HIRI HELMHOLTZ Institute for RNA-based Infection Research

Nano-DMS-MaP enables isoform-resolved RNA structure determination

Patrick Bohn*, Anne-Sophie Gribling-Burrer*, Uddhav Ambi, Redmond P. Smyth

*these authors contributed equally

Our lab:

HIV-1 is a retrovirus that packages an RNA genome, which upon cell entry is reverse transcribed and integrated into the host cell genome. Newly transcribed HIV RNA from this proviral DNA may then be spliced into multiple different transcripts to express viral proteins, or be packaged into the viral genome. Intriguingly, while all HIV RNAs share the majority of the 5' UTR, which contains a major packaging motif, only the unspliced genomic RNA is efficiently packaged into virions. We therefore hypothesized that **structural differences play a major role in regulating this packaging activity**.

Current RNA structural probing methods, such as DMS-mutational profiling sequencing (DMS-MaP), use NGS technologies whose **short reads cannot be reliably mapped to specific transcripts for shared regions of transcript isoforms**. To circumvent this issue we developed and benchmarked Nanopore-based DMS mutational profiling (Nano-DMS-MaP) to simultaneously investigate RNA structures HIV-1 spliced and unspliced transcripts in cells [1]. Our results demonstrate that HIV-1 splice variants have distinct structures that cluster with acceptor site usage. Importantly, our data suggest that in addition to increasing protein diversity, alternative splicing also resultsv in the generation of RNA transcripts with distinct functions through altered RNA structures.

Nanopore DMS-MaP Workflow

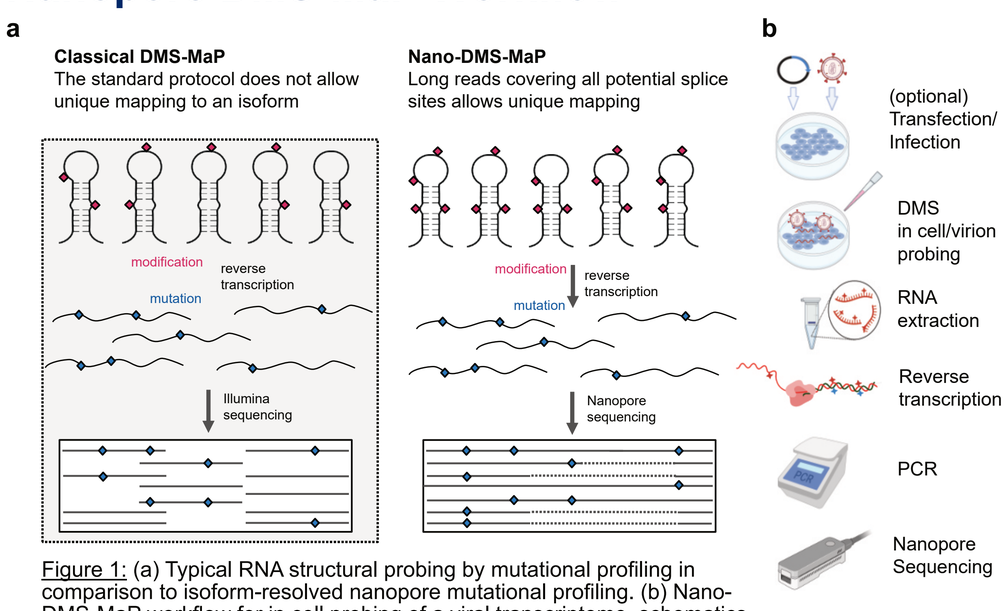

Figure 1: (a) Typical RNA structural probing by mutational profiling in comparison to isoform-resolved nanopore mutational profiling. (b) Nano-DMS-MaP workflow for in cell probing of a viral transcriptome, schematics from Biorender.

Nano-DMS-MaP generates high quality structural probing data

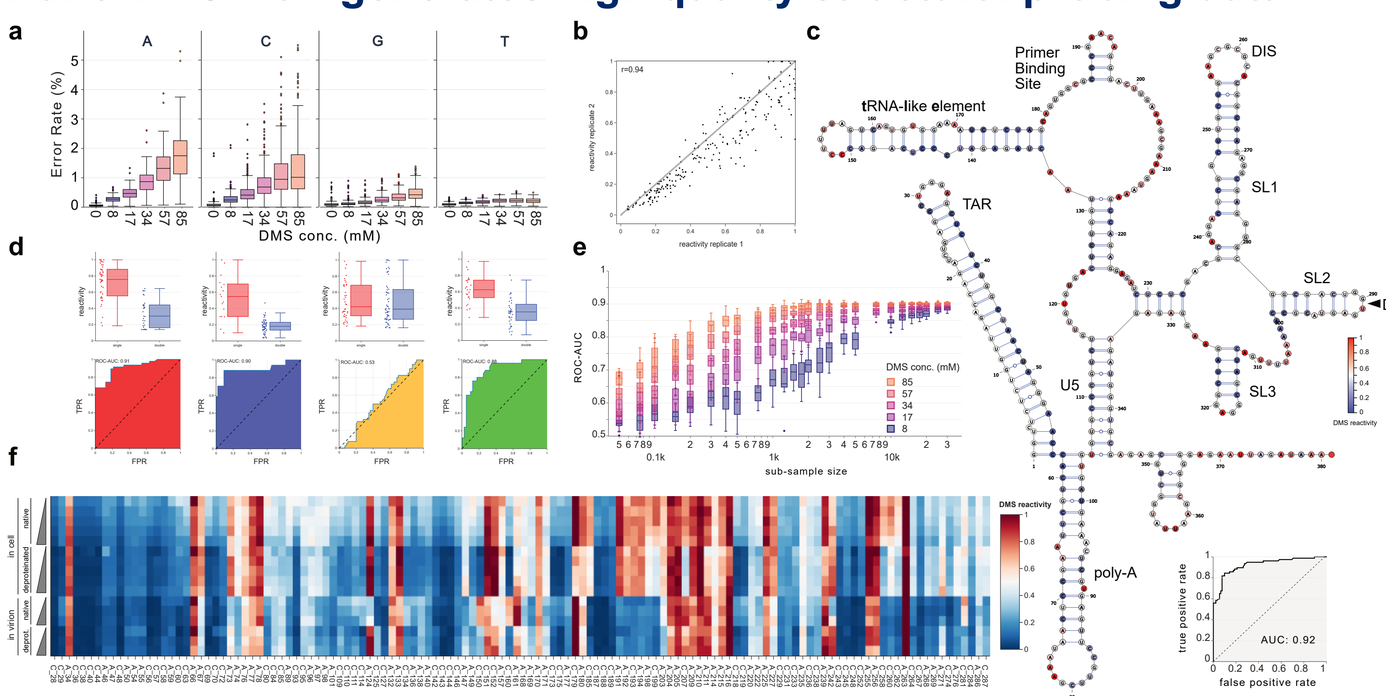

Figure 5: (a) Error rates per nucleotide with increasing DMS concentration. (b) Correlation of reactivity between replictes. (c) Canonical HIV NL4-3 5' UTR structure with reactivity of A, C and U residues colored. (d) Per nucleotide reactivity of single- (ss) and double-stranded (ds) A and C nucleotides and receiver-operator curve (ROC) for single-strandedness vs reactivity. (e) Dependency between coverage and correlation of single-strandedness and high reactivity. (f) Heatmap of reactivity of A and C residues of unspliced RNA in cells and virions in native state or after proteinase K treatment ("deprot.").

The HIV splicing landscape

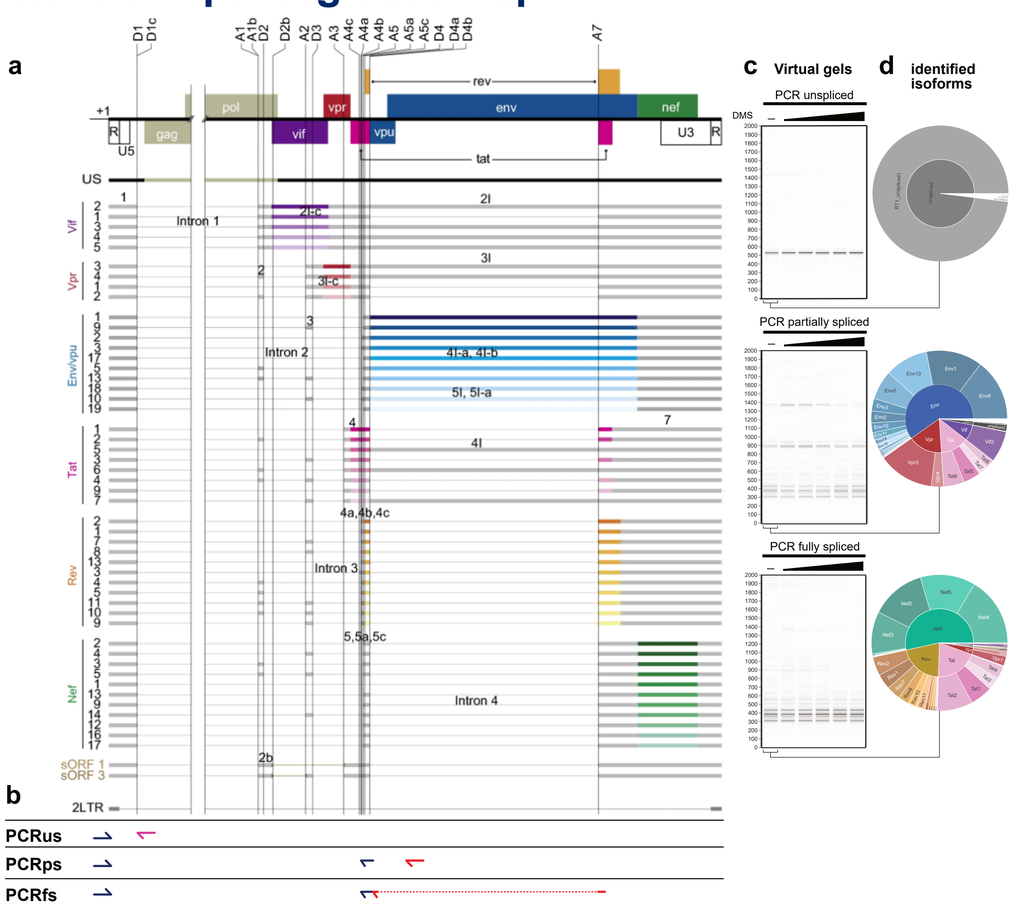

Figure 2: (a) Subset of all identified HIV isoforms, figure reproduced from Nguyen et al. 2020 [2]. (b) Our PCR strategy to amplify all possible isoforms with three PCRs. Reverse transcription primers in red/purple, PCR primers in blue/purple. us: unspliced, ps: partially spliced (contains intron 4), fs: fully spliced (D4A7 spliced) (c)Virtual gels showing read length distriburtions of the sequenced molecules at different DMS concentrations for the PCRs. (d) Relative distribution of identified isoforms for the 0 mM DMS samples as assigned by Isoquant v2.0.0 [3].

Isoform-resolved probing reveals splicing-dependent structural changes

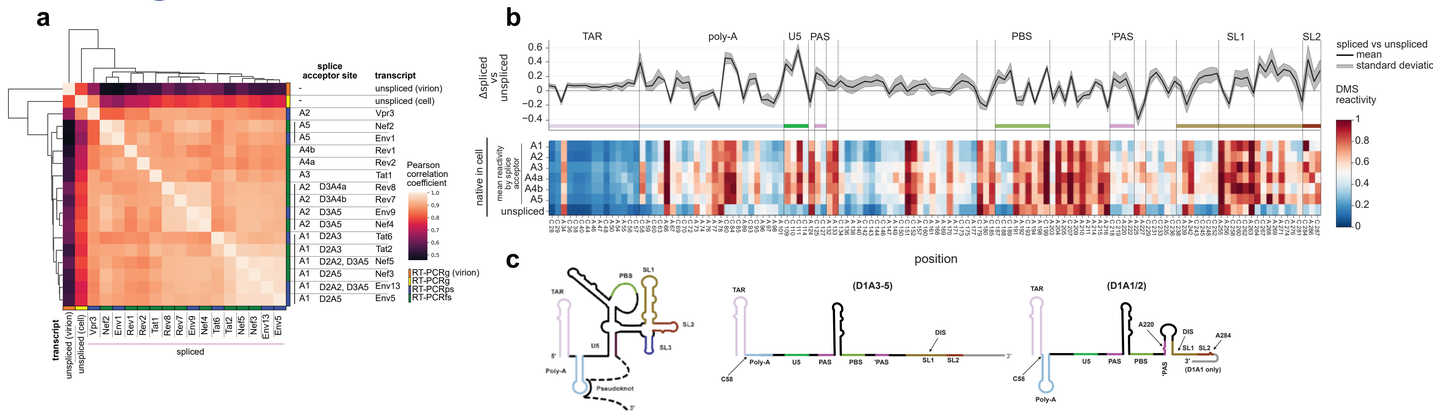

Figure 6: (a) Correlation matrix and clustering of common 5' UTR sequence DMS reactivity per isoform (57 mM native in cell, at least 4000-fold coverage). (b) Heatmap and graph showing DMS reactivity of spliced transcripts (mean per first acceptor site) and unspliced HIV RNA. (c) DMS-guided Eternafold-predicted RNA structures of unspliced and spliced HIV RNAs.

Near-full length HIV genome probing

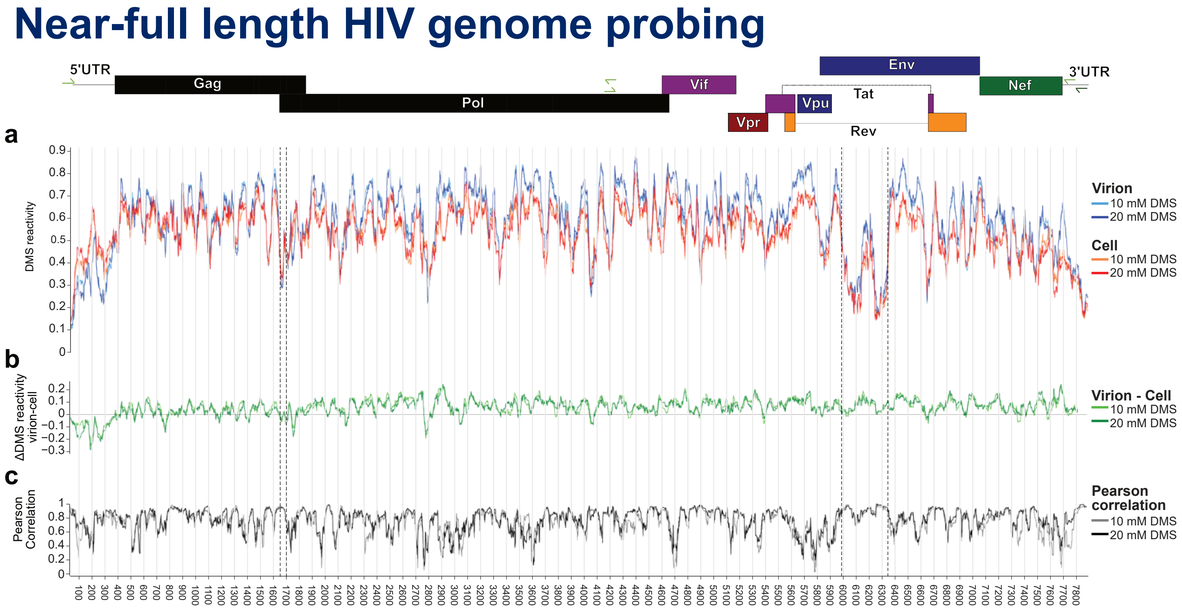

Figure 7: (a) DMS reactivity (mean of 50 nt sliding window) across the whole HIV genomic RNA in cell and in virion. (b) Difference in DMS reactivity between cell and virion. (c) Pearson correlation between DMS reactivities in cell and virion in 50 nt sliding window.

Platform-specific error rates

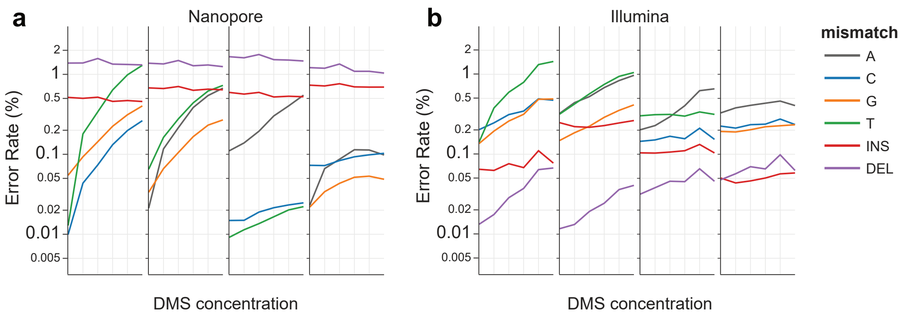

Figure 4: Empirical mean mismatch rates of unspliced HIV 1 RNA amplicons probed at increasing DMS concentrations, sequenced with (a) nanopore Kit 12 chemistry or (b) Illumina NovaSeq. Mismatch rates are quantified from LAST (Nanopore) or bowtie2 (Illumina) alignments using perbase with minimal per base Phred qscore 22.

References

[1] Bohn, P., Gribling-Burrer A., Ambi U., Smyth RP *Nature Methods* (in print 2023).

[2] Nguyen Quang, N. , Goudey S., Ségéral, E. *et al. Retrovirology* **17**, 25 (2020). PMID: 32807178

[23 Prjibelsk A. *et al.* https://github.com/ablab/IsoQuant/releases/tag/v2.3.0

[4] Incarnato D., Morandi E., Simon LM., Oliviero S., *Nucleic Acids Res.* (2018). PMID: 29893890

In cooperation with

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

A site of

HZI HELMHOLTZ Centre for Infection Research

www.helmholtz-hiri.de