Defining expansions and perturbations to the RNA polymerase III transcriptome and epitranscriptome by modified direct RNA nanopore sequencing

Authors: Ruth Verstraten[1,2], Pierina Cetraro[1], Amy H. Fitzpatrick[1#], Yasmine Alwie[1], Yannick Noah Frommeyer[3], Elene Loliashvili[1], Saskia C. Stein[1], Susanne Häussler[3,4,5,6], Werner J.D. Ouwendijk[7], Daniel P. Depledge[1,2,6*]

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

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

[3]Institute for Molecular Bacteriology, TWINCORE GmbH, Center of Clinical and Experimental Infection Research, a joint venture of the Hannover Medical School and the Helmholtz Center for Infection Research, Hannover, Germany

[4]Department of Molecular Bacteriology, Helmholtz Center for Infection Research, Braunschweig, Germany

[5]Department of Clinical Microbiology, Copenhagen University Hospital – Rigshospitalet, 2100 Copenhagen, Denmark

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

[7] Department of Viroscience,Erasmus Medical Center, Rotterdam, The Netherlands

[#]Present address: School of Veterinary Medicine, University College Dublin, Belfield, Dublin 4, Ireland

*Corresponding author: email: depledge.daniel@mh-hannover.de

ORCIDs

RV – 0009-0008-6046-6335

PC – 0009-0007-8072-0858

AHF – 0000-0002-1883-0489

YA – 0000-0002-7001-5641

YF – 0009–0001–2485-3724

EL – 0009-0006-4098-5010

SCS – 0000-0002-0550-7543

SH – 0000-0001-6141-9102

WJDO – 0000-0001-8393-296X

DPD - 0000-0002-4292-0599

ABSTRACT

RNA polymerase III (Pol III) transcribes cytosolic transfer RNAs (tRNAs) and other non-coding RNAs (ncRNAs) essential to cellular function. However, many aspects of Pol III transcription and processing, including RNA modifications, remain poorly understood, mainly due to a lack of available sensitive and systematic methods for their analysis. Here, we present DRAP3R (Direct Read and Analysis of Polymerase III transcribed RNAs), a modified nanopore direct RNA sequencing approach and analysis framework that enables the specific and sensitive capture of nascent Pol III transcribed RNAs. Applying DRAP3R to distinct cell types, we identify previously unconfirmed tRNA genes and other novel Pol III transcribed RNAs, thus expanding the known Pol III transcriptome. Critically, DRAP3R also enables discrimination between co- and post-transcriptional RNA modifications such as pseudouridine (Ψ) and N^6-methyladenosine (m^6A) at single-nucleotide resolution across all examined transcript types and reveals differential Ψ installation patterns across tRNA isodecoders and other ncRNAs. Finally, applying DRAP3R to epithelial cells infected with Herpes Simplex Virus Type 1 reveals an extensive remodelling of both the Pol III transcriptome and epitranscriptome. Our findings thus establish DRAP3R as a powerful tool for systematically studying Pol III transcribed RNAs and their modifications in diverse cellular contexts.

INTRODUCTION

The Pol III transcriptome comprises in excess of 400 tRNA genes (1) and at least 12 gene families that may comprise large numbers of multi-copy genes and pseudogenes (reviewed in (2)). The ncRNAs encoded by these genes effect critical roles across diverse cellular processes including translation regulation, splicing, tRNA processing, RNA stability, and immune signalling. Crucially, recent ChIP-Seq based studies have inferred the existence of additional Pol III

transcribed genes, most of which are of unknown function (3, 4), indicating that the full extent of the Pol III transcriptome has yet to be determined.

3' end processing is crucial for successful gene expression, and this extends to both Pol II and Pol III ncRNAs. Pol III transcription termination initiates with recognition of a short nucleotide (nt) poly(T) tract located within the coding DNA strand (5). In mammalian genomes, the poly(T) tract is usually 4-6 nt in length (6) and may contain degenerate elements (e.g. TTATTT) (7). These degenerate elements are generally associated with tRNA genes and induce low-level disruption of 3' end processing and result in Pol III transcribed RNAs with extended 3' trailing sequences (7, 8). Following transcription termination, the 3' poly(U) tract of Pol III transcribed RNAs is rapidly bound by the N-terminal domain of the La protein (9) to facilitate clearance of newly-terminated RNA from the transcription complex (10–12), trimming of the poly(U) tail, and the recruitment of additional proteins involved in RNA processing (9).

The role of RNA modifications in the biology of tRNAs and other Pol III transcribed RNAs remains an area of intense study. While the addition of modifications to such ncRNAs is often described as a post-transcriptional process (2), there remains ample evidence that many such modifications are installed either co-transcriptionally or during the association of nascent RNAs with the La protein (reviewed in (13, 14)). The list of modifications that have been reported on select La-associating pre-tRNAs include pseudouridine (Ψ), N2,N2-dimethylguanosine ($m^2{}_2G$), N^1-methyladenosine (m^1A), and m^5-cytosine (m^5C) (15–19). The extent to which nascent Pol III RNAs are modified prior to or during binding of La protein remains unknown as a systematic examination of the nascent Pol III epitranscriptome has yet to be achieved.

The direct RNA sequencing (DRS) methodology developed by Oxford Nanopore Technologies (ONT) enables the end-to-end sequencing of native RNA molecules (20) and, through subsequent interrogation of signal data obtained during sequencing, permits the identification of specific chemically modified nucleotides, including m^6A and Ψ, within the sequence itself (21–25). While primarily designed for the selective capture and profiling of poly(A) RNAs, alternative strategies have been developed and successfully applied to the capture and sequencing of (i) individual RNAs (23), (ii) mature tRNAs (26, 27), and (iii) the global transcriptome (28). While the latter was achieved by size-separation and polyadenylation of all RNAs present in a population, the targeted sequencing of mature tRNAs and, on an individual basis, other non-adenylated RNAs, was achieved through modification of the existing ONT adaptors that are ligated onto RNA targets.

Reasoning that a similar approach would enable of a full systematic survey of the Pol III transcriptome and epitranscriptome, we developed the Direct Read and Analysis of Polymerase III transcribed RNAs sequencing (DRAP3R) methodology. In contrast to the above methodologies, we apply a subtractive approach to deplete non-poly(U) RNAs prior to ligation with a custom poly(U)-targeting adapter designed for the latest RNA004 chemistry, thus enabling the capture of nascent poly(U)-tailed RNAs. We highlight specific computational challenges associated with basecalling, adapter trimming, alignment, and the robust identification of m^6A and Ψ modifications within individual reads. We demonstrate our ability to sensitively and reproducibly obtain high yields of reads that span, and expand, the known Pol III transcriptome across multiple cell types including primary cells. We further demonstrate that discrete patterns of site-specific installation of pseudouridine is a common feature of nascent pre-tRNA and that perturbations to the Pol III transcriptome and epitranscriptome may be induced by external stressors such as viral infection.

RESULTS

Subtractive selection and RNAse A treatment enables efficient capture of poly(U) RNAs

To enable capture of poly(U) RNAs, we developed DRAP3R which uses a custom nanopore adaptor to bind 3' poly(U) tracts of four nucleotides or longer. Our custom adaptor replaces the 10 nt poly(T) overhang of standard nanopore DRS adapters and replaces it with a 6 nt overhang sequence of 3' NNAAAA 5' (Supplementary Fig. S1a), thus enabling the capture of any RNA with a poly(U) tract of 4 nt or longer. The standard DRS approach relies on the isolation of poly(A)$^+$ RNA from total RNA prior to adaptor ligation, reverse transcription, motor protein ligation, and sequencing (20), a selective approach designed to minimize overloading of the flowcell. In the absence of a robust poly(U) enrichment approach, we opted for a subtractive approach in which we deplete poly(A)+ RNAs and ribosomal RNAs (rRNAs) prior to adaptor ligation (Fig. 1a). We stabilize adaptor-ligated RNAs by reverse transcription and subsequently introduce RNAse A to specifically degrade unadapted ssRNAs (circRNAs, mature tRNAs, nascent RNAs, miRNAs, etc.) prior to a final cleanup and motor protein ligation. Resulting libraries are then loaded onto RNA specific (RA) flowcells for sequencing (Fig. 1a). To validate the DRAP3R approach, we generated and sequenced libraries from five samples representing three distinct cell lines: ARPE-19, a spontaneously arising retinal pigment epithelia cell line; NHDF, normal human dermal fibroblasts; and CRO-AP5, a B cell lymphoma cell line containing

latent Epstein Barr Virus (EBV) and Kaposi's sarcoma associated Herpesvirus (KSHV) that expresses high-levels of the well-characterized Pol III transcribed EBER2 RNA.

Each library was sequenced for 24 hrs each using RA flowcells on a MinION Mk.1b and yielded between 0.925 – 2.611 million reads (Fig. 1b) that, following basecalling without trimming by Dorado v0.7.0, showed a modal read length (inclusive of the ~70 nt adaptor sequence) of ~140 nucleotides for the ARPE-19 and NHDF samples (Fig. 1b). For the CRO-AP5 dataset the modal value was ~220 nt which corresponds to the length of the highly abundant Pol III transcribed EBER2 RNA (29). We supplemented this with two runs using the DRAP3R adaptor to sequence *in vitro* transcribed (IVT) RNA generated from cellular RN7SK and EBV EBER2 templates (Supplementary Fig. S1b) and a standard DRS run on the poly(A) fraction retained from one of the ARPE-19 samples (Fig. 1b).

Optimized processing strategies required for poly(A) adapter removal and read alignment

As previously reported (27), minimap2 (30) is suboptimal for aligning shorter read lengths (< 200 nt) and a parameter sweep using minimap2 and BWA (31–33) enabled us to identify the optimal alignment parameters (bwa mem –W 13 –k 6 –T 20 –x ont2d) and increase the proportion of aligned reads from 33.2% (minimap2 –ax map-ont) to > 99.9% (Supplementary Fig. S1c). A comparison of read length vs. alignment length for our optimized alignment strategy confirmed that the optimal bwa mem parameters generally aligned the full length of each read (Supplementary Fig. S1d).

DRS datasets benefit from the removal (trimming) of the 3' adaptor sequence prior to alignment, a procedure typically performed during basecalling with Dorado. While highly effective for standard poly(A) datasets, we observed a significant reduction in the ability of Dorado to remove adaptors from our poly(U) datasets (Supplementary Fig. S1e) with more than half of the reads remaining untrimmed and a further fraction being under-, or over-trimmed (Supplementary Fig. S1f). We attribute this to an inability of the Dorado algorithm to effectively identify and segregate the poly(U) signal. While the adaptor sequence is a fixed length, RNA basecalling remains inaccurate, particularly where DNA is passing through the pore, as evidenced by the variable degree of trimming (30-150 nt, modal value ~65 nt) on both poly(U) and poly(A) datasets (Supplementary Fig. S1f). This variability precludes the use of a fixed distance trim strategy as a proportion of reads would still be over- or under-trimmed. There is thus an unmet need for improved adaptor trimming strategies when using non-poly(A) adaptors.

For this study, we omitted the trimming step entirely, observing that many adaptor sequences were removed as soft-clipped sequences during BWA alignment.

DRAP3R is highly selective for poly(U) RNAs

To evaluate the specificity of DRAP3R, we performed intersect analyses between the genome alignments and existing annotations for RNA Pol II and RNA Pol III transcripts (Gencode v45 (34)) and tRNA genes (Genomic tRNA Database, GtRNADB (1) (Fig. 1c). We required alignments to overlap with at least 25% of an annotation to be classified and counted. Approximately 60-70% of read alignments could be assigned to known RNA Pol III transcribed genes (and pseudogenes) including *RNY5*, *7SK*, *RMRP*, *RPPH1*, *RNA5S*, and vault RNAs. The triple repeat of vault RNA genes (*VTRNA1-1*, *VTRNA1-2*, and *VTRNA1-3*) located on chromosome 5 are well characterized in terms of expression with *VTRNA1-1* the most highly expressed and *VTRNA1-2* and *VTRNA1-3* expressed at increasingly lower levels (35), an observation mirrored in our data (Fig. 1d). As expected, given the generally short lengths of Pol III transcribed RNAs, most read alignments extend to within 5-10 nt of the defined 5' end of transcripts. In general, only a small fraction (0.8-1%) of reads aligned to Pol III pseudogenes. For tRNA genes, GtRNADB predicts 619 tRNA genes within the human genome, 429 of which are reported with high confidence. Restricting our analysis to alignments with mapping quality (mapQ) scores ≥ 10 and tRNA genes with at least five reads aligned, we observe expression of 362 distinct tRNA genes with read alignment counts between 5 to >30000. Of these, 322 tRNA genes were previously classified as high confidence. Of the 40 low-confidence tRNA genes, 24 were supported by 10 or more unique read alignments with six of these supported by 100s to 1000s of read alignments, thus providing strong evidence that these are authentic transcription targets (representative examples shown in Supplementary Fig. S1g). The presence of untrimmed and incorrectly aligned adapter sequences provides a visual indication that the 3' end of these tRNA genes are not accurately recorded in GtRNADB (Supplementary Fig. S1g). We further observed the presence of T -> C mutations at several positions in these tRNAs. Such mutations are often artefacts derived from the presence of pseudouridine which can cause errors during standard basecalling (25).

Of the remaining reads, ~2-3% could be assigned the defined RNA pol II transcripts (Supplementary Fig. S1h), principally the U1 and U2 genes that, similar to RNA Pol III genes, are characterized by short poly(U) tracts at 3' end (36, 37). 1-3% of alignments overlapped with existing Pol III annotations but were excluded from further analysis for being too short (<25%

length of Pol III gene). Finally, ~5% of aligned reads could not be assigned to any existing tRNA or Gencode annotations and thus represent potentially novel RNA Pol III transcribed RNAs. Similar distributions were observed for all DRAP3R datasets except for CRO-AP5 which was dominated by the EBER2 RNA (Fig. 1c). Importantly, very high correlation scores were observed between biological replicates (Fig. 1e) indicating the robustness of DRAP3R.

DRAP3R expands the RNA polymerase III transcriptome

Recent studies reported the identification of eight putative Pol III transcribed genes based on ChIP-Seq binding profiles of select Pol III subunits (3, 4), four of which have since been classified as high-confidence tRNA genes (1) and two as putative tRNA genes. For the two putative genes not classified as tRNAs, one comprises an Alu/SINE element (AluJb_SINE_Alu) reported as a Pol III binding target by Canella *et al*, while the other was named as nc111 by Rajendra *et al*. Here, DRAP3R identified RNAs originating from each of these locations for all three cell lines, confirming the two putative tRNA genes and novel genes each as active sites of transcription (Supplementary Fig. S2a-b, Supplementary Table 1), Sequence analysis confirmed the presence of sequence elements that show high similarity to A and B box motifs, along with short 3' poly(T) tracts between 100-250 nt downstream of the relevant Pol III subunit binding site (Supplementary Fig. S2a-b). This combination of sequence elements is characteristic of type 2 promoter architecture (38, 39).

Further examination of our own datasets subsequently identified six additional novel putative Pol III genes, again expressed in all three cell lines, with similar sequence elements (Supplementary Fig. S2b, Supplementary Table 1). Sequence homology analyses identified no matches to known RNA species. Given that some Pol II transcribed RNAs also terminate at short poly(U) tracts, we treated cultured ARPE-19 cells with differing concentrations (50 – 250 µg/ml) of α-amanitin, an inhibitor of Pol II and Pol III transcription, the latter in a dose-dependent manner. We subsequently assayed expression levels of the 8 candidate Pol III RNAs alongside four defined Pol III RNAs and four defined Pol II RNAs relative to the Pol I transcribed 18s rRNA. At 50µg/ml we observe a specific reduction in expression of all Pol II transcribed RNAs along with two of the candidate Pol III RNAs (novel tx 1 and 6) while all Pol III transcribed RNA levels, including those of the remaining six novel RNAs, were unchanged (Supplementary Fig. S2c). These observations remained consistent at higher α-amanitin concentrations (150 and 250 µg/ml) although the latter showed, as expected, a relative reduction in Pol III transcripts levels, consistent with increasing levels of α-amanitin induced transcription inhibition

(Supplementary Fig. S2c). Together, these data demonstrate that the Pol III transcriptome comprises additional uncharacterized RNAs, the expression of which is conserved across all three cell types.

Pseudouridine modifications are characteristic of nascent RNA Pol III transcripts

Sequencing RNAs in their native state enables the detection of select RNA modifications. The Dorado basecaller (v0.7.0) released by ONT supports native detection of (all-context) *N*^{6}-methyladenosine (m^{6}A) and pseudouridine (Ψ) and enabled us to test the hypothesis that Ψ installation on Pol III transcribed RNAs occurs prior to post-transcriptional processing of the poly(U) tail. Following super accuracy (sup) basecalling, we plotted the basecall accuracy probability distributions for A vs. m^{6}A and U vs. Ψ for each dataset and observed near-identical distributions for canonical A and U nucleotides across all datasets (Fig. 2a, Supplementary Fig. S3). High confidence Ψ basecalls accounted for 0.5-1% of all U sites and were enriched in DRAP3R datasets relative to standard poly(A) DRS and IVT datasets, the latter of which do not contain any modified ribonucleotides. By contrast, high confidence m^{6}A basecalls accounted for 0.5-1% of all A sites in the standard poly(A) DRS dataset but were significantly reduced in the DRAP3R (and IVT) datasets. Two unexpected peaks, around 0.57 and 0.68) in m^{6}A modification probability were observed in the DRAP3R datasets but not in the standard DRS or IVT datasets, potentially suggesting the presence of other adenosine modifications that are misclassified as m^{6}A during basecalling (Fig. 2a, Supplementary Fig. S3).

We reasoned that filtering to retain only high-confidence basecalls would be required to prevent false-positive modification calls and set a baseline filter at 0.98 for both Ψ and m^{6}A. To examine the effect of this filter, we compared the location and stoichiometry of Ψ and m^{6}A in nascent RN7SK RNAs. We observed similar coverage profiles between DRAP3R and IVT datasets (Fig. 2b) but discordant patterns in Ψ predictions (Fig. 2c). Here, many putative Ψ sites are consistently predicted in the unfiltered data but are absent from our filtered dataset containing high confidence Ψ calls. We characterized these as false positives, supported by their also being reported in the IVT dataset. Notably, the well-characterized Ψ250 in RN7SK (40) was observed in all filtered DRAP3R datasets and absent in the IVT dataset. Additional, lower stoichiometry positions were observed at Ψ243 and Ψ247, also in agreement with a prior observation (41). These observations thus provide confidence in the specificity and sensitivity of Ψ detection following our filtering strategy. An analysis of putative m^{6}A sites within RN7SK showed an absence of any installation at this nascent stage (Supplementary Fig. S4a),

consistent with previous reports that m^6A is post-transcriptionally installed on RN7SK RNA (23). A similar analysis was conducted on the EBV EBER2 RNA for which Ψ114 and Ψ160 have been discordantly reported (42, 43). Here, we obtained good coverage of EBER2 in DRAP3R and IVT datasets (Supplementary Fig. S4b) and detected Ψ at position 114 but not 160 (Supplementary Fig. S4c). We observed no evidence of m^6A installation on the nascent RNA (Supplementary Fig. S4d).

We next examined the distribution of Ψ and m^6A stoichiometries across all filtered datasets (Fig 2d, Supplementary Fig. S4e), segregated into two distinct groups, Pol III transcribed pre-tRNAs and all other Pol III transcribed RNAs. For both Ψ and m^6A, a majority of putatively modified positions had stoichiometries below 10%. For Ψ, ~32% of Ψ installation sites had stoichiometries above 10% while for m^6A this value was reduced to ~5%. Notably, larger numbers of Ψ and m^6A sites were predicted in pre-tRNAs relative to other nascent Pol III RNAs. These observations remained consistent across all datasets, leading to the conclusion that high stoichiometry Ψ, but not m^6A, is a feature of nascent Pol III transcribed RNAs.

Metaplots visualize the density of a given modification across the body of RNAs. To validate the accuracy of the high-confidence Ψ and m^6A basecalling, we observed that, consistent with previous studies (43–45), m^6A and Ψ distributions across mRNAs in our poly(A) DRS dataset showed characteristic enrichment of m^6A at the boundary of the CDS and 3' UTR and a general enrichment of Ψ across the CDS (Supplementary Fig. S4f). For the DRAP3R datasets, we observed that m^6A and Ψ installation is enriched at specific regions in pre-tRNAs and more generally toward the 5' end of other nascent Pol III transcripts (Fig. 2e) and, in the case of Ψ, regardless of whether filtering for only high stoichiometry positions (>10%) (Supplementary Fig. S4g).

DRAP3R enables identification of Ψ and m^6A on nascent small nuclear RNAs

While some prior reports have identified individual pre-tRNAs with modified ribonucleotides(15–19), the extent to which the nascent Pol III transcriptome is modified prior to or during binding of La protein has yet to be systematically evaluated. To address this, we examined the installation profiles for m^6A and Ψ across the nascent Pol III transcribed RNAs captured by DRAP3R. We restricted our analysis to sites at which m^6A/ Ψ stoichiometry exceeded 10% in at least one dataset. Excluding pre-tRNAs, we detected 11 m^6A sites (Supplementary Fig. S5a) and 24 Ψ sites (Fig. 3a) across fifteen transcript isoforms. The

methyltransferase(s) responsible for m^6A installation on nascent Pol III transcribed RNAs is not known and no sequence motif could be derived from these sites (Supplementary Fig. S5b). We detected three (Ψ31, Ψ40, and Ψ86) of four Ψ modifications previously reported for U6 RNA (46–50) and two sites in U6atac (Ψ9, Ψ83), the latter of which was also previously reported (51) (Fig. 3a). Interestingly, while Ψ211 is reported for RN7SL RNAs (43, 52), we observed no evidence of Ψ in these RNAs in any of our datasets. Here, we additionally report the discovery of Ψ sites located in RNA Y, RMRP, BC200, RNA5S, and the conserved 5' sequences of VTRNA1-1, VTRNA1-2, and VTRNA1-3 (Fig. 3a). No motif could be established for the Ψ sites (Supplementary Fig. S5c), many of which are known to be installed by distinct pseudouridine synthases (reviewed in (53)). Notably, we observed a complete absence of Ψ sites on the nascent Pol II transcribed U1 and U2 RNAs that are also captured by DRAP3R, indicating that Ψ installation on nascent poly(U) RNAs is generally restricted to pre-tRNAs and select Pol III genes.

Pseudouridine installation is a characteristic marker of pre-tRNAs

We next applied DRAP3R to generate a quantitative map of Ψ installation across all pre-tRNAs (Fig. 3b). Ψ sites at high stoichiometry were consistently detected at positions 13, 27-28, 31/32, 35/36, 55, and 65 with positions 55 and 13 having the highest stoichiometry across all datasets. Discrete combinations of Ψ installation were observed for different pre-tRNAs, both within and between isotypes, as shown for glutamic acid tRNA genes (Fig. 3c). These data are in agreement with a recent study of Ψ sites across the human transcriptome (43) while also providing greater specificity by identifying discrete patterns at the level of individual tRNA genes and showing that significant pseudouridylation of pre-tRNAs occurs during transcription and/or La binding (Fig. 3d). This ability of DRAP3R to discriminate between pre-tRNAs at the gene level provides a significant advantage when compared to methods that require collapsing to the level of isodecoder or isotype (27, 43). Based on the presence vs. absence of Ψ at positions 13 and 55, we ascertained consensus motifs for pre-tRNA pseudouridylation by Pus7 (UVΨAR) and Pus4 (VNΨCR) (Fig. 3e), respectively, the former matching that of a previous report (43).

Having established that pre-tRNAs carry pseudouridine modifications, we next asked whether other known modifications of tRNAs are evidenced at the nascent pre-tRNA level or are only manifest as part of post-transcriptional processing. While modifications such as 1-methylguanosine (m^1G) are not yet available for native identification during basecalling, prior studies have demonstrated that RNA modifications are frequently associated with increased

basecalling errors around the site of the modification (26, 27). To enable comparisons between pre-tRNA and mature tRNAs we adapted the nano-tRNA-Seq (27) method for the latest DRS RNA004 chemistry, sequenced mature tRNAs from the same cultured ARPE-19 cells and observed near-identical Ψ installation profiles across glutamic acid isodecoders (Fig. 3f). We then examined differences in basecall error rates for DRAP3R and nano-tRNA-Seq datasets at each individual position across the glutamic acid tRNA-TTC-1 gene (Fig. 3g) and observed notable differences in error rates at positions 9 (m^1G), 20 (dihydrouridine, D), 35 (5-methoxycarbonylmethyl-2-thiouridine, mcm^5s^2U) and 50 (m^1G), indicating that these positions are post-transcriptionally modified. By contrast, error rates at positions 13, 32, 55, and 59 (all Ψ) were similar in both datasets, confirming this modification is installed co-transcriptionally or while bound by La. Analysis of all glutamic acid isodecoders yielded similar overall results (Fig. 3h), while also demonstrating isodecoder specific differences e.g. between Glu-TTC-2 and Glu-TTC-4 (Supplementary Fig. S5d). This demonstrates the ability of DRAP3R, in combination with nano-tRNA-Seq, to distinguish between nascent and post-transcriptional tRNA modifications, even where basecalling models are lacking the ability to natively detect specific modifications.

DRAP3R reveals infection with HSV-1 increases pre-tRNA levels while reducing pseudouridine stoichiometries

We and others have previously demonstrated the ability of HSV-1 to regulate m^6A installation on mRNAs (54, 55) and that HSV-1 infection leads to changes in pre-tRNA and mature tRNA levels (56), but no study has addressed whether pseudouridine installation is also regulated. To address this, we applied DRAP3R to RNA collected from two biological replicates of HSV-1 infected ARPE-19s collected at 6 hours post infection (hpi) and one at 12 hpi. We obtained similar numbers of reads and read length distributions to the other DRAP3R datasets (Fig. 1b, Fig. 4a) while classification showed a slight increase in the relative proportions of pre-tRNA and other Pol III transcripts captured (Fig. 1c, Fig. 4b). The biological replicates again showed significant correlation (Supplementary Fig. 7a) while basecall accuracy probability distributions for A vs. m^6A and U vs. Ψ (Supplementary Fig. 7c) and Ψ stoichiometry distributions (Supplementary Fig. 7d) were also consistent with the previous DRAP3R runs (Fig. 2d). A scatter plot analysis showed that HSV-1 infection reduces expression of detected Pol II transcribed RNAs, principally U1 and U2, in a time-dependent manner (Fig. 4c). A differential gene expression analysis for Pol III transcripts showed, in agreement with a previous study (56), a significant upregulation in expression of ~50 pre-tRNAs and significant downregulation of ~50

other Pol III transcribed genes (Supplementary Fig. 7b). We further observed strong correlations in site-specific Ψ frequencies between both uninfected (Supplementary Fig. 7e) and infected (Supplementary Fig. 7f) biological replicates. Notably, comparisons of site-specific Ψ stoichiometries between uninfected and HSV-1 infected ARPE-19s revealed decreasing Ψ stoichiometries that correlate with the length of the infection (Fig. 4d) and that this is not a result of interferon induction (Supplementary Fig. 7g).

DISCUSSION

Pioneering studies in the 1970s/1980s provided the first evidence that nascent pre-tRNAs originating from Pol III transcription are modified either co-transcriptionally or while bound to protein La as a first step in post-transcriptional processing (15–19). The role of such modifications in orchestrating the first steps of processing remain poorly characterized, limited by the lack of a systematic method for examining RNA modifications on individual nascent RNAs. This lack is driven by the challenges in capturing small RNAs for analysis and the difficulty in distinguishing between such RNAs at different stages of post-transcriptional maturation. A direct consequence of this is that only a minority of human tRNAs, and far fewer pre-tRNAs, have been characterized in terms of their RNA modifications (57). Further, efforts to generate an atlas of the full Pol III transcriptome has been hindered by the general reliance on ChIP-Seq based studies to putatively identify previously unreported Pol III transcribed RNAs(3, 4). This similarly impacts the ability to screen for changes induced in the Pol III transcriptome and epitranscriptome by internal and external stressors.

To address these limitations, we developed DRAP3R, the first nanopore DRS based method that specifically and sensitively captures RNAs with short 3' poly(U) tails, a unifying characteristic of all known RNAs transcribed by Pol III. In contrast to the limited number of existing nanopore DRS-based methods for profiling ncRNAs, which generally rely on adding poly(A) tails in vitro prior to adapter ligation, we have devised a custom adaptor that specifically targets poly(U) tails and combined this with a subtractive approach to remove non-poly(U) tailed RNAs. In combination with native basecalling and stringent filtering of Ψ and m^6A modifications, we demonstrate our ability to delineate and measure changes in the discreet patterns of modifications installed on nascent ncRNAs, thus enabling a clear discrimination of modifications installed co-transcriptionally or during La-binding from those which are added during subsequent post-transcriptional processing steps. We further identify multiple novel RNAs that are generally insensitive to low concentrations of α-amanitin treatment, indicating they are

386 transcribed by Pol III. The two identified RNAs that were sensitive to α-amanitin also require
387 further study to deduce whether these are transcribed RNAs by Pol II or simply require Pol II
388 activity for Pol III transcription to occur, as has been described for U6 RNAs (58).

389 DRAP3R is simple to implement with the entire protocol, including sequencing, requiring
390 2-3 days. The high data yields obtained, combined with the promise of recently proposed DRS
391 multiplexing approaches (59, 60), further increase the potential of our method. We should,
392 however, note two current limitations that still need resolving. Firstly, a recurrent issue observed
393 is that adapter-trimming during basecalling of DRAP3R datasets with Dorado frequently leads to
394 over-trimming while retaining the adapter sequence can lead to misalignments that extend
395 beyond the 3' end of sequenced RNAs. There is thus an unmet need for improved adapter
396 identification and removal approaches that are not reliant on the presence of poly(A) tails.
397 Secondly, while short poly(U) tails serve as markers of nascent RNAs transcribed by Pol III and
398 are rapidly removed upon La binding, the multi-copy U6 RNA exhibits a unique processing
399 pathway. Here, nascent U6 RNAs undergo polyuridylation to generate an extended poly(U) tail,
400 which is subsequently trimmed during post-transcriptional processing, ultimately yielding mature
401 U6 RNAs with poly(U) tails ranging from 5 to 12 nucleotides in length (61). It is thus likely that a
402 proportion of U6 RNAs captured by the DRAP3R protocol are not purely representative of
403 nascent RNAs.

404 The recent development of other nanopore-based methods for targeted analysis of
405 mature tRNAs (26, 27) or the entire cellular transcriptome (28) offer significant potential in
406 combination with DRAP3R, allowing for the comparison of relative abundances and the
407 modification status of nascent and mature versions of tRNAs and other ncRNAs transcribed by
408 Pol III. This is demonstrated by our integration of DRAP3R and nano-tRNA-Seq which reveals
409 that pseudouridine is generally installed co-transcriptionally or during La binding while
410 modifications such as m^1G are added at a later post-transcriptional stage. This finding highlights
411 the value of comparative sequencing of nascent and mature Pol III transcribed RNAs in
412 determining the timing of modification installation and we anticipate the value of this approach
413 will only increase in the future as nanopore basecallers are updated to include native detection
414 of additional RNA modifications.

415 Another potential application of DRAP3R is in the study of Pol III-mediated Pol II
416 transcription. Rajendra *et al* (3) recently reported on the transcription of short-lived poly(U)-tailed
417 Pol III transcribed RNAs occurring from a subset of Pol II promoters as a necessary step for Pol
418 II transcription initiation. While such RNAs were not observed in our datasets, presumably due

to their instability and rapid degradation, modified experimental designs that prevent rapid turnover of such RNAs could provide sufficient targets for DRAP3R. Similarly, where internal (i.e. inborn errors) or external stressors (i.e. viral infection) may exert impacts on Pol III transcriptome and RNA modification installation, the combination of DRAP3R and nano-tRNA-Seq would allow researchers to establish at which co- and post-transcriptional steps the presence of specific modifications are required.

In summary, DRAP3R is a fast, sensitive, and accurate method for quantitative profiling of the RNA Pol III transcriptome and the detection of RNA modifications at single-nucleotide resolution on individual RNAs. We anticipate a wide range of applications for this method in broadening our understanding of Pol III transcription and modification dynamics, including the role of the Pol III transcriptome and epitranscriptome in infection biology and the study of human disease induced by inborn errors in ncRNAs and Pol III subunits.

Funding/Acknowledgements

DPD is supported by a German Centre for Infection Research (DZIF) Associate Professorship and the NIAID grants R01-AI170583 and R01-AI152543. EL is supported by the NIAID grant R01-AI170583. DPD and SH receive funding from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2155 - project number 390874280.” PC, EL, and YF were supported by the Hannover Biomedical Research School (HBRS) and the Center for Infection Biology (ZIB). WJDO is supported by the National Institute of Allergy and Infectious Diseases (NIAID) of the National Institutes of Health (NIH) under Award Number R01-AI151290. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors would additionally like to thank Thomas Hennig and Lars Dölken for useful discussions about this project.

Methods

Cell culture, RNA extraction, and virus preparation

ARPE-19 cells were grown in DMEM/F12 with 8% FCS + 1% P/S +1% L-glutamine. Cells at 70% confluency were treated with 50, 150 or 250 µg / ml of α-amanitin for 24 h. Cells were harvested by adding Trizol and RNA extracted according to manufacturer’s protocol with

addition of GlycoBlue (Thermo-Fisher) during the first precipitation step. cDNA was synthesized using the Quantitect reverse transcription kit (Qiagen) followed by q-PCR using the QuantiNova SYBR green (Qiagen) and primers (table X) using 18sRNA to normalize. The qPCR program consisted of 40 cycles of 10 sec at 95^0C and 30 sec at 60^0C and a melting curve at the end to check for product specificity. For HSV-1 infections, ARPE-19 cells were grown to 90% confluency and cell-free HSV-1 strain KOS added at an MOI of 10 in 2% FCS media and incubated for 1h at 37°C with 5%CO_2. After 1 h the infection media was removed and fresh media with 8% FCS was added and cells were harvested at 6 or 12 hpi. ARPE-19 cells grown to 70% confluency were treated with interferon-α2a and interferon-γ (from ImmunoTools) 500 and 100 U respectively for 24h before harvesting. CRO-AP5 cells were obtained from the DSMZ (ACC 215) and grown in RPMI 1640 (Gibco) with 20% FCS (Sigma).

DRAP3R

To remove poly-adenylated RNAs from 20 - 25 µg of total RNA, we used the Dynabeads™ mRNA Purification Kit (Invitrogen) and retained the supernatant after the first pulldown of the magnetic beads as the non-poly(A) fraction and re-isolated the RNA by ethanol precipitation. Following poly(A) depletion, 10 µg of remaining RNA was used for rRNA depletion using the RiboMinus ™ Transcriptome Isolation Kit (Invitrogen), also followed by an ethanol precipitation. The concentration of the resulting poly(A)$^-$/rRNA$^-$ RNA was determined by Qubit. We used 500ng of this RNA for Direct RNA Sequencing library preparation using the RNA-SQK004 chemistry (Oxford Nanopore Technologies), with some modifications to the standard protocol. Here, we replaced the RTA adapter with our custom DRAP3R adapter (Supplementary Fig. 1a). The DRAP3R adapter replaces the 10nt poly(T) sequence with 3' NNAAAA 5' for specific capture of poly(T) RNAs. To prepare the adapter, we combined 1.4 µM each of RNAse-free HPLC purified Oligo A and DRAP3R Oligo B in 100 µl of buffer (10 mM Tris-HCL pH 7.5, 50mM NaCL), heated this to 95^0C for 2 min followed by cooling (0.1°C /sec) to 4°C. During the first adapter ligation we also omitted the polyadenylated RNA calibration strand (RCS). Following reverse transcription with SuperScript III (Invitrogen), RNAse A (43 µg / ml final concentration) in the presence of 0.4M NaCl was added and incubated at 37°C for 15 min to remove all remaining single-stranded RNA. RNAse A was subsequently inactivated by proteinase K treatment for 15 min at 37°C, followed by 70°C for 15 min. Ligated of the RLA adapter was performed according to the ONT protocol with all purification steps using AMPure XP beads at 2X concentration. The resulting sequencing library was loaded onto an RNA-specific (RA)

minION flowcell and sequencing performed for 24h with MinKNOW parameters modified to disable q-score filtering and reduce the minimum read length for retention to 20 nt.

In vitro transcription of RN7SK and EBER

To produce templates from which *in vitro* transcribed (IVT) RNAs could be generated, we used genomic DNA extracted from MeWo or CRO-AP5 cells on which a PCR was performed using Q5-high-fidelity polymerase and custom primers where the forward primer contains the T7 promoter (Supplementary Table 2). This was then followed by transcription using the TranscriptAid T7 High Yield Transcription Kit (Thermo-Fisher) following the manufacturers instruction for high yield. Note that only canonical dNTPs (A, C, G, and U) were provided. The resulting RNA was isolated using the Monarch RNA clean up kit 50 µg (NEB) and yielded ~100 µg of IVT RNA. 500 ng was used as input for the DRAP3R DRS library preparation, omitting the poly(A) and rRNA depletion steps.

Nanopore Sequencing of poly(A) RNA

The poly(A) fraction was isolated from 25 ug of total input RNA using the Dynabeads™ mRNA Purification Kit (Invitrogen), according to manufacturer's protocol. 300ng of poly(A) RNA was used as input for the standard nanopore DRS RNA004 protocol with 45 ng loaded onto an RNA specific RA flowcell prior to sequencing on a MinION Mk.1b. for 24 hours.

Genomes and databases

The human genome (GRCh38.p14) was obtained via Ensembl (62) while the human transcriptome (v45) was obtained via GENCODE (63). The human tRNA database (GRCh38 Dec 2013) was downloaded from GtRNADB (1). The HSV-1 strain KOS genome (KT899744, (64)) and EBV strain B95-8 genome (AJ507799.2, (65)) were downloaded from Genbank. Hybrid genome indexes (e.g. HG38 + HSV-1 strain KOS) were generated by merging fasta files and indexing using BWA [*bwa index -a bwtsw hybrid.fasta hybrid.fasta*] or minimap2 [*default parameters*] (30).

Nanopore basecalling

Basecalling was performed using Dorado v0.7.0 (Oxford Nanopore Technologies) in super-accuracy mode with modification detection enabled for Ψ and m^6A. Each dataset was processed with and without adapter trimming with the latter used for all major downstream analyses. Resulting unaligned BAM files were parsed using SAMtools (66) to generate fastq files with modification tags for each read included in the fastq read headers [*samtools fastq -TMM,ML*].

Genome and transcriptome level alignments

To define an optimal alignment strategy fastq reads were aligned against the HG38 genome assembly using distinct strategies including BWA SW v0.7.17 (33) [*bwa bwasw, bwa bwasw -z 10 -a2 -b1 -q2 -r1*], BWA MEM v0.7.17 (31) [*bwa mem -W 13 -k 6 -x ont2d, bwa mem -W 13 -k 6 -T 20 -x ont2d, bwa mem -W 13 -k 6 -T 10 -x ont2d, bwa mem -W 9 -k 5 -T 10 -x ont2d*], and Minimap2 v2.26 (30) [*minimap2 -ax map-ont, minimap2 -ax splice -k15 -uf -L*]. Prior to alignment, the fastq dataset was filtered to remove all reads longer than 1,000 nt and divided into bins comprising 250k reads. Resulting SAM files were processed (view, merge, sort, index) using SAMtools v1.18 with only primary alignments [-F 2308] retained. For all downstream analysis, the optimal genome alignment strategy was defined as *bwa mem -W 13 -k 6 -T 20 -x ont2d*.

For alignment of standard nanopore DRS datasets again the human transcriptome we used minimap2 [*-ax map-ont -y -L -p 0.99*]. SAM files were subsequently processed (view, merge, sort, index) using SAMtools (66) with only primary alignments [-F 2324].

Processing of RNA modification calls using Modkit

Sample probability distributions were obtained using sample-probs subcommand from Modkit v0.4.1 (https://github.com/nanoporetech/modkit) with the following parameters [*--hist --only-mapped --percentiles 0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.85,0.9,0.95 --num-reads 250000*]. Note that the –num-reads value was reduced to 20000 for each IVT dataset. Relevant columns were extracted from the probabilities.tsv output files and plotted. To generate extended bedMethyl files for which base modifications counts from every sequencing across each reference genomic position are tabulated, we used the Modkit pileup subcommand [*modkit pileup in.bam out.bed --*

filter-threshold 0.8 --mod-thresholds a:0.98 --motif A 0 --ref HG38.fasta]. Here, we retain canonical nucleotide calls (ACGU) with >80% confidence scores while retaining only Ψ and m^6A calls (17802 and a, respectively) with > 98% confidence scores, as informed by our analysis of sample probability distributions (Fig. 2a). We further restricted modification call analysis to positions matching the canonical nucleotide in the reference genome (e.g. A for m^6A and T for Ψ).

Intersect analyses and generation of count data

BEDtools v2.26.1 (67) intersect [*-s –F 0.25 –wao –a alignment.bed –b specific.database.bed*] was used to classify read alignments in bed and bedMethyl datasets as deriving from transcription of Pol III primary genes, Pol III tRNA genes, Pol III pseudogenes, or Pol II genes. Read alignments with no overlap against a combined Pol II & Pol III gene database were identified using exclusion parameters [*-s –v –a alignment.bed –b combined.database.bed*]. Novel Pol III transcription was only considered where 50 or more read alignments against the same unannotated locus were reported. Reads aligning to HSV-1 were identified by filtering for the HSV-1 strain KOS ID in the first (chromosome) column of the alignment.bed files. Count data for each gene was collected by summing the total number of primary alignments against each gene.

Analysis and visualization of modification stoichiometries

Stoichiometry distribution histograms, metaplots, transcript-level plots, and scatter plots were generated using Rstudio/Posit and R v4.3.3 with the following libraries: data.table, patchwork, scales, tidyr, Dplyr, readr, tools, gdata, stringr, Biostrings, EnhancedVolcano, DESeq2 (68), GenomicFeatures (69), ggplot2 (70), Gviz (71). Access to scripts used for analysis and visualization is detailed below.

Data availability

Raw pod5 datasets generated as part of this study are available via the ENA/SRA under the accession number PRJEB81484. R scripts used in the analysis and visualization of data presented herein are available from the https://github.com/DepledgeLab/DRAP3R repository.

References

1. Chan PP, Lowe TM. 2009. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. Nucleic Acids Res 37:D93–D97.

2. Zhou S, Van Bortle K. 2023. The Pol III transcriptome: Basic features, recurrent patterns, and emerging roles in cancer. Wiley Interdiscip Rev RNA 14:e1782.

3. Rajendra K, Cheng R, Zhou S, Lizarazo S, Smith DJ, Van Bortle K. 2024. Evidence of RNA polymerase III recruitment and transcription at protein-coding gene promoters. Mol Cell 84:4111-4124.e5.

4. Canella D, Praz V, Reina JH, Cousin P, Hernandez N. 2010. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. Genome Res 20:710–721.

5. Bogenhagen DF, Brown DD. 1981. Nucleotide sequences in Xenopus 5S DNA required for transcription termination. Cell 24:261–270.

6. Braglia P, Percudani R, Dieci G. 2005. Sequence context effects on oligo(dT) termination signal recognition by Saccharomyces cerevisiae RNA polymerase III. J Biol Chem 280:19551–19562.

7. Orioli A, Pascali C, Quartararo J, Diebel KW, Praz V, Romascano D, Percudani R, van Dyk LF, Hernandez N, Teichmann M, Dieci G. 2011. Widespread occurrence of non-canonical transcription termination by human RNA polymerase III. Nucleic Acids Res 39:5499–5512.

8. Turowski TW, Leśniewska E, Delan-Forino C, Sayou C, Boguta M, Tollervey D. 2016. Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. Genome Res 26:933–944.

9. Stefano JE. 1984. Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts. Cell 36:145–154.

10. Gottlieb E, Steitz JA. 1989. The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription in vitro. EMBO J 8:841–850.

11. Gottlieb E, Steitz JA. 1989. Function of the mammalian La protein: evidence for its action in transcription termination by RNA polymerase III. EMBO J 8:851–861.

12. Maraia RJ, Kenan DJ, Keene JD. 1994. Eukaryotic transcription termination factor La mediates transcript release and facilitates reinitiation by RNA polymerase III. Mol Cell Biol 14:2147–2158.

13. Maraia RJ. 1996. Transcription termination factor La is also an initiation factor for RNA polymerase III. Proc Natl Acad Sci U S A 93:3383–3387.

14. Maraia RJ, Intine RV. 2002. La protein and its associated small nuclear and nucleolar precursor RNAs. Gene Expr 10:41–57.

15. Calvo O, Cuesta R, Anderson J, Gutiérrez N, García-Barrio MT, Hinnebusch AG, Tamame M. 1999. GCD14p, a repressor of GCN4 translation, cooperates with Gcd10p and Lhp1p in the maturation of initiator methionyl-tRNA in Saccharomyces cerevisiae. Mol Cell Biol 19:4167–4181.

16. Anderson J, Phan L, Cuesta R, Carlson BA, Pak M, Asano K, Björk GR, Tamame M, Hinnebusch AG. 1998. The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. Genes Dev 12:3650–3662.

17. Harada F, Matsubara M, Kato N. 1984. Stable tRNA precursors in HeLa cells. Nucleic Acids Res 12:9263–9269.

614 18. Nishikura K, De Robertis EM. 1981. RNA processing in microinjected Xenopus oocytes. Sequential
615 addition of base modifications in the spliced transfer RNA. J Mol Biol 145:405–420.

616 19. Rinke J, Steitz JA. 1985. Association of the lupus antigen La with a subset of U6 snRNA molecules.
617 Nucleic Acids Res 13:2617–2629.

618 20. Garalde DR, Snell EA, Jachimowicz D, Sipos B, Lloyd JH, Bruce M, Pantic N, Admassu T, James P,
619 Warland A, Jordan M, Ciccone J, Serra S, Keenan J, Martin S, McNeill L, Wallace EJ, Jayasinghe L,
620 Wright C, Blasco J, Young S, Brocklebank D, Juul S, Clarke J, Heron AJ, Turner DJ. 2018. Highly parallel
621 direct RNA sequencing on an array of nanopores. Nat Methods 15:201–206.

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

624 22. Liu H, Begik O, Novoa EM. 2021. EpiNano: Detection of m6A RNA Modifications Using Oxford
625 Nanopore Direct RNA Sequencing. Methods Mol Biol 2298:31–52.

626 23. Leger A, Amaral PP, Pandolfini L, Capitanchik C, Capraro F, Miano V, Migliori V, Toolan-Kerr P, Sideri
627 T, Enright AJ, Tzelepis K, van Werven FJ, Luscombe NM, Barbieri I, Ule J, Fitzgerald T, Birney E,
628 Leonardi T, Kouzarides T. 2021. RNA modifications detection by comparative Nanopore direct RNA
629 sequencing. 1. Nat Commun 12:7198.

630 24. Huang S, Wylder AC, Pan T. 2024. Simultaneous nanopore profiling of mRNA m6A and
631 pseudouridine reveals translation coordination. Nat Biotechnol 42:1831–1835.

632 25. Tavakoli S, Nabizadeh M, Makhamreh A, Gamper H, McCormick CA, Rezapour NK, Hou Y-M,
633 Wanunu M, Rouhanifard SH. 2023. Semi-quantitative detection of pseudouridine modifications and

type I/II hypermodifications in human mRNAs using direct long-read sequencing. Nat Commun 14:334.

26. Thomas NK, Poodari VC, Jain M, Olsen HE, Akeson M, Abu-Shumays RL. 2021. Direct Nanopore Sequencing of Individual Full Length tRNA Strands. ACS Nano 15:16642–16653.

27. Lucas MC, Pryszcz LP, Medina R, Milenkovic I, Camacho N, Marchand V, Motorin Y, Ribas de Pouplana L, Novoa EM. 2024. Quantitative analysis of tRNA abundance and modifications by nanopore RNA sequencing. Nat Biotechnol 42:72–86.

28. Saville L, Wu L, Habtewold J, Cheng Y, Gollen B, Mitchell L, Stuart-Edwards M, Haight T, Mohajerani M, Zovoilis A. 2024. NERD-seq: a novel approach of Nanopore direct RNA sequencing that expands representation of non-coding RNAs. Genome Biol 25:233.

29. Lerner MR, Andrews NC, Miller G, Steitz JA. 1981. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A 78:805–809.

30. Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34:3094–3100.

31. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997. arXiv https://doi.org/10.48550/arXiv.1303.3997.

32. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

33. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595.

654 34. Frankish A, Diekhans M, Ferreira A-M, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J,
655 Armstrong J, Barnes I, Berry A, Bignell A, Carbonell Sala S, Chrast J, Cunningham F, Di Domenico T,
656 Donaldson S, Fiddes IT, García Girón C, Gonzalez JM, Grego T, Hardy M, Hourlier T, Hunt T, Izuogu
657 OG, Lagarde J, Martin FJ, Martínez L, Mohanan S, Muir P, Navarro FCP, Parker A, Pei B, Pozo F,
658 Ruffier M, Schmitt BM, Stapleton E, Suner M-M, Sycheva I, Uszczynska-Ratajczak B, Xu J, Yates A,
659 Zerbino D, Zhang Y, Aken B, Choudhary JS, Gerstein M, Guigó R, Hubbard TJP, Kellis M, Paten B,
660 Reymond A, Tress ML, Flicek P. 2019. GENCODE reference annotation for the human and mouse
661 genomes. Nucleic Acids Res 47:D766–D773.

662 35. van Zon A, Mossink MH, Schoester M, Scheffer GL, Scheper RJ, Sonneveld P, Wiemer EA. 2001.
663 Multiple human vault RNAs. Expression and association with the vault complex. J Biol Chem
664 276:37715–37721.

665 36. Madore SJ, Wieben ED, Pederson T. 1984. Eukaryotic small ribonucleoproteins. Anti-La human
666 autoantibodies react with U1 RNA-protein complexes. J Biol Chem 259:1929–1933.

667 37. Busch H, Reddy R, Rothblum L, Choi YC. 1982. SnRNAs, SnRNPs, and RNA processing. Annu Rev
668 Biochem 51:617–654.

669 38. Schramm L, Hernandez N. 2002. Recruitment of RNA polymerase III to its target promoters. Genes
670 Dev 16:2593–2620.

671 39. Dieci G, Fiorino G, Castelnuovo M, Teichmann M, Pagano A. 2007. The expanding RNA polymerase
672 III transcriptome. Trends Genet 23:614–622.

673 40. Zhao Y, Karijolich J, Glaunsinger B, Zhou Q. 2016. Pseudouridylation of 7SK snRNA promotes 7SK
674 snRNP formation to suppress HIV-1 transcription and escape from latency. EMBO Rep 17:1441–
675 1451.

676 41. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. 2014. Pseudouridine
677 profiling reveals regulated mRNA pseudouridylation in yeast and human cells. Nature 515:143–146.

678 42. Henry BA, Marchand V, Schlegel BT, Helm M, Motorin Y, Lee N. 2022. Pseudouridylation of Epstein-
679 Barr virus noncoding RNA EBER2 facilitates lytic replication. RNA 28:1542–1552.

680 43. Xu H, Kong L, Cheng J, Al Moussawi K, Chen X, Iqbal A, Wing PAC, Harris JM, Tsukuda S, Embarc-Buh
681 A, Wei G, Castello A, Kriaucionis S, McKeating JA, Lu X, Song C-X. 2024. Absolute quantitative and
682 base-resolution sequencing reveals comprehensive landscape of pseudouridine across the human
683 transcriptome. Nat Methods 21:2024–2033.

684 44. Chan F-F, Kwan KK-L, Seoung D-H, Chin DW-C, Ng lO-L, Wong CC-L, Wong C-M. 2024. N6-
685 Methyladenosine modification activates the serine synthesis pathway to mediate therapeutic
686 resistance in liver cancer. Mol Ther 32:4435–4447.

687 45. Dominissini D, Moshitch-Moshkovitz S, Schwartz S, Salmon-Divon M, Ungar L, Osenberg S, Cesarkas
688 K, Jacob-Hirsch J, Amariglio N, Kupiec M, Sorek R, Rechavi G. 2012. Topology of the human and
689 mouse m6A RNA methylomes revealed by m6A-seq. Nature 485:201–206.

690 46. Gong J, Shao D, Xu K, Lu Z, Lu ZJ, Yang YT, Zhang QC. 2018. RISE: a database of RNA interactome
691 from sequencing experiments. Nucleic Acids Res 46:D194–D201.

692 47. Kiss AM, Jády BE, Bertrand E, Kiss T. 2004. Human box H/ACA pseudouridylation guide RNA
693 machinery. Mol Cell Biol 24:5797–5807.

694 48. Lestrade L, Weber MJ. 2006. snoRNA-LBME-db, a comprehensive database of human H/ACA and
695 C/D box snoRNAs. Nucleic Acids Res 34:D158-162.

696 49. Yamauchi Y, Nobe Y, Izumikawa K, Higo D, Yamagishi Y, Takahashi N, Nakayama H, Isobe T, Taoka M.
697 2016. A mass spectrometry-based method for direct determination of pseudouridine in RNA.
698 Nucleic Acids Res 44:e59.

699 50. Zerby DB, Patton JR. 1997. Modification of human U4 RNA requires U6 RNA and multiple
700 pseudouridine synthases. Nucleic Acids Res 25:4808–4815.

701 51. Massenet S, Branlant C. 1999. A limited number of pseudouridine residues in the human atac
702 spliceosomal UsnRNAs as compared to human major spliceosomal UsnRNAs. RNA 5:1495–1503.

703 52. Yamaki Y, Nobe Y, Koike M, Yamauchi Y, Hirota K, Takahashi N, Nakayama H, Isobe T, Taoka M.
704 2020. Direct Determination of Pseudouridine in RNA by Mass Spectrometry Coupled with Stable
705 Isotope Labeling. Anal Chem 92:11349–11356.

706 53. Borchardt EK, Martinez NM, Gilbert WV. 2020. Regulation and Function of RNA Pseudouridylation in
707 Human Cells. Annu Rev Genet 54:309–336.

708 54. Srinivas KP, Depledge DP, Abebe JS, Rice SA, Mohr I, Wilson AC. 2021. Widespread remodeling of
709 the m6A RNA-modification landscape by a viral regulator of RNA processing and export. Proc Natl
710 Acad Sci U S A 118:e2104805118.

711 55. Jansens RJJ, Olarerin-George A, Verhamme R, Mirza A, Jaffrey S, Favoreel HW. 2023.
712 Alphaherpesvirus-mediated remodeling of the cellular transcriptome results in depletion of m6A-
713 containing transcripts. iScience 26:107310.

714 56. Dremel SE, Sivrich FL, Tucker JM, Glaunsinger BA, DeLuca NA. 2022. Manipulation of RNA
715 polymerase III by Herpes Simplex Virus-1. Nat Commun 13:623.

716 57. Suzuki T. 2021. The expanding world of tRNA modifications and their disease relevance. Nat Rev Mol
717 Cell Biol 22:375–392.

718 58. Listerman I, Bledau AS, Grishina I, Neugebauer KM. 2007. Extragenic accumulation of RNA
719 polymerase II enhances transcription by RNA polymerase III. PLoS Genet 3:e212.

720 59. Pryszcz LP, Diensthuber G, Llovera L, Medina R, Delgado-Tejedor A, Cozzuto L, Ponomarenko J,
721 Novoa EM. 2024. SeqTagger, a rapid and accurate tool to demultiplex direct RNA nanopore
722 sequencing datasets. bioRxiv 2024.10.29.620808.

723 60. van der Toorn W, Bohn P, Liu-Wei W, Olguin-Nava M, Smyth RP, von Kleist M. 2024. Demultiplexing
724 and barcode-specific adaptive sampling for nanopore direct RNA sequencing. bioRxiv
725 2024.07.22.604276.

726 61. Licht K, Medenbach J, Lührmann R, Kambach C, Bindereif A. 2008. 3′-cyclic phosphorylation of U6
727 snRNA leads to recruitment of recycling factor p110 through LSm proteins. RNA 14:1532–1538.

728 62. Dyer SC, Austine-Orimoloye O, Azov AG, Barba M, Barnes I, Barrera-Enriquez VP, Becker A, Bennett
729 R, Beracochea M, Berry A, Bhai J, Bhurji SK, Boddu S, Branco Lins PR, Brooks L, Ramaraju SB,
730 Campbell LI, Martinez MC, Charkhchi M, Cortes LA, Davidson C, Denni S, Dodiya K, Donaldson S,
731 El Houdaigui B, El Naboulsi T, Falola O, Fatima R, Genez T, Martinez JG, Gurbich T, Hardy M, Hollis Z,
732 Hunt T, Kay M, Kaykala V, Lemos D, Lodha D, Mathlouthi N, Merino GA, Merritt R, Mirabueno LP,
733 Mushtaq A, Hossain SN, Pérez-Silva JG, Perry M, Piližota I, Poppleton D, Prosovetskaia I, Raj S, Salam
734 AIA, Saraf S, Saraiva-Agostinho N, Sinha S, Sipos B, Sitnik V, Steed E, Suner M-M, Surapaneni L,

735 Sutinen K, Tricomi FF, Tsang I, Urbina-Gómez D, Veidenberg A, Walsh TA, Willhoft NL, Allen J,
736 Alvarez-Jarreta J, Chakiachvili M, Cheema J, da Rocha JB, De Silva NH, Giorgetti S, Haggerty L, Ilsley
737 GR, Keatley J, Loveland JE, Moore B, Mudge JM, Naamati G, Tate J, Trevanion SJ, Winterbottom A,
738 Flint B, Frankish A, Hunt SE, Finn RD, Freeberg MA, Harrison PW, Martin FJ, Yates AD. 2025. Ensembl
739 2025. Nucleic Acids Research 53:D948–D957.

740 63. Mudge JM, Carbonell-Sala S, Diekhans M, Martinez JG, Hunt T, Jungreis I, Loveland JE, Arnan C,
741 Barnes I, Bennett R, Berry A, Bignell A, Cerdán-Vélez D, Cochran K, Cortés LT, Davidson C, Donaldson
742 S, Dursun C, Fatima R, Hardy M, Hebbar P, Hollis Z, James BT, Jiang Y, Johnson R, Kaur G, Kay M,
743 Mangan RJ, Maquedano M, Gómez LM, Mathlouthi N, Merritt R, Ni P, Palumbo E, Perteghella T,
744 Pozo F, Raj S, Sisu C, Steed E, Sumathipala D, Suner M-M, Uszczynska-Ratajczak B, Wass E, Yang YT,
745 Zhang D, Finn RD, Gerstein M, Guigó R, Hubbard TJP, Kellis M, Kundaje A, Paten B, Tress ML, Birney
746 E, Martin FJ, Frankish A. 2025. GENCODE 2025: reference gene annotation for human and mouse.
747 Nucleic Acids Res 53:D966–D975.

748 64. Colgrove RC, Liu X, Griffiths A, Raja P, Deluca NA, Newman RM, Coen DM, Knipe DM. 2016. History
749 and genomic sequence analysis of the herpes simplex virus 1 KOS and KOS1.1 sub-strains. Virology
750 487:215–221.

751 65. Arrand JR, Rymo L, Walsh JE, Björck E, Lindahl T, Griffin BE. 1981. Molecular cloning of the complete
752 Epstein-Barr virus genome as a set of overlapping restriction endonuclease fragments. Nucleic Acids
753 Res 9:2999–3014.

754 66. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000
755 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and
756 SAMtools. Bioinformatics 25:2078–2079.

67. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

68. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15:550.

69. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. PLoS Comput Biol 9:e1003118.

70. Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York. https://ggplot2.tidyverse.org.

71. Hahne F, Ivanek R. 2016. Visualizing Genomic Data Using Gviz and Bioconductor, p. 335–351. *In* Mathé, E, Davis, S (eds.), Statistical Genomics: Methods and Protocols. Springer, New York, NY.

Figure Legends

Figure 1: DRAP3R efficiently captures poly(U) RNAs.

(a) DRAP3R utilizes a subtractive approach to remove poly(A) and ribosomal RNAs in a stepwise fashion, prior to ligating a double-stranded DNA adapter to all RNAs with a 3' poly(U) tract. Following cDNA synthesis to generate hybrid RNA:cDNA molecules, all remaining single-stranded RNAs are removed by RNAseA digestion. Resulting libraries are loaded onto RNA (RA) flowcells and sequenced on a MinION mk1B for 24 h. **(b)** Total read count and read length distribution for four DRAP3R runs (2x ARPE-19 biological replicates, NHDF, and CRO-AP5) and a standard poly(A) DRS run (ARPE-19). **(c)** Sequence reads were classified as Pol II transcribed (green) or Pol III transcribed with the latter subdivided into pre-tRNAs (dark blue), primary Pol III genes (orange), and pseudogenes (red). Short reads that could not be unambiguously assigned to a specific gene are shown in beige while reads that did not overlap with any existing annotation are shown in gold. **(d)** Read coverage plots across the VTRNA1 locus on chromosome 5 highlight the specificity and full transcript coverage provided by DRAP3R. **(e)** Scatter plot comparing the abundances of RNAs within distinct classes of RNA

Pol II and RNA Pol III transcripts shows very high correlation between biological replicates of ARPE-19 cells. Colour coding as in (**c**)

Figure 2: Detection of Ψ and m^{6}A and in nascent Pol III transcribed RNAs

(a) Modification probability distributions for U, A, Ψ, and m^{6}A in all datasets generated in this study were calculated using modkit and used to inform a filtering strategy that retained only high-confidence modification calls. **(b)** Coverage plots of nascent RN7SK RNA in the DRAP3R ARPE-19 #1 and RN7SK IVT datasets. Y-axis denotes read depth while X-axis denotes co-ordinates on HG38 chromosome 6. The RN7SK gene annotation is shown as a grey box. **(c)** Ψ stochiometry distributions on nascent RN7SK demonstrates a high-false positive rate in unfiltered datasets (top five rows) but not in filtered datasets (bottom five rows) which show pseudouridine installation at positions 243, 247, and 250 (blue asterisks). **(d)** Stoichiometry distribution histograms for Ψ (blue) and m^{6}A (red) in the filtered DRAP3R ARPE-19 #1 dataset. Pol III transcribed pre-tRNAs are shown separately from other Pol III transcribed RNAs. **(e)** Modification density plots (metaplots) for Ψ and m^{6}A show consistent and specific distributions for pre-tRNAs and other Pol III transcribed RNAs.

Figure 3: Ψ profiles across the Pol III transcriptome

(a) Heatmap showing stoichiometry distributions across Ψ sites in nascent Pol III transcribed ncRNAs vary between cell types. **(b)** Ψ profiles across all nascent pre-tRNAs with coverage ≥ 50 reads in the DRAP3R ARPE-19 #1 dataset. Each row depicts different tRNA genes and rows are grouped by isoacceptor. **(c)** As (b) but showing only detected glutamic acid tRNAs genes. **(d)** Ψ stoichimetry distributions (> 5%) at defined Ψ positions across all pre-tRNAs. **(e)** Consensus sequences motifs for Ψ installation by Pus7 at Ψ 13 and Pus4 at Ψ 55. **(f)** As (**c**) except that the ARPE-19 #1 DRAP3R and nano-tRNA-Seq datasets were aligned against a tRNA isodecoder splint database. **(g)** Distribution of basecall errors for the ARPE-19 #1 DRAP3R and nano-tRNA-Seq datasets across the Glu-TTC-1 isodecoder. Select modified nucleotides are labelled. **(h)** Basecall error rate distributions for the ARPE-19 #1 DRAP3R and nano-tRNA-Seq datasets across each position of all glutamic acid isodecoders.

Figure 4: HSV-1 infection regulates pre-tRNA expression and Ψ installation

(a) Total read count and read length distribution per DRAP3R run using RA flowcells with RNA004 chemistry and a minimum read length cutoff of 20 nt. **(b)** As in Fig. 1c, sequence reads were classified as Pol II transcribed (green) or Pol III transcribed with the latter subdivided into pre-tRNAs (dark blue), primary Pol III genes (orange), and pseudogenes (red). Ambiguous reads are shown in beige while reads with no overlap are shown in gold. **(c)** Scatter plot comparing the abundances of RNAs within distinct classes of RNA Pol II and RNA Pol III transcripts shows decreasing correlation between uninfected and HSV-1 infected (left – 6hpi, right – 12hpi) ARPE-19 cells, colour coding as in (**c**). **(d)** Scatter plot comparing Ψ stochiometries on pre-tRNAs shows decreasing correlation between uninfected and HSV-1 infected (left – 6hpi, right – 12hpi) ARPE-19 cells.

Supplementary Figure S1: DRAP3R efficiently captures poly(U) RNAs.

(a) The standard double-stranded DNA nanopore poly(T) adaptor (left) is shown in contrast to our custom poly(U) targeting DRAP3R adaptor (right). **(b)** Total read count and read length distribution per DRAP3R run for IVT-derived RN7SK and EBER2 RNAs using RA flowcells with RNA004 chemistry and a minimum read length cutoff of 20 nt. **(c)** Sequence reads derived from the ARPE-19 #1 DRAP3R sample were aligned to the HG38 genome using BWA or minimap2 in combination with different flags. Total numbers of primary, secondary/supplementary, and unaligned reads are shown on the left with the read length distribution of aligned reads shown on the right. **(d)** For the optimal alignment strategy (bwa mem –W 13 –k 6 –T 20 –x ont2d), the aligned length of reads was generally 60-70nt shorter than the full read length, indicative of full alignment along the full RNA body and excluding the untrimmed adapter sequence. **(e)** Barplots denoting the proportion (blue/green) of reads that could be trimmed using dorado v0.7.0. Untrimmed reads (grey) were classified as those in which the difference between trimmed and untrimmed read was < 10 nt. **(f)** Trim length distributions for each dataset were derived by calculating the difference in length for each trimmed read relative to its untrimmed state. A normal distribution with a modal value around 65 nt (adapter length) is observed for the ARPE-19 #1 poly(A) DRS dataset while DRAP3R datasets show multi-modal distributions resulting from over-trimming of reads. **(g)** Integrative Genomics Viewer (IGV) screenshots showing extensive primary alignments across three low-confidence tRNA genes; (top) tRNA-Glu-TTC-5-1, (middle) tRNA-Gly-CCC-6-1, and (bottom) tRNA-Gly-CCC-4-1. Soft-clipped (adapter) sequences are shown to the right (top) or left (middle, bottom) of the alignment. (**h**) Breakdown of Pol II transcribed poly(U) RNAs detected by DRAP3R for sample ARPE-19 #1.

Supplementary Figure 2: DRAP3R expands the RNA polymerase III transcriptome

(a) Coverage plots for four recently confirmed and two recently predicted tRNA genes are shown in dark blue with 5' (red) and 3' (grey) alignment ends (indicating putative sites of transcription initiation and 3' end processing) shown as overlaid histograms. Characteristic sequence features including A (blue) and B (green) boxes as well as TATA boxes (pink) and poly(T) tracts ≥ 4 nt (purple) are shown as part of the underlying transcript schematic. Genome co-ordinates are specified for the HG38 assembly while the y-axis denotes read depth. **(b)** Coverage plots for eight putative novel RNA Pol III transcribed RNAs are shown in gold with 5' (red) and 3' (black) alignment ends shown as overlaid histograms. Characteristic sequence features are shown as for (a). **(c)** ARPE-19 cultures were treated with 50, 150, or 250 µg/ml α-amanitin to selectively inhibit Pol II transcription and quantitative PCR performed on extracted RNA to determine expression levels relative to 18s rRNA (expressed as log2 fold difference). Selective inhibition of Pol II transcription was confirmed for all Pol II transcribed targets (POL3RA, GAPDH, C-MYC, and RNU2-1).

Supplementary Figure S3: Modification probability distributions for all datasets

Individual modification probability distributions for U and Ψ (left two columns) as well as A and m^6A (right two columns) in all datasets generated in this study.

Supplementary Figure S4: Detection of m^6A and Ψ in nascent Pol III transcribed RNAs

(a) m^6A stoichiometry plot for nascent RN7SK demonstrates a high-false positive rate in unfiltered datasets (top five rows, unmodified sites shown in pale blue) whereas filtered datasets (bottom five rows) show absence of co-transcriptional m^6A installation. Increasing stoichiometry is shown by increased blue (Ψ) or red (m^6A) shading. **(b)** Coverage plots of EBER2 in the DRAP3R CRO-AP5 and EBER2 IVT datasets. Y-axis denotes read depth while X-axis denotes genome co-ordinates. The EBER2 gene annotation is shown as a black box. **(c)** Ψ and **(d)** m^6A stoichiometry plots for nascent EBER2 also demonstrate a high-false positive rate in unfiltered datasets (upper two rows, unmodified sites shown in pale blue) whereas filtered datasets (bottom two rows) show absence of co-transcriptional Ψ and m^6A installation. **(e)** Stoichiometry distribution histograms for Ψ (blue) and m^6A (red) for all filtered DRAP3R datasets. Pol III

transcribed pre-tRNAs are shown separately from all other Pol III transcribed RNAs. **(f)** Modification density plots for m^6A (top) and Ψ (bottom) recapitulate known distributions of these modifications on polyadenylated RNAs. **(g)** Modification density plots (metaplots) for Ψ sites with ≥ 10% stoichiometry on pre-tRNAs (top) and other Pol III transcribed RNAs (bottom).

Supplementary Figure S5: Ψ and m^6A profiles across the Pol III transcriptome

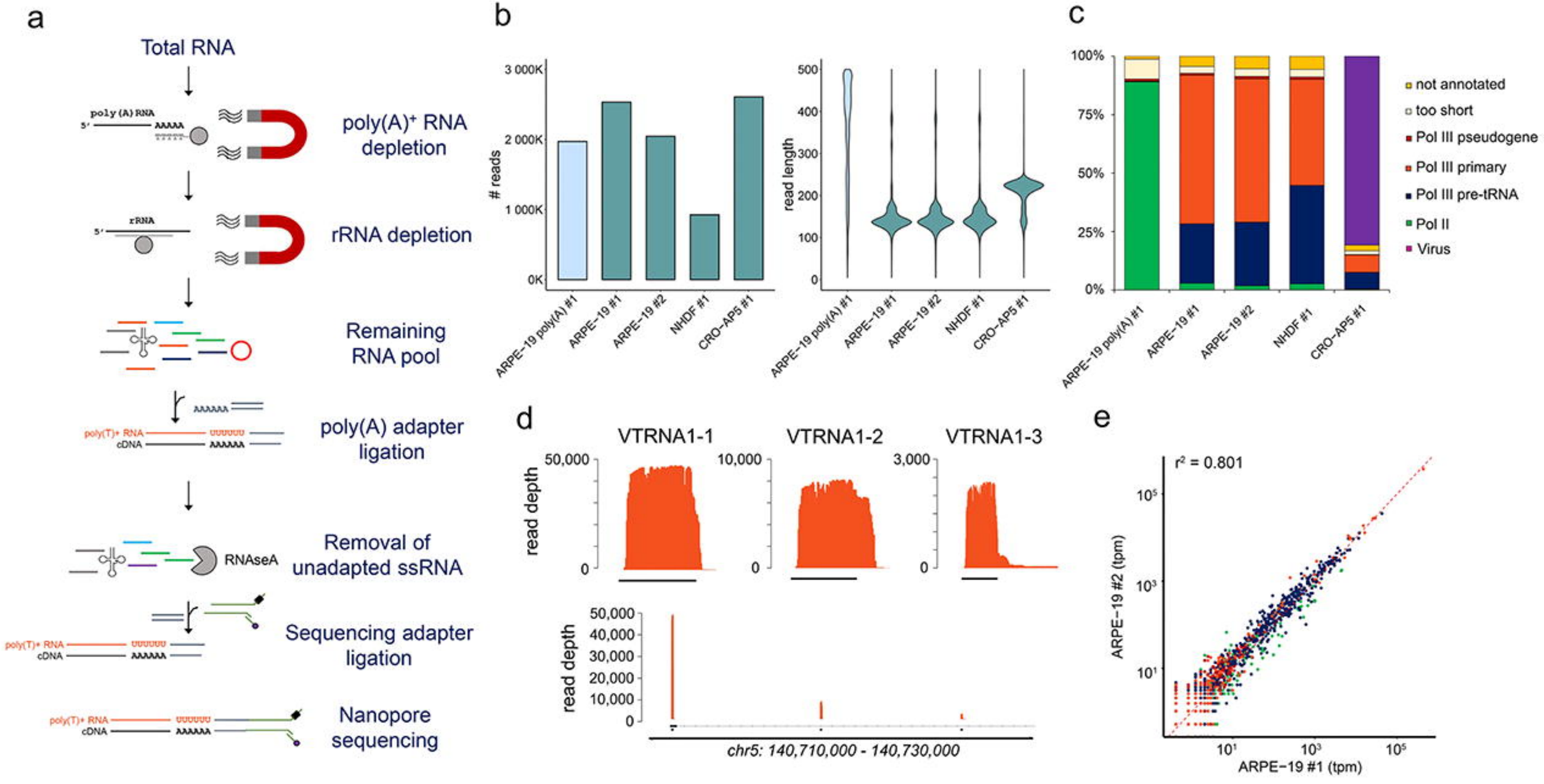
(a) Heatmap showing stoichiometry distributions across m^6A sites in nascent Pol III transcribed ncRNAs vary between cell types. **(b-c)** Motif analysis reveals no consensus sequences motifs for m^6A **(b)** and Ψ **(c)** installation on Pol III transcribed ncRNAs (excluding pre-tRNAs). **(d)** Distribution of basecall errors for the ARPE-19 #1 DRAP3R and nano-tRNA-Seq datasets across four glutamic acid isodecoders.

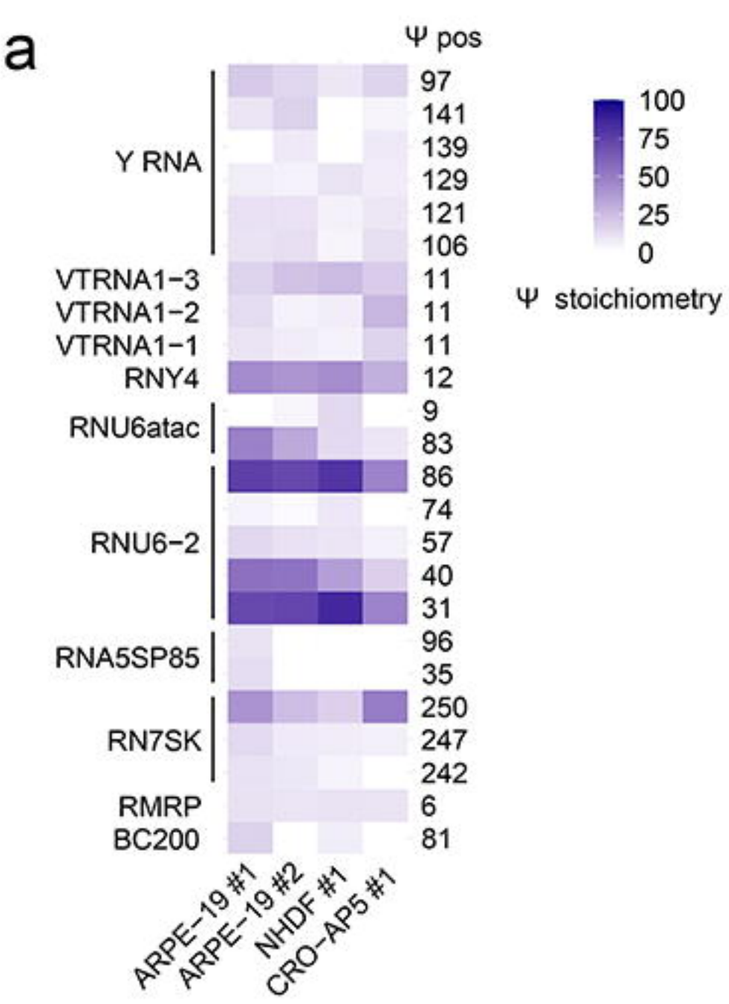
Supplementary Figure S6: Ψ profiles across all nascent pre-tRNAs

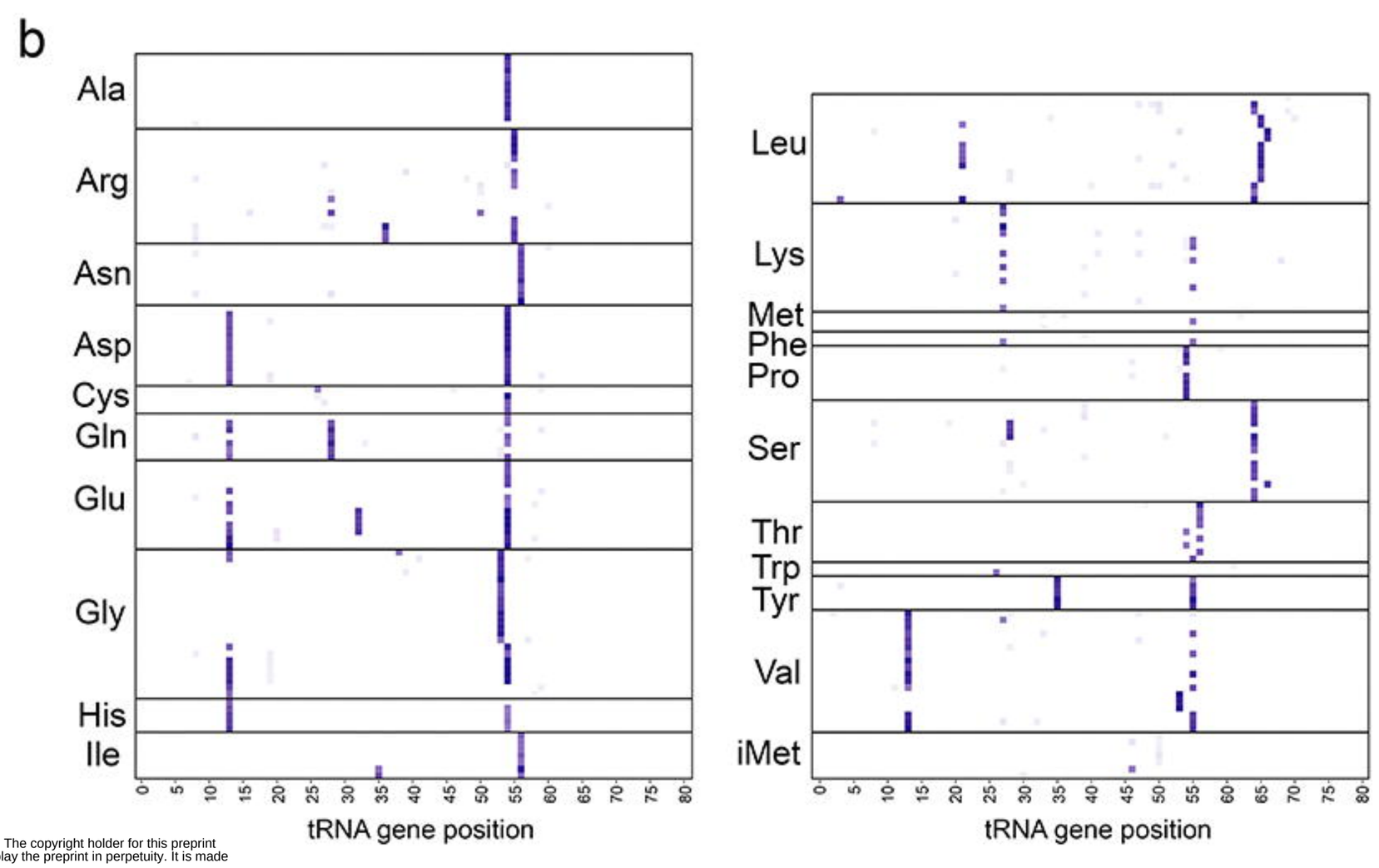
Ψ profiles across all nascent pre-tRNAs with coverage ≥ 50 reads for DRAP3R datasets ARPE-19 #2, ARPE-19 + IFN #3, NHDF #1, and CRO-AP5 #1. Each row depicts different tRNA genes and rows are grouped by isoacceptor.

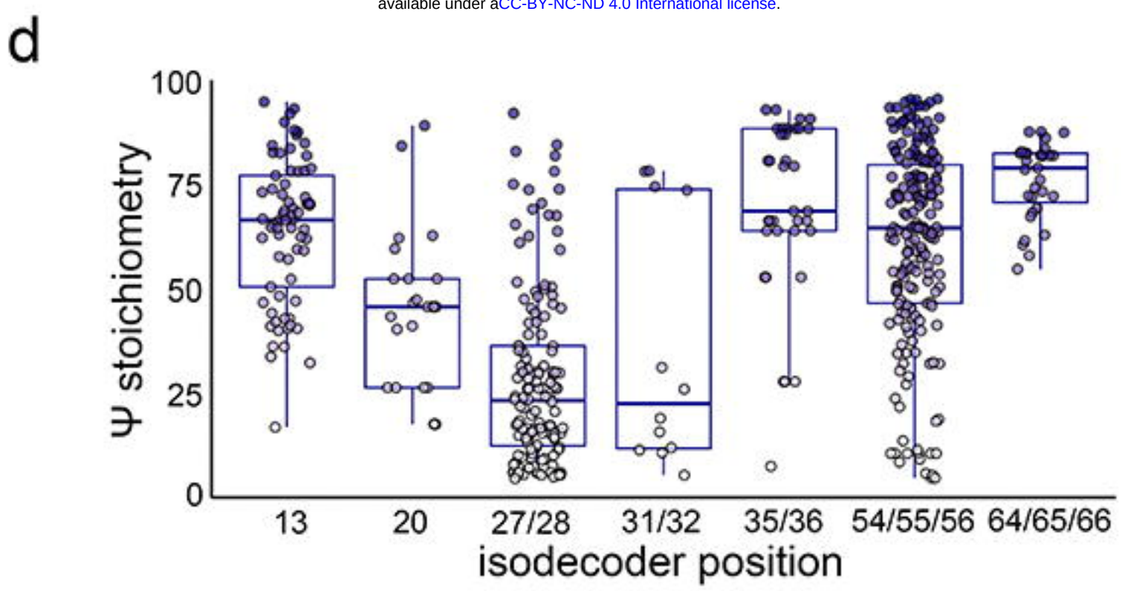
Supplementary Figure S7: Viral infection induced regulation the Pol III transcriptome and Ψ installation

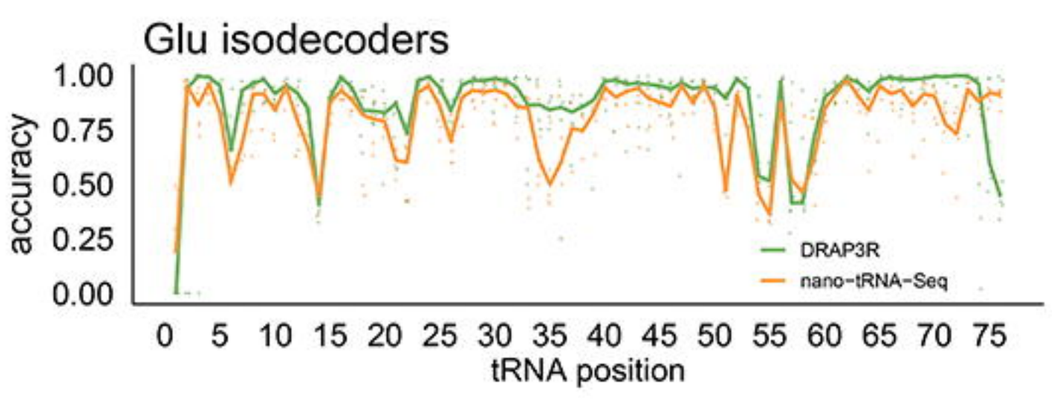
(a) Scatter plot comparing the abundances of RNAs within distinct classes of RNA Pol II and RNA Pol III transcripts shows very high correlation between biological replicates of HSV-1 infected ARPE-19 cells harvested at 6hpi. **(b)** Volcano plot showing significantly differentially regulated Pol III transcribed RNAs (red). **(c)** Individual modification probability distributions for U, A, Ψ and m^6A (right columns) for each infected cell dataset. **(d)** Stoichiometry distribution histograms for Ψ in the filtered HSV-1 infected ARPE-19 datasets. Pol III transcribed pre-tRNAs are shown separately from other Pol III transcribed RNAs. **(e-g)** Scatter plot comparing Ψ stochiometries on pre-tRNAs shows very high correlation between biological replicates of **(e)** uninfected ARPE-19 cells, **(f)** HSV-1 infected ARPE-19 cells harvested at 6hpi, and **(g)** uninfected ARPE-19 cells versus uninfected ARPE-19 cells treated with a combination of interferons α and γ.

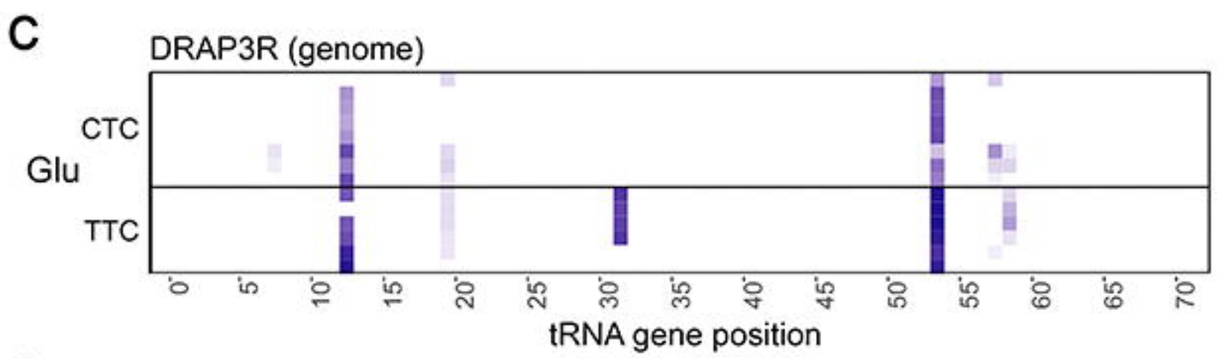

a
Total RNA
poly(A) RNA
poly(A)+ RNA depletion
rRNA
rRNA depletion
Remaining RNA pool
poly(A) adapter ligation
RNAseA
Removal of unadapted ssRNA
Sequencing adapter ligation
poly(T)+ RNA
cDNA
Nanopore sequencing
b
reads
3 000K
2 000K
1 000K
0K
read length
500
400
300
200
100
0
ARPE−19 poly(A) #1
ARPE−19 #1
ARPE−19 #2
NHDF #1
CRO−AP5 #1
c
100%
75%
50%
25%
0%
not annotated
too short
Pol III pseudogene
Pol III primary
Pol III pre-tRNA
Pol II
Virus
d
VTRNA1-1
VTRNA1-2
VTRNA1-3
read depth
50,000
10,000
3,000
0
40,000
30,000
20,000
chr5: 140,710,000 - 140,730,000
e
r² = 0.801
ARPE−19 #2 (tpm)
ARPE−19 #1 (tpm)
10¹
10³
10⁵

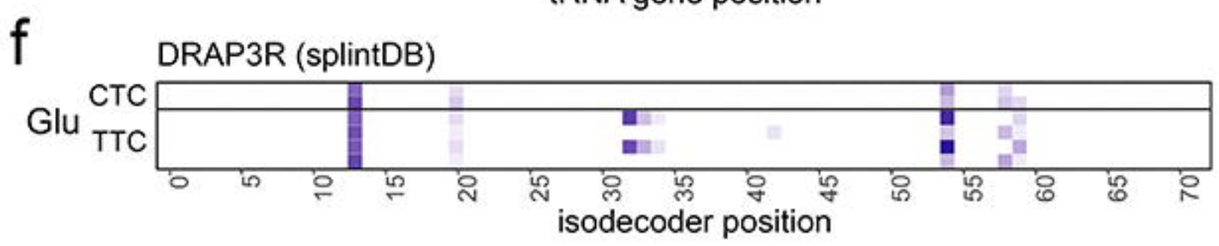


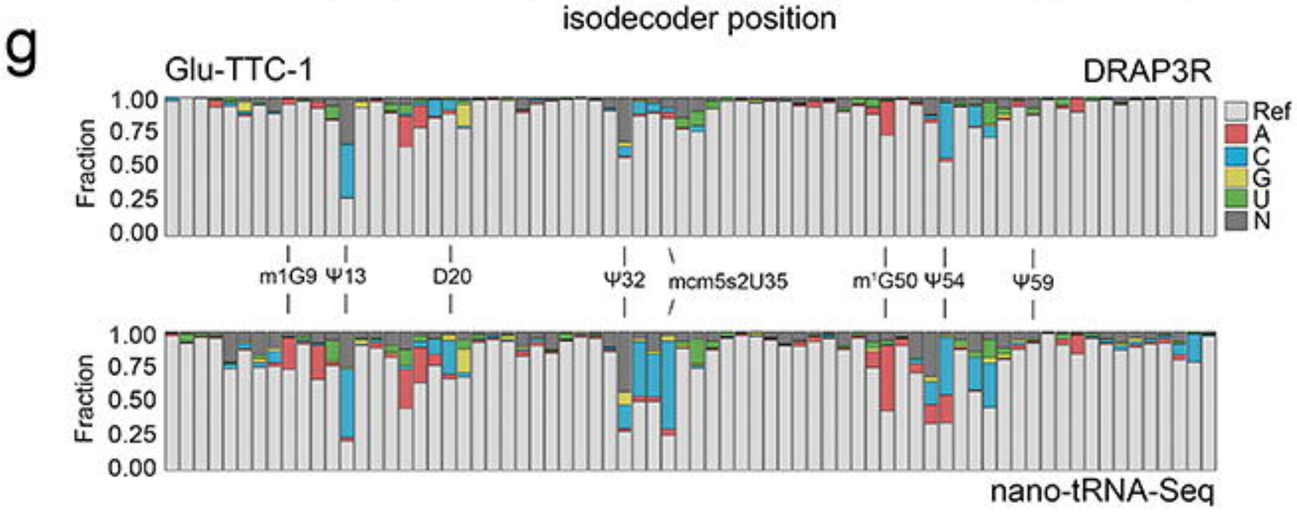




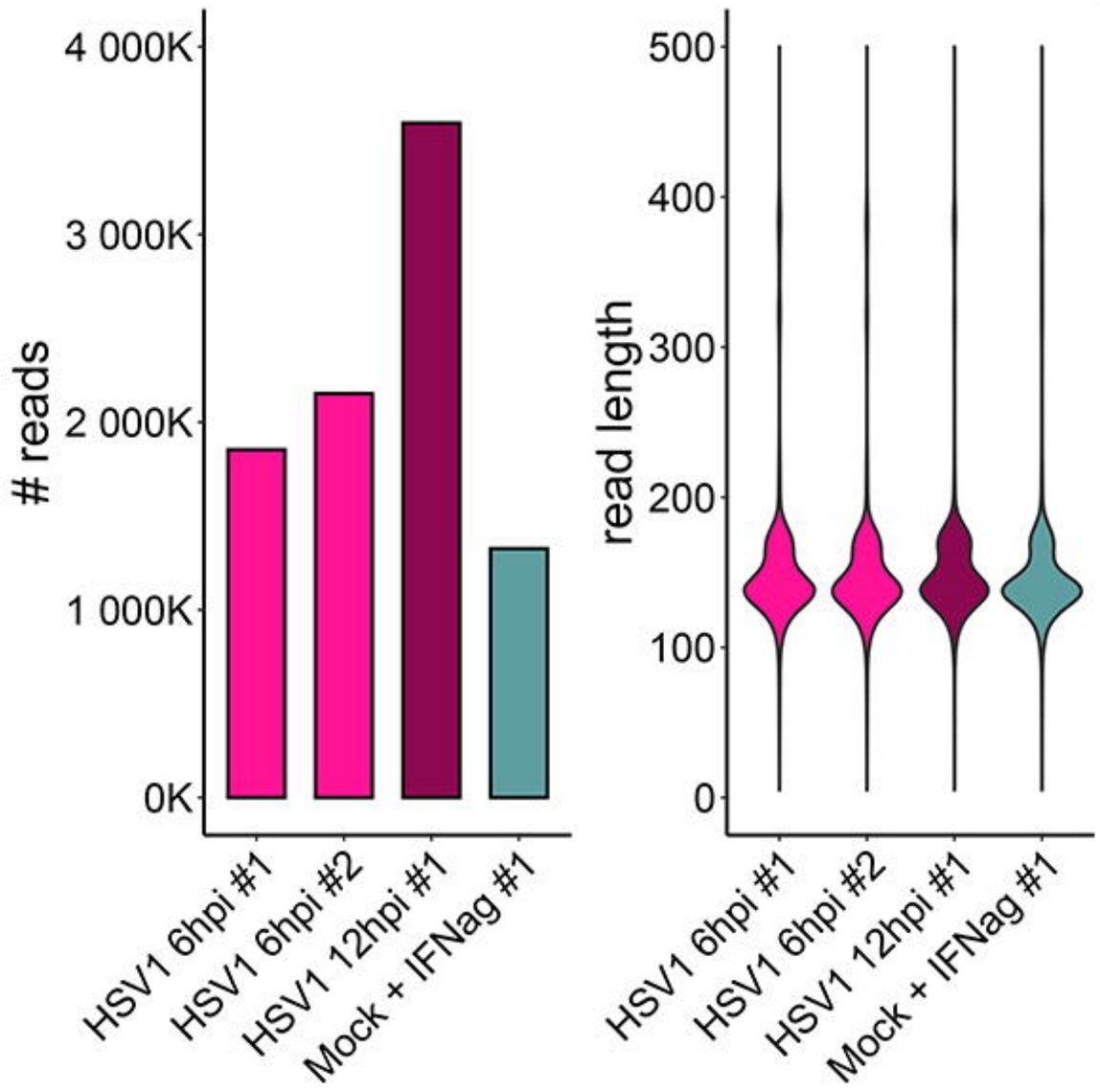




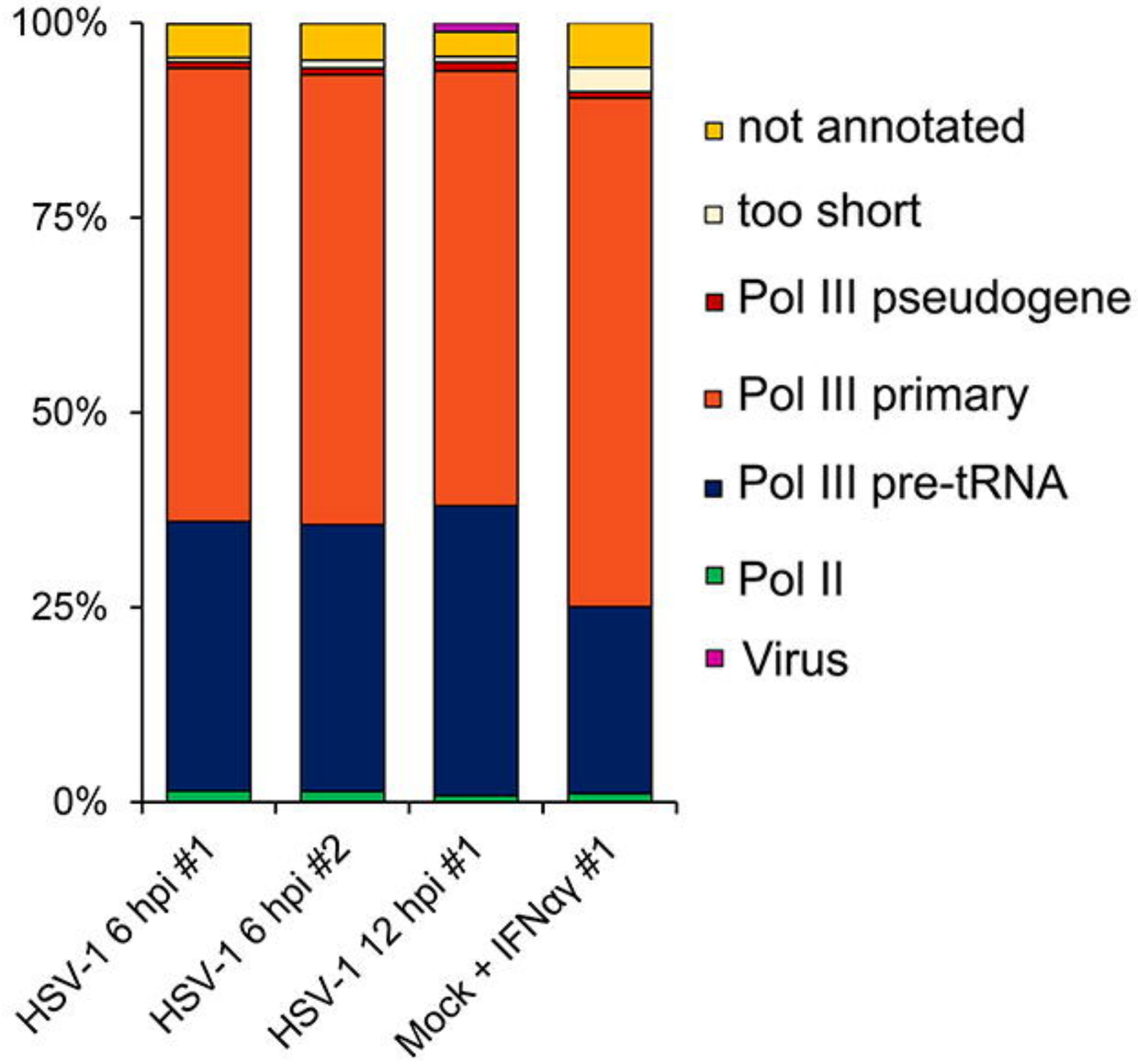




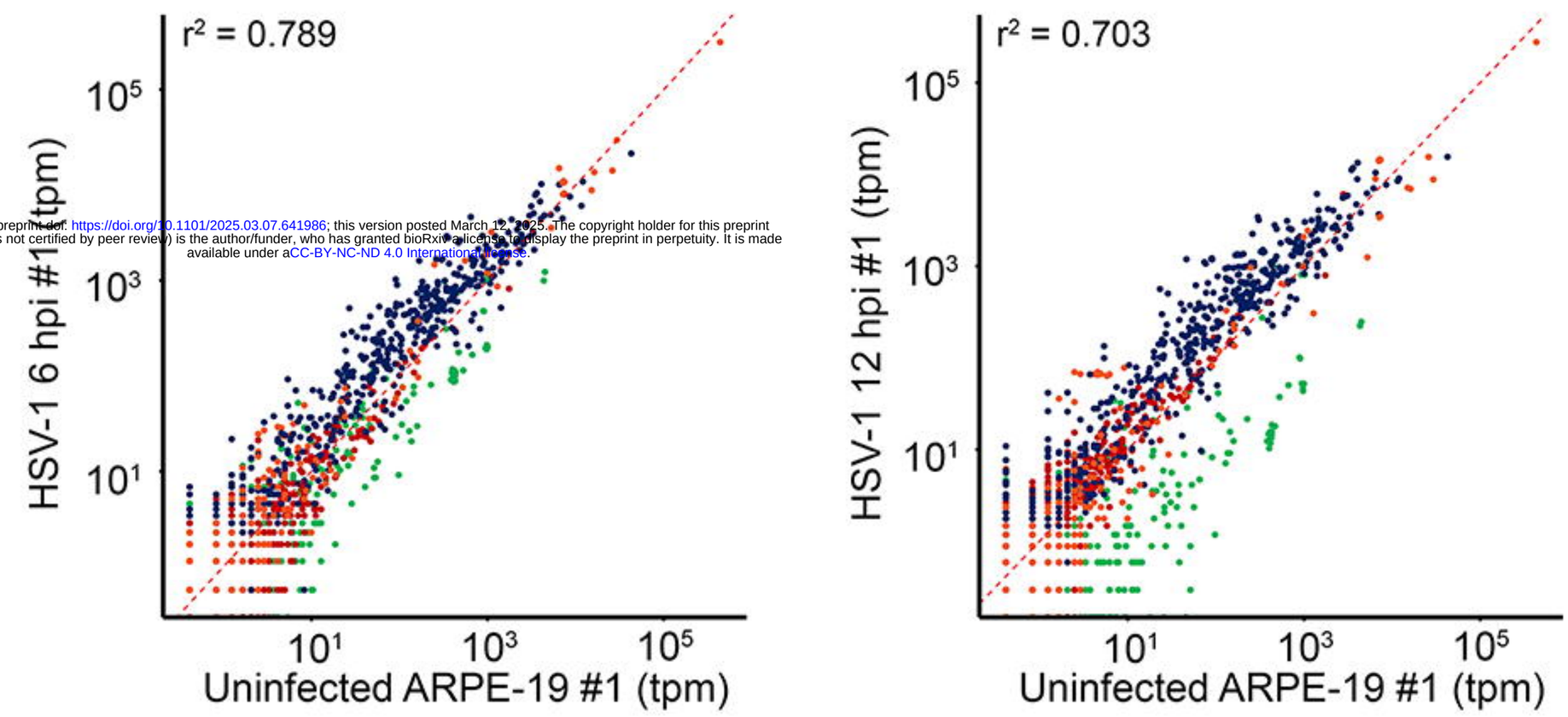
a



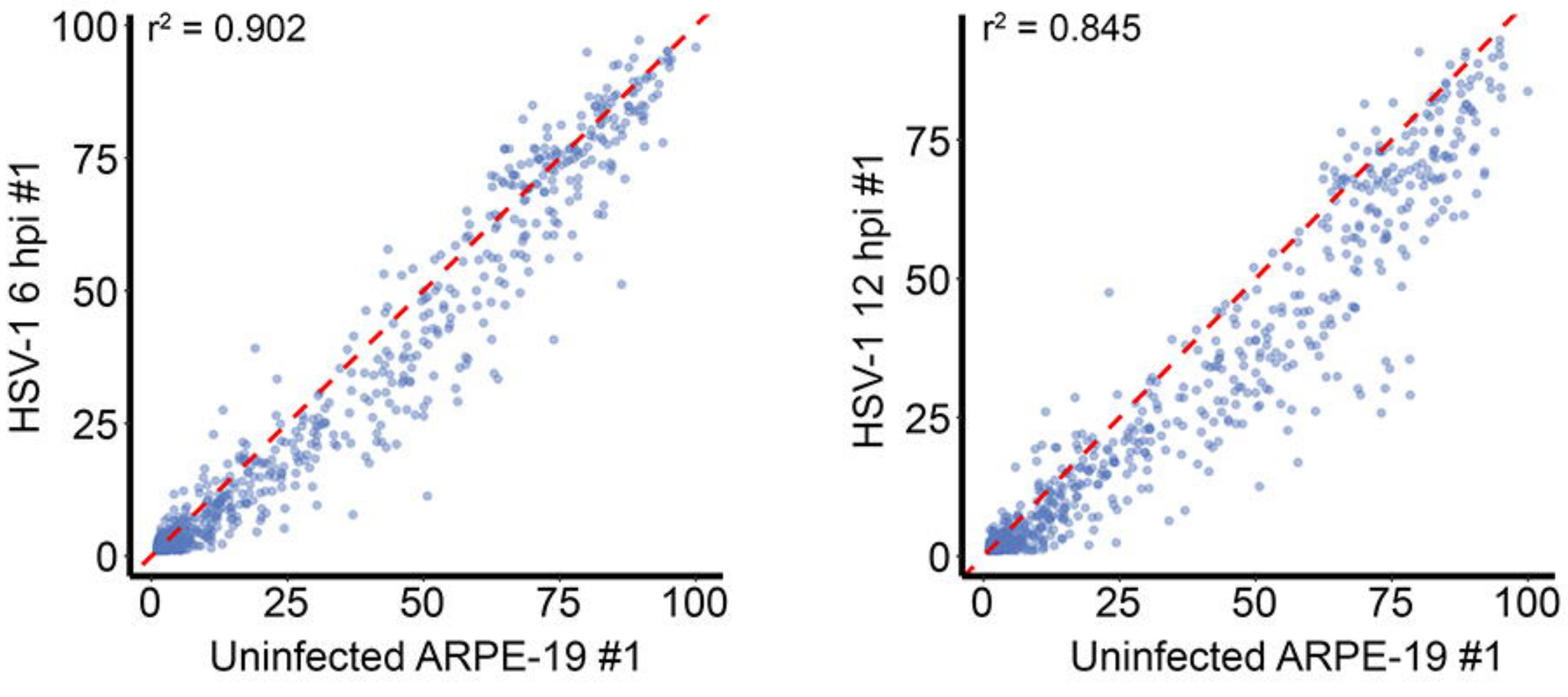
b



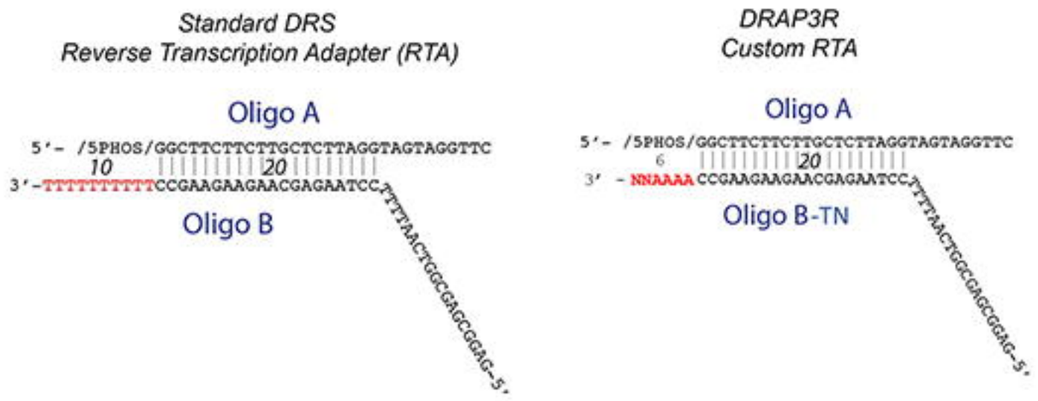
c



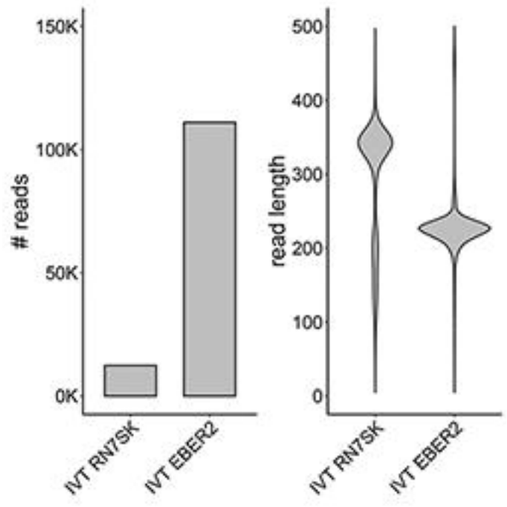
d



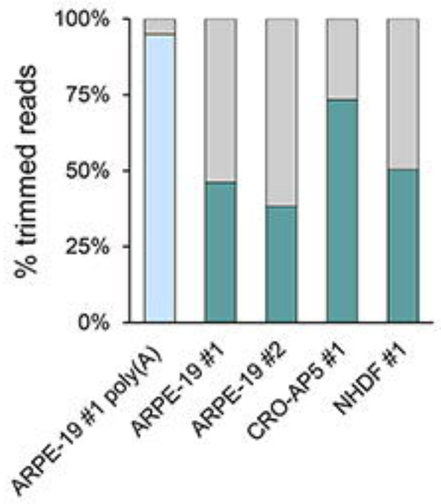
a



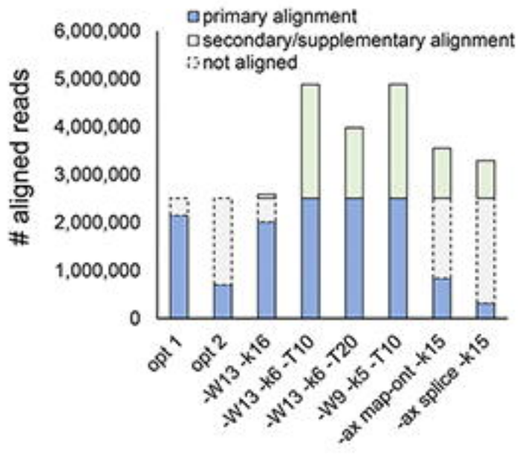
b

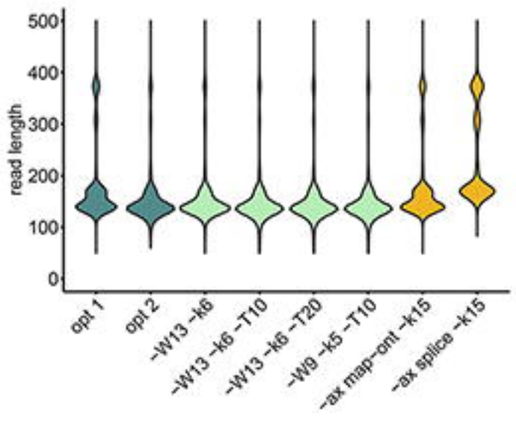


e

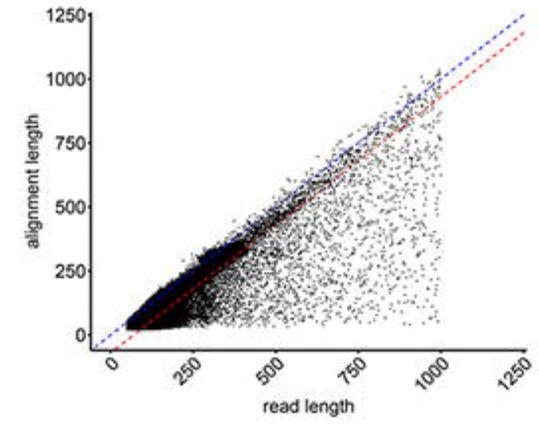


c

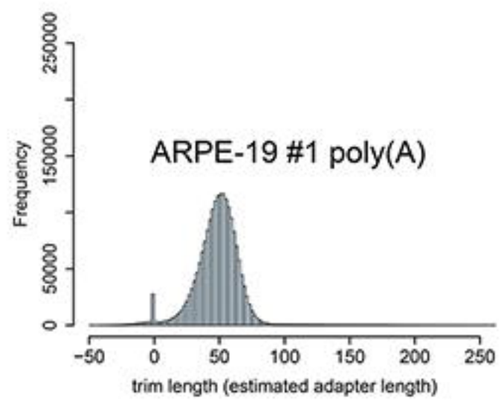


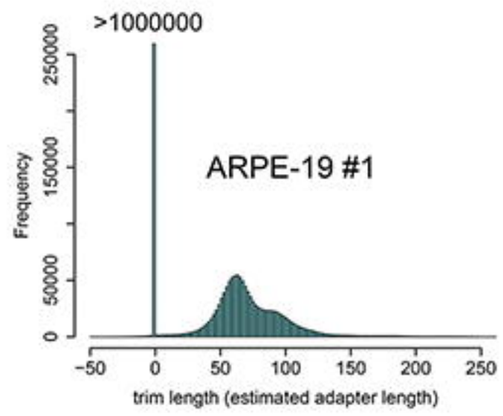


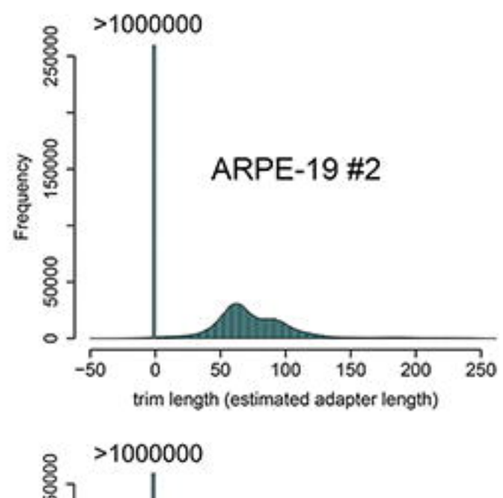
d

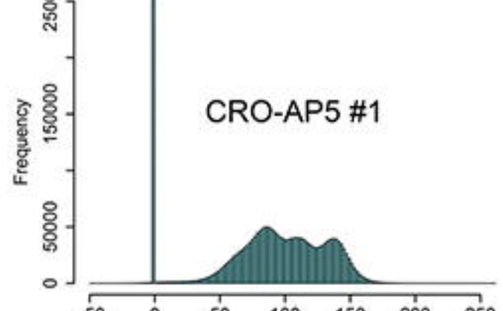


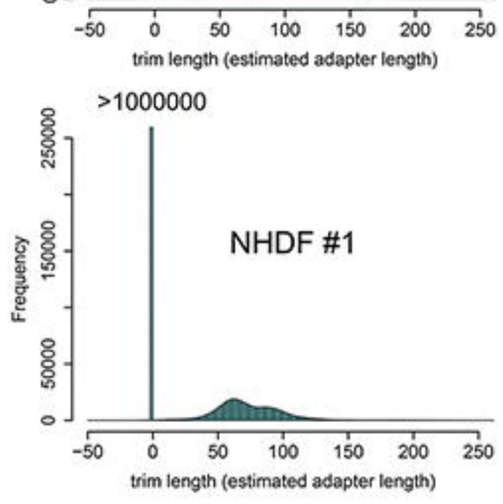
f



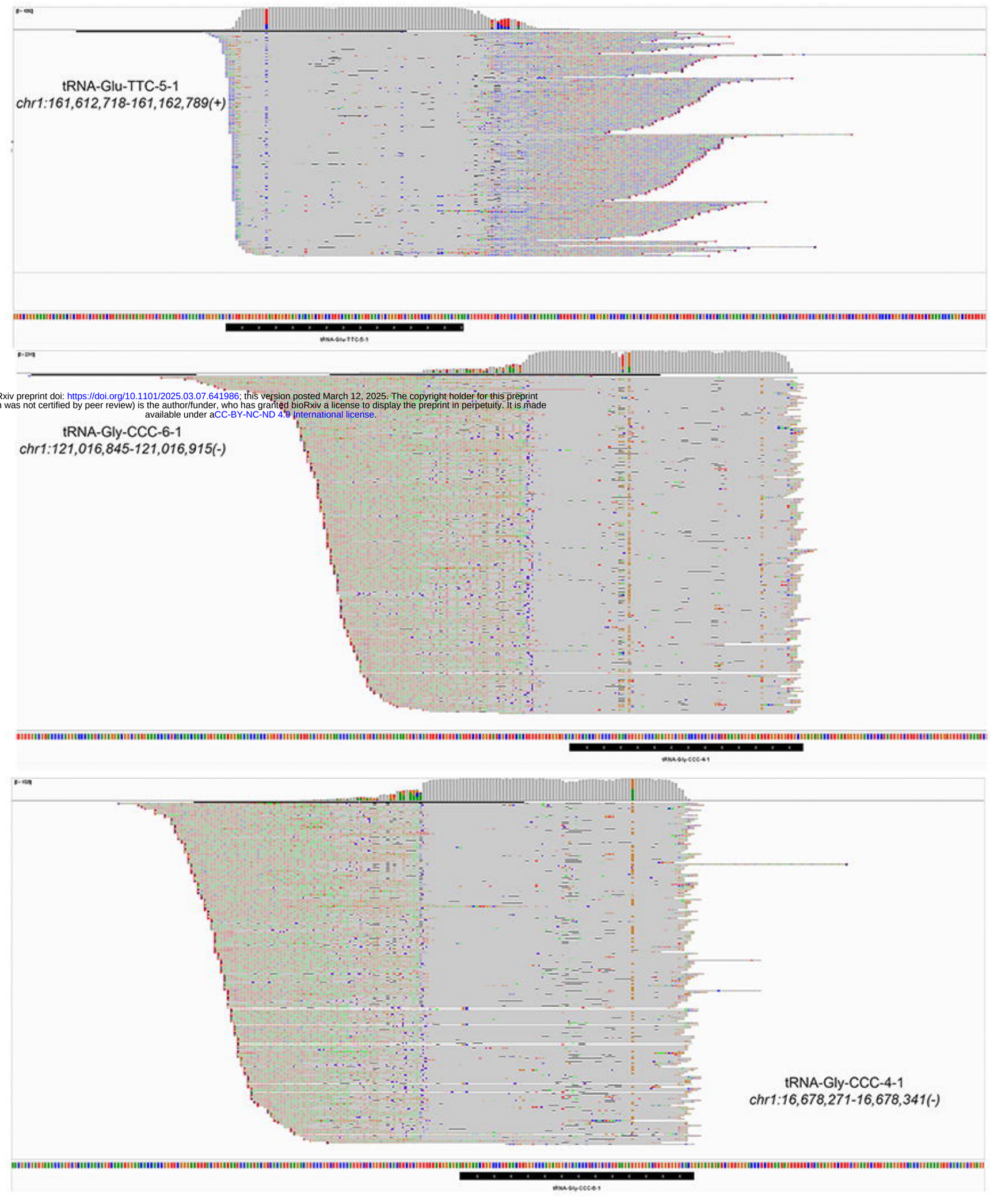




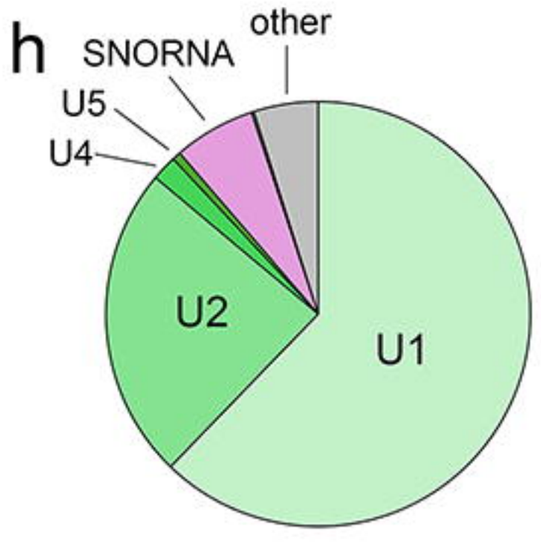


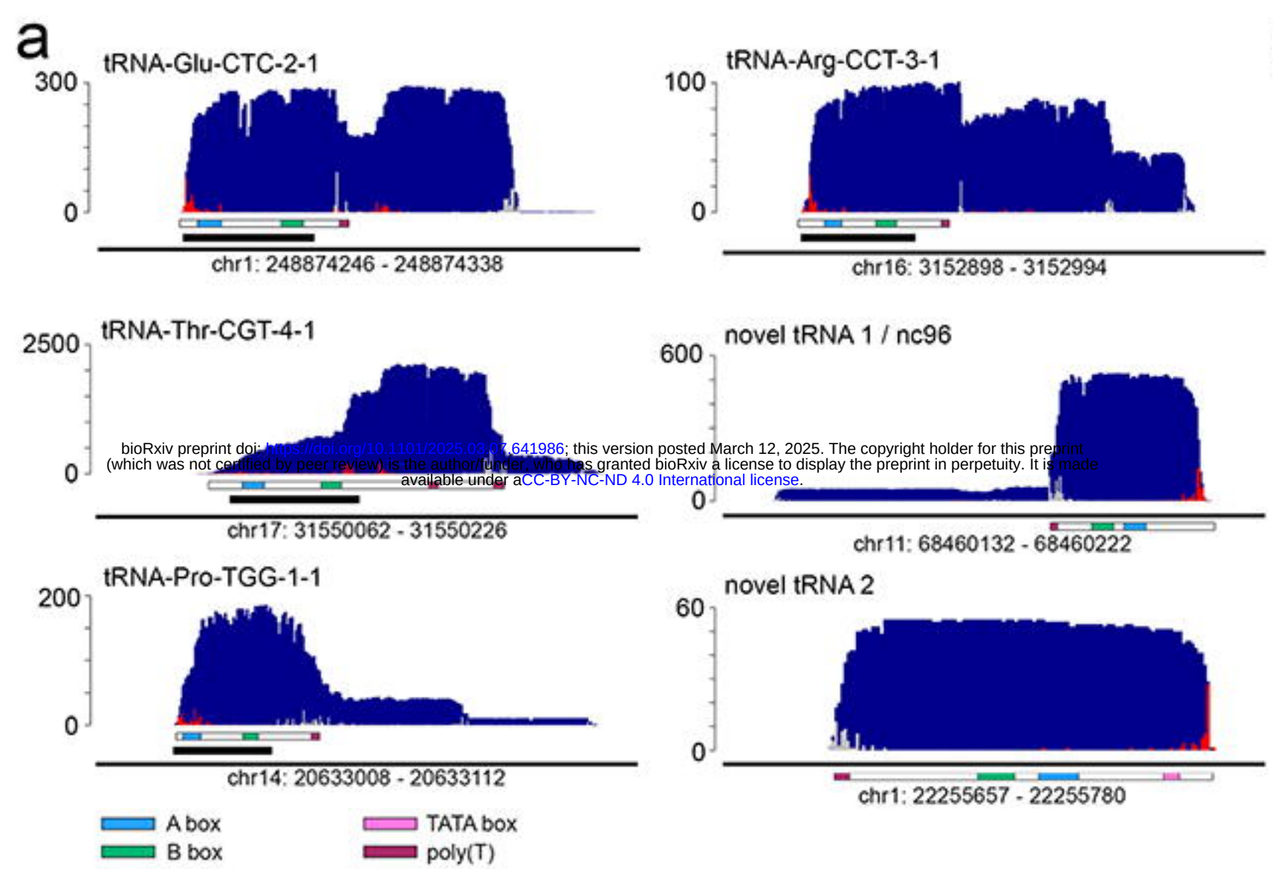


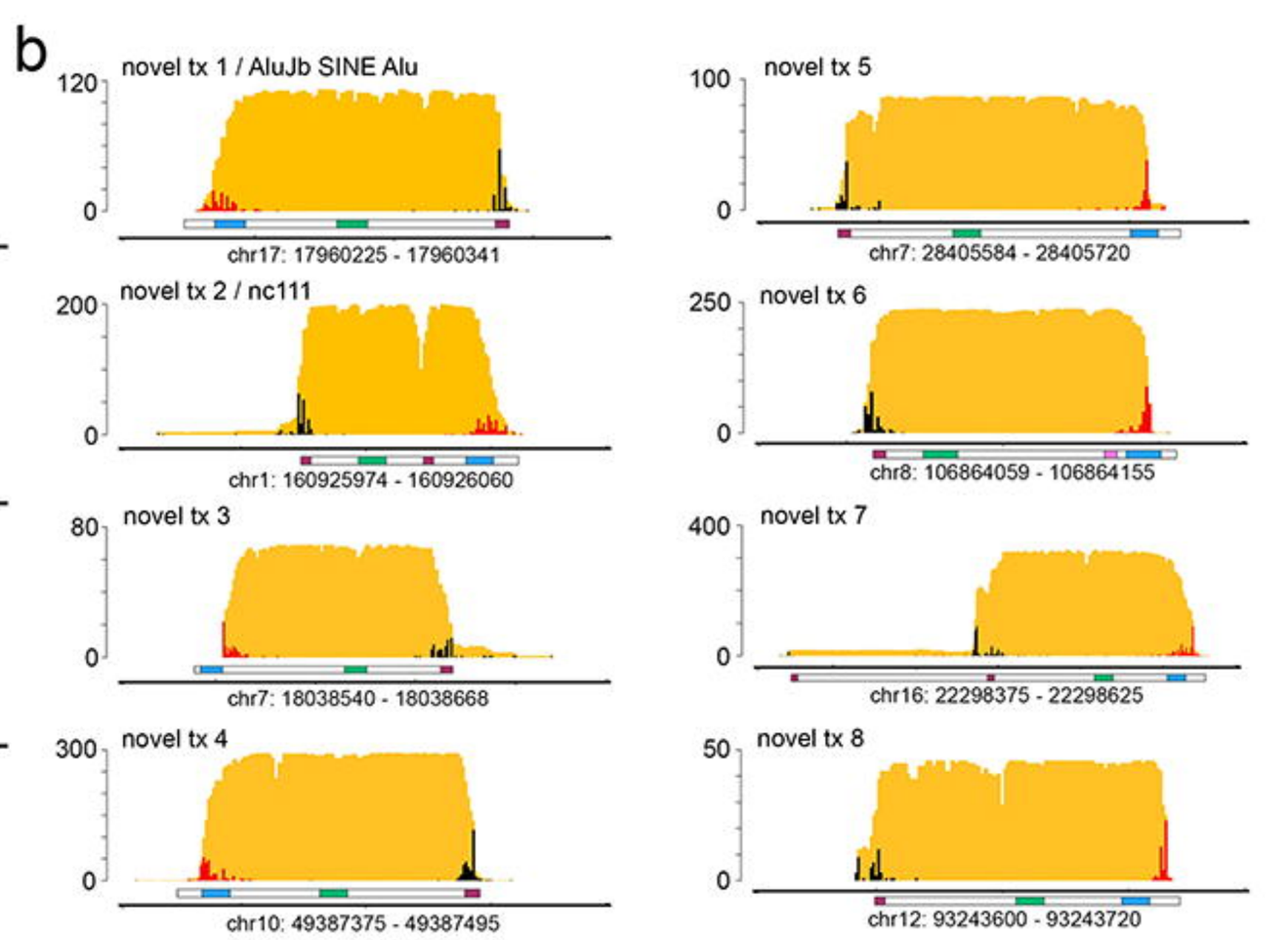
g

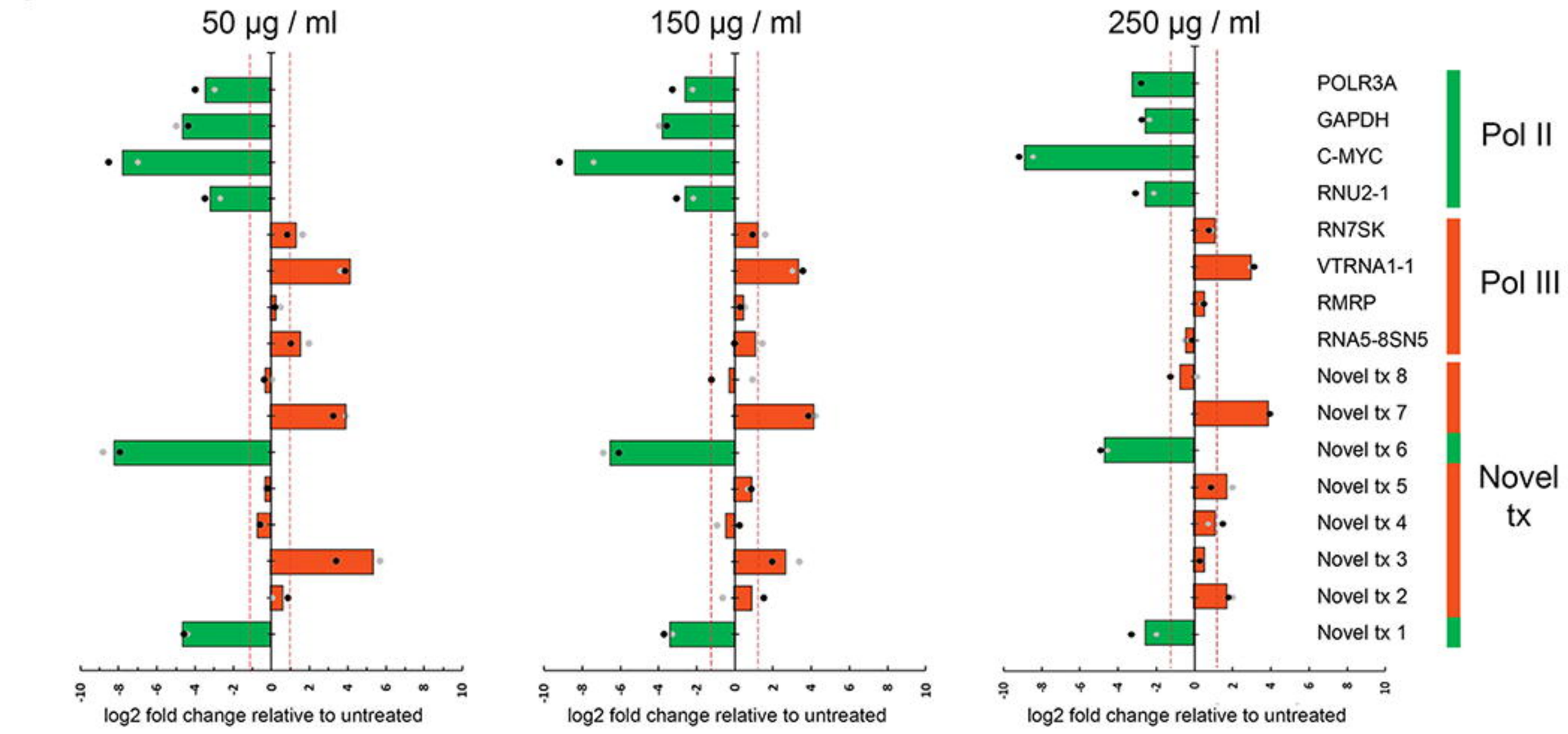


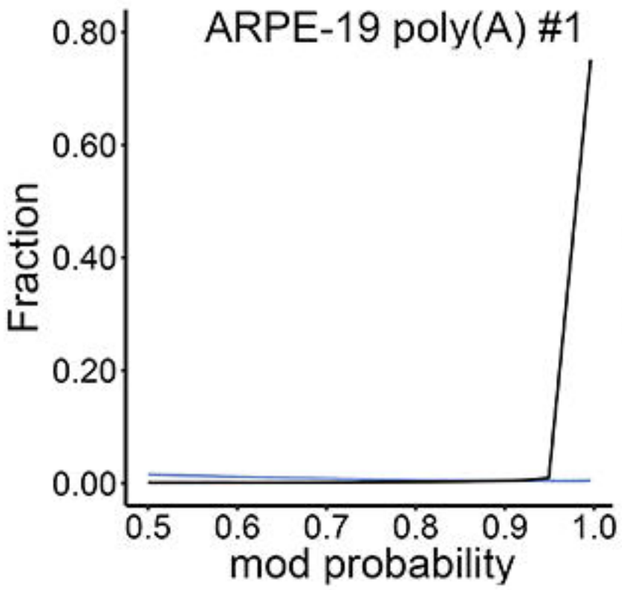
h

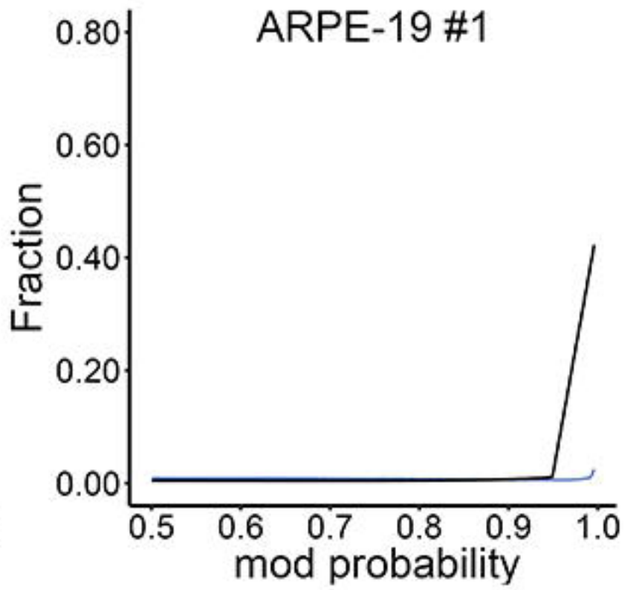


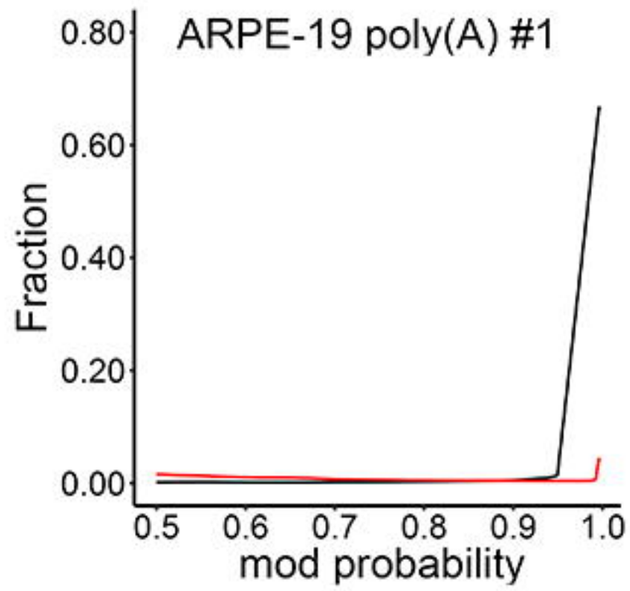


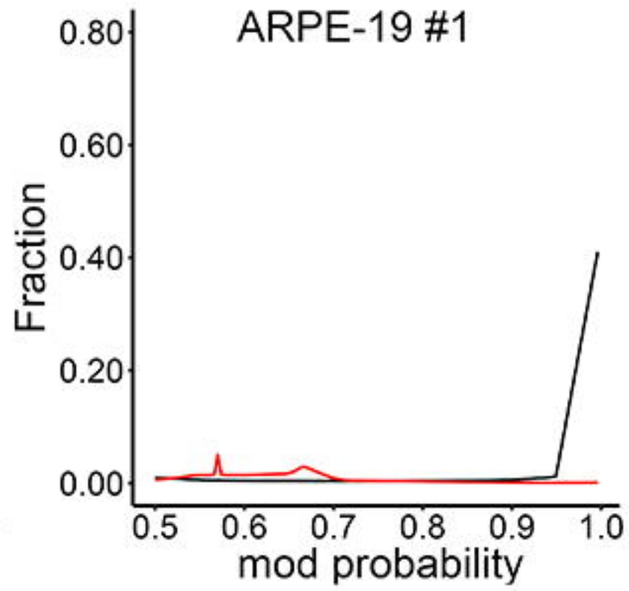


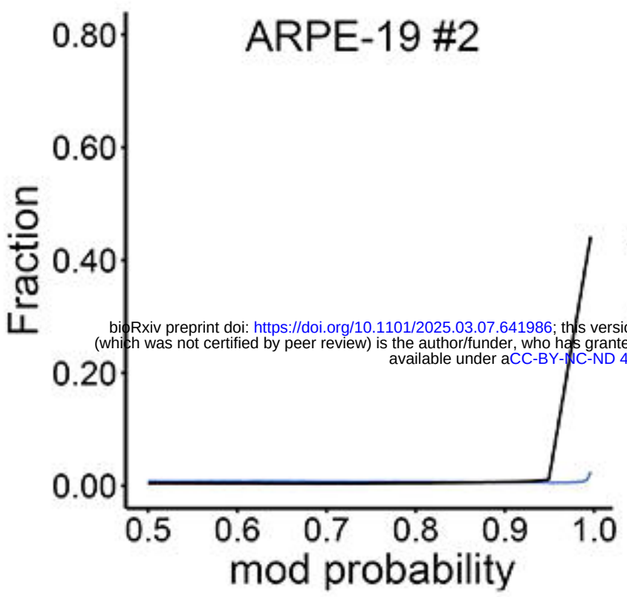


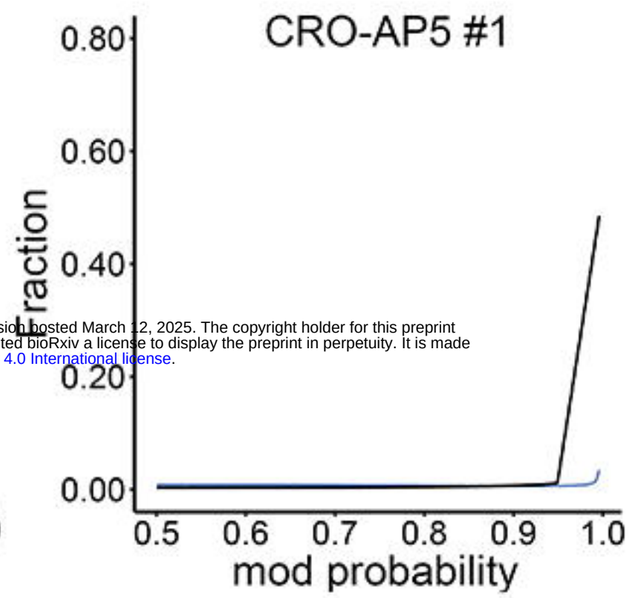


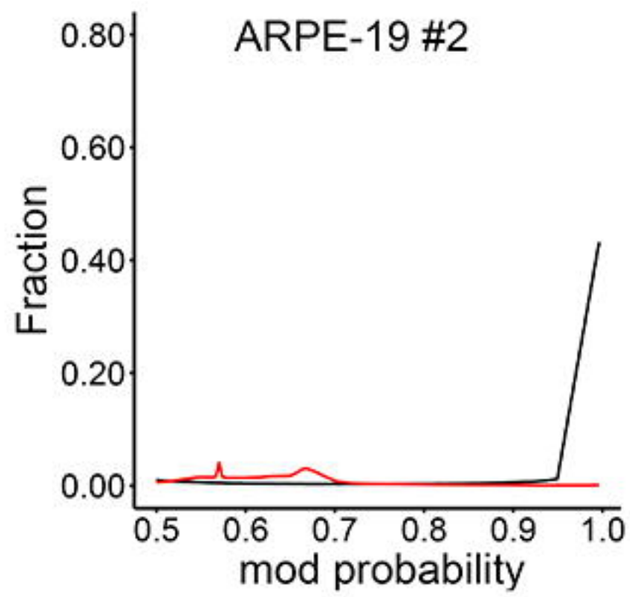


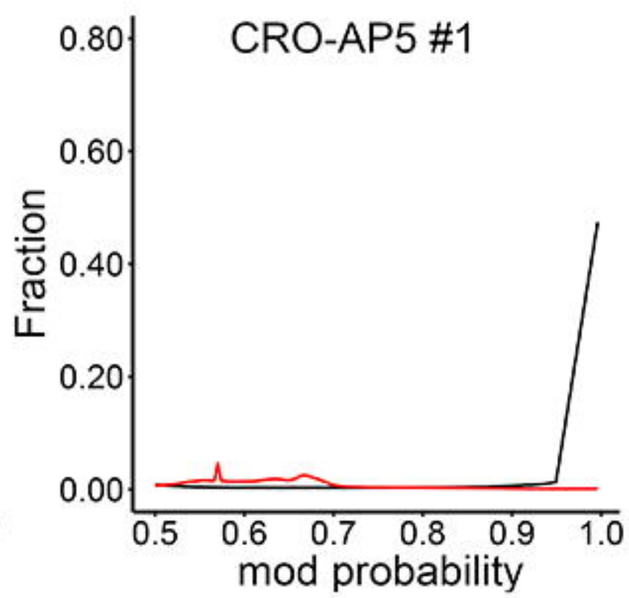


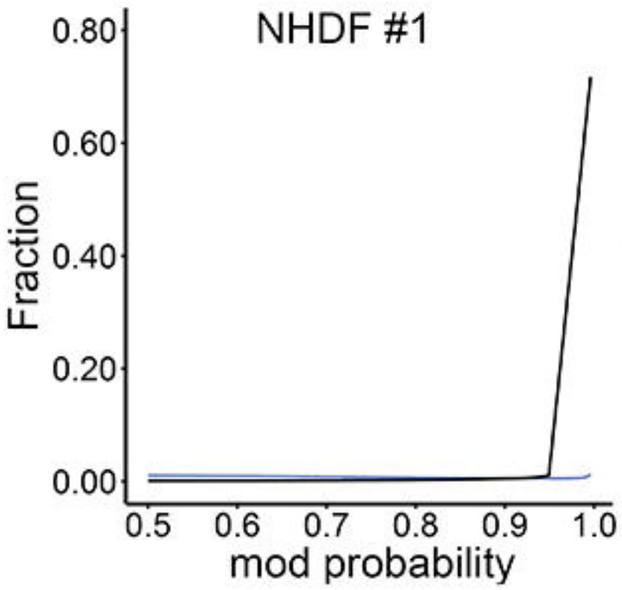


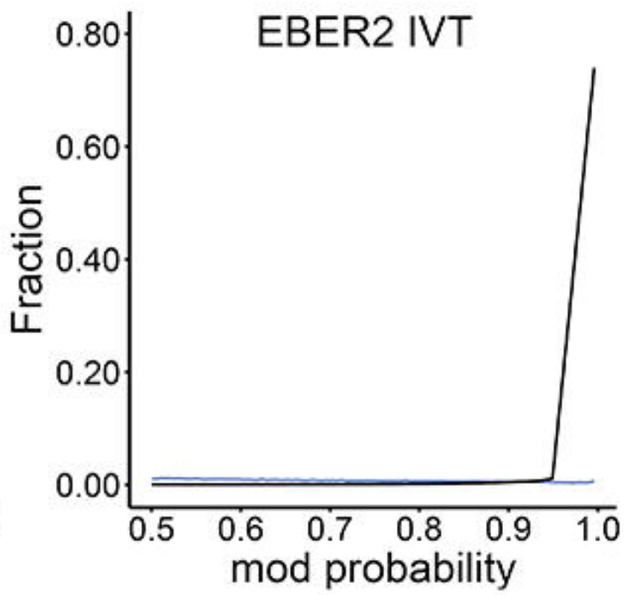


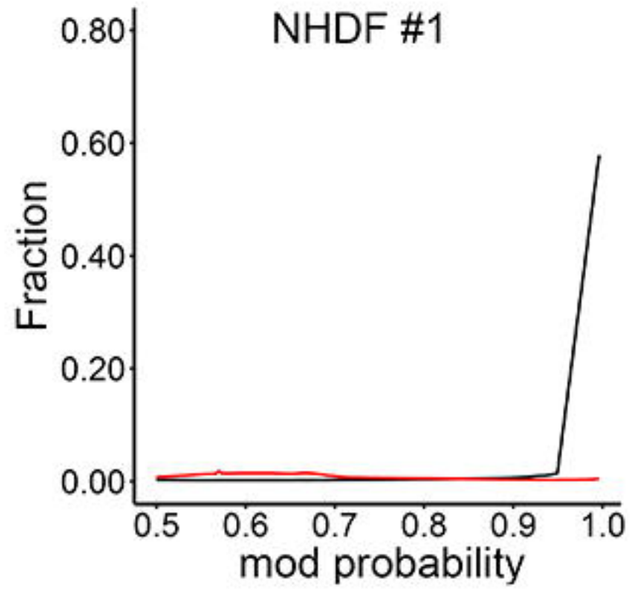


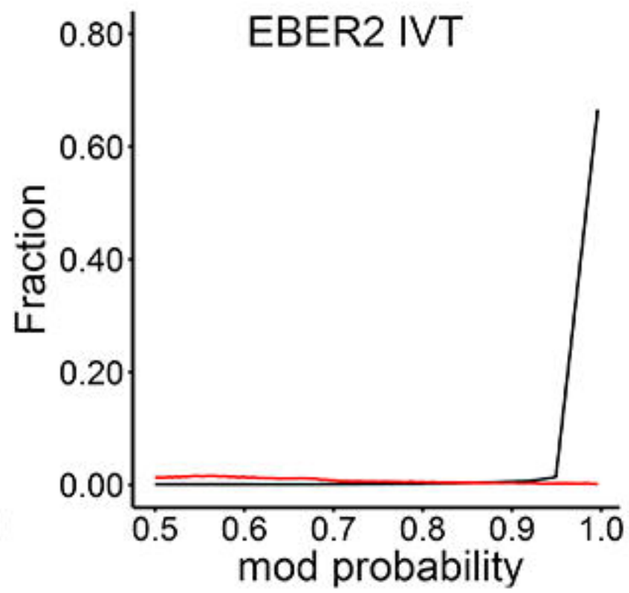


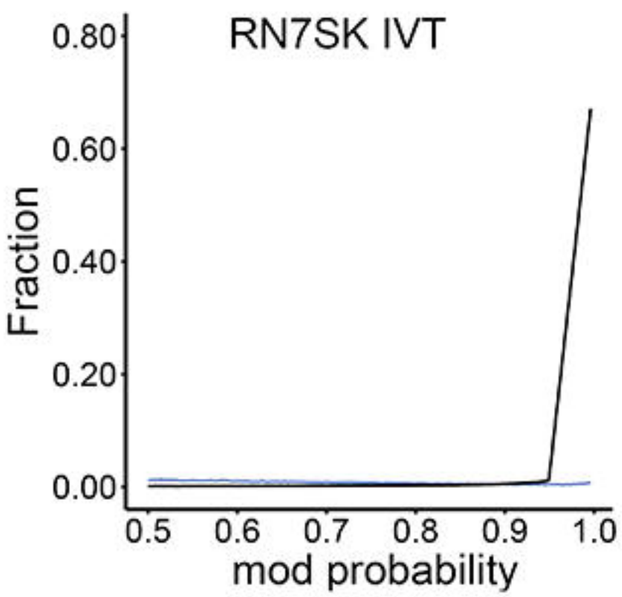


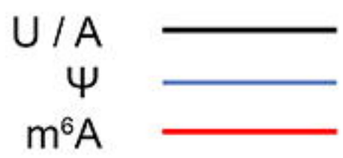


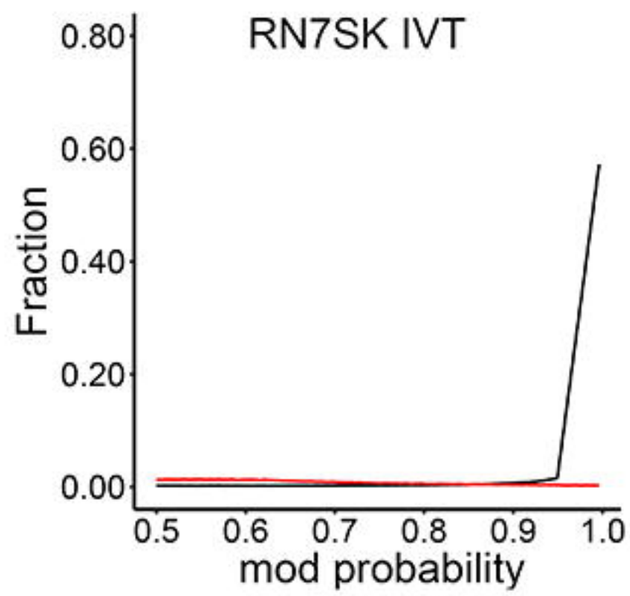


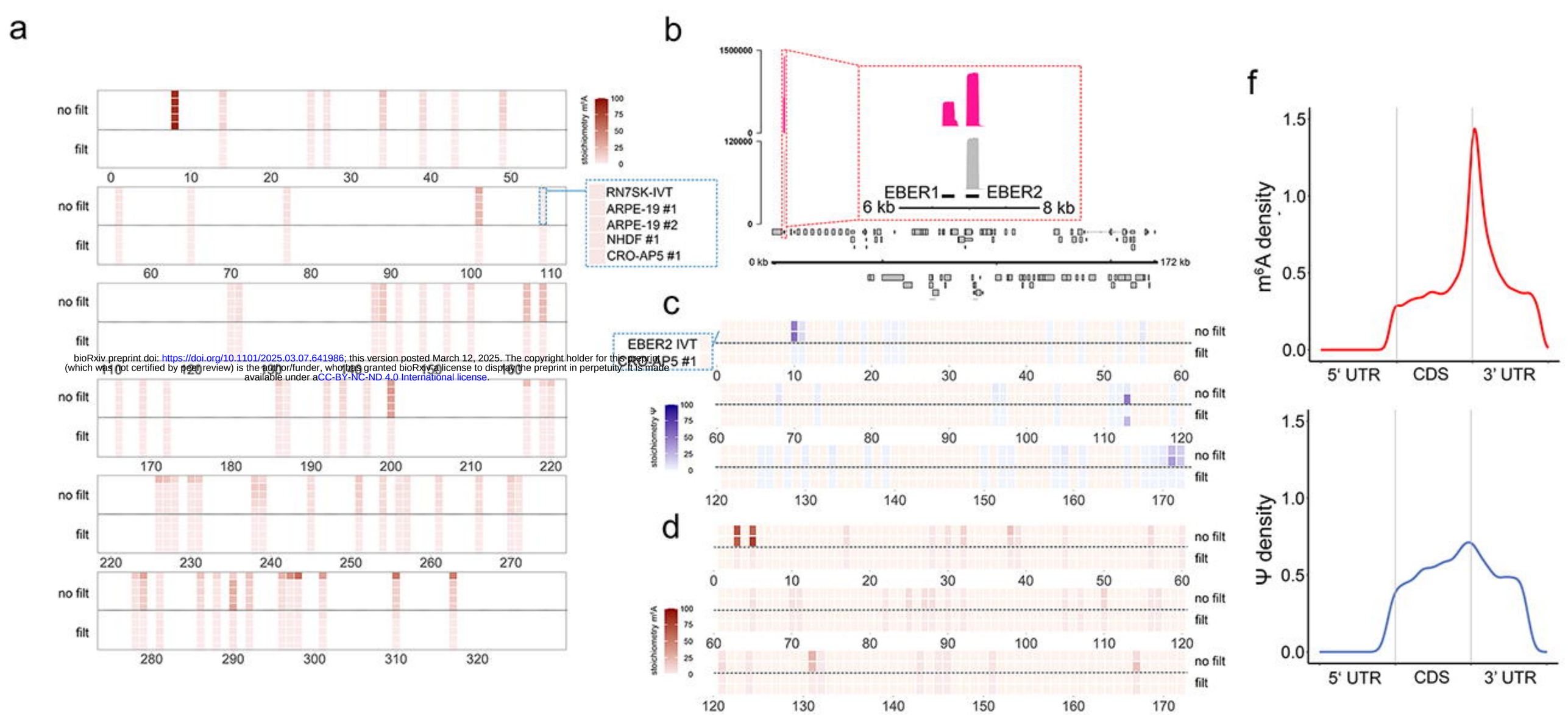


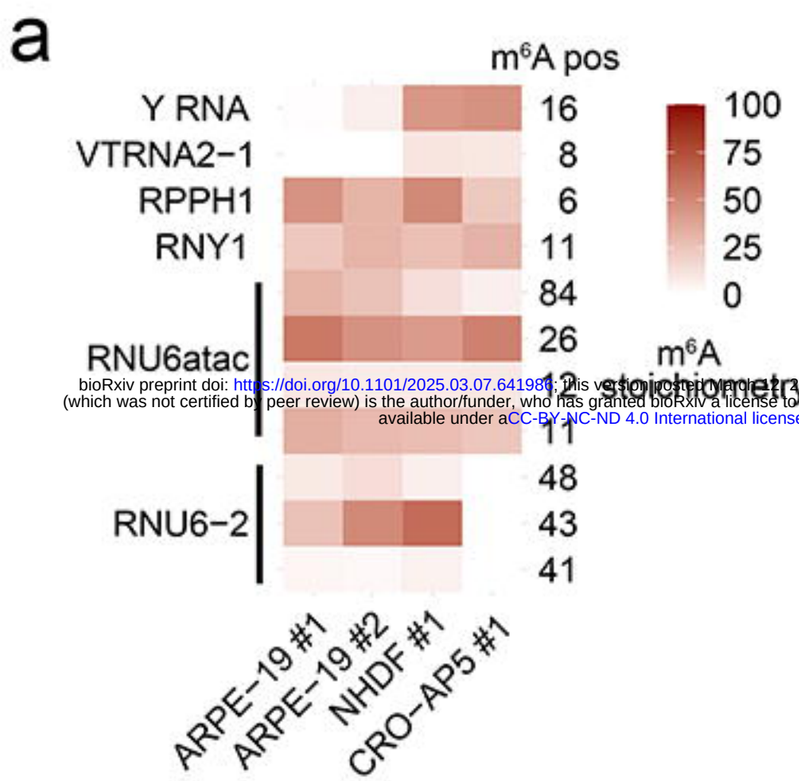


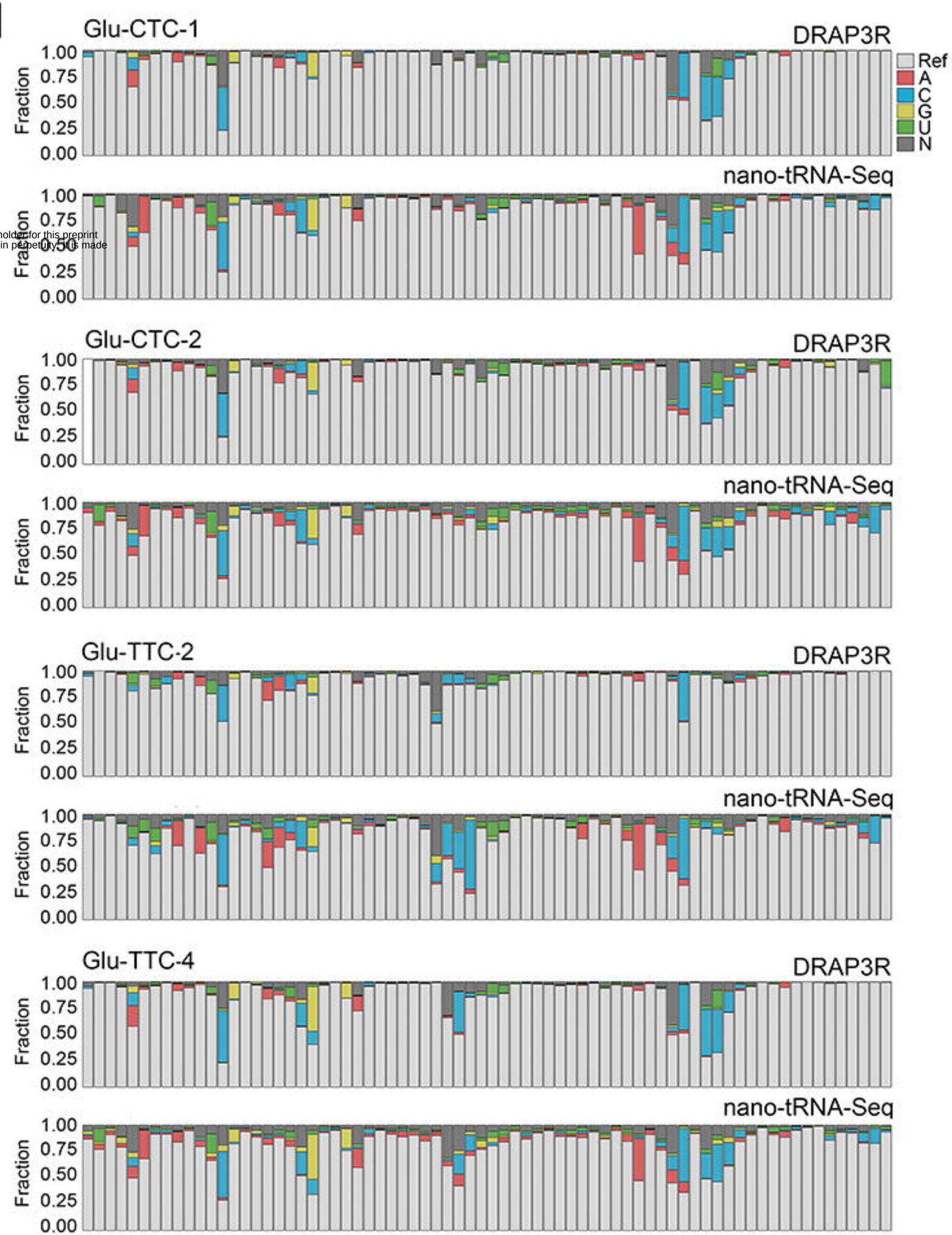


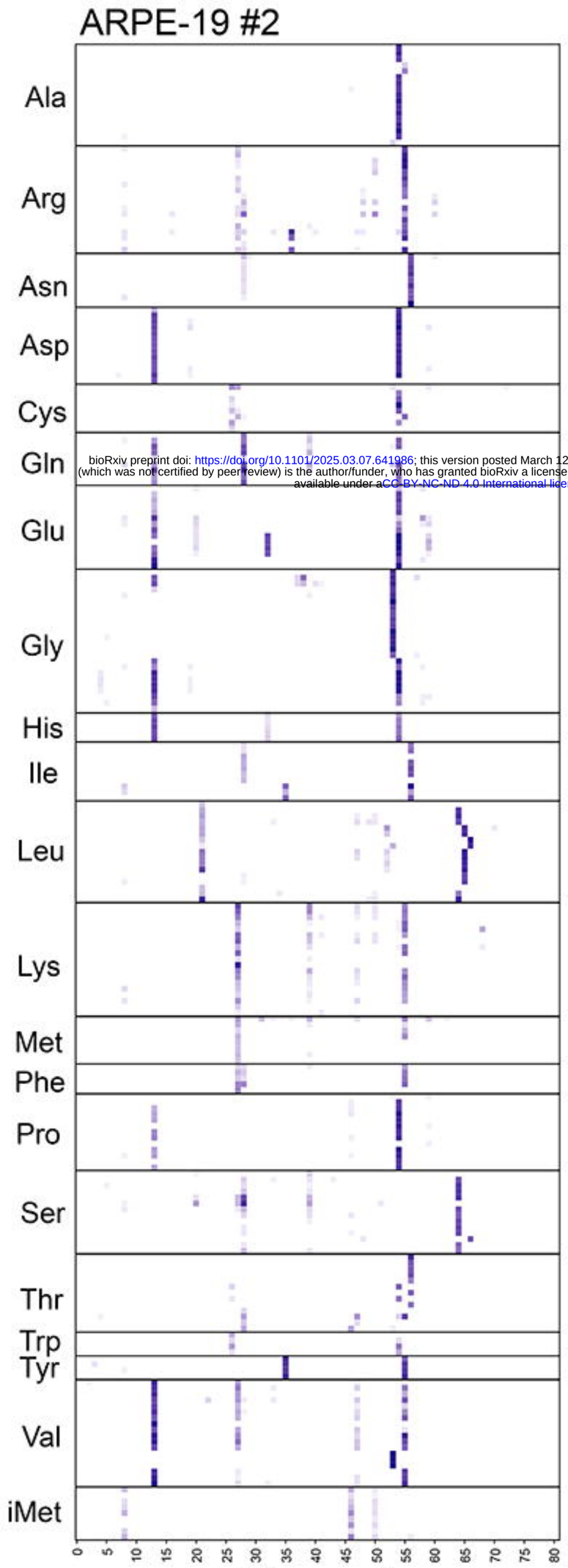


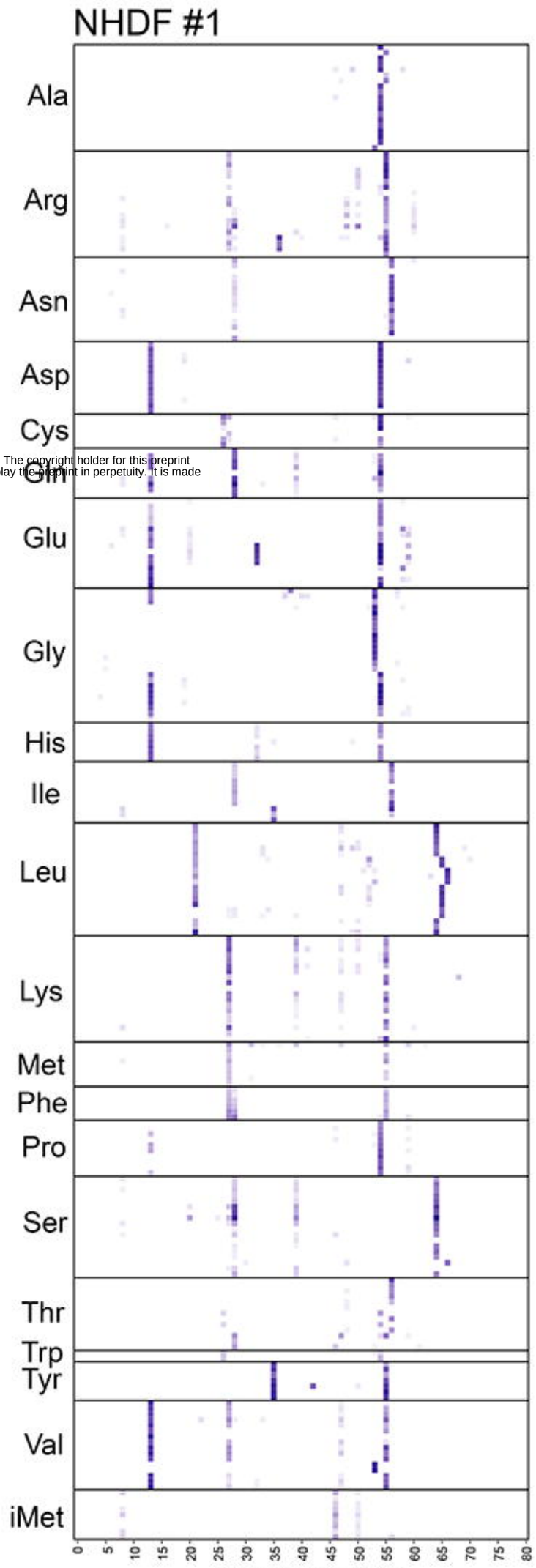


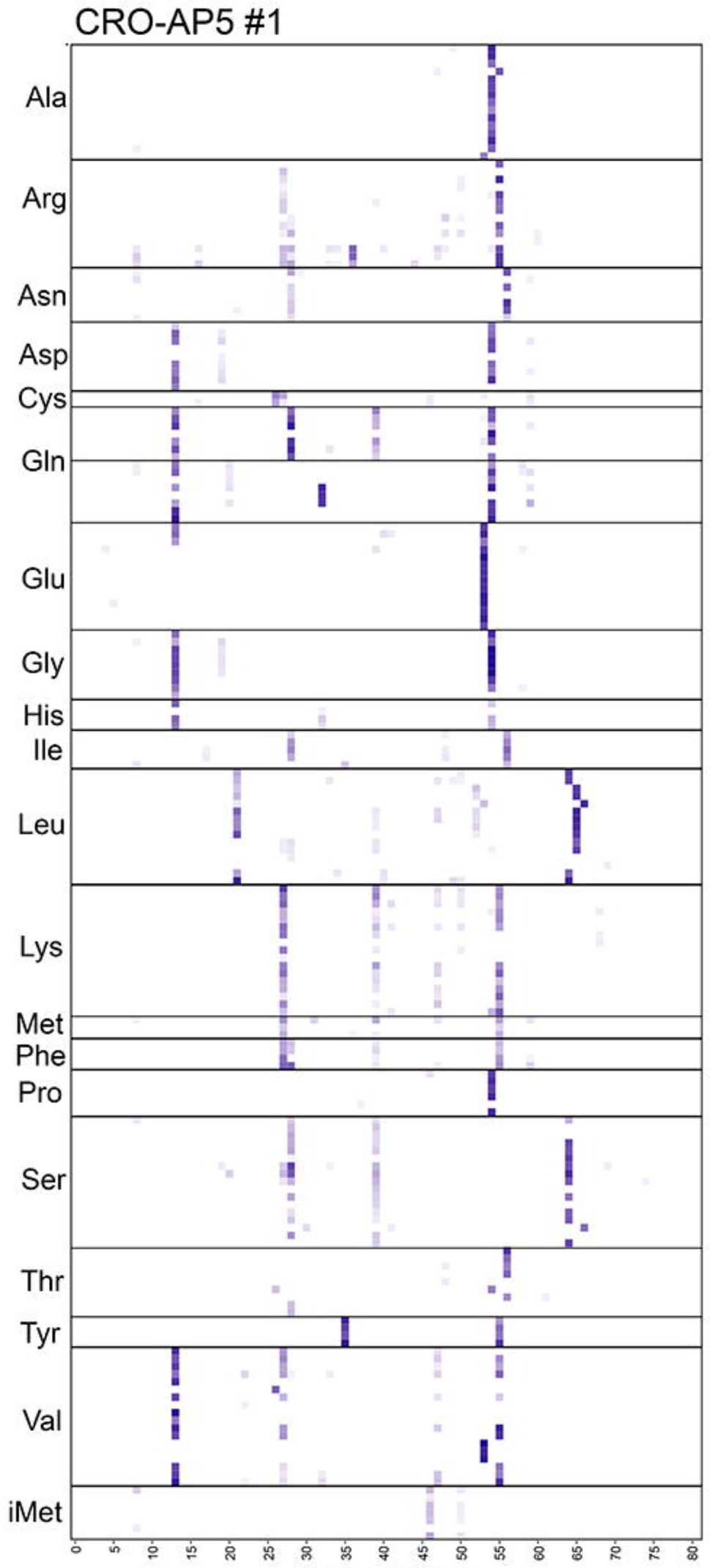




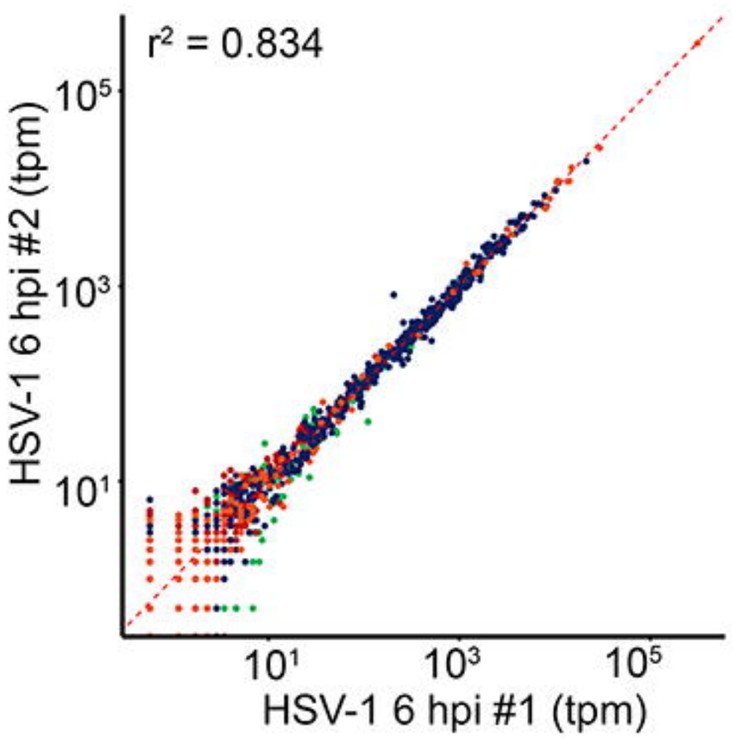




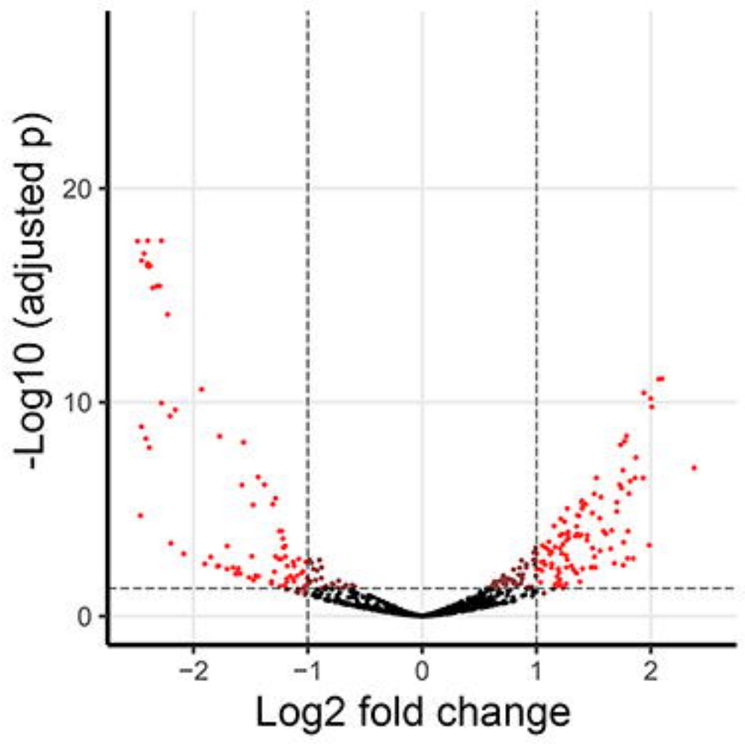




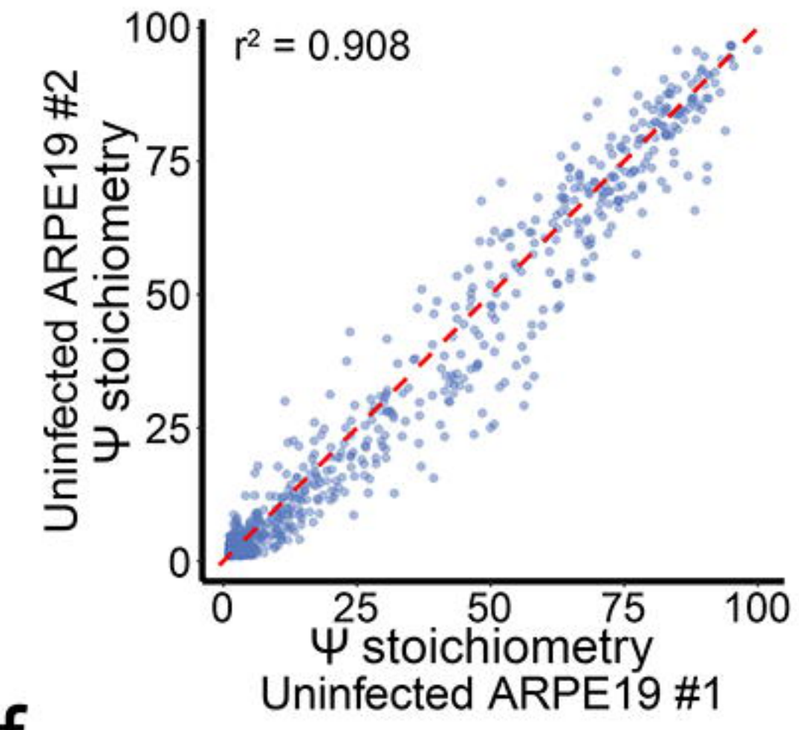
a



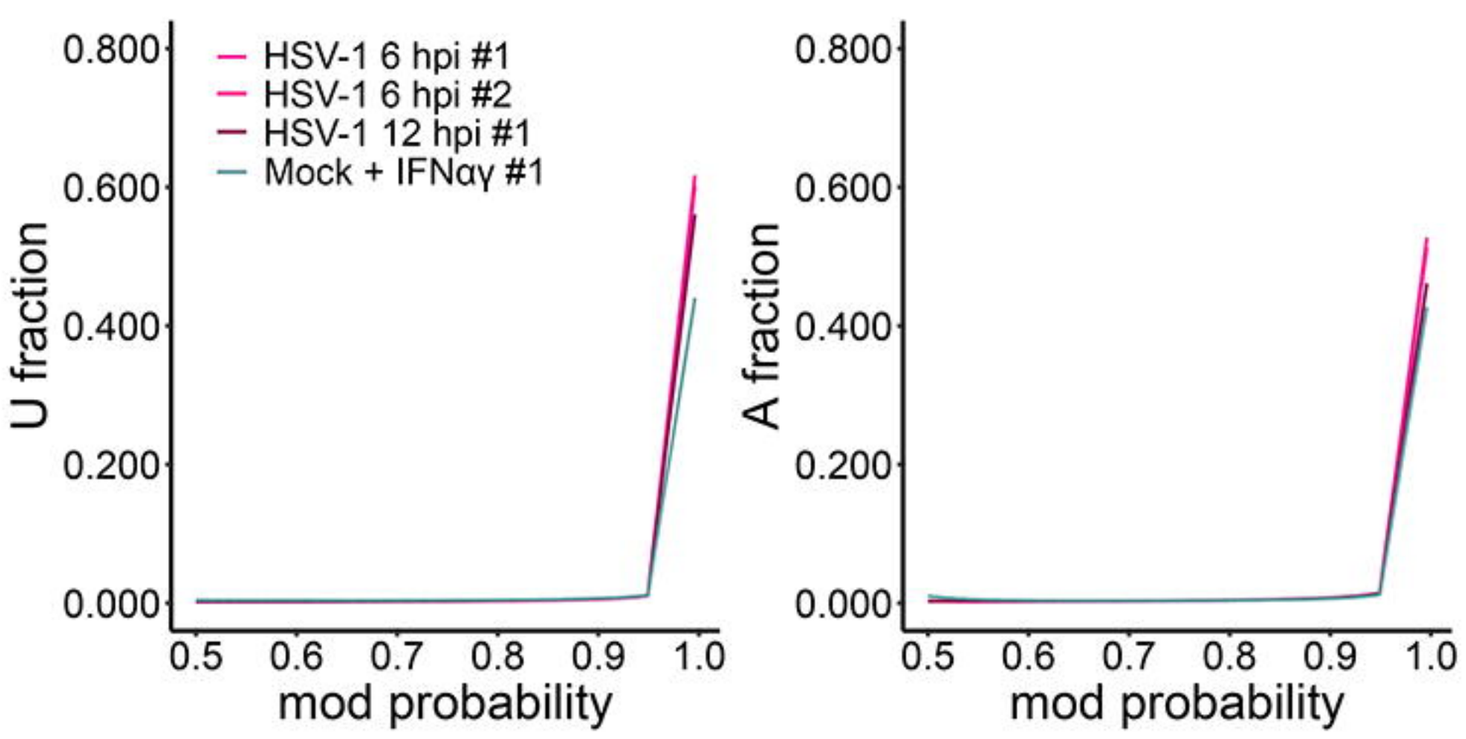
b

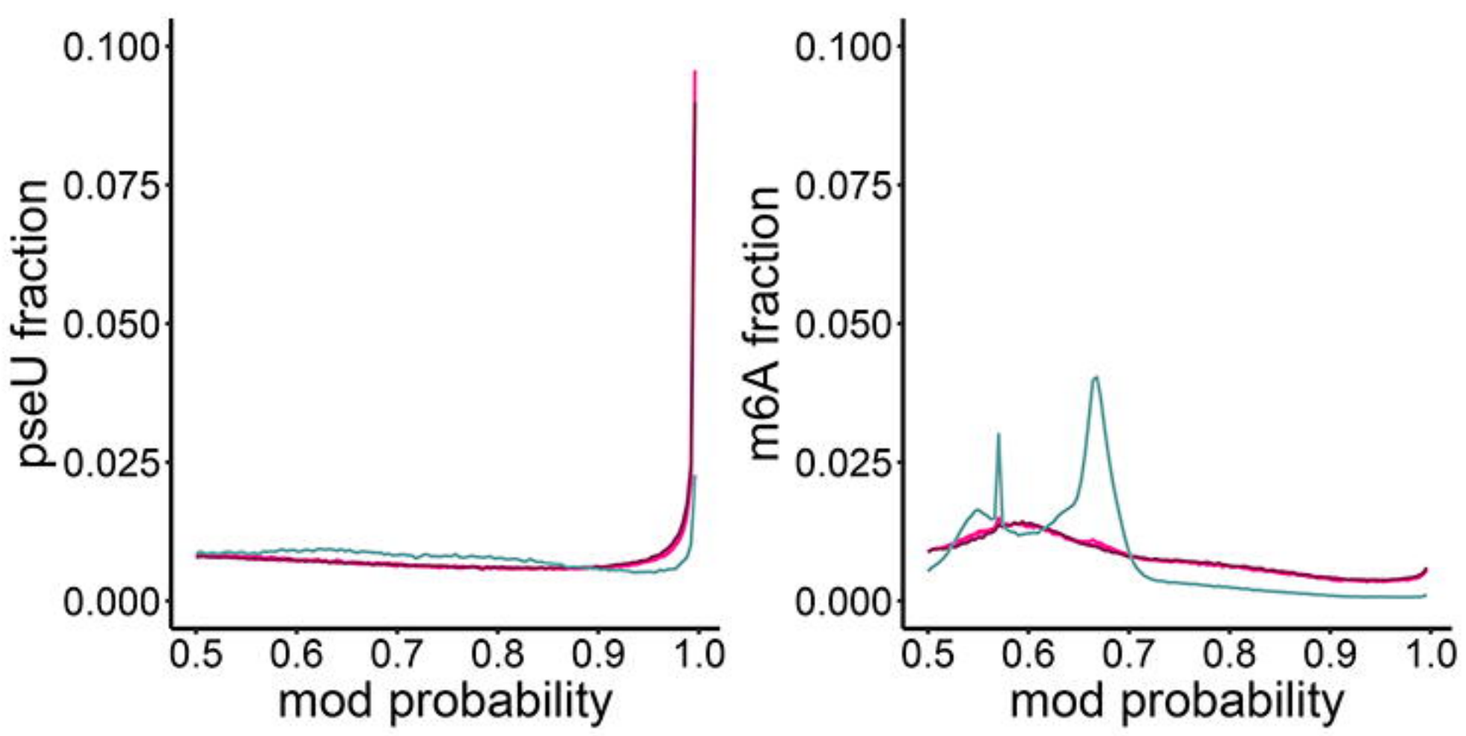


e

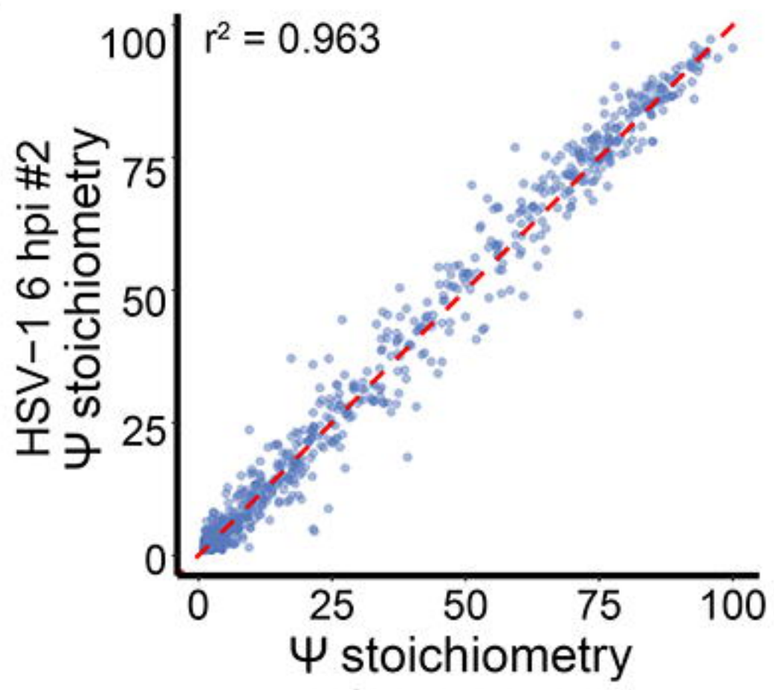


c

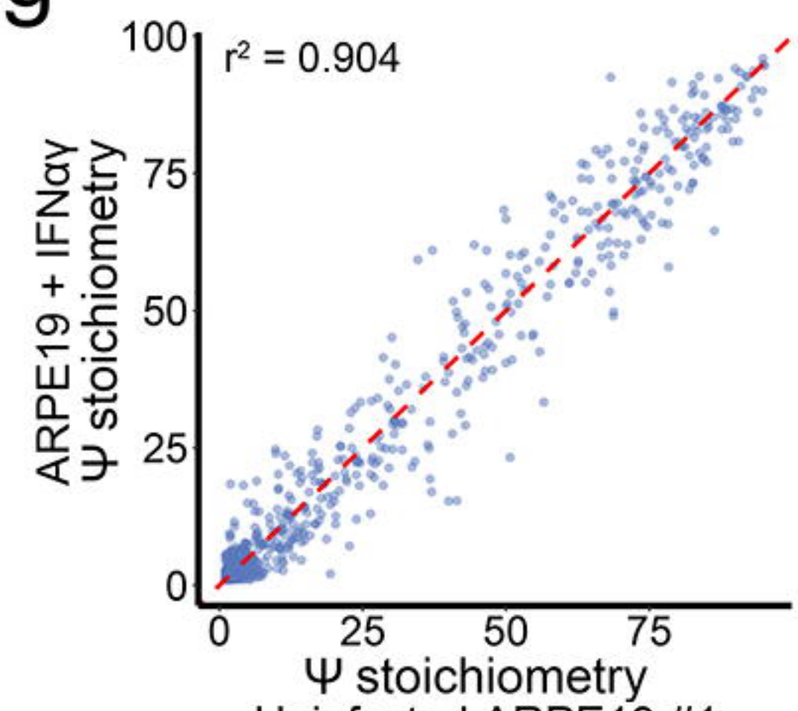




f



g



d