RESEARCH ARTICLE

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

Denise Ohnezeit[1†], Elene Loliashvili[2†], Gregory Putzel[3], Ruth Verstraten[2,4], Adam T Silhavy[1], Jianheng Fox Liu[5], Luke Nicholson[5], Alejandro Pironti[3], Samie R Jaffrey[5], Daniel Depledge[1,2,4,6*‡], Angus C Wilson[1*‡]

[1]Department of Microbiology, New York University School of Medicine, New York, United States; [2]Institute of Virology, Hannover Medical School, Hannover, Germany; [3]Anmicrobial-Resistant Pathogens Program, New York University Grossman School of Medicine, New York, United States; [4]German Center for Infection Research (DZIF), partner site Hannover-Braunschweig, Hannover, Germany; [5]Department of Pharmacology, Weill Medical College, Cornell University, New York, United States; [6]Cluster of Excellence RESIST (EXC 2155), Hannover Medical School, Carl-Neuberg-Straße 1, Hannover, Germany

***For correspondence:** depledge.daniel@mh-hannover.de (DD); angus.wilson@nyulangone.org (ACW)

[†]These authors contributed equally to this work
[‡]These authors also contributed equally to this work

Competing interest: See page 20

Funding: See page 20

Preprint posted
18 January 2026
Sent for Review
04 February 2026
Reviewed preprint v1
05 May 2026
Reviewed preprint v2
22 October 2026
Reviewed preprint v3
29 October 2026
Version of Record
20 November 2026

Reviewing Editor: Pedro Batista, National Institutes of Health, National Cancer Institute, United States

© Copyright Ohnezeit, Loliashvili *et al*. This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

eLife Assessment

This **valuable** study compares orthogonal approaches for detecting RNA chemical modifications and provides a helpful framework for improving the reliability of RNA modification identification by direct RNA sequencing. The evidence supporting the technical benchmarking claims is **solid**, with multiple complementary controls, including in vitro-transcribed RNA, chemical inhibition of METTL3, and orthogonal m6A mapping. However, claims regarding relative method accuracy and cell-type- and isoform-specific biological differences remain less well supported due to limited biological replication and validation.

Abstract Nanopore direct RNA sequencing (DRS) coupled with Dorado modification-aware base calling enables mapping of epitranscriptomic modifications including N^6-methyladenosine (m^{6}A) at the level of individual RNAs. However, the sensitivity, specificity, and reproducibility of this method remain unclear and have only recently begun to be addressed through systematic benchmarking studies. Here, we aimed to establish a best-practice workflow for DRS-based epitranscriptomic analyses. Specifically, we evaluated multiple Dorado versions and models using RNA isolated from primary cells and unmodified in vitro transcribed RNAs. We further utilized an m^{6}A methyltransferase inhibitor as a specificity control. We established that stringent filtering is necessary to reduce the contribution of background and found that Dorado predictions captured an increasing proportion of GLORI sites detected at high m^{6}A/A proportions. Further, by applying DRS to a human primary fibroblast and a HD10.6 neuron dataset, we detected differences in the predicted m^{6}A/A proportions at conserved sites between these datasets. Our study thus presents the first systematic comparison of Dorado and GLORI from the same input RNA and expands characterization of the m^{6}A 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., 2024***). Of these, at least 13 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 N6-methyladenosine (m^6A), 5-methylcytosine (m^5C), and N7-methylguanosine (m^7G; ***Yang et al., 2024***). 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***). The short-read length inherent to these approaches also limits the ability to determine whether multiple modifications occur on the same individual RNA molecule. 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***). This enables modification detection in an isoform-resolved, single-molecule context and, in principle, allows examination of the co-occurrence of multiple modifications along individual RNA molecules. 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, the introduction of a modification-aware basecaller (Dorado) provides an all-in-one toolkit for basecalling, RNA modification detection, and poly(A) tail length estimation. An early RNA modification-aware Dorado model (v3.0.1) supported m^6A detection exclusively within the DRACH context, while a later model (v5.0.0) introduced detection of Ψ and all-context m^6A. A more recent model (v5.1.0) supports both DRACH-context and all-context m^6A detection, together with Ψ, inosine, and m^5C (https://github.com/nanoporetech/Dorado, ***Bicknell et al., 2026***).

In a recent overview of this emerging technology, Cruciani and Novoa highlighted the importance of rigorous evaluation of Dorado's limitations and advocated for the establishment of best practices to ensure consistency and accuracy going forward (***Cruciani and Novoa, 2026***). Furthermore, Diensthuber and Novoa emphasized the need for direct comparisons between Dorado and orthogonal approaches such as GLORI using matched input material, as most benchmarking studies rely on previously published datasets rather than side-by-side experimental comparisons (***Diensthuber and Novoa, 2025***).

To date, several studies have evaluated Dorado in a benchmarking context. Two studies compared Dorado with m6Anet using SQK-RNA004 chemistry and reported improved performance of Dorado, while also incorporating in vitro transcribed (IVT) controls to assess false-positive predictions (***Zou et al., 2025***; ***Ghohabi Esfahani et al., 2026***). Incorporating selective METTL3 inhibition as a complementary, modification-specific negative control could further strengthen the assessment of m^{6}A detection specificity. Both studies also applied predicted m^{6}A/A proportion-based filtering to define high-confidence Dorado modification sites, using thresholds of 10% in Zou et al. and 20% in Ghobabi Esfahani et al.; however, the effect of the chosen predicted m^{6}A/A proportion threshold on site detection and false-positive reduction was not systematically evaluated. Consequently, a comprehensive set of best-practice parameters for Dorado-based modification detection has yet to be established. Notably, although Zou et al. compared Dorado-derived predictions with GLORI-identified sites, the GLORI data were obtained from previously published datasets rather than generated from matched input material, leaving direct comparisons between Dorado and GLORI on the same RNA samples unaddressed (***Cruciani et al., 2025***; ***Diensthuber and Novoa, 2025***; ***Zou et al., 2025***; ***Ghohabi Esfahani et al., 2026***). By directly comparing m^{6}A detection by Dorado and GLORI on the same RNA inputs, our study overcomes this limitation. Here, we systematically benchmark Dorado using poly(A) RNA from primary normal human dermal fibroblasts (NHDFs) and functional human neurons produced through terminal differentiation of HD10.6 cells (***Raymon et al., 1999***), along with IVT RNA. We included treatment with STM2457, an enzymatic inhibitor of the primary m^{6}A methyltransferase METTL3 (***Yankova et al., 2021***), and compared the modification-detection performance of multiple Dorado models released since the tool's introduction. While our analysis is centered on the sensitivity and specificity of m^{6}A detection in DRS datasets, we also benchmark the ability of Dorado to reliably detect m^{5}C, Ψ, and inosine. Together, our data provide a robust workflow for DRS-based modification detection that reliably predicts transcriptome-wide, isoform-level, site-specific proportions of RNA modifications. By expanding m^{6}A mapping on primary human fibroblasts and functional human neurons, we uncover mostly conserved m^{6}A profiles but reveal increased predicted m^{6}A/A proportions in neuronal cells.

Results

Systematic benchmarking of ONT's Dorado within human fibroblast and neuronal datasets

To evaluate the performance of ONT's basecalling tool Dorado across different human cell types, we selected primary NHDFs and differentiated HD10.6 cells as a neuron-like model. Prior to isolation of total RNA, cells were treated for 48 hr with a small-molecule inhibitor of METTL3 (STM2457) or with vehicle (DMSO). We ensured that this treatment had no effect on cell viability by performing a CellTiter Glo assay (***Figure 1—figure supplement 1***). We then isolated poly(A) RNA and subjected the same input material to both DRS and GLORI. The same material was also used to generate IVT RNA (***Figure 1A***). Dorado basecalling using software version v.0.9.0 and basecalling model version v.5.1.0 resulted in over 13 M reads using poly(A) RNA from NHDFs treated with DMSO or STM2457, while we obtained over 5 M and 10 M reads from DMSO- and STM2457-treated HD10.6 cells, respectively. The rate of primary alignment to the human genome (hg38) was over 98% for all the sequenced samples. A detailed overview of read statistics generated by DRS can be found in ***Supplementary file 1***.

To systematically benchmark Dorado, we performed basecalling using software versions v0.5.0 through v0.9.0. Reads were aligned to the human genome and transcriptome to obtain comprehensive information on isoform-level modification detection. Further, we explored different aspects of detected modifications using the outputs generated by ONT's Modkit tool (***Rand et al., 2026***). We assessed multiple filtering strategies using base and modification probabilities, as well as predicted modification proportion distributions, and evaluated them based on the IVT data and sensitivity to METTL3 inhibition. By filtering out background signal using multiple control layers, we determined high-confidence m^{6}A sites (***Figure 1B***). Finally, we compared these high-confidence predictions with sites identified by GLORI performed on the same input RNA. Using these results, we developed a benchmarked pipeline that filters m^{6}A sites predicted by Dorado to minimize background signal detection without impacting sensitivity (***Figure 1C***).

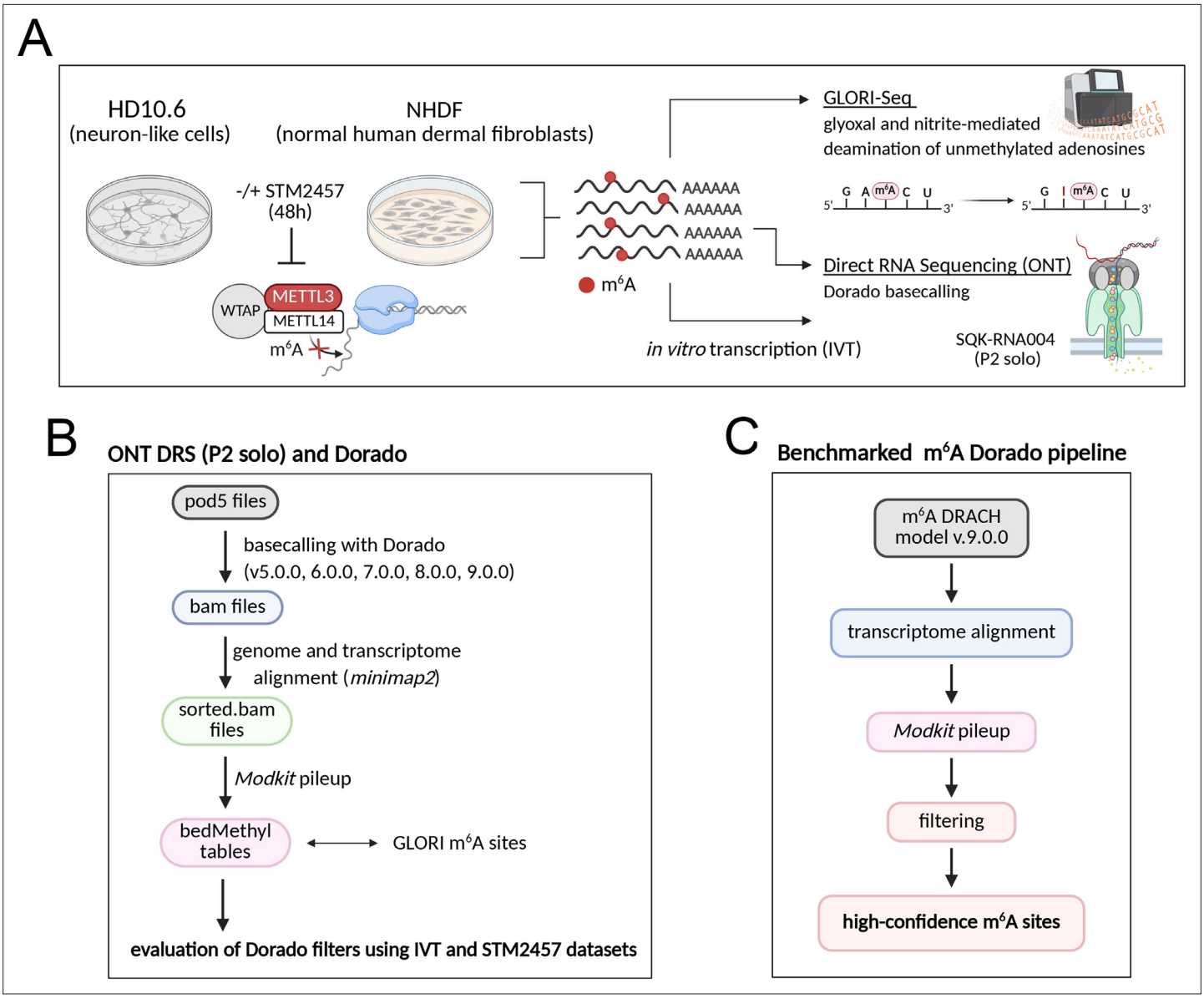

Figure 1. Workflow for m^{6}A mapping in different human cell types using orthogonal sequencing approaches. (**A**) Schematic of the experimental setup. RNA was isolated from normal human dermal fibroblasts (NHDFs) and differentiated HD10.6 cells treated with STM2457 or DMSO for 48 hr. Poly(A) RNA was subjected to direct RNA sequencing (DRS) and GLORI from the same input material. In vitro transcribed (IVT) RNA was generated in parallel to assess modification calls arising from background signal. (**B**) Workflow for benchmarking Dorado basecalling. Reads were processed with Dorado versions v.0.5.0 through v0.9.0, aligned to the human genome or transcriptome, and analyzed with ONT's (Oxford Nanopore Technologies) Modkit tool. Multiple filtering strategies were tested against IVT controls and METTL3 inhibition (STM2457). Resulting m^{6}A sites were compared with GLORI data from the same input RNA. (**C**) Final benchmarked Dorado pipeline to detect high confidence m^{6}A sites.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Cell viability assays in STM2457-treated normal human dermal fibroblasts (NHDFs) and HD10.6 cells.

Stringent filtering of modification probabilities removes false positives

During modification-aware basecalling, each nucleotide is assigned a basecall probability score ranging from 0 to 1. Similarly, individual modified nucleotides are assigned a modification probability score in the same range. To establish a baseline for background signal, we compared basecall and m^{6}A modification probability distributions between a human genome-aligned IVT poly(A) RNA dataset and DMSO- and STM2457-treated NHDF datasets, all processed using Dorado v0.9.0 (***Figure 2A and B***, ***Figure 2—figure supplement 1***). Basecall probabilities for adenosine and m^{6}A

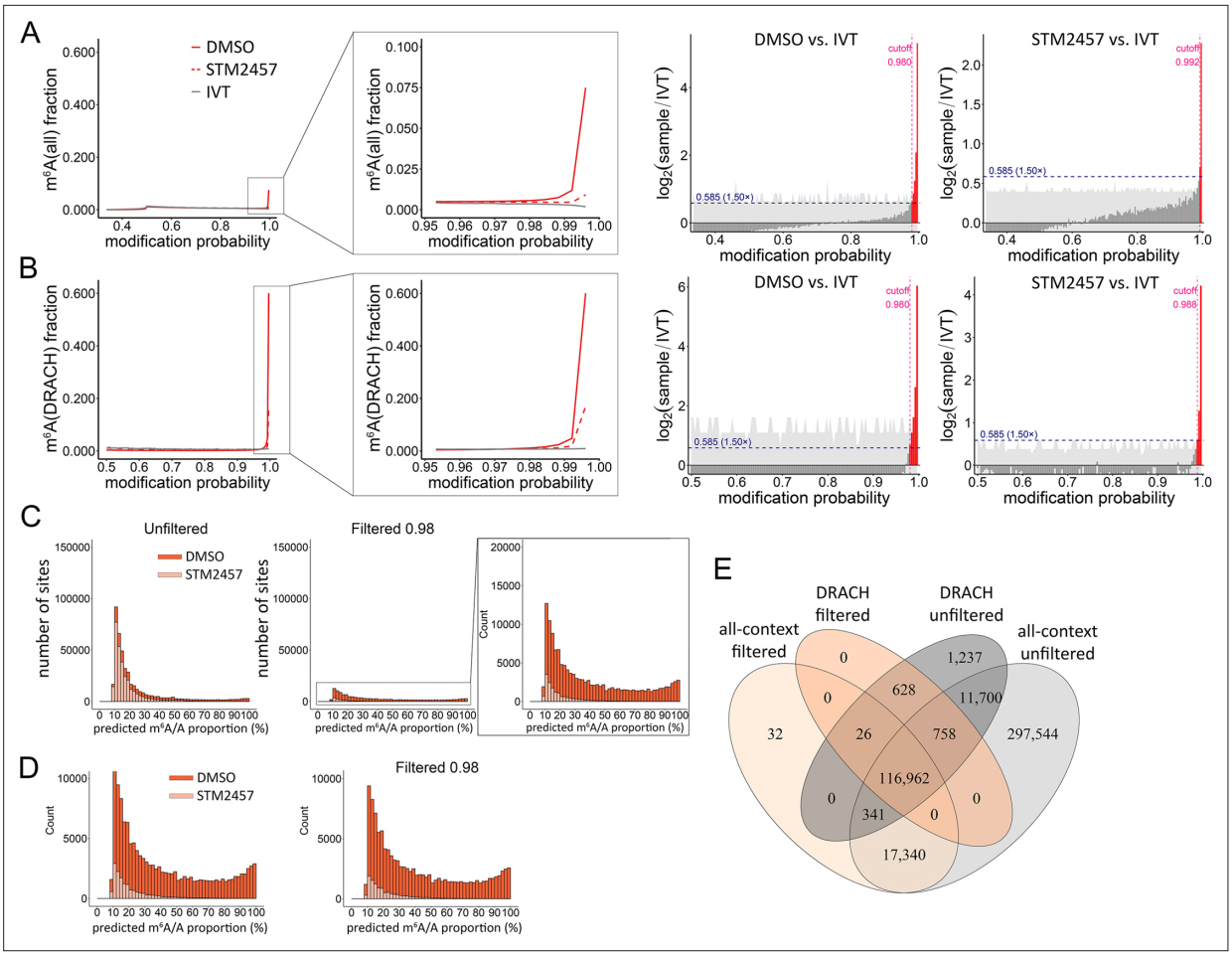

Figure 2. Modification probability of 0.98 and predicted m^{6}A/A proportion score of ≥10% are necessary cutoffs to identify m^{6}A sites accurately. (**A**) (left) Dorado v0.9.0-generated modification probability distributions for all-context m^{6}A from genome-aligned DMSO- and STM2457-treated normal human dermal fibroblasts (NHDFs) plotted against the distributions obtained from the in vitro transcribed (IVT) NHDF poly(A) RNA. DMSO probability score distribution is shown as a red solid line, STM2457 – red dashed line, and IVT – grey solid line. (right) Bar plots showing the per-bin log2(sample/IVT) ratio of modification probability fractions for the DMSO- and STM2457-treated NHDF conditions relative to the unmodified IVT reference. The horizontal navy dashed line marks the fold-enrichment threshold used to call enrichment. Bars are colored by significance relative to the fold-enrichment threshold (red, log2≥1.5 fold; dark grey, not significant). The light grey ribbon indicates the 95% bootstrap confidence interval around the log2 ratio estimate at each bin. The vertical pink dashed line indicates the derived modification probability cutoff above which sites are considered genuinely modified. (**B**) Same as in (**A**), but for DRACH-context m^{6}A sites. (**C**) Predicted m^{6}A/A proportion (≥10%) distributions for all-context m^{6}A from DMSO- (in red) or STM2457-treated (in pink) NHDF datasets (genome-aligned) are plotted using a 0.98 modification probability cutoff and a filter for ≥20 reads. (**D**) Same as in (**C**), but for DRACH-context m^{6}A sites. (**E**) Overlap of m^{6}A sites between all-context and DRACH models (A basecall accuracy ≥80%, coverage ≥20 reads, predicted m^{6}A/A proportion ≥10%), comparing non-filtered to filtered (modification probability >0.98) outputs.

The online version of this article includes the following figure supplement(s) for figure 2:

Figure supplement 1. Evaluation of Dorado by base and modification probability distributions.

Figure supplement 2. In vitro transcribed (IVT)-based estimation of background signal-derived m^{6}A detection across modification probability thresholds and ranges in normal human dermal fibroblasts (NHDFs).

Figure supplement 3. In vitro transcribed (IVT)-based estimation of m^{6}A detection derived from background signal across modification probability thresholds and ranges in HD10.6 cells.

Figure supplement 4. Modification probability cutoffs are necessary to identify Ψ and inosine sites more accurately.

Figure supplement 5. In vitro transcribed (IVT)-based estimation of Ψ, inosine, and m^{5}C detection derived from background signal across modification probability thresholds and ranges in normal human dermal fibroblasts (NHDFs).

Figure supplement 6. In vitro transcribed (IVT)-based estimation of Ψ, inosine, and m^{5}C derived from background signal detection across modification probability thresholds and ranges in HD10.6 cells.

Figure supplement 7. Modification probability cutoffs are necessary to identify modification sites accurately, shown for transcriptome-aligned data.

Figure 2 continued on next page

Figure 2 continued

Figure supplement 8. Modification probability distribution profiles in genome-aligned DMSO- and STM2457-treated HD10.6 cells.

Figure supplement 9. Modification probability distribution profiles in transcriptome-aligned DMSO- and STM2457-treated HD10.6 cells.

Figure supplement 10. Modification site probabilities detected with different versions and models of Dorado and compared to RNA calibration strand (RCS).

Figure supplement 11. Evaluation of Dorado by the distributions of predicted m^{6}A/A proportions.

Figure supplement 12. Distributions of predicted m^{6}A/A, Ψ/U, and inosine/A proportions in in vitro transcribed (IVT) datasets.

Figure supplement 13. Evaluation of Dorado by predicted modification proportion distributions in transcriptome-aligned normal human dermal fibroblasts (NHDFs).

Figure supplement 14. Evaluation of Dorado by distributions of predicted m^{6}A/A proportions in transcriptome-aligned HD10.6 cells.

Figure supplement 15. Distributions of predicted Ψ/U and inosine/A proportions in normal human dermal fibroblasts (NHDFs).

Figure supplement 16. Distributions of predicted Ψ/U and inosine/A proportions in HD10.6 cells.

Figure supplement 17. Relationship between predicted RNA modification proportions and nucleotide basecalling errors in normal human dermal fibroblast (NHDF) and HD10.6 datasets.

Figure supplement 18. Predicted m^{6}A/A and Ψ/U proportion distributions in different versions of Dorado.

Figure supplement 19. Comparison of modification sites detected with different versions of Dorado.

modification probabilities from both the all-context and DRACH-context models were grouped into probability bins for comparison.

We observed that 75–90% of A basecalls (in all datasets) have probability scores above 0.8 (***Figure 2—figure supplement 1A***). Since all A bases are considered during all-context m^{6}A calling, the fraction of m^{6}A sites is substantially smaller when compared to DRACH-context m^{6}A sites (***Figure 2A and B***, ***Figure 2—figure supplement 1B and C***). The modification probability distributions are effectively identical for DMSO-treated, STM2457-treated, and IVT datasets below 0.98, while the profiles differ substantially at probabilities ≥0.98. This suggests that sites detected below 0.98 are likely background signal, indicating that only high-probability m^{6}A calls can be considered reliable. Consistent with this interpretation, METTL3 inhibition with STM2457 markedly reduced the number of detected m^{6}A sites with modification probabilities ≥0.98, compared with DMSO-treated samples, in both the all-context and DRACH-only datasets (***Figure 2A and B***, ***Figure 2—figure supplement 1B and C***).

To identify specific modification probability cutoffs for each comparison between biological samples and the IVT dataset, we established a statistical framework in which modification probabilities were divided into bins, and enrichment of modification calls in the biological sample relative to IVT was expressed as the log2-transformed sample/IVT ratio. The Pruned Exact Linear Time (PELT) algorithm was then used to identify the point at which this ratio shifted from the baseline distribution, indicating divergence of the biological signal from the IVT background. The modification probability cutoff was subsequently defined as the first bin following the detected change point that exceeded the specified fold-enrichment threshold. This analysis independently supported the use of stringent modification probability cutoffs to distinguish modification-associated signal from the IVT background (***Figure 2A and B***).

To assess the contribution of background signal to modification detection across different modification probability thresholds (0.5, 0.7, 0.9, 0.98, and 0.99), 200,000 reads were randomly subsampled from each dataset in 10 iterations, and modification probability distributions were determined for each subsample. Background signal was estimated by comparing modification counts in biological samples with those detected in the corresponding IVT controls at each threshold. In NHDF samples, the m^{6}A DRACH sites showed a relatively low contribution of background signal (~10%) even at a probability threshold of 0.5, which decreased further at more stringent thresholds (***Figure 2—figure supplement 2A***). By contrast, the m^{6}A all-context model showed substantially higher background signal (~35% at ≥0.5), decreasing to <5% at the highest thresholds. Analysis of discrete modification probability ranges showed a similar trend (***Figure 2—figure supplement 2B***), with background signal decreasing significantly at higher modification probabilities. Statistical significance was assessed using pairwise comparisons derived from linear mixed-effects models, with sample included as a random effect and Benjamini–Hochberg correction for multiple comparisons. The reduction in background signal in all-context m^{6}A detection was less pronounced following STM2457 treatment, indicating that

background signal becomes proportionally more prominent as the number of genuine modifications decreases. Similar patterns were observed in HD10.6 cells (***Figure 2—figure supplement 3A and B***).

Furthermore, we examined modification probability distributions for Ψ, inosine, and m^5C, each of which can be basecalled with Dorado v0.9.0 (***Figure 2—figure supplement 4***). For Ψ and I, the modification probability distributions for the STM2457 and DMSO datasets only diverge from the IVT dataset at ≥0.99 modification probability, supporting the use of this stringent threshold to reduce the impact of background signal (***Figure 2—figure supplement 4A and B***). Consistent with this, the estimated background signal decreased with increasing modification probability, although they remained comparatively high (~20–30%) even at stringent thresholds (***Figure 2—figure supplement 4A and B***, ***Figure 2—figure supplement 5A***, ***Figure 2—figure supplement 6A***). Analysis across discrete modification probability intervals showed the same trend, with progressively lower error rates at higher probabilities (***Figure 2—figure supplement 4A and B***, ***Figure 2—figure supplement 5B***, ***Figure 2—figure supplement 6B***). By contrast, m^5C modification probability distributions were similar between all three datasets, and no filter could be adequately established (***Figure 2—figure supplement 4C***). While this likely reflects a low rate of m^5C installation, it also indicates that the m^5C model utilized by Dorado requires further refinement.

Modification probability distributions have similar profiles for transcriptome-aligned NHDF data (***Figure 2—figure supplement 7***) as well as genome- and transcriptome-aligned data from HD10.6 cells (***Figure 2—figure supplement 8***, ***Figure 2—figure supplement 9***). Of note, the fraction of high-probability inosine sites is increased in STM2457-treated cells, compared to DMSO-treated ones in genome-aligned data (***Figure 2—figure supplement 4B***, ***Figure 2—figure supplement 8D***), which points to increased numbers of retained introns in the METTL3-inhibited sample. Accordingly, this effect is absent in transcriptome-aligned data (***Figure 2—figure supplement 7D***, ***Figure 2—figure supplement 9D***).

While IVT transcriptomes serve as excellent negative controls, we also considered whether the RNA calibration strand (RCS) that is added during DRS library preparation could serve a similar function. The RCS is a synthetic Enolase II RNA derived by IVT that, similarly to the IVT transcriptomes, is devoid of RNA modifications. Despite the reduced diversity of sequence in the RCS data, the overall results were comparable to those from the IVT transcriptome data (***Figure 2A and B***, ***Figure 2—figure supplement 10A and B***), demonstrating that the RCS can be used as a 'built-in' negative control for modification detection with DRS. Finally, we determined that basecall and modification probability distributions remain generally consistent between different Dorado models, although more m^6A and Ψ sites are detected with high modification probability in Dorado model v5.1.0 compared to v5.0.0 (***Figure 2—figure supplement 10C–E***).

Based on our observations, we propose a base probability cutoff of 0.8 and a modification probability cutoff of 0.98 for m^6A (and 0.99 for Ψ and inosine) to reduce background signal.

Filtering by predicted m^6A/A proportion further reduces the impact of background signal

We next examined the predicted m^6A/A proportion distributions for all m^6A sites in both modification probability-unfiltered and filtered NHDF datasets (***Figure 2—figure supplement 11***). In addition to applying base and modification probability cutoffs, we restricted analyses to sites corresponding to adenosine residues in the reference sequence used for alignment, an approach also reported by Esfahani and colleagues (***Ghohabi Esfahani et al., 2026***) thereby further reducing the likelihood of background signal. We additionally required a minimum valid coverage of 20 reads per site, which is also consistent with the criteria applied in the same study. Valid coverage was defined as the sum of reads supporting the canonical adenosine base and reads supporting an m^6A modification call (with ≥0.98 modification probability score). Unfiltered datasets showed that Dorado identifies 11.5x more m^6A sites with the all-context model than with the DRACH model (***Supplementary file 2***), consistent with the prior study by Cruciani and colleagues (***Cruciani and Novoa, 2026***), reporting 20x more all-context m^6A sites than DRACH-context. Furthermore, we showed that the vast majority of m^6A sites, 92% for all-context and 75% for DRACH-context, have predicted m^6A/A proportion values <10%, with less than 1% of sites showing high (>50%) predicted proportions.

A similar analysis of the IVT transcriptome revealed large numbers of m^6A sites derived from background signal with predicted m^6A/A proportion values below 10%, most of which were eliminated by

application of our filters (***Figure 2—figure supplement 12A and B***). However, the remaining presence of low-proportion sites derived from background signal indicated that an additional layer of filtering to exclude sites with <10% proportion is beneficial. Notably, this filter was also adopted in the benchmarking study by Zou and colleagues (***Zou et al., 2025***). An additional argument supporting our filtering strategy is the observation that the effect of treatment with the METTL3-specific inhibitor STM2457 is much more pronounced at m^{6}A sites with ≥10% predicted m^{6}A/A proportion (***Figure 2C and D***; ***Supplementary file 2***), regardless of the motif context. Applying modification probability and per-site proportion cutoffs substantially reduced the difference between the numbers of m^{6}A sites detected with all-context and DRACH-context modes of Dorado (***Supplementary file 2***). Application of our modification probability filtering strategy removed over 70% of the sites (***Figure 2E***), regardless of the context, almost all of which showed predicted m^{6}A/A proportions of <10%, without significantly changing the numbers of high-proportion sites detected. Similar results were also observed for the transcriptome-aligned NHDF and HD10.6 distributions (***Figure 2—figure supplement 13***, ***Figure 2—figure supplement 14***; ***Supplementary file 2***).

To determine whether our filtering strategy could be applied to other RNA modifications detected by Dorado, we examined the distributions of predicted Ψ/U and inosine/A proportions in human genome-aligned DMSO- and STM2457-treated NHDF and HD10.6 datasets, before and after applying the ≥0.99 modification probability threshold. Similar to m^{6}A (***Figure 2—figure supplement 11***), the majority of detected Ψ and inosine sites exhibited low predicted modification proportions, with 92% of Ψ sites and 96% of inosine sites showing proportions below 10% (***Figure 2—figure supplement 15*** for NHDF cells, ***Figure 2—figure supplement 16*** for HD10.6 cells). This predominance of low-proportion Ψ sites is consistent with a recent study using 2-bromoacrylamide-assisted cyclization sequencing, which reported that 82% of Ψ modifications in human mRNAs occurred at proportions below 20% (***Xu et al., 2024***).

Importantly, analysis of the corresponding IVT datasets further supported the use of the ≥0.99 modification probability threshold for Ψ and inosine (***Figure 2—figure supplement 12C and D***). Applying this threshold substantially reduced the number of sites detected in the IVT controls compared with the unfiltered data. For inosine, predictions derived from background signal were present in the unfiltered IVT data even at high (>50%) predicted modification proportions and were markedly reduced following modification probability filtering (***Figure 2—figure supplement 12D***). These findings indicate that stringent modification probability filtering is necessary to reduce background predictions for Ψ and inosine before interpreting their estimated per-site proportions.

We have thus demonstrated that setting thresholds for basecall modification probabilities provides a robust method for identifying and filtering against the influence of background signal.

Ψ and inosine sites exhibit the highest degree of modification-specific basecalling mismatches

An important consideration when interpreting the predicted modification proportions is that these estimates do not account for all reads overlapping a given position. Modification proportions are calculated from reads in which the corresponding reference base is basecalled as the expected canonical nucleotide and is subsequently assigned as either canonical or modified (above the selected modification probability threshold). Reads in which the position is basecalled as a different nucleotide, contains a deletion, is assigned another modification, or does not meet the selected modification probability threshold are therefore excluded from the proportion calculation. For the DRACH-context m^{6}A model, modification calling additionally depends on preservation of the required sequence context, such that reads containing a non-matching base within the motif may also be excluded. Consequently, modification-associated effects on canonical basecalling could introduce bias into the reported modification proportions.

To assess the extent of basecalling mismatches, we examined alternative nucleotide calls at identified modification sites in the filtered DMSO-treated NHDF and HD10.6 datasets and compared, for each site, the percentage of reads containing a non-reference basecall with the predicted modification proportion (***Figure 2—figure supplement 17***). For most sites, non-reference basecalls represented <25% of overlapping reads. Sites at which ≥25% of reads contained a non-reference basecall were comparatively uncommon and were predominantly associated with low predicted modification proportions (<10%). We examined the composition of alternative nucleotide calls at these sites

across the different modification types. Notably, in both NHDF and HD10.6 cells, C accounted for the majority of alternative basecalls at Ψ sites, whereas G predominated at inosine sites (***Figure 2—figure supplement 17***).

These observations indicate that basecalling errors can exclude a subset of reads from modification proportion calculations, with the extent and nature of this effect differing between modification types. The predicted modification proportions may therefore underestimate the true proportion of modified molecules at reported sites, as reads containing a modification but basecalled as an alternative nucleotide would not contribute to the modification estimate. This dependency may also result in some modification sites being missed altogether when basecalling errors reduce the number of reads eligible for downstream modification calling. Thus, the reported proportions should be interpreted as only estimates derived from reads satisfying the respective basecalling and modification-calling criteria.

RNA modification detection rates differ between Dorado models

Dorado basecalling relies on models generated using Remora (https://github.com/nanoporetech/remora, ***Stoiber et al., 2024***) that are frequently updated. This raises an important question as to whether these updates produce notably different outputs that could impact biological interpretation. To address this, we utilized multiple Dorado versions and Remora models to perform modification calling of our NHDF dataset. We restricted our analyses to m^6A and Ψ as these are the only modifications that could be detected in previous models. Here, we observe that predicted m^6A and Ψ proportion distributions are generally consistent between different Dorado versions and models (***Figure 2—figure supplement 18***), albeit with greater numbers of low predicted proportion (<10%) sites detected in older releases. There is a large overlap of the detected ≥10%-proportion of m^6A (DRACH and all-context) and Ψ sites between versions/models, although there are a lot of unique all-context m^6A and Ψ sites identified with just model v5.0.0 and not v5.1.0, that are reduced after applying our modification probability filter (***Figure 2—figure supplement 19A, B and C***). Ψ sites uniquely detected with different models are found in slightly different sequence contexts (***Figure 2—figure supplement 19C and D***). Thus, while RNA modification detection with Dorado is sensitive to changes in the underlying models/versions, this can be effectively controlled through the application of filters that only retain sites with high-confidence modification probability calls.

GLORI and Dorado converge at sites with high predicted modification proportions

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

As shown in the violin plots, Dorado identified similar numbers of m^6A sites using all-context and DRACH-context modification calling (81,012 and 72,372 sites, respectively), with predicted modification proportions strongly reduced upon STM2457 treatment in both cases (***Figure 3B***). The two Dorado approaches differ in the sequence context in which m^6A modifications are called: all-context calling permits m^6A prediction at all adenosines, whereas DRACH-context calling restricts predictions to adenosines within DRACH motifs (DRACH, where D denotes A, G, or U; R denotes A and G, and H denotes A, C, or U). Motif analysis of sites identified using all-context calling revealed nearly 9,000 sites outside DRACH motifs. Accordingly, when all-context calls were restricted to DRACH motifs, the resulting number of sites was similar to that obtained using DRACH-context calling (***Figure 3B***). Unexpectedly, GLORI detects a substantially higher number of sites when using all motifs (140,934) compared to DRACH motifs alone (27,926; ***Figure 3C***). The predominance of non-DRACH sites among the additional GLORI calls may reflect a higher contribution of background signal. Consistent with this interpretation, the overlap between Dorado and GLORI datasets remains virtually unchanged when

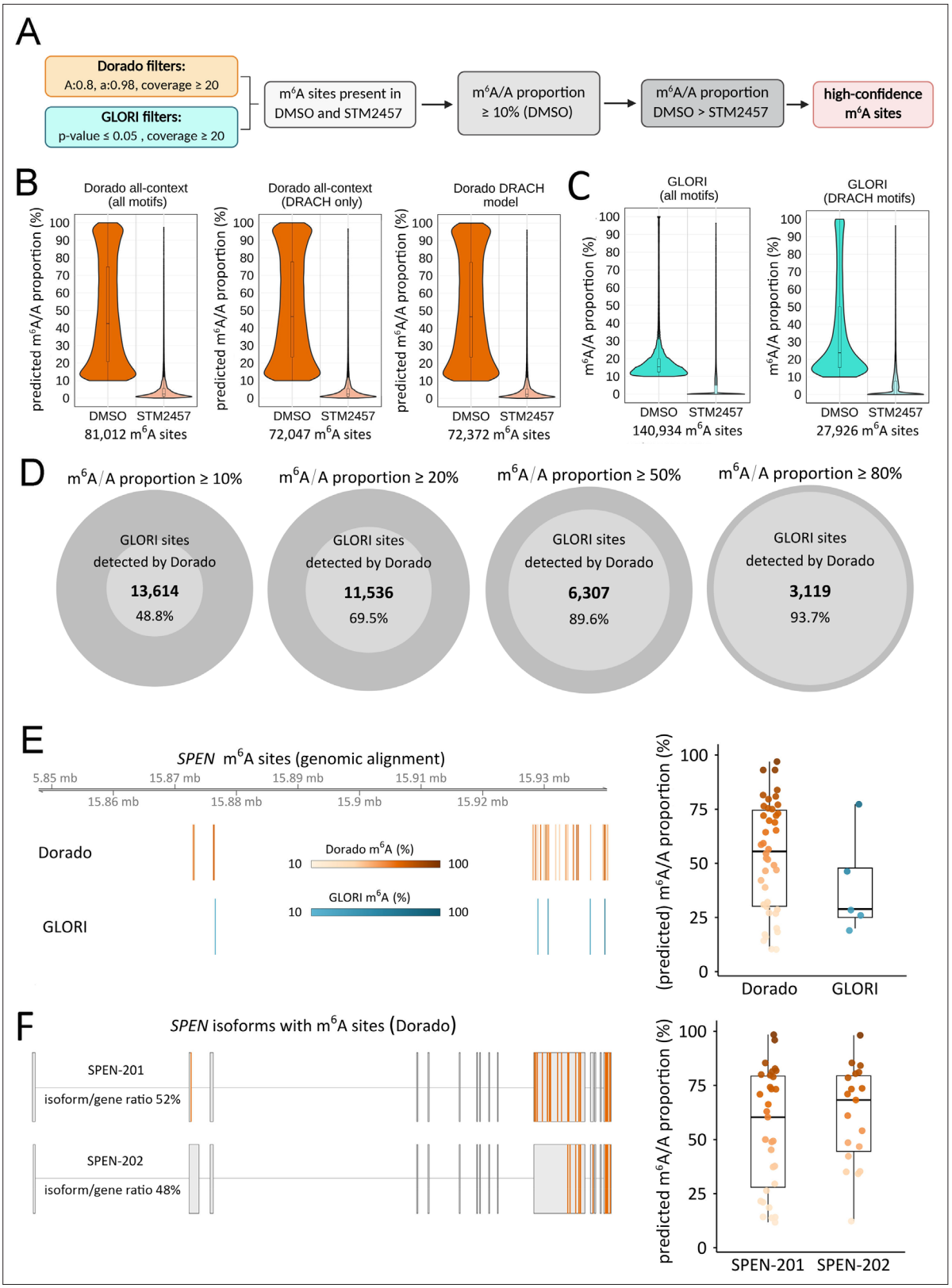

Figure 3. Comparison of m⁶A sites detected by Dorado and GLORI. (**A**) Filtering strategy for Dorado vs GLORI to obtain a high-confidence list of m⁶A sites in normal human dermal fibroblasts (NHDFs). (**B**) Predicted m⁶A/A proportion distributions are plotted from genome-aligned DMSO- or STM2457-treated NHDF datasets using the indicated Dorado models and the filtering strategy shown in (**A**). (**C**) Predicted m⁶A/A proportion distributions from GLORI performed on DMSO- or STM2457-treated NHDFs are plotted, using the filtering strategy shown in (**A**). In addition, m⁶A sites were filtered for

Figure 3 continued on next page

Figure 3 continued

DRACH motifs (right plot). (**D**) Overlap of high-confidence m^{6}A sites detected by Dorado vs GLORI. GLORI and Dorado datasets were filtered using increasing predicted m^{6}A/A proportion cutoffs (from left to right). GLORI sites overlapping with Dorado-identified sites are shown in light grey, whereas sites unique to GLORI are shown in darker grey. (**E**) m^{6}A sites in the *SPEN* gene detected by GLORI and Dorado. Boxplots (right) represent the predicted m^{6}A/A proportions per m^{6}A site shown on the left. (**F**) *SPEN* isoforms with m^{6}A sites, determined by Dorado. Boxplots (right) show the distribution of predicted m^{6}A/A proportions on the m^{6}A sites on *SPEN-201* and *SPEN-202*.

The online version of this article includes the following figure supplement(s) for figure 3:

Figure supplement 1. Comparison of m^{6}A sites detected by Dorado and GLORI-sequencing in normal human dermal fibroblasts (NHDFs) and HD10.6s.

Figure supplement 2. Isoform-level m^{6}A prediction using the Dorado DRACH model in normal human dermal fibroblast (NHDF) data.

comparing GLORI 'all motif' and 'DRACH motifs' (n=13,614) site sets, underscoring that the additional non-DRACH sites in GLORI (n=113,008) contributed little to the observed overlap (***Figure 3—figure supplement 1A***).

We obtained similar results when comparing m^{6}A sites obtained with Dorado and GLORI from HD10.6 cells. Overall, we detected a similar number of m^{6}A sites in HD10.6 cells using the all-context and DRACH-specific Dorado models (77,728 and 67,222, respectively) (***Figure 3—figure supplement 1B***). By contrast, we detected 174,282 vs 27,965 m^{6}A sites using GLORI without and with DRACH filtering, respectively (***Figure 3—figure supplement 1C***).

To evaluate how predicted m^{6}A/A proportion thresholds affect the agreement between GLORI and Dorado, we compared the overlap of high-confidence GLORI-identified sites with Dorado m^{6}A (DRACH context) calls across increasing proportion cutoffs (***Figure 3D***). At the lowest threshold applied (≥10%), 13,614 sites overlapped, representing less than half (48.8%) of GLORI sites. However, as the stringency increased, the relative overlap increased sharply: at ≥20%, nearly 70% of GLORI sites overlapped (11,536 sites), at ≥50% almost 90% overlapped (6307 sites), and at the highest cutoff (≥80%), over 93% of GLORI sites (3119 sites) were also detected by Dorado. These results demonstrate that, while GLORI detects a broad set of potential m^{6}A sites, Dorado increasingly recovers GLORI sites with higher m^{6}A/A proportions.

GLORI relies on the high-efficiency conversion of adenosine to inosine, with overall gene A-to-I conversion rates needing to exceed 95–98% (***Xu et al., 2024***; ***Xie et al., 2025***). Similarly, we observed the conversion rates for our NHDF and HD10.6 datasets to be 95% (***Supplementary file 3***), but we found that small variations in gene conversion rates can strongly influence the sensitivity and specificity of GLORI, as only high-proportion m^{6}A sites overlapped with DRS-derived predictions (***Figure 3D***). When we analyzed the distribution of read coverage among m^{6}A sites detected by both Dorado and GLORI or by GLORI alone (all motifs and DRACH motifs), we did not observe a striking difference. This suggests that read coverage does not account for the large number of sites detected only by GLORI (***Figure 3—figure supplement 1D***). In contrast, comparison of m^{6}A/A proportion distributions revealed an enrichment of high-proportion m^{6}A sites among shared sites, whereas GLORI-only sites were predominantly in the 10–20% range (***Figure 3—figure supplement 1E***). At low m^{6}A/A proportions, the overall gene conversion rate may make it more difficult to distinguish background from true signal. In summary, given the technical limitations of the GLORI dataset, it provides complementary, rather than complete, orthogonal validation.

DRS coupled with Dorado enables m^{6}A prediction at RNA isoform level

Next, we wanted to highlight the advantage of long-read sequencing in terms of mapping reads with isoform-precision to the human transcriptome, which is not possible using GLORI due to its reliance on short-read sequencing. We first confirmed that mapping Dorado-basecalled sites to the human transcriptome results in the detection of similar numbers of m^{6}A sites in both NHDFs and HD10.6 cells (***Figure 3—figure supplement 1F and G***). To assess if DRS coupled with Dorado identifies distinct m^{6}A profiles on different transcript isoforms, we analyzed the *SPEN* gene locus, which was previously shown to contain more than 100 m^{6}A sites (***Liu et al., 2023***). On a genomic level, we detected 42 m^{6}A sites using DRS, while GLORI only detected 5 m^{6}A sites (***Figure 3E***). Out of the seven total isoforms of *SPEN*, *SPEN-201*, and *SPEN-202* were detectable in DRS-derived data, considering a minimal coverage of 20 (***Supplementary file 4***). We found that both isoforms were expressed at equal ratios in our NHDFs. When plotting the m^{6}A sites detected by transcriptome-aligned DRS data, we found that the *SPEN-201* isoform exhibited more m^{6}A sites, albeit with a lower median predicted proportion

compared to the *SPEN-202* isoform (***Figure 3F***). To further illustrate the ability of DRS to resolve isoform-specific m^{6}A patterns, we included three more examples in the supplementary data. The monoexonic *KRTAP1-5*, with a single detectable isoform, demonstrates that DRS identifies more sites than GLORI, especially those with lower predicted m^{6}A/A proportions (***Figure 3—figure supplement 2A***). For *HSPA8*, which has five detectable m^{6}A-containing isoforms, DRS reveals additional m^{6}A sites present in two isoforms that cannot be resolved by GLORI (***Figure 3—figure supplement 2B***). Similarly, *CITED2*, with three detectable m^{6}A-containing isoforms, exhibits distinct isoform-specific m^{6}A modification patterns detectable only by DRS (***Figure 3—figure supplement 2C***). These examples highlight that DRS provides RNA-variant-resolved predictions that are not readily assignable using short-read GLORI data, as illustrated at several loci.

Dorado profiling reveals m^{6}A differences between fibroblast and neuronal datasets

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

We next aimed to conduct a comparative motif analysis of m^{6}A sites detected in NHDFs and HD10.6 cells. Sequence logos confirmed strong enrichment of the canonical DRACH consensus motif for both cell types in modification probability-filtered data, with the GGACU pentamer being the most prominent (***Figure 4D and E***). Quantification of 5-mer distributions revealed highly similar motif preferences across datasets. While the global motif architecture of m^{6}A deposition is conserved between NHDFs and HD10.6 cells, we observed a trend toward higher predicted m^{6}A/A proportions in the HD10.6 dataset, particularly within the most abundant DRACH motifs (***Figure 4F and G***).

Intriguingly, even after filtering, Dorado called m^{6}A at more than 100 motifs that deviate from the canonical DRACH consensus. To understand what these represent, we examined the 15 most frequently detected non-DRACH motifs in NHDFs using the DRACH model. AAAAC and AAAAA were the most abundant, followed by ACACU and GGACG. Most sites within these motifs showed a strong reduction in predicted m^{6}A/A proportions upon STM2457 treatment, indicating that they were deposited by METTL3 (***Figure 4—figure supplement 1A***). Within the top 15 motifs, we found two motifs that conformed to DRAC but not DRACH. Because DRAC sites have been reported as authentic m^{6}A sites (***Körtel et al., 2021***; ***Zhou et al., 2024***; ***Kang et al., 2025***), these calls are likely biologically real (***Figure 4—figure supplement 1A***). For the remaining motifs, the discrepancy that they were sensitive to STM2457 treatment but neither in a DRACH or DRAC context prompted us to ask whether these calls instead reflect systematic limitations of nanopore sequencing and basecalling. Many non-DRACH sites lie one or two nucleotides from a genuine DRACH motif. For example, m^{6}A sites within ACACU motifs are located in proximity to DRACH sites, shifted by one or two positions (***Figure 4—figure supplement 1B***). Another category that contributes to the detection of non-DRACH sites are SNPs. One example is represented by the GGACG motif, in which the final G position of the 5-mer is basecalled as A in 49% and G in 51% of all reads. This makes the sequence context a DRACH motif in half of the reads, and a non-DRACH motif in half of the other reads (***Figure 4—figure supplement 1C***). Another category reflected in the non-DRACH motifs are sites that lie within homopolymer sequences. For example, several sites within AAAAA and AAAAC are reported by Dorado, but because the nanopore signal is poorly resolved within homopolymers, Dorado appears unable to localize the modified adenosine precisely, assigning it to neighboring positions rather than to the adjacent DRACH motif (***Figure 4—figure supplement 1D***).

In parallel, we investigated the occurrence of non-DRACH sites using the all-context model, and as expected, Dorado predicted more than 3000 sites. Among the top 15 motifs, GGACG, GUACU, AGACG, and GGAUU were most abundant (***Figure 4—figure supplement 2A***). As with the DRACH

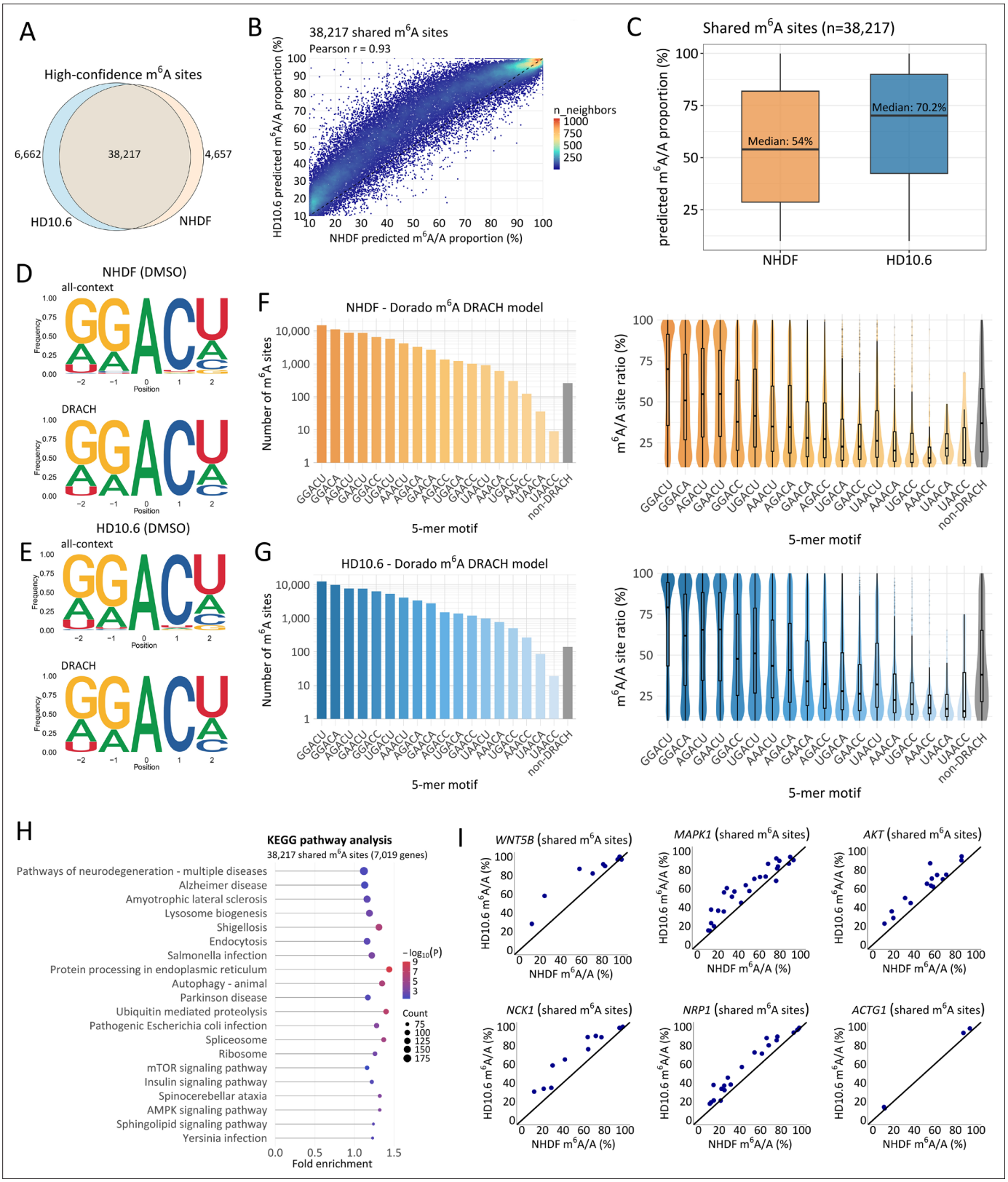

Figure 4. Comparative analysis of m⁶A modification patterns in normal human dermal fibroblasts (NHDFs) and HD10.6 cells using Dorado v0.9.0. (**A**) Overlap of high-confidence m⁶A sites between NHDFs and HD10.6 cells. (**B**) Correlation of predicted m⁶A/A proportions at shared m⁶A sites (n=38,217). (**C**) Boxplots showing the predicted m⁶A/A proportion distributions of the 38,217 detected in both NHDFs and HD10.6 cells. (**D,E**) Motif analysis showing m⁶A (all-context and DRACH-context) motif logos, 5-mer distributions, and predicted modification proportions in NHDFs. (**F,G**) Same as in (**D,E**) but

Figure 4 continued on next page

Figure 4 continued

for HD10.6 cells. (**H**) KEGG pathway analysis (DAVID Bioinformatics Resources) of shared m^{6}A-containing genes. (**I**) Predicted m^{6}A/A proportions of representative examples of individual genes from enriched pathways (**H**). *ACTG1* serves as a housekeeping control.

The online version of this article includes the following figure supplement(s) for figure 4:

Figure supplement 1. Non-DRACH m^{6}A calls predicted by the Dorado DRACH model in normal human dermal fibroblast (NHDF) data.

Figure supplement 2. Non-DRACH m^{6}A calls predicted by the Dorado all-context model in normal human dermal fibroblast (NHDF) data.

Figure supplement 3. Metagene plots showing m^{6}A (all-context and DRACH-context) distribution profiles in normal human dermal fibroblast (NHDF) and HD10.6 data filtered at different predicted m^{6}A/A proportions.

model, GGACG calls derived largely from SNPs or represented genuine m^{6}A sites in a DRAC context. GGAUU, the second most abundant motif, behaved differently, as we found no evidence of miscalling or problematic sequence context. m^{6}A sites within this motif were STM2457-sensitive, suggesting authentic METTL3-dependent methylation; however, this requires further validation (***Figure 4—figure supplement 2B***). Notably, the all-context calls separated into two classes: (i) STM2457-sensitive sites and (ii) STM2457-insensitive sites (***Figure 4—figure supplement 2A***). The insensitive sites were enriched for motifs resembling the METTL16 consensus. ACAGA is a clear example, recurring in *MAT2A*, which is a well-established METTL16 target. Their insensitivity to METTL3 inhibition is consistent with METTL16-dependent deposition, although direct validation, for example in METTL16-depleted cells, will be required to confirm this (***Figure 4—figure supplement 2C***).

Metagene analysis showed, as known previously, that m^{6}A is enriched near the stop codons and in the 3′ UTRs (***Figure 4—figure supplement 3A***). Several studies have examined the relationship between exon architecture and m^{6}A distribution on mRNAs (***Yang et al., 2022***; ***He et al., 2023***; ***Luo et al., 2023***), and so we also examined m^{6}A sites in monoexonic transcripts separately and found that monoexonic transcripts are more broadly m^{6}A-modified across the entire CDS (***Figure 4—figure supplement 3A***). m^{6}A (both all-context and DRACH-context) distribution profiles look less typical for sites with predicted m^{6}A/A proportions lower than 10% in both NHDF and HD10.6 datasets (***Figure 4—figure supplement 3B and C***). For all-context m^{6}A (***Figure 4—figure supplement 3B***), the modification probability filter also proves to be deciding for the distribution profile, together with predicted m^{6}A/A proportion of ≥10%.

To connect shared m^{6}A sites with their functional context and illustrate their distribution on individual transcripts, we conducted a KEGG pathway analysis with the genes that contain shared m^{6}A sites between NHDFs and HD10.6 cells. We used an overlap of all genes that were expressed in both cell types as a background gene set. We observed a significant enrichment in pathways related to neurodegeneration, infection, and intracellular signaling (***Figure 4H***). To further illustrate these findings, we examined representative genes contained in one or more of these pathways. *WNTB5*, *MAPK1*, *AKT1*, *NCK1*, and *NRP1* all showed robust methylation in both NHDFs and HD10.6 cells, with consistently higher predicted m^{6}A/A proportions observed in the neuronal-like context. In contrast, *ACTG1*, included as a housekeeping control not represented in the enriched pathways, displayed stable and comparatively uniform methylation across both cell types (***Figure 4I***).

Together, these results demonstrate that while the global architecture of m^{6}A deposition is largely conserved between primary fibroblasts and neuronal-like cells, neuronal-like cells exhibit globally higher predicted m^{6}A/A proportions, a difference that is readily detectable in individual genes. However, the biological significance of these differences remains unclear, and whether they are associated with differences in RNA abundance, stability, translation, or other cell-type-specific features will require further functional investigation.

Discussion

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

direct comparison with GLORI, we showed that combining DRS datasets with Dorado modification-aware basecalling provides greater scalability, isoform resolution, and reproducibility across different cell types. Further, the simplicity of DRS library preparation enhances consistency, whereas GLORI is constrained by a laborious and technically challenging protocol that is extremely sensitive to adenosine conversion rates (***Shen et al., 2024***). Despite these limitations, concordance with Dorado increased at higher GLORI-reported m^{6}A/A proportions, with the majority of high-proportion GLORI sites also detected in our filtered DRS datasets. By expanding our analysis across multiple human cell types, we further determine that the sites of m^{6}A installation are conserved for the most part, but that the predicted m^{6}A/A proportions show cell-type-specific differences.

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

Prior studies have emphasized the need for isoform-resolved m^{6}A quantification by demonstrating widespread heterogeneity in m^{6}A levels between genes and their isoforms (***Guo et al., 2025***; ***Xie et al., 2025***). Here, we showed that aligning DRS datasets to the human transcriptome reveals differential m^{6}A installation patterns between isoforms, underscoring the importance of accurate, isoform-specific m^{6}A quantification at single-molecule resolution and highlighting the advantage of DRS against short-read sequencing approaches such as GLORI (***Figure 3E and F***, ***Figure 3—figure supplement 2***).

Recently, the ability of Dorado and m6Anet (***Hendra et al., 2022***) to identify m^{6}A in DRS datasets from 293T and HeLa cells was systematically evaluated using IVT RNA and comparisons to published datasets obtained by GLORI and eTAM-seq (***Zou et al., 2025***). In this case, the authors found that Dorado (v0.7.2) outperforms m6Anet but also advocated for the importance of applying filters to reduce false-positive calls. Likewise, we found that IVT RNA served as a valuable negative control to detect background signal. Furthermore, using the RCS as a negative control provided a convenient and simpler alternative to IVT RNA, yielding comparable probability distributions and thereby enabling a built-in control for routine experiments (***Viehweger et al., 2019***). While m^{6}A detection is generally reproducible across different Dorado modification-aware basecalling models, model-specific differences are present, that is model v5.0.0 produces substantially more low-confidence all-context m^{6}A and Ψ calls than v5.1.0 (***Figure 2—figure supplement 10C–E***), highlighting that the models remain under active development and that benchmarking should be repeated when switching to a new model. Recent work further highlights that continued improvements in both nanopore sequencing chemistry and basecalling models significantly enhance RNA modification detection accuracy, emphasizing the importance of revisiting benchmarking and filtering strategies as the models evolve (***Hewel et al., 2025***).

Notably, Dorado v1.1.0 was included in a systematic comparison of computational approaches for the detection of six RNA modification types (***Luo et al., 2026***). While this study provides an important broad assessment of currently available computational methods, the workflow and filtering criteria used to derive high-confidence modification sites from Dorado predictions were not explicitly described. Our findings demonstrate that appropriate filtering and the inclusion of modification-free controls substantially reduce Dorado predictions derived from background signal. Thus, broader

comparisons of modification-detection tools would benefit from consideration of the criteria used to define modification sites, particularly as Dorado models continue to evolve. In this respect, the framework established here complements such benchmarking efforts by providing practical guidance for obtaining high-confidence modification calls from Dorado output.

Focusing on other modifications that are detected by Dorado, we observe that stringent probability and predicted m^6A/A proportion filtering substantially reduces background predictions for inosine and Ψ, with similar filtering strategies also supported by recent studies (***Verstraten et al., 2025***; ***Ghohabi Esfahani et al., 2026***). For inosine, we observed an enrichment of high-probability (>0.99) inosine sites in genome-aligned data from STM2457-treated samples. Further studies are needed to determine whether this reflects increased intron retention or misclassification of modifications by the current inosine model. Nevertheless, very recent systematic benchmarking of Dorado modification-aware basecalling models highlights important limitations that warrant caution when interpreting these predictions (***Diensthuber et al., 2026***). Modification-aware models in particular were shown to exhibit cross-reactivity with chemically related modifications, including substantial cross-reactivity of the Ψ model with $m^1Ψ$ and of the m5C model with hm^5C and ac^4C. Moreover, neighboring modifications can alter the nanopore signal and contribute to predictions arising from background signal. Thus, even high-confidence predictions cannot necessarily be considered unambiguous evidence of the corresponding chemical modification, particularly in the absence of orthogonal validation. These findings further emphasize the importance of modification-free controls and stringent filtering, while also highlighting the need for continued improvement and benchmarking of modification-aware basecalling models.

Emerging evidence indicates that m^6A can occur outside the DRACH consensus in vivo, and our attribution of most non-DRACH GLORI calls to background signal should not be interpreted as a claim that cellular m^6A is restricted to this motif. Our dissection of non-DRACH DRS calls in the filtered datasets revealed that technical and biological explanations coexist. Several classes were clearly artefactual, arising from positional imprecision relative to neighboring DRACH sites, from SNPs that render the context DRACH in only a fraction of reads, or from homopolymeric stretches in which the modified adenosine cannot be localized unambiguously (***Figure 4—figure supplement 1***, ***Figure 4—figure supplement 2***). However, we also found two abundant motifs that conformed to DRAC, but not DRACH, consistent with reports of authentic m^6A at DRAC sites, and STM2457-sensitive GGAUU sites showed no evidence of miscalling, which reflects previous findings (***Körtel et al., 2021***; ***Zhou et al., 2024***; ***Kang et al., 2025***). A recent study showing measurable activity of recombinant METTL3-METTL14 on non-DRACH substrates in vitro further supports our interpretation that the observed GGAUU sites represent valid non-canonical m^6A sites (***Zhang et al., 2024***). Conversely, a distinct class of STM2457-insensitive sites was enriched for METTL16-like motifs, including recurrent ACAGA sites in *MAT2A* (***Su et al., 2022***), pointing to writers other than METTL3 rather than to error. Our results suggest that DRS is in principle well suited to detecting non-consensus m^6A sites, as we included STM2457, which readily makes METTL3-dependent vs. -independent sites discernible. Establishing individual non-canonical sites, nonetheless, requires read-level inspection of the surrounding sequence, awareness of local genetic variation, and future orthogonal validation.

While DRS is the only approach that allows for robust isoform-specific m^6A mapping, comparisons with orthogonal approaches remain essential to validate findings. When comparing Dorado and GLORI using the same input RNA, we unexpectedly observed that only 50% of m^6A sites with predicted proportions of above 10% that were identified by Dorado were also present in the GLORI data, while Dorado predictions captured an increasing proportion of GLORI sites detected at high m^6A/A proportions (***Figure 3D***). One factor that may contribute to this discrepancy is the suboptimal gene-wide adenosine conversion rate of approximately 95% achieved in our GLORI datasets, compared with the 98–99% reported in the original study (***Shen et al., 2024***). Incomplete adenosine conversion could increase background signal and potentially reduce the ability to distinguish m^6A sites with lower predicted proportions. The large number of GLORI-detected sites outside DRACH motifs may similarly reflect increased background signal (***Figure 3C***), although differences in detection sensitivity, sequencing depth, and other methodological biases cannot be excluded based on the current analyses. Further methodological updates to GLORI (***Sun et al., 2025***) may help improve its feasibility and facilitate more robust comparisons with DRS.

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

In summary, using matched RNA samples from different experimental conditions and from different human cell types analyzed with both Dorado and GLORI orthogonal methodologies, we establish a robust pipeline for precise and quantitative transcriptome-wide detection of m^{6}A in DRS datasets. Importantly, we recommend the inclusion of a modification-free control as a best practice for accurate and reliable DRS-based RNA modification detection. Future studies should aim at extending this approach using orthogonal sequencing approaches to benchmark Dorado-based detection of Ψ, m^{5}C, and inosine, as well as the newly released models for 2'-O-methylation.

Methods

Cell culturing and STM2457 treatment

NHDFs (purchased from Lonza Biosciences without donor information) were cultured in Dulbecco's Modified Eagle Medium (DMEM, Gibco), supplemented with 5% heat-inactivated FBS (HI-FBS, Gibco), and 1% penicillin-streptomycin (PS, Lonza) under standard culture conditions. The HD10.6 cell line (a gift from Anna Cliffe, University of Virginia School of Medicine) was originally derived from fetal human dorsal root ganglia and immortalized by overexpression of the v-myc oncogene under tetracycline regulation. All cells are routinely tested for mycoplasma and only used if they test negative. Upon addition of tetracycline and neurotrophins, HD10.6 cells can be differentiated into neuron-like cells that show typical neuronal morphology and nociceptive properties (***Raymon et al., 1999***). For our experiments, we used HD10.6 cells that have been matured for 10 days. For maintenance of undifferentiated HD10.6 cells, T-75 flasks were coated with 17 µg/mL fibronectin in PBS for 15 min at 37 °C before adding the cell suspension. Cells were grown in Advanced DMEM/F12 (Gibco), supplemented with 1 x Neurocult SM1 (STEMCELL Technologies), 2 mM L-glutamine (Gibco), 1 x Primocin (Invivogen), and 10 ng/mL prostaglandin E1 (Sigma). 0.5 ng/mL bFGF (PeproTech) was added freshly to each media change. For maturation, HD10.6 cells were seeded at a density of 25,000 cells/cm² in six-well plates which were pre-coated overnight with 50 µg/mL poly-L-ornithine hydrobromide (Sigma) in 0.5 M borate buffer pH 8.5 (Boston Bioproducts), and 1 µg/mL fibronectin (overnight). After cell attachment overnight, media was changed to Neurobasal media (Gibco), supplemented with 1x Neurocult SM1, 2 mM L-glutamine, 1 x Primocin, 1 µg/mL doxycycline, 50 ng/mL NGF (Alomone Labs), and 25 ng/mL CNTF, GDNF, and NT-3 (PeproTech). Media was changed every 3 days until HD10.6 cells showed complete neuronal morphology after 12 days. NHDFs and matured HD10.6 cells were treated with 30 µM STM2457 (Sigma) or DMSO for 48 hr, harvested in TRIzol (Invitrogen), and stored at –80 °C until RNA isolation.

Cell viability testing

Cell viability was assessed using the CellTiter-Glo assay (Promega), performed according to the manufacturer's protocol. NHDF and HD10.6 cells were seeded in 96-well plates and treated with either DMSO or 30 µM STM2457 for 48 hr, using 10 wells per condition. Following treatment, the culture medium was aspirated and replaced with 50 µL of assay reagent. After a 2-min mixing step and a 10-min incubation at room temperature, luminescence was recorded using a PerkinElmer EnVision 2103 Multilabel Reader. Statistical significance was assessed using an unpaired, two-tailed Student's *t*-test.

RNA isolation

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

GLORI-sequencing

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

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

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

Generation of a transcriptome-wide IVT RNA library

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

Nanopore sequencing of poly(A) RNA

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

Nanopore basecalling

Basecalling was performed using Dorado versions v0.5.0, v0.6.0, v0.7.0, v0.8.0, and v0.9.0 (ONT) in super accuracy mode with modification detection enabled for m^6A (DRACH context) in v.0.5.0 and 0.6.0 (model rna004_130bps_sup@v3.0.1_m6A_DRACH@v1) and 0.8.0 and 0.9.0 (model rna004_130bps_sup@v5.1.0), m^6A (all context) in v0.7.0 (model rna004_130bps_sup@v5.0.0_m6A@v1), 0.8.0 and 0.9.0 (model rna004_130bps_sup@v5.1.0), Ψ in v0.7.0 (model rna004_130bps_sup@v5.0.0_pseU@v1), v0.8.0 and v0.9.0 (model rna004_130bps_sup@v5.1.0), and inosine and m^5C in v.0.9.0 (model rna004_130bps_sup@v5.1.0). Each dataset was processed with adapter trimming. Resulting unaligned BAM files were converted to FASTQ files containing modification tags for every read in the FASTQ read headers, using SAMtools v1.18 (***Li et al., 2009***) [*samtools fastq -TMM,ML*].

Genome and transcriptome level alignments

For genome-level alignment, reads were aligned against a combined human genome reference (GRCh38.p14), and the RCS gene sequence, using minimap2 v2.26 [*minimap2 -ax splice -k14 -y -uf* `--secondary=no`]. For transcriptome-level alignment, the reads were aligned against the human transcriptome (v47), obtained via GENCODE (***Mudge et al., 2025***) [*minimap2 -t 8 ax map-ont -y -L -p*

0.99]. Using SAMtools v1.18, resulting genome- and transcriptome-aligned SAM files were converted to BAM files retaining only primary alignments [*samtools view -b -F2308* and *samtools view -b -F2324*, respectively]. BAM files were sorted [*samtools sort*] and indexed [*samtools index*].

Subsampling aligned BAM files

Aligned BAM files were randomly subsampled to 200,000 reads using *reformat.sh* from BBMap v39.01 [samplereadstarget = 200,000]. Ten independent subsampling iterations were performed for each dataset using different random seeds (sampleseed = 1–10).

Processing of RNA modification calls using Modkit

Probability distributions of the detected sites in human genome- and transcriptome-aligned reads were generated using sample-probs subcommand from Modkit v0.4.1 [sample-probs `--hist --only-mapped --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. The resulting .tsv files were imported into RStudio (v. 2024.12.1+563) and plotted using a custom R (v. 4.4.2) script (Dorado_mod_probability_distributions.R), with the following packages: dplyr (v1.1.4), ggplot2 (v3.5.1), reshape2 (v1.4.4), patchwork (v1.3.0).

To generate extended bedMethyl files with m^{6}A-modified and unmodified base counts from the reads mapping against the potential modification sites, we used Modkit pileup subcommand with the thresholds [`--filter-threshold` *A:0.8* `--filter-threshold` *T:0.8* `--filter-threshold` *C:0.8* `--filter-threshold` *G:0.8* `--motif` *A 0 --ref $REF*] with and without modification threshold filter [e.g. *--mod-thresholds a:0.98*] to generate our 'filtered' and 'unfiltered' datasets, respectively. Canonical nucleotide calls (A, C, G, U) were retained with >80% probability scores for all modification datasets. Modification thresholds were defined by our statistical framework (see below) and were set at 98% for m^{6}A and 99% for all the other modifications. `--motif` was adjusted to the base for which modification status was examined, which further restricted the analysis to only the positions that matched canonical nucleotides in the reference genome. The resulting sites were further filtered according to the read number per site (≥20 reads) for all the downstream analysis.

Statistical framework for modification probability cutoff determination

Modification probability distributions were analyzed using a two-stage statistical framework implemented in R to identify the probability threshold above which sites are considered genuinely modified relative to an unmodified IVT reference. For each sample, the per-bin log2 ratio of modification probability fractions was computed as log2 (sample/IVT), with a pseudocount of 1×10^{-6} added to each value prior to transformation to avoid undefined values. This ratio quantifies fold-enrichment of the sample relative to IVT at each probability bin and is insensitive to differences in absolute signal magnitude across samples. For the first stage, change-point detection was applied to the log2 ratio series using the PELT algorithm (***Killick et al., 2012***) implemented in *changepoint* R package (***Killick and Eckley, 2014***), with a manual penalty value of 1.0 and a minimum segment length of three bins. The first detected change-point defines the end of the baseline region (i.e. the range of probability bins over which sample and IVT distributions are statistically indistinguishable) and the start of the signal region where enrichment above baseline begins. For the second stage, the modification probability cutoff was defined as the first bin at or after the first change-point at which the log2 (sample/IVT) ratio exceeded a pre-specified minimum fold-enrichment threshold. This cutoff represents the lowest modification probability at which a site shows at least 1.5-fold more signal (i.e. log2 ratio >0.585) than would be expected from an unmodified transcript. To characterize uncertainty in the log2 ratio estimates, bootstrap confidence intervals (95%, 2000 iterations) were computed by resampling bins with replacement and recalculating the log2 ratio distribution.

KEGG pathway enrichment analysis

KEGG pathway over-representation analysis was performed using DAVID (***Huang et al., 2009***). Genes harboring shared m^{6}A sites in both NHDF and HD10.6 cells were used as input. All genes detected in both datasets were used as a custom background gene set (~13,000 genes). Pathways were considered significant at a p-value ≤0.05. For visualization, the top 20 KEGG pathways containing the highest number of genes were selected.

Analysis and visualization of predicted modification proportions and distributions

Histograms of predicted modification proportion distributions, as well as metaplots, were generated using RStudio and R v4.4.2 with the following libraries: data.table, patchwork, scales, tidyr, dplyr, ggplot2[4]. Access to scripts used for analysis and visualization is detailed below.

Acknowledgements

We thank our colleagues in New York and Hannover for their continuous input on these studies and to Dr. Shino Murakami for access to an unpublished assay. DPD is supported by a German Centre for Infection Research (DZIF) Associate Professorship and also receives funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2155, project number 390874280. DPD also received support from the NIAID grants R01-AI170583 and R01-AI152543. EL was supported by the NIAID grant R01-AI170583 as well as the Hannover Biomedical Research School (HBRS) and the Center for Infection Biology (ZIB). DO is supported by the Walter Benjamin Programme of the DFG, project 554758329. ACW is supported by NIAID grants R01AI176335 and R01-AI170583.

Additional information

Competing interests

Samie R Jaffrey: Co-founder and/or has equity in Chimerna Therapeutics, 858 Therapeutics, and Lucerna Technologies. The other authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
German Center for Infection Research		Daniel Depledge
National Institute of Allergy and Infectious Diseases	R01-AI170583	Angus C Wilson
Deutsche Forschungsgemeinschaft	EXC 2155 project number 390874280	Daniel Depledge
Deutsche Forschungsgemeinschaft	project 554758329	Denise Ohnezeit
National Institute of Allergy and Infectious Diseases	R01-AI176335	Angus C Wilson
National Institutes of Health	RM1-HG011563	Samie R Jaffrey
National Institutes of Health	R35-NS111631	Samie R Jaffrey
National Institutes of Health	R01-CA186702	Samie R Jaffrey
National Institute of Allergy and Infectious Diseases	R01-AI152543	Daniel Depledge

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Denise Ohnezeit, Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Elene Loliashvili, Conceptualization, Data curation, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing; Gregory Putzel, Data curation, Software;

Ruth Verstraten, Validation, Methodology; Adam T Silhavy, Luke Nicholson, Methodology; Jianheng Fox Liu, Formal analysis, Methodology; Alejandro Pironti, Supervision; Samie R Jaffrey, Supervision, Funding acquisition, Methodology; Daniel Depledge, Conceptualization, Data curation, Formal analysis, Supervision, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review and editing; Angus C Wilson, Conceptualization, Formal analysis, Supervision, Funding acquisition, Investigation, Writing – original draft, Project administration, Writing – review and editing

Author ORCIDs
Denise Ohnezeit https://orcid.org/0000-0002-2346-2545
Elene Loliashvili https://orcid.org/0009-0006-4098-5010
Gregory Putzel https://orcid.org/0000-0002-5211-2608
Ruth Verstraten https://orcid.org/0009-0008-6046-6335
Adam T Silhavy https://orcid.org/0009-0003-4308-9420
Jianheng Fox Liu https://orcid.org/0000-0003-0216-1951
Luke Nicholson https://orcid.org/0009-0005-8203-7249
Alejandro Pironti https://orcid.org/0000-0002-7759-6776
Samie R Jaffrey https://orcid.org/0000-0003-3615-6958
Daniel Depledge https://orcid.org/0000-0002-4292-0599
Angus C Wilson https://orcid.org/0000-0002-5016-4164

Peer review material
Version of Record:
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110672.4.sa1
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110672.4.sa2
Author response https://doi.org/10.7554/eLife.110672.4.sa3

Reviewed preprint v3
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110672.3.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110672.3.sa1
Author response: https://doi.org/10.7554/eLife.110672.3.sa0
eLife Assessment https://doi.org/10.7554/eLife.110672.3.sa3

Reviewed preprint v2
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110672.2.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110672.2.sa1
Author response: https://doi.org/10.7554/eLife.110672.2.sa0
eLife Assessment https://doi.org/10.7554/eLife.110672.2.sa3

Reviewed preprint v1
Reviewer #1 (Public review): https://doi.org/10.7554/eLife.110672.1.sa2
Reviewer #2 (Public review): https://doi.org/10.7554/eLife.110672.1.sa1
Reviewer #3 (Public review): https://doi.org/10.7554/eLife.110672.1.sa0
eLife Assessment https://doi.org/10.7554/eLife.110672.1.sa3

Additional files

Supplementary files
Supplementary file 1. Direct RNA sequencing (DRS) read statistics.
Supplementary file 2. Comparative evaluation of filtering strategies for Dorado.
Supplementary file 3. GLORI adenosine (A) conversion rates.
Supplementary file 4. Direct RNA sequencing (DRS) read coverage of SPEN isoforms.
MDAR checklist

Data availability
Raw fastq and pod5 datasets generated in this study are available via the European Nucleotide Archive (ENA) under the accession number PRJEB101380. R scripts used in the analysis and visualization

of data presented herein are available in the https://github.com/DepledgeLab/HOLDEN repository (copy archived at ***DepledgeLab, 2026***).The Snakemake workflow implementing the GLORI analysis pipeline is available in the https://github.com/gp-micro/glori repository (copy archived at ***gp-micro, 2025***).

The following datasets were generated: ***Ohnezeit et al., 2026***

References

Abebe JS, Verstraten R, Depledge DP. 2022. Nanopore-based detection of viral RNA modifications. *mBio* **13**:e0370221. DOI: https://doi.org/10.1128/mbio.03702-21, PMID: 35579392

Anreiter I, Mir Q, Simpson JT, Janga SC, Soller M. 2021. New twists in detecting mRNA modification dynamics. *Trends in Biotechnology* **39**:72–89. DOI: https://doi.org/10.1016/j.tibtech.2020.06.002, PMID: 32620324

Bicknell M, Harris R, Lee S, blawrence-ont, malton-ont, isovic-ont. 2026. Dorado. 8b8fc5d. Github. https://github.com/nanoporetech/Dorado

Cappannini A, Ray A, Purta E, Mukherjee S, Boccaletto P, Moafinejad SN, Lechner A, Barchet C, Klaholz BP, Stefaniak F, Bujnicki JM. 2024. MODOMICS: a database of RNA modifications and related information. 2023 update. *Nucleic Acids Research* **52**:D239–D244. DOI: https://doi.org/10.1093/nar/gkad1083, PMID: 38015436

Cruciani S, Delgado-Tejedor A, Pryszcz LP, Medina R, Llovera L, Novoa EM. 2025. De novo basecalling of RNA modifications at single molecule and nucleotide resolution. *Genome Biology* **26**:38. DOI: https://doi.org/10.1186/s13059-025-03498-6, PMID: 40001217

Cruciani S, Novoa EM. 2026. The new era of single-molecule RNA modification detection through nanopore base-calling models. *Nature Reviews. Molecular Cell Biology* **27**:10–18. DOI: https://doi.org/10.1038/s41580-025-00896-3, PMID: 41083713

DepledgeLab. 2026. HOLDEN. swh:1:rev:9ff98febfb218702e6b103dc95956e0ff058334b. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:aa844eb7da1959018d3dbe6fbb8379eb29c8cad9;origin=https://github.com/DepledgeLab/HOLDEN;visit=swh:1:snp:bf14a6e081fb5c99051b1eb1f639f0b1153c6af4;anchor=swh:1:rev:9ff98febfb218702e6b103dc95956e0ff058334b

Diensthuber G, Novoa EM. 2025. Charting the epitranscriptomic landscape across RNA biotypes using native RNA nanopore sequencing. *Molecular Cell* **85**:276–289. DOI: https://doi.org/10.1016/j.molcel.2024.12.014, PMID: 39824168

Diensthuber G, Milenkovic I, Llovera L, Milovanovic A, Pelizzari F, Novoa EM. 2026. Systematic benchmarking of dorado basecalling models for RNA modification detection with highly multiplexed nanopore sequencing. *Nucleic Acids Research* **54**:gkag349. DOI: https://doi.org/10.1093/nar/gkag349, PMID: 42306945

Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas K, Jacob-Hirsch J, Amariglio N, Kupiec M, Sorek R, Rechavi G. 2012. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**:201–206. DOI: https://doi.org/10.1038/nature11112, PMID: 22575960

Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P, Warland A, Jordan M, Ciccone J, Serra S, Keenan J, Martin S, McNeill L, Wallace EJ, Jayasinghe L, Wright C, Blasco J, et al. 2018. Highly parallel direct RNA sequencing on an array of nanopores. *Nature Methods* **15**:201–206. DOI: https://doi.org/10.1038/nmeth.4577, PMID: 29334379

Ghohabi Esfahani N, Stein AJ, Akeson S, Tzadikario T, Powell C, Daneshvar Kakhaki P, Jain M. 2026. Assessment of nanopore RNA modification calling in human cell lines and synthetic systems. *Genome Biology* **27**:190. DOI: https://doi.org/10.1186/s13059-026-04096-w, PMID: 42098827

Gilbert WV, Nachtergaele S. 2023. mRNA regulation by RNA modifications. *Annual Review of Biochemistry* **92**:175–198. DOI: https://doi.org/10.1146/annurev-biochem-052521-035949, PMID: 37018844

Gleeson J, Madugalle SU, Wan CY, McLean C, Bredy TW, De Paoli-Iseppi R, Clark MB. 2025. Isoform-level profiling of m^{6}A epitranscriptomic signatures in human brain. *Science Advances* **11**:eadp0783. DOI: https://doi.org/10.1126/sciadv.adp0783, PMID: 40779635

gp-micro. 2025. Glori. swh:1:rev:232b3631b4a7a634147052fcb1b1f0f7f657e9aa. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:fda7aa46c5532decad75809fbc47bb5536045cbf;origin=https://github.com/gp-micro/glori;visit=swh:1:snp:f38f85b3e22531d688f961e78a7cbcddc4930b0d;anchor=swh:1:rev:232b3631b4a7a634147052fcb1b1f0f7f657e9aa

Guo W, Ren Z, Huang X, Liu J, Shao J, Ma X, Wei C, Cun Y, He J, Zhang J, Wu Z, Guo Y, Zhang Z, Feng Z, He J, Wang J. 2025. Single-molecule m^{6}A detection empowered by endogenous labeling unveils complexities across RNA isoforms. *Molecular Cell* **85**:1233–1246. DOI: https://doi.org/10.1016/j.molcel.2025.01.014, PMID: 39922195

He PC, He C. 2021. m6A RNA methylation: from mechanisms to therapeutic potential. *The EMBO Journal* **40**:e105977. DOI: https://doi.org/10.15252/embj.2020105977

He PC, Wei J, Dou X, Harada BT, Zhang Z, Ge R, Liu C, Zhang LS, Yu X, Wang S. 2023. Exon architecture controls mRNA m(6)A suppression and gene expression. *Science* **379**:677–682. DOI: https://doi.org/10.1126/science.abj9090

Hendra C, Pratanwanich PN, Wan YK, Goh WSS, Thiery A, Göke J. 2022. Detection of m6A from direct RNA sequencing using a multiple instance learning framework. *Nature Methods* **19**:1590–1598. DOI: https://doi.org/10.1038/s41592-022-01666-1, PMID: 36357692

Hewel C, Wierczeiko A, Miedema J, Friedrich J, Hofmann F, Weißbach S, Dietrich V, Holthöfer L, Haug V, Mündnich S, Schartel L, Lehmann L, Jensen KL, Diederich S, Sys S, Butto T, Paul NW, Koch J, Lyko F, Kraft F, et al. 2025. Direct RNA sequencing enables improved transcriptome assessment and tracking of RNA modifications for medical applications. *Nucleic Acids Research* **53**:gkaf1314. DOI: https://doi.org/10.1093/nar/gkaf1314, PMID: 41325774

Huang DW, Sherman BT, Lempicki RA. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**:44–57. DOI: https://doi.org/10.1038/nprot.2008.211, PMID: 19131956

Jain M, Olsen HE, Paten B, Akeson M. 2016. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biology* **17**:239. DOI: https://doi.org/10.1186/s13059-016-1103-0, PMID: 27887629

Kang G, Hwang H, Jeon H, Choi H, Chang HR, Yeo N, Park J, Son N, Jeon E, Lim J, Yun J, Choi W, Jo J-Y, Kim J-S, Park S, Kim YK, Baek D. 2025. Comprehensive discovery of m^{6}A sites in the human transcriptome at single-molecule resolution. *Nature Communications* **17**:664. DOI: https://doi.org/10.1038/s41467-025-67417-w, PMID: 41398146

Killick R, Fearnhead P, Eckley IA. 2012. Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association* **107**:1590–1598. DOI: https://doi.org/10.1080/01621459.2012.737745

Killick R, Eckley IA. 2014. Changepoint: an r package for changepoint analysis. *Journal of Statistical Software* **58**:1–19.

Körtel N, Rücklé C, Zhou Y, Busch A, Hoch-Kraft P, Sutandy FXR, Haase J, Pradhan M, Musheev M, Ostareck D, Ostareck-Lederer A, Dieterich C, Hüttelmaier S, Niehrs C, Rausch O, Dominissini D, König J, Zarnack K. 2021. Deep and accurate detection of m6A RNA modifications using miCLIP2 and m6Aboost machine learning. *Nucleic Acids Research* **49**:e92. DOI: https://doi.org/10.1093/nar/gkab485, PMID: 34157120

Köster J, Rahmann S. 2012. Snakemake--a scalable bioinformatics workflow engine. *Bioinformatics* **28**:2520–2522. DOI: https://doi.org/10.1093/bioinformatics/bts480, PMID: 22908215

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**:2078–2079. DOI: https://doi.org/10.1093/bioinformatics/btp352, PMID: 19505943

Linder B, Grozhik AV, Olarerin-George AO, Meydan C, Mason CE, Jaffrey SR. 2015. Single-nucleotide-resolution mapping of m6A and m6Am throughout the transcriptome. *Nature Methods* **12**:767–772. DOI: https://doi.org/10.1038/nmeth.3453, PMID: 26121403

Liu C, Sun H, Yi Y, Shen W, Li K, Xiao Y, Li F, Li Y, Hou Y, Lu B, Liu W, Meng H, Peng J, Yi C, Wang J. 2023. Absolute quantification of single-base m^{6}A methylation in the mammalian transcriptome using GLORI. *Nature Biotechnology* **41**:355–366. DOI: https://doi.org/10.1038/s41587-022-01487-9, PMID: 36302990

Liu JF, Hawley BR, Nicholson LS, Jaffrey SR. 2025. Decoding m^{6}Am by simultaneous transcription-start mapping and methylation quantification. *eLife* **13**:RP104139. DOI: https://doi.org/10.7554/eLife.104139, PMID: 40162895

Luo Z, Ma Q, Sun S, Li N, Wang H, Ying Z, Ke S. 2023. Exon-intron boundary inhibits m6A deposition, enabling m6A distribution hallmark, longer mRNA half-life and flexible protein coding. *Nature Communications* **14**:4172. DOI: https://doi.org/10.1038/s41467-023-39897-1

Luo T, Xu M, Wang M, Chen F, Shi J. 2026. Systematic evaluation of computational tools for multitype RNA modification detection using nanopore direct RNA sequencing. *Nature Methods* **23**:438–450. DOI: https://doi.org/10.1038/s41592-025-02974-y, PMID: 41372661

Meyer KD, Saletore Y, Zumbo P, Elemento O, Mason CE, Jaffrey SR. 2012. Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and near stop codons. *Cell* **149**:1635–1646. DOI: https://doi.org/10.1016/j.cell.2012.05.003, PMID: 22608085

Meyer KD. 2019. DART-seq: an antibody-free method for global m6A detection. *Nature Methods* **16**:1275–1280. DOI: https://doi.org/10.1038/s41592-019-0570-0

Moshitch-Moshkovitz S, Sevilla-Sharon M, Ashwal-Fluss R, Glick-Saar E, Rechavi G, Dominissini D. 2024. mRNA m6A detection. *Nature Reviews Methods Primers* **4**:365-9. DOI: https://doi.org/10.1038/s43586-024-00365-9

Mudge JM, Carbonell-Sala S, Diekhans M, Martinez JG, Hunt T, Jungreis I, Loveland JE, Arnan C, Barnes I, Bennett R, Berry A, Bignell A, Cerdán-Vélez D, Cochran K, Cortés LT, Davidson C, Donaldson S, Dursun C, Fatima R, Hardy M, et al. 2025. GENCODE 2025: reference gene annotation for human and mouse. *Nucleic Acids Research* **53**:D966–D975. DOI: https://doi.org/10.1093/nar/gkae1078, PMID: 39565199

Murakami S, Jaffrey SR. 2022. Hidden codes in mRNA: Control of gene expression by m6A. *Molecular Cell* **82**:2236–2251. DOI: https://doi.org/10.1016/j.molcel.2022.05.029

Ohnezeit D, LoliashviliE, Depledge D. 2026. Establishing benchmarks for quan ta ve mapping of m6A by Nanopore Direct RNA Sequencing. *European Nucleotide Archive*. https://www.ebi.ac.uk/ena/browser/view/PRJEB101380

Rand A, Srinivasan S, Teyssier N. 2026. Modkit. 5cecc3f. Github. https://github.com/nanoporetech/modkit

Raymon HK, Thode S, Zhou J, Friedman GC, Pardinas JR, Barrere C, Johnson RM, Sah DWY. 1999. Immortalized human dorsal root ganglion cells differentiate into neurons with nociceptive properties. *The Journal of Neuroscience* **19**:5420–5428. DOI: https://doi.org/10.1523/JNEUROSCI.19-13-05420.1999, PMID: 10377351

Sendinc E, Shi Y. 2023. RNA m6A methylation across the transcriptome. *Molecular Cell* **83**:428–441. DOI: https://doi.org/10.1016/j.molcel.2023.01.006, PMID: 36736310

Shen W, Sun H, Liu C, Yi Y, Hou Y, Xiao Y, Hu Y, Lu B, Peng J, Wang J, Yi C. 2024. GLORI for absolute quantification of transcriptome-wide m^{6}A at single-base resolution. *Nature Protocols* **19**:1252–1287. DOI: https://doi.org/10.1038/s41596-023-00937-1, PMID: 38253658

Stoiber M, Rand A, Leger A. 2024. Remora. 0787dae. Github. https://github.com/nanoporetech/remora

Su CH, DhananjayaD, Tarn WY. 2018. Alternative splicing in neurogenesis and brain development. *Frontiers in Molecular Biosciences* **5**:12. DOI: https://doi.org/10.3389/fmolb.2018.00012, PMID: 29484299

Su R, Dong L, Li Y, Gao M, He PC, Liu W, Wei J, Zhao Z, Gao L, Han L, Deng X, Li C, Prince E, Tan B, Qing Y, Qin X, Shen C, Xue M, Zhou K, Chen Z, et al. 2022. METTL16 exerts an m^{6}A-independent function to facilitate translation and tumorigenesis. *Nature Cell Biology* **24**:205–216. DOI: https://doi.org/10.1038/s41556-021-00835-2, PMID: 35145225

Sun H, Li K, Liu C, Yi C. 2023. Regulation and functions of non-m6A mRNA modifications. *Nature Reviews Molecular Cell Biology* **24**:714–731. DOI: https://doi.org/10.1038/s41580-023-00622-x

Sun H, Lu B, Zhang Z, Xiao Y, Zhou Z, Xi L, Li Z, Jiang Z, Zhang J, Wang M, Liu C, Ma Y, Peng J, Wang XJ, Yi C. 2025. Mild and ultrafast GLORI enables absolute quantification of m^{6}A methylome from low-input samples. *Nature Methods* **22**:1226–1236. DOI: https://doi.org/10.1038/s41592-025-02680-9, PMID: 40325216

Verstraten R, Cetraro P, Fitzpatrick AH, Alwie Y, Frommeyer YN, Loliashvili E, Stein SC, Häussler S, Ouwendijk WJD, Depledge DP. 2025. Defining expansions and perturbations to the RNA polymerase III transcriptome and epitranscriptome by modified direct RNA nanopore sequencing. *bioRxiv*. DOI: https://doi.org/10.1101/2025.03.07.641986

Viehweger A, Krautwurst S, Lamkiewicz K, Madhugiri R, Ziebuhr J, Hölzer M, Marz M. 2019. Direct RNA nanopore sequencing of full-length coronavirus genomes provides novel insights into structural variants and enables modification analysis. *Genome Research* **29**:1545–1554. DOI: https://doi.org/10.1101/gr.247064.118, PMID: 31439691

Xiao YL, Liu S, Ge R, Wu Y, He C, Chen M, Tang W. 2023. Transcriptome-wide profiling and quantification of N^6-methyladenosine by enzyme-assisted adenosine deamination. *Nature Biotechnology* **41**:993–1003. DOI: https://doi.org/10.1038/s41587-022-01587-6, PMID: 36593412

Xie Y-Y, Zhong Z-D, Chen H-X, Ren Z-H, Qiu Y-T, Lan Y-L, Wu F, Kong J-W, Luo R-J, Zhang D, Liu B-D, Shu Y, Yin F, Wu J, Li Z, Zhang Z, Luo G-Z. 2025. Single-molecule direct RNA sequencing reveals the shaping of epitranscriptome across multiple species. *Nature Communications* **16**:5119. DOI: https://doi.org/10.1038/s41467-025-60447-4, PMID: 40456740

Xu H, Kong L, Cheng J, Al Moussawi K, Chen X, Iqbal A, Wing PAC, Harris JM, Tsukuda S, Embarc-Buh A, Wei G, Castello A, Kriaucionis S, McKeating JA, Lu X, Song C-X. 2024. Absolute quantitative and base-resolution sequencing reveals comprehensive landscape of pseudouridine across the human transcriptome. *Nature Methods* **21**:2024–2033. DOI: https://doi.org/10.1038/s41592-024-02439-8, PMID: 39349603

Yang X, Triboulet R, Liu Q, Sendinc E, Gregory RI. 2022. Exon junction complex shapes the m6A epitranscriptome. *Nature Communications* **13**:7904. DOI: https://doi.org/10.1038/s41467-022-35643-1

Yang L, Tang L, Min Q, Tian H, Li L, Zhao Y, Wu X, Li M, Du F, Chen Y, Li W, Li X, Chen M, Gu L, Sun Y, Xiao Z, Shen J. 2024. Emerging role of RNA modification and long noncoding RNA interaction in cancer. *Cancer Gene Therapy* **31**:816–830. DOI: https://doi.org/10.1038/s41417-024-00734-2, PMID: 38351139

Yankova E, Blackaby W, Albertella M, Rak J, De Braekeleer E, Tsagkogeorga G, Pilka ES, Aspris D, Leggate D, Hendrick AG, Webster NA, Andrews B, Fosbeary R, Guest P, Irigoyen N, Eleftheriou M, Gozdecka M, Dias JML, Bannister AJ, Vick B, et al. 2021. Small-molecule inhibition of METTL3 as a strategy against myeloid leukaemia. *Nature* **593**:597–601. DOI: https://doi.org/10.1038/s41586-021-03536-w, PMID: 33902106

Yu J, She Y, Ji SJ. 2021. m6A modification in mammalian nervous system development, functions, disorders, and injuries. *Frontiers in Cell and Developmental Biology* **9**:2021. DOI: https://doi.org/10.3389/fcell.2021.679662

Zaccara S, Ries RJ, Jaffrey SR. 2019. Reading, writing and erasing mRNA methylation. *Nature Reviews. Molecular Cell Biology* **20**:608–624. DOI: https://doi.org/10.1038/s41580-019-0168-5, PMID: 31520073

Zhang C, Scott RL, Tunes L, Hsieh MH, Wang P, Kumar A, Khadgi BB, Yang YY, Doxtader Lacy KA, Herrell E, Zhang X, Evers B, Wang Y, Xing C, Zhu H, Nam Y. 2024. Cancer mutations rewire the RNA methylation specificity of METTL3-METTL14. *Science Advances* **10**:eads4750. DOI: https://doi.org/10.1126/sciadv.ads4750

Zhou Y, Ćorović M, Hoch-Kraft P, Meiser N, Mesitov M, Körtel N, Back H, Naarmann-de Vries IS, Katti K, Obrdlík A, Busch A, Dieterich C, Vaňáčová Š, Hengesbach M, Zarnack K, König J. 2024. m6A sites in the coding region trigger translation-dependent mRNA decay. *Molecular Cell* **84**:4576–4593. DOI: https://doi.org/10.1016/j.molcel.2024.10.033, PMID: 39577428

Zou Y, Ahsan MU, Chan J, Meng W, Gao SJ, Huang Y, Wang K. 2025. A comparative evaluation of computational models for RNA modification detection using nanopore sequencing with RNA004 chemistry. *Briefings in Bioinformatics* **26**:baf404. DOI: https://doi.org/10.1093/bib/bbaf404