High-resolution HIV-1 m^{6}A epitranscriptome reveals splicing-dependent methylation clusters and unique 2-LTR transcript modifications

Delphine Naquin*, Sandra Blanchet*, Erwin van Dijk, Lisa Bertrand, Sylvie Grégoire, Cecilia Bertha Ramirez, Rania Ouazahrou, Yan Jaszczyszyn, Arnaud Moris, Claude Thermes, Céline Hernandez, Olivier Namy

ABSTRACT

The *N*^{6}-methyladenosine (m^{6}A) modification of HIV-1 RNAs plays an essential role in regulating viral infection. This modification has been widely studied but the number and precise positions of the m^{6}A sites remain unclear due to the lack of precision of detection methods. We used the latest Nanopore chemistry and direct m^{6}A base-calling option to identify 18 m^{6}As, 14 of which were located at the 3' end of the genome, the other four being located in central regions. Our data reveal differential methylation of these positions between splicing isoforms. Eleven of these sites are clustered in two short segments (197 nt and 43 nt) with peak-shaped methylation profiles. Single-molecule analysis revealed that a very small number of transcripts were unmethylated in both clusters (≤ 1.5% of spliced transcripts and ≤ 5.4% of unspliced viral RNA). We also identified a ~732 nt RNA species resulting from the transcription of non-integrated viral DNA circles closed by two long terminal repeats (2-LTR circles). These transcripts started in the first LTR, terminated at the polyA site of the second LTR and harbored six m^{6}A sites. Five of these sites were present in other transcripts and, remarkably, had the highest methylation rates. The sixth site was methylated only in this transcript, suggesting a role for this RNA in HIV-1 infection. These findings reveal a new landscape of HIV m^{6}A transcriptome modifications and pave the way for studies deciphering their role in the viral life cycle.

INTRODUCTION

Epitranscriptomic RNA modifications play crucial roles in diverse cellular and viral processes and diseases (1). The N6-methyladenosine (m^{6}A) modification occurs frequently, in all types of cellular RNA, and consists of the covalent addition of a methyl group to the N6 position of adenosine. This methyl group addition is catalyzed by

a multi-subunit methyltransferase complex with a catalytic core composed of methyltransferase-like 3 (METTL3) and methyltransferase-like 14 (METTL14) (2). This writer complex preferentially targets adenosines in the consensus sequence 5′-DRACH-3′ (D = A/G/U, R = A/G, H = U/A/C) in mammalian mRNAs (3). DRACH motifs are frequent in mRNA but are not always targeted, as only about 10% of these motifs were found to be m^6A-methylated in cells. This suggests that there are probably other context-dependent signals for the methylation of specific DRACH motifs. The m^6A modification is reversible and can be removed by erasers, such as alkB homolog 5 (ALKBH5) in particular (4). Reader proteins recognize m^6A and determine the functional consequences of the modification for the fate of the RNA. The readers capable of binding m^6A include the YTH domain-containing family of proteins, the most frequently studied members of which are YTHDC-1, YTHDF1,2,3 (5, 6). However, discrepancies remain concerning the function of these proteins (7). In mammals, m^6A is the most prevalent modification involved in gene expression regulation and RNA metabolism, including mRNA transport, localization, stability, translation efficiency and processing (8-13), but this modification is also present in viral RNA (14). Several studies have detected methylated regions (m^6A) in the HIV-1 genome with techniques based on methylated RNA immunoprecipitation (15-21). However, it was not possible to localize the methylated positions with single-nucleotide precision because the resolution of the techniques used ranged from 20 nt to more than 200 nt. There were also differences between studies, with some reporting peaks distributed along most of the length of the genome (15, 18), whereas others detected peaks only at the 5’ and 3’ ends (17, 19) or only at the 3’ end (16). These discrepancies may reflect differences between the cell types or infection protocols (cell transfection with proviral plasmids or infection with viruses) used or a lack of reproducibility of antibody-based methods (22). In principle, Oxford Nanopore Technologies (ONT) direct RNA sequencing (23) would allow the detection of modified RNA nucleotides at single-molecule level. Several studies have provided diverse bioinformatics tools for the detection of RNA modifications based on various strategies (base-call errors induced by modifications, altered current intensity, modification-aware base-calling models, reviewed in (24)). However, improvements in data quality and recovery rates for full-length transcripts are required to enable direct base-calling for the modified nucleotides (25). Here, we considered the full-length genome, including the 5’LTR and 3’LTR sequences. A recent study using the RNA002 kit resulted in the development of a modified library protocol and major improvements in the recovery of full-length RNA molecules (26). Combining several bioinformatics tools based on different properties of the ionic current, the authors identified three m^6A sites at the 3’ end of the viral genome. We made use of an advance developed by ONT — RNA004 chemistry (new nanopore and new motor protein) — that greatly

57 improves overall yield and data quality (27), and new base-calling models making it possible to call m^6As
58 directly with the four standard bases. Using these developments, we characterized the HIV-1 m^6A methylation
59 pattern of HIV-1 mRNAs in infected $CD4^+$ T cells, including a $CD4^+$ T cell line (SupT1) and, importantly,
60 primary $CD4^+$ T cells. We extended the detection to 18 m^6A sites at single-nucleotide resolution, single-
61 molecule level and in different splicing isoforms, providing the most complete and precise description of the
62 landscape of HIV-1 m^6A methylation to date.
63
64 **RESULTS**
65 **Nanopore sequencing of full-length HIV-1 transcripts and identification of splicing isoforms**
66 For the detection of HIV-1 m^6A sites, cellular transcripts from SupT1 and primary $CD4^+$ T cells infected with
67 HIV-1 were extracted and sequenced by ONT direct RNA sequencing (DRS, Materials and Methods). It has
68 been reported that it is very difficult to obtain reads corresponding to full-length HIV RNA molecules with the
69 standard DRS library preparation protocol (26). In our previous DRS experiments with the standard protocol,
70 we obtained less than 0.008% and 0.06% of total reads for our first two samples, corresponding to the ~9.6 kb
71 unspliced HIV-1 genome. In this study, we used a modified protocol described elsewhere (26) to increase
72 recovery rates for full-length transcripts. We used two biological replicates of SupT1 cells (rep1 and rep2)
73 harvested 20 h and 30 h post-infection and primary $CD4^+$ T cells isolated from the blood of two different blood
74 donors collected 48 h and 72 h post-infection (Figure S1). PolyA RNA was isolated from each of these six
75 infected samples and sequenced to generate 7.3 to 28.9 million total reads (Table S1) and between 33,587
76 and 164,360 reads corresponding to full-length HIV transcripts (Table S1). Full-length reads were examined for
77 the presence of splice junctions. We identified eight known donor sites (four major sites — D1, D2, D3, D4 —
78 together with D1c, D2b, D4a and D4b) and eleven acceptor sites (four major sites — A1, A2, A5, A7 —
79 together with A1b, A3, A4a, A4b, A4c, A5a and A5c) (28-31) (Table S2A). The different isoforms were identified
80 by the succession of donor and acceptor sites (Materials and Methods); we identified 51 isoforms, as
81 previously described (31, 32) (Table S2B). These isoforms are typically classified into three groups on the
82 basis of size: (i) unspliced ~9 kb transcripts (US), (ii) partially spliced transcripts (~4 kb, PS, lacking intron 1
83 and retaining intron 4) including Vif2, Vpr3,4, Env/Vpu and Tat5, and (iii) completely spliced transcripts (~2 kb,
84 CS, lacking introns 1 and 4) including Vif1, Vpr1,2, Tat1-4, Rev and Nef. The relative abundances of the
85 various isoforms were similar in all samples (correlation R>0.86 in all comparisons, Table S3).
86

87 **Detection of m^{6}A sites in all HIV-1 transcripts**

88 The first step was the detection of m^{6}A sites in all HIV-1 reads. The positions of N6-methyladenosine
89 nucleotides were determined directly at read level during the base-calling process, providing a first estimate of
90 methylation rates at each position in a set of transcripts. These methylation rates were compared with those
91 for the corresponding positions in the unmodified RNAs transcribed *in vitro* (Figure 1A). We performed this
92 analysis for all HIV-1 transcripts from the six samples, resulting in a set of 14 putative m^{6}A sites, all located in
93 the 3' region of the HIV-1 genome common to all transcripts (Tables 1, S4; Figure 1B). The numbering used in
94 this paper corresponds to that of the full-length genome, including the 5'LTR and 3'LTR sequences
95 (transcription start site in position +454). The corresponding methylation landscape presented two clusters of
96 sites. The first cluster consisted of eight sites (A8436 to A8632) with two highly methylated central sites (64-
97 80%), and the second cluster consisted of three sites (A9400, A9428, A9442) with a highly methylated central
98 site (79-84%). Three additional sites displaying low rates of methylation were detected between these two
99 clusters. Remarkably, the segment containing the first cluster (8436 to 8632) was found to harbor five DRACH
100 sequences (3, 33), all of which were methylated. The other three sites of this segment correspond to 1-nt
101 variants of the DRACH motif (found here at positions A8487, A8555, and A8572) interspersed with the
102 canonical DRACH sites (Tables 1, S4; Figure 1B). It should be noted that m^{6}As have also been detected in
103 non-DRACH motifs, especially the variant GGACG (found here at positions A8487 and A8572), which has
104 been found to be highly methylated in human mRNAs (34, 35).

105

m^{6}A position	Sequence	Methylation rate (%)
A8436	AGACA	9..7
A8442	AGACA	10.3
A8487	GGACG	30
A8532	AGACT	79.3
A8555	GGATT	19.3
A8563	GAACT	79.3
A8572	GGACG	23.7
A8632	GAACT	10.3
A9024	TGACT	7.3
A9074	GGACT	8.7
A9163	GAACT	17.3
A9400	GAACT	29.7
A9428	GGACT	84.3
A9442	GGACT	53.3

106

Table 1. *List of the 14 m^{6}A sites detected in the 3' region (8360-9700) common to all transcripts. Methylation rates were measured in the SupT1_20h_pi rep1 sample. The underlined nucleotides are single-nucleotide variants of the consensus DRACH sequence.*

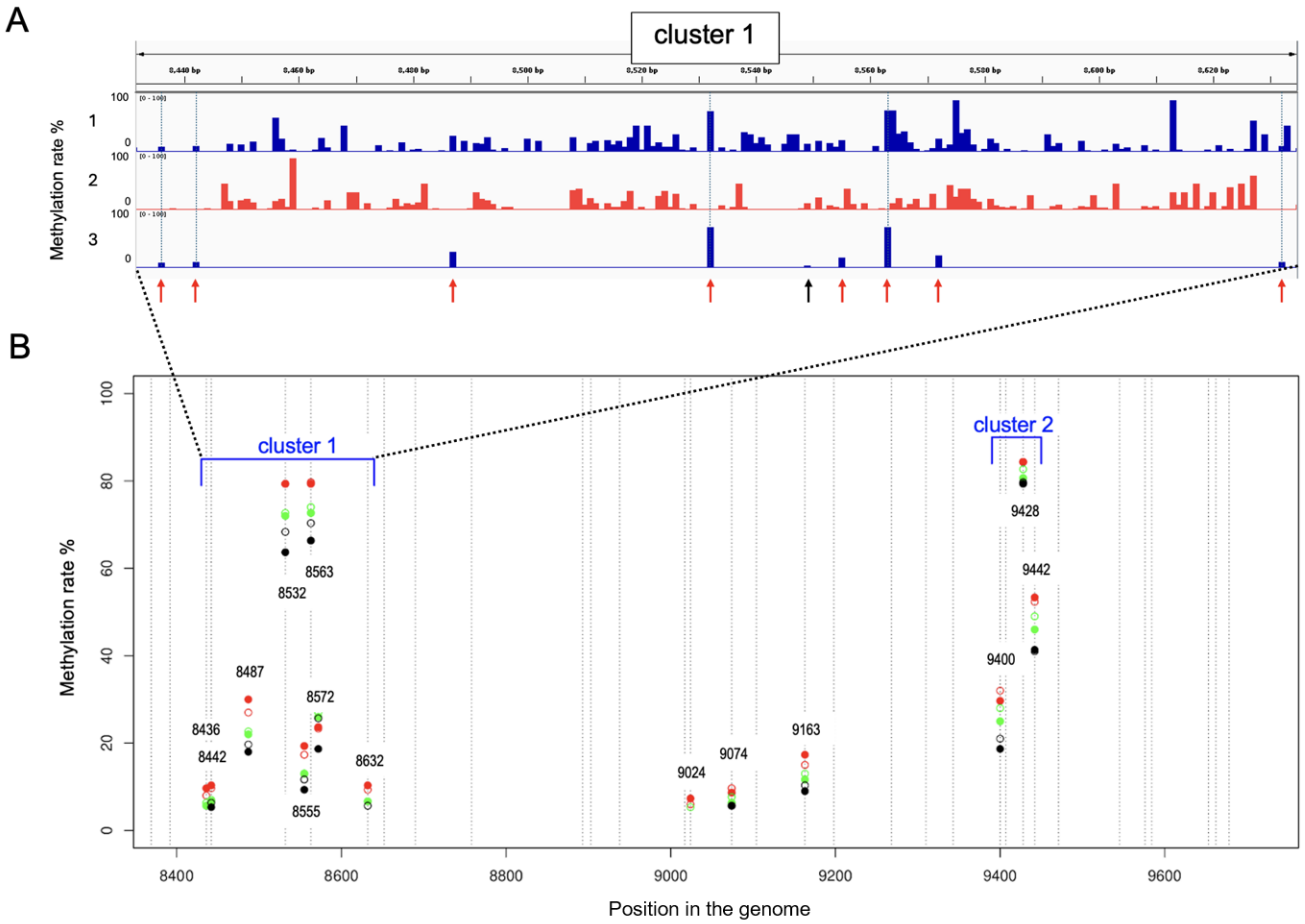
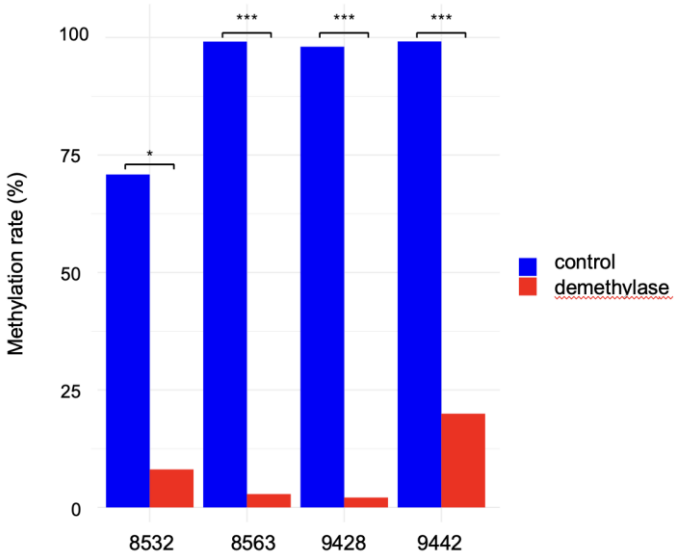


Figure 1. *Detection of m^{6}A sites in all HIV-1 transcripts.* ***(A)*** *The IGV screenshot shows the methylation rate measured in the cluster 1 region of the HIV-1 genome including sites A8436 to A8632. Line 1, methylation rate measured directly during the m^{6}A base-call of SupT1_20h_pi rep1. Line 2, methylation rate measured directly for the in vitro sample in the same region. Line 3, Difference (line 1) - (line 2); red arrows indicate the eight sites detected; negative values are set to zero; sites with rates < 3% were not retained (black arrow). Vertical lines, positions of As in the five DRACH sequences of the region.* ***(B)*** *Methylation profile of all HIV-1 transcripts in the 3' region (8360-9700). Closed red circles, SupT1_20h_pi rep1; empty red circles, SupT1_30h_pi rep1; closed green circles, SupT1_20h_pi rep2; empty green circles, SupT1_30h_pi rep2; closed black circles, CD4_48h_pi; empty black circles, CD4_72h_pi. Vertical lines, positions of As in the 33 DRACH sequences of the region.*

123 **Efficiency of m^{6}A detection in synthetic oligonucleotides**

124 We assessed the efficiency of m^{6}A site detection by subjecting two synthetic RNA oligonucleotides with
125 sequences corresponding to HIV-1 segments but harbouring m^{6}A at positions 8532 and 8563 (oli1-mod) and
126 at positions 9400, 9428 and 9432 (oli2-mod) to the detection procedure. As controls, we analyzed the same
127 HIV RNA oligonucleotides without m^{6}A modifications. Remarkably, all but one (position A9400) of the
128 methylated As were identified (Figure 2, blue bars). The m^{6}A site at position A9400 is located too close (9 nt)
129 to the 5' end of oli2-mod, which prevents its accurate detection (25). The methylated A8563, A9428 and A9442
130 positions were detected with a methylation rate of ~100% in both oligomers (Figure 2). The lower methylation
131 rate (~75%) observed for A8532 is probably due to its position close (13 nt away) to the 5' end of oli1-mod.
132 These observations confirm the specificity and high level of efficiency of m^{6}A detection with RNA004 chemistry
133 and the m^{6}A base-calling option. We further assessed the quality of m^{6}A detection by treating oli1-mod and
134 oli2-mod harboring m^{6}A modifications with the ALKBH5 demethylase and then performing the detection
135 process. Methylation rates were significantly lower at all sites, confirming the efficiency of the m^{6}A detection
136 process (Figure 2, red bars).

137



138

139 ***Figure 2.*** *Efficiency of m^{6}A site detection with methylated synthetic oligonucleotides. In blue, methylation*
140 *rates, measured at the indicated m^{6}A positions, comparing unmodified and m^{6}A-modified oli1-nmod and oli2-*
141 *nmod. In red, the same measurement after the demethylation of synthetic m^{6}A-modified oligonucleotides with*
142 *ALKBH5 (Materials and Methods). *, p-value < 10^{-6}; ***, p-value < 10^{-149}.*

143

144 **Validation of the candidate m^{6}A sites with the ALKBH5 demethylase**

145 For further validation of the detection of m^6A sites in viral RNAs from HIV-1-infected cells, replicates 1 and 2 of
146 SupT1 cells infected for 20 h (SupT1_20h_pi rep 1 and rep2) were treated with the ALKBH5 demethylase,
147 sequenced and the results compared with those of the untreated samples. Remarkably, all 14 sites had
148 significantly lower methylation rates (Table S5). Overall, these results, obtained with synthetic oligonucleotides
149 and viral mRNA from HIV-1 infected cells, highlight the quality of our detection method and confirm that the 14
150 m^6A modification sites identified in HIV-1 transcripts are *bona fide* m^6A modification sites.

151 **Detection of m^6A sites in HIV-1 splicing isoforms**

152 We then investigated the distribution of m^6A sites in each of the different splicing isoforms. The 14 m^6A sites
153 found in the pool of all transcripts were also detected in each of the 51 isoforms identified in the six samples
154 (Figure 3A, Table S6, Figures S2A-F, S3A-E). The methylation rate of these sites was similar to that for all
155 transcripts, characterized by cluster 1 centered around the two highly methylated A8532 and A8563 sites and
156 cluster 2 centered around the highly methylated A9428. An unsupervised clustering of viral transcripts based
157 on the methylation rates observed at the 14 m^6A sites showed that the isoforms clustered into three groups.
158 The first group corresponded to the US transcripts. The second group consisted mostly of PS transcripts and
159 the third group consisted of the CS transcripts (Figure 3A, Figure S3A-E). Methylation rates differed
160 significantly between these classes and were highest for CS and lowest for US (CS rates > PS rates > US
161 rates, Figure 3B, Figure S4A-E). These results are consistent with previous observations showing that
162 methylation levels are higher in CS than in PS transcripts, and higher in PS than in US transcripts (17, 19, 26).
163

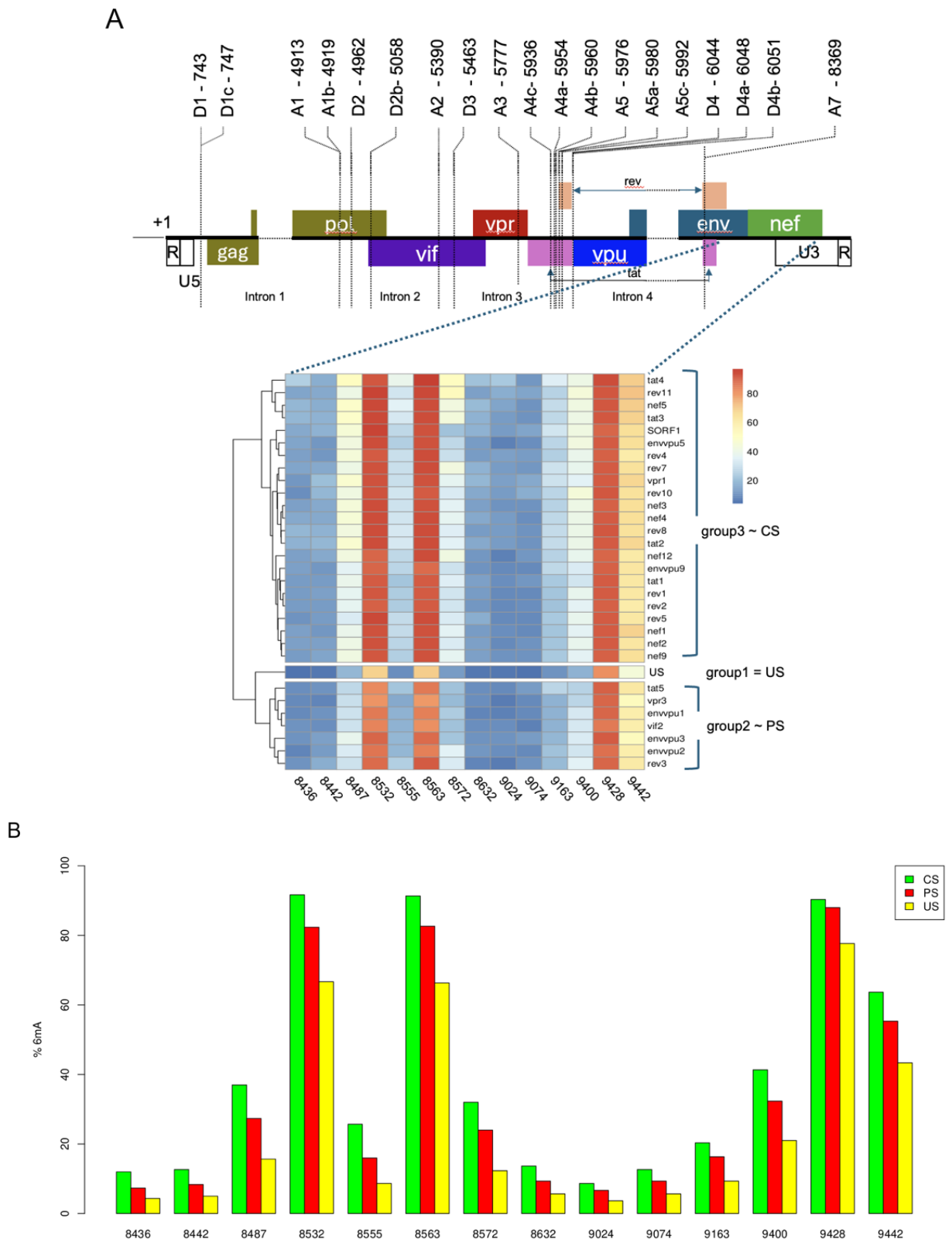


Figure 3. *Methylation rates for the 14 m^{6}A sites detected in SupT1_20h_pi rep1.* ***(A)*** *Unsupervised heatmap of the methylation rates of HIV-1 transcripts (Materials and Methods). Isoforms are clustered based on the*

169 *methylation rates of the 14 m^{6}A sites. The clusters essentially correspond to US, PS and CS transcripts.* ***(B)***
170 *Methylation rates of m^{6}A sites in the CS, PS and US transcripts of the SupT1_20h_pi rep1 sample.*
171
172 In parallel to the 14 sites common to all viral mRNA isoforms, we detected four additional m^{6}A sites specific to
173 only a few isoforms. Three of these sites, A5613, A5654 and A5703, were detected in the Vif and Vpr coding
174 regions (between the D3 and A3 splice sites) and are specific to the isoforms encoding Vif and Vpr. The fourth
175 site, A5887, was detected in the Tat coding region (between A3 and A4c) and is specific to the Vif, Vpr, Tat and
176 SORF1 isoforms (Figure 4A, Table 2, Table S7). The size of the corresponding RNAs depends strongly on the
177 absence (CS, -i4) or presence (PS, +i4) of intron 4 (i4: D4A7, 2325 nt-long). As for the global trend observed
178 for the 14 sites common to all isoforms presented in Figure 3, the methylation rates of these four additional
179 m^{6}A sites was higher for the CS (-i4) than for the PS (+i4) isoforms, but this difference was more pronounced
180 (Tables 2, S7). For example, in the SupT1_20h_pi rep1 sample, the methylation rates of the four sites in the
181 Vif1 isoform (-i4) were 3 to 9 times higher than those in the Vif2 isoform (+i4) (Table 2). Similarly, the
182 methylation rates in Vpr1 and Vpr2 were much higher than those in Vpr3 and Vpr4, respectively. There was
183 one notable exception to this tendency: the rate of A5887 methylation was significantly higher in Tat 5 (PS,
184 +i4) than in Tat1 (CS, -i4) (Figure 4B). Importantly, all four sites were unmethylated or had very low rates of
185 methylation in unspliced transcripts (≤ 1.3%). A similar trend was observed for the other five SupT1 and CD4
186 samples, for the CS and PS transcripts of the Vif, Vpr and Tat isoforms (Table S7).

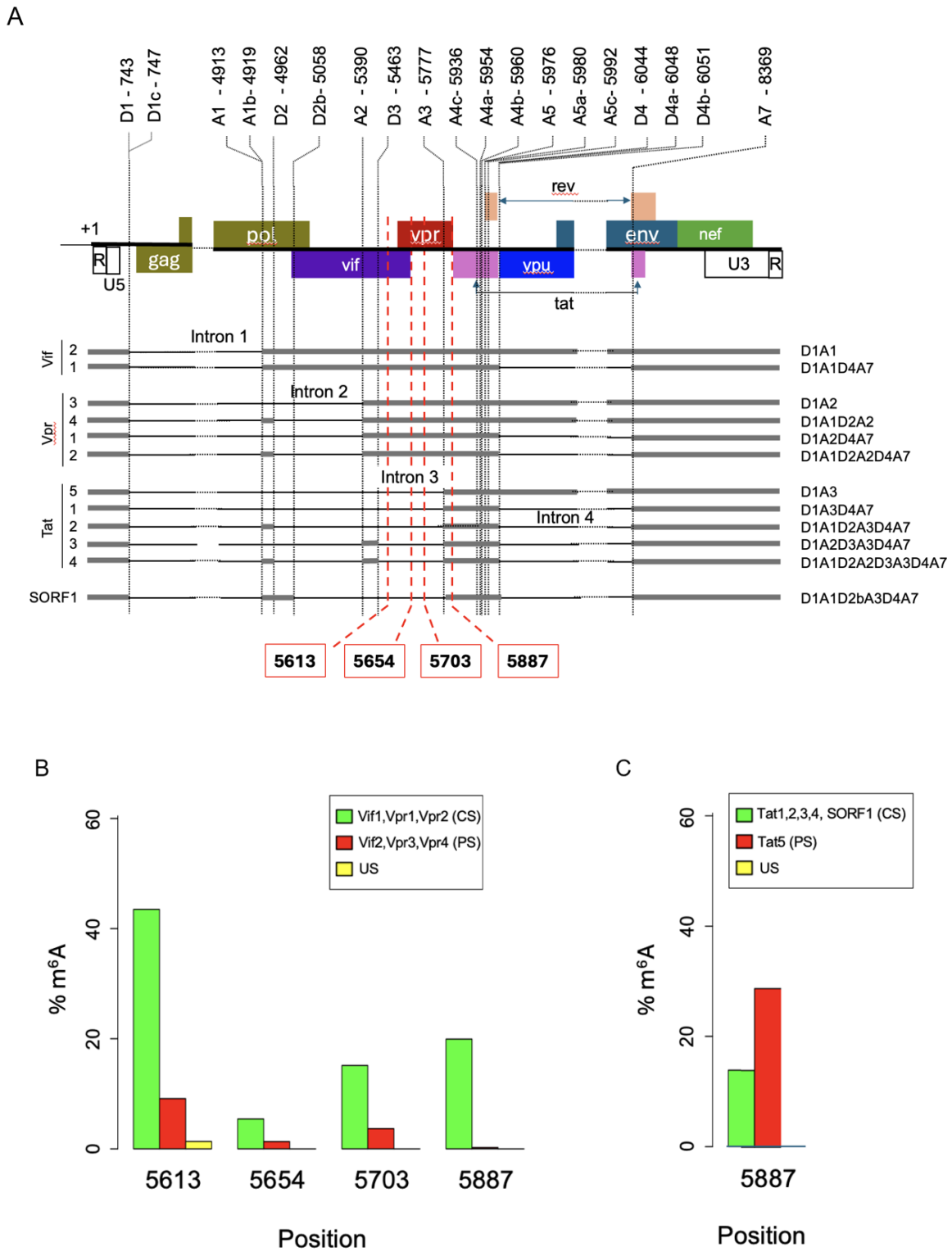


Figure 4. *Position and methylation rates of the four isoform-specific m^{6}A sites.* ***(A)*** *The A5613, A5654 and A5703 sites are common to the Vif and Vpr isoforms; A5887 is common to the Vif, Vpr, Tat and SORF1 isoforms.* ***(B)*** *Methylation rates in SupT1_20h_pi rep1, calculated for the four sites in the CS (- intron 4, Vif1, Vpr1, Vpr2), PS (+ intron 4, Vif2, Vpr3, Vpr4) and US isoforms.* ***(C)*** *Methylation rates calculated for the A5887 site for the CS (Tat1, 2, 3, 4, SORF1), PS (Tat5) and US isoforms.*

Isoform		SupT1_20h_pi rep1			
CS - i4	PS + i4	% A5613 GGACA	% A5654 AGACA	% A5703 AAACT	% A5887 AAACT
Vif1		35.2	2.4	11.4	18.6
	Vif2	9.3	0.7	3.3	2
Vpr1		58	9.3	24	25.3
	Vpr3	12.3	1	4	0.3
Vpr2		37.3	4.6	10	15.9
	Vpr4	5.7	2.3	3.7	0
Tat1					18.3
	Tat5				28.7
Tat2					12.3
Tat3					12.3
Tat4					9.5
SORF1					16.7
US		1.3	0	0	0

Table 2. *Methylation rates measured for the A5613, A5654, A5703 and A5887 m^{6}A sites in the Vif, Vpr, Tat and sORF1 isoforms in the SupT1_20h_pi rep1 sample. The first two columns correspond to the isoforms completely spliced (CS, - i4) or partially spliced (PS, + i4). The last row indicates the methylation rates in the unspliced transcripts.*

Single-molecule analysis of m^{6}A sites

We investigated the distribution of the methylated m^{6}A sites in individual transcripts by counting, for US, PS and CS transcripts, the numbers of methylated As in the eight sites of cluster 1 (from 0 to 8 methylated As) (Figure 5A) and the three sites of cluster 2 (from 0 to 3 methylated As) (Figure 5B) of the 3'UTR region of HIV-1, as defined in Figure 1. We examined the two clusters either separately (Figure 5A, 5B, Figure S6A-D) or simultaneously on each read (Figure 5C-E, Figure S7A-O, Table S8). The distribution of m^{6}A sites was similar for rep1 and rep2 samples, for the SupT1 samples at 20 h and 30 h post-infection, and also in SupT1 and primary CD4$^+$ T cells. We found that the number of sites modified per molecule increased from US to PS to CS. When considering the two clusters simultaneously, we observed that the peaks for the US transcripts corresponded to two methylated sites in cluster 1 and one or two methylated sites in cluster 2 (Figure 5A,B, Figure S6A-D). By contrast, for the PS and CS transcripts, the peaks corresponded to two or three methylated

sites in cluster 1 and two methylated sites in cluster 2. In the six samples, a small proportion (3.7% to 5.4%) of US transcripts were totally unmethylated simultaneously in the two clusters. The proportion of unmethylated transcripts decreased to 1.5% for PS transcripts and 0.3% for CS transcripts (Table S8).

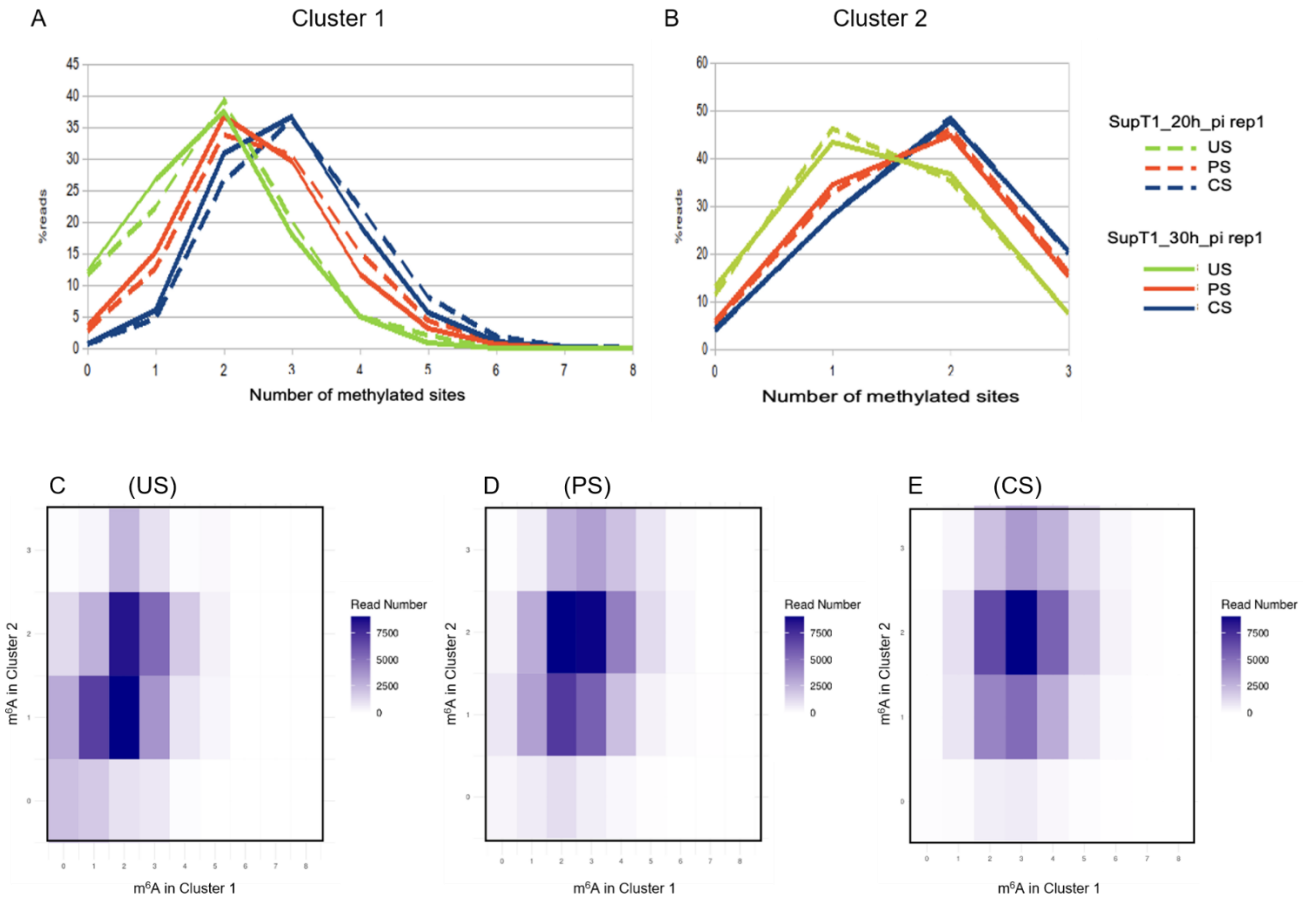


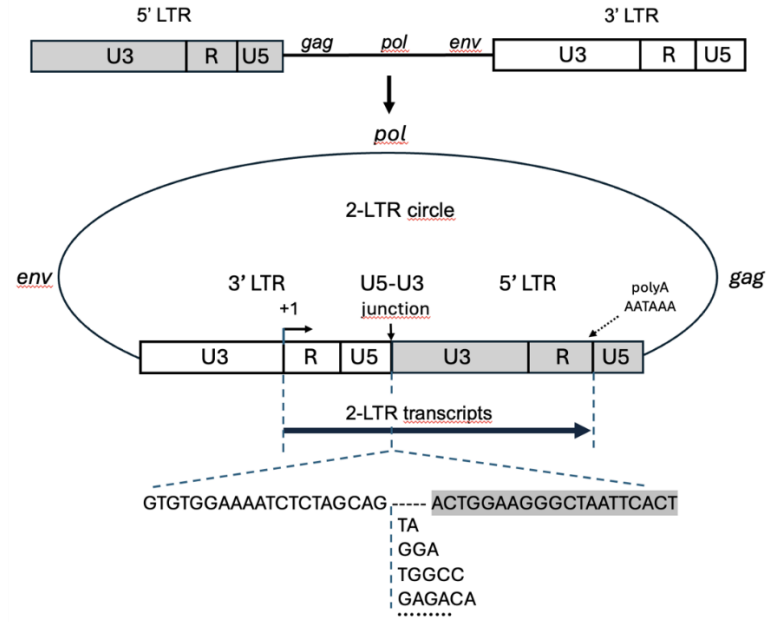
Figure 5. *Single-molecule analysis of methylated m^{6}A sites.* ***(A)*** *Cluster 1.* ***(B)*** *Cluster 2.* ***(C-E)*** *Correlation between the number of methylated sites in cluster 1 and in cluster 2 of the same transcript from the SupT1_20h_pi rep1 sample. The intensity of the color on the heatmaps indicates the number of reads showing the indicated numbers of methylated As in cluster 1 (0 to 8) and in cluster 2 (0 to 3) (Materials and Methods).*

Detection and analysis of 2-LTR transcripts

During the analysis of the splicing isoforms, we identified a class of RNA transcripts mapping to the long terminal repeats (LTR, constituted by the U3-R-U5 segments) located at either end of the viral genome. They result from the transcription of 2-LTR circles, a non-integrated circular form of the reverse-transcribed viral DNA formed early after infection during the conversion of the viral genome into DNA by the junction of the two LTRs located at the ends of the genome (36) (Figure 6A, 6B). They begin at the promoter at the 5' end of the R

235 segment of the 3' LTR and terminate at the polyA site at the 3' end of the R segment of the 5' LTR. They create
236 a precise join between the 3' end of the RU5 segment of the 3' LTR and the 5' end of the U3R segment of the
237 5' LTR. Studies of the 2-LTR U5-U3 junction (the junction between the two LTRs) have shown this junction to
238 be highly heterogeneous in infected cells and between infected individuals, frequently due to insertions (37,
239 38). Transcripts encompassing this junction have been detected by RT-PCR, but were not precisely identified.
240 A recent study of the HIV-1 transcriptome by Nanopore cDNA sequencing identified RNA species that could
241 correspond to 2-LTR transcripts and were detected in infected $CD4^+$ T and HeLa cells (31). Here, the 2-LTR
242 transcripts were detected in both infected SupT1 and infected primary $CD4^+$ T cells (0.16% to 1.69 % of HIV-1
243 transcripts, Table S2B). They displayed either the precise U5-U3 junction or a number of variants mostly
244 carrying short (1-nt to 10-nt) insertions. The detection of m^6A sites in these 2-LTR transcripts revealed the
245 presence of six sites that were detected in all six samples (Figure 6B, 6D, Table S9). The first site, 2L125, is
246 homologous to the A578 position of the viral genome (Figure 6C) common to all HIV-1 transcripts.
247 Remarkably, this 2L125 site was methylated in 24 to 38% of 2-LTR transcripts (Table S9) but was not detected
248 as methylated in any other HIV-1 transcript, in all six samples. The other five sites are located in the U3
249 segment of the 5' LTR and are homologous to m^6A sites A9074, A9163, A9400, A9428 and A9442, which are
250 common to all splicing isoforms (Fig. 6B-D). Remarkably, in the 2-LTR transcripts, the methylation rates of all
251 five of these sites were higher than those of the homologous sites in the viral genome (Table S9). For
252 example, the 2L537 site attained methylation rates of 91.3% and 94.1% in SupT1 and $CD4^+$ T-cell samples,
253 respectively (Table S9), whereas the homologous site A9428 had methylation rates of 84.3% and 79.7% in the
254 corresponding samples (Table S4).
255

A



B

2-LTR transcript: R-U5-U3-R

```
1
R-->
GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTT

           <--R U5-->                  2L125(578)
GCCTTGAGTGCTCAA AGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGT

           <--U5||U3-->
                 2L183(9074*)
GGAAAATCTCTAGCAG  ACTGGAAGGGCTAATTCACTCCCAAAGAAGACAAGATATCCTTGATCTGTGGATCTACCACACACAA

                  2L272(9163)
GGCTACTTCCCTGATTGGCAGAACTACACACCAGGGCCAGGGGTCAGATATCCACTGACCTTTGGATGGTGCTACAAGCTAGT


ACCAGTTGAGCCAGATAAGGTAGAAGAGGCCAATAAAGGAGAGAACACCAGCTTGTTACACCCTGTGAGCCTGCATGGAATGG


ATGACCCTGAGAGAGAAGTGTTAGAGTGGAGGTTTGACAGCCGCCTAGCATTTCATCACGTGGCCCGAGAGCTGCATCCGGAG

       2L509(9400)                2L537(9428)   2L551(9442)
TACTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGGGAGGCGTGGCCTGGGCGGGAC

                                                  <--U3 R-->
TGGGGAGTGGCGAGCCCTCAGATGCTGCATATAAGCAGCTGCTTTTTGCCTGTACT GGGTCTCTCTGGTTAGACCAGATCTGA

                                                                      R<-- -->polyA
GCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCAA AAAAAAA...
```

C

Position in 2-LTR transcript (736 nt)	Homologous position in HIV-1 genome
2L125	A578
2L183	A9074
2L272	A9163
2L509	A9400
2L537	A9428
2L551	A9442

D

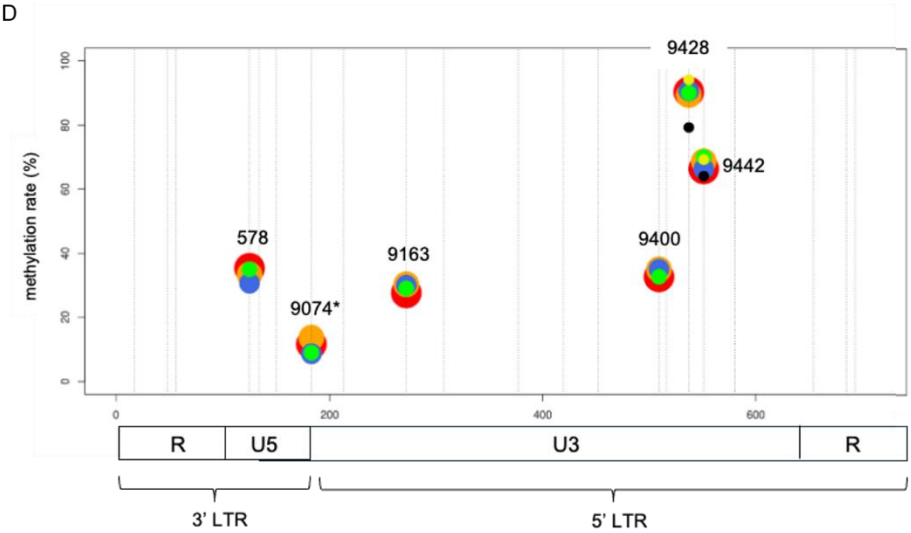


Figure 6. *Transcription and methylation of the 2-LTR transcripts.* ***(A)*** *Diagram of the 2-LTR circles and of the transcripts joining the RU5 segment of the 3' LTR to the U3R segment of the 5' LTR. Several examples of insertion variants at the U5-U3 junction are shown.* ***(B)*** *Sequence of the 2-LTR transcripts. The successive transcribed regions — R, U5, U3, R — are indicated; m^{6}A sites are shown in bold typeface and underlined and are named according to the position of the methylated A in the 2-LTR transcript; in brackets, the homologous position in the HIV-1 reference genome; for example, for 2L125(578), 125 is the position of the m^{6}A in the 2-LTR transcript and (578) is the homologous position in the HIV-1 reference genome; * indicates that the sequence context of 2L183 in the 2-LTR transcript (AGACT, see B) differs from that at the homologous position 9074 in the reference genome (GGACT).* ***(C)*** *Correspondence table for a position in the 2-LTR transcript and the homologous position in the reference genome.* ***(D)*** *Sites detected in the 2-LTR transcripts of infected SupT1 and CD4$^+$ T-cell samples (Materials and Methods). Large red circles, SupT1_20h_pi rep1; orange circles, SupT1_30h_pi rep1; blue circles, SupT1_20h_pi rep2; green circles, SupT1_30h_pi rep2; yellow circles, CD4_48h_pi; black circles, CD4_72h_pi. Vertical lines indicate the positions of the As in all the DRACH motifs (21) of the transcript.*

DISCUSSION

We report here the detection of m^{6}A nucleotides in HIV-1 transcripts from infected human cells, with comparison to previous studies (15-19, 26). We included cells of two different origins: CD4$^+$ T cells from the SupT1 cell line and CD4$^+$ T cells from the peripheral blood mononuclear cells (PBMC) of blood donors. These cells, which are natural HIV-1 targets, were infected for different periods of time (20 h and 30 h for SupT1

284 cells; 48 h and 72 h for blood donor cells). Using the new RNA004 chemistry and a new base-calling model,
285 we detected 14 m^6A sites common to the unspliced HIV-1 RNA and to all the splicing isoforms encoding viral
286 proteins.
287
288 The m^6A positions were clustered in two regions of the 3' end of the HIV-1 genome, both having a methylation
289 profile featuring highly methylated central sites. These sites are partly included in the methylated regions
290 identified in previous studies based on the immunoprecipitation of m^6A-containing RNA (reviewed in (9)) or on
291 m^6A-SAC-seq (39). Unlike these studies, we found no methylation sites in the 5' region of the reference HIV-1
292 genome, either in the unspliced RNA or in the diverse splicing isoforms. Overall, we observed a remarkable
293 specificity of the 14 m^6A sites in the 245 DRACH sites of the HIV-1 genome. The saturation of the five DRACH
294 sites located in the 200 nt cluster 1 and the additional methylation of three single-nucleotide DRACH variants
295 in the same region further increase the specificity of HIV-1 methylation. However, despite the proposal of
296 hypotheses involving the recruitment of m^6A writers by RNA-binding proteins or by RNA polymerase II close to
297 DRACH sites during transcription (7, 40), there is currently no explanation for the specificity of the methylation
298 of some DRACH sites but not others. Interestingly, on examination of the secondary structure of the HIV-1
299 RNA, most of the 14 m^6A sites were found to be located in single-stranded portions of the structure, or within
300 the final nucleotides at the 3' end of a helix (Figure S8). This finding is fully consistent with studies showing
301 that RRACH sequences located in a double-stranded RNA duplex prevent efficient m^6A methylation by the
302 METTL3–METTL14 complex (41). The 14 m^6A sites include three sites (A8532, A9428, A9442) previously
303 detected as peaks in RNA immunoprecipitation experiments (16-18) and as methylated nucleotides (m^6As) by
304 direct RNA sequencing (26). This last study also identified A8563 as a weakly methylated m^6A site, contrasting
305 with our results, as A8563 was one of the highly methylated sites of cluster 1 (Figure 1). We found that the 14
306 m^6A sites had higher methylation rates in CS isoforms than in PS isoforms and US transcripts (CS rates > PS
307 rates > US rates, Figure 3B, Figure S4A-E). At single-molecule level, we observed that most CS transcripts
308 had three methylated sites in cluster 1, and two in cluster 2. By contrast, most US transcripts had two
309 methylated sites in cluster 1, and one in cluster 2. In addition, only very small proportions of CS transcripts had
310 no methylated site in either cluster (0.3%). This proportion increased only slightly, from 3.7% to 5.4%, in US
311 transcripts. These observations are consistent with a previous study (26) of three methylation sites (A8532,
312 A9428, A9442) reporting that 2.5% and 8.5% of CS and US transcripts, respectively, were unmethylated.
313 Overall, these results are consistent with previous observations that reducing m^6A methylation inhibits splicing

314 and stimulates genomic RNA packaging (17, 19, 26). They also strongly suggest that the methylation of these
315 sites plays an important role in splicing regulation and, ultimately, in virus infectivity.
316
317 Four additional sites were detected specifically in several isoforms (Figure 4A, Tables 2, S7). The methylation
318 rates of these additional sites were higher for CS than for PS isoforms, but the ratio of CS transcript
319 methylation to PS transcript methylation was much higher than for the 14 sites common to all isoforms (Figure
320 4B and Figure 3B). This finding provides additional support for a role for methylation in splicing regulation. It
321 also suggests that m^6A sites located at different positions within a specific isoform may be involved in different
322 processes, such as translation or the maintenance of stability. Further experiments involving the mutation of
323 one or several of these sites would help to address these questions.
324
325 In addition to the splicing isoforms, we detected ≈736 nt-long RNAs resulting from the transcription of 2-LTR
326 unintegrated DNA circles joining the R segments of the two LTRs (RU5-U3R, Figure 6A). These transcripts
327 contained six m^6A sites, five of which were common to all spliced isoforms and had higher methylation rates
328 than any of the other transcripts. An additional site, 2L125, was identified in the U5 segment, with a
329 methylation rate ranging from 24% to 35%. We observed that (i) positions homologous to this position (2L125)
330 in the other transcripts were not considered methylated by our detection process and (ii) the other five sites
331 had methylation rates higher than that at the homologous sites in all the other transcripts. These results
332 strongly suggest that, despite not encoding canonical viral proteins, 2-LTR transcripts play a functional role in
333 the viral life cycle. Consistent with this idea, 2-LTR circles were initially thought to be dead-end products in the
334 viral life cycle (42), but many studies subsequently showed that they could be transcribed and might constitute
335 a reserve supply of HIV-1 genomes for proviral integration (43-48). In addition, the large difference in 2L125
336 methylation rate between 2-LTR transcripts and the other transcripts (same sequence context) confirms that
337 this site is not an artifact of the detection process but, indeed, a *bona fide* methylation site.
338
339 It is generally though that methylation occurs cotranscriptionally and that exclusion of m^6A from splice-site
340 proximal regions by the exon junction complex results in low methylation rates close to splice junctions (49,
341 50). This hypothesis is not consistent with our observation that CS transcripts have higher methylation levels

342 than PS and US transcripts. It has recently been shown that m^6A methylation can occur posttranscriptionally,
343 especially in transcripts that are highly methylated (51). This raises the possibility that the methylated sites
344 observed at highly specific positions within HIV-1 transcripts are methylated posttranscriptionally and are
345 potentially associated with an epitranscriptomic role in the virus life cycle.
346
347 In conclusion, we provide an unbiased assessment of m^6A modifications in HIV-1 transcripts isolated from
348 infected $CD4^+$ T cells. We reveal the existence of novel m^6A sites in the 3' end common to all transcripts, in
349 specific isoforms, and in 2-LTR transcripts. This study also highlights the complex interplay between m^6A
350 modifications, splicing events and the fate of HIV-1 transcripts. These results pave the way for efforts to
351 decipher the role of m^6A modifications in the HIV life cycle.
352
353 **MATERIALS AND METHODS**
354
355 **Cell infection**
356 VSV-G pseudotyped $HIV_{NL4-3\ XCS}$ was produced by transfection of 293T cells using CAPHOS kit (Sigma-
357 Aldrich) as in (52). Viruses were pseudotyped to favor viral entry and to achieve a high infection rate within a
358 short time. SupT1 $CD4^+$ T cells were cultured with RPMI GlutaMax 1640 (Gibco) complemented with 10% FBS
359 (Dutscher) and 1% penicillin/streptomycin.
360 2 x 10^8 SupT1 $CD4^+$ T cells were infected with 800 ng of VSV-G-$HIV_{NL4-3\ XCS}$ HIV-Gagp24 for 3h in IMDM plus
361 10mM HEPES (Gibco/ThermoFisher Scientific) supplemented with 2 µg/ml of DEAE-dextran (Sigma). Two
362 biological replicates were performed corresponding to 100 million cells infected at time 0. Eighteen hours post-
363 infection a fraction of the cells (2 x 10^5) was harvested, stained with CD4-Vio770 antibody (Miltenyi) and a
364 Live/Dead (LVD) fixable violet dye (Invitrogen), fixed with 4% paraformaldehyde (ChemCruz) and
365 permeabilized (PBS1X, BSA 0,5%, saponin 0,05%). Cells were then stained with an HIV-1 Gag-p24 specific
366 antibody (KC57-PE, Beckman Coulter). Sample acquisition and analysis were performed using a BD
367 LSRFortessa and FlowJo v10.8 Software, respectively (both from BD Life Sciences). Mock-infected cells were
368 used as negative control. Twenty and thirty hours post-infection, cells were harvested, washed in cold PBS
369 and lysed at 4°C in RLT (from RNeasy Mini Kit (Qiagen) buffer containing ß-mercaptoethanol and 1% NP40.
370 PBMCs from anonymous blood donors (n=2, donor 864 and 880) were purchased from the blood bank at EFS
371 (Établissement Français du sang) under the agreement number 15/EFS/022. PBMCs were isolated using a
372 Ficoll gradient and monocytes depleted by overnight incubation in lying cell culture flasks containing

373 RPMI+10%FBS supplemented with 10 IU/ml of recombinant human IL-2 (rhIL-2, Miltenyi Biotec). $CD4^+$ T cells
374 were then isolated using $CD4^+$ magnetic beads (Miltenyi Biotec) and activated with Phytohemagglutinin (PHA,
375 1 µg/ml, PAA Laboratories) in RPMI GlutaMax 1640 complemented with 10% FBS, 1% penicillin/streptomycin,
376 nonessential amino acids and sodium pyruvate (all from Gibco/ThermoFisher Scientific), so called T-cell
377 medium. 48 h post-activation, $CD4^+$ T cells were harvested and resuspended in T-cell medium containing 100
378 IU/ml of recombinant human IL-2 (rhIL-2, Miltenyi Biotec). 7-days post activation, 300 x 10^6 and 130 x 10^6
379 $CD4^+$ T cells, were infected (for donor 864 and donor 880, respectively) at a multiplicity of infection (moi) 0,01
380 with VSV-G-$HIV_{NL4-3\ XCS}$ (representing around 1000 ng of HIVGagp24 / 10^7 cells) for 3h in IMDM plus 10mM
381 HEPES supplemented with 2 µg/ml of DEAE-dextran. $CD4^+$ T cells were then maintained in T-cell medium
382 containing 100 IU/ml of rhIL-2 and HIV infection was monitored using flow cytometry, as for SupT1 $CD4^+$ T
383 cells. For donor 864 and donor 880, cells were harvested and lysed 48 h and 72 h post-infection, respectively.
384

385 **In vitro transcription**

386 HIV-1 genomic RNA was transcribed in two parts, from the beginning of 5'LTR to the end of Pol (5124 bases)
387 and from the end of Pol to the end of 3'LTR (4829 bases), with 225 overlapping bases. Each part of the
388 genome was PCR amplified from pNL4.3XCS (53) and cloned in a pUC19 vector containing T7 promoter,
389 using NEBuilder cloning kit (New England Biolabs). The plasmids were then linearized with SmaI restriction
390 enzyme during 3 hours at 30°C, to perform run-off transcription. In vitro transcription was performed using
391 RiboMAX T7 Large Scale RNA Production System (Promega) with incubation 4 hours at 37°C , followed by
392 DNAseI treatment during 15 min and RNA column clean-up (ZymoResearch).
393

394 **RNA purification**

395 Total RNA was extracted from the cell pellets using RNeasy Mini Kit (Qiagen), according to the manufacturer's
396 protocol. Briefly, cells were lysed in RLT buffer containing b-mercaptoethanol and homogenized using
397 QIAshredder homogenizer. The lysate was then purified on RNeasy spin columns and RNA was eluted in
398 RNase-free water. PolyA RNA were then purified from the total RNA using two rounds on $Oligo(dT)_{25}$
399 Dynabeads (Thermo Fisher Scientific). In vitro transcripts were polyadenylated using E. coli PolyA polymerase
400 kit from NEB and purified on $Oligo(dT)_{25}$ Dynabeads.
401

402 **Library preparation and sequencing**

403 The polyA RNA was used for ONT direct RNA sequencing using the RNA004 kit. The reverse transcription
404 (RT) step was modified as described by (26); a pool of 111 oligonucleotides complementary to the HIV-1
405 genome was added during library preparation in an estimated 30-fold molar excess to the viral RNA, which
406 resulted in a much more efficient reverse transcription process.
407 Sequencing was processed on P2solo through the GridION (MinKNOW version: 24.02.10 for SupT1_20h_pi
408 rep1, 24.02.16 for SupT1_30h_pi rep1 and IVT, 24.06.15 for SupT1_20h_pi+demeth rep1 and rep2,
409 SupT1_20h_pi rep2, SupT1_30h_pi rep2, CD4_48h_pi and CD4_72h_pi, IVT). After sequencing, base-calling
410 has been done with dorado-0.7.2 and rna004_130bps_sup@v5.0.0 model with –modified-bases m^6A
411 parameter (for all samples).
412
413 **Identification of HIV-1 splicing isoforms**
414 Pass reads were mapped with minimap2 (v2.17-r941) (with -x splice and –secondary=no options) on the
415 PNL4-3XCS HIV-1 genome. BamToBed tools with -splice option have been used to select full-length
416 transcripts using a homemade script python. With this script, each read is analyzed and its splice junctions
417 must match (with a tolerance of one base) the splice junctions of a specific isoform. In each full-length read we
418 identified the donor (D) and acceptor (A) sites that match the already known HIV-1 splice sites (eight donor
419 and eleven acceptor sites, Nguyen Quang et al., Emery et al., Ocwieja et al., O'Reilly et al., Widera et al.)
420 (Table S2A). Some rare isoforms were detected with only few reads and only those detected in at least five of
421 the six samples were considered. Finally, 51 isoforms were identified by their succession of splice junctions
422 and were named as in (Nguyen Quang et al. 2020) (Table S2B). Between 4.4% (CD4_48h_pi) and 8.0%
423 (SupT1_20h_pi rep1) of full-length reads did not fully match these sites and were not considered.
424
425 **Detection of m^6A sites, computation of methylation rates**
426 Nucleotide position of m^6A's was obtained thanks to the specific base-calling process that can detect m^6A's
427 (version 5.1.0). For each sample, the methylation rate is computed by modkit pileup with –filter-threshold
428 A:0.8, --mod-thresholds a:0.9 and –edge-filter 20. The nucleotides that are called as m^6A are counted at each
429 position in a set of reads to determine the corresponding modification rate. The methylation rates computed at
430 each position of the transcripts in a sample were compared to those computed at the corresponding positions
431 of the IVT sample using the DMR model and scoring details - Modkit command that computes a final
432 modification rate and a corresponding p-value at each position. The positions presenting p-values $p \leq 10^{-5}$ and

433 a number of methylated reads ≥ 30 were retained as putative m^6A sites. For the clustering analyses in Figure
434 3 and Figure S3, only the isoforms identified by more than 100 full-length reads were retained.
435
436 **Single molecule analysis of methylated m^6A sites**
437 Modkit extract enables access to read-level base modification calls. Parameters –mapped-only and –include-
438 bed were used (bed file containing eight positions from cluster1 and three positions from cluster2). The output
439 tsv files were parsed with the R language. During this analysis, the threshold of probability of modification to
440 call 6mA was 0.9 (as with modkit pileup).
441
442 **Synthetic oligonucleotides**
443 We used two synthetic RNA oligonucleotides identical to HIV-1 regions with methylated As and the same two
444 oligonucleotides differing only at extremities, without methylated As.
445 Oli2-mod harbors m^6A at positions 9400, 9428 et 9432:
446 CTTCAAGA(m^6A)CTGCTGACATCGAGCTTGCTACAAGGG(m^6A)CTTTCCGCTGGGG(m^6A)CTTTCCAGG
447 (60 nt-long).
448 Oli2-nmod harbors non-modified A's and 6nt tag at 3' end (**TCATAG**):
449 CTTCAAGAACTGCTGACATCGAGCTTGCTACAAGGGACTTTCCGCTGGGGACTTTCCAGG**TCATAG** (66 nt-
450 long).
451 Oli1-mod harbors m^6A at positions 8532 et 8563 + 5nt tag at 3' end (**TATGC)**:
452 CCGCTTGAGAG(m^6A)CTTACTCTTGATTGTAACGAGGATTGTGGA(m^6A)CTTCTGGGAC**TATGC** (58 nt-
453 long).
454 Oli1-nmod harbors non-modified A's and 5nt tag at 3' end (**GTACG**):
455 CCGCTTGAGAGACTTACTCTTGATTGTAACGAGGATTGTGGAACTTCTGGGAC**GTACG** (58 nt-long).
456
457 **Oligonucleotides demethylation and sequencing**
458 As recommended by the manufacturer, less than 3 µM of m^6A was incubated with 1 µM ALKBH5 at room
459 temperature. The composition of the reaction mixture was as follows: 3 µL enzyme (0.9 µg, 19.8 pmol), a mix
460 of four oligos at 0.13 µM each (oli1-mod, oli1-nmod, oli2-mod and oli2-nmod). A total of (3 x 0.13) + (2 x 0.13)
461 = 0.65 µM m^6A was used in this step; the buffer contained 100 mM KCl, 2 mM MgCl2, 2 mM L-ascorbic acid,
462 300 µM alpha-ketoglutarate, 150 µM (NH4)2Fe((SO4)2 and 50 mM Hepes buffer (pH 6.5), in a total volume of

463 20 µL. The reactions were stopped after 1.5 h incubation by adding EDTA to a final concentration of 5 mM; the
464 oligos were purified using SPRIselect beads following a modified protocol as follows. Water was added to a 25
465 µL total volume, followed by the addition of 45 µL SPRI beads and mixing. After 5 minutes of room
466 temperature incubation, the beads were magnetized and the supernatant containing the small RNA was
467 transferred to a new tube containing 45 µL SPRI beads and 135 µL isopropanol, and mixed. After another 5
468 minutes of room temperature incubation, the beads were magnetized and the supernatant was discarded. The
469 beads were washed twice with 86% ethanol, allowed to dry for 5 minutes and resuspended in 10 µL water. The
470 purified oligos were then used for ONT direct RNA004 library preparation and sequenced on two MinION flow
471 cells (MinKNOW version: 24.06.14). Distinction between oligos has been made by mapping to the reference
472 with these four oligos added by eight As in 3'.
473

474 **ACKNOWLEDGMENTS**

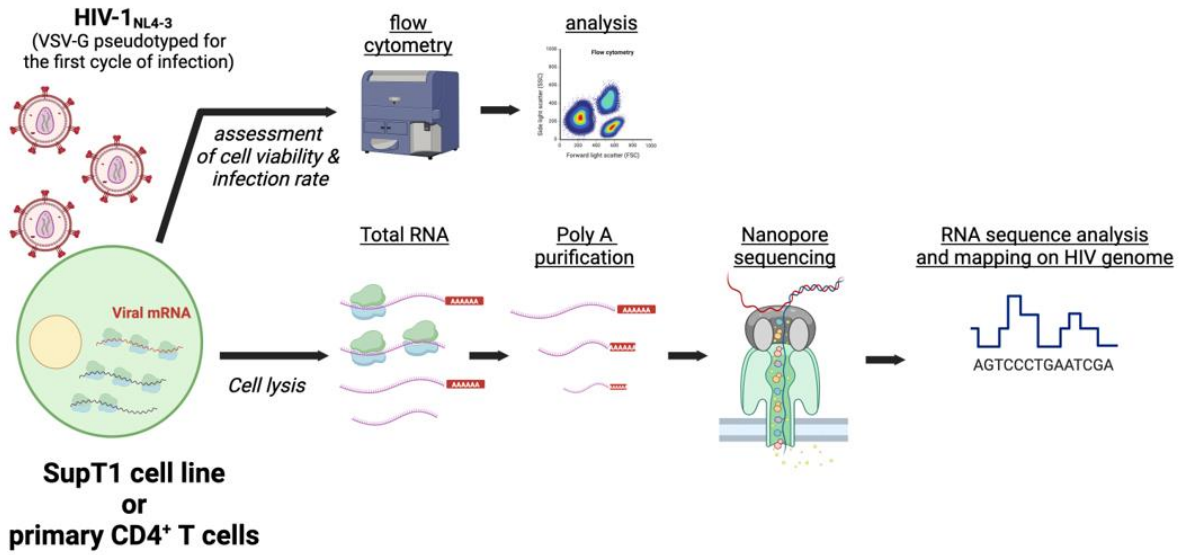
475 This work was granted by the Agence Nationale de Recherche sur le SIDA (ANRS-Maladies infectieuses
476 émergentes) for fundings to A.M. and O.N., S.B salary was supported by ANRS. L.B PhD salary was
477 supported by ANRS-MIE. This study was initiated thanks to fundings from the cross-disciplinary thematic axes
478 of I2BC through the epiRNA programs. The HTS facility of the I2BC acknowledges the support of France
479 Génomique (funded by the French National Program "Investissement d'Avenir" ANR-10-INBS-09).

480

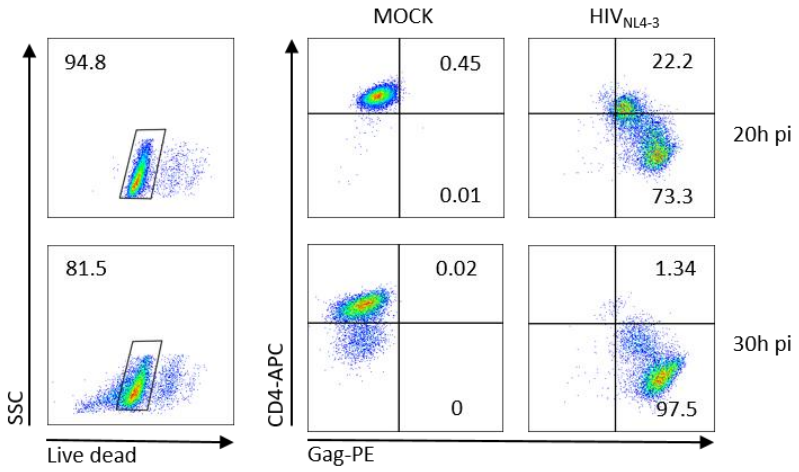
481 SUPPLEMENTARY FIGURES

482

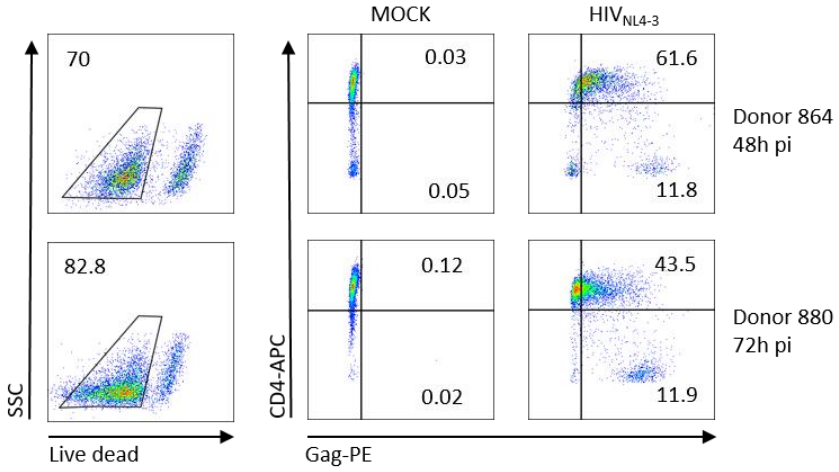
A



B



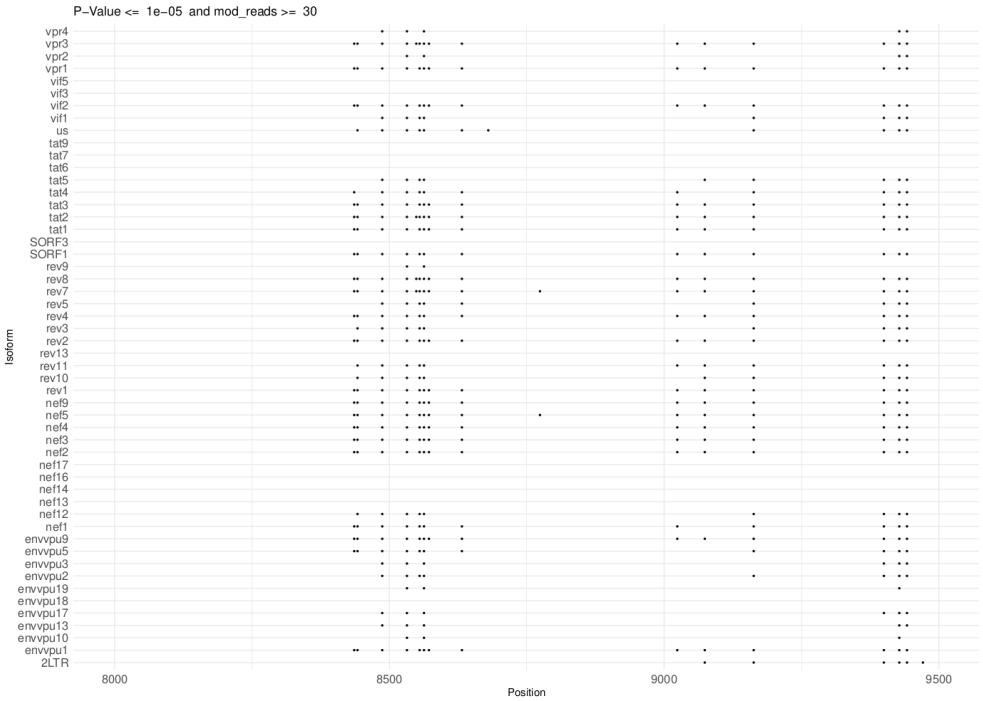
C



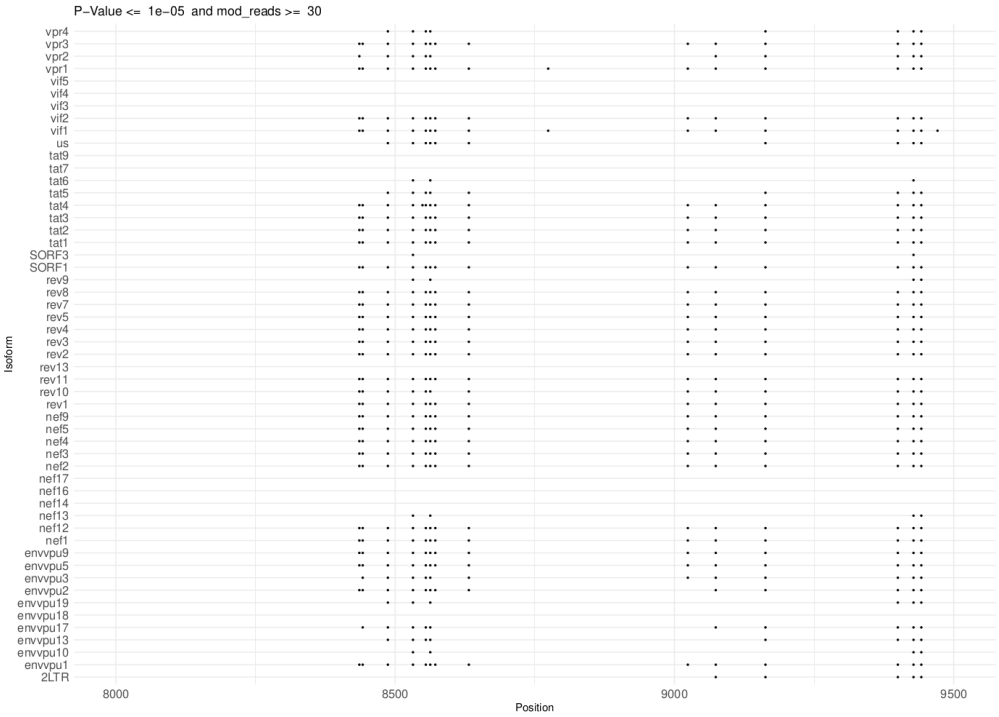
483

484 **Figure S1. (A) Outline of the infection and Nanopore sequencing protocols**. SupT1 T cells or primary
485 $CD4^+$ T cells were infected with HIV_{NL4-3} pseudotyped with VSV-G to enhance viral infection. At 20 or 30 hours
486 post-infection (pi) for HIV $_{NL4-3}$-infected SupT1 cells and at 48 h or 72 h pi for HIV $_{NL4-3}$-infected primary $CD4^+$ T
487 cells, we collected a fraction of the cell culture. The viability and percentages of HIV-Gag^+ and $CD4^+$ cells were
488 assessed by flow cytometry. During HIV infection, cells start expressing HIV-Gag protein and progressively
489 lose their membrane expression of CD4 due to expression of the HIV accessory proteins Nef and Vpu, and
490 HIV-Env. The cells were harvested and polyA RNA was purified and sequenced with Nanopore technology.
491 Reads were mapped onto the HIV_{NL4-3} genome sequence. Created with BioRender. Moris, A. (2025)
492 https://BioRender.com/o40u811 **(B) Analysis of the infection rate of HIV_{NL4-3}-infected $CD4^+$ SupT1 cells.**
493 After 20 h or 30 h, cells infected with HIV_{NL4-3} or mock-treated were labeled with a viability dye and stained
494 with antibodies specific for CD4 and HIV-Gag. Left panels: the percentage viable cells. Middle and right
495 panels: percentages of $CD4^-$HIV-Gag^+ and $CD4^+$HIV-Gag^+ cells among mock-treated and HIV_{NL4-3}-infected
496 $CD4^+$ SupT1 cells, respectively. **(C) Analysis of the infection rate of HIV_{NL4-3}-infected primary $CD4^+$ T**
497 **cells.** After 48 h or 72 h, cells from donors 864 and 880 were harvested and stained as in (B). Left panels: the
498 percentage viable cells. Middle and right panels: percentages of $CD4^-$HIV-Gag^+ and $CD4^+$HIV-Gag^+ cells
499 among mock-treated and HIV_{NL4-3}-infected primary $CD4^+$ T cells, respectively, from donors 864 (top panels)
500 and 880 (bottom panels). Numbers indicated the percentage of cells within the quadrant. MOCK: cells mock-
501 treated without infection used as a negative control.
502

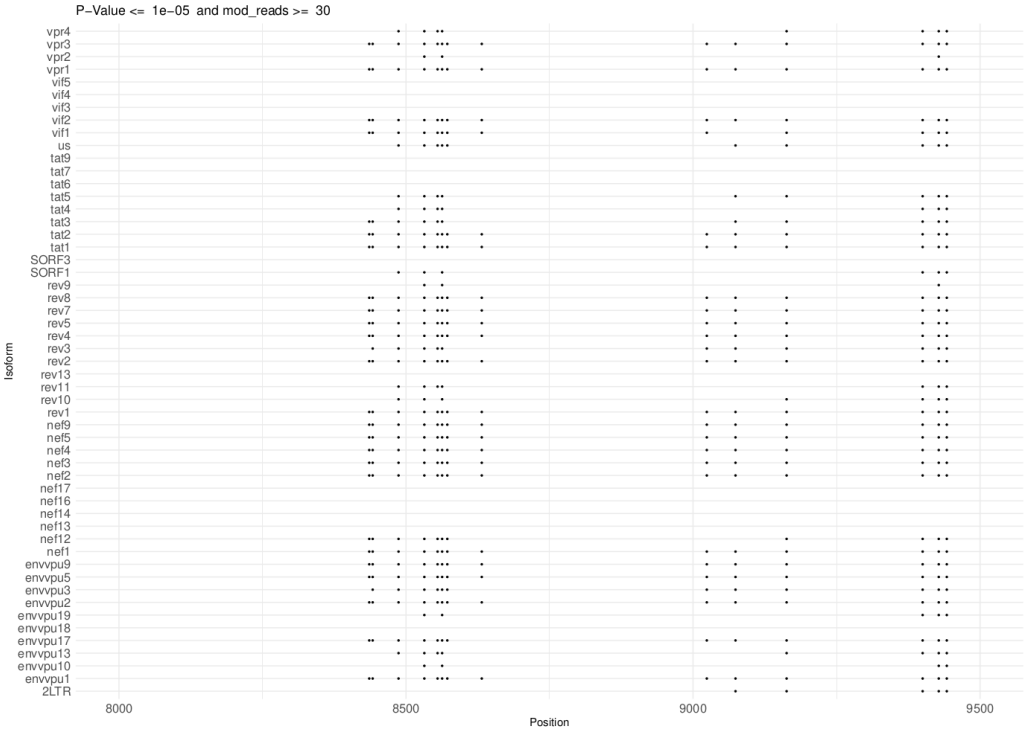
A


P−Value <= 1e−05 and mod_reads >= 30
Isoform
vpr4
vpr3
vpr2
vpr1
vif5
vif3
vif2
vif1
us
tat9
tat7
tat6
tat5
tat4
tat3
tat2
tat1
SORF3
SORF1
rev9
rev8
rev7
rev5
rev4
rev3
rev2
rev13
rev11
rev10
rev1
nef9
nef5
nef4
nef3
nef2
nef17
nef16
nef14
nef13
nef12
nef1
envvpu9
envvpu5
envvpu3
envvpu2
envvpu19
envvpu18
envvpu17
envvpu13
envvpu10
envvpu1
2LTR
8000
8500
9000
9500
Position

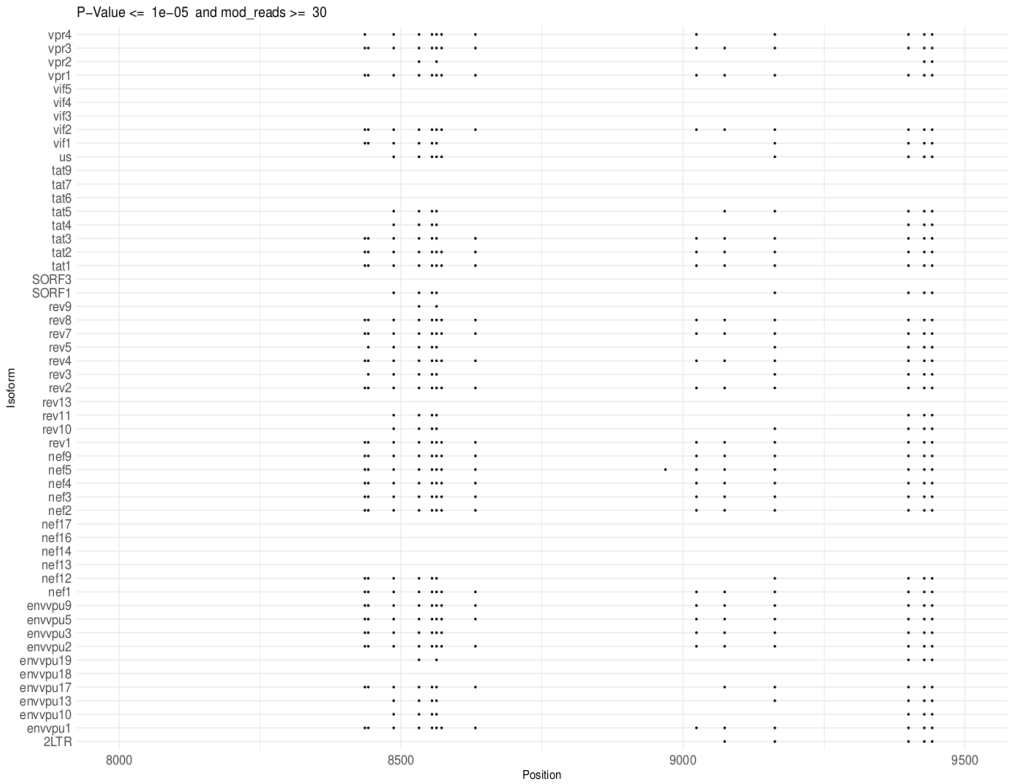
B


P−Value <= 1e−05 and mod_reads >= 30
Isoform
vpr4
vpr3
vpr2
vpr1
vif5
vif4
vif3
vif2
vif1
us
tat9
tat7
tat6
tat5
tat4
tat3
tat2
tat1
SORF3
SORF1
rev9
rev8
rev7
rev5
rev4
rev3
rev2
rev13
rev11
rev10
rev1
nef9
nef5
nef4
nef3
nef2
nef17
nef16
nef14
nef13
nef12
nef1
envvpu9
envvpu5
envvpu3
envvpu2
envvpu19
envvpu18
envvpu17
envvpu13
envvpu10
envvpu1
2LTR
8000
8500
9000
9500
Position

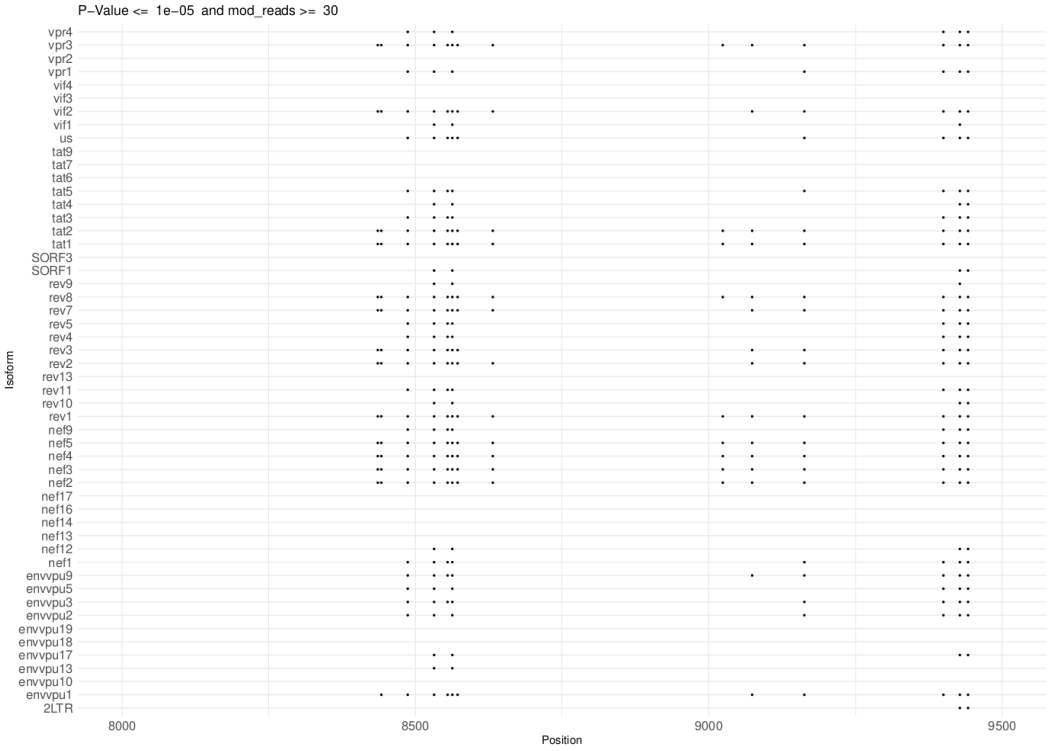
C


P−Value <= 1e−05 and mod_reads >= 30
Isoform
vpr4
vpr3
vpr2
vpr1
vif5
vif4
vif3
vif2
vif1
us
tat9
tat7
tat6
tat5
tat4
tat3
tat2
tat1
SORF3
SORF1
rev9
rev8
rev7
rev5
rev4
rev3
rev2
rev13
rev11
rev10
rev1
nef9
nef5
nef4
nef3
nef2
nef17
nef16
nef14
nef13
nef12
nef1
envvpu9
envvpu5
envvpu3
envvpu2
envvpu19
envvpu18
envvpu17
envvpu13
envvpu10
envvpu1
2LTR
8000
8500
9000
9500
Position

D


P−Value <= 1e−05 and mod_reads >= 30
Isoform
vpr4
vpr3
vpr2
vpr1
vif5
vif4
vif3
vif2
vif1
us
tat9
tat7
tat6
tat5
tat4
tat3
tat2
tat1
SORF3
SORF1
rev9
rev8
rev7
rev5
rev4
rev3
rev2
rev13
rev11
rev10
rev1
nef9
nef5
nef4
nef3
nef2
nef17
nef16
nef14
nef13
nef12
nef1
envvpu9
envvpu5
envvpu3
envvpu2
envvpu19
envvpu18
envvpu17
envvpu13
envvpu10
envvpu1
2LTR
8000
8500
9000
9500
Position

E



F

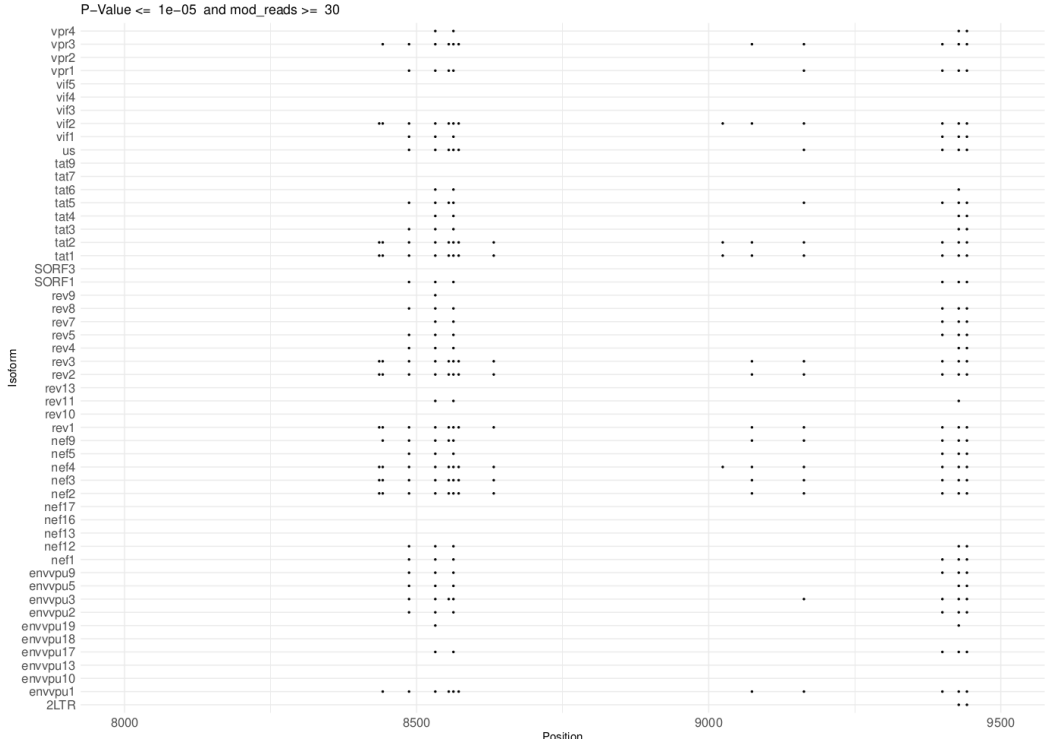
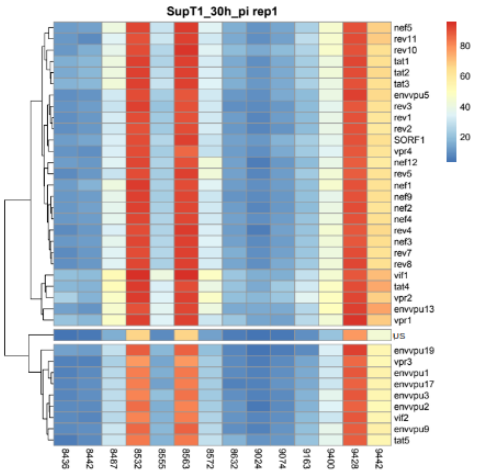


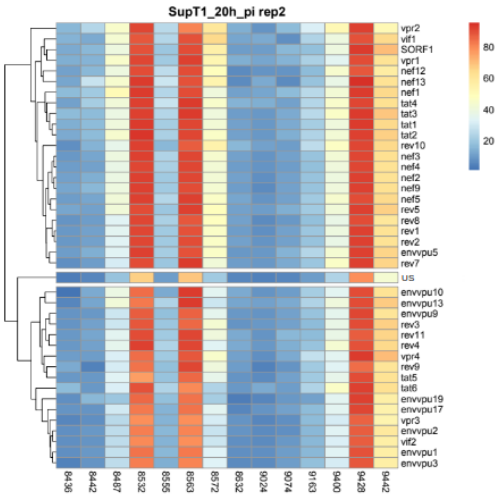
Figure S2. m^{6}A sites detected in splicing isoforms. **(A)** SupT1_20h_pi rep 1 sample; **(B)** SupT1_30h_pi rep1 sample; **(C)** SupT1_20h_pi rep2 sample; **(D)** SupT1_30h_pi rep2 sample; **(E)** CD4_48h_pi sample; **(F)**

CD4_72h_pi sample. Only positions with *p*-values ≤ 1 x 10^{-5} and a number of modified reads ≥ 30 are displayed.

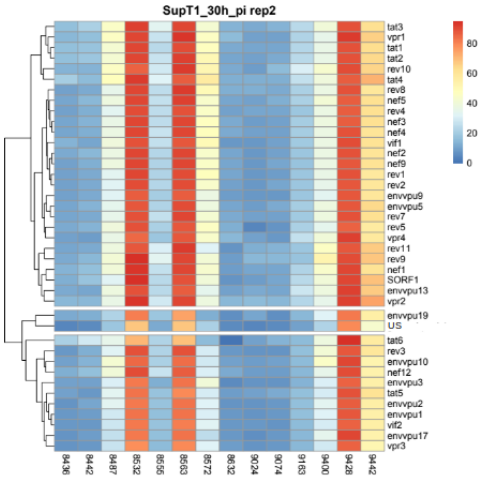
A - SupT1_30h_pi rep1



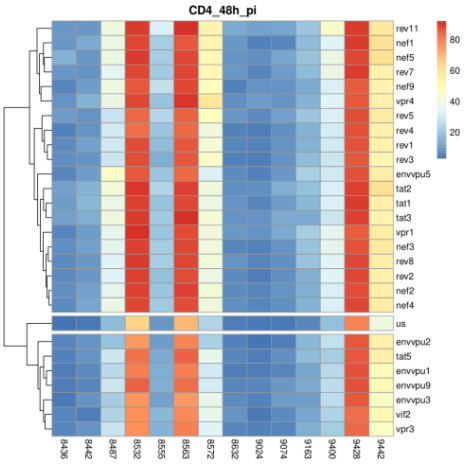
B - SupT1_20h_pi rep2



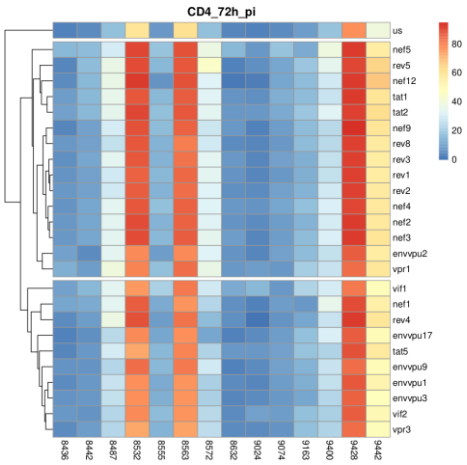
C - SupT1_30h_pi rep2



D - CD4_48h_pi



E - CD4_72h_pi



547 **Figure S3.** Heatmap of the methylation levels of the 14 m^6A sites detected in the various splicing isoforms. **(A-**

548 **E)**, results for the samples indicated. Methylation level was calculated for each set of full-length reads

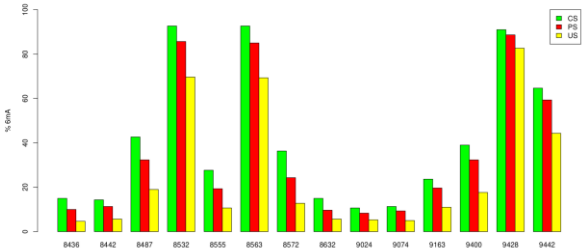
549 corresponding to each isoform represented by ≥100 reads (Materials and Methods). Isoforms are clustered

550 based on the methylation level of the 14 m^6A sites and the clusters obtained correspond essentially to US, PS

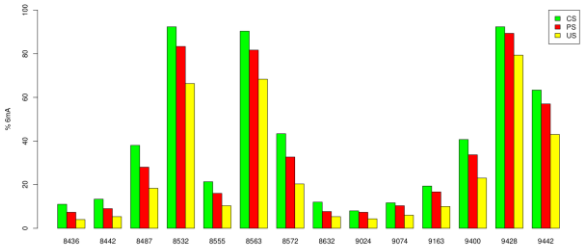
551 and CS isoforms.

552

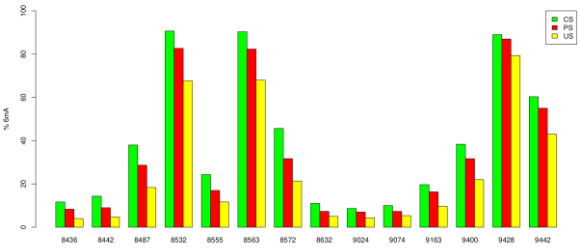
A - SupT1_30h_pi rep1



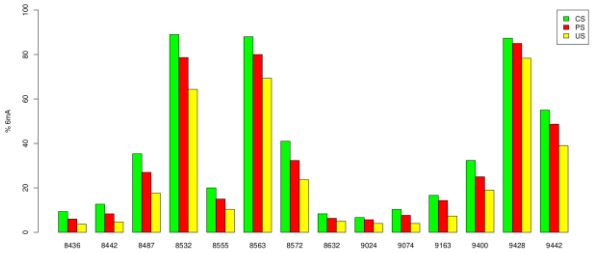
B - SupT1_20h_pi rep2



C - SupT1_30h_pi rep2



D - CD4_48h_pi



E - CD4_72h_pi

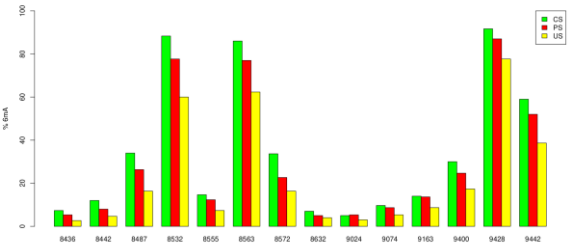
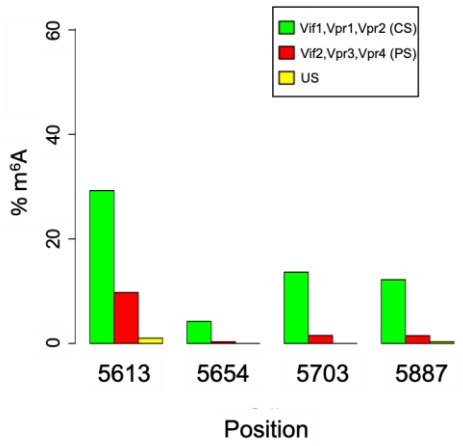


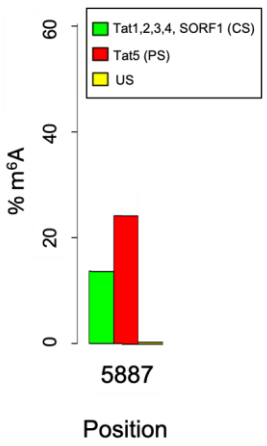
Figure S4. Methylation levels of m^6A sites in CS, PS and US transcripts. (**A-E**), results for the samples indicated.

SupT1_20h_pi rep2

A

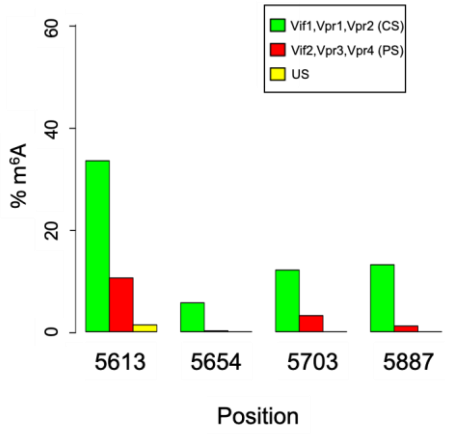


B

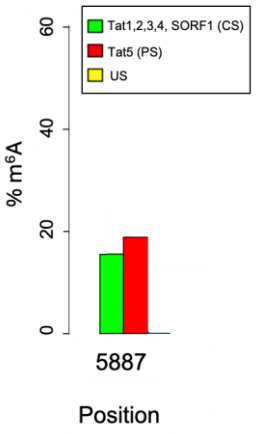


SupT1_30h_pi rep1

C

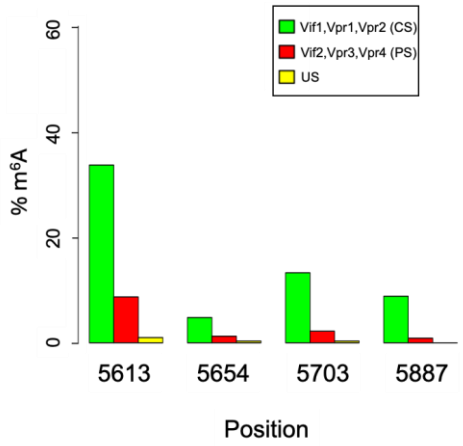


D

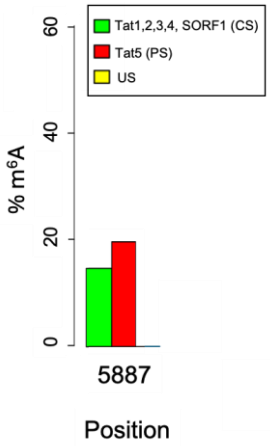


SupT1_30h_pi rep2

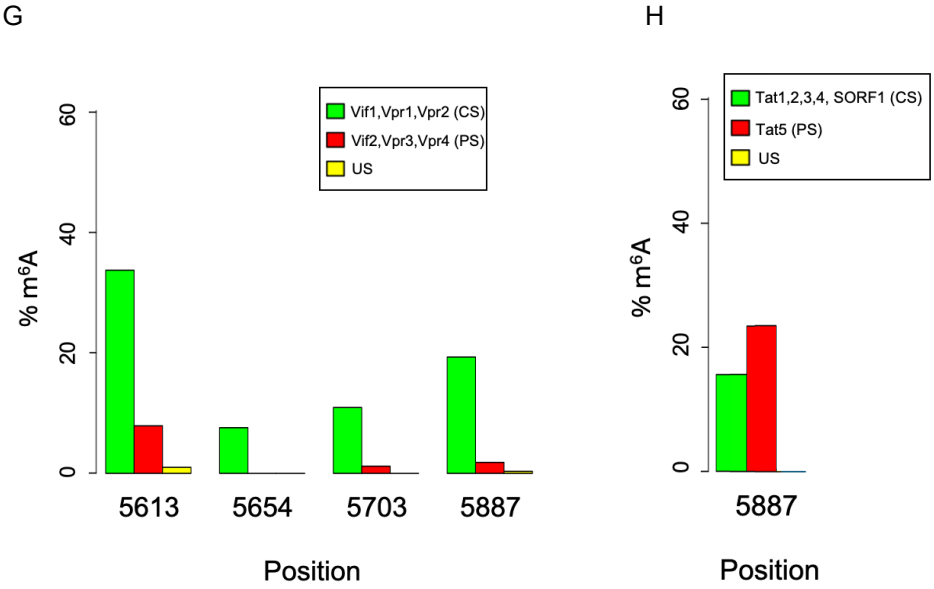
E



F



CD4_48h_pi



CD4_72h_pi

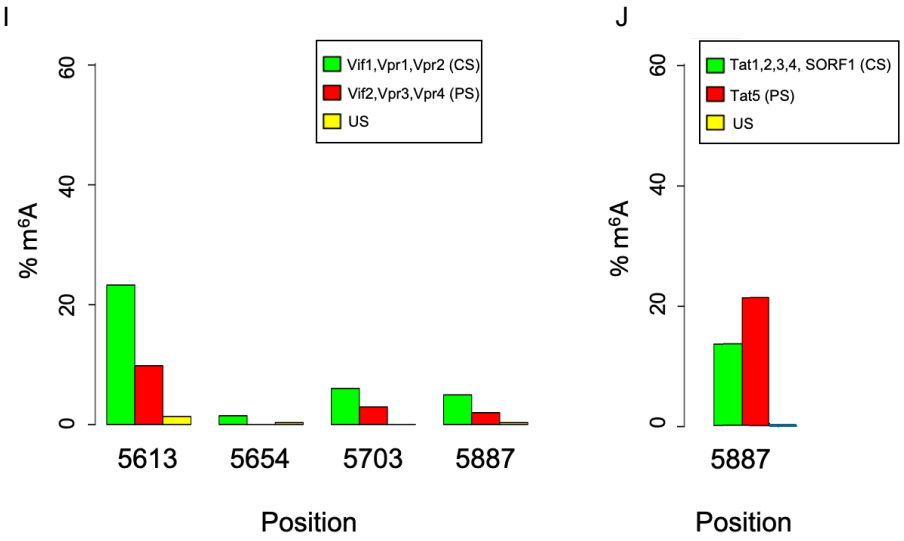


Figure S5. Methylation rates of the four isoform-specific m^6A sites in the samples indicated. The A5613, A5654 and A5703 sites are common to the Vif and Vpr isoforms; A5887 is common to the Vif, Vpr, Tat and SORF1 isoforms. **(A,C,E,G,I)** Methylation rates calculated for the four sites in CS (- intron 4, Vif1, Vpr1, Vpr2), PS (+ intron 4, Vif2, Vpr3, Vpr4) and US isoforms. **(B,D,F,H,J)** Methylation rates calculated for the A5887 site for the CS (Tat1, 2, 3, 4, SORF1), PS (Tat5) and US isoforms.

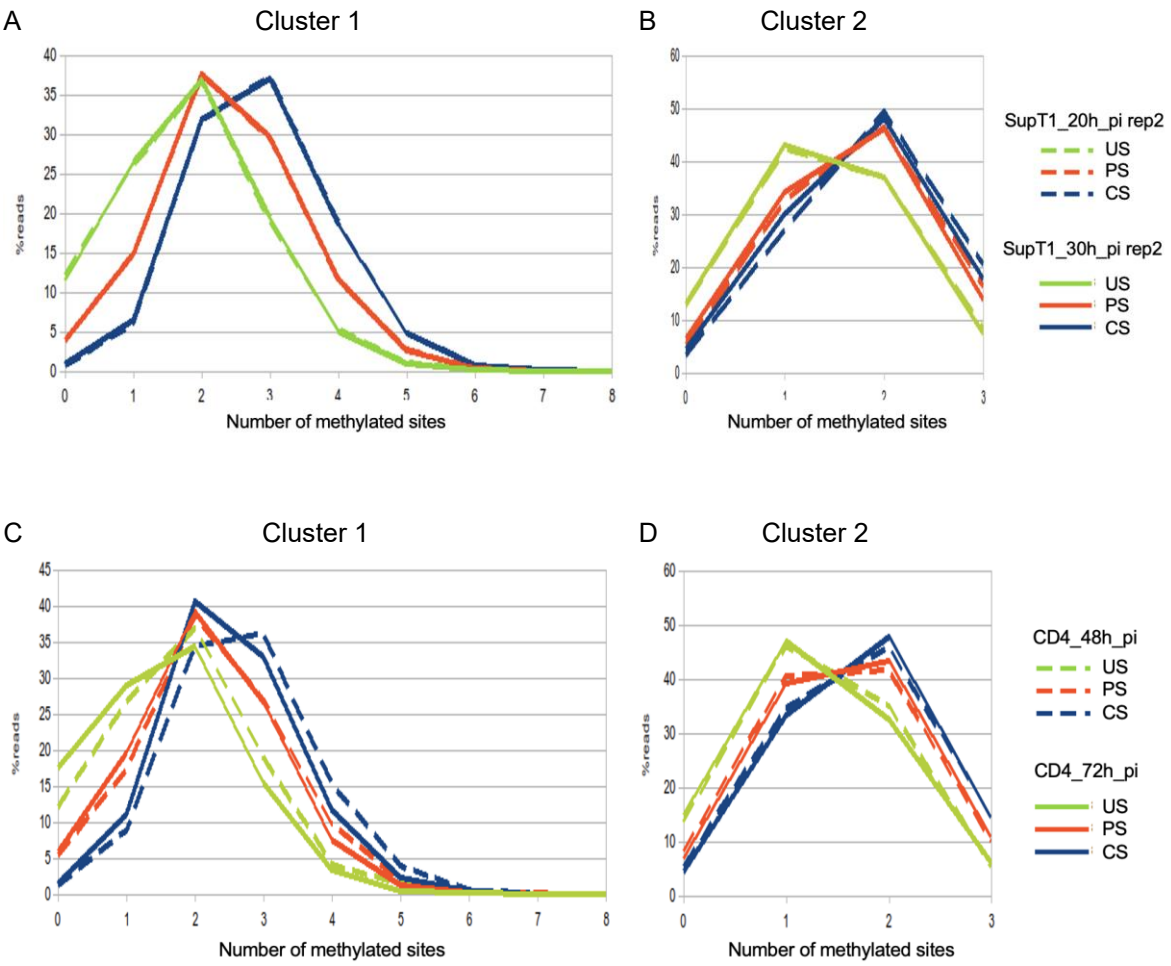
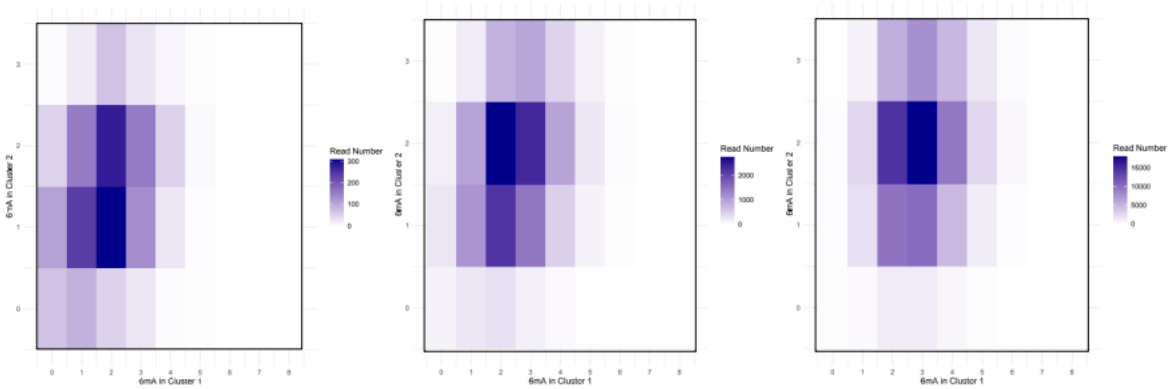


Figure S6. Single-molecule analysis of methylated m^6A sites. **(A)** Cluster 1 and **(B)** cluster 2 of the indicated SupT1 samples. **(C)** Cluster 1 and **(D)** cluster 2 of the indicated CD4 samples.

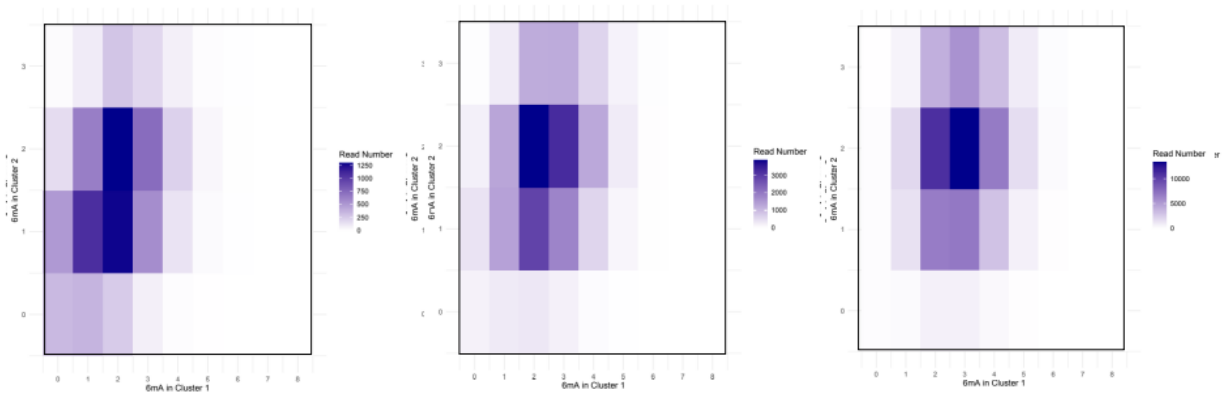
SupT1_30h_pi rep1

A (US) B (PS) C (CS)



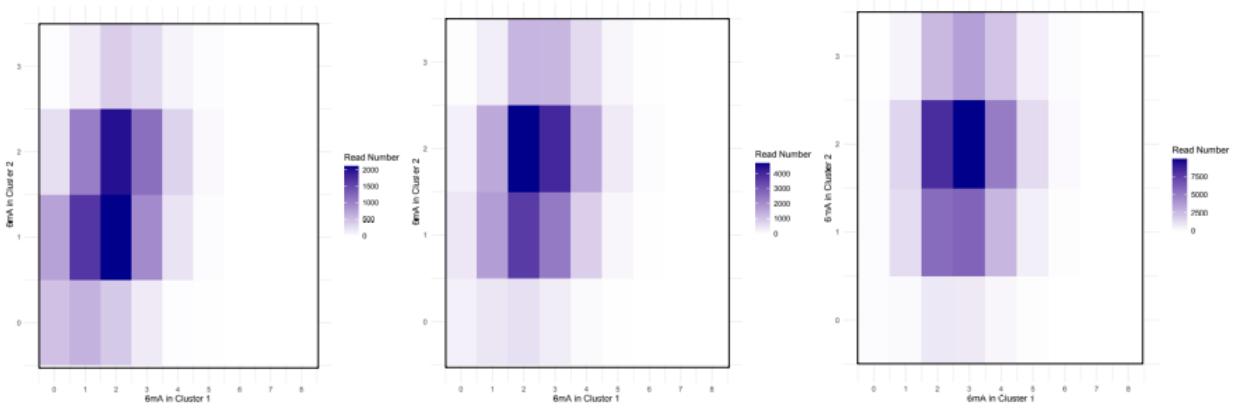
SupT1_20h_pi rep2

D (US) E (PS) F (CS)



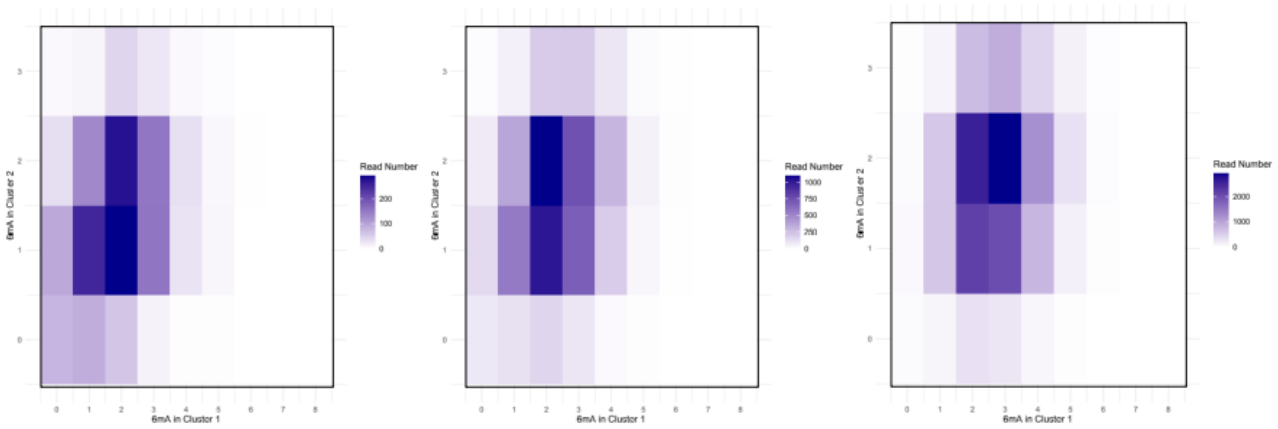
SupT1_30h_pi rep2

G (US) H (PS) I (CS)



CD4_48h_pi

J (US) K (PS) L (CS)



CD4_72h_pi

M (US) N (PS) O (CS)

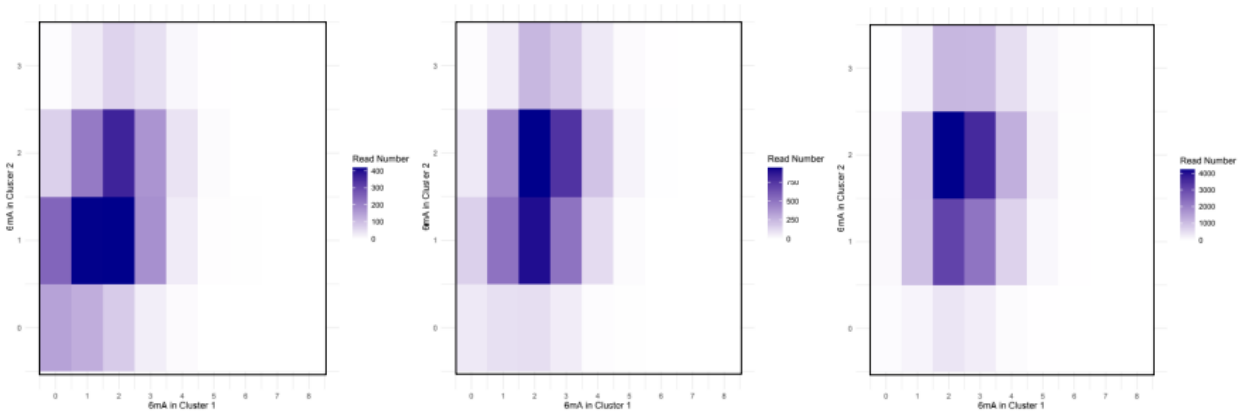


Figure S7. Single-molecule analysis of methylated m^6A sites illustrating the correlation between the methylation patterns for clusters 1 and 2 of the same transcripts. The intensity of the color on the heatmap indicates the number of reads, with the indicated numbers of methylated As in cluster 1 (0 to 8 m^6As) and cluster 2 (0 to 3 m^6As). US, unspliced; PS partially spliced; CS, completely spliced. One m^6A of one read is considered methylated if the probability of base modification is ≥ 0.5 (value in modkit extract), Materials and Methods).

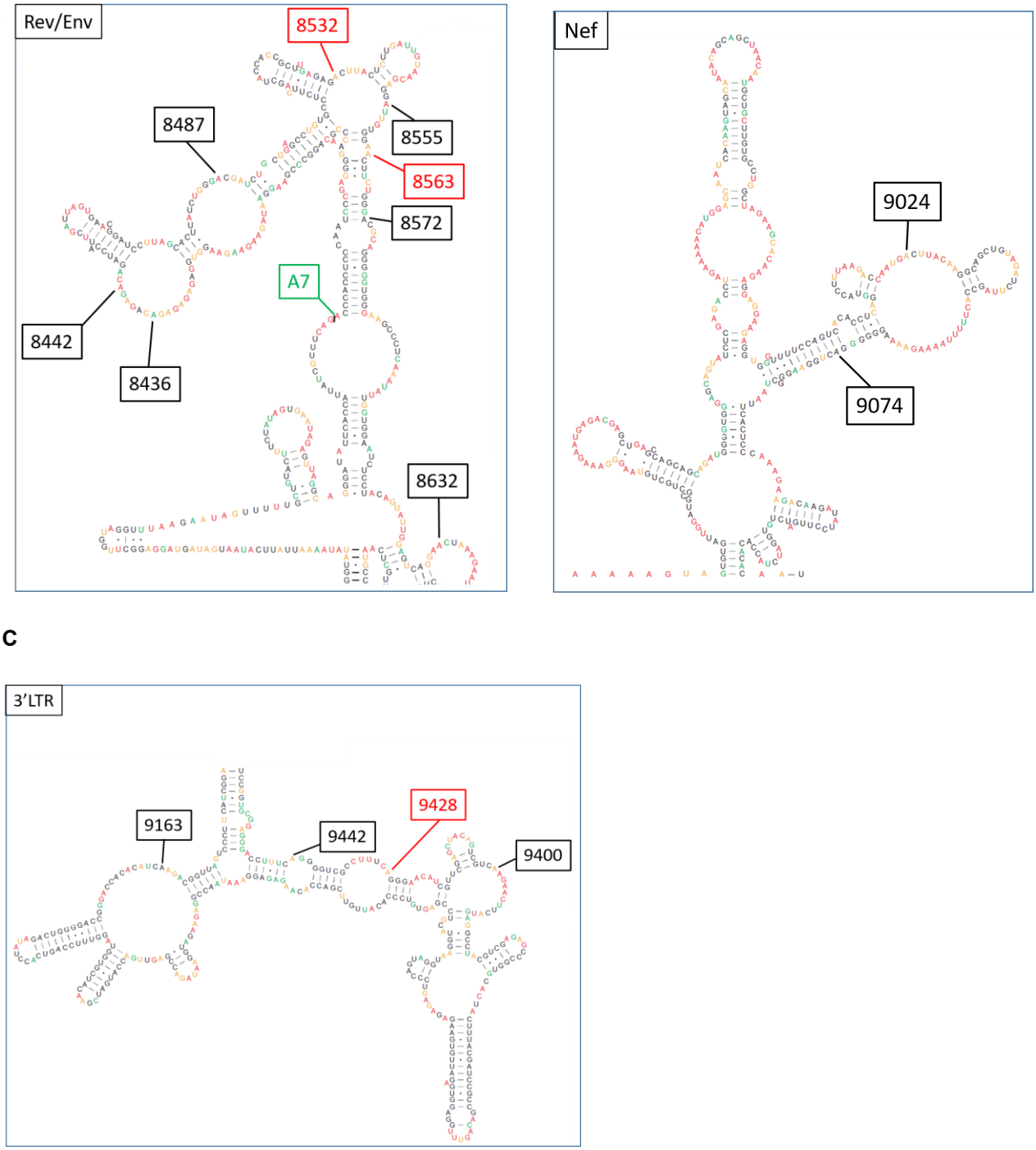


Figure S8. Secondary structures of HIV-1 RNA around the 14 m^6A sites detected in all isoforms. **(A)** Region of the Rev/Env sequence containing the eight m^6A sites of cluster 1; this structure is common to the US and PS transcripts as both retain the end of the unspliced intron 4 at the A7 acceptor site, in green (intron 3' end). In red, the two highly methylated sites (64-80%). **(B)** Region in the Nef sequence containing the methylated sites A9024 and A9074. **(C)** Region in the 3'LTR containing the three m^6A sites of cluster 2. In red, the highly methylated site (79-84%). The structure predictions are adapted from (54).

SUPPLEMENTARY TABLES

	SupT1_20h_pi rep1	SupT1_20h_pi rep2	SupT1_30h_pi rep1	SupT1_30h_pi rep2	CD4_48h_pi	CD4_72h_pi
Total reads	7317460	27288259	15758351	27289373	26172498	28942428
HIV-1 reads	386852	287479	1074370	345351	166733	393666
Infection rate (%)	5.5	1.31	7.1	1.53	0.77	1.9
Full-length reads	80230	134086	164360	127964	33587	39278
Full length (%)	20.74	46.64	15.3	37.05	20.14	9.98

Table S1. Run statistics. Infection rate, number of HIV reads divided by the total of HIV and human reads.

673 **A**

674

Donor sites	Acceptor sites
D1: 743	
D1c: 747	
	A1: 4913
	A1b: 4919
D2: 4962	
D2b: 5058	
	A2: 5390
D3 : 5463	
	A3: 5777
	A4c: 5936
	A4a: 5954
	A4b: 5960
	A5: 5976
	A5a: 5980
	A5c: 5992
D4: 6044	
D4a: 6048	
D4b: 6051	
	A7: 8369

675

676

677 **B**

678

		SupT1_ 20h_pi rep1	%	SupT1_ 20h_pi rep2	%	SupT1_ 30h_pi rep1	%	SupT1_ 30h_pi rep2	%	CD4_ 48h_pi	%	CD4_ 72h_pi	%
US		610	0.83	10494	8.21	2387	1.57	16566	13.75	2286	7.12	3280	8.76
envvpu1	D1A5	3971	5.38	17206	13.47	12515	8.22	16904	14.04	3699	11.52	4054	10.83
envvpu2	D1A4b	214	0.29	757	0.59	659	0.43	901	0.75	216	0.67	151	0.4
envvpu3	D1A4a	139	0.19	462	0.36	370	0.24	451	0.37	269	0.84	348	0.93
envvpu5	D1A1D2A5	281	0.38	756	0.59	672	0.44	943	0.78	111	0.35	93	0.25
envvpu9	D1A2D3A5	619	0.84	1345	1.05	1304	0.86	2367	1.97	397	1.24	195	0.52
envvpu10	D1A2D3A4b	46	0.06	61	0.05	54	0.04	143	0.12	30	0.09	9	0.02
envvpu13	D1A1D2A2D3A5	76	0.1	147	0.12	140	0.09	250	0.21	38	0.12	21	0.06
envvpu17	D1A5a	92	0.12	488	0.38	350	0.23	466	0.39	91	0.28	100	0.27
envvpu18	D1cA4b	2	0	6	0	3	0	9	0.01	3	0.01	2	0.01
envvpu19	D1cA5	49	0.07	127	0.1	120	0.08	140	0.12	30	0.09	40	0.11
nef1	D1A7	286	0.39	463	0.36	663	0.44	492	0.41	175	0.55	228	0.61
nef2	D1A5D4A7	29546	40.01	49676	38.88	65277	42.86	34533	28.67	8482	26.42	13658	36.47

nef3	D1A1D2A5D4A7	6934	9.39	7760	6.07	13894	9.12	7378	6.13	1066	3.32	1214	3.24
nef4	D1A2D3A5D4A7	7816	10.59	6156	4.82	11672	7.66	7242	6.01	1591	4.95	931	2.49
nef5	D1A1D2A2D3A5D4A7	2290	3.1	1512	1.18	3781	2.48	1897	1.58	376	1.17	162	0.43
nef9	D1A5aD4A7	690	0.93	1382	1.08	1540	1.01	857	0.71	205	0.64	357	0.95
nef12	D1cA5D4A7	214	0.29	331	0.26	496	0.33	268	0.22	65	0.2	105	0.28
nef13	D1A5D4aA7	25	0.03	35	0.03	58	0.04	27	0.02	8	0.02	10	0.03
nef14	D1A2D3A5D4aA7	8	0.01	8	0.01	11	0.01	8	0.01	2	0.01	0	0
nef16	D1cA5D4A7	10	0.01	10	0.01	23	0.02	6	0	1	0	3	0.01
nef17	D1A5D4bA7	8	0.01	9	0.01	22	0.01	9	0.01	2	0.01	2	0.01
rev1	D1A4bD4A7	3290	4.46	5743	4.5	7876	5.17	4360	3.62	1645	5.12	1350	3.61
rev2	D1A4aD4A7	3748	5.08	6522	5.11	8662	5.69	4546	3.77	3284	10.23	3979	10.63
rev3	D1A4cD4A7	213	0.29	348	0.27	539	0.35	252	0.21	396	1.23	584	1.56
rev4	D1A1D2A4bD4A7	363	0.49	504	0.39	830	0.55	445	0.37	169	0.53	110	0.29
rev5	D1A1D2A4aD4A7	199	0.27	326	0.26	464	0.3	218	0.18	137	0.43	111	0.3
rev7	D1A2D3A4bD4A7	786	1.06	641	0.5	1111	0.73	742	0.62	387	1.21	90	0.24
rev8	D1A2D3A4aD4A7	717	0.97	524	0.41	1006	0.66	626	0.52	507	1.58	190	0.51
rev9	D1A2D3A4cD4A7	33	0.04	38	0.03	52	0.03	35	0.03	51	0.16	34	0.09
rev10	D1A1D2A2D3A4bD4A7	191	0.26	154	0.12	315	0.21	169	0.14	81	0.25	23	0.06
rev11	D1A1D2A2D3A4aD4A7	202	0.27	161	0.13	335	0.22	151	0.13	130	0.4	42	0.11
rev13	D1cA4bD4A7	23	0.03	29	0.02	38	0.02	29	0.02	8	0.02	11	0.03
SORF1	D1A1D2bA3D4A7	218	0.3	145	0.11	358	0.24	155	0.13	66	0.21	84	0.22
SORF3	D1A1D2bA2D3A3D4A7	7	0.01	10	0.01	35	0.02	13	0.01	6	0.02	4	0.01
tat1	D1A3D4A7	4534	6.14	3746	2.93	4210	2.76	2101	1.74	1584	4.93	2379	6.35
tat2	D1A1D2A3D4A7	2354	3.19	2181	1.71	3641	2.39	1914	1.59	745	2.32	1019	2.72
tat3	D1A2D3A3D4A7	484	0.66	262	0.21	568	0.37	332	0.28	171	0.53	97	0.26
tat4	D1A1D2A2D3A3D4A7	171	0.23	109	0.09	294	0.19	120	0.1	73	0.23	53	0.14
tat5	D1A3	230	0.31	278	0.22	251	0.16	270	0.22	250	0.78	281	0.75
tat6	D1A1D2A3	16	0.02	31	0.02	39	0.03	32	0.03	34	0.11	39	0.1
tat7	D1A2D3A3	7	0.01	15	0.01	10	0.01	13	0.01	12	0.04	7	0.02
tat9	D1cA3D4A7	11	0.01	8	0.01	6	0	10	0.01	9	0.03	6	0.02
vif1	D1A1D4A7	98	0.13	267	0.21	389	0.26	280	0.23	39	0.12	112	0.3
vif2	D1A1	677	0.92	3199	2.5	2108	1.38	5234	4.35	547	1.7	591	1.58
vif3	D1cA1b	2	0	4	0	10	0.01	7	0.01	2	0.01	3	0.01
vif4	D1cA1	0	0	11	0.01	9	0.01	22	0.02	2	0.01	2	0.01

vif5	D1A1b	3	0	10	0.01	15	0.01	19	0.02	0	0	3	0.01
vpr1	D1A2D4A7	332	0.45	290	0.23	435	0.29	426	0.35	139	0.43	186	0.5
vpr2	D1A1D2A2D4A7	61	0.08	50	0.04	148	0.1	76	0.06	21	0.07	16	0.04
vpr3	D1A2	880	1.19	2793	2.19	2386	1.57	5651	4.69	2340	7.29	1022	2.73
vpr4	D1A1D2A2	64	0.09	166	0.13	142	0.09	335	0.28	131	0.41	54	0.14
Total isoforms		73840		127756		152293		120440		32109		37445	
2LTR*		1273	1.69	566	0.44	2391	1.55	878	0.72	52	0.16	109	0.29

679
680
681 **Table S2. (A)** List of the HIV-1 donor and acceptor splice sites. **(B)** Absolute and relative amounts of the HIV-1
682 transcript isoforms. %, amount calculated relative to all identified splicing isoforms (not including 2LTR); *,
683 amount calculated relative to all HIV-1 isoforms including 2LTR. Gray lines, the seven most abundant
684 isoforms, accounting for about 80% of all isoforms.
685

	SupT1_20h_pi rep2	SupT1_30h_pi rep1	SupT1_30h_pi rep2	CD4_48h_pi	CD4_72h_pi
SupT1_20h_pi rep1	0.946	0.991	0.861	0.890	0.925
SupT1_20h_pi rep2		0.972	0.964	0.952	0.978
SupT1_30h_pi rep1			0.891	0.911	0.946
SupT1_30h_pi rep2				0.941	0.933
CD4_48h_pi					0.971

Table S3. Pearson correlation coefficients for the relationship between the relative amounts of the splicing isoforms (Table S2B).

m^6A position	Sequence	SupT1-20h-pi rep1	SupT1-20h-pi rep2	SupT1-30h-pi rep1	SupT1-30h-pi rep2	CD4-48h-pi	CD4-72h-pi
A8436	AGACA	9.7	6.3	8	5.7	4.7	3.7
A8442	AGACA	10.3	7.0	9.7	6.7	6.3	5.3
A8487	GGACG	30	22.7	27	22	19.7	18
A8532	AGACT	79.3	72.7	79.3	72	68.3	63.7
A8555	GGATT	19.3	13	17.3	13	11.7	9.3
A8563	GAACT	79.3	74	79.7	72.7	70.3	66.3
A8572	GGACG	23.7	27	23.3	26	25.7	18.7
A8632	GAACT	10.3	6.7	9.3	6.7	5.7	3.7
A9024	TGACT	7.3	5.3	6.0	5.0	4.3	3.0
A9074	GGACT	8.7	7.7	9.7	6.3	5.7	5.7
A9163	GAACT	17.3	13	15	11.7	10.3	9.7
A9400	GAACT	29.7	28.0	32.0	25.0	21.0	18.7
A9428	GGACT	84.3	82.7	84.3	80.7	79.7	79.3
A9442	GGACT	53.3	49	52.3	46	41	41.3

Table S4. Rates of methylation of the 14 m^6A sites detected in all HIV-1 transcripts in each of the 6 samples. Underlined, 1-nt variants of the consensus DRACH sequences. Gray, the three most methylated sites, A8532, A8563, A9428.

723

Position	Sequence	SupT1_20h _pi rep1	SupT1_20h _pi+demeth rep1		SupT1_20h _pi rep2	SupT1_20h _pi+demeth rep2	
		% methyl.	% methyl.	p-value	% methyl.	% methyl.	*p*-value
A8436	AGACA	9.7	1	2.00E+101	6.3	2	9.60E-38
A8442	AGACA	10.3	1.7	7.90E-147	7	2.7	2.40E-36
A8487	GGACG	30	7.3	1.60E-265	22.7	11.7	3.90E-75
A8532	AGACT	79.3	49.3	0	72.7	51.7	6.70E-191
A8555	GGATT	19.3	11.3	4.00E-56	13	9	2.50E-14
A8563	GAACT	79.3	60	5.30E-174	74	60.7	8.60E-82
A8572	GGACG	23.7	12.3	1.80E-24	27	17.3	1.70E-15
A8632	GAACT	10.3	3	1.10E-80	6.7	2.3	3.30E-36
A9024	TGACT	7.3	1.3	3.00E-87	5.3	1.3	1.00E-43
A9074	GGACT	8.7	5	1.10E-22	7.7	5.3	3.20E-09
A9163	GAACT	17.3	3.3	1.00E-165	13	5.3	4.30E-58
A9400	GAACT	29.7	18.3	3.20E-93	28	21.7	6.90E-24
A9428	GGACT	84.3	22.3	0	82.7	25.3	0
A9442	GGACT	53.3	18.3	0	49	22.3	1.5e-311

724
725
726 **Table S5.** Demethylation efficiency for the 14 m^6A sites detected in all transcripts of SupT1_20h_pi rep1 and
727 SupT1_20h_pi rep2. SupT1_20h_pi+demeth is the SupT1_20h_pi sample demethylated with ALKBH5
728 (Materials and Methods). *P*-values are given for the difference between the numbers of methylated reads in
729 SupT1_20h_pi and in SupT1_20h_pi+demeth (Materials and Methods).

730 **A - SupT1_20h_pi rep1**

	8436	8442	8487	8532	8555	8563	8572	8632	9024	9074	9163	9400	9428	9442
envvpu1	9.00	11.00	31.00	85.33	18.00	84.67	24.33	8.67	7.67	9.00	18.00	32.33	88.67	59.00
envvpu2	8.12	13.37	33.63	86.14	18.83	84.77	35.48	12.00	10.58	9.33	21.21	31.54	87.42	59.77
envvpu3	11.45	14.48	37.81	89.22	18.87	88.22	22.58	11.31	10.83	12.97	18.33	33.51	91.33	52.72
envvpu5	14.94	11.89	43.11	93.81	24.92	92.54	28.20	12.85	7.77	10.32	26.73	42.67	90.46	65.72
envvpu9	14.33	12.00	36.67	89.00	26.67	88.33	30.82	11.33	9.00	9.67	27.67	33.67	89.67	62.67
nef1	14.18	12.26	39.08	92.82	28.73	92.46	36.00	13.40	10.44	9.64	22.14	36.62	90.65	68.79
nef12	13.61	15.46	42.78	88.73	27.75	93.39	41.97	11.77	7.23	12.10	26.31	38.24	89.33	61.10
nef2	14.00	14.00	41.33	92.67	27.33	92.33	33.67	14.33	9.33	11.00	22.00	38.00	91.00	63.00
nef3	14.33	16.00	44.67	93.67	30.33	93.00	36.00	14.33	11.00	11.33	25.67	41.33	90.33	66.33
nef4	16.00	15.67	45.67	93.67	31.67	93.33	37.33	16.33	12.67	10.67	27.67	40.00	91.00	66.33
nef5	18.33	17.00	49.67	94.67	35.67	94.00	42.33	19.33	14.67	11.67	29.00	42.00	92.00	69.33
nef9	13.00	15.00	41.67	92.00	25.00	92.67	37.44	13.67	11.33	11.33	23.00	36.33	90.33	66.33
rev1	12.67	14.00	37.00	92.00	24.33	90.67	32.00	13.00	9.33	9.00	22.67	37.67	91.00	62.00
rev10	10.75	18.48	41.39	92.68	31.41	90.54	33.72	15.48	13.17	16.88	27.25	45.02	91.37	65.40
rev11	15.10	15.91	41.05	95.12	31.18	94.15	52.88	14.10	16.66	15.40	26.08	44.85	91.55	71.86
rev2	12.67	12.67	37.33	90.33	24.33	90.67	31.67	12.00	9.67	10.67	21.33	36.67	90.00	60.67
rev3	10.95	16.66	30.68	88.50	24.78	90.60	33.52	8.55	6.79	11.39	19.84	35.14	88.76	61.13
rev4	13.67	12.33	45.00	90.67	30.00	94.00	29.37	12.67	13.67	12.67	25.33	39.00	90.00	60.67
rev5	11.64	13.28	33.21	94.26	28.68	93.90	33.63	16.46	9.68	9.06	23.33	37.76	90.90	61.83
rev7	14.67	16.00	41.00	93.00	29.67	93.33	42.52	18.00	13.00	12.33	30.33	38.33	91.00	65.00
rev8	18.33	17.33	42.33	93.67	32.67	93.67	34.26	13.33	13.00	11.00	28.67	39.67	92.00	64.67
SORF1	15.61	17.37	45.83	95.40	27.48	92.54	21.55	17.93	13.46	14.60	31.49	42.18	89.19	65.47
tat1	13.33	13.33	38.67	90.67	22.00	91.33	34.00	13.67	10.00	12.33	23.67	31.33	91.00	62.33
tat2	18.00	18.33	47.67	93.67	32.33	93.33	35.67	16.00	13.33	11.33	28.67	39.67	91.00	67.67
tat3	19.00	17.67	49.00	94.33	32.33	93.33	43.70	18.00	12.67	11.00	27.67	41.33	92.00	68.67
tat4	25.31	16.43	51.76	91.88	39.37	96.89	51.67	21.15	23.18	12.30	34.00	40.95	93.42	69.13

tat5	10.23	10.98	27.91	82.54	22.81	85.05	21.41	9.95	9.60	13.55	26.76	26.27	89.54	61.77
us	4.33	5.67	19.00	68.67	10.33	70.00	12.99	5.67	5.33	5.00	11.33	17.33	81.33	44.33
vif2	11.33	10.33	30.67	83.67	18.00	82.00	21.81	10.33	10.00	8.00	18.00	33.00	85.33	59.67
vpr1	12.00	20.67	47.00	92.33	30.00	92.67	37.22	18.33	14.00	16.33	30.73	35.13	91.29	67.36
vpr3	9.33	8.33	29.33	80.33	16.33	80.00	19.42	10.33	6.33	9.00	17.33	27.00	88.67	53.00

B - SupT1_30h_pi rep1

	8436	**8442**	**8487**	**8532**	**8555**	**8563**	**8572**	**8632**	**9024**	**9074**	**9163**	**9400**	**9428**	**9442**
envvpu1	7.00	8.67	27.00	82.33	16.00	82.67	23.00	8.33	6.67	9.33	15.00	32.00	88.00	55.67
envvpu13	13.71	10.78	44.99	89.45	36.43	91.19	41.19	15.18	14.74	15.49	25.68	49.71	92.21	75.56
envvpu17	7.33	10.00	27.67	81.67	13.33	82.00	23.12	7.00	7.00	10.67	15.00	30.67	86.67	54.33
envvpu19	12.84	12.06	29.28	87.28	18.09	86.83	19.76	8.01	3.64	5.53	8.43	33.11	91.81	53.41
envvpu2	8.33	8.67	28.00	84.33	15.67	81.00	23.06	11.33	5.33	8.33	17.00	34.67	86.67	51.33
envvpu3	7.33	9.67	28.33	81.00	19.67	84.00	22.68	9.33	7.00	8.67	17.33	36.67	88.00	54.00
envvpu5	9.67	10.33	32.00	90.33	23.00	88.67	31.31	11.67	9.67	14.67	19.00	40.00	91.67	65.67
envvpu9	8.67	9.33	31.00	86.00	18.00	86.33	27.33	10.00	7.00	10.67	20.00	38.00	87.67	56.67
nef1	12.00	17.00	40.33	93.00	24.33	90.67	35.21	12.00	9.67	12.00	19.67	45.00	89.00	62.33
nef12	12.67	11.67	35.67	91.00	25.00	89.00	42.24	13.33	6.33	11.00	19.00	39.00	87.67	59.00
nef2	11.00	13.33	37.00	92.00	24.67	91.00	31.67	12.67	7.67	11.67	19.00	39.67	90.00	62.00
nef3	11.00	12.67	37.00	91.00	26.33	92.33	34.00	13.67	9.00	12.67	21.33	43.00	88.67	64.33
nef4	11.00	12.00	36.67	91.33	25.00	91.33	31.33	13.33	9.33	12.33	21.00	41.00	89.67	63.33
nef5	12.00	13.00	40.33	91.33	29.33	92.67	35.00	16.33	10.67	14.33	24.33	45.67	89.00	67.67
nef9	11.67	12.33	38.00	91.67	23.33	91.33	30.67	10.67	8.33	11.00	20.00	38.33	89.67	63.33
rev1	10.00	11.67	32.67	90.00	22.00	89.67	31.33	12.33	7.67	12.00	19.00	38.67	89.67	61.00
rev10	11.60	16.33	37.67	90.33	26.00	93.67	35.32	18.87	14.94	18.45	25.64	46.29	89.19	63.75
rev11	11.00	8.67	35.00	89.67	30.33	92.33	31.67	15.00	10.00	13.67	27.63	45.33	90.67	68.48
rev2	10.00	11.00	32.00	89.67	22.33	90.00	32.67	11.67	7.67	10.67	19.00	39.33	89.33	60.33
rev3	10.33	12.33	29.67	91.67	21.00	89.67	31.43	10.00	8.67	11.33	18.33	37.00	88.67	63.00

rev4	8.67	11.67	33.33	90.67	25.00	93.00	33.20	12.33	6.33	13.33	20.33	44.00	88.33	65.33
rev5	9.00	11.00	29.67	90.33	23.67	91.67	42.43	13.33	8.00	12.67	22.33	43.67	87.33	62.67
rev7	9.00	11.33	35.67	90.00	26.33	91.67	32.33	15.67	9.67	13.67	21.67	40.33	90.00	62.33
rev8	10.00	11.67	35.00	91.00	24.67	91.67	32.08	14.00	9.33	14.67	22.67	42.67	91.00	63.33
SORF1	12.00	14.00	33.33	90.33	26.33	92.00	29.76	13.67	11.00	16.67	23.00	37.33	89.33	64.00
tat1	15.67	17.67	41.00	91.67	25.33	91.67	32.67	15.00	9.67	14.67	23.00	40.33	91.33	63.67
tat2	15.00	17.67	40.67	91.67	28.33	92.33	33.67	16.00	10.67	15.00	23.33	43.00	90.00	65.67
tat3	16.67	18.00	42.67	91.67	30.00	91.67	34.77	15.67	11.00	15.33	24.33	40.00	89.67	63.67
tat4	18.01	17.09	50.58	93.05	33.07	90.31	39.89	20.77	15.18	22.33	31.67	47.94	90.13	74.81
tat5	5.98	9.13	30.40	83.34	17.25	83.01	29.00	17.10	10.15	11.25	20.41	37.17	87.73	53.42
us	4.33	5.00	16.00	66.33	8.67	66.67	13.00	6.00	3.67	5.33	9.67	21.00	77.33	43.33
vif1	20.00	19.00	51.33	95.33	40.33	95.33	47.83	21.67	15.67	16.00	28.67	44.00	88.67	71.33
vif2	8.33	8.00	27.33	80.67	17.00	81.67	25.33	10.33	7.67	9.67	19.00	31.00	88.00	53.67
vpr1	20.33	17.67	47.00	92.33	34.00	91.67	31.09	21.33	14.33	13.00	27.00	43.33	93.67	68.33
vpr2	24.48	17.03	47.23	93.24	34.02	92.19	47.05	19.74	18.19	21.23	36.13	49.17	91.18	67.63
vpr3	6.67	6.67	24.00	77.33	14.33	77.33	23.00	9.00	6.33	8.33	14.33	30.00	85.67	52.33
vpr4	10.45	10.04	33.81	90.38	22.91	85.46	29.29	12.18	9.84	15.12	22.33	35.52	88.84	61.83

C - SupT1_20h_pi rep2

	8436	**8442**	**8487**	**8532**	**8555**	**8563**	**8572**	**8632**	**9024**	**9074**	**9163**	**9400**	**9428**	**9442**
envvpu1	7.00	9.33	27.67	83.33	15.67	82.33	32.67	7.33	7.00	10.00	16.33	33.67	89.67	56.67
envvpu13	5.51	11.81	40.84	83.11	23.31	93.33	28.77	12.52	11.52	12.71	23.43	40.38	90.21	67.64
envvpu17	8.33	8.00	30.33	82.33	15.33	84.00	34.27	5.33	8.67	11.33	17.33	35.00	90.67	54.00
envvpu19	7.36	15.50	19.51	85.73	17.85	83.70	21.41	3.94	5.55	9.42	11.26	34.93	91.77	56.19
envvpu2	7.67	9.00	26.67	81.33	16.00	81.00	27.90	8.00	6.67	7.67	16.00	33.67	87.33	59.33
envvpu3	4.67	8.33	26.67	81.67	14.33	80.33	32.57	6.33	6.67	10.67	18.00	31.33	88.67	56.33
envvpu5	10.00	13.67	33.67	92.00	21.33	89.67	45.98	9.33	9.33	15.33	24.67	43.33	92.33	65.33
envvpu9	8.00	9.67	29.33	87.33	18.00	86.00	35.67	9.00	8.00	11.67	18.67	36.00	90.67	61.00

nef1	17.33	20.67	50.67	93.00	24.33	90.33	40.24	11.67	12.00	17.33	25.33	42.00	93.33	64.33
nef12	11.57	15.67	39.67	90.33	26.67	90.33	49.08	6.67	8.33	8.67	16.73	44.47	88.07	56.91
nef2	11.00	13.67	38.33	92.00	21.00	90.67	42.67	11.00	7.67	11.00	18.33	39.67	91.67	61.67
nef3	9.67	12.33	39.00	92.00	23.00	91.33	43.67	11.33	8.67	11.67	19.00	42.00	90.67	62.67
nef4	10.00	11.33	37.33	91.33	22.67	91.00	42.67	10.67	8.00	10.67	17.67	40.67	90.67	62.00
nef5	10.67	12.00	41.33	92.33	24.33	91.67	46.33	12.33	8.67	11.00	20.67	43.67	90.33	64.33
nef9	12.00	16.00	39.00	92.00	20.67	90.33	43.33	10.33	6.67	11.33	18.67	40.00	91.67	63.33
rev1	9.67	11.33	34.67	91.00	20.67	89.33	45.67	10.33	7.67	11.00	19.67	39.67	92.33	61.67
rev10	7.36	14.90	36.50	92.38	16.01	87.69	54.58	15.88	11.23	10.62	26.01	46.63	90.64	65.18
rev11	8.94	12.32	34.53	90.17	20.57	91.44	40.44	12.65	8.76	16.33	16.61	37.49	87.40	66.10
rev2	9.00	11.67	33.33	91.00	20.33	90.00	46.00	11.00	8.00	11.00	18.67	38.00	91.67	61.67
rev3	8.67	11.33	29.33	88.67	20.00	85.67	42.78	8.67	9.67	11.33	16.67	37.33	90.33	66.33
rev4	9.67	11.33	35.33	91.67	23.00	90.67	33.33	12.67	9.00	14.00	18.33	43.33	89.33	64.33
rev5	12.33	11.67	41.33	90.33	19.33	91.67	49.08	11.33	10.00	14.00	20.99	41.33	92.00	67.01
rev7	7.33	10.33	31.00	90.33	20.00	93.00	41.47	10.67	9.33	13.00	19.67	45.67	92.00	65.67
rev8	8.00	10.00	33.67	90.67	22.33	88.67	47.60	11.00	6.33	9.33	18.00	37.33	90.00	59.67
SORF1	12.90	15.75	39.56	89.18	19.86	89.88	55.00	10.23	10.16	16.21	18.23	39.04	92.70	70.88
tat1	14.33	17.00	40.67	93.00	21.33	91.33	42.33	11.67	9.00	12.67	22.00	40.67	92.67	65.00
tat2	13.00	16.67	43.67	93.67	22.67	91.33	45.67	11.33	9.67	14.33	23.67	42.33	92.00	66.33
tat3	17.22	17.31	38.92	92.27	24.60	92.52	40.71	9.86	8.90	12.84	24.96	45.84	92.98	68.09
tat4	10.88	21.89	40.90	92.19	28.41	92.94	45.39	15.05	13.82	11.10	24.26	45.05	90.82	62.67
tat5	9.10	9.26	33.33	76.50	18.85	85.86	36.79	8.72	6.38	10.99	20.31	37.02	90.14	61.92
us	4.00	5.33	18.33	66.33	10.33	68.33	20.33	5.33	4.33	6.00	10.00	23.00	79.33	43.00
vif1	14.69	12.97	43.05	88.56	24.58	89.12	65.00	12.22	12.64	7.87	23.27	39.49	89.69	64.71
vif2	7.00	8.67	26.67	79.67	15.00	79.00	29.33	7.00	6.33	9.33	15.67	31.67	87.67	56.67
vpr1	13.90	14.56	44.85	91.54	20.35	91.78	52.28	13.95	12.57	14.77	22.17	34.28	89.81	62.40
vpr3	6.00	7.67	25.00	78.67	14.67	78.67	30.67	8.67	6.00	7.67	16.67	31.33	86.00	53.67
vpr4	9.92	10.10	32.13	84.33	20.74	87.87	44.67	10.92	7.21	13.30	27.57	34.37	93.30	70.14

737
738
739 **D - SupT1_30h_pi rep2**

	8436	8442	8487	8532	8555	8563	8572	8632	9024	9074	9163	9400	9428	9442
envvpu1	8.00	8.67	29.00	83.33	16.33	82.67	31.00	7.33	6.67	7.33	16.00	31.67	87.33	54.67
envvpu10	10.04	13.43	42.38	84.45	20.74	84.34	28.64	10.04	9.16	9.37	15.80	44.52	90.33	52.34
envvpu13	11.25	10.50	37.82	88.60	23.48	86.85	34.33	9.75	11.90	10.92	27.77	38.29	91.43	67.07
envvpu17	7.00	9.00	25.67	82.00	16.67	83.00	25.00	8.00	6.00	7.33	14.00	34.67	90.67	53.67
envvpu19	11.76	12.37	21.95	80.58	18.20	74.26	12.62	5.77	10.69	6.07	14.18	31.64	81.57	55.50
envvpu2	6.33	11.00	33.33	82.00	17.33	82.67	31.83	7.00	6.33	8.00	16.33	33.33	86.33	55.33
envvpu3	10.00	9.67	31.00	82.00	12.67	82.67	40.27	5.33	6.67	6.67	13.67	31.00	85.33	54.67
envvpu5	11.00	13.00	34.33	87.67	22.33	89.67	38.43	10.33	8.67	11.33	19.67	38.33	87.67	64.00
envvpu9	9.67	11.33	32.67	86.33	20.00	87.00	37.67	8.33	7.33	7.67	18.00	33.33	88.00	58.67
nef1	13.33	16.33	39.67	93.33	25.67	90.33	42.67	9.00	10.67	10.33	21.67	44.00	90.67	60.67
nef12	12.86	16.06	40.30	89.27	26.76	86.66	30.48	11.27	8.87	9.13	14.84	35.55	89.22	53.35
nef2	11.67	14.00	37.67	90.67	23.00	90.33	44.00	10.33	8.00	9.67	18.00	37.33	88.67	60.00
nef3	11.00	13.00	36.67	90.33	25.00	90.67	46.33	11.33	8.33	9.67	20.00	38.00	86.33	62.00
nef4	10.67	12.33	37.00	90.67	26.33	90.67	47.33	11.33	9.33	9.33	20.33	37.67	89.00	60.33
nef5	10.67	14.00	38.67	90.67	25.67	91.67	44.00	11.00	10.00	12.67	21.33	40.67	86.67	62.00
nef9	9.33	12.33	38.67	90.33	22.00	90.67	44.72	10.67	7.33	10.67	17.00	37.00	88.00	60.67
rev1	9.67	11.67	33.67	90.33	22.00	89.67	48.00	10.67	7.67	9.00	18.00	36.67	89.00	58.33
rev10	13.59	13.33	45.58	85.73	28.45	92.12	51.27	11.54	9.16	15.19	31.03	42.81	89.40	64.77
rev11	10.55	11.91	37.78	87.59	30.72	92.47	31.67	6.56	12.77	13.86	17.80	45.54	87.64	66.33
rev2	10.00	11.33	32.33	90.00	22.67	89.33	44.33	9.67	8.00	10.67	18.67	35.67	88.00	58.67
rev3	8.03	12.48	32.47	88.74	21.95	88.51	29.44	7.93	6.87	6.84	15.89	40.58	87.56	58.70
rev4	9.33	12.33	33.67	89.67	25.00	89.00	45.07	9.33	10.33	11.00	18.67	37.67	86.33	61.67
rev5	11.11	13.88	39.50	88.82	20.74	87.76	46.06	12.71	3.55	7.06	21.04	34.63	85.67	57.48
rev7	11.00	12.00	35.33	88.67	25.67	90.33	41.65	11.67	10.00	10.00	19.00	38.67	89.00	62.33
rev8	9.67	12.00	35.67	90.33	25.00	92.67	50.03	12.33	12.67	11.33	20.67	37.00	88.67	59.33

SORF1	12.42	16.89	32.11	94.40	24.17	88.26	40.44	7.05	8.57	10.78	23.63	39.87	90.45	68.20
tat1	16.00	17.33	41.67	90.33	23.67	90.00	47.00	12.00	10.67	11.00	22.67	36.67	89.00	62.67
tat2	16.33	18.00	42.00	90.67	27.00	91.33	50.33	10.67	10.00	11.67	24.00	36.67	88.67	64.33
tat3	16.33	20.67	45.33	91.67	33.33	91.00	42.08	13.00	12.00	13.33	24.00	39.67	88.39	61.75
tat4	19.67	15.32	48.11	91.97	31.46	86.39	57.16	12.60	12.49	12.06	26.10	38.97	87.02	71.48
tat5	9.51	7.87	32.99	83.32	17.43	78.74	30.08	8.93	9.43	11.52	16.67	29.45	89.08	61.24
us	4.00	4.67	18.33	67.67	11.67	68.00	21.33	5.00	4.33	5.33	9.67	22.00	79.33	43.00
vif1	13.63	10.94	39.15	88.35	25.39	88.06	49.40	7.70	7.30	9.21	19.00	36.98	89.76	61.41
vif2	7.67	8.00	27.33	81.67	17.33	81.00	32.67	7.67	7.33	7.00	15.67	29.67	85.33	55.00
vpr1	14.67	16.00	45.67	91.00	26.67	93.33	45.95	15.33	10.00	9.33	21.67	35.67	92.00	65.67
vpr3	7.00	7.33	26.67	78.00	15.67	79.33	32.00	8.33	6.33	7.33	14.33	28.33	85.33	52.00
vpr4	10.00	8.67	36.33	85.00	18.67	88.00	44.12	14.67	9.67	7.67	25.46	39.67	91.33	58.33

E - CD4_48h_pi

	8436	**8442**	**8487**	**8532**	**8555**	**8563**	**8572**	**8632**	**9024**	**9074**	**9163**	**9400**	**9428**	**9442**
envvpu1	5.00	8.00	25.67	79.67	14.67	79.67	30.67	5.33	5.00	7.33	13.33	26.00	86.33	48.67
envvpu2	6.88	9.81	19.92	74.43	11.32	77.99	23.33	5.86	3.03	5.14	16.86	27.63	85.65	53.03
envvpu3	4.56	9.51	22.62	75.73	12.23	74.77	36.21	7.90	5.92	11.27	11.24	26.61	85.33	46.48
envvpu5	7.14	11.93	45.08	85.65	20.36	83.00	31.67	8.24	6.46	10.88	23.21	29.81	86.67	55.11
envvpu9	5.67	6.67	29.67	83.33	15.00	81.67	31.26	6.67	7.33	11.00	14.00	26.67	86.67	50.33
nef1	10.71	14.12	38.38	87.90	22.53	85.38	52.33	10.98	6.18	6.73	17.40	32.87	90.02	61.40
nef2	9.00	12.00	36.00	89.33	20.33	88.67	41.67	8.33	6.67	10.00	15.33	32.33	87.67	54.67
nef3	9.00	12.67	36.00	89.33	22.33	88.33	39.82	7.33	6.33	9.33	19.67	35.33	86.67	57.00
nef4	8.33	13.67	37.67	88.33	21.33	87.67	39.67	8.67	6.67	8.67	16.00	31.67	88.33	55.67
nef5	15.67	10.33	38.67	90.33	22.33	88.33	50.71	9.33	8.67	13.67	19.67	37.67	88.33	61.00
nef9	7.44	11.73	39.31	85.17	27.79	88.61	48.33	6.30	9.77	9.36	13.97	31.99	85.01	52.10
rev1	8.00	12.00	30.33	87.33	18.00	85.67	42.33	7.33	6.00	9.33	16.00	30.33	86.00	53.33

rev11	10.41	10.87	40.27	89.91	31.81	91.85	52.50	14.30	12.47	13.49	20.85	44.22	89.89	58.00
rev2	8.00	10.33	33.00	89.33	19.00	87.33	38.33	8.33	5.33	9.00	15.33	32.67	88.33	54.00
rev3	8.67	10.00	29.67	87.00	14.67	85.33	46.68	6.33	5.67	10.00	17.67	29.67	85.67	51.67
rev4	6.54	9.91	30.21	82.20	20.12	82.49	40.27	5.80	5.09	9.10	16.12	30.78	86.17	53.41
rev5	11.88	10.11	28.55	82.50	24.56	84.00	48.33	10.19	8.26	12.17	20.49	31.64	86.95	56.69
rev7	11.00	9.67	35.00	88.00	21.33	88.00	51.56	11.00	7.00	11.33	17.67	37.33	90.00	53.67
rev8	8.00	10.33	33.33	90.00	22.00	88.00	41.19	9.33	6.67	9.00	19.67	33.67	88.00	55.33
tat1	11.67	16.67	38.67	88.67	19.67	88.00	38.67	10.00	7.33	11.33	17.00	31.67	89.67	54.00
tat2	12.00	15.67	41.67	90.67	23.67	89.33	40.12	9.00	8.67	15.00	19.00	32.33	89.00	61.67
tat3	12.41	15.42	39.63	90.03	21.29	91.69	42.78	9.31	10.10	14.32	14.13	29.71	89.46	52.37
tat5	6.56	9.33	29.17	80.74	14.61	83.35	36.52	7.36	5.60	8.60	17.60	22.48	85.54	54.73
us	3.67	4.67	17.67	64.33	10.33	69.33	23.67	5.00	4.00	4.00	7.33	19.00	78.33	39.00
vif2	6.67	5.67	24.00	76.67	16.67	77.33	36.54	9.67	4.67	7.67	13.00	24.33	82.33	51.00
vpr1	7.48	11.70	36.88	92.20	20.21	87.73	40.27	10.35	9.30	14.51	22.44	35.42	84.85	56.12
vpr3	7.33	9.33	28.67	76.33	15.67	79.00	33.33	7.00	6.33	7.00	15.00	22.00	83.00	46.33
vpr4	9.58	16.33	37.28	88.04	19.93	91.14	58.94	9.19	11.48	9.34	16.39	33.22	85.62	53.77

F - CD4_72h_pi

	8436	**8442**	**8487**	**8532**	**8555**	**8563**	**8572**	**8632**	**9024**	**9074**	**9163**	**9400**	**9428**	**9442**
envvpu1	4.67	8.00	26.67	78.00	11.33	77.00	21.33	4.67	4.33	8.67	12.33	25.67	88.00	52.33
envvpu17	3.04	5.98	23.47	79.26	10.34	77.79	10.83	2.76	4.93	9.04	16.20	30.97	86.44	49.67
envvpu2	9.81	5.26	33.25	79.40	8.52	79.24	31.67	6.47	3.75	9.94	15.54	26.23	86.61	56.40
envvpu3	7.00	7.33	25.67	78.33	12.67	76.67	21.22	4.33	5.00	8.00	10.33	23.33	85.67	47.67
envvpu9	6.45	6.54	22.60	83.44	14.22	82.11	28.89	3.37	5.13	9.50	13.86	26.36	83.67	49.67
nef1	10.19	11.43	33.13	86.54	11.09	77.04	23.33	6.09	3.35	8.38	7.84	34.98	89.44	54.64
nef12	5.22	14.82	35.96	92.74	7.08	88.78	31.67	0.71	1.73	10.30	15.33	28.93	92.42	67.41
nef2	7.33	11.00	33.33	89.00	13.67	86.00	32.33	6.33	4.67	9.00	13.00	29.67	92.00	58.00
nef3	7.67	10.33	33.00	89.33	14.33	86.67	34.60	8.67	4.00	8.67	15.33	30.33	92.33	60.00

nef4	6.33	11.00	33.33	87.33	16.67	86.00	31.96	7.67	6.67	11.00	16.67	31.00	90.33	58.00
nef5	14.43	14.90	29.77	90.81	17.71	88.27	37.22	14.41	8.11	17.00	11.50	38.55	92.33	56.31
nef9	4.33	10.00	29.67	88.67	15.67	86.00	27.37	7.67	3.00	9.00	15.67	24.33	94.67	59.67
rev1	6.33	9.33	29.00	86.00	14.33	84.67	32.95	6.33	4.33	7.67	12.67	26.00	91.00	56.67
rev2	7.33	10.00	29.00	87.00	13.67	83.67	31.33	6.00	5.00	9.67	13.00	28.33	91.67	57.00
rev3	6.00	10.33	33.00	88.00	13.33	83.00	33.01	7.67	4.33	9.67	12.67	26.00	91.00	56.33
rev4	5.27	9.10	36.86	88.56	8.62	82.52	15.72	5.38	-0.33	6.33	11.00	26.69	92.42	59.45
rev5	4.30	15.41	36.27	89.67	13.95	87.63	45.00	3.37	6.21	10.24	17.81	34.24	90.67	65.05
rev8	7.62	9.22	26.06	87.77	13.13	81.19	28.64	5.19	4.04	8.19	15.77	24.06	91.15	57.47
tat1	9.00	14.67	35.33	88.00	14.67	86.33	32.67	7.00	6.00	12.67	15.67	27.00	92.67	59.67
tat2	9.33	14.67	34.67	90.00	16.67	88.00	33.57	7.33	7.33	12.00	18.00	32.67	91.67	63.00
tat5	4.34	7.47	29.48	73.75	14.33	79.96	22.06	7.59	7.59	9.82	16.89	22.01	88.27	57.73
us	2.67	4.67	16.33	60.00	7.33	62.33	16.33	4.00	3.00	5.33	8.67	17.33	77.67	38.67
vif1	9.57	14.21	32.94	75.81	21.43	81.91	28.33	11.70	14.35	8.08	22.00	29.61	81.69	48.69
vif2	7.67	6.67	23.33	79.00	15.33	77.33	23.50	5.33	10.00	9.00	16.67	21.00	86.33	50.67
vpr1	12.31	8.51	38.77	80.33	16.22	83.73	36.11	7.49	8.46	8.05	20.18	26.76	84.21	58.97
vpr3	5.33	8.67	23.00	73.67	13.00	74.33	24.84	5.00	5.00	7.00	16.67	22.00	84.00	48.33

Table S6. Rates of methylation of the 14 m^6A sites detected in all transcripts of the splicing isoforms found in the samples indicated. **A-F**, rates for the isoforms represented by ≥ 100 full-length reads (Materials and Methods).

Isoform		SupT1_30h_pi rep1			
CS - i4	PS + i4	%methyl (5613)	%methyl (5654)	%methyl (5703)	%methyl (5887)
Vif1		21	4	5.7	10.7
	Vif2	8.7	0	2.0	1
Vpr1		40.3	5	12	13.7
	Vpr3	10.3	0.3	1.3	0.3
Vpr2		39.0	8.0	18.5	15
	Vpr4	12.6	0.5	6.1	2.0
Tat1					16.7
	Tat5				18.8
Tat2					13
Tat3					13.3
Tat4					18.3
SORF1					16
US		1.3	0	0	0

752
753

Isoform		SupT1_20h_pi rep2			
CS - i4	PS + i4	%methyl (5613)	%methyl (5654)	%methyl (5703)	%methyl (5887)
Vif1		15.7	1.2	5.9	5.5
	Vif2	7.3	0.3	1.7	1.3
Vpr1		40.1	4.7	15.8	14.8
	Vpr3	8.7	0.3	2	1
Vpr2		32	6.7	19.2	16.3
	Vpr4	13.3	0.3	0.9	2.1
Tat1					17
	Tat5				24.3
Tat2					11.7
Tat3					16.6
Tat4					16.6
SORF1					16.7

US		1.0	0	0	0.3

754
755

Isoform		**SupT1_30h_pi rep2**			
CS - i4	PS + i4	%methyl (5613)	%methyl (5654)	%methyl (5703)	%methyl (5887)
Vif1		18.7	1.4	9.4	7.0
	Vif2	7	0.7	1.7	1
Vpr1		42.3	7.7	17	13.3
	Vpr3	9.3	1	2	0.3
Vpr2		41.0	5.5	13.8	6.4
	Vpr4	10	2.1	3	1.3
Tat1					16.7
	Tat5				19.7
Tat2					12.0
Tat3					14.3
Tat4					14.3
SORF1					16.1
US		1.0	0.3	0.3	0

756
757

Isoform		**CD4_48h_pi**			
CS - i4	PS + i4	%methyl (5613)	%methyl (5654)	%methyl (5703)	%methyl (5887)
Vif1		15.1	5.3	5.2	10.8
	Vif2	6.7	0.3	1.3	2
Vpr1		43.5	5.1	8.8	14.0
	Vpr3	8.7	0.7	1	0.7
Vpr2		42.2	12.3	18.7	33.0
	Vpr4	8.3	0	1.2	2.8
Tat1					19.3
	Tat5				23.5
Tat2					13.7

Tat3					13.5
Tat4					16.6
SORF1					15.3
US		1.0	0	0	0.3

Isoform		**CD4_72h_pi**			
CS - i4	PS + i4	%methyl (5613)	%methyl (5654)	%methyl (5703)	%methyl (5887)
Vif1		9.8	1.0	4.5	6.0
	Vif2	8	1.0	2.3	1.3
Vpr1		27.1	4.4	13.9	9.1
	Vpr3	8	0.7	3	1
Vpr2		32.7	0	0	0
	Vpr4	13.3	0	3.4	3.6
Tat1					13
	Tat5				21.1
Tat2					8
Tat3					9.1
Tat4					20.4
SORF1					16.7
US		1.3	0.3	0	0.3

Table S7. Methylation rates for the A5613, A5654, A5703 and A5887 m^6A sites in the Vif, Vpr, Tat and SORF1 isoforms, as measured in the samples indicated. The first two columns indicate the isoforms completely spliced (CS, lacking intron 4, - i4) or partially spliced (PS, containing intron 4, + i4).

766 **A**

767 **SupT1_20h_pi rep1**

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3
0	19	26	11	1	34	75	39	7	86	82	72	12
1	16	61	24	3	65	313	274	54	149	876	1066	290
2	9	80	75	22	109	679	864	261	561	4310	6522	2136
3	7	33	47	9	54	540	854	313	655	5241	9068	3503
4	1	8	16	1	26	208	433	202	284	2878	5581	2679
5	0	3	4	2	5	45	115	84	100	861	2005	1203
6	0	0	0	0	0	13	24	14	18	162	427	303
7	0	0	0	0	0	2	1	0	1	19	50	57
8	0	0	0	0	0	0	0	0	0	2	4	4

768

769 **B**

770 **SupT1_20h_pi rep2**

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3
0	341	503	167	18	194	396	220	34	147	143	124	29
1	366	1076	644	95	293	1373	1320	287	251	1508	1962	570
2	252	1280	1298	282	346	2877	3872	1212	722	6740	11069	4025
3	78	566	737	191	213	1826	3268	1234	739	6966	13411	5547
4	13	135	222	77	50	605	1268	615	332	3078	6830	3278

5	3	24	41	15	12	137	279	183	74	719	1696	982
6	0	5	7	5	1	19	26	30	4	75	247	168
7	0	0	0	1	0	3	4	0	0	4	19	18
8	0	0	0	0	0	0	0	0	0	0	0	0

C
SupT1_30h_pi rep1

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3
0	71	110	50	4	134	259	165	21	138	170	194	53
1	91	238	156	24	251	1125	961	208	375	2112	2709	881
2	52	308	278	71	316	2232	2757	804	1318	9691	14648	5491
3	29	134	157	31	156	1460	2359	938	1304	10289	17984	7611
4	5	28	52	12	58	481	955	452	596	4745	9458	4762
5	2	3	7	2	13	128	240	142	135	1211	2667	1634
6	0	0	1	0	0	19	43	22	17	170	442	349
7	0	0	0	0	0	1	1	4	1	19	52	47
8	0	0	0	0	0	0	0	0	0	0	3	5

D
SupT1_30h_pi rep2

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3

0	496	740	257	25	265	432	264	42	122	139	143	31
1	609	1681	1048	158	440	1736	1520	300	237	1301	1541	423
2	428	2110	1999	389	558	3630	4691	1300	854	5717	8325	2608
3	160	945	1174	284	319	2405	4023	1292	792	5991	9961	3600
4	22	215	344	94	105	847	1596	643	328	2685	5025	2192
5	6	32	53	27	17	152	371	160	28	584	1332	649
6	0	3	3	3	0	11	60	16	11	68	202	134
7	0	0	0	0	1	0	1	3	0	7	14	11
8	0	0	0	0	0	0	0	0	0	1	0	0

E

CD4_48h_pi

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3
0	81	95	35	7	94	151	84	16	51	78	52	23
1	90	253	134	12	123	562	370	59	120	623	613	125
2	64	293	277	45	172	1015	1097	211	320	2203	2606	736
3	14	156	155	26	103	688	746	215	275	2048	2930	901
4	2	30	34	7	24	208	306	106	104	790	1242	453
5	2	10	9	4	9	36	56	19	18	158	322	152
6	0	0	1	0	1	5	10	6	4	17	42	29
7	0	0	0	0	0	0	1	2	0	0	3	4

8	0	0	0	0	0	0	0	1	0	0	0	0

F
CD4_72h_pi

	US				PS				CS			
	0	1	2	3	0	1	2	3	0	1	2	3
0	148	250	73	4	78	165	77	10	55	111	91	16
1	129	418	214	32	108	513	425	68	168	1000	1021	200
2	82	419	366	68	116	902	943	256	402	3123	4221	1138
3	26	178	170	48	65	510	762	189	273	2264	3585	1129
4	7	31	43	12	9	126	211	78	67	700	1273	508
5	0	3	6	1	5	15	40	16	10	118	244	141
6	0	2	0	0	1	0	4	2	4	10	21	28
7	0	0	0	0	0	0	0	1	0	0	2	1
8	0	0	0	0	0	0	0	0	0	0	1	0

Table S8. Single-molecule analysis of methylated m^6A sites illustrating the correlation between the methylation of clusters 1 and 2 of the same transcript, by sample. **(A-F)** Number of reads in the indicated samples with defined numbers of methylated As in cluster 1 (by line: 0 to 8 m^6As) and in cluster 2 (by column: 0 to 3 m^6As). US, unspliced; PS partially spliced; CS, completely spliced. One m^6A of one read is considered methylated if the probability of base modification is ≥ 0.5 (value in modkit extract, Materials and Methods).

Position	Sequence	SupT1_20h_pi rep1	SupT1_20h_pi rep2	SupT1_30h_pi rep1	SupT1_30h_pi rep2	CD4_48h_pi	CD4_72h_pi
		% methyl.	% methyl.	% methyl.	% methyl.	% methyl.	% methyl.
2L125	UGACU	35.3	30.7	33.3	35	38.4 (19)	24.4 (25)
2L183	AGACU	11.7	9	13.7	9	6.5 (3)	4.1 (4)
2L272	GAACU	27.7	30.7	30.3	29	21.9 (11)	26 (24)
2L509	GAACU	32.7	37.3	35	32.7	26.8 (14)	19.7 (21)
2L537	GGACU	90.3	91.3	89.3	90	94.1	79.2
2L551	GGACU	66.3	66.7	68.7	70	69.2	64

Table S9. Rates of methylation of the six m^6A sites detected in the 2-LTR transcripts. The names of the sites refer to the position of the corresponding A in the HIV-1 reference genome. * indicates that the sequence context of the A9074 in the 2-LTR transcript (AGACT, Figure 6B) differs from the context of the same position, 9074, in the reference genome (GGACT). The number of methylated reads is indicated in brackets if below 30.

REFERENCES

1. Y. Saletore *et al.*, The birth of the Epitranscriptome: deciphering the function of RNA modifications. *Genome Biol* **13**, 175 (2012).
2. J. Liu *et al.*, A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-adenosine methylation. *Nat Chem Biol* **10**, 93-95 (2014).
3. D. Dominissini *et al.*, Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. *Nature* **485**, 201-206 (2012).
4. G. Zheng *et al.*, ALKBH5 is a mammalian RNA demethylase that impacts RNA metabolism and mouse fertility. *Mol Cell* **49**, 18-29 (2013).
5. I. A. Roundtree *et al.*, YTHDC1 mediates nuclear export of N(6)-methyladenosine methylated mRNAs. *Elife* **6** (2017).

809 6. S. Zaccara, S. R. Jaffrey, A Unified Model for the Function of YTHDF Proteins in Regulating m(6)A-
810 Modified mRNA. *Cell* **181**, 1582-1595 e1518 (2020).
811 7. S. Zaccara, R. J. Ries, S. R. Jaffrey, Reading, writing and erasing mRNA methylation. *Nat Rev Mol*
812 *Cell Biol* **20**, 608-624 (2019).
813 8. S. U. Madugalle, K. Meyer, D. O. Wang, T. W. Bredy, RNA N(6)-Methyladenosine and the Regulation
814 of RNA Localization and Function in the Brain. *Trends Neurosci* **43**, 1011-1023 (2020).
815 9. S. Phillips, T. Mishra, S. Huang, L. Wu, Functional Impacts of Epitranscriptomic m(6)A Modification on
816 HIV-1 Infection. *Viruses* **16** (2024).
817 10. I. A. Roundtree, M. E. Evans, T. Pan, C. He, Dynamic RNA Modifications in Gene Expression
818 Regulation. *Cell* **169**, 1187-1200 (2017).
819 11. X. Wang *et al.*, N6-methyladenosine-dependent regulation of messenger RNA stability. *Nature* **505**,
820 117-120 (2014).
821 12. X. Wang *et al.*, N(6)-methyladenosine Modulates Messenger RNA Translation Efficiency. *Cell* **161**,
822 1388-1399 (2015).
823 13. X. Zhao *et al.*, FTO-dependent demethylation of N6-methyladenosine regulates mRNA splicing and is
824 required for adipogenesis. *Cell Res* **24**, 1403-1419 (2014).
825 14. S. Lavi, A. J. Shatkin, Methylated simian virus 40-specific RNA from nuclei and cytoplasm of infected
826 BSC-1 cells. *Proc Natl Acad Sci U S A* **72**, 2012-2016 (1975).
827 15. G. Lichinchi *et al.*, Dynamics of the human and viral m(6)A RNA methylomes during HIV-1 infection of
828 T cells. *Nat Microbiol* **1**, 16011 (2016).
829 16. E. M. Kennedy *et al.*, Posttranscriptional m(6)A Editing of HIV-1 mRNAs Enhances Viral Gene
830 Expression. *Cell Host Microbe* **19**, 675-685 (2016).
831 17. N. Tirumuru *et al.*, N(6)-methyladenosine of HIV-1 RNA regulates viral infection and HIV-1 Gag protein
832 expression. *Elife* **5** (2016).
833 18. S. Cristinelli, P. Angelino, A. Janowczyk, M. Delorenzi, A. Ciuffi, HIV Modifies the m6A and m5C
834 Epitranscriptomic Landscape of the Host Cell. *Frontiers in Virology* **1** (2021).
835 19. C. Pereira-Montecinos *et al.*, Epitranscriptomic regulation of HIV-1 full-length RNA packaging. *Nucleic*
836 *Acids Res* **50**, 2302-2318 (2022).
837 20. S. Chen *et al.*, N6-methyladenosine modification of HIV-1 RNA suppresses type-I interferon induction
838 in differentiated monocytic cells and primary macrophages. *PLoS Pathog* **17**, e1009421 (2021).

839 21. K. Tsai *et al.*, Epitranscriptomic addition of m(6)A regulates HIV-1 RNA stability and alternative
840 splicing. *Genes Dev* **35**, 992-1004 (2021).
841 22. A. B. R. McIntyre *et al.*, Limits in the detection of m(6)A changes using MeRIP/m(6)A-seq. *Sci Rep* **10**,
842 6590 (2020).
843 23. D. R. Garalde *et al.*, Highly parallel direct RNA sequencing on an array of nanopores. *Nat Methods* **15**,
844 201-206 (2018).
845 24. O. Begik, J. S. Mattick, E. M. Novoa, Exploring the epitranscriptome by native RNA sequencing. *RNA*
846 **28**, 1430-1439 (2022).
847 25. M. Jain, R. Abu-Shumays, H. E. Olsen, M. Akeson, Advances in nanopore direct RNA sequencing. *Nat*
848 *Methods* **19**, 1160-1164 (2022).
849 26. A. Baek *et al.*, Single-molecule epitranscriptomic analysis of full-length HIV-1 RNAs reveals functional
850 roles of site-specific m(6)As. *Nat Microbiol* **9**, 1340-1355 (2024).
851 27. W. Liu-Wei *et al.*, Sequencing accuracy and systematic errors of nanopore direct RNA sequencing.
852 *BMC Genomics* **25**, 528 (2024).
853 28. D. F. Purcell, M. A. Martin, Alternative splicing of human immunodeficiency virus type 1 mRNA
854 modulates viral protein expression, replication, and infectivity. *J Virol* **67**, 6365-6378 (1993).
855 29. K. E. Ocwieja *et al.*, Dynamic regulation of HIV-1 mRNA populations analyzed by single-molecule
856 enrichment and long-read sequencing. *Nucleic Acids Res* **40**, 10345-10355 (2012).
857 30. A. Emery, S. Zhou, E. Pollom, R. Swanstrom, Characterizing HIV-1 Splicing by Using Next-Generation
858 Sequencing. *J Virol* **91** (2017).
859 31. N. Nguyen Quang *et al.*, Dynamic nanopore long-read sequencing analysis of HIV-1 splicing events
860 during the early steps of infection. *Retrovirology* **17**, 25 (2020).
861 32. W. Li *et al.*, Characterizing transcripts of HIV-1 different substrains using direct RNA sequencing.
862 *Heliyon* **10**, e39474 (2024).
863 33. K. D. Meyer *et al.*, Comprehensive analysis of mRNA methylation reveals enrichment in 3' UTRs and
864 near stop codons. *Cell* **149**, 1635-1646 (2012).
865 34. C. Liu *et al.*, Absolute quantification of single-base m(6)A methylation in the mammalian transcriptome
866 using GLORI. *Nat Biotechnol* **41**, 355-366 (2023).
867 35. Y. Zhou *et al.*, m6A sites in the coding region trigger translation-dependent mRNA decay. *Mol Cell* **84**,
868 4576-4593 e4512 (2024).

369 36. H. E. Varmus, Viruses, genes, and cancer. I. The discovery of cellular oncogenes and their role in
370 neoplasia. *Cancer* **55**, 2324-2328 (1985).
371 37. J. S. Smith, S. Y. Kim, M. J. Roth, Analysis of long terminal repeat circle junctions of human
372 immunodeficiency virus type 1. *J Virol* **64**, 6286-6290 (1990).
373 38. S. Jurriaans, A. de Ronde, J. Dekker, J. Goudsmit, M. Cornelissen, Analysis of human
374 immunodeficiency virus type 1 LTR-LTR junctions in peripheral blood mononuclear cells of infected
375 individuals. *J Gen Virol* **73 (Pt 6)**, 1537-1541 (1992).
376 39. T. Mishra *et al.*, Epitranscriptomic m(6)A modifications during reactivation of HIV-1 latency in CD4(+) T
377 cells. *mBio* 10.1128/mbio.02214-24, e0221424 (2024).
378 40. B. Slobodin *et al.*, Transcription Impacts the Efficiency of mRNA Translation via Co-transcriptional N6-
379 adenosine Methylation. *Cell* **169**, 326-337 e312 (2017).
380 41. N. Meiser, N. Mench, M. Hengesbach, RNA secondary structure dependence in METTL3-METTL14
381 mRNA methylation is modulated by the N-terminal domain of METTL3. *Biol Chem* **402**, 89-98 (2020).
382 42. J. M. Coffin, S. H. Hughes, H. E. Varmus, "The Interactions of Retroviruses and their Hosts" in
383 Retroviruses, J. M. Coffin, S. H. Hughes, H. E. Varmus, Eds. (Cold Spring Harbor (NY), 1997).
384 43. A. Brussel, P. Sonigo, Evidence for gene expression by unintegrated human immunodeficiency virus
385 type 1 DNA species. *J Virol* **78**, 11263-11271 (2004).
386 44. J. Kelly *et al.*, Human macrophages support persistent transcription from unintegrated HIV-1 DNA.
387 *Virology* **372**, 300-312 (2008).
388 45. R. D. Sloan, M. A. Wainberg, The role of unintegrated DNA in HIV infection. *Retrovirology* **8**, 52 (2011).
389 46. S. Thierry *et al.*, Opposite transcriptional regulation of integrated vs unintegrated HIV genomes by the
390 NF-kappaB pathway. *Sci Rep* **6**, 25678 (2016).
391 47. C. Richetta *et al.*, Two-long terminal repeat (LTR) DNA circles are a substrate for HIV-1 integrase. *J*
392 *Biol Chem* **294**, 8286-8295 (2019).
393 48. X. Wei *et al.*, HIV Replication Under High-Level Cabotegravir Is Associated with the Appearance of 3'-
394 PPT Mutations, Circular DNA Transcription and Recombination. *Viruses* **16** (2024).
395 49. A. Uzonyi *et al.*, Exclusion of m6A from splice-site proximal regions by the exon junction complex
396 dictates m6A topologies and mRNA stability. *Mol Cell* **83**, 237-251 e237 (2023).
397 50. P. C. He *et al.*, Exon architecture controls mRNA m(6)A suppression and gene expression. *Science*
398 **379**, 677-682 (2023).

899 51. P. Tang *et al.*, Nuclear retention coupled with sequential polyadenylation dictates post-transcriptional
900 m(6)A modification in the nucleus. *Mol Cell* **84**, 3758-3774 e3710 (2024).
901 52. L. Bertrand *et al.*, Unveiling conserved HIV-1 open reading frames encoding T cell antigens using
902 ribosome profiling. *Nat Commun* **16**, 1707 (2025).
903 53. F. Bouchonnet *et al.*, Quantification of the effects on viral DNA synthesis of reverse transcriptase
904 mutations conferring human immunodeficiency virus type 1 resistance to nucleoside analogues. *J Virol*
905 **79**, 812-822 (2005).
906 54. J. M. Watts *et al.*, Architecture and secondary structure of an entire HIV-1 RNA genome. *Nature* **460**,
907 711-716 (2009).
908