyping errors from these could result in

erroneous ASM calls. To address this possibility, we verified the genomic DNA sequences near the m6A modification sites using Sanger sequencing (Supplemental Table 4, 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 post-transcriptional and do not reflect genotyping errors.

Taken together, these findings highlight a key strength of our approach based on the ONT direct RNA sequencing technique, which enables the detection of m6A modifications at the individual molecule level, rather than relying on aggregate measurements.

Genetic variants influence allele-specific m6A modification patterns

Having identified sites with ASM, we proceeded to examine the relationship between these sites and the genetic variants that differentiate the two alleles. We hypothesized that local genetic differences could influence methylation efficacy, leading to differential m6A deposition. Accordingly, we categorized ASM-biased sites into two groups based on the proximity of the nearest genetic variation to the canonical m6A 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 m6A 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*).

241 Analysis of the remaining 17 sites (Group 2) revealed that five possessed genetic differences
242 adjacent to the DRACH motif. Specifically, we found six SNPs near five m6A sites: *L3mbtl2* (D-1
243 and H+1); *Trim25* and *Dnpep* (D-1); *Atmin* and *Tmbim6* (H+2) (Fig. 4C). Intriguingly, all instances
244 of SNPs at the D-1 position included a U at the D-2 position (UNUGACU). In this context, a
245 cytosine at the D-1 site correlated with higher m6A levels (*Dnpep*, 0.768; *Trim25*, 0.855; *L3mbtl2*,
246 0.789) compared to an adenine or a guanine at the D-1 site (adenine on *Dnpep*, 0.324 and *Trim25*,
247 0.073; guanine on *L3mbtl2*, 0.157; Fig. 4D). This finding highlights the significant influence of
248 nucleotides adjacent to the DRACH motif on m6A deposition, contingent upon their specific
249 genetic context.
250
251 Among remaining Group 2 ASM sites, eight had a SNP within 100 base pairs of the modified
252 adenine (*Stk38*, -64; *2810004N23Rik*, -51; *Cmtm7*, -12 bp; *Glmp*, +10 bp; *Rsl1d1*, +34 bp; *Gcsh*,
253 +42 bp; Kif11, +59; Armc10 at 1170 position, +99). Despite the limitations imposed by the read
254 length, short-read based m6A detection methods are theoretically capable of detecting SNPs
255 within 50 to 100 bp of the methylated site (Dominissini et al. 2012; Chen et al. 2015). However,
256 our method also identified four ASM sites that had no SNPs within this range hence highlighting
257 the unique strength of long-read sequencing for ASM detection.
258
259 We also noticed that Group 2 ASMs were highly enriched for the UGACU motif sequence over
260 the most commonly observed GAACU instance of the DRACH among m6A modified sites (Fig.
261 4E, p-val = 0.0084, Methods). This finding suggests that ASM may be more prevalent for specific
262 motif sequences distinct from those typically seen in m6A modified sites. In summary, we
263 uncovered differential m6A modification of alleles that may depend on genetic differences that
264 are proximal to the DRACH motif as well as ASM sites which have no nearby genetic differences.

265

266 **Validation of ASM sites through orthogonal approaches**

267 We next assessed the robustness of ASM detection by visualizing read pileups and conducting
268 an orthogonal experimental method. Computational approaches to detect m6A modifications from
269 direct RNA sequencing have been developed to leverage the increased propensity of base-calling
270 errors around modified bases (Liu et al. 2019). To further validate sites that exhibited ASM using
271 the supervised machine learning framework, we assessed the characteristic increase in errors
272 surrounding each candidate site. We visualized sequencing reads that overlap ASM sites,
273 enabling us to verify the expected enrichment of base-calling errors around sites with a higher
274 modification ratio. The phenomenon was observed consistently across both replicates,
275 characterized by a correspondence between base-calling errors and modification ratios
276 (Supplemental Fig. 6).

277

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

283

284 Among the 23 ASM sites, we detected 20 in the MeRIP-seq. Only five out of 20 sites, which
285 contain SNPs within or nearby the DRACH motif had sufficient coverage in our MeRIP-seq data
286 (Supplemental Table 5, Method). The allele bias ratio measured from MeRIP-seq in these five
287 sites demonstrated consistency with the allele bias detected by our approach (Supplemental Fig.
288 7). For example, *Atp5o* (Group 1 ASM) displayed allelic bias consistent with expectation in all
289 three MeRIP-seq replicates (Fig. 4F). Another Group 2 ASM site, *Gcsh* (SNP at position 1,253,
290 42 bp downstream of methylation site), exhibited the same allele-bias pattern in both long-read

291 sequencing and MeRIP-seq data (Fig. 4G). In short, while MeRIP-seq cannot capture all ASM
292 sites detected by the long-read approach due to inherent limitations, we observed consistent allele
293 bias in m6A patterns at five sites with sufficient read coverage.
294
295 The reliance of MeRIP-seq on short-read sequencing can lead to errors in allelic assignment,
296 primarily due to dependence on a limited number of SNPs, which increases susceptibility to
297 reference allele bias, genotyping errors, and systematic biases in library preparation. To assess
298 potential genotyping errors, we examined 37 SNP sites within ASM genes using Sanger
299 sequencing. Of these, 33 sites showed the expected genetic variants with strong peak signals,
300 however, four sites (*Apt5o*, 776; *Psrc1*, 977; *Trim25*, 5007, and 5041) displayed nucleotides from
301 only one allele. This finding suggests potential genotyping errors or limitations in our Sanger
302 sequencing experiments. Importantly, these findings further underscore the challenges of
303 accurate allelic detection especially for the short-read sequencing approach that rely on a one or
304 few SNPs (Supplemental Table 6).
305
306 **Applicability of ONT direct RNA sequencing to detect ASM sites in human cells**
307 The analytical and empirical workflow we developed to detect ASM sites is broadly applicable to
308 any cell type with known genetic information. Given that systematic replication is essential to
309 validate new approaches (Piccolo and Frampton 2016), we next replicated ASM detection using
310 a lymphoblastoid cell line derived from a human with a well characterized genome. Specifically,
311 we analyzed five replicates of ONT direct RNA sequencing data generated using the NA12878
312 cell line (Hansen 2016; Workman et al. 2018), assigned reads to their allelic origin, and quantified
313 m6A modifications for each group of reads (Fig. 5A, Supplemental Table 7).
314
315 Unlike hybrid mouse ESCs, a typical human harbors many fewer heterozygous SNPs per
316 transcript (Rozowsky et al. 2011; Workman et al. 2018). In our long-read sequencing analysis,

317 among 21,569 mouse transcripts, 16,242 contain at least one heterozygous SNP in hybrid
318 mESCs, while only 8,889 human transcripts contain such sites in NA12878. Notably, the
319 transcriptome overall harbors nearly ten times fewer heterozygous SNPs in NA12878 compared
320 to hybrid mESCs (210,004 in mESCs; 27,269 in NA12878; Supplemental Fig. 8). Hence, the
321 percentage of long-reads that can be assigned to their allelic origin with high confidence is
322 reduced (Supplemental Table 2, 4).
323
324 Despite having lower depth of sequencing and fewer informative SNPs per transcript, we
325 identified three ASM sites with reproducible and large effect sizes. These three sites were found
326 on the *BTN3A2* (FDR=0.006), *FCMR* (FDR=0.006), and *TNFSF9* (FDR=0.47) transcripts (Fig.
327 5B, Supplemental Fig. 9-10). *FCMR* contains a SNP on the R site of the DRACH motif. For the
328 other two cases the closest SNP to the ASM site was found 86 (at position 3,175 in *BTN3A2*
329 transcript) and 24 (at position 779 in *TNFSF9* transcript) nucleotides away from the methylation
330 site. All three events were observed within the 3' UTR and exhibited large differences in
331 modification ratio between the two alleles (Fig. 5C; mean difference in modification ratio 0.657,
332 *BTN3A2*; 0.500, *FCMR*; and 0.390, *TNFSF9*). *BTN3A2* plays a crucial role in T cell activation and
333 proliferation (Vantourout et al. 2018; Kabelitz and Dechanet-Merville 2016), *FCMR*, which
334 encodes the IgM Fc receptor, is vital for B cell activation and survival (Wang et al. 2016), and
335 *TNFSF9*, a member of the TNF superfamily, enhances T cell responses by interacting with CD137
336 on activated T lymphocytes (Wang et al. 2016; Hashimoto 2021). Given that the NA12878 cell
337 line is an Epstein-Barr virus (EBV)-transformed lymphoblastoid line, our findings reveal ASM sites
338 within three transcripts related within the immune system. The applicability of our ONT direct
339 RNA sequencing method for ASM detection in human cells supports the wide-ranging utility of
340 our approach in any system with known genetic information.
341

342 **RNA abundance is higher for the allele with higher m6A modification ratio**

343 The role of m6A RNA modification in transcription and translation has been extensively
344 investigated and remains a topic of debate (Akhtar et al. 2021; Jain et al. 2023; Meyer 2019b).
345 Allele-specific differences in m6A modification provides a powerful platform to assess their
346 functional impact on expression dynamics as the genetic background, environmental factors and
347 sample preparation are identical for the two alleles. Hence, we generated matched RNA-seq and
348 ribosome profiling data in hybrid mESCs and leveraged existing measurements for the NA12878
349 cells (Methods) (Cenik et al. 2015). This data enabled us to determine the relative RNA expression
350 and ribosome occupancy on each allele and correlate these with their m6A modification status.

351

352 In hybrid mESCs, transcripts harboring ASM sites demonstrated statistically significant RNA
353 expression bias towards the allele with higher m6A modification. This pattern was consistent
354 across both long-read and short-read sequencing methods (Fig. 6A; Binomial test p-value 0.004
355 and 0.011, respectively). Specifically, the mean proportion of RNA reads from the CAST allele for
356 genes exhibiting CAST-biased ASM were 0.557 and 0.558 for long-read and short-read
357 sequencing. Similarly, genes with B6-biased ASM had higher mean proportion of RNA reads from
358 the B6 allele (0.460 and 0.398, respectively). These observations suggest that ASM is associated
359 with allele-specific expression in the same allelic direction.

360

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

367

368 Recent studies have postulated the role of m6A modification in regulating translation (Mao et al.
369 2019; Meyer 2019b; Jain et al. 2023). In light of this, we employed ribosome profiling to assess
370 the impact of ASM on allele-specific translation. In particular, we analyzed allele-specific ribosome
371 occupancy on genes with ASM in hybrid mESCs. We did not detect significant correlation between
372 allele-specific ribosome occupancy and ASM (Fig. 6A; p-value, 0.83; Supplemental Fig. 11C).
373 Collectively, our results indicate that alleles with higher m6A modification ratios are associated
374 with increased RNA abundance but similar ribosome occupancy.
375
376 **DISCUSSION**
377
378 In this study, we used ONT direct RNA sequencing as a new method to detect allele-specific m6A
379 RNA modifications in both human and mouse cells. Notably, the long-reads generated by the
380 ONT approach revealed ASM sites with no nearby genetic differences, suggesting that m6A
381 modification on a given site may depend on factors beyond local sequence context. These ASM
382 may potentially be governed from long-range interactions that modulate mRNA secondary
383 structure, differences in allele-specific interactions with RNA-binding proteins or the local
384 chromatin context of each allele (Huang et al. 2019; Deng et al. 2022; Berlivet et al. 2019).
385
386 The deposition of m6A modification on mRNA is dependent on the presence of a specific motif
387 (DR**A**CH) surrounding the adenosine that is modified (Linder et al. 2015). Our analysis of ASM
388 sites revealed that nucleotide identity of the positions that surround this canonical motif may also
389 influence m6A deposition in particular contexts. Specifically, we found that alleles containing
390 cytosine at the D-1 site followed by uracil at the D-2 site exhibit higher methylation levels
391 (Supplemental Fig. 6).
392

393 A major strength of our approach over short-read based methods is its ability to cover many more
394 informative SNPs to assign reads to their allelic origin (Supplemental Fig. 12, maximum SNP
395 count per read, 12 with short-read; 78 with long-read in mESC). Hence, long-read technology has
396 higher confidence in allelic assignment (Wu et al. 2023; Glinos et al. 2022). In contrast to short-
397 read methods, which rely heavily on single SNPs within a read and are thus more susceptible to
398 errors from misannotated variants, our approach minimizes the impact of incorrect or missing
399 genetic variant annotations. Furthermore, in samples with less genetic variation, long-reads
400 increase the chance of linking genetic variants that may be far away from the site of interest which
401 would not be detectable by short-read based approaches.
402
403 A recent study leveraged previously generated MeRIP-seq data and claimed to detect numerous
404 ASM sites (Cao et al. 2023). Their approach involved calculating p-values from Fisher's exact test
405 on tables of reads per kilobase of transcript per million mapped reads for each allele from the
406 input control and IP. They interpreted the resulting p-values as evidence of ASM. However, this
407 method is fundamentally flawed. Fisher's exact test is specifically designed for categorical data,
408 particularly in 2x2 contingency tables, and applying it to continuous data in this context is
409 inappropriate. This misuse of the test raises serious concerns about the validity of their
410 conclusions.
411
412 Furthermore, MeRIP-seq suffers from the additional limitations of antibody-based enrichment.
413 Antibody-based approaches introduce specificity artifacts which result in variability in the number
414 and location of peaks detected across experiments (Helm et al. 2019). Similarly, the
415 immunoprecipitation step creates variable yields, limiting quantitative measurements among
416 experiments (McIntyre et al. 2020). Therefore, the large number of sites reported by Cao et al.
417 are inflated with a large number of false positives. In our study, we focused on large effect
418 differences using a bootstrap resampling strategy and minimum effect size threshold to reduce

statistical artifacts as previously recommended (Castel et al. 2020; Mohammadi et al. 2017). 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 m6A 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 m6A sites within DRACH motifs ending in ACA, constituting approximately 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 m6A 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; Cesaro et al. 2023; Min et al. 2018) Specifically, m6A reader proteins such as YTHDC1 and YTHDC2, which interact with m6A sites on 3' UTRs, are recognized for their role in enhancing mRNA stability and, consequently, increasing RNA abundance at steady-state (Lee et al. 2021; Wang et al. 2014). 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 m6A reader proteins that subsequently stabilize m6A-enriched mRNAs.

In contrast, we did not observe an association between ASM and allele-specific ribosome occupancy. Given that ribosome occupancy is a composite measurement of RNA expression and translation efficiency, this observation may indicate that alleles with higher modification ratios are less efficiently translated despite having higher steady-state RNA abundance. Such a mechanism

444 would be in agreement with a previously proposed model of coupling between co-transcriptional
445 m6A deposition and translation (Slobodin et al. 2017).
446
447 Our method has several important limitations. First, the supervised machine learning framework
448 we adopted is predicated on the assumption that modifications occur exclusively within DRACH
449 motifs. Consequently, our analysis does not account for genetic variations that alter the motif into
450 sequences not matching the DRACH pattern, which are presumed to result in methylation loss.
451 Second, the limitation in the number of reads generated by ONT direct RNA sequencing
452 constrains our method's ability to detect ASM sites in lowly expressed transcripts. Hence, ASM
453 sites identified in this study occur in genes within the top 30th percentile of RNA expression
454 (Supplemental Fig. 4).
455
456 In summary, we present a novel method for identifying allele-specific m6A modification using ONT
457 direct RNA sequencing. Our analyses emphasize the benefits of long-read sequencing and direct
458 detection of RNA modifications for ASM analysis. Future ASM studies are likely to extend the
459 catalog of allelic variants that influence RNA modifications, and characterize the mechanisms
460 leading to ASM and its functional consequences on gene expression.
461
462 **METHODS**
463
464 **Cell culture**
465 The C57BL/6J-CAST/EiJ F1 Hybrid mESCs were generously provided by Dr. David Spector
466 (Balasooriya and Spector 2022). Cells were cultured in 2i medium, comprising Knockout DMEM
467 (Gibco, Cat. No. 10829-018) supplemented with 15% Fetal Bovine Serum (FBS) (Millipore, Cat.
468 No. ES-009-B), 1X Glutamine (Millipore, Cat. No. TMS-002-C), 1X non-essential amino acids
469 (Millipore, Cat. No. TMS-001-C), 0.15 mM 2-Mercaptoethanol (Millipore, Cat. No. ES-007-E),

470 100 U/ml Penicillin-Streptomycin (Gibco, Cat. No.15140-122), 100 U/ml Lif (Millipore, Cat. No.
471 ESG 1107), 1 μM PD0325901 (Sigma Aldrich, Cat. No. 444968), and 3 μM CHIR99021 (Sigma
472 Aldrich, Cat. No. 361571). The culture plates were coated with 0.1% gelatin (Millipore, ES-006-
473 B). The cells were cultured at 37 °C under 5% CO_2 and passaged at a 70-80% sub-confluent
474 state.
475
476 **Generation of *Mettl3* knockout mESCs**
477 *Mettl3* knockout cells were generated by introducing Cas9/sgRNA ribonucleoprotein (RNP)
478 complexes into mESCs via nucleofection (Kirton et al. 2013). The sgRNA was synthesized by
479 Synthego (Supplemental Table 2). To form the RNP, 300 pmol of Cas9 protein (NEB, Cat. No.
480 M0386M) and 600 pmol of sgRNA were incubated in Cas9 Buffer (150 mM KCl, 1 mM $MgCl_2$,
481 10% v/v Glycerol, 20 mM HEPES–KOH [pH 7.5]) at room temperature for 30 minutes.
482 Subsequently, 65 μL of 4D-Nucleofector X Solution was added to the RNP solution. Nucleofection
483 was performed using the optimized protocol recommended by the manufacturer (SF Cell Line 4D-
484 NucleofectorTM X Kit L). A cell pellet was collected from 2 x 10^6 cells, resuspended in the RNP
485 solution, and transferred into a 100 μL Nucleocuvette Vessel. Electroporation was carried out
486 using the 4D-Nucleofector X Unit (Lonza) with the FF120 program. Post-nucleofection, cells were
487 equilibrated at room temperature for 8 minutes, then transferred to a gelatin coated culture dish
488 containing prewarmed 2i media. The cells were allowed to recover at 37 °C for 72 hours, followed
489 by the isolation of single clones using serial dilution. The genomic DNA was isolated from cells
490 grown from single clones and mutations were confirmed using the primers listed in Supplemental
491 Table 2.
492
493 **Oxford Nanopore direct RNA sequencing**
494 mESCs were grown in a 10 mm petri dish and collected from two different numbers of passages
495 on separate days, considered as two biological replicates. The cells were lysed in TRIzol reagent

496 (Zymo Research, Cat. No. R2050) and RNA was extracted according to the manufacturer's
497 instructions (Zymo Direct-zol RNA Kits, Cat. No. R2061). 5 µg of total RNA without poly(A) RNA
498 isolation was used for direct RNA sequencing (Viscardi and Arribere 2022). The library was
499 generated using the Oxford Nanopore Direct RNA Sequencing Kit (Nanopore Cat. No. SQK-
500 RNA002) following the manufacturer's protocol. The RNA sequencing from each RNA replicate
501 was performed on four MinION MkIb with R9.4 flow cells (Oxford Nanopore Technologies Ltd.)
502 with a 24-h runtime for each run.
503
504 **Human long-read data for method validation**
505 We utilized published ONT direct RNA sequencing data from the human cell line NA12878
506 downloaded from https://github.com/nanopore-wgs-consortium/NA12878 (Hansen 2016;
507 Workman et al. 2018). Five replicates were generated using RNA obtained from different institutes
508 (JHU, Johns Hopkins University; OICR, Ontario Institute for Cancer Research; UBC, University
509 of British Columbia; UCSC, University of California Santa Cruz; UN, University of Nottingham).
510 Raw fast5 files were used directly for analysis (558,005 reads from JHU; 1,226,344 reads from
511 OICR; 2,073,885 reads from UBC; 2,059,045 reads from UCSC; 1,686,124 reads from UN).
512
513 **m6A detection from ONT direct RNA sequencing**
514 Following sequencing, we used Guppy v. 6.3.2 (quality score cutoff = 7) for base-calling from
515 fast5 files and verified error rates with Pomoxis v0.3.15 (Buttler and Drown 2022). Reads were
516 aligned to the transcriptome with minimap 2.1 (minimap2 '-ax map-ont). To reduce alignment
517 biases, we used a transcriptome reference in which SNPs were masked with Ns as previously
518 described (Ozadam et al. 2023) A mouse VCF was downloaded from the Mouse Genome Project
519 (https://www.mousegenomes.org/), and the NA12878 VCF file was obtained from
520 https://hgdownload.soe.ucsc.edu/gbdb/hg38/platinumGenomes/.
521

522 To identify m6A modifications, we first used Nanopolish v0.11.3 to generate an index with the '--
523 scale-events' and '--signal-index' options, aligning events to the N-masked transcriptome
524 reference. Detection of m6A RNA modifications was conducted using m6Anet v-2.0.0 and a
525 pretrained model (Hct116_RNA002) (Hendra et al. 2022). A minimum of 20 reads per site was
526 required to call modification sites.
527
528 **Assignment of reads to their allelic origin**
529 To assign aligned reads to their allelic origin, we first identified the positions on each read that
530 correspond to a SNP, adjusting for any deletions and insertions in the read with respect to the
531 reference. Initially, we selected reads that intersect a predefined number of SNPs. The number
532 of heterozygous SNPs in mESC and NA12878 transcriptomes was 210,004, and 27,269,
533 respectively (Supplemental Fig. 8). Consequently, we required at least three SNPs per read for
534 mESCs, and one SNP for NA12878.
535
536 Then, we calculated the number of matches to each allele and defined a read-level allele-bias
537 ratio:
538 $$Allele\ A\ bias\ ratio\ = \frac{SNP\ count\ from\ Allele\ A\ in\ \ a\ read}{total\ number\ of\ detected\ SNPs\ in\ a\ read}$$
539
540 We assigned each read into one of three groups based on this ratio: allele A (bias ratio exceeding
541 0.7), allele B (bias ratio less than 0.3), and undefined (bias ratio between 0.3 and 0.7). The
542 process was implemented in a python script that is provided on Github: allele_assignment.py.
543 Subsequently, the three groups of reads underwent processing through m6Anet separately to
544 predict m6A probability and modification ratios as described above.
545
546 **Identification of allele-specific m6A modifications**

We first selected m6A sites with a high probability of modification (prob > 0.85) using all sequenced 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). We select the case in which SNPs on the motif do not make changes in the DRACH motif. In mESCs, 29 out of 178 in replicate 1 and 29 out of 145 in replicate 2 were further evaluated, as the SNP overlapping the motif led to different instances of the DRACH sequence (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), 2 in 22 (UN).

Given that m6Anet sets a threshold of 20 reads for determining modification sites and ratios, sites with fewer than 20 reads in one of the allelic groups are excluded during the detection phase. This results in a discrepancy between the number of reads assigned to alleles A, B, the undefined category, and the total count of reads. To correct this disparity and obtain an accurate modification ratio for these sites, we initially identified modification sites in transcripts that harbor at least one heterozygous SNP and at least 20 reads assigned to one of the alleles. If both alleles had more than 20 reads, the modification ratios were used directly as calculated by m6Anet (63% of total instance). However, when one allele has read counts less than 20, we recalculated the modification ratio leveraging modification information from all reads, without distinguishing the two alleles. Without loss of generality, let's 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:

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

572 If at least 10 reads were assigned to allele A, we retained this site for further analysis and
573 recalculated the modification ratio of allele A using the following formula:
574
575 $Allele\ A\ mod\ ratio$
576 $$= \frac{(Total\ read\ count\ *\ mod\ ratio\ from\ total\ reads)\ -\ (Allele\ B\ read\ count\ *\ Allele\ B\ mod\ ratio)}{Allele\ A\ read\ count}$$
577
578 See Supplemental Fig. 14 for a schematic description of this procedure.
579
580 We identified statistically significant ASM sites using a bootstrapping-based statistical test. First,
581 for each allele, methylated read counts were derived by multiplying modification ratios with total
582 read numbers. We then sampled the number of methylated reads for each allele with replacement
583 and calculated the difference between the modification ratios using the resampled read counts
584 (McLachlan and Rathnayake 2014; Banjanovic and Osborne 2016). This resampling procedure
585 was repeated 1,000,000 times and a one-sided p-value was calculated by using a effect size
586 threshold (*T*) of 0 or 0.1 as follows:
587 $$p\ value\ =\ 1 - \frac{|\{\ b: (Mod\ Ratio\ {}^{A}_{b}\ -\ Mod\ Ratio\ {}^{B}_{b}\ )\ >\ (T)\ \}|}{1{,}000{,}000}$$
588 where, $Mod\ Ratio\ {}^{A}_{b}$ represents the bootstrap resampling value for the allele with the higher
589 observed m6A modification ratio. An aggregate p-value was calculated by combining the p-values
590 from each replicate using the harmonic mean method (Wilson 2019). False Discovery Rate (FDR)
591 was calculated by the Benjamini & Hochberg method (Yoav Benjamini 1995). Finally, statistically
592 significant ASM sites were defined if adjusted harmonic mean p-values (FDR) were below 0.1.
593 For instance, with an effect size threshold of 0.1 (T = 0.1), if none of the randomizations exceed
594 this cutoff, it would suggest that the modification ratios of the two alleles from resampled reads

595 are highly similar. Consequently, the probability of this site being an ASM would be very low,
596 corresponding to a p-value of 1.
597
598 To assess the significance of UGACU being the most common DRACH variant among Group 2
599 ASM sites, we randomly resampled 17 motifs 10,000 times. For the resampling, we used the
600 observed frequency of each of the 15 instances of the DRACH motif among sites with a high
601 probability of modification. In these random samples, only 84 instances had UGACU as the most
602 frequent motif hence corresponding to a p-value of 0.0084.
603
604 **mESC MeRIP-seq experiments and analyses**
605 MeRIP-seq libraries were prepared with EpiNext CUT&RUN RNA m6A-Seq Kit (EpiGentek). The
606 three replicates of mESCs were collected from different numbers of passages on separate dates.
607 The total RNA was extracted with Direct-zol RNA Purification Kits (Zymo Research, Cat. No.
608 R2050). 7 µg of total RNA was subjected to immunoprecipitation with an m6A antibody (P9016,
609 EpiGentek, 1:100 dilution) and digested with cleavage enzyme on beads. The beads were then
610 washed three times with a wash buffer and protein digestion buffer, and RNA was eluted in 13 µl
611 of the elution buffer. The sequencing libraries were generated using the Diagenode small RNA
612 sequencing kit following the manufacturer's protocol (Diagenode, Cat. No. C05030001). The
613 libraries were sequenced on a NovaSeq6000 system (Illumina).
614
615 Adaptor sequences were trimmed from raw reads with cutadapt v4.7 (Martin 2011) using following
616 parameters: -a `AAAAAAAAAACAAAAAAAAAA` -G `^TTTTTTTTTGTTTTTTTTTT` -A
617 `AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT` -n 2 --overlap=4 --trimmed-only --maximum-
618 length=150 --minimum-length=31 --quality-cutoff=28. Trimed reads were aligned to the N-masked
619 mouse transcriptome with STAR v2.7.10b (Dobin et al. 2013). Reads with low mapping quality

620 were discarded (mapping quality less than 10) and indexed with samtools v1.15.1 (Danecek et
621 al. 2021; Bonfield et al. 2021).
622
623 To compute the allele bias ratio, we counted the number of allelic reads that harbor at least one
624 SNP within 100 bp of the ASM sites. Out of 23 ASM sites, four did not have genetic differences
625 within 100 base pairs of the methylated position and 15 had fewer than 40 reads across the three
626 replicates (Supplemental Table 5). Allele bias for the remaining five sites were calculated as:
627

628 $$Allele\ bias\ ratio\ = \frac{B6\ allele\ read\ count}{B6\ allele\ read\ count\ +\ CAST\ allele\ read\ count}$$

629
630 **mESC RNA-seq and ribosome profiling library preparation**
631 Five million mESCs were scraped and transferred to 1.5 mL tubes in lysis buffer (20 mM Tris
632 HCl pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, 100 µg/mL Cycloheximide, 1% Triton-X).
633 All experiments were done in two replicates. Cells were lysed on ice by pipetting up and down
634 ~5 times every 5 min for a total of 10 min. The lysate was clarified by centrifugation at 1300 x g
635 for 10 min at 4 °C. Ten percent of the clarified lysate by volume was reserved for RNA
636 extraction using Direct-zol RNA Purification Kits (Zymo Research, Cat. No. R2050). The RNA-
637 seq libraries were prepared using the NEBNext Ultra™ II RNA Library Prep Kit for Illumina
638 following manufacturer's protocol by Novagene.
639
640 The rest of the supernatant was immediately processed for ribosome profiling. Briefly, 7 µL of
641 RNaseI (Invitrogen, Cat. No. AM2249) was added to the clarified lysates and digestion was
642 carried out for 1 h at 4 °C. The digestion was stopped with ribonucleoside vanadyl complex (NEB,
643 Cat. No. S1402S) at a final concentration of 20 mM. Digested lysates were layered on a sucrose
644 cushion (20 mM Tris HCl pH 7.4, 150 mM NaCl, 5 mM $MgCl_2$, 34% sucrose, 1 mM DTT) and the

645 ribosomes were pelleted by centrifugation in a SW 41 Ti rotor (Beckman Coulter) at 38,000 rpm
646 for 2.5 h at 4°C. RNA was isolated with the RNeasy Mini RNA Kit (Qiagen, Cat. No. 74104) and
647 size-selected by running 5 µg of each sample on a 15% polyacrylamide TBE-UREA gel. The 21–
648 34 nt RNA fragments were excised and extracted by crushing the gel fragment in 400 µL of RNA
649 extraction buffer (300 mM sodium acetate pH 5.5, 1 mM EDTA, 0.25% SDS) followed by a 30 min
650 incubation on ice and an overnight incubation at room temperature. The sample was passed
651 through a Spin X filter (Corning, Cat. No. 8160) and the flowthrough was ethanol precipitated in
652 the presence of 5 mM $MgCl_2$ and 1 µL GlycoBlue (Invitrogen, Cat. No. AM9516). The RNA pellet
653 was resuspended in 10 µL of RNase-free water and immediately processed for library preparation.
654
655 For ribosome profiling library preparation, the D-Plex Small RNA-seq kit (Diagenode, Cat. No.
656 C05030001) was used with slight modifications. The dephosphorylation reaction was
657 supplemented with 0.5 µl T4 PNK (NEB, Cat. No. M0201S), and the reaction was incubated for
658 25 minutes. Subsequently, the complementary DNA (cDNA) was amplified for 12 PCR cycles.
659 We used AMPure XP bead cleanup (1.8X), followed by size selection using 3% agarose, dye-free
660 gel cassettes with internal standards (Sage Science, Cat. No. BDQ3010) on the BluePippin
661 platform. Sequencing was performed on a Novaseq 6000 platform.
662
663 **Read processing of RNA-seq and ribosome profiling**
664 For mESC, RNA-seq and Ribo-seq data were processed using RiboFlow v0.0.1 (Ozadam et al.
665 2020). For the Ribo-seq library, Unique Molecular Identifier (UMI) sequences were isolated
666 employing the following parameters: "umi_tools extract -p
667 '^(?P<umi_1>.{12})(?P<discard_1>.{4}).+$' --extract-method=regex". Subsequently, reads
668 underwent clipping with the parameters "-a `AAAAAAAAAACAAAAAAAAAA` --overlap=4 --trimmed-
669 only". Trimmed reads were then filtered by alignment to mouse rRNA and tRNA sequences with
670 bowtie2 version 7.3.0 and utilizing unaligned reads for subsequent alignment to the N-masked

transcriptome (Langmead and Salzberg 2012). Following transcriptome alignment, reads with mapping quality greater than two were preserved and deduplicated utilizing UMI-tools directional adjacency method with the parameter “--read-length” (Smith et al. 2017).

In mESC RNA-seq analysis, we clipped NEB Read adaptors using cutadapt, v1.18 (Martin 2011) with following parameters: “-a `AGATCGGAAGAGCACACGTCTGAACTCCAGTCA` -A `AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT` -O 8 -m 20 --cores=8”. The reads were aligned to the N-masked transcriptome using bowtie2 (Langmead and Salzberg 2012). The read count for RNA-seq and Ribo-seq were obtained from .ribo files with RiboR (Ozadam et al. 2020).

For the NA12878 sample, we analyzed RNA-seq and ribosome profiling data (NCBI Gene Expression Omnibus (GEO) under accession number GSE65912) based on the study by Cenik et al. (Cenik et al. 2015). We processed reads from both methods using cutadapt v1.18 with the parameters "-a `AGATCGGAAGAGCACACGTCTGAACTCCAGTCA` -A `AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGT` -O 8 -m 20" for trimming. The trimmed reads were then filtered by aligning to human rRNA and tRNA sequences with bowtie2 v7.3.0. Reads that did not align were subsequently mapped to the N-masked human transcriptome.

Allele-specific RNA expression and ribosome occupancy analysis

Utilizing the aligned BAM files obtained from RNA-seq and Ribo-seq, ASE counts were acquired using GATK (version 3.8.1) 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).

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

699 $$Allele\ A\ bias\ ratio\ = \frac{Allele\ A\ read\ count}{Allele\ A\ read\ count\ +\ Allele\ B\ read\ count}$$

700

701 To quantify the relationship between the allele-bias ratio from long-read sequencing and short-
702 read sequencing, we calculate the weighted Spearman correlation using long-read sequencing
703 read counts as weights. The correlation was calculated by using R package "boot", version 1.3-
704 30.

705

706 **Genomic DNA extraction and PCR for genetic variant verification**

707 Genomic DNA from mESCs was extracted using the Quick-DNA Miniprep Plus Kit (Zymo, Cat.
708 No. D4068) following the manufacturer's protocol. 17 primer pairs were designed to detect genetic
709 variants at genomic ASM and SNP sites (Supplemental Table 8). The target regions were
710 amplified by PCR using Q5 High-Fidelity DNA Polymerase (NEB, Cat. No. M0491S). The thermal
711 cycling conditions were set as follows: initial denaturation at 95°C for 90 seconds, followed by 32
712 cycles of denaturation at 95°C for 10 seconds, annealing at the (primer melting temperature -
713 2°C) for 15 seconds, extension at 72°C for 20-40 seconds, and a final extension at 72°C for 5
714 minutes. The resulting PCR products were purified using the NucleoSpin Gel and PCR Clean-up
715 Kit (Takara, Cat. No. 740609.250) and sequenced by Sanger sequencing (ACGT, Inc. DNA
716 sequencing service).

717

718 **DATA ACCESS**

719 All mESC short-read sequencing data sets presented in this paper have been deposited in the
720 Sequence Read Archive under BioProject accession number PRJNA1071025 (SRP486746

). The ONT direct RNA sequencing data is available on Zenodo under the following record numbers: mESC replicate 1 (10815502, 13255832, 13256383), mESC replicate 2 (13257639, 13259594, 13273847, 13275906, 13278114, 13277067), and Mettl3 knockout cells (13257082). All custom scripts used to perform bioinformatics analyses available on Github: https://github.com/DayeaPark/Allele-specific-m6A-modification.git

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 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.

735 **REFERENCES**

736 Akhtar J, Lugoboni M, Junion G. 2021. m6A RNA modification in transcription regulation.
737 *Transcription* **12**: 266–276.

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

741 Balasooriya GI, Spector DL. 2022. Allele-specific differential regulation of monoallelically
742 expressed autosomal genes in the cardiac lineage. *Nat Commun* **13**: 5984.

743 Banjanovic ES, Osborne JW. 2016. Confidence intervals for effect sizes: Applying bootstrap
744 resampling. *Practical Assessment, Research, and Evaluation*.
745 https://openpublishing.library.umass.edu/pare/article/id/1604/.

746 Benton MC, Lea RA, Macartney-Coxson D, Sutherland HG, White N, Kennedy D, Mengersen K,
747 Haupt LM, Griffiths LR. 2019. Genome-wide allele-specific methylation is enriched at gene
748 regulatory regions in a multi-generation pedigree from the Norfolk Island isolate.
749 *Epigenetics Chromatin* **12**: 60.

750 Berlivet S, Scutenaire J, Deragon J-M, Bousquet-Antonelli C. 2019. Readers of the mA
751 epitranscriptomic code. *Biochim Biophys Acta Gene Regul Mech* **1862**: 329–342.

752 Bokar JA, Rath-Shambaugh ME, Ludwiczak R, Narayan P, Rottman F. 1994. Characterization
753 and partial purification of mRNA N6-adenosine methyltransferase from HeLa cell nuclei.
754 Internal mRNA methylation requires a multisubunit complex. *J Biol Chem* **269**: 17697–
755 17704.

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

759 Buttler J, Drown DM. 2022. Accuracy and Completeness of Long Read Metagenomic
760 Assemblies. *Microorganisms* **11**. http://dx.doi.org/10.3390/microorganisms11010096.

761 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
762 RNA N 6-methyladenosine modifications reveal functional genetic variants in human
763 tissues. *Genome Res* **33**: 1369–1380.

764 Castel SE, Aguet F, Mohammadi P, GTEx Consortium, Ardlie KG, Lappalainen T. 2020. A vast
765 resource of allelic expression data spanning human tissues. *Genome Biol* **21**: 234.

766 Castel SE, Levy-Moonshine A, Mohammadi P, Banks E, Lappalainen T. 2015. Tools and best
767 practices for data processing in allelic expression analysis. *Genome Biol* **16**: 195.

768 Cenik C, Cenik ES, Byeon GW, Grubert F, Candille SI, Spacek D, Alsallakh B, Tilgner H, Araya
769 CL, Tang H, et al. 2015. Integrative analysis of RNA, translation, and protein levels reveals
770 distinct regulatory variation across humans. *Genome Res* **25**: 1610–1621.

771 Cenik C, Chua HN, Singh G, Akef A, Snyder MP, Palazzo AF, Moore MJ, Roth FP. 2017. A
772 common class of transcripts with 5'-intron depletion, distinct early coding sequence

773 features, and N1-methyladenosine modification. *RNA* **23**: 270–283.

774 Cesaro B, Tarullo M, Fatica A. 2023. Regulation of Gene Expression by m6Am RNA
775 Modification. *Int J Mol Sci* **24**. http://dx.doi.org/10.3390/ijms24032277.

776 Chen K, Luo G-Z, He C. 2015. High-Resolution Mapping of N^6-Methyladenosine in
777 Transcriptome and Genome Using a Photo-Crosslinking-Assisted Strategy. *Methods*
778 *Enzymol* **560**: 161–185.

779 Chen Z, Lei C, Wang C, Li N, Srivastava M, Tang M, Zhang H, Choi JM, Jung SY, Qin J, et al.
780 2019. Global phosphoproteomic analysis reveals ARMC10 as an AMPK substrate that
781 regulates mitochondrial dynamics. *Nat Commun* **10**: 104.

782 Cho H, Davis J, Li X, Smith KS, Battle A, Montgomery SB. 2014. High-resolution transcriptome
783 analysis with long-read RNA sequencing. *PLoS One* **9**: e108095.

784 Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T,
785 McCarthy SA, Davies RM, et al. 2021. Twelve years of SAMtools and BCFtools.
786 *Gigascience* **10**. http://dx.doi.org/10.1093/gigascience/giab008.

787 David J. Adams, Anthony G. Doran, Jingtao Lilue & Thomas M. Keane. 2015. The Mouse
788 Genomes Project: a repository of inbred laboratory mouse strain genomes. *Mamm Genome*
789 **26**: 403–412.

790 de la Chapelle A. 2009. Genetic predisposition to human disease: allele-specific expression and
791 low-penetrance regulatory loci. *Oncogene* **28**: 3345–3348.

792 Deng S, Zhang J, Su J, Zuo Z, Zeng L, Liu K, Zheng Y, Huang X, Bai R, Zhuang L, et al. 2022.
793 RNA mA regulates transcription via DNA demethylation and chromatin accessibility. *Nat*
794 *Genet* **54**: 1427–1437.

795 DeVeale B, van der Kooy D, Babak T. 2012. Critical Evaluation of Imprinted Gene Expression
796 by RNA–Seq: A New Perspective. *PLoS Genet* **8**: e1002600.

797 Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras
798 TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15–21.

799 Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S,
800 Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, et al. 2012. Topology of the human
801 and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**: 201–206.

802 Fan J, Hu J, Xue C, Zhang H, Susztak K, Reilly MP, Xiao R, Li M. 2020. ASEP: Gene-based
803 detection of allele-specific expression across individuals in a population by RNA
804 sequencing. *PLoS Genet* **16**: e1008786.

805 Fournier C, Goto Y, Ballestar E, Delaval K, Hever AM, Esteller M, Feil R. 2002. Allele-specific
806 histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO J* **21**:
807 6560–6570.

808 Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T,
809 James P, Warland A, et al. 2018. Highly parallel direct RNA sequencing on an array of
810 nanopores. *Nat Methods* **15**: 201–206.

811 Garcia-Campos MA, Edelheit S, Toth U, Safra M, Shachar R, Viukov S, Winkler R, Nir R,
812 Lasman L, Brandis A, et al. 2019. Deciphering the “m6A Code” via Antibody-Independent
813 Quantitative Profiling. *Cell* **178**: 731–747.e16.

814 Gicquel C, Rossignol S, Cabrol S, Houang M, Steunou V, Barbu V, Danton F, Thibaud N, Le
815 Merrer M, Burglen L, et al. 2005. Epimutation of the telomeric imprinting center region on
816 chromosome 11p15 in Silver-Russell syndrome. *Nat Genet* **37**: 1003–1007.

817 Glinos DA, Garborcauskas G, Hoffman P, Ehsan N, Jiang L, Gokden A, Dai X, Aguet F, Brown
818 KL, Garimella K, et al. 2022. Transcriptome variation in human tissues revealed by long-
819 read sequencing. *Nature* **608**: 353–359.

820 Hansen NF. 2016. Variant Calling From Next Generation Sequence Data. In *Statistical*
821 *Genomics: Methods and Protocols* (eds. E. Mathé and S. Davis), pp. 209–224, Springer
822 New York, New York, NY.

823 Hashimoto K. 2021. CD137 as an Attractive T Cell Co-Stimulatory Target in the TNFRSF for
824 Immuno-Oncology Drug Development. *Cancers* **13**.
825 http://dx.doi.org/10.3390/cancers13102288.

826 Helm M, Lyko F, Motorin Y. 2019. Limited antibody specificity compromises epitranscriptomic
827 analyses. *Nat Commun* **10**: 5669.

828 Hendra C, Pratanwanich PN, Wan YK, Goh WSS, Thiery A, Göke J. 2022. Detection of m6A
829 from direct RNA sequencing using a multiple instance learning framework. *Nat Methods* **19**:
830 1590–1598.

831 He PC, He C. 2021. m6A RNA methylation: from mechanisms to therapeutic potential. *EMBO J*
832 **40**: e105977.

833 He PC, Wei J, Dou X, Harada BT, Zhang Z, Ge R, Liu C, Zhang L-S, Yu X, Wang S, et al. 2023.
834 Exon architecture controls mRNA m6A suppression and gene expression. *Science* **379**:
835 677–682.

836 Huang H, Weng H, Zhou K, Wu T, Zhao BS, Sun M, Chen Z, Deng X, Xiao G, Auer F, et al.
837 2019. Histone H3 trimethylation at lysine 36 guides m6A RNA modification co-
838 transcriptionally. *Nature* **567**: 414–419.

839 Inoue A. 2023. Noncanonical imprinting: intergenerational epigenetic inheritance mediated by
840 Polycomb complexes. *Curr Opin Genet Dev* **78**: 102015.

841 Jain S, Koziej L, Poulis P, Kaczmarczyk I, Gaik M, Rawski M, Ranjan N, Glatt S, Rodnina MV.
842 2023. Modulation of translational decoding by m6A modification of mRNA. *Nat Commun* **14**:
843 4784.

844 Jiang X, Liu B, Nie Z, Duan L, Xiong Q, Jin Z, Yang C, Chen Y. 2021. The role of m6A
845 modification in the biological functions and diseases. *Signal Transduct Target Ther* **6**: 74.

846 Joglekar A, Prjibelski A, Mahfouz A, Collier P, Lin S, Schlusche AK, Marrocco J, Williams SR,
847 Haase B, Hayes A, et al. 2021. A spatially resolved brain region- and cell type-specific
848 isoform atlas of the postnatal mouse brain. *Nat Commun* **12**: 463.

849 Kabelitz D, Dechanet-Merville J. 2016. *Recent Advances in γδ T Cell Biology: New Ligands,*

850 *New Functions, and New Translational Perspectives*. Frontiers Media SA.

851 Kirton HM, Pettinger L, Gamper N. 2013. Transient Overexpression of Genes in Neurons Using
852 Nucleofection. In *Ion Channels: Methods and Protocols* (ed. N. Gamper), pp. 55–64,
853 Humana Press, Totowa, NJ.

854 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:
855 357–359.

856 Lee J-H, Wang R, Xiong F, Krakowiak J, Liao Z, Nguyen PT, Moroz-Omori EV, Shao J, Zhu X,
857 Bolt MJ, et al. 2021. Enhancer RNA m6A methylation facilitates transcriptional condensate
858 formation and gene activation. *Mol Cell* **81**: 3368–3385.e9.

859 Lee Y, Choe J, Park OH, Kim YK. 2020. Molecular Mechanisms Driving mRNA Degradation by
860 m6A Modification. *Trends Genet* **36**: 177–188.

861 Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-
862 nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nat*
863 *Methods* **12**: 767–772.

864 Lin S, Gregory RI. 2014. Methyltransferases modulate RNA stability in embryonic stem cells.
865 *Nat Cell Biol* **16**: 129–131.

866 Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, Schwartz S, Mattick JS, Smith
867 MA, Novoa EM. 2019. Accurate detection of m6A RNA modifications in native RNA
868 sequences. *Nat Commun* **10**: 4079.

869 Liu J, Yue Y, Han D, Wang X, Fu Y, Zhang L, Jia G, Yu M, Lu Z, Deng X, et al. 2013. A
870 METTL3–METTL14 complex mediates mammalian nuclear RNA N6-adenosine
871 methylation. *Nat Chem Biol* **10**: 93–95.

872 Liu Z, Dong X, Li Y. 2018. A Genome-Wide Study of Allele-Specific Expression in Colorectal
873 Cancer. *Front Genet* **9**: 570.

874 Mao Y, Dong L, Liu X-M, Guo J, Ma H, Shen B, Qian S-B. 2019. mA in mRNA coding regions
875 promotes translation via the RNA helicase-containing YTHDC2. *Nat Commun* **10**: 5332.

876 Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads.
877 *EMBnet journal*. http://journal.embnet.org/index.php/embnetjournal/article/view/200.

878 Mauer J, Luo X, Blanjoie A, Jiao X, Grozhik AV, Patil DP, Linder B, Pickering BF, Vasseur J-J,
879 Chen Q, et al. 2017. Reversible methylation of mA in the 5' cap controls mRNA stability.
880 *Nature* **541**: 371–375.

881 McIntyre ABR, Gokhale NS, Cerchietti L, Jaffrey SR, Horner SM, Mason CE. 2020. Limits in the
882 detection of m6A changes using MeRIP/m6A-seq. *Sci Rep* **10**: 6590.

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

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

888 Meyer KD. 2019a. DART-seq: an antibody-free method for global m6A detection. *Nat Methods*
889 **16**: 1275–1280.

890 Meyer KD. 2019b. m6A-mediated translation regulation. *Biochim Biophys Acta Gene Regul*
891 *Mech* **1862**: 301–309.

892 Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. 2012. Comprehensive
893 analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell*
894 **149**: 1635–1646.

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

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

900 Morcos L, Ge B, Koka V, Lam KCL, Pokholok DK, Gunderson KL, Montpetit A, Verlaan DJ,
901 Pastinen T. 2011. Genome-wide assessment of imprinted expression in human cells.
902 *Genome Biol* **12**: R25.

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

905 Ozadam H, Geng M, Cenik C. 2020. RiboFlow, RiboR and RiboPy: an ecosystem for analyzing
906 ribosome profiling data at read length resolution. *Bioinformatics* **36**: 2929–2931.

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

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

914 Pelizzola M, Baranov PV, Dassi E. 2021. *Computational Epitranscriptomics: Bioinformatic*
915 *Approaches for the Analysis of RNA Modifications*. Frontiers Media SA.

916 Piccolo SR, Frampton MB. 2016. Tools and techniques for computational reproducibility.
917 *Gigascience* **5**: 30.

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

921 Robles-Espinoza CD, Mohammadi P, Bonilla X, Gutierrez-Arcelus M. 2021. Allele-specific
922 expression: applications in cancer and technical considerations. *Curr Opin Genet Dev* **66**:
923 10–19.

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

927 Santini L, Halbritter F, Titz-Teixeira F, Suzuki T, Asami M, Ma X, Ramesmayer J, Lackner A,
928 Warr N, Pauler F, et al. 2021. Genomic imprinting in mouse blastocysts is predominantly
929 associated with H3K27me3. *Nat Commun* **12**: 3804.

930 Sergeeva A, Davydova K, Perenkov A, Vedunova M. 2023. Mechanisms of human DNA
931 methylation, alteration of methylation patterns in physiological processes and oncology.
932 *Gene* **875**: 147487.

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

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

940 Slobodin B, Han R, Calderone V, Vrielink JAFO, Loayza-Puch F, Elkon R, Agami R. 2017.
941 Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-
942 adenosine Methylation. *Cell* **169**: 326–337.e12.

943 Smith T, Heger A, Sudbery I. 2017. UMI-tools: modeling sequencing errors in Unique Molecular
944 Identifiers to improve quantification accuracy. *Genome Res* **27**: 491–499.

945 Song H, Song J, Cheng M, Zheng M, Wang T, Tian S, Flavell RA, Zhu S, Li H-B, Ding C, et al.
946 2021. METTL3-mediated mA RNA methylation promotes the anti-tumour immunity of
947 natural killer cells. *Nat Commun* **12**: 5522.

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

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

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

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

961 Viscardi MJ, Arribere JA. 2022. Poly(a) selection introduces bias and undue noise in direct
962 RNA-sequencing. *BMC Genomics* **23**: 530.

963 Wang H, Coligan JE, Morse HC 3rd. 2016. Emerging Functions of Natural IgM and Its Fc
964 Receptor FCMR in Immune Homeostasis. *Front Immunol* **7**: 99.

965 Wang S, Lv W, Li T, Zhang S, Wang H, Li X, Wang L, Ma D, Zang Y, Shen J, et al. 2022.

966 Dynamic regulation and functions of mRNA m6A modification. *Cancer Cell Int* **22**: 48.

967 Wang X, Lu Z, Gomez A, Hon GC, Yue Y, Han D, Fu Y, Parisien M, Dai Q, Jia G, et al. 2014.
968 N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**: 117–
969 120.

970 Wilson DJ. 2019. The harmonic mean *p*-value for combining dependent tests. *Proc Natl Acad*
971 *Sci U S A* **116**: 1195–1200.

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

976 Wu J, Hu W, Li S. 2023. Long-read transcriptome sequencing reveals allele-specific variants at
977 high resolution. *Trends Genet* **39**: 31–33.

978 Yang X, Triboulet R, Liu Q, Sendinc E, Gregory RI. 2022. Exon junction complex shapes the
979 m6A epitranscriptome. *Nat Commun* **13**: 7904.

980 Yoav Benjamini YH. 1995. Controlling the False Discovery Rate: A Practical and Powerful
981 Approach to Multiple Testing. *J R Stat Soc* **57**. https://www.jstor.org/stable/2346101.

982 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
983 al. 2021. Systematic calibration of epitranscriptomic maps using a synthetic modification-
984 free RNA library. *Nat Methods* **18**: 1213–1222.

985 Zitovsky JP, Love MI. 2019. Fast effect size shrinkage software for beta-binomial models of
986 allelic imbalance. *F1000Res* **8**: 2024.

987

988 **FIGURE LEGENDS**

989

990 **Figure. 1 | Allelic read assignment and m6A modification analysis using ONT direct RNA**
991 **sequencing in hybrid mESCs**

992 A) M6A modification ratio and locations detected from m6Anet using all reads (top, green, WT;
993 bottom, red, *Mettl3* knockout). The relative m6A locations within the transcript body were
994 determined. It presents modification ratios after high probability selection (> 0.85). The color
995 darkness represents the counts of the ratio on the position. B) Comparison of the frequencies of
996 instances of DRACH motif sequences (green, WT; red, *Mettl3* knockout). C) Schematic overview
997 of the strategy used for allelic long-read assignment for allele-specific m6A modification analysis.
998 Total RNA from hybrid mESC (C57BL/6J x CAST/EiJ) underwent ONT direct RNA sequencing.
999 To avoid reference bias, we used an N-masked transcriptome for alignment. Reads were then
1000 allocated to each allele. This process classified reads into Allele A (B6), Allele B (CAST), and
1001 undefined categories, enabling m6A detection within each group individually.

1002

1003 **Figure. 2 | Comparative analysis of allelic modifications in wild-type and *Mettl3* knockout**
1004 **mESCs**

1005 A) Allelic impartiality while allelic read assignment and m6A detection. The left pane outlines the
1006 schematic of the data procedural steps for allele-specific m6A modifications analysis. The right
1007 panels display the counts for allelic reads, candidate m6A sites, and high probability m6A sites
1008 selected through our criteria (red circles, B6; blue plusses, CAST; and gray triangles, undefined
1009 group). The left top two plots show the counts from mESC wild-type (WT) replicates and the
1010 bottom plot exhibits the numbers from mESC *Mettl3* knockout. B) Spearman correlation of
1011 modification ratio between alleles from wild-type and *Mettl3* knockout cells (rep1: mESC WT
1012 replicate 1; rep2: mESC WT replicate 2; *Mettl3*: *Mettl3* knockout). C-D) Distribution of sites with
1013 high probability of m6A modification (prob > 0.85) is displayed in a metagene plot by calculating

1014 the relative positions of these sites within gene regions. The color scale represents the number
1015 of m6A sites with the given modification ratio inferred from reads assigned to either of the two
1016 alleles in mESC wild-type (C) or *Mettl3* knockout (D) cells (red, B6 allele; blue, CAST allele).
1017

1018 **Figure. 3 | Identification and classification of ASM sites**

1019 A) Schematic of statistical procedure for ASM detection (Methods). Reads overlapping the site
1020 under consideration were resampled, and the modification ratio was estimated in each bootstrap
1021 sample. A statistically significant ASM site was defined as adjusted harmonic p-value (FDR < 0.1;
1022 Methods). B) An ASM site within *Trim25*, exhibits distinct modification ratio samples. Conversely,
1023 a non-ASM site within *Trim59* displays substantial overlap in modification ratios between
1024 bootstrap sampling distributions. C) Modification ratios of each allele across mESC wild-type
1025 replicates. Y-axis displays the name of the gene and m6A position in the transcript. D-E) ASM
1026 sites were classified into two groups. Group 1 is defined by genetic variants within the DRACH
1027 motif, and Group 2 is characterized by variants adjacent to or distal from the DRACH motif (D).
1028 The modification differences of the defined ASM were represented by color according to their
1029 classification (Group 1 in magenta, Group 2 in blue, and non-ASM in gray). Each axis is the
1030 modification ratio, where negative values denote CAST allele bias and positive values indicate
1031 B6 allele bias in m6A modification (E).
1032

1033 **Figure. 4 | Characterization of ASM sites and orthogonal detection with MeRIP-seq**

1034 A) SNP distribution in Group 1 ASM. B) Motif frequencies and modification ratios of motif
1035 sequences. The top bar plot illustrates motif sequence frequencies in all m6A instances, while the
1036 bottom heatmap indicates modification ratios. The first row presents modification ratio of all
1037 instances and the following six rows represents modification ratio on each motif sequence
1038 differentiated by SNPs from two alleles of Group 1 ASM sites. C) Information content of the
1039 extended DRACH motif in 17 Group 2 ASM sites shown in DNA sequence. The D-1 site has three

1040 SNPs, while the H+1 and H-2 sites each have one SNP. D) Extended motif sequences where the
1041 D-1 site possesses SNPs. The gray box represents the DRACH motif, in which all three genes
1042 share the same sequence (UGACU) followed by U on the D-2 site. E) Motif prevalence in Group
1043 2 ASM. The UGACU motifs are predominantly observed, contrasting with the common m6A
1044 motifs, which are typically represented by GGACU. F-G) Orthogonal detection of ASM through
1045 MeRIP-seq and long-read sequencing. On the top panel, points illustrate the allele-bias m6A ratio
1046 (proportion of reads from CAST allele) derived from three replicates of MeRIP-seq analysis. The
1047 Integrated Genome Viewer browser displays MeRIP-seq reads on SNPs adjacent to m6A sites,
1048 which correspond to the MeRIP-seq allele-bias ratio. The points in the bottom panel indicate the
1049 modification ratio of each allele from long-read sequencing, with gray color pairs representing
1050 data from two replicates. Two examples from *Apt5o*, Group 1 ASM (F) and *Gcsh*, Group 2 ASM
1051 (G) are shown.
1052
1053 **Figure. 5 | Reproducibility of ASM profiling procedure in human cells**
1054 A) Number of detected candidate m6A modification sites among five replicates (blue, Allele A;
1055 orange, Allele B; gray, undefined read classification). B) Modification ratios from each allele,
1056 including three ASM sites (blue) and non-ASM sites (gray) in UCSC (left) and UN (right) datasets,
1057 the highest depth datasets among five replicates. The X-axiswepresents the modification ratio of
1058 Allele A reads, while the Y-axis represents the modification ratio of Allele B reads. C) Resampled
1059 modification ratios from bootstrapping. Each color represents an allele (blue for Allele A, orange
1060 for Allele B), and the gray gradient indicates each replicate.
1061
1062 **Figure 6 | Effects of ASM on allele specific RNA expression and ribosome occupancy**
1063 A) Allele bias ratio of genes containing ASM sites (sky blue for B6 biased; pink for CAST biased).
1064 Different shapes represent the replicates (circle for replicate 1, triangle for replicate 2) The Y-axis
1065 displays the allele bias ratio obtained from long-read (left), short-read (middle) sequencing, and

1066 ribosome profiling (right). The X-axis shows the difference in m6A modification ratios between the
1067 two alleles (CAST - B6). The red dashed horizontal bar indicates allele-bias ratio 0.5, an allele-
1068 bias cutoff point. The gray bar represents the mean allele bias ratio for genes with B6 or CAST
1069 biased ASM sites. B) Model for regulation of ASM and allele-specific expression.

Figure 1

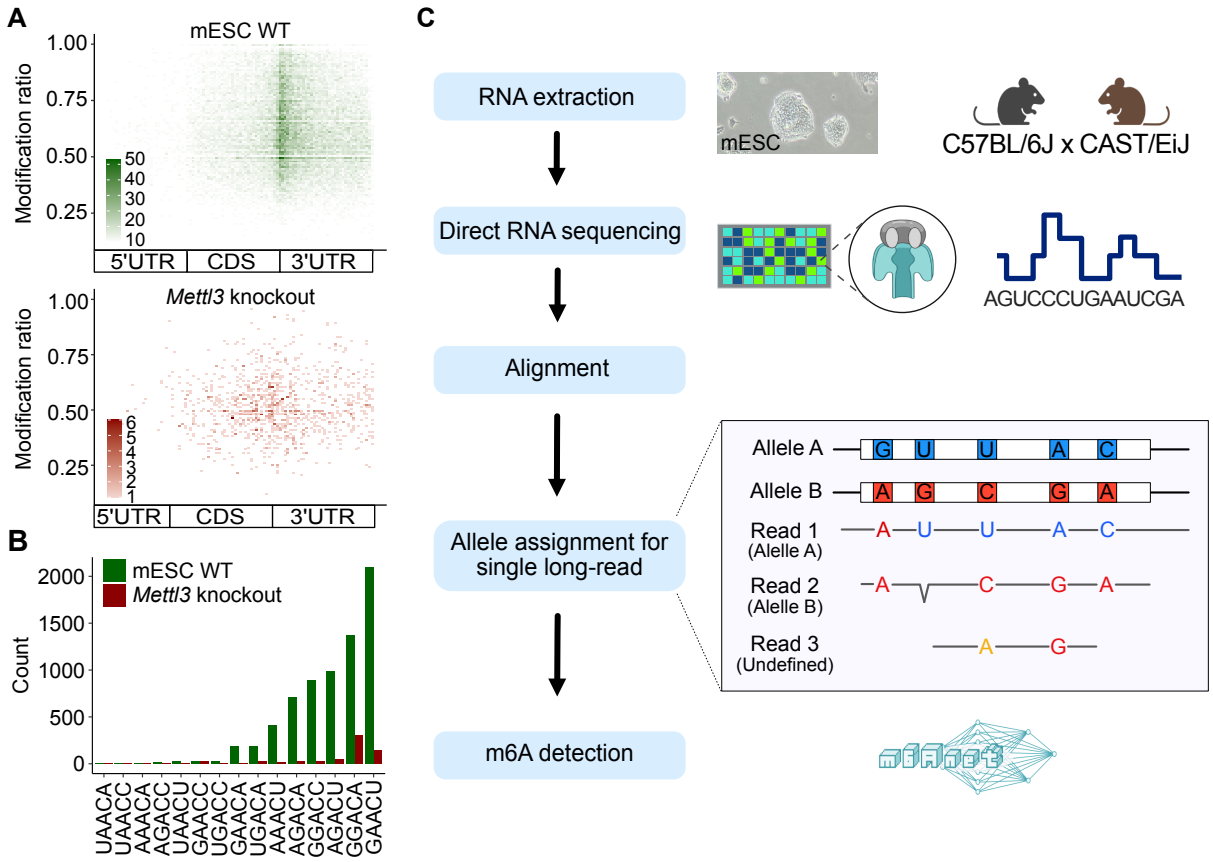

A
mESC WT
Modification ratio
1.00
0.75
0.50
0.25
50
40
30
20
10
5'UTR
CDS
3'UTR
Mettl3 knockout
6
5
4
3
2
1
B
mESC WT
Mettl3 knockout
Count
2000
1500
1000
500
0
UAACA
UAACC
AAACA
AGACC
UAACU
GAACC
UGACC
GAACA
UGACA
AAACU
AGACA
GGACC
AGACU
GGACA
GAACU
C
RNA extraction
mESC
C57BL/6J x CAST/EiJ
Direct RNA sequencing
AGUCCCUGAAUCGA
Alignment
Allele assignment for single long-read
Allele A
Allele B
Read 1 (Alelle A)
Read 2 (Alelle B)
Read 3 (Undefined)
m6A detection

Figure 2

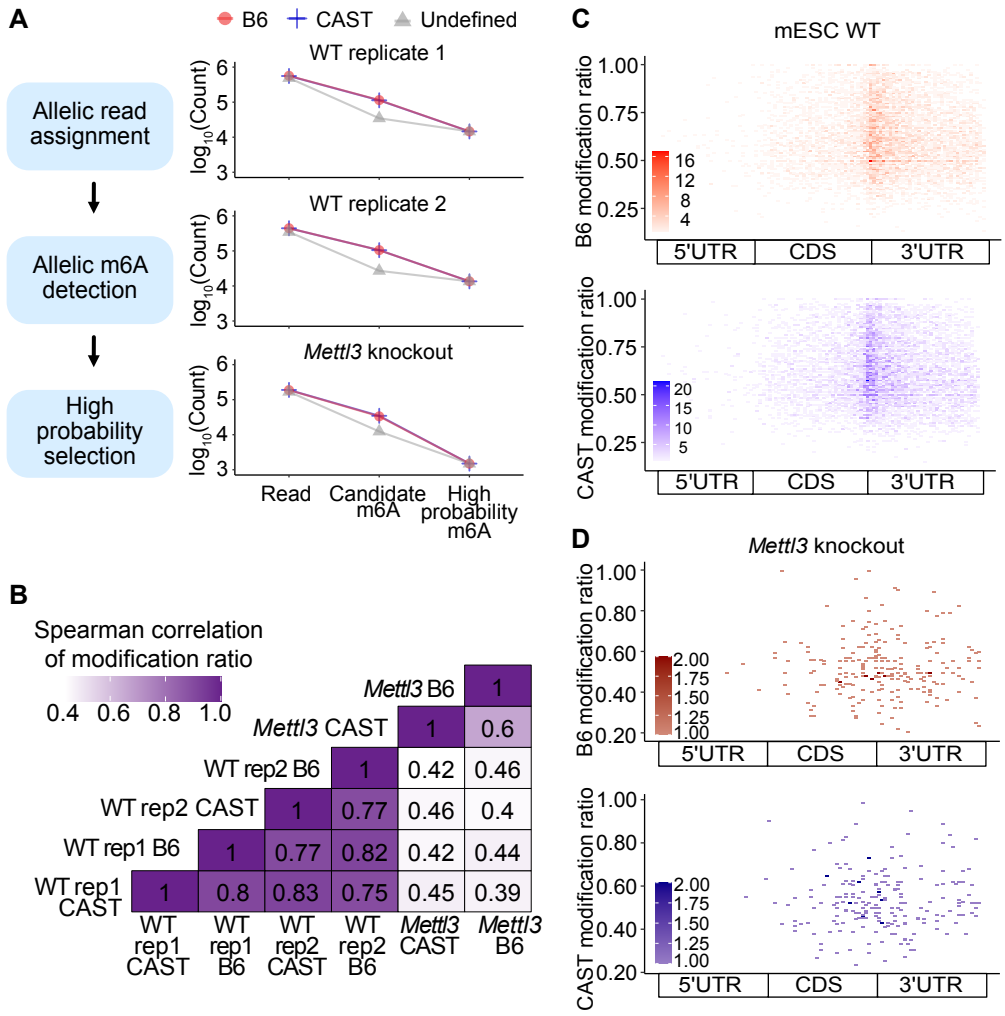

A
B6
CAST
Undefined
Allelic read assignment
Allelic m6A detection
High probability selection
WT replicate 1
WT replicate 2
Mettl3 knockout
log10(Count)
Read
Candidate m6A
High probability m6A
B
Spearman correlation of modification ratio
0.4 0.6 0.8 1.0
Mettl3 B6
Mettl3 CAST
WT rep2 B6
WT rep2 CAST
WT rep1 B6
WT rep1 CAST
WT rep1 CAST
WT rep1 B6
WT rep2 CAST
WT rep2 B6
Mettl3 CAST
Mettl3 B6
C
mESC WT
B6 modification ratio
CAST modification ratio
5'UTR
CDS
3'UTR
D
Mettl3 knockout
B6 modification ratio
CAST modification ratio
5'UTR
CDS
3'UTR

Figure 3

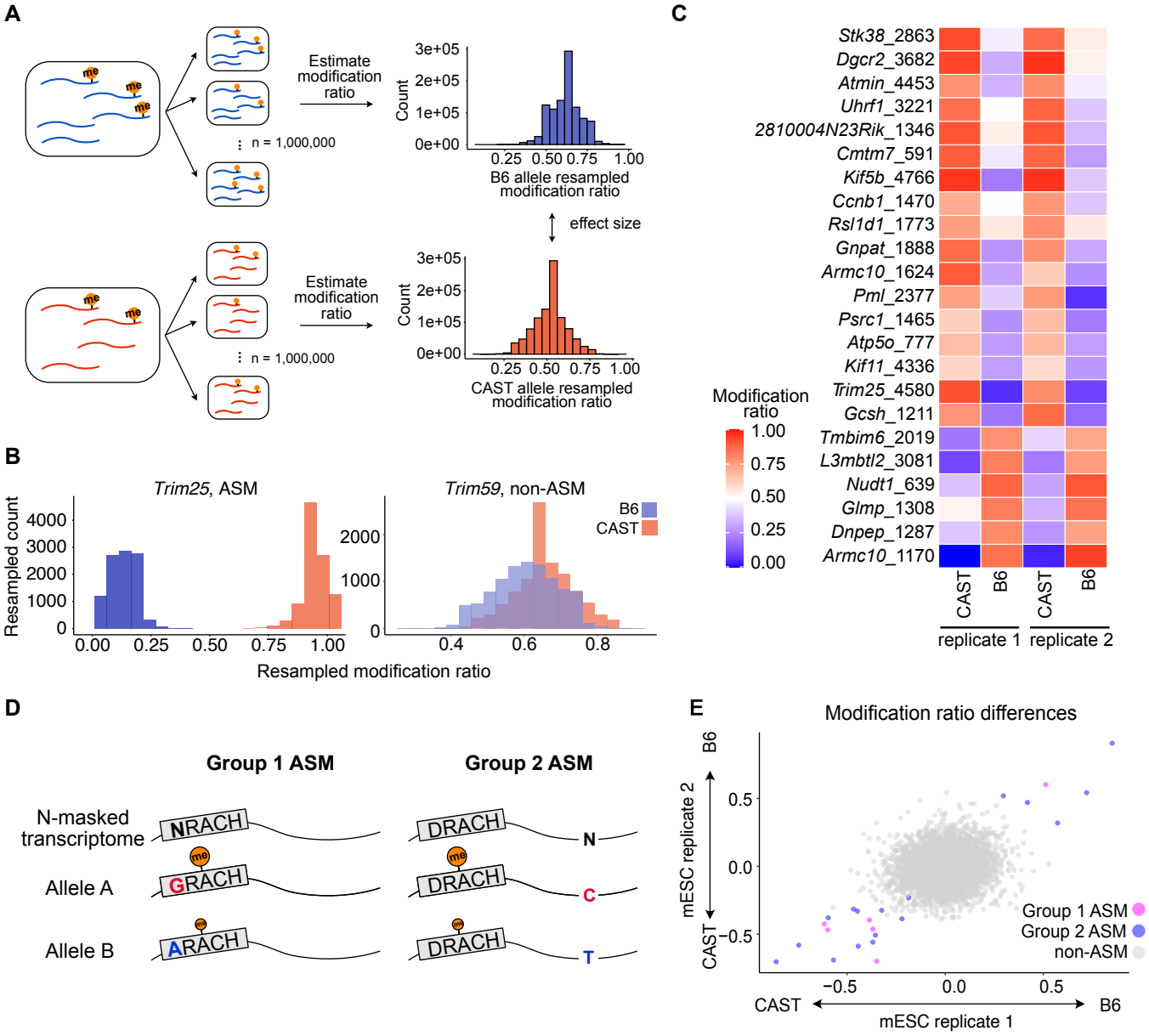

A
me
Estimate modification ratio
n = 1,000,000
Count
3e+05
2e+05
1e+05
0e+00
0.25 0.50 0.75 1.00
B6 allele resampled modification ratio
effect size
CAST allele resampled modification ratio
B
Trim25, ASM
Trim59, non-ASM
B6
CAST
Resampled count
4000
3000
2000
1000
0
0.00 0.25 0.50 0.75 1.00
0.4 0.6 0.8
Resampled modification ratio
C
Stk38_2863
Dgcr2_3682
Atmin_4453
Uhrf1_3221
2810004N23Rik_1346
Cmtm7_591
Kif5b_4766
Ccnb1_1470
Rsl1d1_1773
Gnpat_1888
Armc10_1624
Pml_2377
Psrc1_1465
Atp5o_777
Kif11_4336
Trim25_4580
Gcsh_1211
Tmbim6_2019
L3mbtl2_3081
Nudt1_639
Glmp_1308
Dnpep_1287
Armc10_1170
Modification ratio
1.00
0.75
0.50
0.25
0.00
CAST
B6
CAST
B6
replicate 1
replicate 2
D
Group 1 ASM
Group 2 ASM
N-masked transcriptome
Allele A
Allele B
NRACH
GRACH
ARACH
DRACH
DRACH
DRACH
N
C
T
E
Modification ratio differences
B6
CAST
mESC replicate 2
0.5
0.0
−0.5
Group 1 ASM
Group 2 ASM
non-ASM
−0.5
0.0
0.5
CAST
B6
mESC replicate 1

Figure 4

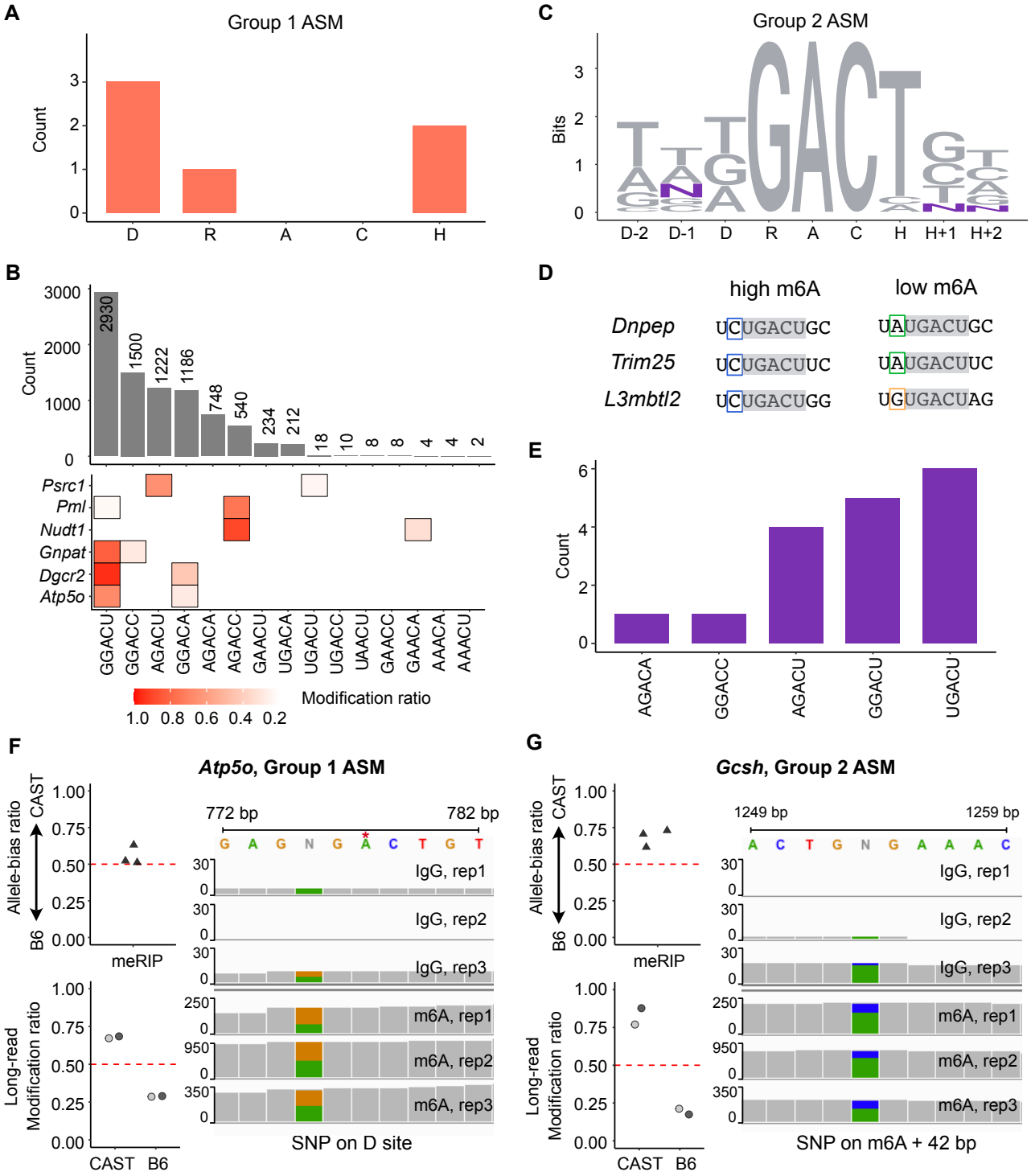

A
Group 1 ASM
Count
D
R
A
C
H
C
Group 2 ASM
Bits
D-2
D-1
D
R
A
C
H
H+1
H+2
B
Count
2930
1500
1222
1186
748
540
234
212
18
10
8
8
4
4
2
Psrc1
Pml
Nudt1
Gnpat
Dgcr2
Atp5o
GGACU
GGACC
AGACU
GGACA
AGACA
AGACC
GAACU
UGACA
UGACU
UGACC
UAACU
GAACC
GAACA
AAACA
AAACU
Modification ratio
1.0 0.8 0.6 0.4 0.2
D
high m6A
low m6A
Dnpep
UCUGACUGC
UAUGACUGC
Trim25
UCUGACUUC
UAUGACUUC
L3mbtl2
UCUGACUGG
UGUGACUAG
E
Count
AGACA
GGACC
AGACU
GGACU
UGACU
F
Atp5o, Group 1 ASM
Allele-bias ratio
CAST
B6
meRIP
Long-read Modification ratio
CAST
B6
772 bp
782 bp
G A G N G A C T G T
IgG, rep1
IgG, rep2
IgG, rep3
m6A, rep1
m6A, rep2
m6A, rep3
SNP on D site
G
Gcsh, Group 2 ASM
Allele-bias ratio
CAST
B6
meRIP
Long-read Modification ratio
CAST
B6
1249 bp
1259 bp
A C T G N G A A A C
IgG, rep1
IgG, rep2
IgG, rep3
m6A, rep1
m6A, rep2
m6A, rep3
SNP on m6A + 42 bp

Figure 5

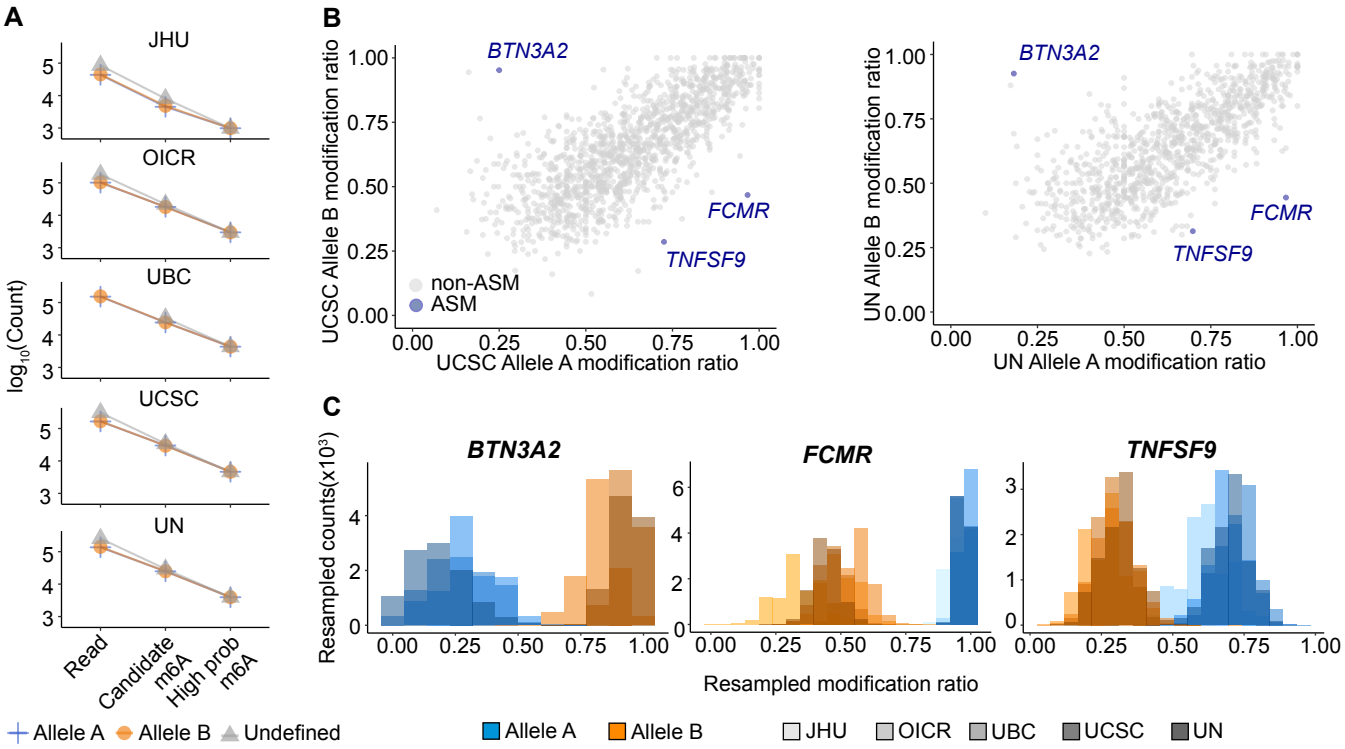
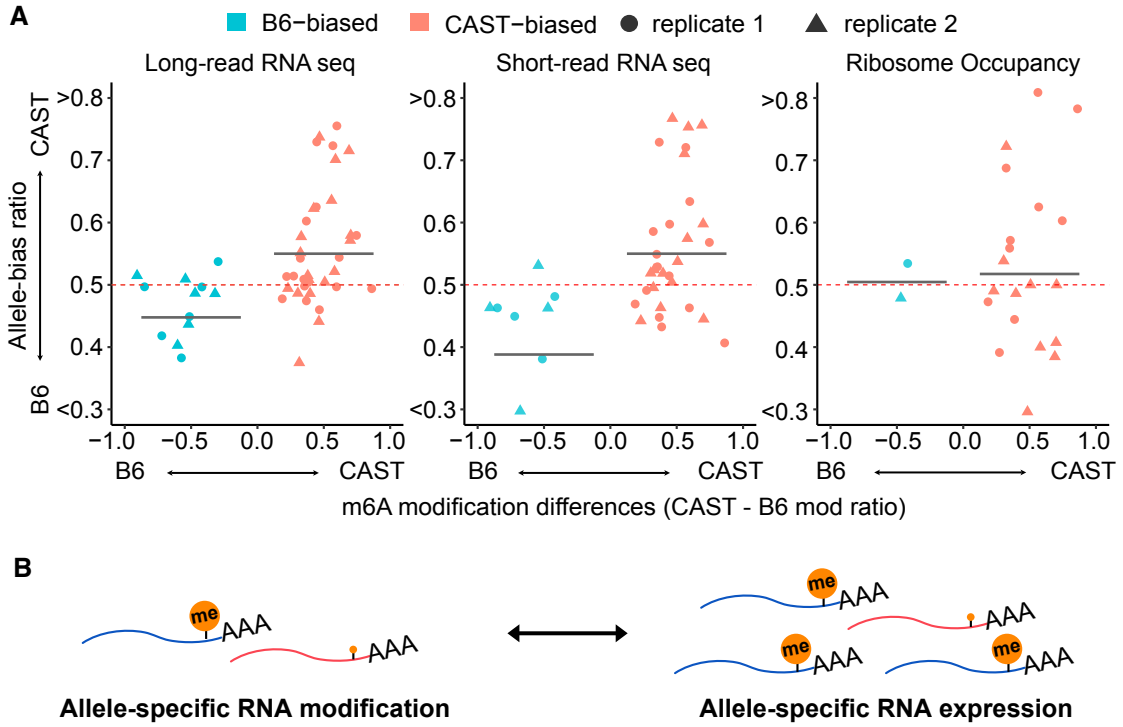

A
JHU
OICR
UBC
UCSC
UN
log10(Count)
Read
Candidate m6A
High prob m6A
Allele A
Allele B
Undefined
B
BTN3A2
FCMR
TNFSF9
non-ASM
ASM
UCSC Allele B modification ratio
UCSC Allele A modification ratio
UN Allele B modification ratio
UN Allele A modification ratio
C
BTN3A2
FCMR
TNFSF9
Resampled counts(x10³)
Resampled modification ratio
Allele A
Allele B
JHU
OICR
UBC
UCSC
UN

Figure 6


A
B6−biased
CAST−biased
replicate 1
replicate 2
Long-read RNA seq
Short-read RNA seq
Ribosome Occupancy
Allele-bias ratio
CAST
B6
>0.8
0.7
0.6
0.5
0.4
<0.3
−1.0
−0.5
0.0
0.5
1.0
B6
CAST
m6A modification differences (CAST - B6 mod ratio)
B
me
AAA
Allele-specific RNA modification
Allele-specific RNA expression