Calibrated analysis framework for nanopore direct RNA sequencing uncovers cell-specific m^{6}A stoichiometry at conserved sites

Authors

Denise Ohnezeit*[1], Elene Loliashvili*[2], Gregory Putzel[3], Ruth Verstraten[2,4], Jianheng Liu[5], Luke S. Nicholson[5], Alejandro Pironti[3], Samie R. Jaffrey[5], Daniel P. Depledge[#1,2,4,6], Angus C. Wilson[#1]

Affiliations

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

[2]Institute of Virology, Hannover Medical School, Hannover, Germany

[3]Antimicrobial-Resistant Pathogens Program, New York University Grossman School of Medicine, New York, NY, USA

[4]German Center for Infection Research (DZIF), partner site Hannover-Braunschweig, Hannover, Germany

[5]Department of Pharmacology, Weill Medical College, Cornell University, New York, NY, USA

[6]Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany

* Joint first authors: contributed equally

[#] Joint senior authors: contributed equally

ORCIDs

DO = 0000-0002-2346-2545

EL = 0009-0006-4098-5010

GP = 0000-0002-5211-2608

RV = 0009-0008-6046-6335

JL = 0000-0003-0216-1951

LSN = 0009-0005-8203-7249

AP= 0000-0002-7759-6776

SRJ = 0000-0003-3615-6958

DPD = 0000-0002-4292-0599

ACW = 0000-0002-5016-4164

Abstract

Nanopore direct RNA sequencing (DRS) coupled with Dorado modification-aware basecalling enables mapping of epitranscriptomic modifications including N^6-methyladenosine (m^6A) at the level of individual RNAs. However, a lack of systematic benchmarking continues to raise questions regarding the sensitivity, specificity, and reproducibility of this method. To address this and to establish a best-practice workflow, we evaluated multiple Dorado versions using *in vitro* transcribed RNA and an m^6A methyltransferase inhibitor as specificity controls. We established that stringent filtering is necessary to reduce false-positive calls and found strong concordance at high-stoichiometry sites when compared to an orthogonal m^6A mapping method (GLORI). Further, by applying DRS to human primary fibroblasts and HD10.6 neurons, we uncovered cell type-specific differences in m^6A stoichiometry, indicating a finely tuned epitranscriptomic regulation. Our study thus presents the first systematic comparison of Dorado and GLORI from the same input RNA and expands characterization of the m^6A epitranscriptome to fibroblasts and neurons.

Introduction

Epitranscriptomics describes the occurrence of chemical modifications on RNA that regulate stability, localization, splicing, translation or decay (Gilbert and Nachtergaele 2023). To date, over 170 RNA modifications are known to occur across all classes of RNA including eukaryotic mRNA (Cappannini et al. 2023). Of these, at least thirteen different chemical modifications have been characterized in mRNAs, comprising both cap-proximal and internal modifications (Anreiter et al. 2021). By contrast, current evidence indicates that lncRNAs harbor a more limited repertoire of modifications, primarily N^6-methyladenosine (m^6A), 5-methylcytosine (m^5C), and N7-methylguanosine (m^7G) (Yang et al. 2024a). One of the best studied and most abundant internal mRNA modifications is m^6A. Functional studies of m^6A methylases, demethylases and m^6A binding proteins, as well as sequencing-based mapping

methods have increased our understanding of m^6A as a reversible and dynamic modification that fine-tunes nearly all aspects of mRNA metabolism (Zaccara et al. 2019; He and He 2021; Murakami and Jaffrey 2022; Gilbert and Nachtergaele 2023; Sendinc and Shi 2023).

In addition to m^6A, other widespread and functionally important marks include m^5C, which is less abundant but influences mRNA export and stability, pseudouridine (Ψ), which contributes to mRNA stability, translation and splicing, and adenosine-to-inosine (A to I) editing, which is involved in splicing, stability and recoding events that contribute to protein diversity (Sun et al. 2023).

Techniques based on short-read (Illumina) sequencing have advanced in recent years to achieve single-nucleotide resolution and quantitative mapping of RNA modifications. The approaches that have emerged to map m^6A can be broadly classified as antibody-based, enzyme-assisted or chemical conversion methods (Moshitch-Moshkovitz et al. 2024).

Early antibody-based techniques (Dominissini et al. 2012; Meyer et al. 2012; Linder et al. 2015) provided the first transcriptome-wide insights into m^6A architecture, albeit at limited resolution. Enzyme-assisted methods (Meyer 2019; Xiao et al. 2023) have improved precision but can introduce enzymatic biases. These biases are avoided with GLORI-Sequencing (GLORI), which selectively deaminates unmethylated adenosines, enabling theoretically unbiased, quantitative, single-nucleotide mapping of m^6A (Liu et al. 2023).

Despite advances, the broad utility of these approaches is limited by a lack of isoform specificity, as well as biases associated with amplification, specificity of antibodies or variable adenosine conversion rates (Diensthuber and Novoa 2025). Oxford Nanopore Technologies (ONT) direct RNA sequencing (DRS) offers an alternative to these methods by analyzing native RNA molecules that have retained their modifications (Garalde et al. 2018). As single RNAs pass through nanopores, signal changes are interpreted to determine canonical and modified residues at single-nucleotide resolution on individual transcripts (Jain et al. 2016; Viehweger et al. 2019). For the now discontinued DRS SQK-RNA002 chemistry used with R9 flowcells, multiple research groups developed algorithms for detecting RNA modifications. These methods can be divided into error rate-based and current signal-based (Abebe et al. 2022). While the recent release of the DRS SQK-RNA004 chemistry and RNA-specific flowcells reduced the utility of many of these tools, introduction of a modification-aware basecaller (Dorado) provides an all-in-one toolkit for basecalling, RNA modification detection, and poly(A) tail length estimation. Early versions of Dorado (e.g. v0.5.0 and v0.6.0) enabled m^6A (DRACH-context only) detection, while version

84 v0.7.0 enabled detection of Ψ and all-context m^6A. More recent versions (v0.8.0 and v0.9.0) allow both
85 DRACH- and all-context m^6A detection, together with Ψ, inosine and m^5C (Zou et al. 2025).

86 In a recent overview of this emerging technology, Cruciani and Novoa highlighted the importance of
87 rigorous evaluation of the limitations of Dorado and advocated forcefully for the establishment of best
88 practices to ensure consistency and accuracy going forward (Cruciani et al. 2025). To date, only a handful
89 of studies of this kind have been published. A series of studies from Novoa and colleagues showed that
90 Dorado performs robustly on synthetic modified and unmodified RNA molecules but also found that
91 overlaps in the identification of specific modifications and false-positive calls remain a challenge but can
92 be mitigated by the inclusion of *in vitro* transcribed (IVT) controls (Cruciani et al. 2025; Diensthuber and
93 Novoa 2025). Independently, Esfahani and colleagues used poly(A) RNA from a B-lymphocyte cell line to
94 compare SQK-RNA002 and SQK-RNA004-derived DRS data. While they concluded that Dorado
95 outperformed other RNA modification detection tools, IVT controls were necessary to reduce false-
96 positive calls (Esfahani et al. 2025). Recent benchmarking efforts comparing Dorado and m6Anet on SQK-
97 RNA004 chemistry have shown that both tools achieve high recall and strong stoichiometry correlations
98 at DRACH motifs but exhibit higher false discovery rates on IVT controls and lower recalls of low-
99 stoichiometry genomic m^6A sites. However, these evaluations relied on previously published GLORI and
100 eTAM-Seq datasets, rather than matched input material (Diensthuber and Novoa 2025; Zou et al. 2025).
101 By directly comparing m^6A detection by DRS coupled with Dorado and GLORI on the same RNA inputs,
102 our study overcomes this limitation. Specifically, we systemically benchmark multiple versions of Dorado
103 using poly(A) RNA from primary (normal) human dermal fibroblasts (NHDFs) and functional human
104 neurons produced through terminal differentiation of HD10.6 cells (Raymon et al. 1999), along with IVT
105 RNA. We included treatments with STM2457, an enzymatic inhibitor of the primary m^6A
106 methyltransferase METTL3 (Yankova et al. 2021), to robustly assess false-positive m^6A detection rates
107 and also compared the DRS outputs with those obtained from GLORI (Yang et al. 2024b) using the same
108 RNA as input. While our analysis is centered on the sensitivity and specificity of m^6A detection in DRS
109 datasets, we also benchmark the ability of Dorado to reliably detect m^5C, Ψ, and inosine. Together, our
110 data provides a robust workflow for DRS-based modification detection that reliably captures
111 transcriptome-wide, isoform-level, site-specific stoichiometries of RNA modifications.By expanding m^6A
112 mapping on primary human fibroblasts and functional human neurons, we uncover mostly conserved
113 m^6A profiles but reveal increased site-specific m^6A stoichiometries in neuronal cells.

114

115 # Results

116 **Systematic benchmarking of ONT's Dorado across human cell types.**

117 To evaluate the performance of ONT's basecalling tool Dorado across different human cell types, we
118 selected primary NHDFs and differentiated HD10.6 cells as a neuron-like model. Prior to isolation of total
119 RNA, cells were treated for 48h with a small molecule inhibitor of METTL3 (STM2457) or with vehicle
120 (DMSO). We then isolated poly(A) RNA and subjected the same input material to both DRS and GLORI.
121 The same material was also used to generate IVT RNA (**Fig. 1A**). Dorado basecalling using version v.0.9.0
122 resulted in over 13M reads using poly(A) RNA from NHDFs treated with DMSO or STM2457, while we
123 obtained over 5M and 10M reads from DMSO- and STM2457-treated HD10.6 cells, respectively. The rate
124 of primary alignment to the human genome (hg38) was over 98% for all the sequenced samples. A
125 detailed overview of read statistics generated by DRS can be found in **Supplementary Table 1**.

126 To systematically benchmark Dorado, we performed basecalling using versions v0.5.0 through v0.9.0.
127 Reads were aligned to the human genome and transcriptome to obtain comprehensive information on
128 isoform-level modification detection. Further, we explored different aspects of detected modifications
129 using the outputs generated by ONT's Modkit tool. We assessed multiple filtering strategies using base
130 and modification probabilities, as well as stoichiometry distributions, and evaluated them based on the
131 IVT data and sensitivity to METTL3 inhibition. By filtering out false-positive calls using multiple control
132 layers, we determined high-confidence m^6A sites (**Fig. 1B**). Finally, we compared these high-confidence
133 predictions with sites identified by GLORI performed on the same input RNA. Using these results, we
134 developed a benchmarked pipeline that filters m^6A sites predicted by Dorado to minimize false-positive
135 detection without impacting sensitivity (**Fig. 1C**).

136 **Stringent filtering of modification probabilities removes false positives**

137 During modification-aware basecalling, each nucleotide is assigned a basecall probability score ranging
138 from 0 to 1. Similarly, individual modified nucleotides are assigned a modification probability score in
139 the same range. To establish a baseline for false-positive m^6A predictions, we plotted basecall and
140 modification probability distributions for an IVT poly(A) RNA dataset processed using Dorado v0.9.0 and
141 aligned to human genome (**Fig.2**). We binned basecall and modification probability values for A and m^6A
142 (all-context and DRACH) and compared this data from DMSO- and STM2457-treated NHDF datasets.

143 We observed that 75-90% of A basecalls (in all datasets) have probability scores above 0.8 **(Fig. S1A)**.
144 Since all A bases are considered during all-context m^6A calling, the fraction of m^6A sites is substantially
145 smaller when compared to DRACH-context m^6A sites (**Fig.2A,B, Fig.S1B,C**). The modification probability
146 distributions are effectively identical for DMSO-treated, STM2457-treated and IVT datasets below 0.98,
147 while the profiles differ substantially at probabilities ≥ 0.98. This suggests that sites detected below 0.98
148 are likely false positives, indicating that only high-probability m^6A calls can be considered reliable. Here,
149 STM2457 treatment shows a marked reduction in the number of reported m^6A sites in both all-context
150 and DRACH-only datasets when compared to the DMSO treatment (**Fig. 2A,B**).

151 In addition to m^6A, we examined modification probability distributions for Ψ, inosine and m^5C, each of
152 which can be basecalled with Dorado v0.9.0 (**Fig. S2**). For Ψ and I, the modification probability
153 distributions for the STM2457 and DMSO datasets only diverge from the IVT dataset at ≥ 0.99
154 modification probability, thus providing a relevant threshold for filtering out false positives (**Fig. S2A,B**).
155 By contrast, m^5C modification probability distributions were similar between all three datasets and no
156 filter could be adequately established (**Fig. S2C**). While this likely reflects a low rate of m^5C installation, it
157 also indicates that the m^5C model utilized by Dorado requires further refinement.

158 Modification probability distributions have similar profiles for transcriptome-aligned NHDF data (**Fig. S3)**
159 as well as genome- and transcriptome-aligned data from HD10.6 cells **(Fig. S4, Fig. S5**). Of note, the
160 fraction of high-probability inosine sites is increased in STM2457-treated cells, compared to DMSO-
161 treated ones in genome-aligned data (**Fig. S2B, Fig.S4D**), which points to increased numbers of retained
162 introns in the METTL3-inhibited sample. Accordingly, this effect is absent in transcriptome-aligned data
163 (**Fig. S3D, Fig. S5D**).

164 While IVT transcriptomes serve as excellent negative controls, we also considered whether the RNA
165 calibration strand (RCS) that is added during DRS library preparation could serve a similar function. The
166 RCS is a synthetic Enolase II RNA derived by IVT[13] that, similarly to the IVT transcriptomes, are devoid of
167 RNA modifications. Despite the reduced diversity of sequence in the RCS data, the overall results were
168 comparable to that from the IVT transcriptome data (**Fig. S6A,B, Fig. 2A,B**), demonstrating that the RCS
169 can be used as a "built-in" negative control for modification detection with DRS. Finally, we determined
170 that basecall and modification probability distributions remain generally consistent between Dorado
171 versions although more m^6A and Ψ sites are detected with high modification probability in Dorado
172 versions v0.8.0/v0.9.0 compared to v0.7.0 (**Fig. S6C-E)**.

173 **Filtering by stoichiometry further eliminates potential false positives**

174 We next examined the stoichiometry distributions for all m^6A sites in both unfiltered and filtered NHDF
175 datasets (**Fig. S7**). Here, unfiltered datasets showed that the vast majority of m^6A sites, 92% for all-
176 context and 75% for DRACH-context, have stoichiometry values < 10% with less than 1% sites showing
177 high (> 50%) stoichiometry. Application of our filtering strategy removed over 70% of the sites (**Fig. 2E)**,
178 regardless of the context, almost all of which showed stoichiometries of < 10%, without significantly
179 changing the numbers of high stoichiometry sites detected (**Fig. S7**). A similar analysis of the IVT
180 transcriptome revealed large numbers of false positive m^6A sites with stoichiometry values below 10%,
181 most of which were eliminated by application of our filters (**Fig. S8A,B**). However, the remaining
182 presence of low stoichiometry false-positive sites indicates that an additional layer of filtering to exclude
183 sites with < 10% stoichiometry is beneficial. An additional argument in support of our modification
184 probability filtering strategy is the observation that treatment with the METTL3 specific inhibitor
185 STM2457 dramatically reduced m^6A installation at sites with ≥ 10% stoichiometry in untreated datasets
186 in a context-independent manner (**Fig. 2C,D, Supplementary Table 2**). Similar results were observed for
187 the transcriptome aligned NHDF and HD10.6 m^6A stoichiometry distributions as well (**Fig. S9, Fig. S10,**
188 **Supplementary Table 2**).

189 To determine whether our filtering strategy could be applied to other RNA modifications detected by
190 Dorado, we generated stoichiometry distributions for Ψ and inosine in unfiltered and filtered (≥ 0.99
191 modification probability) human genome-aligned DMSO- and STM2457-treated NHDFs and HD10.6s. As
192 with m^6A (**Fig. S7**), 92% of detected Ψ and 96% of inosine have less than 10% stoichiometry (**Fig. S11,**
193 **Fig. S12)**. A recent study that explores Ψ in human mRNAs using 2-bromoacrylamide-assisted cyclization
194 sequencing (BACS) corroborates our results, showing that 82% of Ψ modifications are detected at low
195 levels (below 20%) (Xu et al. 2024).

196 Above 10% stoichiometry, in our data, there are more inosine sites in STM2457-treated cells than DMSO-
197 treated cells, reflected also in the stoichiometry distributions (**Fig. S11, Fig. S12, Supplementary Table 2**).
198 Ψ and inosine stoichiometry distributions in our IVT data **(Fig.S8C,D)** further validates our filtering
199 strategy of ≥ 0.99 modification probability, as it substantially reduces the number of detected false-
200 positive sites in the unfiltered data even at higher (> 50%) stoichiometries in case of inosine (**Fig.S8D**).

201 We have thus demonstrated that setting thresholds for basecall modification probabilities provides a
202 robust method for identifying and filtering false-positive modification calls. For the purposes of this

203 study, we further evaluated the effect of placing the modification threshold on the modification data and
204 restricted the downstream analyses to m^6A sites with probability scores ≥ 0.98.

205 **RNA modification detection rates differ between Dorado versions**

206 Dorado basecalling relies on models generated using Remora (github.com/nanoporetech/remora) that
207 are frequently updated. This raises an important question as to whether these updates produce notably
208 different outputs that could impact biological interpretation. To address this, we utilized multiple Dorado
209 versions and Remora models to perform modification calling of our NHDF dataset. We restricted our
210 analyses to m^6A and Ψ as these are the only modifications that could be detected in previous versions.
211 Here, we observe that m^6A and Ψ stoichiometry distributions are generally consistent between different
212 Dorado versions (**Fig. S13**), albeit with greater numbers of low stoichiometry (< 10%) sites detected in
213 older Dorado versions. There is a large overlap of the detected ≥ 10%-stoichiometry m^6A (DRACH and
214 all-context) and Ψ sites between versions, although there are a lot of unique all-context m^6A and Ψ sites
215 identified with just v0.7.0 and not v0.8.0-v0.9.0, that are reduced after applying our modification
216 probability filter (**Fig. S14A,B**). Ψ sites uniquely detected with different versions are found in slightly
217 different sequence contexts (**Fig. S14C,D**). Thus, while RNA modification detection with Dorado is
218 sensitive to changes in the underlying models/versions, this can be effectively controlled through the
219 application of filters that only retain sites with high-confidence modification probability calls.

220 **GLORI exhibits reduced specificity but converges with Dorado predictions at high stoichiometry sites**

221 To assess the agreement of DRS Dorado-derived predictions with an orthogonal m^6A detection method,
222 we conducted GLORI on the same NHDF RNA preparations. To ensure a fair comparison of m^6A sites
223 across the two methods, we applied the following filtering strategies: For Dorado (DRACH model), this
224 included thresholds on modification probability (A ≥ 0.8, m^6A ≥ 0.98) and read coverage (≥ 20). For
225 GLORI, we used filters based on coverage (≥ 20) and statistical significance (p-value ≤ 0.05). Next, we
226 extracted m^6A sites that were present in both DMSO and STM2457 treated samples and removed all
227 sites with ≤ 10% stoichiometry in DMSO (**Fig. 3A**). This stepwise strategy reduced false positives and
228 enabled a fair comparison of the high confidence m^6A sites identified by both approaches.

229 As shown in the violin plots, both the all-context and DRACH-context Dorado models identify similar
230 numbers of m^6A sites (81,012 (all-context model) and 72,372 (DRACH model)), the stoichiometries of
231 which were strongly reduced upon STM2457 treatment (**Fig. 3B**). When we analyzed the motifs of the
232 m^6A sites detected with the all-context model, we observed nearly 9,000 sites outside of DRACH motifs.

233 Importantly, the total number of m^6A sites within DRACH motifs was nearly identical to the number of
234 m^6A sites reported using the DRACH model (**Fig.3B**). Unexpectedly, GLORI detects a substantially higher
235 number of sites when using all motifs (140,934) compared to DRACH motifs alone (27,926) (**Fig. 3C**). The
236 fact that most GLORI-identified sites fall outside DRACH motifs suggests these are false positives. This
237 interpretation is reinforced by the observation that the overlap between Dorado and GLORI datasets
238 remains virtually unchanged when comparing GLORI "all motif" and "DRACH motifs" (n=13,614) site
239 sets, underscoring that the additional non-DRACH sites in GLORI (n= 113,008) are largely spurious (**Fig.**
240 **S15A**).

241 We obtained similar results when comparing m^6A sites obtained with Dorado and GLORI from HD10.6
242 cells. Overall, we detected a similar number of m^6A sites in HD10.6 cells using the all-context and
243 DRACH-specific Dorado models (77,728 and 67,222, respectively) (**Fig. S15B**). By contrast, we detected
244 174,282 vs 27,965 m^6A sites using GLORI without or with DRACH filtering, respectively (**Fig. S15C**).

245 To evaluate how stoichiometry thresholds affect the agreement between GLORI and Dorado, we
246 compared the overlap of high-confidence GLORI-identified sites with Dorado m^6A (DRACH context) calls
247 across increasing m^6A stoichiometry cutoffs (**Fig. 3D**). At the lowest stoichiometry threshold applied (≥
248 10%), 13,614 sites overlapped, representing less than half (48.8%) of GLORI sites. However, as the
249 stringency increased, the relative overlap increased sharply: at ≥ 20%, nearly 70% of GLORI sites
250 overlapped (11,536 sites), at ≥ 50% almost 90% overlapped (6,307 sites), and at the highest cutoff (≥
251 80%), over 93% of GLORI sites (3,119 sites) were also detected by Dorado. These results demonstrate
252 that while GLORI detects a broad set of potential m^6A sites, only the sites with high stoichiometries
253 overlap with Dorado, indicating high signal to noise ratios in the GLORI datasets.

254 GLORI relies on the high-efficiency conversion of adenosine to inosine with overall gene A to I conversion
255 rates needing to exceed 95-98% (Xu et al. 2024; Xie et al. 2025). Similarly, we observed the conversion
256 rates for our NHDF and HD10.6 datasets to be 95% (Supplementary **Table 3**), but we found that small
257 variations in gene conversion rates can strongly influence the sensitivity and specificity of GLORI, as only
258 high-stoichiometry m^6A sites overlapped with DRS-derived predictions (**Fig. 3D**).

259 **DRS coupled with Dorado outperforms GLORI in isoform-specific m^6A detection**

260 Next, we wanted to highlight the advantage of long-read sequencing in terms of mapping reads with
261 isoform-precision to the human transcriptome, which is not possible using GLORI due to its reliance on
262 short-read sequencing. We first confirmed that mapping Dorado-basecalled sites to the human

263 transcriptome results in the detection of similar numbers of m^6A sites in both NHDFs and HD10.6 cells
264 (**Fig. S15D,E**). To assess if DRS coupled with Dorado identifies distinct m^6A profiles on different transcript
265 isoforms, we analyzed the *SPEN* gene locus which was previously shown to contain more than 100 m^6A
266 sites (Liu et al. 2023). On a genomic level, we detected 42 m^6A sites using DRS, while GLORI only
267 detected 5 m^6A sites (**Fig. 3E**). Out of the 7 total isoforms of *SPEN*, *SPEN-201* and *SPEN-202* were
268 detectable in DRS-derived data, considering a minimal coverage of 20 (**Supplementary Table 4**). When
269 plotting the m^6A sites detected by transcriptome-aligned DRS data, we found that the *SPEN-201* isoform
270 exhibited more m^6A sites, albeit with a lower median stoichiometry compared to the *SPEN-202* isoform
271 (**Fig. 3F**). This example highlights that DRS is superior compared with GLORI for isoform-level m^6A
272 quantification.

273 **Dorado profiling reveals cell type differences in m^6A distribution and stoichiometry**

274 To assess the reproducibility of Dorado across DRS datasets derived from different cell types, we
275 compared the m^6A methylomes of NHDFs and HD10.6 cells. To ensure a fair comparison, we extracted
276 genes that were expressed in both cell types. Next, we compared the total numbers of high confidence
277 m^6A sites in NHDFs and HD10.6 cells within the genes expressed in both cell types and found that 87% of
278 HD10.6-derived m^6A sites overlap with those detected in NHDFs (**Fig. 4A**). Site-specific m^6A
279 stoichiometries from the shared gene set correlated strongly across cell types (Pearson r = 0.93),
280 although we observed a global shift in the HD10.6 dataset, where m^6A stoichiometries were generally
281 higher than in NHDFs (**Fig. 4B**). Plotting the m^6A stoichiometry distributions of the 38,217 m^6A sites
282 shared between both cell types showed a strong increase in the median m^6A stoichiometry of HD10.6
283 cells (70.2%), compared to NHDFs (54%) (**Fig. 4C**).

284 We next aimed to conduct a comparative motif analysis of m^6A sites detected in NHDFs and HD10.6
285 cells. Sequence logos confirmed strong enrichment of the canonical DRACH consensus motif for both cell
286 types in modification probability-filtered data, with the GGACU pentamer being the most prominent (**Fig.**
287 **4D,E, Fig.S16**). Quantification of 5-mer distributions revealed highly similar motif preferences across
288 datasets. While the global motif architecture of m^6A deposition is conserved between NHDFs and
289 HD10.6 cells, we observed a trend toward higher stoichiometries in the HD10.6 dataset, particularly
290 within the most abundant DRACH motifs, suggesting enhanced modification efficiency in the neuron-like
291 context (**Fig. 4F, G**).

292 Metagene analysis showed, as known previously, that m^6A is enriched near the stop codons and in the 3'
293 UTRs (**Fig. S17A**). Several studies have examined the relationship between exon architecture and m^6A
294 distribution on mRNAs[1,2,3], and so we also examined m^6A sites in monoexonic transcripts separately and
295 found that monoexonic transcripts are more broadly m^6A-modified across entire CDS (**Fig. S17A**). m^6A
296 (both all-context and DRACH-context) distribution profiles look less typical for all-stoichiometry or ≤ 10%
297 stoichiometry sites in both NHDF and HD10.6 datasets (**Fig. S17B+C**). For all-context m^6A (**Fig. S17B**),
298 modification filter also proves to be deciding for the distribution profile, together with ≥ 10%
299 stoichiometry.

300 To connect shared m^6A sites with their functional context and illustrate their distribution on individual
301 transcripts, we conducted a KEGG pathway analysis with the genes that contain shared m^6A sites
302 between NHDFs and HD10.6 cells. We observed a significant enrichment in pathways related to
303 neurodegeneration, infection, and intracellular signaling, including axon guidance and mTOR signaling
304 (**Fig. 4H**). To further illustrate these findings, we examined representative genes contained in one or
305 more of these pathways. *WNTB5*, *MAPK1*, *AKT1*, *NCK1* and *NRP1* all showed robust methylation in both
306 NHDFs and HD10.6 cells, with consistently higher stoichiometries observed in the neuronal-like context.
307 In contrast, *ACTG1*, included as a housekeeping control not represented in the enriched pathways,
308 displayed stable and comparatively uniform methylation across both cell types (**Fig. 4I**).

309 Together, these results demonstrate that while the global architecture of m^6A deposition is largely
310 conserved between primary fibroblasts and neuronal-like cells, but unexpectedly the neuronal context is
311 characterized by a global shift toward higher m^6A stoichiometries. This difference is readily detectable in
312 single genes. These findings suggest that cell type-specific regulation of m^6A may fine-tune core cellular
313 pathways relevant to that cell type.

314

315 ## Discussion

316 High rates of false positives remain a major problem in epitranscriptomic mapping tools, even with the
317 advent of modification-aware basecalling models (Cruciani et al. 2025). Comparing the outputs of
318 orthogonal methods performed on the same input material, combined with use of IVT RNA and/or
319 specific methyltransferase inhibition, offers an effective mitigation strategy (Cruciani et al. 2025;
320 Diensthuber and Novoa 2025). The results presented here establish a robust framework that minimizes
321 false positives and enables isoform-specific and quantitative m^6A mapping. Through a direct comparison

322 with GLORI, we showed that combining DRS datasets with Dorado modification-aware basecalling
323 provides greater scalability, isoform resolution, and reproducibility across different cell types. Further,
324 the simplicity of DRS library preparation enhances consistency whereas GLORI is constrained by a
325 laborious and technically challenging protocol that is extremely sensitive to adenosine conversion rates
326 (Shen et al. 2024). Despite these limitations, we determined that the majority of m^6A sites in our filtered
327 DRS datasets could be confirmed by GLORI. By expanding our analysis across multiple human cell types,
328 we further determine that the sites of m^6A installation are conserved for the most part but that the
329 stoichiometries show cell-type specific differences.

330 Crucially, the inclusion of transcriptome-wide IVT datasets and the use of a selective METTL3 inhibitor
331 (STM2457) provides a robust approach to developing filters that dramatically decrease false-positive m^6A
332 detection. The impact of both approaches was readily visualized by plotting modification probability
333 distributions, enabling our determination that retaining only m^6A calls with confidence scores of 0.98 or
334 greater retained almost all STM2457-sensitive m^6A sites while eliminating the large numbers of false-
335 positive m^6A sites reported for the IVT datasets. This approach is supported by a recent study which
336 highlighted that Dorado introduces large numbers of false positives outside canonical DRACH contexts
337 (Xie et al. 2025). Our data confirms that in the unfiltered datasets, the number of false positives is higher
338 in the all-context model compared to the DRACH model. This may reflect the degenerate nature of the
339 canonical DRACH consensus and understudied differences in sequence preference between different cell
340 types and organisms. Retaining only sites at which m^6A stoichiometries exceeded 10% further reduced
341 false-positive rates while also focusing the results on identification of sites that are more likely to be
342 biologically significant. The combined effect of these filters was to provide a strong agreement between
343 the m^6A all-context and m^6A DRACH model in the detection of high confidence m^6A sites, as observed by
344 filtering for DRACH motifs from m^6A sites detected with the all-context model (**Fig.3B**).

345 Prior studies have emphasized the need for isoform-resolved m^6A quantification by demonstrating
346 widespread heterogeneity in m^6A levels between genes and their isoforms (Guo et al. 2025; Xie et al.
347 2025). Here, we showed that aligning DRS datasets to the human transcriptome reveals differential m^6A
348 installation patterns between isoforms, underscoring the importance of accurate, isoform-specific m^6A
349 quantification at single-molecule resolution and highlighting the advantage of DRS against short-read
350 sequencing approaches such as GLORI (**Fig. 3E,F**).

351 Recently, the ability of Dorado and m6Anet (Hendra et al. 2022) to identify m^6A in DRS datasets from
352 293T and HeLa cells were systematically evaluated using IVT RNA and comparisons to published datasets

353 obtained by GLORI and eTAM-seq (Zou et al. 2025). In this case the authors found that Dorado (v0.7.2)
354 outperforms m6Anet but also advocated for the importance of applying filters to reduce false positive
355 calls. Likewise, we found that IVT RNA served as a valuable negative control to detect false positive calls.
356 Furthermore, using the RNA calibration strand (RCS) as a negative control provided a convenient and
357 simpler alternative to IVT RNA, yielding comparable probability distributions and thereby enabling a
358 built-in control for routine experiments (Viehweger et al. 2019). While m^6A detection was generally
359 reproducible across Dorado versions, model-specific differences are present, i.e. v0.7.0 produces
360 substantially more low-confidence all-context m^6A and Ψ calls than v0.8.0–0.9.0, highlighting that the
361 models remain under active development and that benchmarking should be repeated when switching to
362 a new model. Recent work further highlights that continued improvements in both nanopore sequencing
363 chemistry and basecalling models significantly enhance RNA modification detection accuracy,
364 emphasizing the importance of revisiting benchmarking and filtering strategies as models evolve (Hewel
365 et al. 2025).

366 Focusing on other modifications that are detected by Dorado, we observe that inosine and Ψ detection
367 appears robust when applying similar filtering strategies as for m^6A, the latter supported by other
368 studies (Esfahani et al. 2025; Verstraten et al. 2025). For inosine, we observed an enrichment of high-
369 probability (> 0.99) inosine sites in genome-aligned data from STM2457-treated samples. Further studies
370 are needed to determine whether this reflects increased intron retention or misclassification of
371 modifications by the current inosine model. While DRS is the only approach that allows for robust
372 isoform-specific m^6A mapping, comparisons with orthogonal approaches remain essential to validate
373 findings. When comparing Dorado and GLORI using the same input RNA, we unexpectedly observed that
374 only 50% of m^6A sites with stoichiometries above 10% that were identified by Dorado were also present
375 in the GLORI data although increasing the stoichiometries led to increased overlap (**Fig.3D**). We attribute
376 this to a sub-optimal gene-wide adenosine conversion rate of 95% during GLORI, compared to the 98-
377 99% rate reported in the original study (Shen et al. 2024). Here, unconverted adenosines result in high
378 signal-to-noise ratios which result in lower sensitivity, explaining why the overlap with Dorado increases
379 at higher-stoichiometry sites. This is also reflected in the relatively high amount of false positive sites
380 detected outside of DRACH motifs in the GLORI datasets (**Fig. 3C**). The sensitivity of GLORI to small
381 fluctuations thus limits its utility in comparison to DRS. Whether this can be resolved with improvements
382 to the GLORI methodology (Sun et al. 2025) requires further study.

383 Prior analyses of m^6A stoichiometry distributions across immortalized human cell lines, including cancer
384 cell lines, have revealed broadly similar m^6A stoichiometry distributions (Guo et al. 2025). By contrast,
385 our analysis of primary fibroblasts (NHDFs) and neuronal cells (differentiated HD10.6 cells) revealed an
386 overall increase in m^6A stoichiometries in neurons compared to fibroblasts, especially in genes
387 associated with neuronal function. While further studies are warranted to establish the underlying
388 mechanisms and functional consequences of these differences, this aligns with the importance of m^6A in
389 neuronal function (Yu et al. 2021). This includes m^6A maps from synaptic and nuclear RNA isolated from
390 mouse dorsal hippocampus which showed that synapse-specific RNAs showed generally higher
391 methylation levels (Sun et al. 2025). Intriguingly, neuronal regulation pathways utilize differentially
392 spliced isoforms (Su et al. 2018) and a recent study revealed isoform-specific differences in m^6A sites and
393 stoichiometries in the human brain that are linked to m^6A machinery expression and cell type context
394 (Gleeson et al. 2025).

395 In summary, using matched RNA samples from different experimental conditions and from different
396 human cell types analyzed with both Dorado and GLORI orthogonal methodologies, we establish a
397 robust pipeline for the precise and quantitative transcriptome-wide detection of m^6A in DRS datasets.
398 Future studies should aim at extending this approach using orthogonal sequencing approaches to
399 benchmark Dorado-based detection of Ψ, m^5C and inosine, as well as the newly released models for 2'O-
400 methylation.

401

402 # Methods

403 **Cell culturing and STM2457 treatment**

404 NHDFs (Lonza) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco), supplemented with
405 5% heat-inactivated FBS (HI-FBS, Gibco) and 1% penicillin-streptomycin (PS, Lonza) under standard
406 culture conditions. The HD10.6 cell line was originally derived from fetal human dorsal root ganglia
407 (DRG) and immortalized by overexpression of the v-myc oncogene under tetracycline regulation. Upon
408 addition of tetracycline and neurotrophins, these cells can be differentiated into neuron-like cells that
409 show typical neuronal morphology and nociceptive properties(Raymon et al. 1999). For our experiments,
410 we used HD10.6 cells that have been matured for 10 days. For maintenance of undifferentiated HD10.6
411 cells (a gift from Anna Cliffe, University of Virginia School of Medicine), T-75 flasks were coated with 17
412 µg/mL fibronectin in PBS for 15 min at 37°C before adding the cell suspension. Cells were grown in

413 Advanced DMEM/F12 (Gibco), supplemented with 1x Neurocult SM1 (STEMCELL Technologies), 2 mM L-
414 glutamine (Gibco), 1x Primocin (Invivogen) and 10 ng/mL prostaglandin E1 (Sigma). 0.5 ng/mL bFGF
415 (PeproTech) were added freshly to each media change. For maturation, HD10.6 cells were seeded at a
416 density of 25.000 cells/cm² in 6-well plates which were pre-coated overnight with 50 µg/mL poly-L-
417 ornithine hydrobromide (Sigma) in 0.5M borate buffer pH 8.5 (Boston Bioproducts), and 1 µg/mL
418 fibronectin (overnight). After cell attachment overnight, media was changed to Neurobasal media
419 (Gibco), supplemented with 1x Neurocult SM1, 2mM L-glutamine, 1x Primocin, 1 µg/mL doxycycline, 50
420 ng/mL NGF (Alomone Labs), and 25 ng/mL CNTF, GDNF and NT-3 (PeproTech). Media was changed every
421 three days until HD10.6 cells showed complete neuronal morphology after 12 days. NHDFs and matured
422 HD10.6 cells were treated with 30 µM STM2457 (Sigma) or DMSO for 48h, harvested in TRIzol
423 (Invitrogen), and stored at -80°C until RNA isolation.

424 **RNA isolation**

425 RNA was extracted by the TRIzol method according to the manufacturer's instructions. Briefly, 0.2 mL
426 chloroform per 1 mL TRIzol were added and mixed by vortexing. After centrifugation for 30 min at
427 12,000xg at 4°C, the aqueous phase was transferred to isopropanol (0.5 mL per 1 mL TRIzol), mixed and
428 incubated for 10 min at room temperature. Subsequently, the samples were centrifuged for 10 min at
429 12,000xg at 4°C and the RNA pellets were washed twice in 75% fresh ethanol. After air-drying for 5-10
430 min, RNA was dissolved in nuclease-free H_2O and stored at -80°C until further processing. One third of
431 each sample was used for DRS and the remaining same-input material was processed for GLORI. For both
432 DRS and GLORI, poly(A) RNA was selected using the Dynabeads™ mRNA Purification Kit (Invitrogen),
433 according to the manufacturer's instructions.

434 **GLORI-Sequencing**

435 GLORI reactions were largely performed following the published protocol for RNA protection,
436 deamination and deprotection with several minor adjustments (Shen et al. 2024). Briefly, poly(A) RNA
437 was fragmented at 94°C for 2 min (NEB E6150S) and subjected to glyoxal protection, sodium-nitrite-
438 mediated deamination (750 mM $NaNO_2$ at 16°C for 8h and 4°C for 6h), and deprotection in
439 triethylammonium acetate-formamide buffer at 95°C for 10 min. Library construction was conducted
440 using a ligation-based strategy as described by Liu and colleagues incorporating 5'adapters with 11-nt
441 UMIs (Liu et al. 2025).

442 Indexed libraries were amplified with NEBNext Dual Primers Set 2 (E7780S), purified using AMPure XP
443 beads, and validated on an Agilent TapeStation (D1000 High Sensitivity). Sequencing was performed on
444 an Illumina NovaSeq 6000 platform (150 bp paired-end mode; > 80 million reads per sample).

445 For data processing, we employed the same computational pipeline described in Liu et al (Liu et al.
446 2025), but implemented within a Snakemake-based framework that automates the complete analysis
447 starting from raw fastq files (Köster and Rahmann 2012). Details on software versions, workflow
448 availability, and repository links are provided in the Data and code availability section below.

449 **Generation of a transcriptome-wide IVT RNA library**

450 We sequenced the IVT RNA from both NHDFs and HD10.6 cells along with the native poly(A) RNA to
451 evaluate potential false-positive predictions by Dorado. This was achieved by converting the poly(A) RNA
452 to cDNA using a poly(A) primer and a T7-incorporating strand switch primer and using the cDNA as a
453 template for IVT synthesis.

454 **Nanopore Sequencing of poly(A) RNA**

455 250-300 ng of isolated poly(A) RNA and 25 ng RNA calibration strand (RCS) were used as input for the
456 standard DRS SQK-RNA004 protocol. Resulting libraries were loaded onto individual PromethION RNA
457 flow cells before sequencing on the PromethION 2 Solo instrument for 48h.

458 **Nanopore basecalling**

459 Basecalling was performed using Dorado v0.5.0, v0.6.0, v0.7.0, v0.8.0 and v0.9.0 (ONT) in super accuracy
460 mode with modification detection enabled for m^6A (DRACH context) in v.0.5.0, 0.6.0, 0.8.0 and 0.9.0,
461 m^6A (all context) in v0.7.0, 0.8.0 and 0.9.0, Ψ in v0.7.0, v0.8.0 and v0.9.0, and inosine and m^5C in v.0.9.0.
462 Each dataset was processed with adapter trimming. Resulting unaligned BAM files were converted to
463 FASTQ files containing modification tags for every read in the FASTQ read headers, using SAMtools v1.18
464 (Li et al. 2009) [*samtools fastq -TMM,ML*].

465 **Genome and transcriptome level alignments**

466 For genome-level alignment, reads were aligned against a combined human genome reference
467 (GRCh38.p14), and the RCS gene sequence, using minimap2 v2.26 [*minimap2 -ax splice -k14 -y -uf --*
468 *secondary=no*]. For transcriptome-level alignment, the reads were aligned against the human
469 transcriptome (v47), obtained via GENCODE (Mudge et al. 2025) [*minimap2 -t 8 -ax map-ont -y -L -p*
470 *0.99*]. Using SAMtools v1.18, resulting genome- and transcriptome-aligned SAM files were converted to

471 BAM files retaining only primary alignments [*samtools view -b -F2308* and *samtools view -b -F2324,*
472 respectively]. BAM files were sorted [*samtools sort*] and indexed [*samtools index*].

473 **Processing of RNA modification calls using Modkit**

474 Probability distributions of the detected sites in human genome- and transcriptome-aligned reads were
475 generated using sample-probs subcommand from Modkit v0.4.1 [sample-probs --hist --only-mapped --
476 percentiles 0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.85,0.9 -f 0.1]. For RCS reads -f value was increased to 0.25.
477 The resulting .tsv files were imported into RStudio (v. 2024.12.1+563) and plotted using a custom R(v.
478 4.4.2) script (Dorado_mod_probability_distributions.R), with the following packages: dplyr (v1.1.4),
479 ggplot2 (v3.5.1), reshape2 (v1.4.4), patchwork (v1.3.0).

480 To generate extended bedMethyl files with m^6A-modified and unmodified base counts from the reads
481 mapping against the potential modification sites, we used Modkit pileup subcommand with the
482 thresholds [*--filter-threshold A:0.8 --filter-threshold T:0.8 --filter-threshold C:0.8 --filter-threshold G:0.8 --*
483 *motif A 0 --ref $REF*] with and without modification threshold filter [*--mod-thresholds a:0.98*] to generate
484 our "filtered" and "unfiltered" datasets, respectively. Canonical nucleotide calls (A, C, G, U) were
485 retained with > 80% probability scores for all modification datasets. Modification thresholds were set at
486 98% for m^6A and 99% for all the other modifications. *--motif* was adjusted to the base for which
487 modification status was examined, which further restricted the analysis to only the positions that
488 matched canonical nucleotides in the reference genome. The resulting sites were further filtered
489 according to the read number per site (≥ 20 reads) for all the downstream analysis.

490 **Gene Ontology analysis**

491 GO analysis was performed using DAVID (Huang da et al. 2009). Genes harboring m^6A sites were filtered
492 for genes that were generally present in both datasets. Next, they were divided into NHDF-specific,
493 HD10.6-specific, and shared categories, and enrichment was assessed for biological processes. GO terms
494 were considered significant at FDR ≤ 0.05 with a minimum of 20 genes per GO term. For plotting, the top
495 20 GO terms that contained the highest gene numbers were extracted from the analysis of the shared
496 gene sets.

497 **Analysis and visualization of modification stoichiometries and distributions**

498 Stoichiometry distribution histograms and metaplots were generated using Rstudio and R v4.4.2 with the
499 following libraries: data.table, patchwork, scales, tidyr, dplyr, ggplot2[4]. Access to scripts used for analysis
500 and visualization is detailed below.

501 **Data and code availability**

502 Raw fastq and pod5 datasets generated in this study are available via the European Nucleotide Archive
503 (ENA) under the accession number PRJEB101380.

504 R scripts used in the analysis and visualization of data presented herein are available in the
505 https://github.com/DepledgeLab/HOLDEN repository.

506 The Snakemake workflow implementing the GLORI analysis pipeline is available in the
507 https://github.com/gp-micro/glori repository.

508

509 **Acknowledgements and funding**

510 DPD is supported by a German Centre for Infection Research (DZIF) Associate Professorship and the
511 NIAID grants R01-AI170583 and R01-AI152543. DPD also receives funding from the Deutsche
512 Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy -
513 EXC 2155, project number 390874280. EL is supported by the NIAID grant R01-AI170583 as well as the
514 Hannover Biomedical Research School (HBRS) and the Center for Infection Biology (ZIB). DO is supported
515 by the Walter Benjamin Programme of the DFG, project 554758329. ACW is supported by NIAID grants
516 R01AI176335 and R01-AI170583.

517 ## References

518 https://github.com/nanoporetech/modkit.
519 Abebe JS, Verstraten R, Depledge DP. 2022. Nanopore-Based Detection of Viral RNA Modifications. *mBio*
520 **13**: e03702-03721.
521 Anreiter I, Mir Q, Simpson JT, Janga SC, Soller M. 2021. New Twists in Detecting mRNA Modification
522 Dynamics. *Trends Biotechnol* **39**: 72-89.
523 Cappannini A, Ray A, Purta E, Mukherjee S, Boccaletto P, Moafinejad SN, Lechner A, Barchet C, Klaholz
524 Bruno P, Stefaniak F et al. 2023. MODOMICS: a database of RNA modifications and related
525 information. 2023 update. *Nucleic Acids Research* **52**: D239-D244.
526 Cruciani S, Delgado-Tejedor A, Pryszcz LP, Medina R, Llovera L, Novoa EM. 2025. De novo basecalling of
527 RNA modifications at single molecule and nucleotide resolution. *Genome Biology* **26**: 38.
528 Diensthuber G, Novoa EM. 2025. Charting the epitranscriptomic landscape across RNA biotypes using
529 native RNA nanopore sequencing. *Mol Cell* **85**: 276-289.
530 Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K,
531 Jacob-Hirsch J, Amariglio N, Kupiec M et al. 2012. Topology of the human and mouse m6A RNA
532 methylomes revealed by m6A-seq. *Nature* **485**: 201-206.
533 Esfahani NG, Stein AJ, Akeson S, Tzadikario T, Jain M. 2025. Evaluation of Nanopore direct RNA
534 sequencing updates for modification detection. *bioRxiv* doi:10.1101/2025.05.01.651717.
535 Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland
536 A et al. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* **15**:
537 201-206.
538 Gilbert WV, Nachtergaele S. 2023. mRNA Regulation by RNA Modifications. *Annu Rev Biochem* **92**: 175-
539 198.
540 Gleeson J, Madugalle SU, Wan CY, McLean C, Bredy TW, De Paoli-Iseppi R, Clark MB. 2025. Isoform-level
541 profiling of m(6)A epitranscriptomic signatures in human brain. *Sci Adv* **11**: eadp0783.
542 Guo W, Ren Z, Huang X, Liu J, Shao J, Ma X, Wei C, Cun Y, He J, Zhang J et al. 2025. Single-molecule m6A
543 detection empowered by endogenous labeling unveils complexities across RNA isoforms.
544 *Molecular Cell* doi:https://doi.org/10.1016/j.molcel.2025.01.014.
545 He PC, He C. 2021. m(6) A RNA methylation: from mechanisms to therapeutic potential. *Embo j* **40**:
546 e105977.
547 Hendra C, Pratanwanich PN, Wan YK, Goh WSS, Thiery A, Göke J. 2022. Detection of m6A from direct
548 RNA sequencing using a multiple instance learning framework. *Nature Methods* **19**: 1590-1598.
549 Hewel C, Wierczeiko A, Miedema J, Friedrich J, Hofmann F, Weißbach S, Dietrich V, Holthöfer L, Haug V,
550 Mündnich S et al. 2025. Direct RNA sequencing enables improved transcriptome assessment and
551 tracking of RNA modifications for medical applications. *Nucleic Acids Research* **53**.
552 Huang da W, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using
553 DAVID bioinformatics resources. *Nat Protoc* **4**: 44-57.
554 Jain M, Olsen HE, Paten B, Akeson M. 2016. The Oxford Nanopore MinION: delivery of nanopore
555 sequencing to the genomics community. *Genome Biology* **17**: 239.
556 Köster J, Rahmann S. 2012. Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics* **28**:
557 2520-2522.
558 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The
559 Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
560 Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-nucleotide-
561 resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods* **12**: 767-
562 772.

563 Liu C, Sun H, Yi Y, Shen W, Li K, Xiao Y, Li F, Li Y, Hou Y, Lu B et al. 2023. Absolute quantification of single-
564 base m(6)A methylation in the mammalian transcriptome using GLORI. *Nat Biotechnol* **41**: 355-
565 366.
566 Liu JF, Hawley BR, Nicholson LS, Jaffrey SR. 2025. Decoding m(6)Am by simultaneous transcription-start
567 mapping and methylation quantification. *Elife* **13**.
568 Meyer KD. 2019. DART-seq: an antibody-free method for global m6A detection. *Nature Methods* **16**:
569 1275-1280.
570 Meyer Kate D, Saletore Y, Zumbo P, Elemento O, Mason Christopher E, Jaffrey Samie R. 2012.
571 Comprehensive Analysis of mRNA Methylation Reveals Enrichment in 3′ UTRs and near Stop
572 Codons. *Cell* **149**: 1635-1646.
573 Moshitch-Moshkovitz S, Sevilla-Sharon M, Ashwal-Fluss R, Glick-Saar E, Rechavi G, Dominissini D. 2024.
574 mRNA m6A detection. *Nature Reviews Methods Primers* **4**.
575 Mudge JM, Carbonell-Sala S, Diekhans M, Martinez JG, Hunt T, Jungreis I, Loveland JE, Arnan C, Barnes I,
576 Bennett R et al. 2025. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic*
577 *Acids Res* **53**: D966-d975.
578 Murakami S, Jaffrey SR. 2022. Hidden codes in mRNA: Control of gene expression by m(6)A. *Mol Cell* **82**:
579 2236-2251.
580 Raymon HK, Thode S, Zhou J, Friedman GC, Pardinas JR, Barrere C, Johnson RM, Sah DWY. 1999.
581 Immortalized Human Dorsal Root Ganglion Cells Differentiate into Neurons with Nociceptive
582 Properties. *The Journal of Neuroscience* **19**: 5420-5428.
583 Sendinc E, Shi Y. 2023. RNA m6A methylation across the transcriptome. *Mol Cell* **83**: 428-441.
584 Shen W, Sun H, Liu C, Yi Y, Hou Y, Xiao Y, Hu Y, Lu B, Peng J, Wang J et al. 2024. GLORI for absolute
585 quantification of transcriptome-wide m6A at single-base resolution. *Nature Protocols* **19**: 1252-
586 1287.
587 Su C-H, D D, Tarn W-Y. 2018. Alternative Splicing in Neurogenesis and Brain Development. *Frontiers in*
588 *Molecular Biosciences* **Volume 5 - 2018**.
589 Sun H, Li K, Liu C, Yi C. 2023. Regulation and functions of non-m(6)A mRNA modifications. *Nat Rev Mol*
590 *Cell Biol* **24**: 714-731.
591 Sun H, Lu B, Zhang Z, Xiao Y, Zhou Z, Xi L, Li Z, Jiang Z, Zhang J, Wang M et al. 2025. Mild and ultrafast
592 GLORI enables absolute quantification of m6A methylome from low-input samples. *Nature*
593 *Methods* **22**: 1226-1236.
594 Verstraten R, Cetraro P, Fitzpatrick AH, Alwie Y, Frommeyer YN, Loliashvili E, Stein SC, Häussler S,
595 Ouwendijk WJD, Depledge DP. 2025. Defining expansions and perturbations to the RNA
596 polymerase III transcriptome and epitranscriptome by modified direct RNA nanopore
597 sequencing. *bioRxiv* doi:10.1101/2025.03.07.641986.
598 Viehweger A, Krautwurst S, Lamkiewicz K, Madhugiri R, Ziebuhr J, Hölzer M, Marz M. 2019. Direct RNA
599 nanopore sequencing of full-length coronavirus genomes provides novel insights into structural
600 variants and enables modification analysis. *Genome Res* **29**: 1545-1554.
601 Xiao YL, Liu S, Ge R, Wu Y, He C, Chen M, Tang W. 2023. Transcriptome-wide profiling and quantification
602 of N(6)-methyladenosine by enzyme-assisted adenosine deamination. *Nat Biotechnol* **41**: 993-
603 1003.
604 Xie YY, Zhong ZD, Chen HX, Ren ZH, Qiu YT, Lan YL, Wu F, Kong JW, Luo RJ, Zhang D et al. 2025. Single-
605 molecule direct RNA sequencing reveals the shaping of epitranscriptome across multiple species.
606 *Nat Commun* **16**: 5119.
607 Xu H, Kong L, Cheng J, Al Moussawi K, Chen X, Iqbal A, Wing PAC, Harris JM, Tsukuda S, Embarc-Buh A et
608 al. 2024. Absolute quantitative and base-resolution sequencing reveals comprehensive
609 landscape of pseudouridine across the human transcriptome. *Nature Methods* **21**: 2024-2033.

610 Yang L, Tang L, Min Q, Tian H, Li L, Zhao Y, Wu X, Li M, Du F, Chen Y et al. 2024a. Emerging role of RNA
611 modification and long noncoding RNA interaction in cancer. *Cancer Gene Ther* **31**: 816-830.
612 Yang Y, Lu Y, Wang Y, Wen X, Qi C, Piao W, Jin H. 2024b. Current progress in strategies to profile
613 transcriptomic m(6)A modifications. *Front Cell Dev Biol* **12**: 1392159.
614 Yankova E, Blackaby W, Albertella M, Rak J, De Braekeleer E, Tsagkogeorga G, Pilka ES, Aspris D, Leggate
615 D, Hendrick AG et al. 2021. Small-molecule inhibition of METTL3 as a strategy against myeloid
616 leukaemia. *Nature* **593**: 597-601.
617 Yu J, She Y, Ji S-J. 2021. m6A Modification in Mammalian Nervous System Development, Functions,
618 Disorders, and Injuries. *Frontiers in Cell and Developmental Biology* **Volume 9 - 2021**.
619 Zaccara S, Ries RJ, Jaffrey SR. 2019. Reading, writing and erasing mRNA methylation. *Nature Reviews*
620 *Molecular Cell Biology* **20**: 608-624.
621 Zou Y, Ahsan MU, Chan J, Meng W, Gao S-J, Huang Y, Wang K. 2025. A comparative evaluation of
622 computational models for RNA modification detection using nanopore sequencing with RNA004
623 chemistry. *Briefings in Bioinformatics* **26**.

624

625

626 **Figure legends**

627

628 **Figure 1. Workflow for m^{6}A mapping in different human cell types using orthogonal sequencing**
629 **approaches.**

630 **(A)** Schematic of the experimental setup. RNA was isolated from NHDFs and differentiated HD10.6 cells
631 treated with STM2457 or DMSO for 48h. Poly(A) RNA was subjected to DRS and GLORI from the same
632 input material. IVT RNA was generated in parallel to assess potential false-positive modification calls. **(B)**
633 Workflow for benchmarking Dorado basecalling. Reads were processed with Dorado versions v.0.5.0
634 through v0.9.0, aligned to the human genome or transcriptome, and analyzed with ONT's Modkit tool.
635 Multiple filtering strategies were tested against IVT controls and METTL3 inhibition (STM2457). Resulting
636 m^{6}A sites were compared with GLORI data from the same input RNA. **(C)** Final benchmarked Dorado
637 pipeline to detect high confidence m^{6}A sites.

638

639 **Figure 2. Modification probability of 0.98 and stoichiometry score of ≥ 10% are necessary cutoffs to**
640 **identify m^{6}A sites accurately.**

641 **(A)** Dorado v0.9.0-generated modification probability distributions for all-context m^{6}A from genome-
642 aligned DMSO- and STM2457-treated NHDFs plotted against the ones from IVT NHDF poly(A) RNA.

643 DMSO probability score distribution is shown as a red solid line, STM2457 – red dashed line, IVT – grey
644 solid line. **(B)** Same as in (A), but for DRACH-context m^6A sites. **(C)** Stoichiometry (≥ 10%) distributions for
645 all-context m^6A from DMSO- (in red) or STM2457-treated (in pink) NHDF datasets (genome-aligned) are
646 plotted using a 0.98 modification probability cutoff and a filter for ≥ 20 reads. **(D)** Same as in (C), but for
647 DRACH-context m^6A sites. **(E)** Overlap of m^6A sites between all-context and DRACH models (A basecall
648 accuracy ≥ 80%, coverage ≥ 20reads, stoichiometry ≥ 10%), comparing non-filtered to filtered
649 (modification probability > 0.98) outputs. **(F)** Metagene plots showing the density of m^6A (all-context –
650 dark red, DRACH-context – bright red) sites across all annotated poly(A) transcripts (left) and
651 monoexonic poly(A) transcripts (right) from genome-aligned DMSO-treated NHDF datasets filtered with
652 0.98 modification probability, ≥ 20 coverage and ≥ 10% stoichiometry.

653

654 **Figure 3. Comparison of m^6A sites detected by Dorado and GLORI.**

655 **(A)** Filtering strategy for Dorado vs GLORI to obtain a high-confidence list of m^6A sites in NHDFs. **(B)**
656 Stoichiometry distributions are plotted from genome-aligned DMSO- or STM2457-treated NHDF datasets
657 using the indicated Dorado models and the filtering strategy shown in (A). **(C)** Stoichiometry distributions
658 from GLORI performed on DMSO- or STM2457-treated NHDFs are plotted, using the filtering strategy
659 shown in (A). In addition, m^6A sites were filtered for DRACH motifs (right plot). **(D)** Overlap of high
660 confidence m^6A sites detected by Dorado vs GLORI. GLORI datasets were filtered using increasing m^6A
661 stoichiometry cutoffs (from left to right). **(E)** m^6A sites in the *SPEN* gene detected by GLORI and Dorado.
662 Boxplots (right) represent the m^6A stoichiometries per m^6A site shown on the left. **(F)** *SPEN* isoforms
663 with m^6A sites, determined by Dorado. Boxplots (right) show the distribution of m^6A stoichiometries on
664 the m^6A sites on *SPEN-201* and *SPEN-202*.

665

666 **Figure 4. Comparative analysis of m^6A modification patterns in NHDFs and HD10.6 cells using Dorado**
667 **v0.9.0.**

668 **(A)** Overlap of high confidence m^6A sites between NHDFs and HD10.6 cells. **(B)** Correlation of shared m^6A
669 stoichiometries (n=38,217). **(C)** Boxplots showing the m^6A stoichiometry distributions of the 38,217
670 detected in both NHDFs and HD10.6 cells. **(D+E)** Motif analysis showing m^6A (all-context and DRACH-
671 context) motif logos, 5-mer distributions and modification stoichiometries in NHDFs. **(F+G)** Same as in
672 (D+E) but for HD10.6 cells. **(H)** KEGG pathway analysis (DAVID Bioinformatics Resources) of shared m^6A-

673 containing genes. **(I)** m^6A stoichiometries of representative examples of individual genes from enriched
674 pathways (H). *ACTG1* serves as a housekeeping control.
675

676 **Figure S1. Evaluation of Dorado by base and modification probability distributions.**

677 **(A)** Dorado v0.9.0-generated basecall accuracy probability values for A (all-context and DRACH-context
678 basecaller models) from genome-aligned DMSO- and STM2457-treated NHDF samples against IVT
679 poly(A) NHDF RNA. A probability score distribution is shown as a black solid line for DMSO, STM2457 –
680 black dashed line, IVT – grey solid line. **(B)** Dorado v0.9.0-generated modification probability
681 distributions, visualized as histograms, for DRACH-context m^6A from genome-aligned DMSO- and
682 STM2457-treated NHDFs and IVT NHDFs. IVT modification probability score distribution is shown in grey,
683 DMSO – bright red, STM2457 – dark red. **(C)** Same as in (B), but for all-context m^6A sites.

684

685 **Figure S2. Modification probability cutoffs are necessary to identify Ψ and inosine sites more**
686 **accurately.**

687 **(A)** Dorado v0.9.0-generated basecall accuracy probability values for U (left) and modification probability
688 distributions for Ψ (right) from genome-aligned DMSO- and STM2457-treated NHDF samples are plotted
689 against IVT poly(A) RNA. Ψ probabilities are visualized in blue solid and dashed lines for DMSO and
690 STM2457, respectively and in grey solid line for IVT. **(B)** Same as in (A) but showing A base accuracy
691 probabilities and inosine sites. Inosine probabilities are visualized in purple solid and dashed lines for
692 DMSO and STM2457, respectively and in grey solid line for IVT. **(C)** Same as in (A) but showing C base
693 accuracy probabilities and m^5C sites. m^5C probabilities are visualized in green solid and dashed lines for
694 DMSO and STM2457, respectively and in grey solid line for IVT. **(D+E)** DRACH-context and all-context m^6A
695 modification probability distributions (red) from genome-aligned DMSO- and STM2457-treated NHDFs
696 are plotted against RCS negative control (grey).

697

698 **Figure S3. Modification probability cutoffs are necessary to identify modification sites accurately,**
699 **shown for transcriptome-aligned data.**

700 Dorado v0.9.0-generated basecall accuracy probability values for A, U and C (left, **A-E**) and modification
701 probability distributions for m^6A (all-context), m^6A (DRACH-context), Ψ, inosine and m^5C (right, **A-E**) in

702 transcriptome-aligned DMSO-and STM2457-treated NHDFs, plotted against the values from IVT NHDF
703 poly(A) RNA.

704

705 **Figure S4. Modification probability distribution profiles in genome-aligned DMSO-and STM2457-**
706 **treated HD10.6 cells.**

707 Dorado v0.9.0-generated basecall accuracy probability values for A, U and C (left, **A-E**) and modification
708 probability distributions for m^6A (all-context), m^6A (DRACH-context), Ψ, inosine and m^5C (right, **A-E**) in
709 human genome-aligned DMSO-and STM2457-treated HD10.6 cells, plotted against the values from IVT
710 NHDF poly(A) RNA.

711

712 **Figure S5. Modification probability distribution profiles in transcriptome-aligned DMSO-and STM2457-**
713 **treated HD10.6 cells.**

714 Dorado v0.9.0-generated basecall accuracy probability values for A, U and C (left, **A-E**) and modification
715 probability distributions for m^6A (all-context), m^6A (DRACH-context), Ψ, inosine and m^5C (right, **A-E**) in
716 human transcriptome-aligned DMSO-and STM2457-treated HD10.6 cells, plotted against the values from
717 IVT NHDF poly(A) RNA.

718

719 **Figure S6. Modification site probabilities detected with different versions of Dorado and compared to**
720 **RCS.**

721 **(A+B)** DRACH-context and all-context m^6A modification probability distributions from genome-aligned
722 DMSO (solid red line)- and STM2457 (dashed red line)-treated NHDFs are plotted against RCS negative
723 control (grey). **(C)** Modification probability distributions for all-context m^6A (Dorado v0.7.0, v0.8.0 and
724 v0.9.0), the bottom panel showing the distributions in 0.95-1.00 modification range. **(D)** Same as in (C),
725 but for DRACH-context m^6A (Dorado v0.5.0, v0.6.0, v0.8.0 and v0.9.0). **(E)** Same as in (C), but for Ψ
726 (Dorado v0.7.0, v0.8.0, v0.9.0). All the plots in C-E are generated from genome-aligned DMSO-treated
727 NHDF datasets and are plotted against RCS negative controls.

728

729 **Figure S7. Evaluation of Dorado by stoichiometry distributions.**

730 **(A)** Stoichiometry distributions for all-context m^6A from DMSO- (in red) or STM2457-treated (in pink)
731 NHDF datasets (genome-aligned and filtered for ≥ 20 reads) are plotted, comparing 0.98 modification
732 probability-unfiltered (left) and filtered (right) sites. **(B)** Same as in (A), but for DRACH-context m^6A sites.

733

734 **Figure S8. Stoichiometry distributions of m^6A, Ψ and inosine in IVT datasets.**

735 **(A)** All-context m^6A stoichiometry distributions (all stoichiometries (top) and ≥ 10% stoichiometries
736 (bottom)) in (genome-aligned) IVT RNA from DMSO-treated NHDFs, comparing modification probability-
737 unfiltered and filtered data. The bottom panel shows stoichiometry distributions only for ≥ 10%
738 stoichiometry sites. **(B)** As in (A), but for DRACH-context m^6A. **(C)** As in (A), but for Ψ. **(D)** As in (A), but
739 for inosine.

740

741 **Figure S9. Evaluation of Dorado by stoichiometry distributions in transcriptome-aligned NHDFs.**

742 **(A)** Stoichiometry distributions for all-context m^6A from DMSO- (in red) or STM2457-treated (in pink)
743 NHDF datasets (transcriptome-aligned and filtered for ≥ 20 reads) are plotted, assessing the 0.98
744 modification probability cutoff. The bottom panel shows stoichiometry distributions only for ≥ 10%
745 stoichiometry sites. **(B)** Same as in (A), but for DRACH-context m^6A sites.

746

747 **Figure S10. Evaluation of Dorado by stoichiometry distributions in transcriptome-aligned HD10.6 cells.**

748 **(A)** Stoichiometry distributions for all-context m^6A from DMSO- (in red) or STM2457-treated (in pink)
749 HD10.6 datasets (transcriptome-aligned and filtered for ≥ 20 reads) are plotted, assessing the 0.98
750 modification probability cutoff. The bottom panel shows stoichiometry distributions only for ≥ 10%
751 stoichiometry sites. **(B)** Same as in (A), but for DRACH-context m^6A sites.

752

753 **Figure S11 Stoichiometry distributions of Ψ and inosine in NHDFs.**

754 **(A)** Distributions of all (left) and ≥ 10% (right) Ψ stoichiometries in the 0.99 modification probability-
755 unfiltered and filtered data from genome-aligned DMSO- (darker color) and STM2457-treated (lighter
756 color) NHDFs (filtered for ≥ 20 reads). **(B)** Same as in (A), but for inosine.

757

758 **Figure S12. Stoichiometry distributions of Ψ and inosine in HD10.6 cells.**

759 **(A)** Distributions of all and ≥ 10% Ψ stoichiometries in the 0.99 modification probability-unfiltered and
760 filtered data from genome-aligned DMSO- (darker color) and STM2457-treated (lighter color) HD10.6
761 cells (filtered for ≥ 20 reads). **(B)** Same as in (A), but for inosine.

762

763 **Figure S13. m^6A and Ψ stoichiometry distributions in different versions of Dorado.**

764 Stoichiometry distributions (all stoichiometries – left, ≥ 10% stoichiometries – right) in a line graph
765 format for **(A)** all-context m^6A detected in genome-aligned DMSO-treated NHDF dataset, with Dorado
766 versions 0.7.0, 0.8.0 and 0.9.0 colored in gradients of red, darkness increasing with the later versions. **(B)**
767 DRACH-context m^6A detected with Dorado versions 0.5.0, 0.6.0, 0.8.0 and 0.9.0, colored in gradients of
768 red, darkness increasing with the later versions. **(C)** Ψ detected with Dorado versions 0.7.0, 0.8.0 and
769 0.9.0, colored in gradients of blue, darkness increasing with the later versions. All the datasets were
770 filtered for ≥ 20 reads.

771

772 **Figure S14. Comparison of modification sites detected with different versions of Dorado.**

773 Upset plots showing unique and shared detected modifications (filtered for ≥ 20 reads, ≥ 10%
774 stoichiometry) in modification probability-unfiltered and filtered genome-aligned DMSO-treated NHDF
775 data. The versions are ordered on the plot by the number of detected sites, in ascending order (on the
776 left). Results are shown for the following models and versions: **(A)** all-context m^6A detected with v.0.7.0,
777 v0.8.0 and v0.9.0, **(B)** DRACH-context m^6A detected with v.0.5.0, v0.6.0, v0.8.0 and v0.9.0, **(C)** Ψ detected
778 with v.0.7.0, v0.8.0 and v0.9.0. **(D)** Nucleotide frequencies in the 5-base motif of Ψ sites for uniquely
779 v0.7.0-, v0.8.0-0.9.0- and v0.7.0-0.8.0-0.9.0-detected Ψ sites. Position 0 marks the modification site.

780

781 **Figure S15. Comparison of m^6A sites detected by Dorado and GLORI-Sequencing in NHDFs and**
782 **HD10.6s.**

783 **(A)** Upset plot showing the number of m^6A sites detected by GLORI and Dorado in NHDFs and their
784 overlaps. Dorado sites represent high-confidence m^6A sites obtained after filtering DRACH-model

785 detected data. Similarly, only high confidence m^6A sites detected by GLORI are shown but were
786 additionally filtered for DRACH sites. **(B)** Stoichiometry distributions are plotted from genome-aligned
787 DMSO- or STM2457-treated HD10.6 datasets using the indicated Dorado models. **(C)** Stoichiometry
788 distributions from GLORI performed on DMSO- or STM2457-treated NHDFs are plotted. In addition, m^6A
789 sites were filtered for DRACH motifs (right plot). **(D)** Stoichiometry distributions are plotted from
790 transcriptome-aligned DMSO- or STM2457-treated NHDF datasets using the indicated Dorado models.
791 **(E)** Stoichiometry distributions are plotted from transcriptome-aligned DMSO- or STM2457-treated
792 HD10.6 datasets using the indicated Dorado models.

793

794 **Figure S16. m^6A (all-context and DRACH-context) motif logos for modification probability-unfiltered**
795 **NHDF and HD10.6 data.**

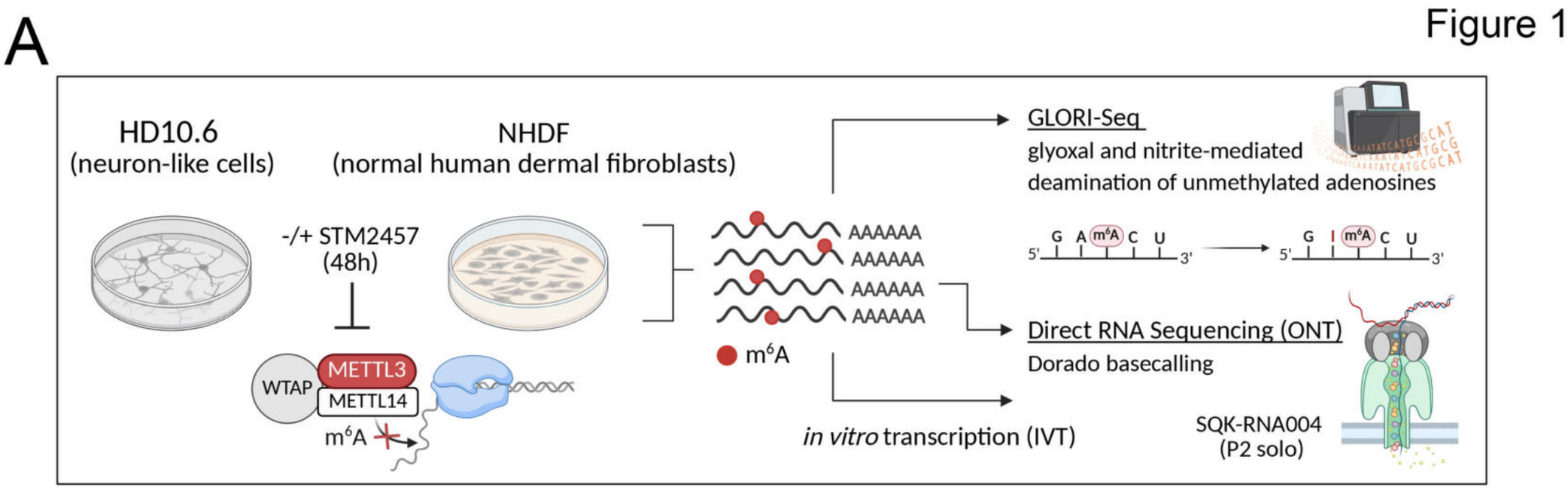
796 **(A)** 5-mer motif logo (showing nucleotide frequences at each position) of m^6A-modified sites (position 0
797 = modified A) from genome-aligned DMSO-treated NHDFs (all-context – left, DRACH-context – right). All
798 the datasets were filtered for ≥ 20 coverage & ≥ 10% stoichiometry. **(B)** Same as (A), but for HD10.6 cells.

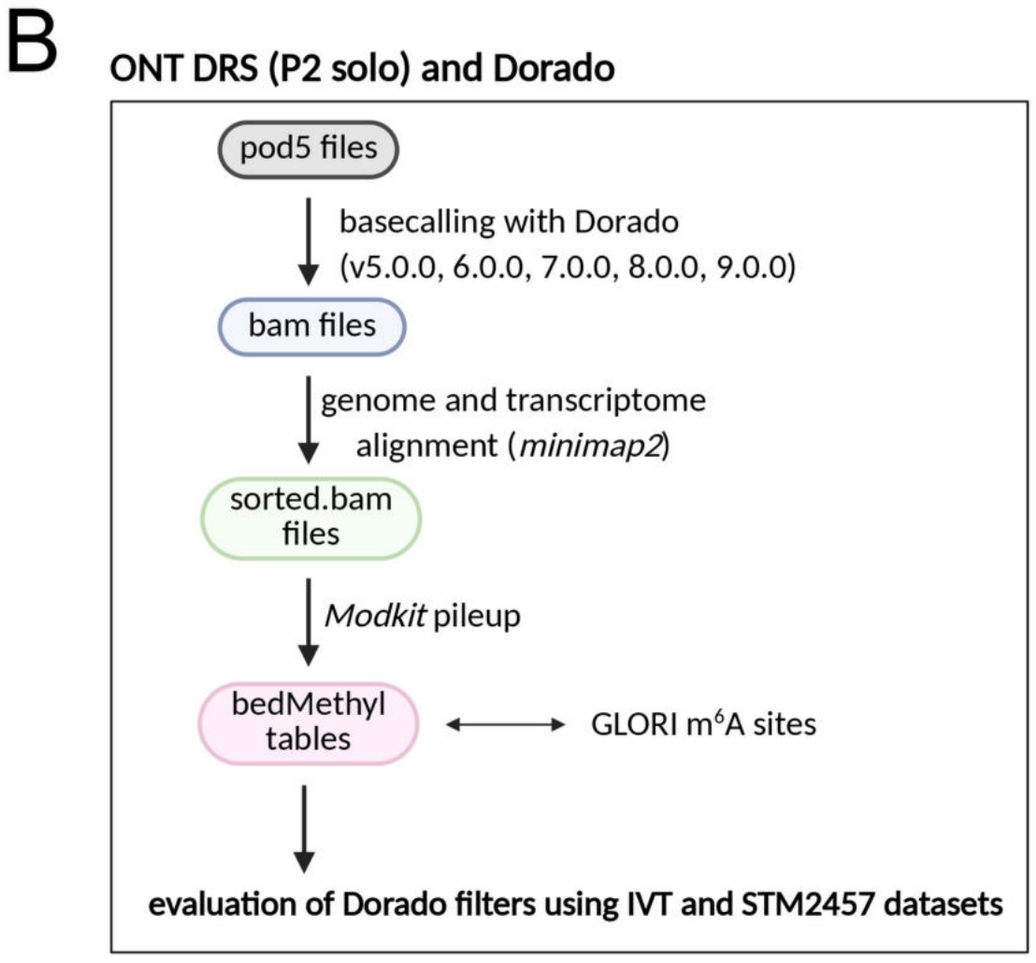
799

800 **Figure S17. Metagene plots showing m^6A (all-context and DRACH-context) distribution profiles in**
801 **NHDF and HD10.6 data filtered at different stoichiometries.**

802 **(A)** Metagene plots showing the density of m^6A (all-context – dark red, DRACH-context – bright red) sites
803 across all annotated poly(A) transcripts (upper plot) and monoexonic poly(A) transcripts (lower plot)
804 from genome-aligned DMSO-treated NHDF datasets filtered with 0.98 modification probability, ≥ 20
805 coverage and ≥ 10% stoichiometry. **(B)** The density of all-context m^6A sites (0.98 modification probability-
806 unfiltered – left, filtered – right) is plotted across all annotated poly(A) transcripts of NHDF (dark red) and
807 HD10.6 (bright red) datasets, filtered with different stoichiometries (descending: all stoichiometries, ≥
808 50%, ≥ 10% and < 10%. All datasets were filtered additionally for ≥ 20 coverage & ≥ 10% stoichiometry.
809 **(C)** Same as (A), but for DRACH-context m^6A sites.

Figure 1


A
HD10.6
(neuron-like cells)
NHDF
(normal human dermal fibroblasts)
-/+ STM2457
(48h)
WTAP
METTL3
METTL14
m6A
AAAAAA
m6A
GLORI-Seq
glyoxal and nitrite-mediated
deamination of unmethylated adenosines
G A m6A C U
G I m6A C U
Direct RNA Sequencing (ONT)
Dorado basecalling
in vitro transcription (IVT)
SQK-RNA004
(P2 solo)


B
ONT DRS (P2 solo) and Dorado
pod5 files
basecalling with Dorado
(v5.0.0, 6.0.0, 7.0.0, 8.0.0, 9.0.0)
bam files
genome and transcriptome
alignment (minimap2)
sorted.bam
files
Modkit pileup
bedMethyl
tables
GLORI m6A sites
evaluation of Dorado filters using IVT and STM2457 datasets

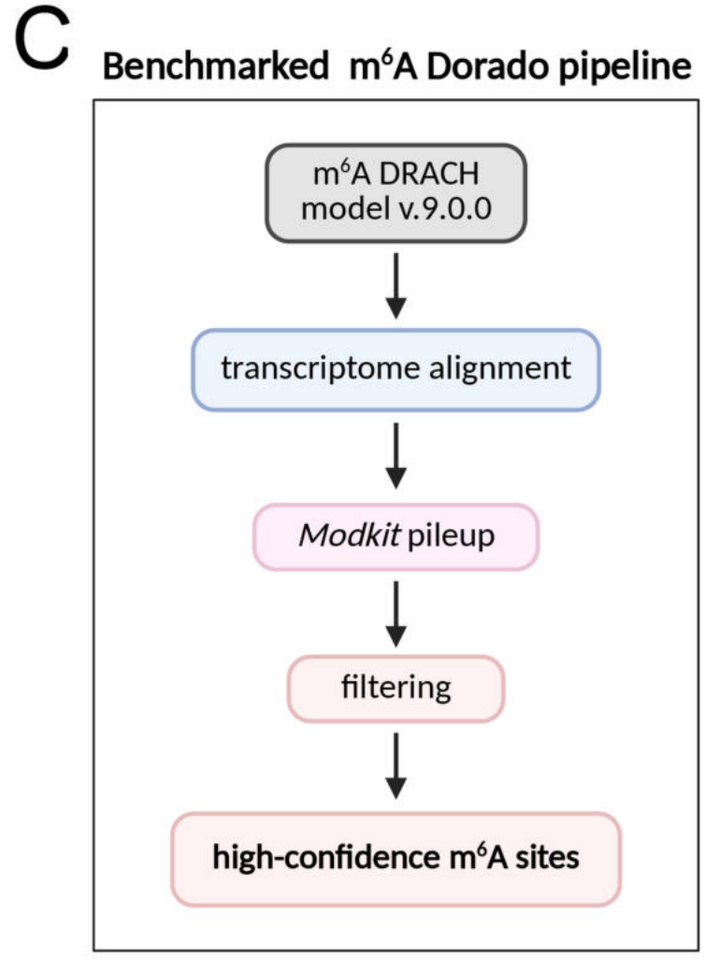
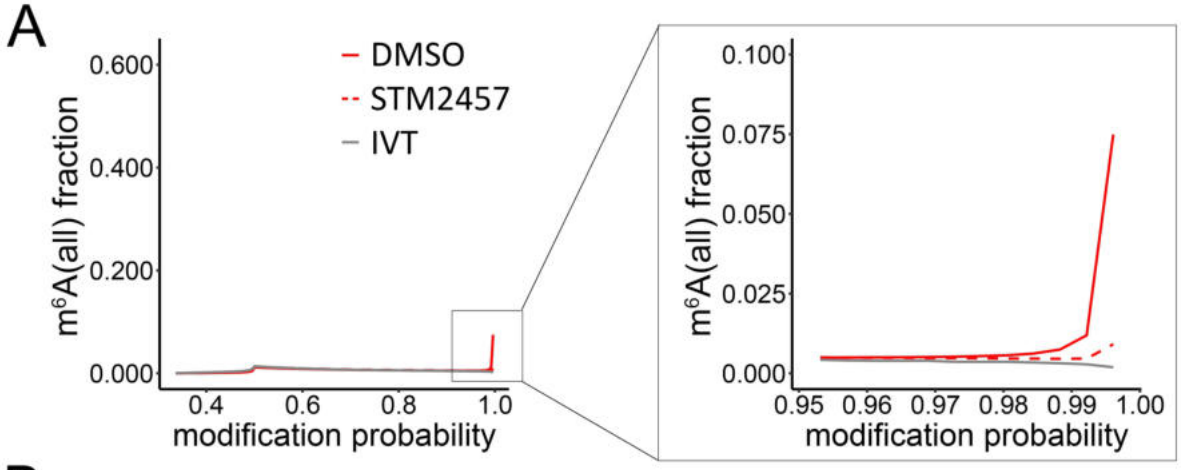
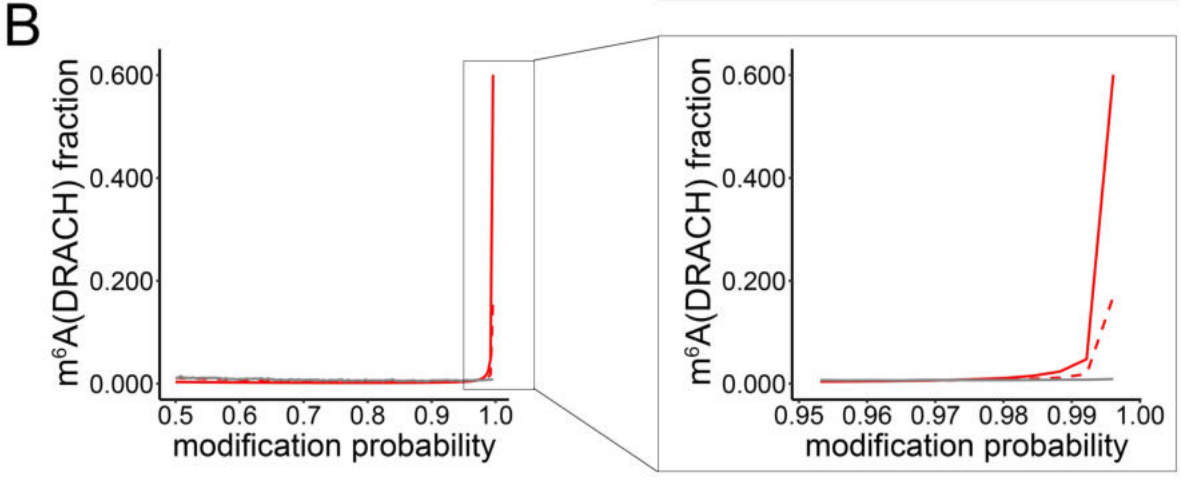
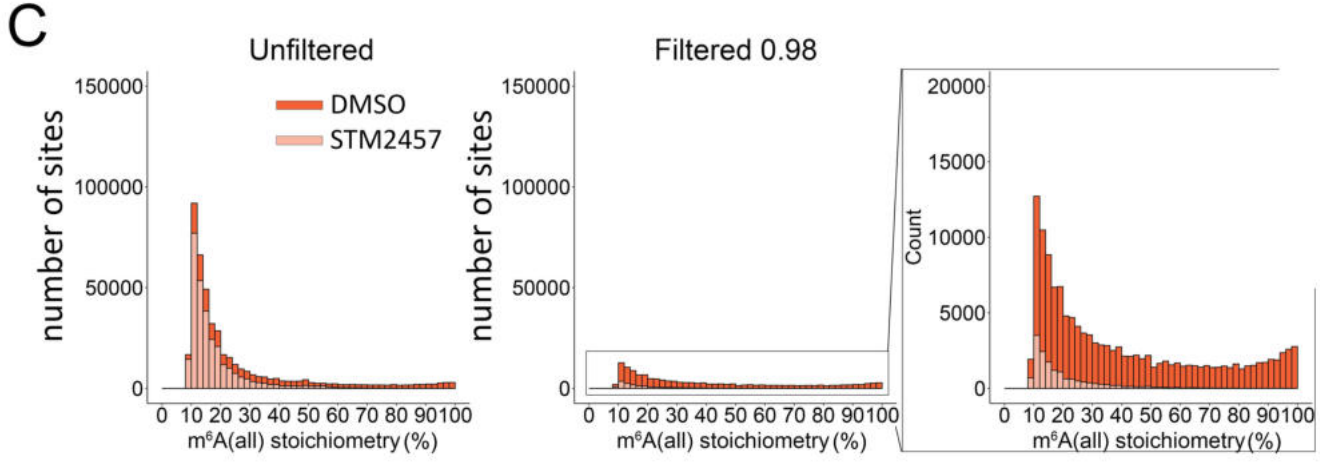
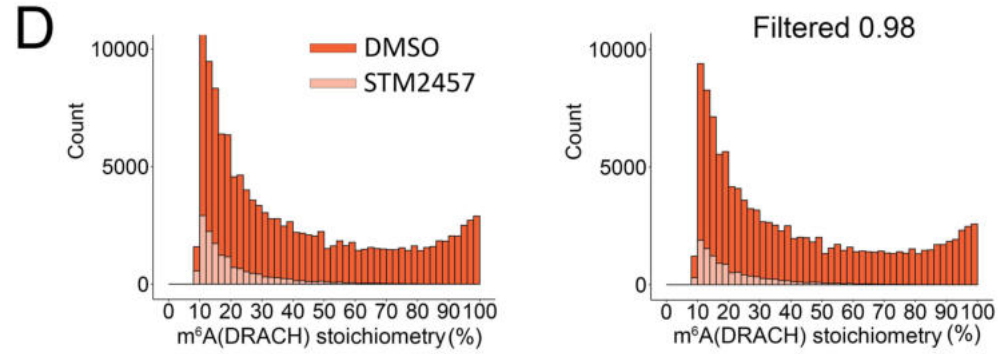

C
Benchmarked m6A Dorado pipeline
m6A DRACH
model v.9.0.0
transcriptome alignment
Modkit pileup
filtering
high-confidence m6A sites

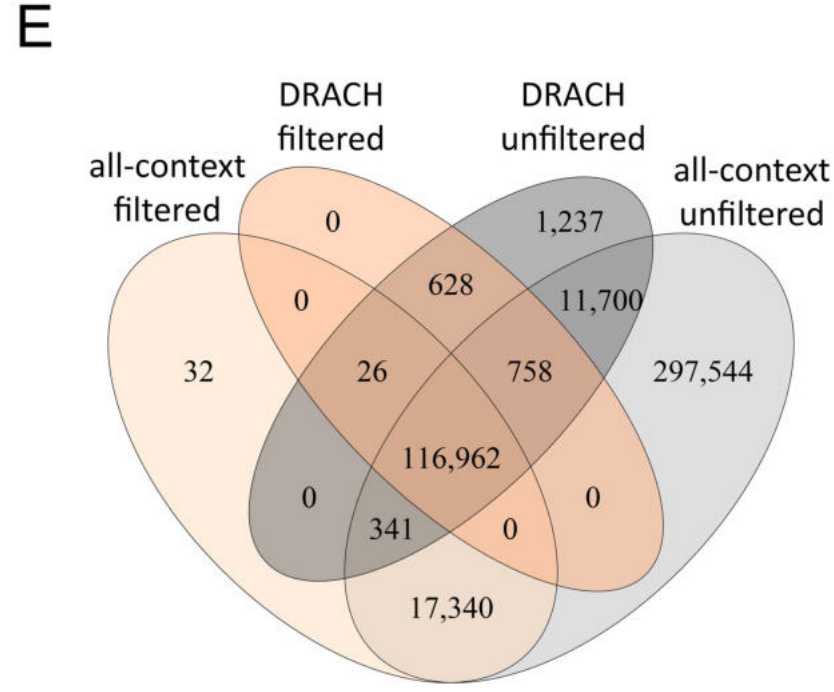
Figure 2


A
DMSO
STM2457
IVT
m6A(all) fraction
modification probability


B
m6A(DRACH) fraction
modification probability


C
Unfiltered
Filtered 0.98
DMSO
STM2457
number of sites
Count
m6A(all) stoichiometry (%)


D
DMSO
STM2457
Filtered 0.98
Count
m6A(DRACH) stoichiometry (%)


E
all-context filtered
DRACH filtered
DRACH unfiltered
all-context unfiltered
0
1,237
628
11,700
0
32
26
758
297,544
116,962
0
0
341
0
17,340

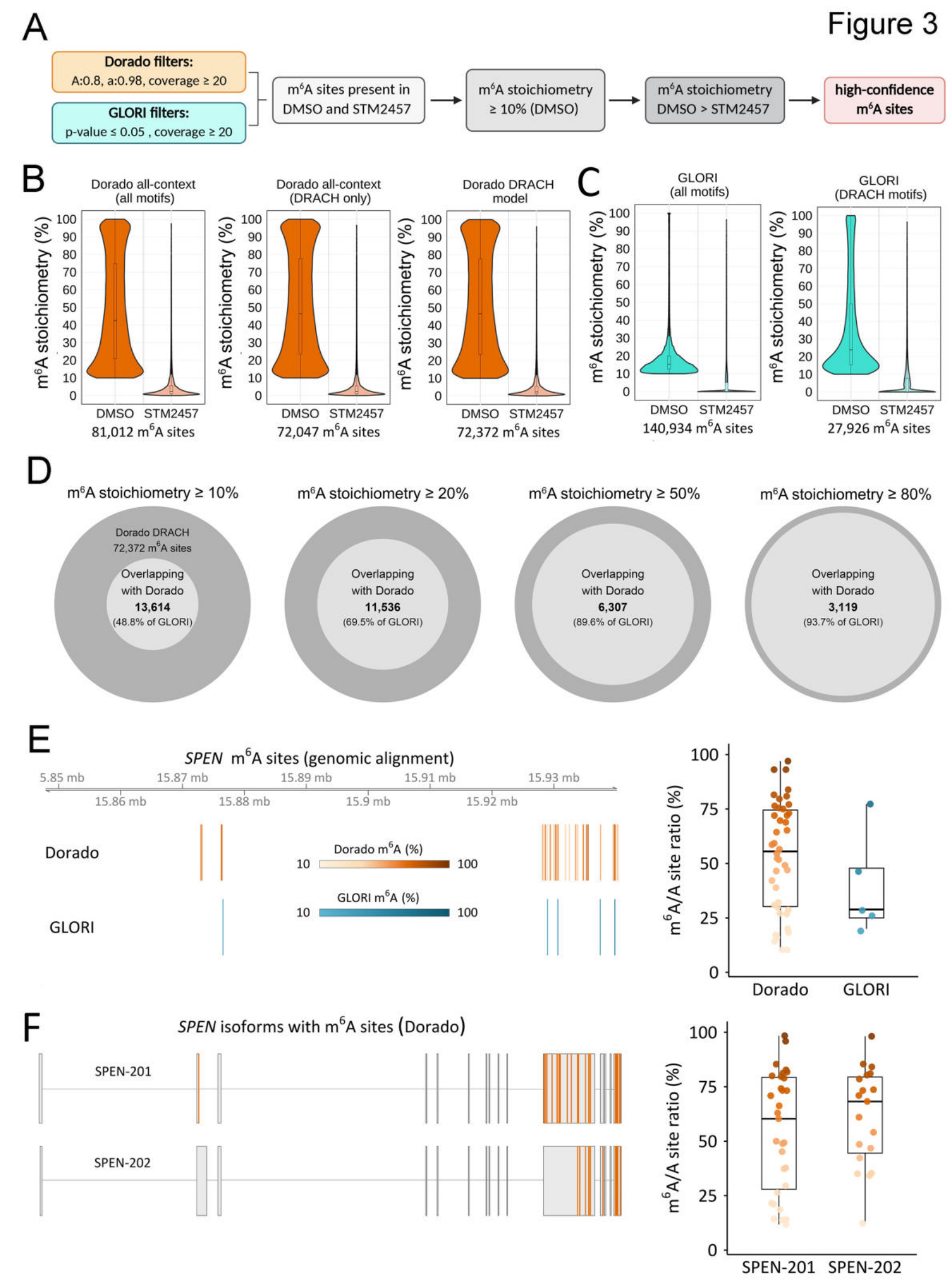
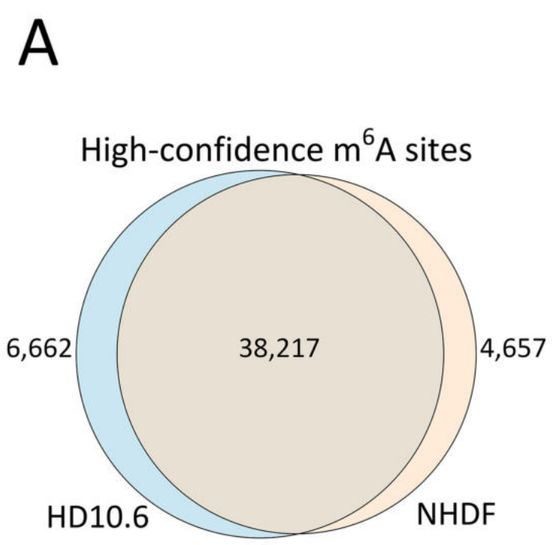
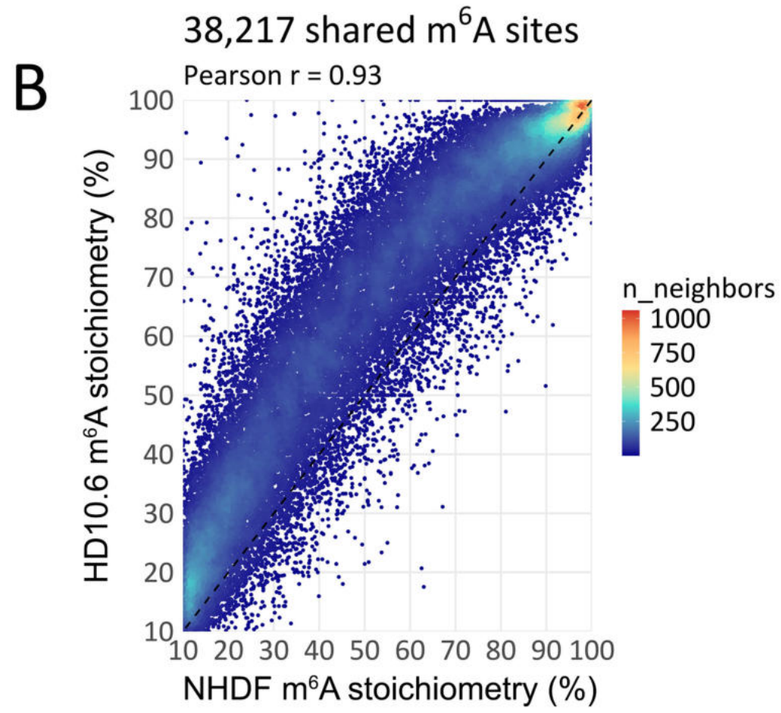
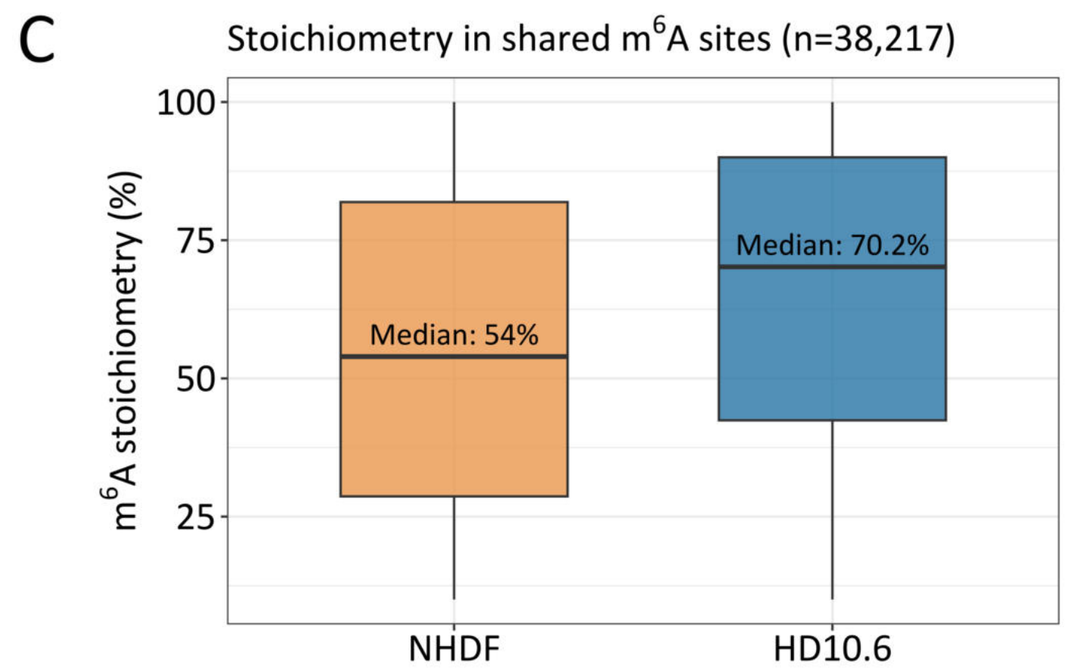


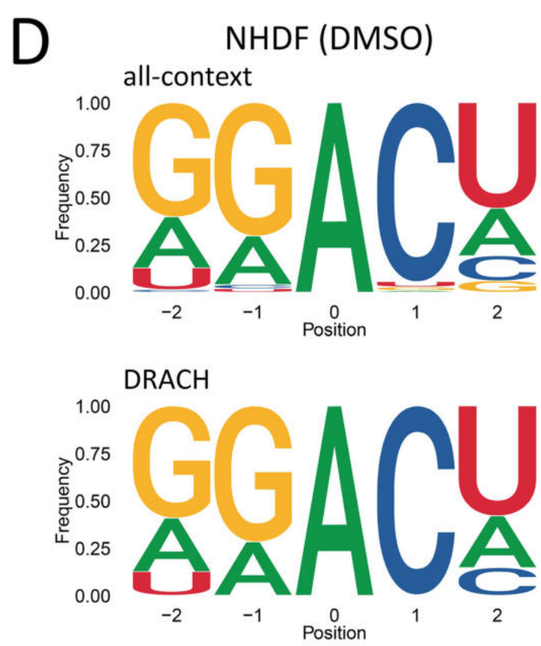
Figure 3
A
Dorado filters:
A:0.8, a:0.98, coverage ≥ 20
GLORI filters:
p-value ≤ 0.05 , coverage ≥ 20
m6A sites present in DMSO and STM2457
m6A stoichiometry ≥ 10% (DMSO)
m6A stoichiometry DMSO > STM2457
high-confidence m6A sites
B
Dorado all-context (all motifs)
Dorado all-context (DRACH only)
Dorado DRACH model
m6A stoichiometry (%)
DMSO STM2457
81,012 m6A sites
72,047 m6A sites
72,372 m6A sites
C
GLORI (all motifs)
GLORI (DRACH motifs)
140,934 m6A sites
27,926 m6A sites
D
m6A stoichiometry ≥ 10%
m6A stoichiometry ≥ 20%
m6A stoichiometry ≥ 50%
m6A stoichiometry ≥ 80%
Dorado DRACH 72,372 m6A sites
Overlapping with Dorado 13,614 (48.8% of GLORI)
Overlapping with Dorado 11,536 (69.5% of GLORI)
Overlapping with Dorado 6,307 (89.6% of GLORI)
Overlapping with Dorado 3,119 (93.7% of GLORI)
E
SPEN m6A sites (genomic alignment)
5.85 mb 15.87 mb 15.89 mb 15.91 mb 15.93 mb
15.86 mb 15.88 mb 15.9 mb 15.92 mb
Dorado
GLORI
Dorado m6A (%)
GLORI m6A (%)
m6A/A site ratio (%)
Dorado GLORI
F
SPEN isoforms with m6A sites (Dorado)
SPEN-201
SPEN-202
m6A/A site ratio (%)
SPEN-201 SPEN-202

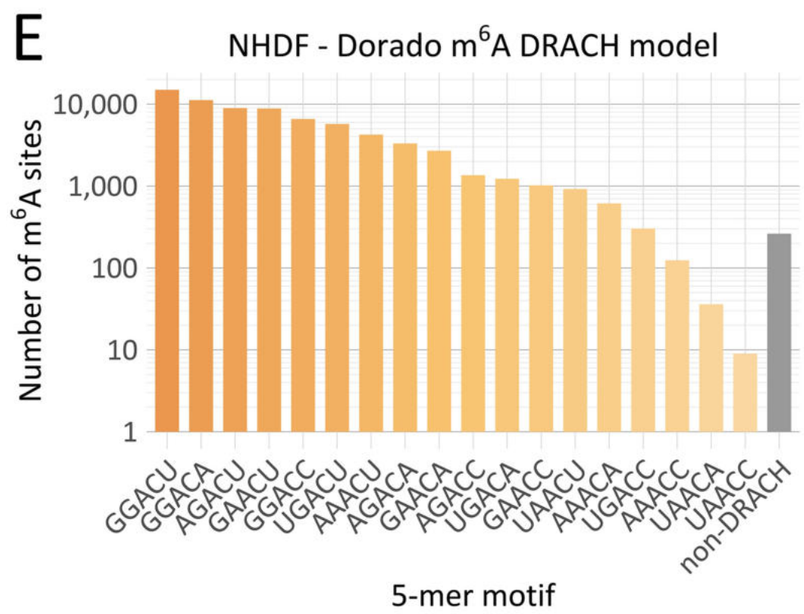
Figure 4

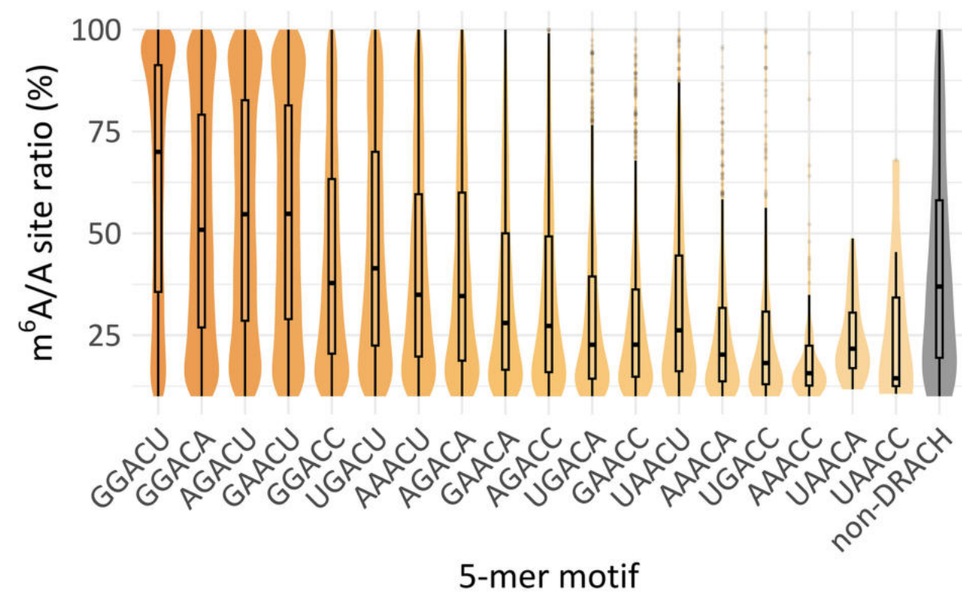

A
High-confidence m6A sites
6,662
38,217
4,657
HD10.6
NHDF

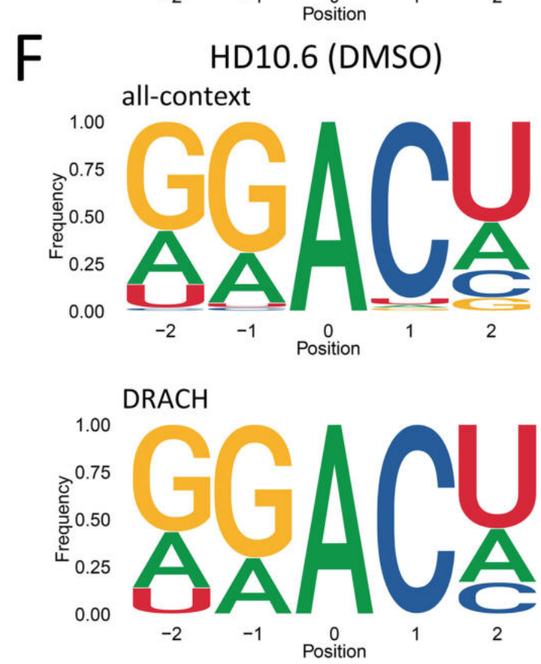

B
38,217 shared m6A sites
Pearson r = 0.93
HD10.6 m6A stoichiometry (%)
NHDF m6A stoichiometry (%)
n_neighbors
1000
750
500
250

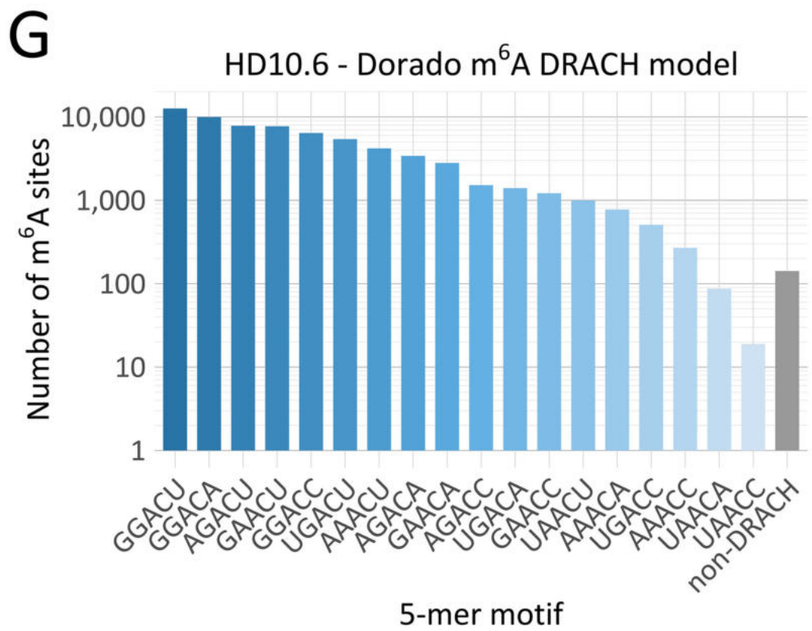

C
Stoichiometry in shared m6A sites (n=38,217)
m6A stoichiometry (%)
Median: 54%
Median: 70.2%
NHDF
HD10.6

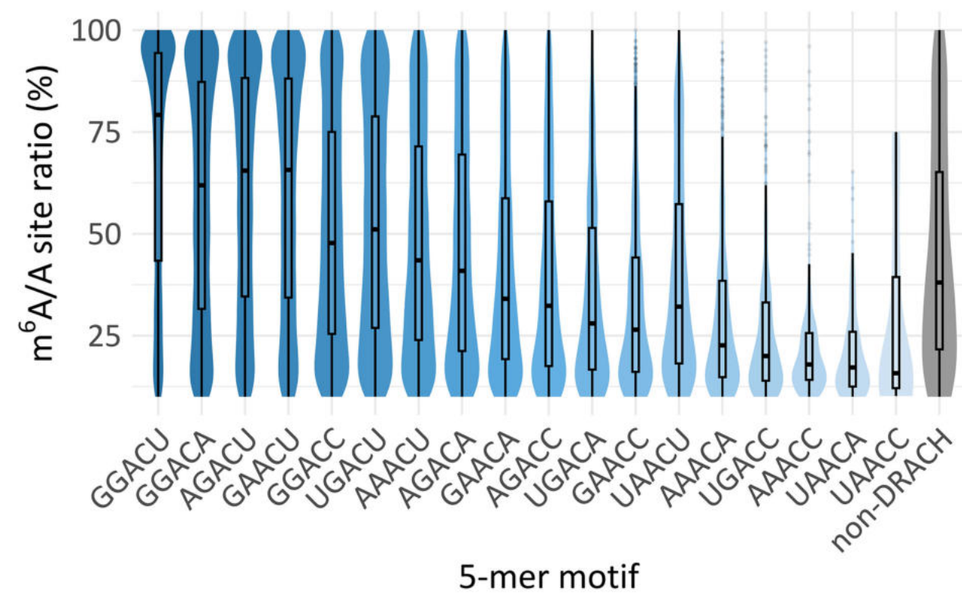

D
NHDF (DMSO)
all-context
DRACH
Frequency
Position

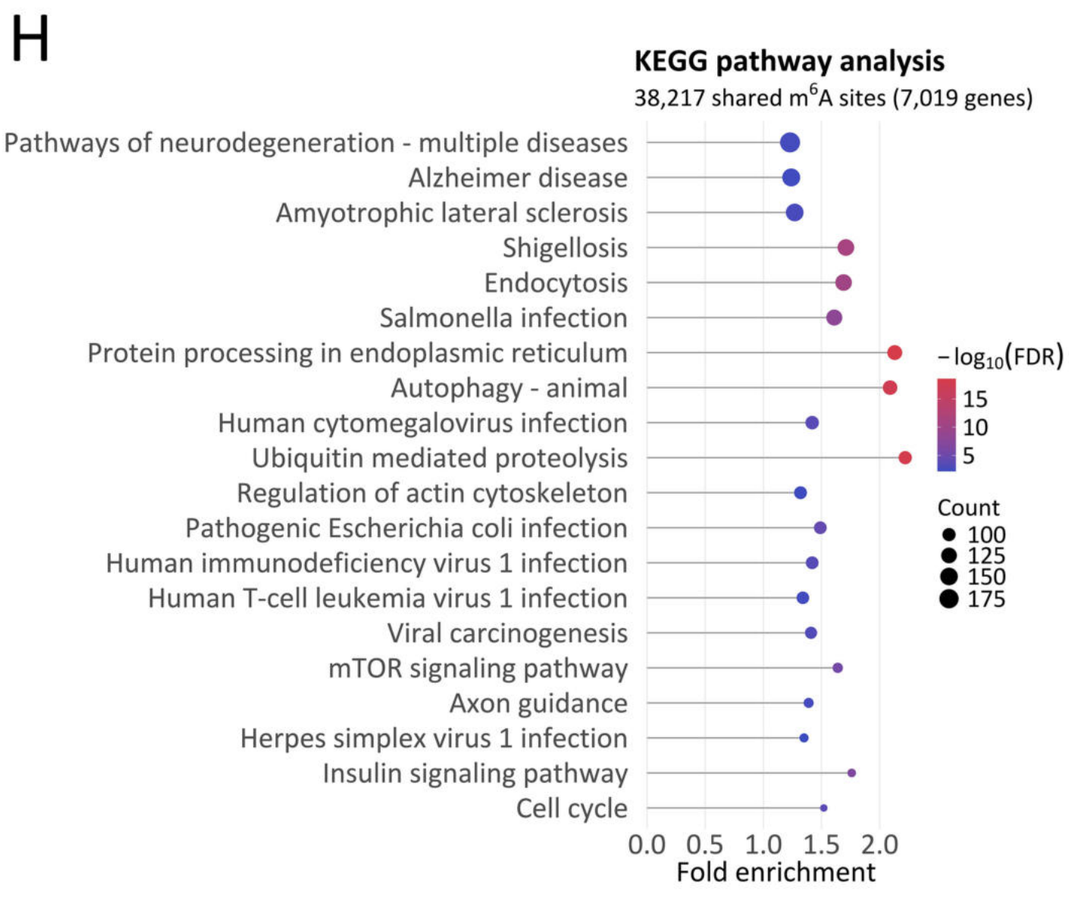

E
NHDF - Dorado m6A DRACH model
Number of m6A sites
GGACU
GGACA
AGACU
GAACU
GGACC
UGACU
AAACU
AGACA
GAACA
AGACC
UGACA
GAACC
UAACU
AAACA
UGACC
AAACC
UAACA
UAACC
non-DRACH
5-mer motif


m6A/A site ratio (%)
5-mer motif


F
HD10.6 (DMSO)
all-context
DRACH
Frequency
Position


G
HD10.6 - Dorado m6A DRACH model
Number of m6A sites
GGACU
GGACA
AGACU
GAACU
GGACC
UGACU
AAACU
AGACA
GAACA
AGACC
UGACA
GAACC
UAACU
AAACA
UGACC
AAACC
UAACA
UAACC
non-DRACH
5-mer motif


m6A/A site ratio (%)
5-mer motif


H
KEGG pathway analysis
38,217 shared m6A sites (7,019 genes)
Pathways of neurodegeneration - multiple diseases
Alzheimer disease
Amyotrophic lateral sclerosis
Shigellosis
Endocytosis
Salmonella infection
Protein processing in endoplasmic reticulum
Autophagy - animal
Human cytomegalovirus infection
Ubiquitin mediated proteolysis
Regulation of actin cytoskeleton
Pathogenic Escherichia coli infection
Human immunodeficiency virus 1 infection
Human T-cell leukemia virus 1 infection
Viral carcinogenesis
mTOR signaling pathway
Axon guidance
Herpes simplex virus 1 infection
Insulin signaling pathway
Cell cycle
Fold enrichment
−log10(FDR)
Count

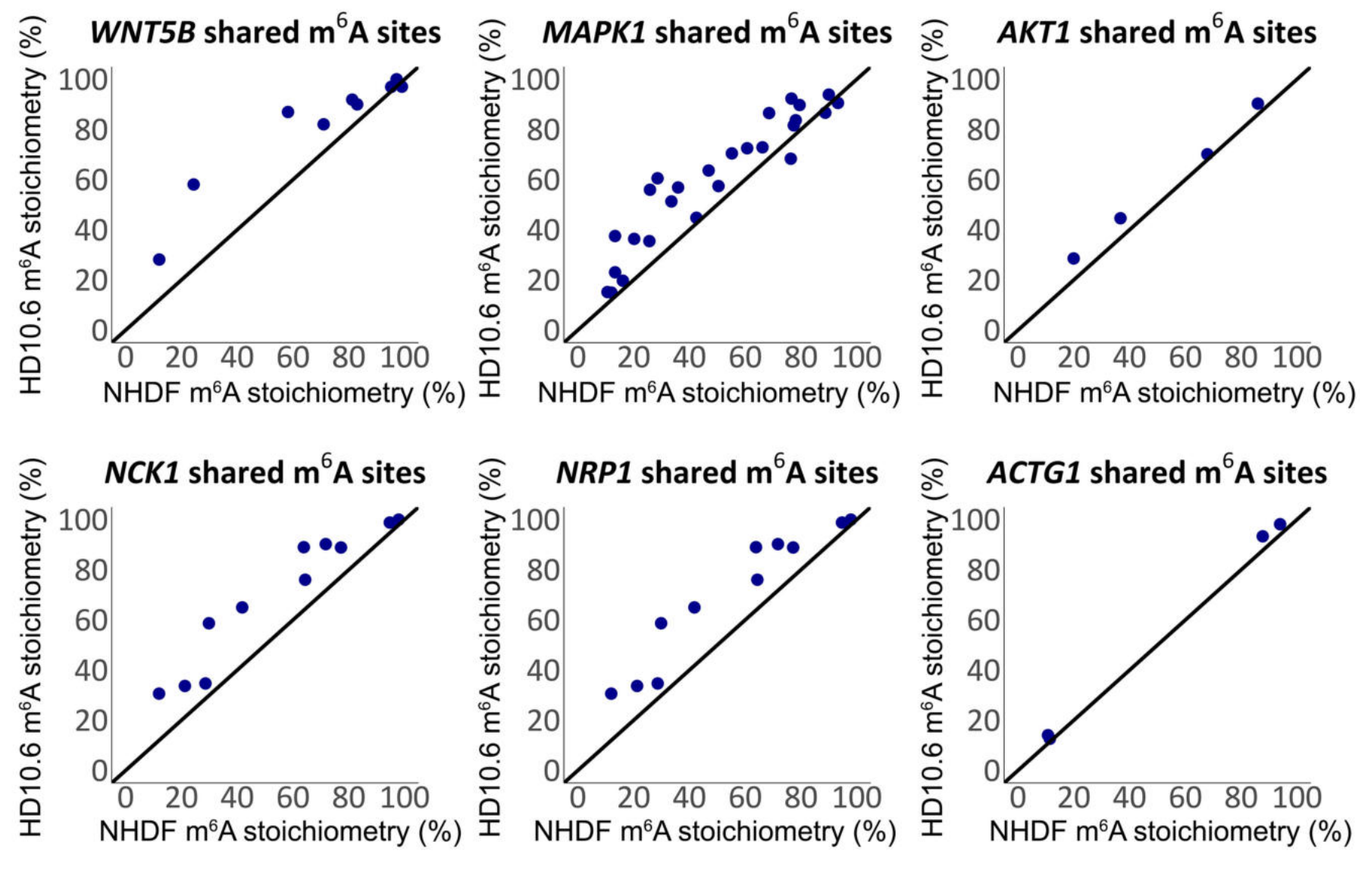

I
WNT5B shared m6A sites
MAPK1 shared m6A sites
AKT1 shared m6A sites
NCK1 shared m6A sites
NRP1 shared m6A sites
ACTG1 shared m6A sites
HD10.6 m6A stoichiometry (%)
NHDF m6A stoichiometry (%)

Figure S1

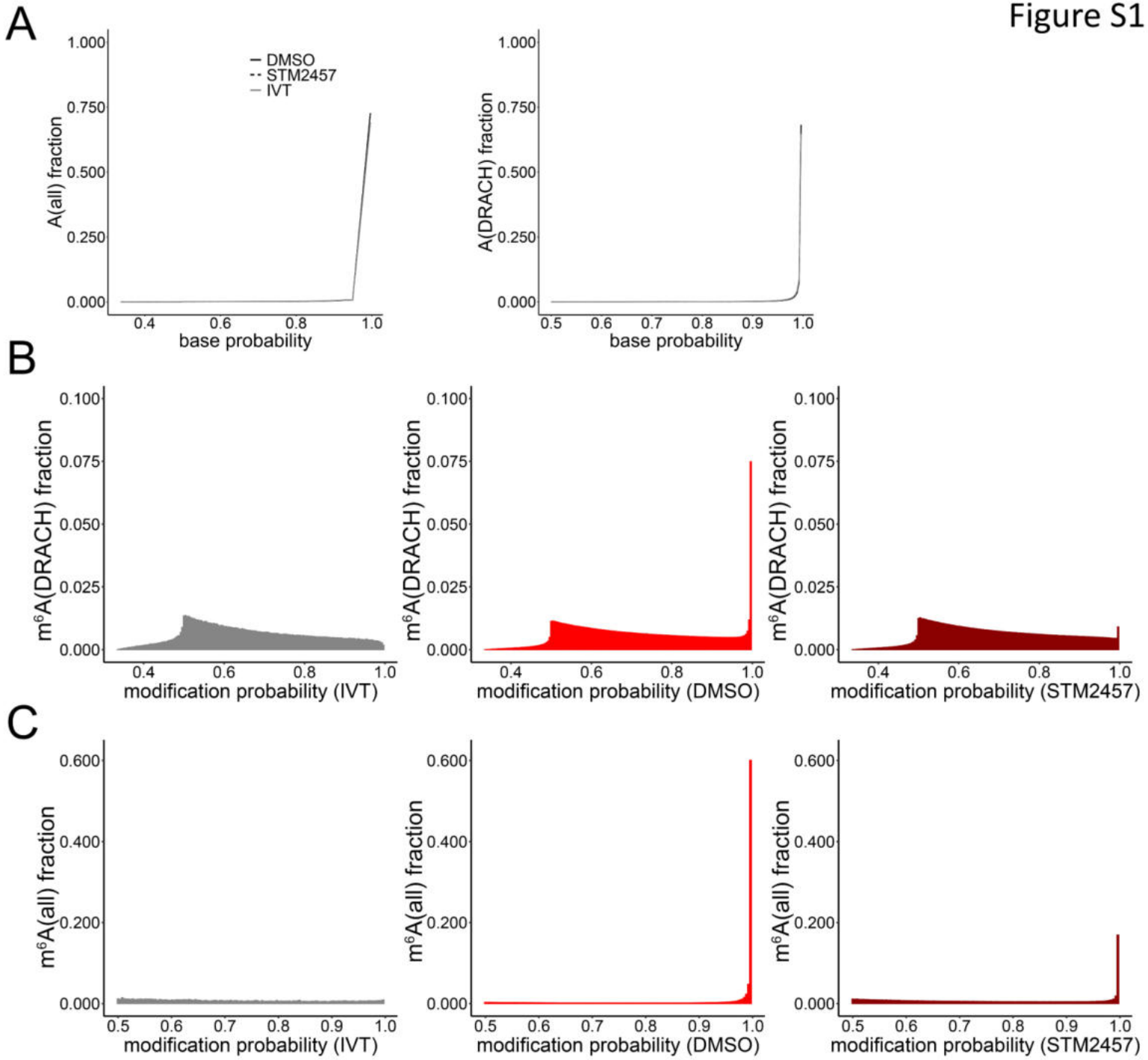

A
DMSO
STM2457
IVT
A(all) fraction
base probability
A(DRACH) fraction
base probability
B
m6A(DRACH) fraction
modification probability (IVT)
modification probability (DMSO)
modification probability (STM2457)
C
m6A(all) fraction
modification probability (IVT)
modification probability (DMSO)
modification probability (STM2457)

Figure S2

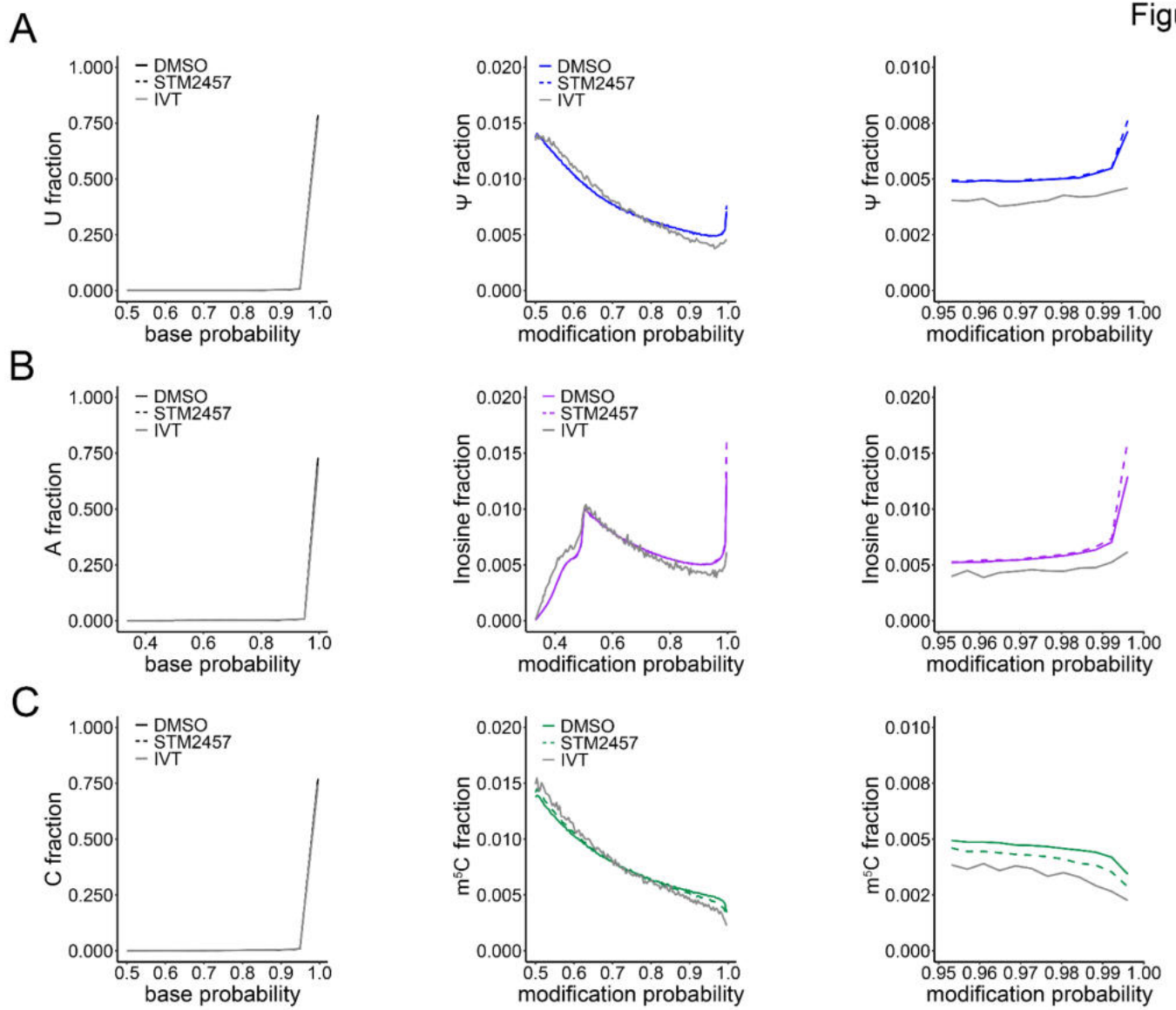

A
DMSO
STM2457
IVT
U fraction
base probability
ψ fraction
modification probability
B
A fraction
Inosine fraction
C
C fraction
m5C fraction

Figure S3

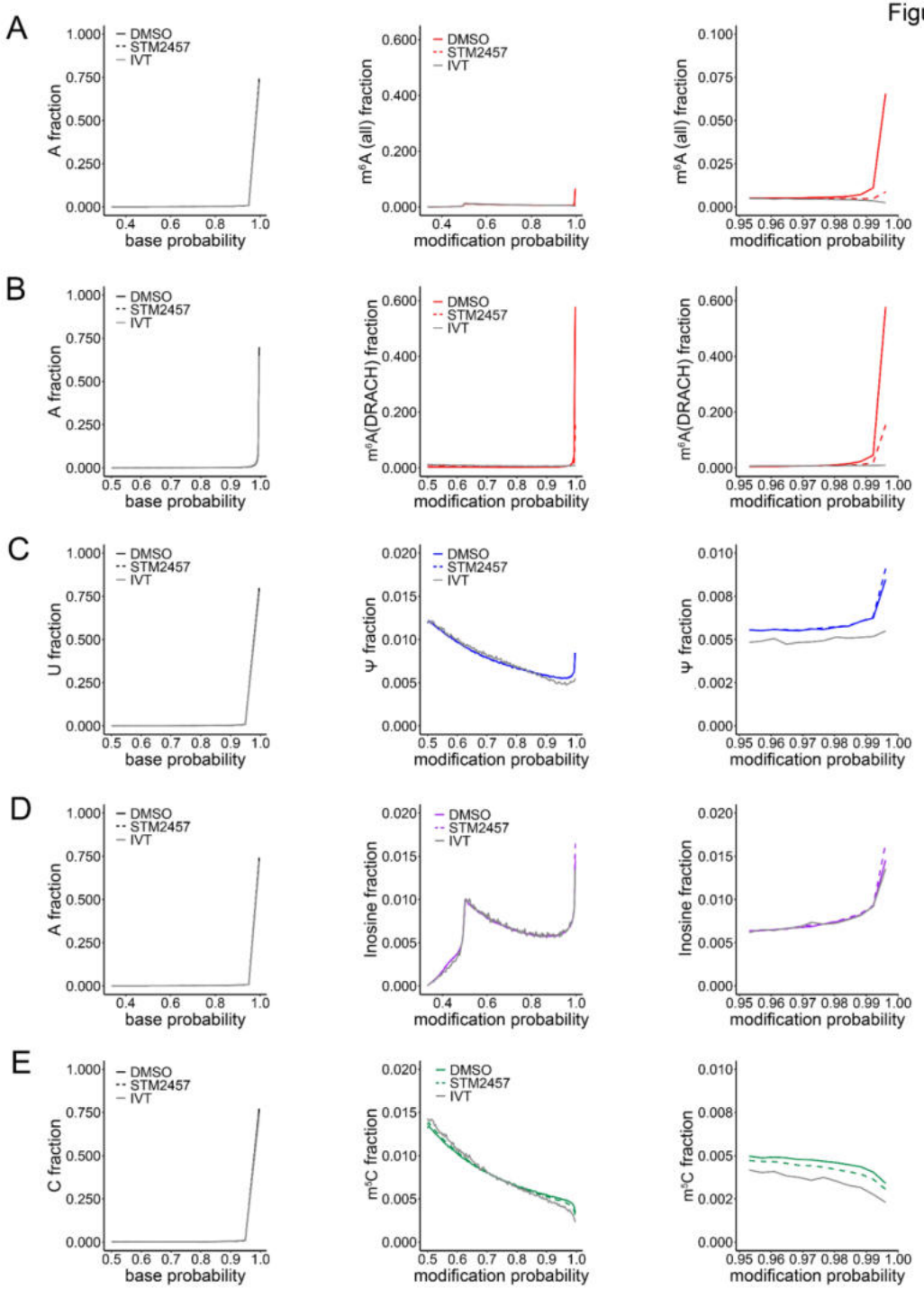

A
B
C
D
E
DMSO
STM2457
IVT
A fraction
m6A (all) fraction
m6A(DRACH) fraction
U fraction
Ψ fraction
Inosine fraction
C fraction
m5C fraction
base probability
modification probability

Figure S4

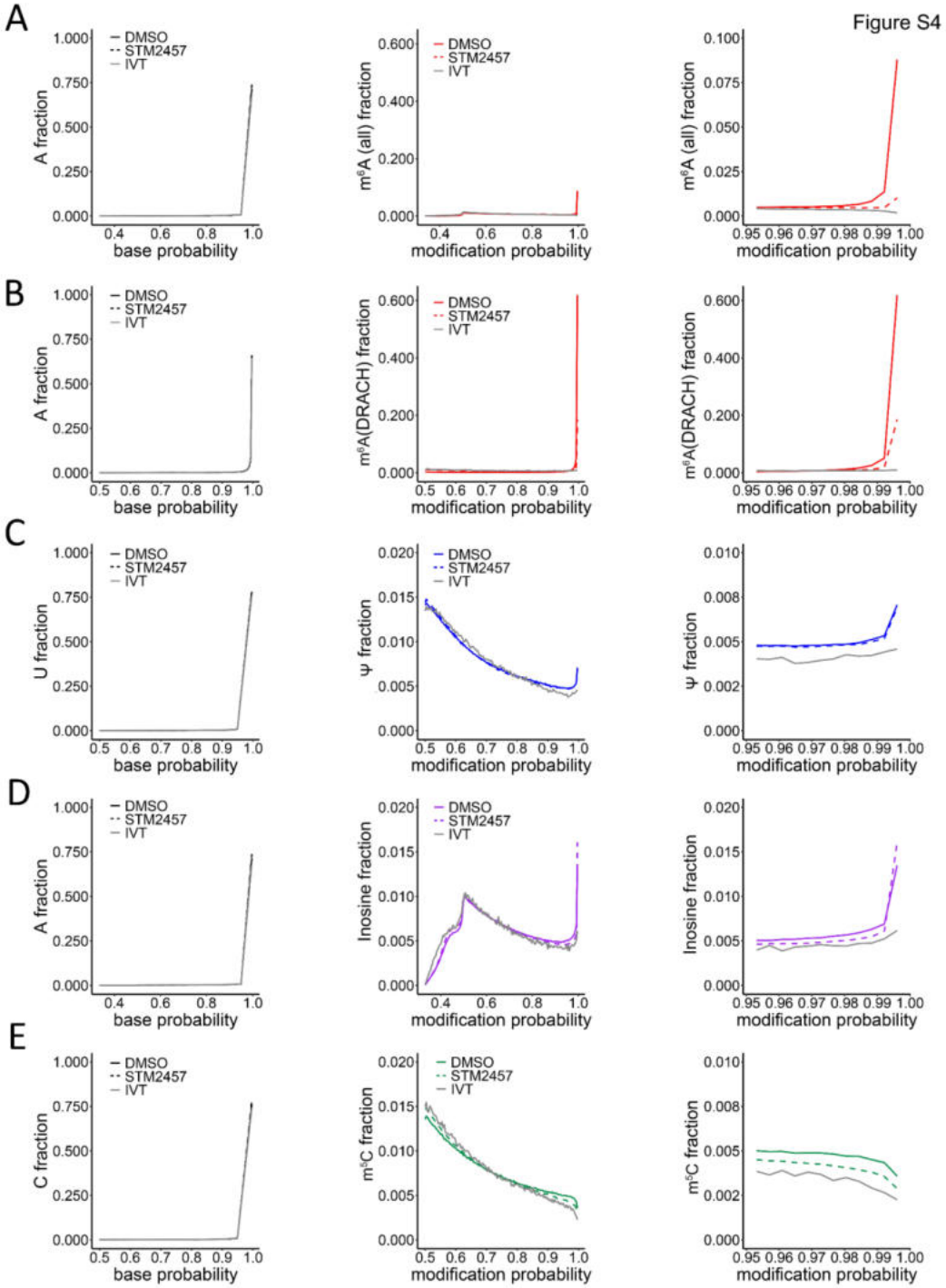

A
B
C
D
E
DMSO
STM2457
IVT
A fraction
m6A (all) fraction
m6A(DRACH) fraction
U fraction
Ψ fraction
Inosine fraction
C fraction
m5C fraction
base probability
modification probability

Figure S5

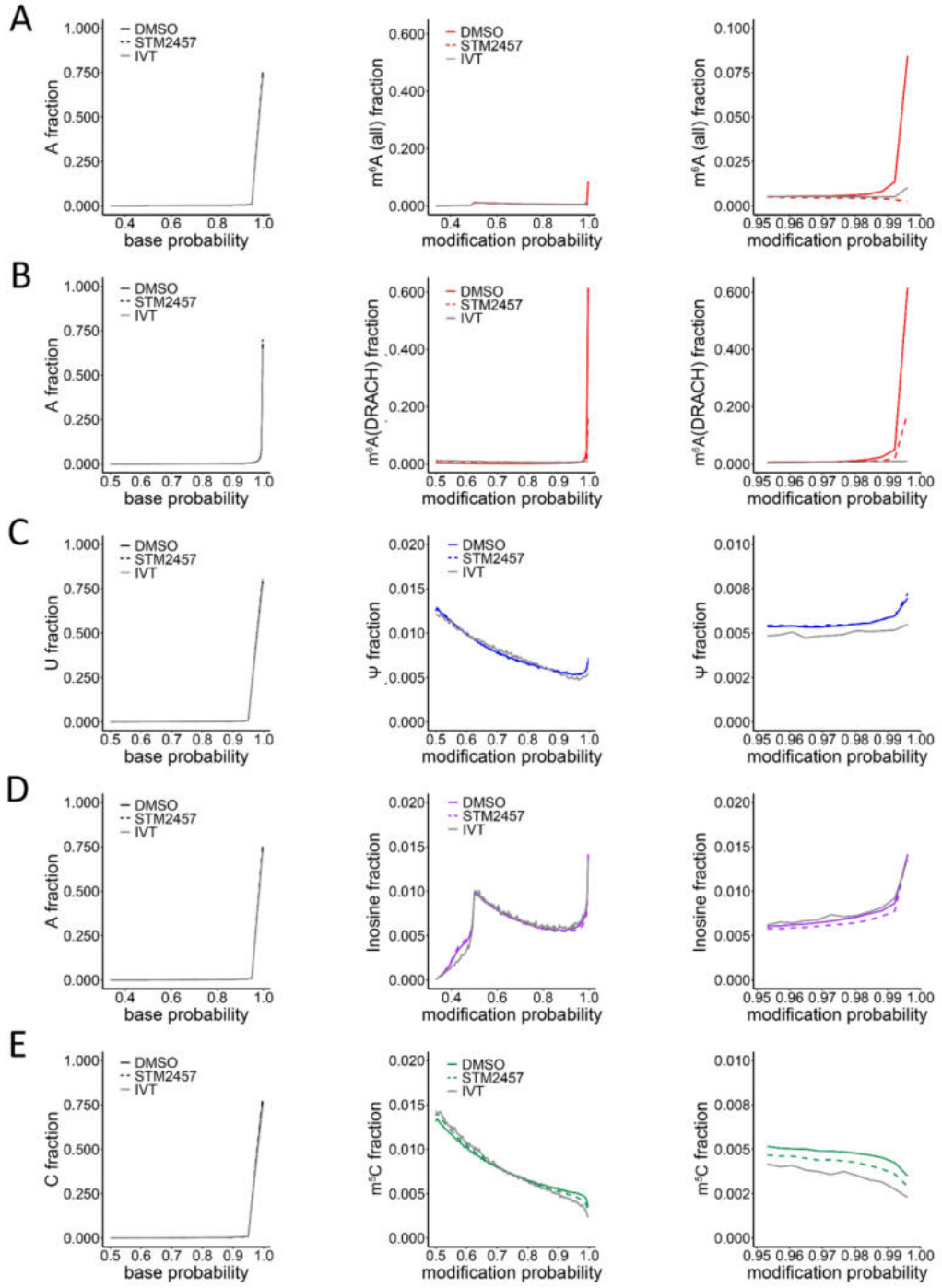

A
DMSO
STM2457
IVT
A fraction
base probability
m6A (all) fraction
modification probability
B
m6A(DRACH) fraction
C
U fraction
Ψ fraction
D
Inosine fraction
E
C fraction
m5C fraction

Figure S6

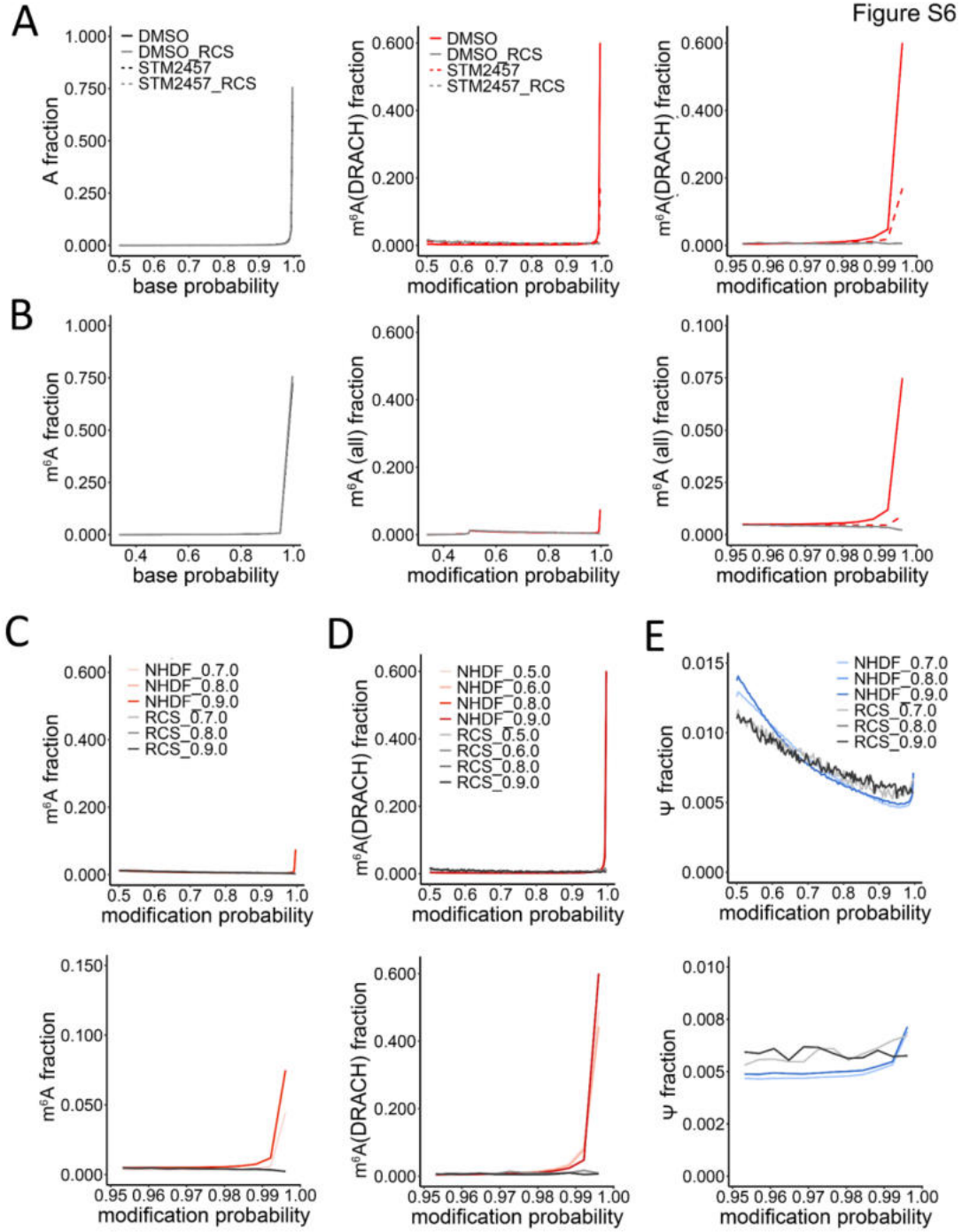

A
DMSO
DMSO_RCS
STM2457
STM2457_RCS
A fraction
base probability
m6A(DRACH) fraction
modification probability
B
m6A fraction
m6A (all) fraction
C
NHDF_0.7.0
NHDF_0.8.0
NHDF_0.9.0
RCS_0.7.0
RCS_0.8.0
RCS_0.9.0
D
NHDF_0.5.0
NHDF_0.6.0
RCS_0.5.0
RCS_0.6.0
E
ψ fraction

Figure S7

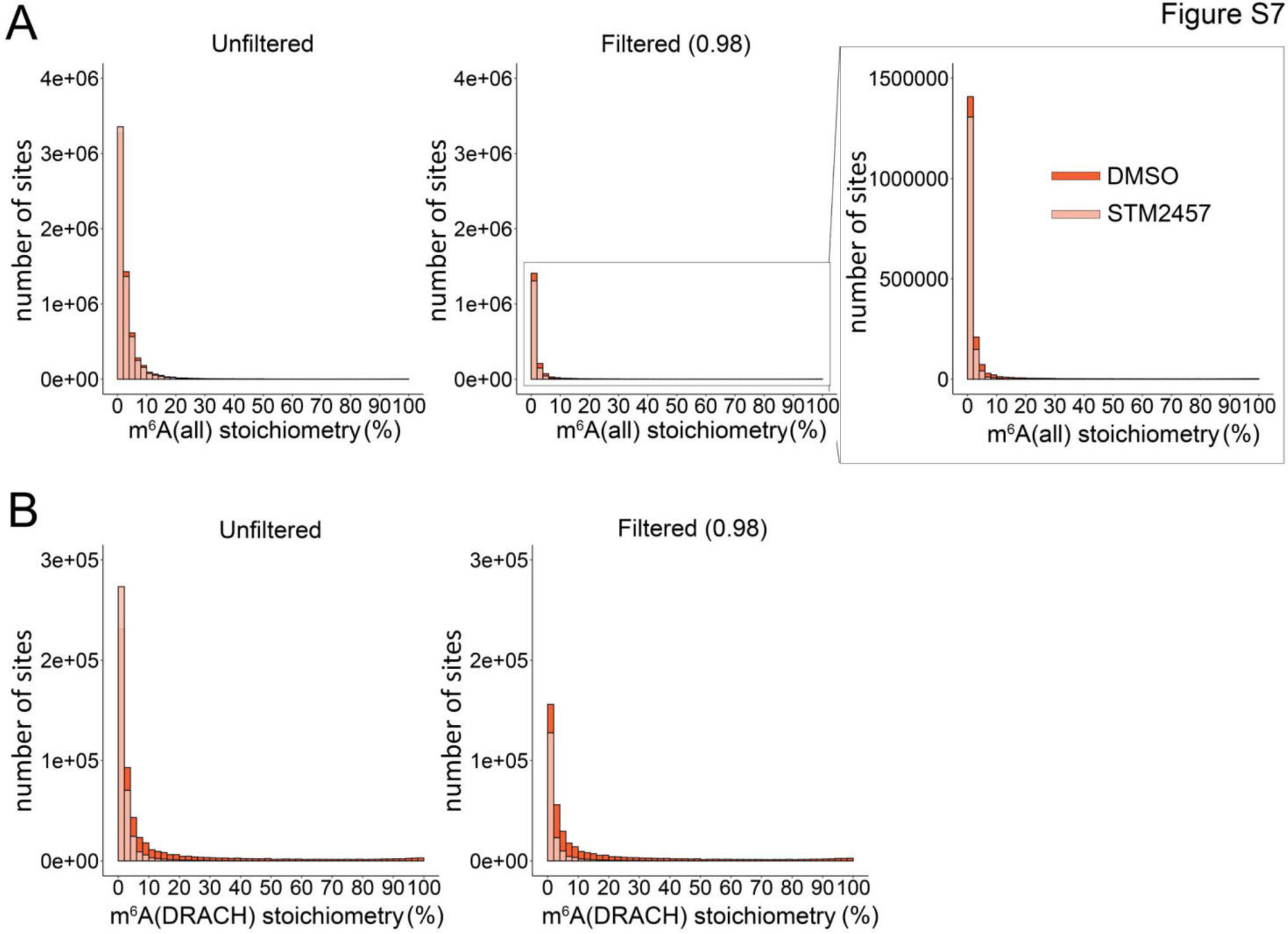

A
Unfiltered
Filtered (0.98)
number of sites
4e+06
3e+06
2e+06
1e+06
0e+00
1500000
1000000
500000
0
0 10 20 30 40 50 60 70 80 90 100
m6A(all) stoichiometry (%)
DMSO
STM2457
B
3e+05
2e+05
1e+05
m6A(DRACH) stoichiometry (%)

Figure S8

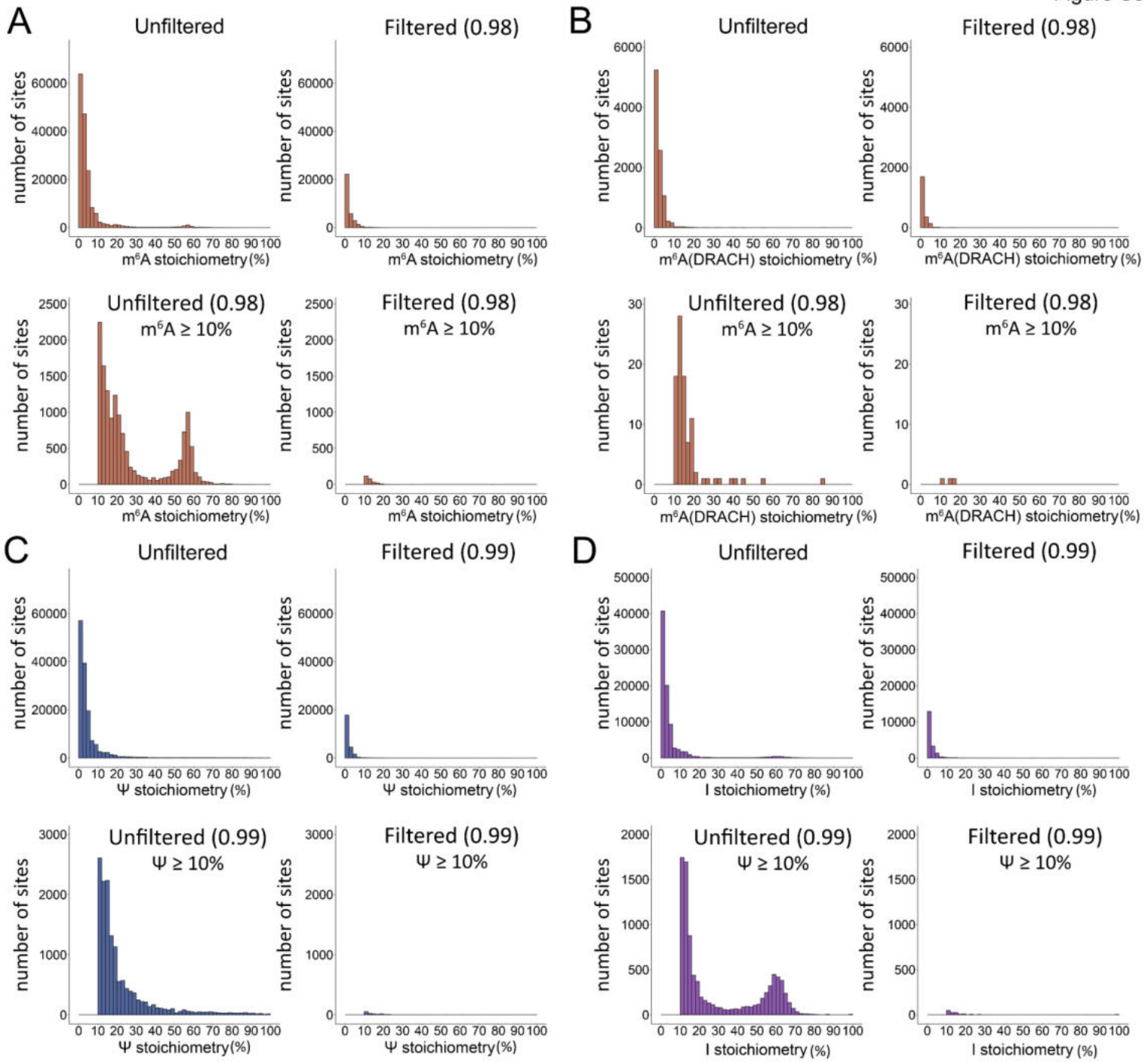

A
Unfiltered
Filtered (0.98)
Unfiltered (0.98)
m6A ≥ 10%
Filtered (0.98)
m6A ≥ 10%
number of sites
m6A stoichiometry (%)
B
Unfiltered
Filtered (0.98)
Unfiltered (0.98)
m6A ≥ 10%
Filtered (0.98)
m6A ≥ 10%
number of sites
m6A(DRACH) stoichiometry (%)
C
Unfiltered
Filtered (0.99)
Unfiltered (0.99)
Ψ ≥ 10%
Filtered (0.99)
Ψ ≥ 10%
number of sites
Ψ stoichiometry (%)
D
Unfiltered
Filtered (0.99)
Unfiltered (0.99)
Ψ ≥ 10%
Filtered (0.99)
Ψ ≥ 10%
number of sites
I stoichiometry (%)

Figure S9

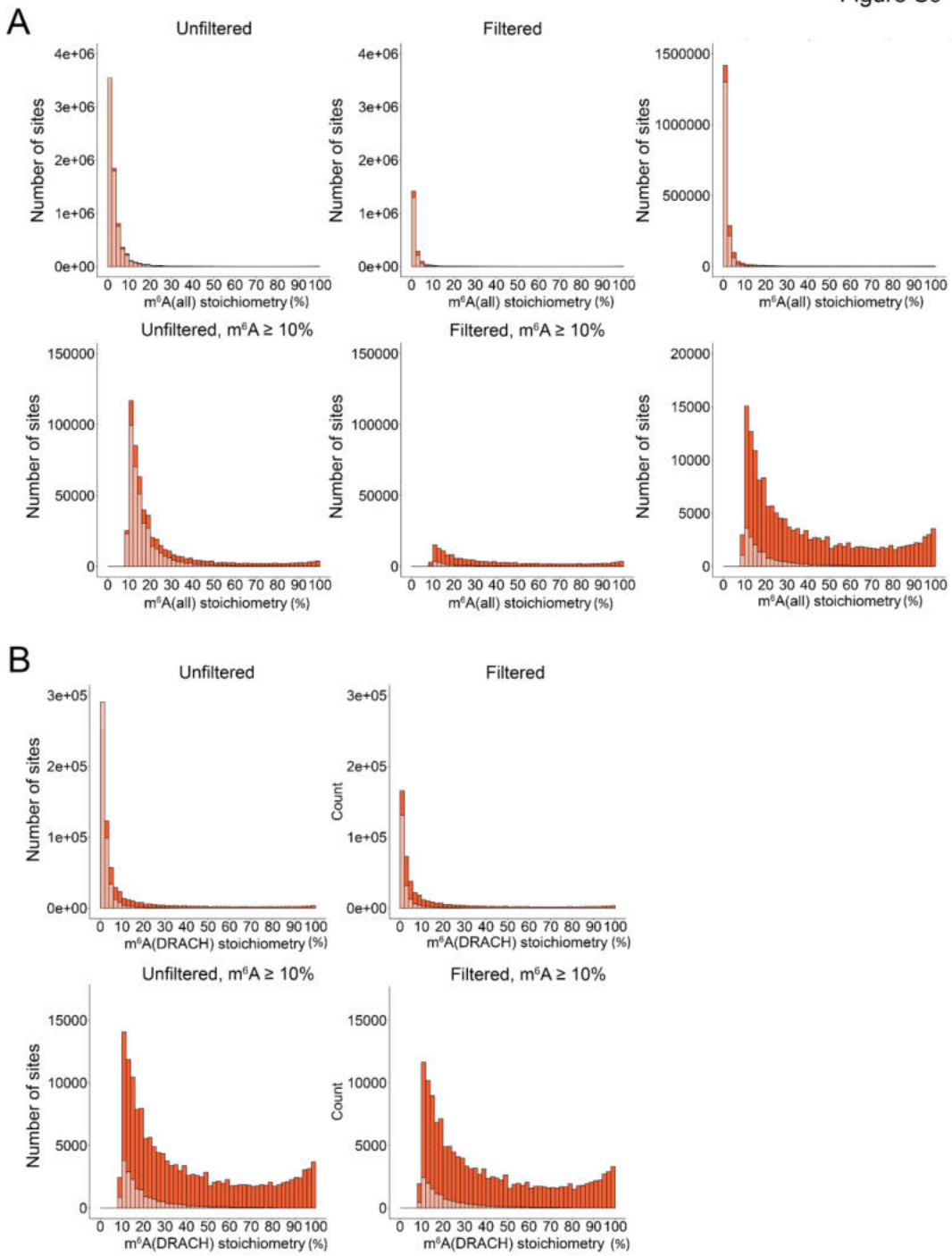

A
Unfiltered
Filtered
Unfiltered, m6A ≥ 10%
Filtered, m6A ≥ 10%
Number of sites
m6A(all) stoichiometry (%)
B
Unfiltered
Filtered
Unfiltered, m6A ≥ 10%
Filtered, m6A ≥ 10%
Number of sites
Count
m6A(DRACH) stoichiometry (%)

Figure S10

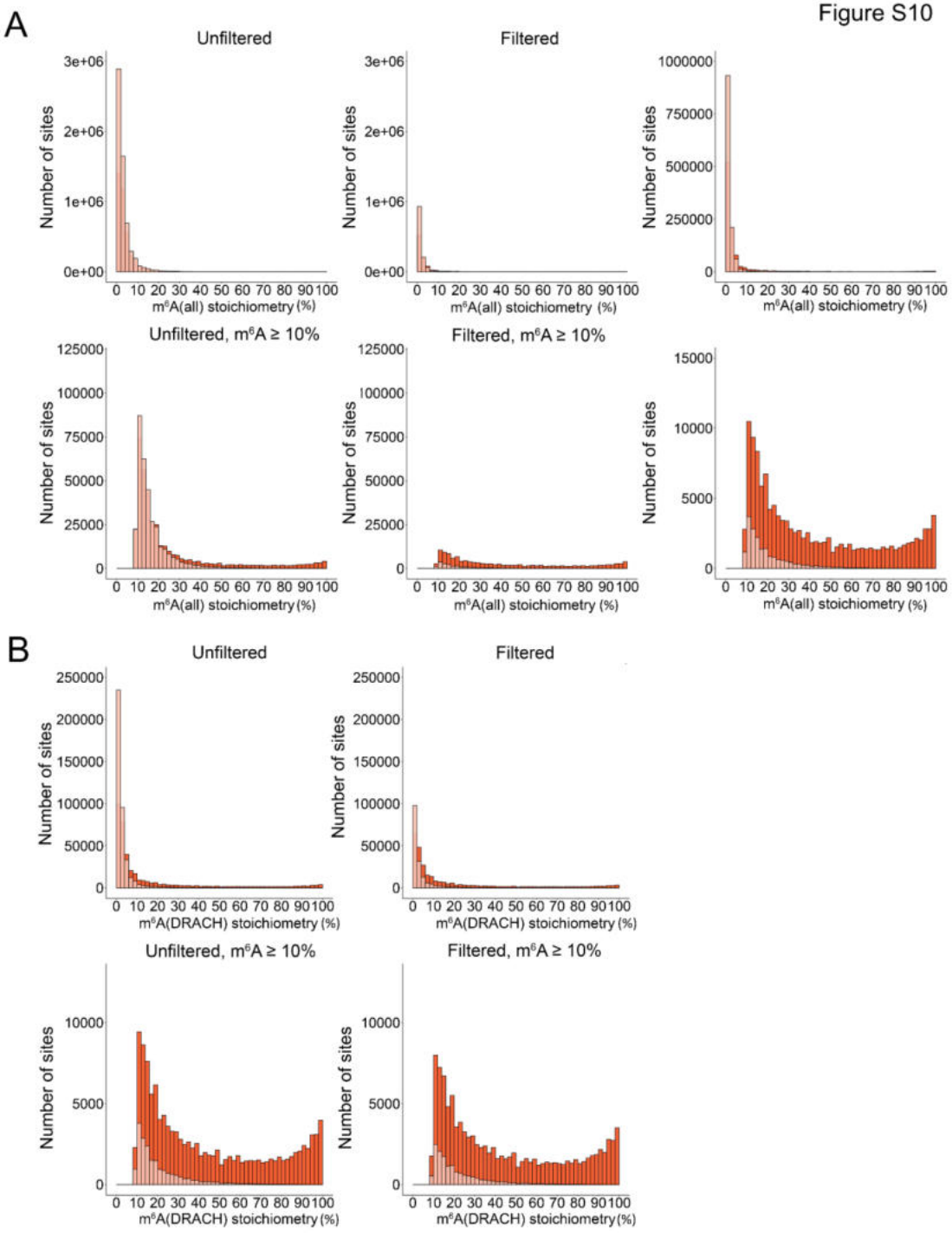

A
Unfiltered
Filtered
Unfiltered, m6A ≥ 10%
Filtered, m6A ≥ 10%
Number of sites
m6A(all) stoichiometry (%)
B
Unfiltered
Filtered
Unfiltered, m6A ≥ 10%
Filtered, m6A ≥ 10%
Number of sites
m6A(DRACH) stoichiometry (%)

Figure S11

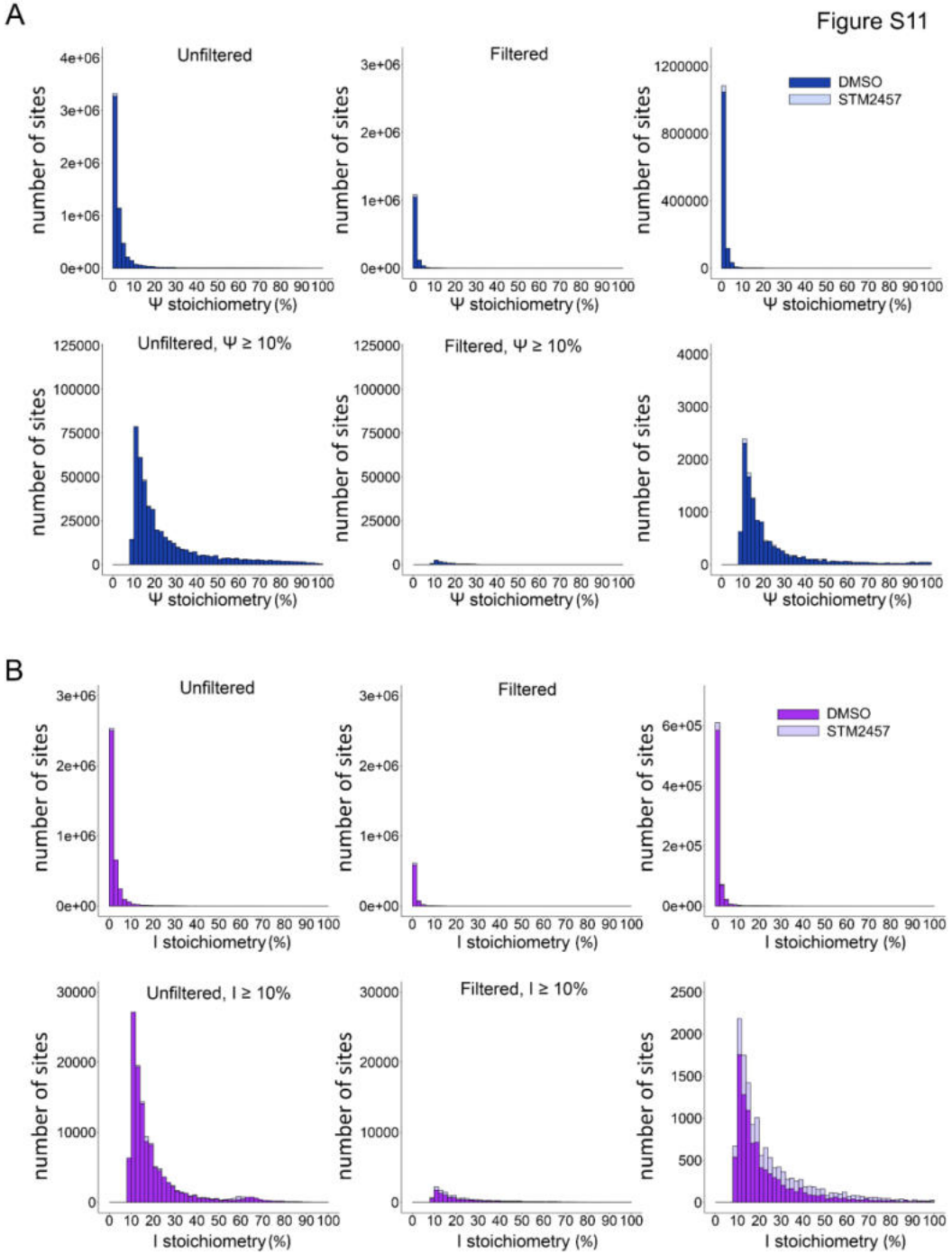

A
Unfiltered
Filtered
DMSO
STM2457
number of sites
Ψ stoichiometry (%)
Unfiltered, Ψ ≥ 10%
Filtered, Ψ ≥ 10%
B
I stoichiometry (%)
Unfiltered, I ≥ 10%
Filtered, I ≥ 10%

Figure S12

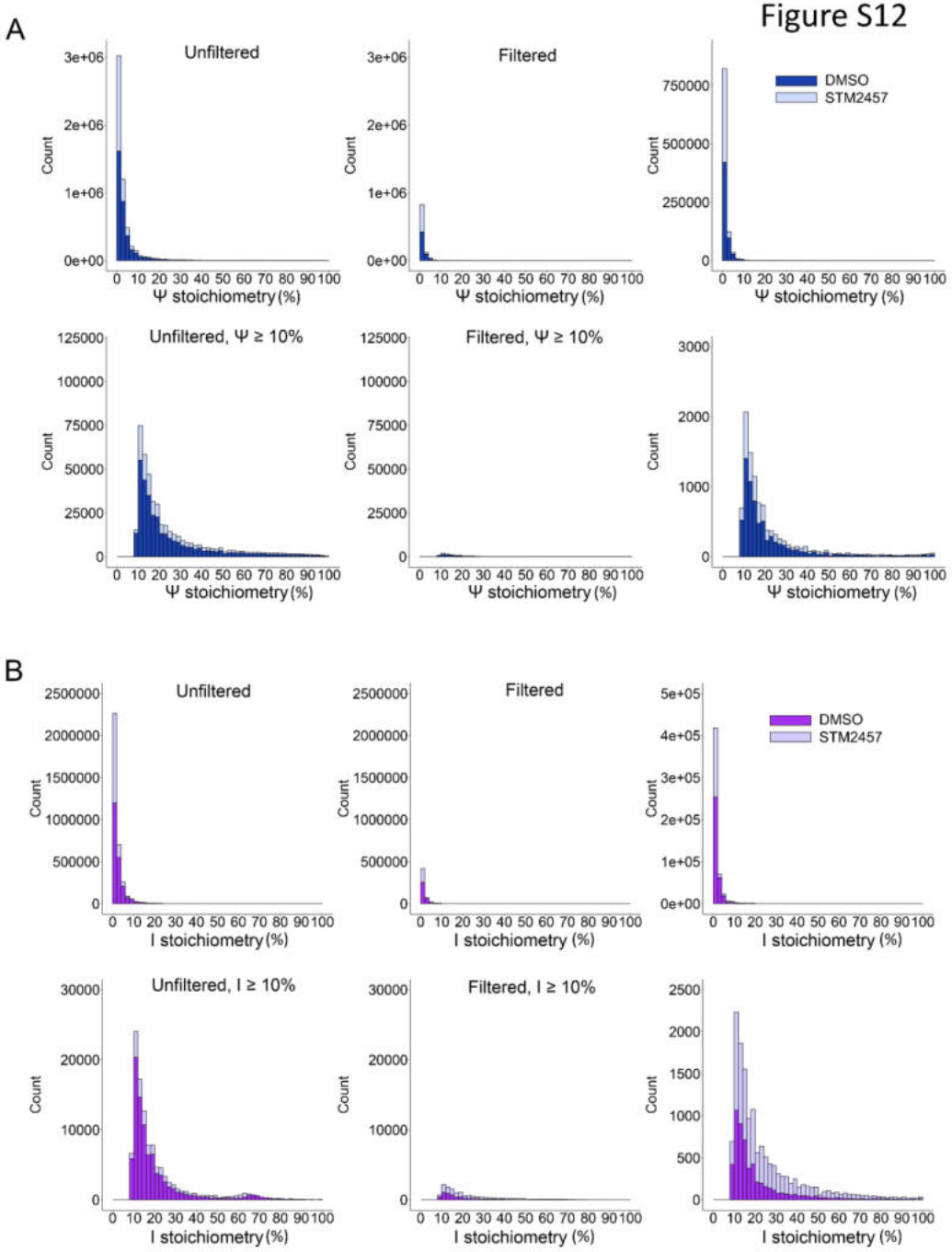

A
Unfiltered
Filtered
DMSO
STM2457
Unfiltered, Ψ ≥ 10%
Filtered, Ψ ≥ 10%
Count
Ψ stoichiometry (%)
B
Unfiltered
Filtered
DMSO
STM2457
Unfiltered, I ≥ 10%
Filtered, I ≥ 10%
Count
I stoichiometry (%)

Figure S13

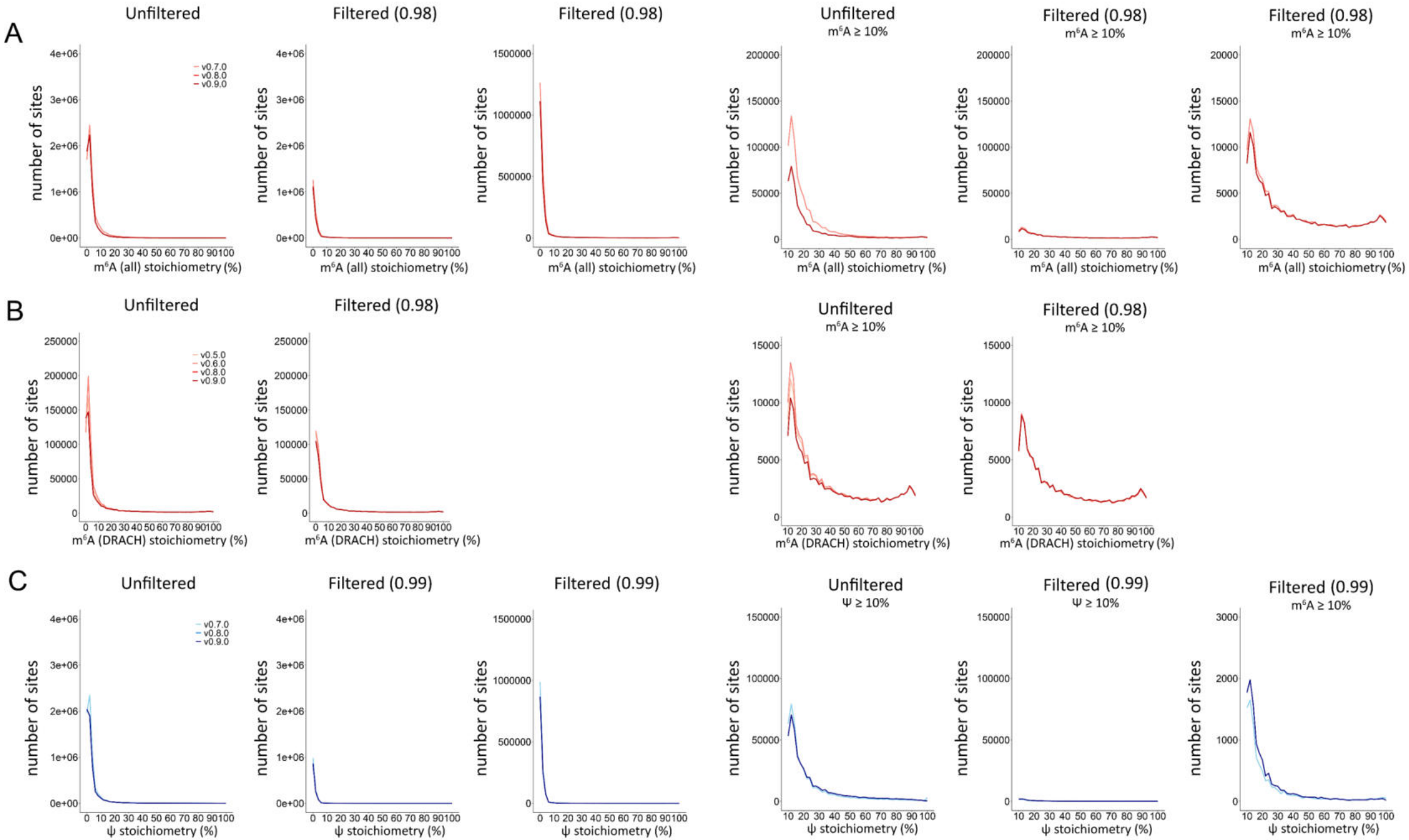

A
Unfiltered
Filtered (0.98)
Filtered (0.98)
Unfiltered
m6A ≥ 10%
Filtered (0.98)
m6A ≥ 10%
Filtered (0.98)
m6A ≥ 10%
v0.7.0
v0.8.0
v0.9.0
number of sites
m6A (all) stoichiometry (%)
B
Unfiltered
Filtered (0.98)
Unfiltered
m6A ≥ 10%
Filtered (0.98)
m6A ≥ 10%
v0.5.0
v0.6.0
v0.8.0
v0.9.0
number of sites
m6A (DRACH) stoichiometry (%)
C
Unfiltered
Filtered (0.99)
Filtered (0.99)
Unfiltered
Ψ ≥ 10%
Filtered (0.99)
Ψ ≥ 10%
Filtered (0.99)
m6A ≥ 10%
v0.7.0
v0.8.0
v0.9.0
number of sites
ψ stoichiometry (%)

Figure S14

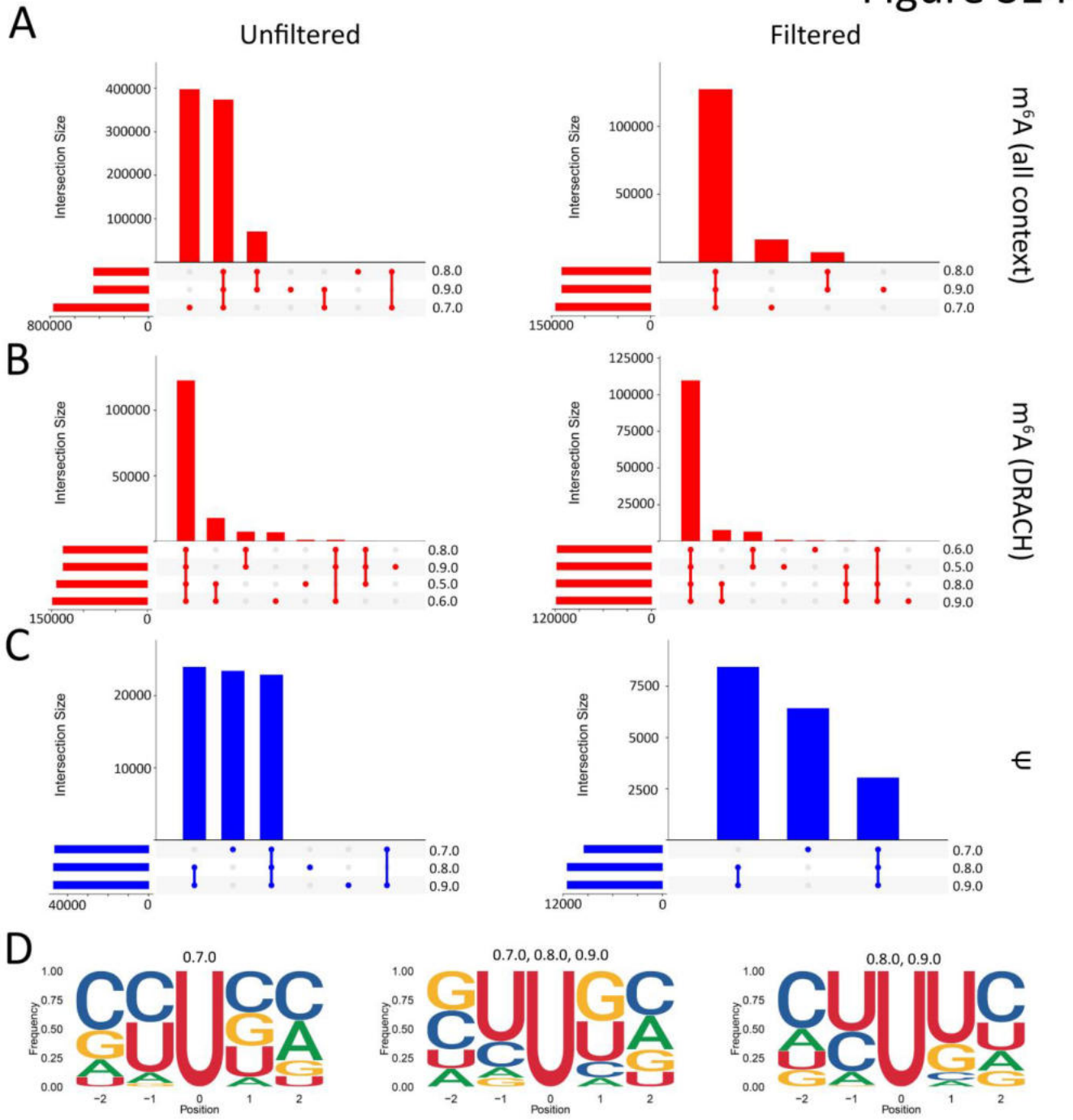

A
Unfiltered
Filtered
m6A (all context)
Intersection Size
400000
300000
200000
100000
800000
0
0.8.0
0.9.0
0.7.0
150000
B
m6A (DRACH)
125000
75000
50000
25000
120000
0.6.0
0.5.0
C
Ψ
20000
10000
7500
5000
2500
40000
12000
D
0.7.0
0.7.0, 0.8.0, 0.9.0
0.8.0, 0.9.0
Frequency
Position

Figure S15

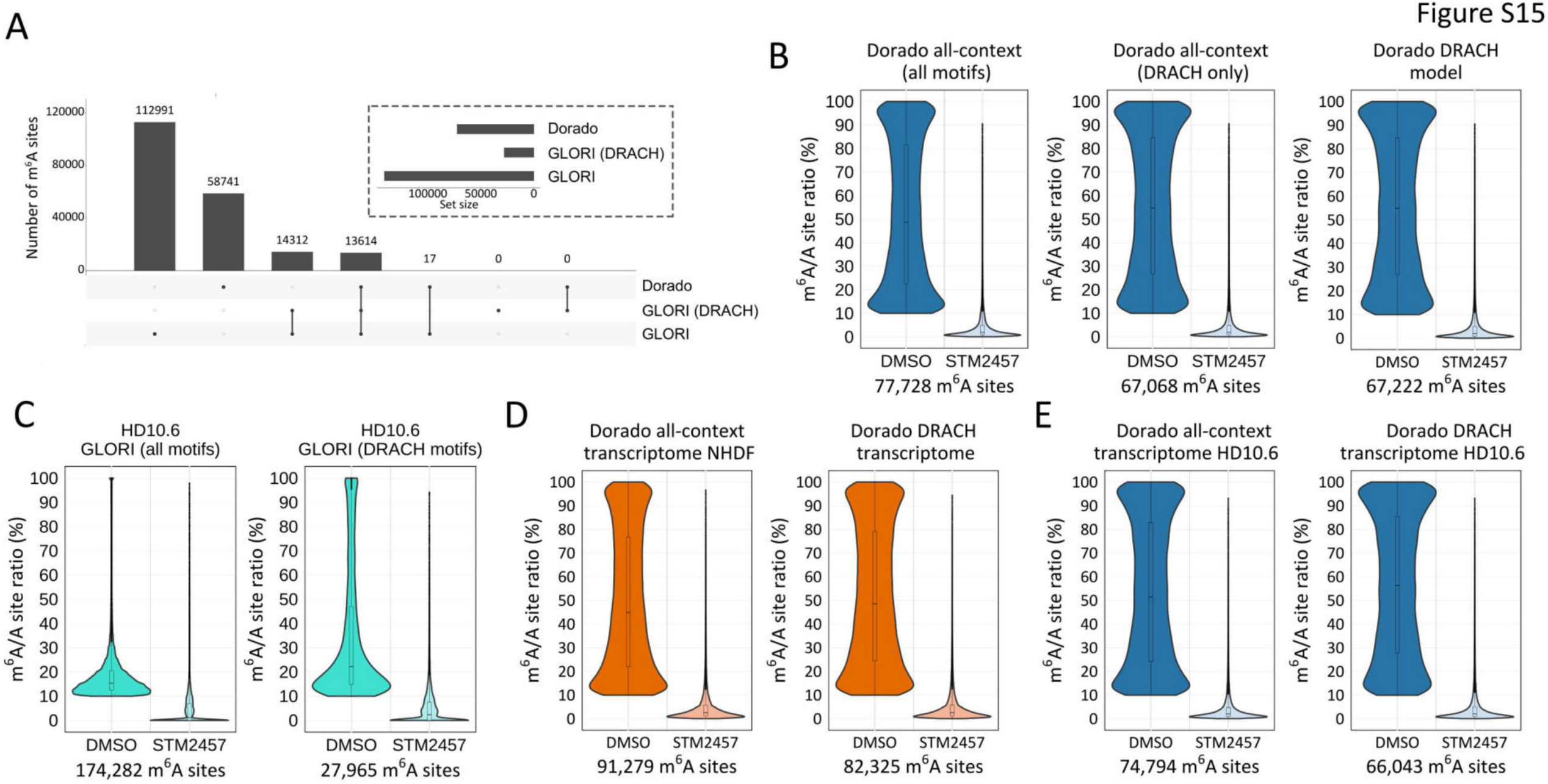
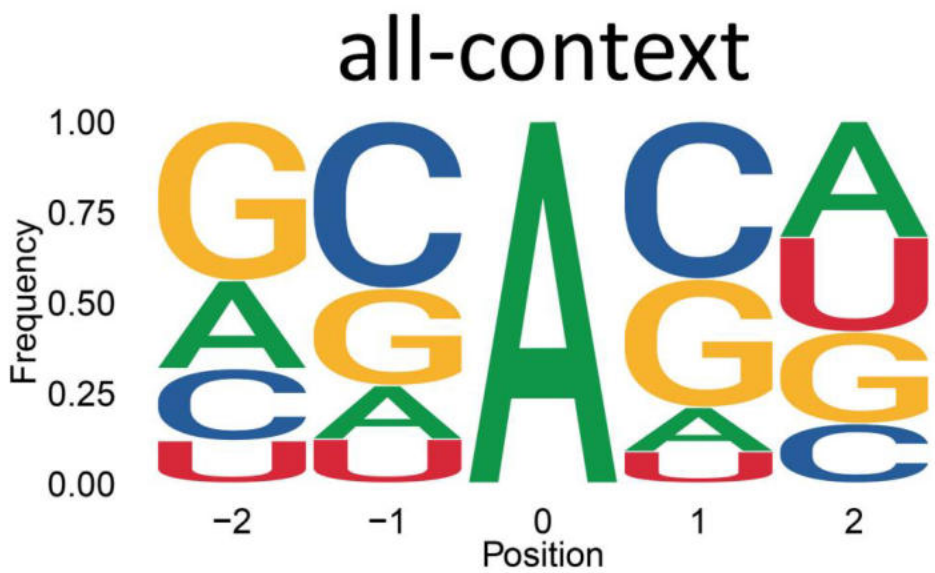
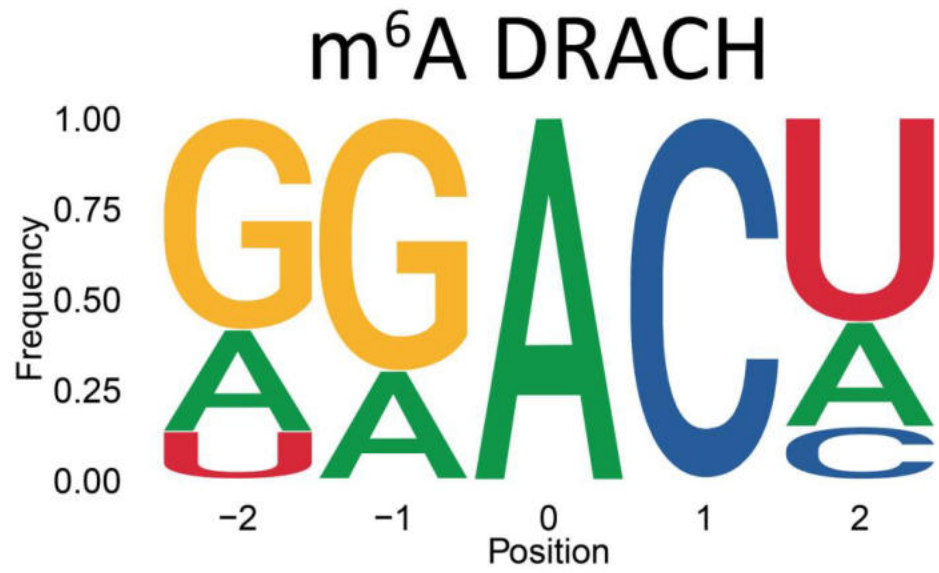

A
Number of m6A sites
120000
80000
40000
0
112991
58741
14312
13614
17
0
0
Dorado
GLORI (DRACH)
GLORI
100000
50000
0
Set size
B
Dorado all-context (all motifs)
Dorado all-context (DRACH only)
Dorado DRACH model
m6A/A site ratio (%)
DMSO
STM2457
77,728 m6A sites
67,068 m6A sites
67,222 m6A sites
C
HD10.6 GLORI (all motifs)
HD10.6 GLORI (DRACH motifs)
174,282 m6A sites
27,965 m6A sites
D
Dorado all-context transcriptome NHDF
Dorado DRACH transcriptome
91,279 m6A sites
82,325 m6A sites
E
Dorado all-context transcriptome HD10.6
Dorado DRACH transcriptome HD10.6
74,794 m6A sites
66,043 m6A sites

Figure S16

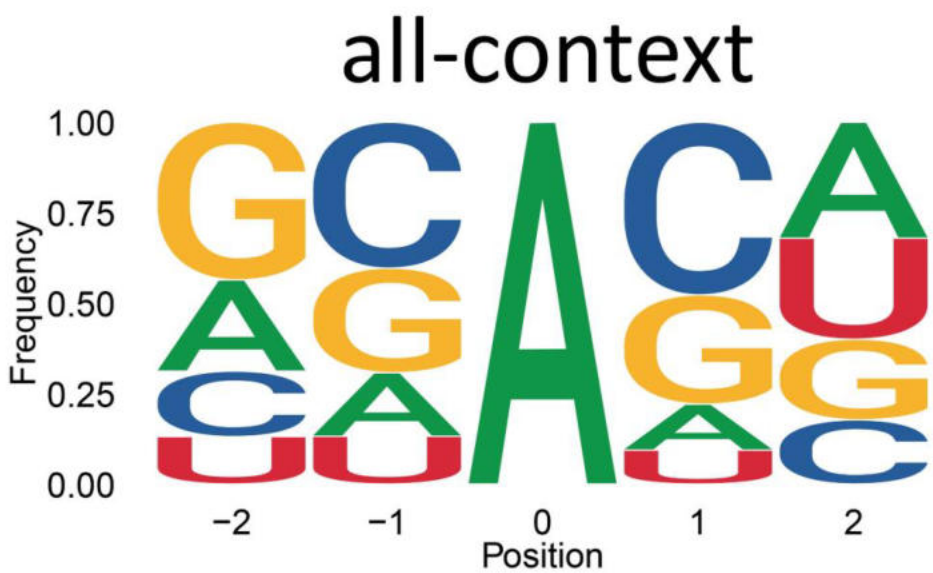
A

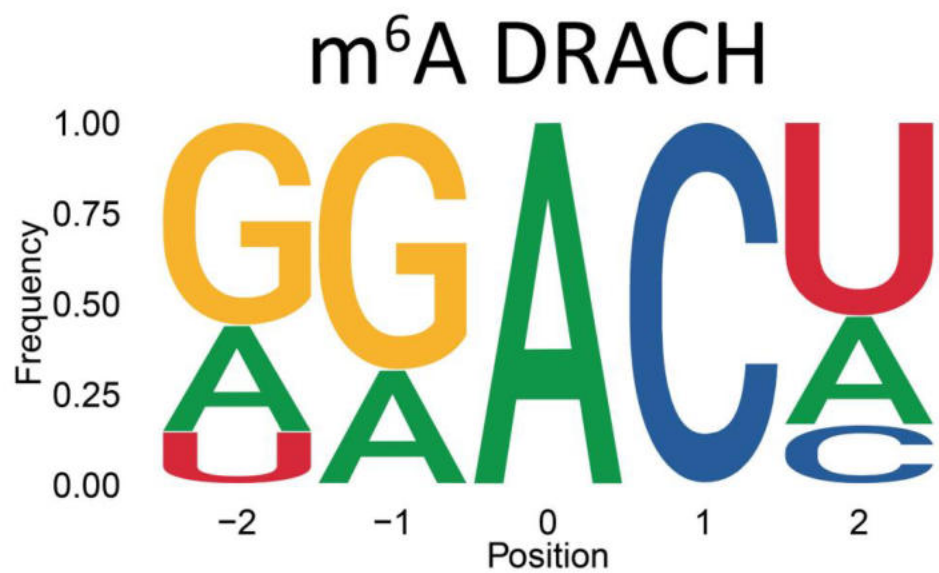




NHDF

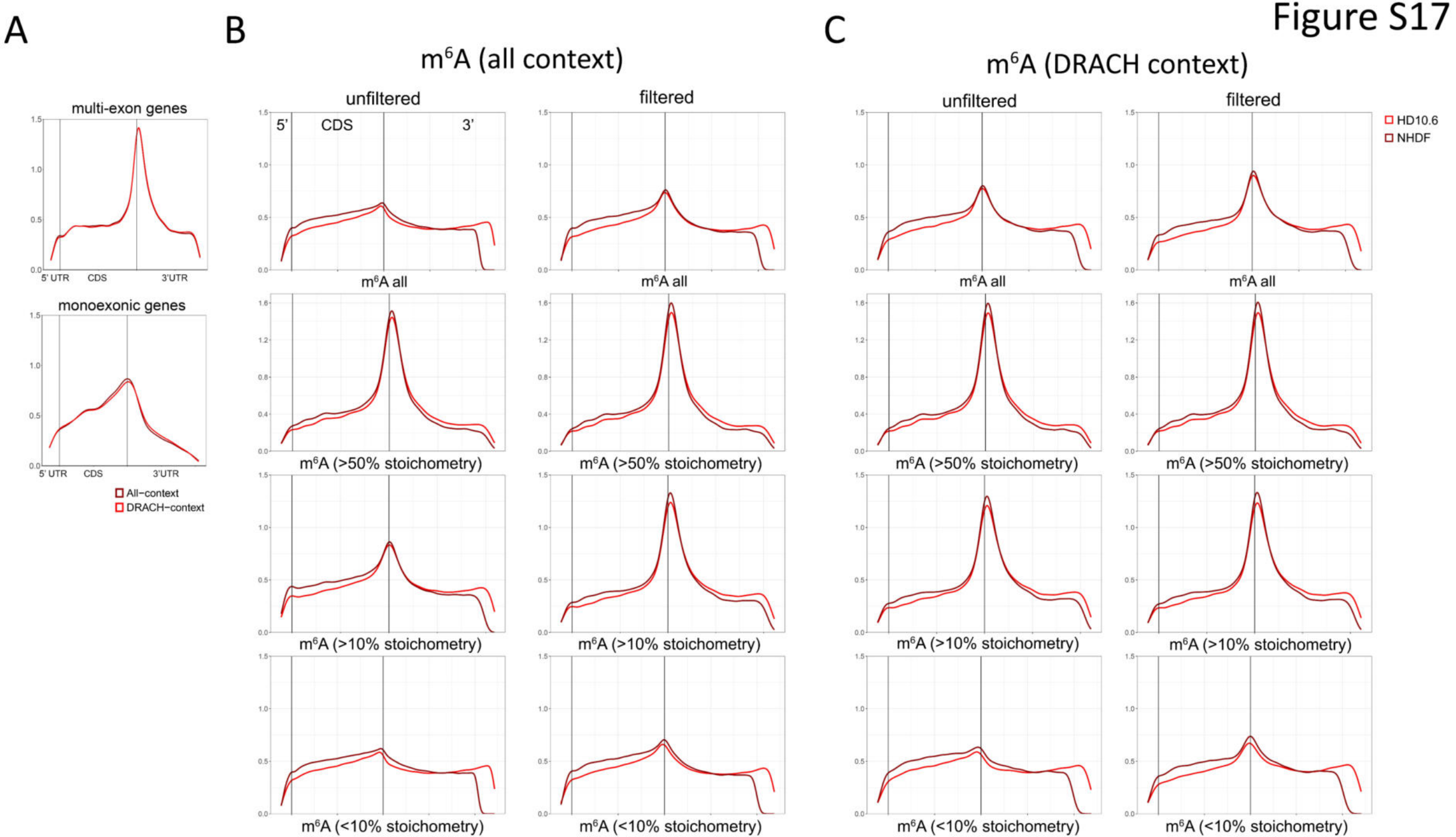
B





HD10.6

Figure S17


A
multi-exon genes
5' UTR
CDS
3'UTR
monoexonic genes
All−context
DRACH−context
B
m6A (all context)
unfiltered
filtered
5'
CDS
3'
m6A all
m6A (>50% stoichometry)
m6A (>10% stoichometry)
m6A (<10% stoichometry)
C
m6A (DRACH context)
HD10.6
NHDF