Journal of Visualized Experiments

Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells
--Manuscript Draft--

Darja Pollpeter, PhD
Department of Infectious Diseases
School of Immunology &
Microbial Sciences

KING'S
College
LONDON

20th August, 2018

Vineeta Bajaj, PhD
Review Editor
JoVE

Re: Revised submission of invited manuscript "Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells"
by Pollpeter et al.

Dear Vineeta,

Please find attached our revised manuscript describing a deep sequencing method to identify the 3'-termini and mutational profiles of single stranded DNA molecules. We thank the editor and the reviewers for their helpful comments, which we have addressed individually in our rebuttal letter. We believe the revisions have further improved the manuscript.

Thank you for considering this method for the Journal of Visualized Experiments, and we look forward to hearing from you.

With best wishes,

Darja Pollpeter

www.kcl.ac.uk

1   **TITLE:**
2   Determining 3'-Termini and Sequences of Nascent Single-Stranded Viral DNA Molecules During
3   HIV-1 Reverse Transcription in Infected Cells
4
5   **AUTHORS AND AFFILIATIONS:**
6   Darja Pollpeter[1], Andrew Sobala[1], Michael H. Malim[1]
7
8   [1]Department of Infectious Diseases, School of Immunology & Microbial Sciences, King's College
9   London, United Kingdom
10
11 **Corresponding Author:**
12 Darja Pollpeter   (darja.pollpeter@kcl.ac.uk)
13 +44 (0)2071885398
14
15 **Email Addresses of Co-authors:**
16 Michael Malim   (michael.malim@kcl.ac.uk)
17 Andrew Sobala   (andrew@aes.me.uk)
18

19 **KEYWORDS:**
20 3'-terminus mapping, DNA polymerization, viral replication intermediates, HIV-1 reverse
21 transcripts, unbiased ssDNA ligation, adaptor ligation, deep sequencing
22

23 **SUMMARY:**
24 Here we present a deep sequencing approach that provides an unbiased determination of
25 nascent 3'-termini as well as mutational profiles of single-stranded DNA molecules. The main
26 application is the characterization of nascent retroviral complementary DNAs (cDNAs), the
27 intermediates generated during the process of retroviral reverse transcription.
28

29 **ABSTRACT:**
30 Monitoring of nucleic acid intermediates during virus replication provides insights into the effects
31 and mechanisms of action of antiviral compounds and host cell proteins on viral DNA synthesis.
32 Here we address the lack of a cell-based, high-coverage, and high-resolution assay that is capable
33 of defining retroviral reverse transcription intermediates within the physiological context of virus
34 infection. The described method captures the 3'-termini of nascent complementary DNA (cDNA)
35 molecules within HIV-1 infected cells at single nucleotide resolution. The protocol involves
36 harvesting of whole cell DNA, targeted enrichment of viral DNA *via* hybrid capture, adaptor
37 ligation, size fractionation by gel purification, PCR amplification, deep sequencing, and data
38 analysis. A key step is the efficient and unbiased ligation of adaptor molecules to open 3'-DNA
39 termini. Application of the described method determines the abundance of reverse transcripts
40 of each particular length in a given sample. It also provides information about the (internal)
41 sequence variation in reverse transcripts and thereby any potential mutations. In general, the
42 assay is suitable for any questions relating to DNA 3'-extension, provided that the template
43 sequence is known.
44

INTRODUCTION:

In order to dissect and understand viral replication fully, increasingly refined techniques that capture replication intermediates are required. In particular, the precise definition of viral nucleic acid species within the context of infected cells can provide new insights, since many viral replication mechanisms have to date been examined in isolated *in vitro* reactions. A prime example is the process of reverse transcription in retroviruses, such as human immunodeficiency virus 1 (HIV-1). The various steps of HIV-1 reverse transcription, during which the viral enzyme reverse transcriptase (RT) copies the single-stranded RNA genome into double-stranded DNA, have been studied primarily in primer extension assays with purified proteins and nucleic acids[1-5]. While fundamental principles were established, such assays do not incorporate all viral and cellular components and may not reflect biologically relevant stoichiometries of involved factors. Therefore, we designed a powerful technique to determine the spectra of reverse transcription intermediates with their precise cDNA 3'-termini (*i.e.,* determining their exact lengths) and nucleotide sequences in the context of infections of living cells[6]. Collection of data from time course experiments can be utilized to compare the profile of transcripts under various conditions, such as the presence of antiviral molecules or proteins, that may influence the efficiency and processivity of DNA synthesis and accumulation. This allows a more detailed understanding of the natural pathogen life cycle, which is often the basis for targeted drug design and successful therapeutic intervention.

HIV-1 reverse transcription comprises a series of successive events initiated by annealing of a tRNA primer to the genomic RNA template, which is then extended by RT to produce a short single-stranded cDNA transcript called a minus-strand strong-stop (-sss) (see **Figure 1**). Subsequently, the -sss cDNA is transferred from the 5'-long terminal repeat (LTR) to the 3'-LTR of the genomic RNA, where it anneals and serves as the primer for continued RT mediated elongation of the minus strand DNA (see reviews on reverse transcription[1-4]). This first strand transfer is one of the rate-limiting steps of reverse transcription; hence, the -sss cDNA is known to accumulate. The basic workflow and library design to capture the reverse transcription products in infected cells is outlined in **Figure 2a**. The specific primers and analysis settings that are used in the protocol and listed in **Table 1** target all early reverse transcription cDNA intermediates within the length range of 23 to ~650 nt, which includes the 180-182 nt -sss DNA. However, appropriate minor adaptations to the strategy will allow application to not only late reverse transcription products but also other viruses and systems, where the objective is to detect 3'-OH containing DNA ends. Important limitations to consider include the length range of the final PCR product in the library; in particular, templates in which the distance between the adaptor on the open 3'-terminus and the upstream primer site exceed ~1000 nt will likely be less efficiently sequenced, potentially introducing misleading technical biases during library preparation (see discussion for more details and adaptation suggestions).

Previously reported techniques for the systematic determination of 3'-termini of nucleic acid strands have focused on RNA, not DNA, molecules. One example is 3' RACE (rapid amplification of cDNA ends)[7], which relies on the polyadenylation of mRNA. In addition, adaptor ligation-based strategies employing RNA ligases were developed, which have included RLM-RACE (RNA ligase-mediated RACE)[8] or LACE (ligation-based amplification of cDNA ends)[9]. It is important to

89 emphasize that ligation-based amplifications are sensitive to any bias introduced by the ligation
90 reaction itself. For example, ligation may be more or less efficient depending on a particular
91 nucleotide in the 3' position, the sequence, total molecule length, or local structure. Such ligase
92 preferences lead to incomplete capture of molecules and misrepresentation in the readout,
93 which we and others have observed[9,10]. To minimize ligation bias during the adaptor addition
94 steps in the protocol described herein, we tested a number of ligation strategies and found the
95 use of T4 DNA ligase with a hairpin single-stranded DNA adaptor (as described by Kwok *et al*.[11])
96 to be the only procedure with near quantitative ligation that did not result in significant
97 differences in ligation efficiency when assessed with a specifically selected set of control
98 oligonucleotides[6]. The choice of this ligation strategy is, therefore, a key feature in the success of
99 this protocol.
100
101 To date, monitoring of HIV-1 RT progression in infected cells has primarily been accomplished by
102 measuring reverse transcription products of various length with quantitative PCR (qPCR) using
103 primer-probe sets that uniquely measure shorter or longer (early and late, respectively) cDNA
104 products[12-14]. While this qPCR approach is appropriate to determine intrinsic efficiencies of the
105 reverse transcription process in cellular systems, the output is of relatively low resolution, with
106 no sequence information being derived. Our new approach, based on optimized adaptor ligation,
107 PCR-mediated library generation, and deep sequencing addresses the technology gap and offers
108 an opportunity to monitor reverse transcription during HIV-1 infection quantitatively and at
109 single nucleotide resolution.
110
111 We have illustrated the utility of this method in a study that distinguished between two proposed
112 models for the capacity of the HIV-1 restriction factor APOBEC3G (apolipoprotein B mRNA editing
113 enzyme catalytic polypeptide-like 3G) to interfere with the production of viral reverse
114 transcripts[6].
115
116 **PROTOCOL:**
117 Note: Please refer to the **Table of Materials** for specific reagents and equipment used in this
118 protocol.
119
120 **1. Virus Production and Cell Infection**
121
122 CAUTION: Infectious HIV-1 should only be handled in approved biosafety containment
123 laboratories.
124
125 Note: Production of HIV-1 particles by transient transfection of human embryonic kidney (HEK)
126 293T cells, as outlined in step 1.1, is a standard procedure and has been described previously[15,16].
127 General cell culture procedures are described previously[17].
128
129 1.1) HIV-1 virus production.
130
131 1.1.1 Maintain 293T cells in Dulbecco's modified Eagle medium (DMEM) supplemented with 10%
132 fetal bovine serum (FBS) and 1% penicillin/streptomycin (full DMEM) in a standard cell culture

133   incubator at 37 °C and 5% CO_2 as described previously[17].

134

135   1.1.2 In a standard laminar flow tissue culture hood remove the growth media and add 3 mL of
136   pre-warmed (37 °C) trypsin to a near-confluent 10 cm cell culture dish (~1.2×10^7 cells) of 293T
137   cells. Put the dish back in the incubator for 2-3 min.

138

139   1.1.3 Take the dish from the incubator back into the tissue culture hood and add 7 mL of full
140   medium. Pipette up and down within the dish several times to resuspend the cells. Split the cells
141   1:4 by adding 2.5 mL of the cell suspension to a new 10 cm dish and fill it with 7.5 mL of full
142   medium.

143

144   1.1.4 The next day, mix 10 µg of proviral HIV-1 plasmid DNA (such as pNL4.3) with 1 mL of serum-
145   free minimal essential medium and add polyethylenimine (PEI) solution (25,000 mW, 1 mg/mL
146   pH 7) at 4.5 µL per 1 µg DNA. Incubate for 10 min at RT and add dropwise to the 293T cells.

147

148   1.1.5 24 h after transfection, remove the medium and replace it with 6 mL of full DMEM
149   containing RNase-free DNase at 20 U/mL medium. After 6 h, replace the medium with 10 mL of
150   full DMEM.

151

152   1.1.6 48 h after transfection, harvest the supernatant and filter it through a 0.22 µm filter, using
153   a 10 mL syringe, into a 15 mL polypropylene tube.

154

155   1.1.7 Add 2 mL of sterile 20% sucrose in 1x phosphate buffered saline (PBS) to an open-top thin-
156   walled ultracentrifuge tube. Slowly overlay the sucrose with the filtered cell supernatant.

157

158   1.1.8 Centrifuge for 1 h 15 min at 134,000 x g at 4 °C using an ultracentrifuge.

159

160   1.1.9 Remove the tubes from the ultracentrifuge carefully. Slowly take off both the supernatant
161   and sucrose using a suction or a pipette. Use a smaller pipette and tilt the tube when taking the
162   last sucrose solution out. Leave the pelleted virus in the bottom of the tube.

163

164   Note: The pellet will not be visible.

165

166   1.1.10 Add 200 µL of 1x PBS, leave in the fridge for 4 to 12 h, resuspend, and freeze in 20 µL
167   aliquots at -80 °C.

168

169   1.1.11 Determine p24^{Gag} content using the a p24 HIV-1 antigen ELISA kit (following
170   manufacturer's instructions).

171

172   1.2) T-cell line infection.

173

174   1.2.1 Culture an immortalized T-cell line (*e.g.*, CEM-SS cells) in Roswell Park Memorial Institute
175   (RPMI) 1640 medium supplemented with 10% FBS and 1% penicillin/streptomycin (full RPMI).

176 Count the cells using a hemocytometer[18] and seed 1 well per sample with 1 mL of full RPMI
177 with 2 x 10^6 cells/mL in a 12-well format cell culture plate.

179 1.2.2 Add HIV-1 particles equivalent to 150 ng p24^{Gag} and place the plate into a swinging
180 centrifuge bucket with biocontainment lids to spin-infect by centrifugation in a benchtop
181 centrifuge for 2 h at 2,000 x g at 30 °C.

183 1.2.3 Remove plate from the centrifuge and let it rest for 1 h in a standard tissue culture incubator
184 at 37 °C and 5% CO_2.

186 1.2.4 To wash off input virus, collect cells by transferring the cell suspensions to microcentrifuge
187 tubes and centrifuging in a microcentrifuge at RT (RT) at 500 x g for 2 min. Take off the
188 supernatant without disturbing the cell pellet.

190 1.2.5 Resuspend the cell pellets in 1 mL of pre-warmed (37 °C), sterile 1x PBS. Repeat the
191 centrifugation, supernatant removal, and resuspension steps two more times.

193 1.2.6 Centrifuge again, remove the supernatant and resuspend the cell pellets in 1 mL of full
194 RPMI. Add each suspension to one well in a new 12 well plate.

196 1.2.6 At 6 h post-initial addition of virus (4 h post-centrifugation), harvest the cells by
197 centrifugation as done in step 1.2.4. Remove and discard the supernatant. The cell pellets can be
198 frozen at -80 °C or processed directly for DNA extraction.

200 **2. DNA extraction, HIV-1 DNA Quantification, and Enrichment by Hybrid Capture**

202 2.1) Extract whole cell DNA with a blood and tissue total DNA extraction kit by following the kit
203 manual for tissue culture cells. The only change is elution in 200 µL of nuclease-free H_2O instead
204 of the provided elution buffer.

206 Note: After addition of the chaotropic lysis buffer (the kit's "AL buffer") and proteinase, samples
207 can be removed from the biosafety containment laboratory and handled in a standard safety
208 level laboratory for the rest of the protocol.

210 2.2) Determine the copy number of HIV-1 cDNA by qPCR.

212 2.2.1 Take 17 µL of the eluate from step 2.1 and add 2 µL of 10x restriction enzyme buffer
213 together with 1 µL of DpnI restriction enzyme. Incubate for 1 h at 37 °C to remove any potential
214 residual input plasmid DNA from transfection.

216 2.2.2 Carry out qPCR for minus-strand strong-stop cDNA using the following primer probe set:
217 oHC64 (5'-taactagggaacccactgc-3') and oHC65 (5'-gctagagattttccacactg-3') and probe oHC66 (5'-
218 FAM- acacaacagacgggcacacacta-TAMRA-3'). qPCR setup and exact conditions can be found in the
219 references[6,13]. Carry along samples with a serial dilution of pNL4.3 proviral plasmid as a standard

220   curve to determine copy numbers of cDNA molecules.
221
222   Note: See discussion for expected quantities.
223
224   2.3) HIV-1 DNA enrichment by hybrid capture.
225
226   Note: From this step forward, it is preferable to use microcentrifuge tubes with low nucleic acid
227   binding properties as well as aerosol filter pipette tips for all DNA samples. If possible, work in a
228   PCR workstation. All steps and reagents are at RT (RT) unless otherwise stated.
229
230   2.3.1 To prepare a master mix of magnetic streptavidin beads, pipette 100 µL beads per sample
231   into a single microcentrifuge tube. Place the tube on a magnet suitable for microcentrifuge tubes.
232
233   2.3.2 After the beads have settled towards the magnet side of the tube (~1 min), take off the
234   storage buffer, remove the tube from the magnet and resuspend beads in 500 µL of bind and
235   wash buffer (BW buffer, 5 mM Tris-HCL pH 7.5, 0.5 mM EDTA, 1 M NaCl) to wash.
236
237   2.3.3 Place the tube back on magnet, remove the supernatant and add 500 µL casein solution.
238   Take of the magnet, resuspend and incubate for 10 min at RT, then wash with BW buffer.
239
240   Note: A wash refers to placing the tube on the magnet, taking off the supernatant, taking the
241   tube off the magnet, adding the buffer, and resuspending.
242
243   2.3.4 Place tube back on the magnet, take off the supernatant and resuspend beads in 500 µL of
244   BW buffer. Add 50 pmol of each capturing biotinylated oligonucleotides (see **Table 1,** three oligos
245   in this case) per sample. (For example, if 5 DNA samples are to be processed, use 500 µL of
246   magnetic beads from step 2.3.1 and 250 pmol of each oligonucleotide).
247
248   2.3.5 Incubate for 30 min at RT while rocking in an end-over-end mixer.
249
250   2.3.6 Wash beads with the immobilized oligonucleotides two times with 500 µL of 1x TEN buffer
251   (10 mM Tris HCl pH 8.0, 1 mM EDTA, 100 mM NaCl).
252
253   2.3.7 Resuspend beads in 10 µL of 1x TEN buffer per sample.
254
255   2.3.8 For each sample label one microcentrifuge tube and add 10 µL of beads suspension, 170 µL
256   of DNA (from step 2.1) and 90 µL of 3x TEN buffer. Incubate in a dry heat block at 92 °C for 2 min
257   to denature the DNA.
258
259   2.3.9 Move the tubes to a different dry heat block, which is set to 52 °C, and incubate for 1 h.
260   Invert to mix regularly (~every 10 min) during this incubation.
261
262   2.3.10 Wash once with 500 µL of 1x TEN buffer and resuspend in 35 µL nuclease-free H_2O.
263

264 2.3.11 To elute, incubate the tubes at 92 °C in a dry heat block for 2 min. Then, quickly move the
265 tubes onto the magnet (one tube at a time). Once beads are bound to the side of the tube,
266 transfer the supernatant containing the HIV-1 DNA to a fresh tube.

268 2.3.12 Optional: Repeat qPCR (as done in step 2.2.2) to determine the recovered HIV-1 cDNA.

270 Note: See discussion for expected quantities.

272 **3. Adaptor Ligation**

274 3.1) Preparing the adaptor

276 3.1.2 Resuspend lyophilized adaptor (see **Table 1** "full Kwok + MiSeq") at 100 µM in nuclease-
277 free H_2O.

279 3.1.2 Per sample plus one control sample, combine 0.45 µL of 10x T4 DNA ligase buffer, 4 µL of
280 adaptor, and 0.05 µL of nuclease-free H_2O. Heat to 92 °C for 2 min and let each cool down slowly.

282 Note: If the option is available use a PCR machine with adjustable cooling rate (use 2% rate). This
283 takes about 30 min from 92 °C to 16 °C. Alternatively, use a dry heat block at 92 °C and turn off.
284 Take the adaptor mastermix out when the heat block is back at RT. This is to let the adaptor form
285 a hairpin structure (see **Figure 2b**).

287 3.2) Prepare a control reaction with a set of synthesized oligonucleotides (see **Table 2**) instead of
288 DNA extracted from cells.

290 3.2.1 Make stocks of 100 µM of each oligonucleotide. Mix 1 µL of each of the 17 oligonucelotides
291 and add 8 µL of H_2O for an equimolar ratio in a final volume of 25 µL.

293 3.2.2 Dilute the mix 1:2,500 in nuclease-free H_2O in a serial dilution. Combine 1 µL of the mix with
294 17.3 µL of nuclease-free H_2O to use in the control sample ligation in step 3.3.1 so that each
295 oligonucleotide is present at 1.6 fmol (equivalent to 0.026 nM in the 60 µL reaction).

297 3.3) Setting up ligations

299 3.3.1 For 60 µL final volume reactions in PCR tubes combine 6 µL of 10x T4 DNA ligase buffer, 24
300 µL of 40% PEG, 6 µL of 5 M betaine, 4.5 µL (400 pmol) of adaptor (pre-annelead as in step 3.1.2),
301 1.2 µL of T4 DNA ligase (2,000,000 Units/mL) and 18.3 µL of DNA (from step 2.3.11)

303 Note: Take special care with viscous solutions such as 40 % PEG to maintain accurate volumes.
304 Do not make a mastermix.

306 3.3.2 Set up the same reaction as performed in step 3.3.1 but with the control oligonucleotide
307 mix prepared in step 3.2.2.

3.3.3) Mix the reactions well and incubate in a PCR machine at 16 °C overnight.

4. Adaptor Removal and Size Separation

4.1) Denaturing gel electrophoresis

4.1.1 Add 30 µL of formamide-containing DNA gel loading buffer to each ligation reaction. Mix well by pipetting.

4.1.2 Heat for 2 min at 94 °C in PCR machine, then immediately put on ice.

4.1.3 Place a precast 6% Tris/borate/EDTA (TBE) denaturing urea polyacrylamide gel (10-well comb) in an appropriate gel tank. Add 1x TBE (89 mM Tris-base, 89 mM boric acid, 2 mM EDTA) running buffer and pre-run the gel for 20 min at 250 V/max constant.

4.1.4 Wash out the gel pockets with running buffer using a syringe and 21G needle.

4.1.5 Load each 90 µL sample into three wells (30 µL per well) and run for 20 min (250 V/max) until the dark-blue dye front is about halfway through the gel.

4.2) Staining and cutting nucleic acids from the gel

4.2.1 Prepare 3 small microcentrifuge tubes (0.5 mL) per sample by poking holes into the bottom using a 21 G syringe needle (take caution while working with sharps). Insert each of the prepared tubes into a 2.0 mL microcentrifuge tube and label them with the sample name plus "low", "mid" or "high".

4.2.2 Take out and pry open the gel cassette. Cut the gel vertically with a razor blade to generously excise the strip with the 3 wells of loaded samples. Add the gel strip to a container with 1x TBE (about 30 mL) and 5 µL of SYBR Gold stain. Incubate for 3-5 min.

Note: The gel extraction step is particularly sensitive to cross-contamination. It is advisable to only run 1 sample per gel and using a separate, clean container for each gel staining. Gloves should be changed if gel particles contact gloved fingers.

4.2.3 Clean the surface of a blue light transilluminator thoroughly with ddH$_2$O. Take the gel piece out of the fluorescent green staining container and add it to the light box.

4.2.4 Turn on the light box and inspect the stained nucleic acids through the orange filter.

Note: The adaptor typically appears overloaded and runs as big "blob" with ligated HIV-1 DNA running above as a streak.

4.2.5 Using a new razor blade, cut away the sides of the gel if there are areas with no sample loaded still present. Next, cut just above the adaptor to remove the adaptor and lower gel parts. Finally, cut away the very top of the gel including about 1 mm of the gel pockets, which often have a sharp intense signal of higher molecular weight DNA.

4.2.6 Divide the remaining gel piece containing the sample, which is typically ~2 x 2 cm in size, horizontally into three even pieces: "low", "mid" and "high" molecular weight areas.

Note: Each piece will now be handled separately [*i.e.*, there will be three tubes (low, mid, and high)] per original sample.

4.2.7 Cut each of the three gel fragments into smaller pieces (~2 x 2 mm particles) and transfer them into the prepped 0.5 mL microcentrifuge tubes (step 4.2.1).

4.2.8 Spin at top speed with open lids for 1 min to squeeze the gel pieces through the hole into the 2 mL tube to create a gel slush. If any gel particles remain in the bottom of the 0.5 mL tube, transfer them to the 2 mL tube manually using a needle or pipette tip.

4.3) DNA extraction

4.3.1 Add 1 mL of urea gel extraction buffer (0.5 M $NH_4CH_3CO_2$, 1 mM EDTA, 0.2% SDS) to the gel slush. Rotate tubes for a minimum of 3 h (overnight is acceptable) at RT with an end-over-end mixer.

4.3.2 Use a clean set of tweezers to add one small round glass fiber filter to centrifuge columns with cellulose acetate membrane filters (0.2 µm), which avoids membrane clogging.

4.3.3 Briefly spin the 2 mL tubes with gel slush and extraction buffer in a microcentrifuge and transfer 700 µL of the supernatant to the prepared filter columns. Keep the gel slush and remaining supernatant.

4.3.4 Centrifuge the filter columns in a microcentrifuge at top speed for 1 min. Transfer the flowthrough into a new 2.0 mL microcentrifuge tube.

4.3.5 Reload the columns with the remaining supernatant. Try to obtain as much liquid as possible from the extraction slush. Transfer of gel pieces is not a concern. Spin again and combine flowthroughs of the same extraction samples.

4.4) DNA precipitation

4.4.1 Add 3 µL of polyA RNA (1 µg/µL; as a carrier), 1 µL of glycogen, and 0.7 mL of isopropanol to the flowthrough from step 4.3.5. Vortex briefly and freeze at -80 °C overnight.

4.4.2 Take samples out of the -80 °C freezer and let them thaw briefly. Put them in a cooled (4

396   °C) microcentrifuge and spin for 30 min at top speed.

398   4.4.3 Remove and discard the supernatant. Be very careful not to remove the pellet. Leave 30 to
399   50 µL of liquid if it is uncertain that pellet would be removed otherwise.

401   Note: Typically all "high" samples show a more visible pellet than "mid" and "low" samples.

403   4.4.4 Add 800 µL of 80% ethanol. Invert tubes and spin again for 1 min at top speed. Remove the
404   majority of ethanol with a pipette, briefly spin the tubes again, and remove more ethanol with a
405   smaller volume pipette.

407   4.4.5 Let any remaining ethanol evaporate by placing the tubes with an open lid into a 55 °C dry
408   heat block. When samples are dry (2-4 min) add 20 µL of nuclease-free H_2O and spread around
409   the bottom of the tube to ensure the DNA pellet is dissolved. The DNA sample can be stored at
410   -20 °C.

412   **5. PCR Amplification and Library Preparation**

414   5.1) Set up a 40 µL PCR reaction with 20 µL of DNA polymerase pre-mix, 18 µL of precipitated and
415   redissolved DNA from step 4.4.5, 1 µL of forward primer "MP1.0 + 22HIV" (10µM) (see **Table 1**),
416   and 1 µL of multiplex oligo primers (index primers 1 to 24) (see **Table of Materials**).

418   Note: Run the three reactions (low, mid, high) of each sample in separate PCR reactions, but with
419   the same indexed primer. Use a different index for each of the original infection samples.

421   5.1.1 Run the PCR reactions under the following conditions: 2 min at 94 °C denaturation, then 18
422   cycles of 3-step PCR; 15 s at 94 °C denaturation, 15 s annealing at 55 °C, and 30 s extension at 68
423   °C.

425   5.2) As a quality control option, analyze PCR reactions with high sensitivity automated gel
426   electrophoresis system. Take 2 µL of a low, mid, and high sample to run as per manufacturer's
427   instruction.

429   Note: The two primers should be visible and often run at a calculated length of about 45 and 95
430   nt (actual length differs). Additionally, DNA should be detected between 150 to 500 nt. If no signal
431   is present, it is advisable to add additional PCR cycles, between 2 and 10 additional cycles. Do not
432   add additional cycles for the oligonucleotide control samples created in step 3.3.2.

434   5.3) To remove the primers use a paramagnetic bead-based PCR clean-up system.

436   5.3.1 Take 20 µL of each PCR reaction and pool the samples together (mix all the samples at this
437   point). Freeze the remaining 20 µL reactions as backups at -20 °C.

439   5.3.2 Let the paramagnetic beads come to RT and mix the pooled PCR reactions with 1.8x the

440   volume of the bead solution. Mix by pipetting and incubate for 5 min.

441

442   Note: As an example, if 4 samples were prepared and each have low, mid, and high reactions, the

443   volume would be 4 x 3 x 20 µL = 240 µL PCR reactions with a 432 µL bead solution.

444

445   5.3.3 Put the tubes on a microcentrifuge tube magnet, let the beads bind for ~1 min, and take off

446   the supernatant to discard. Leave the tubes on the magnet and add 500 µL of 80% ethanol.

447

448   5.3.4 Leave the ethanol for 30 s, then take off thoroughly and let the beads airdry for ~5 min. Add

449   40 µL of nuclease-free H_2O, take the tubes off the magnet, and pipette up and down multiple

450   times.

451

452   5.3.5 Leave the suspension for 5 min. Put the tube back on the magnet, let the beads settle to

453   the side, and transfer the supernatant to a new tube. This is the library. Take a 10 µL aliquot for

454   quality controls and freeze the rest at -20 °C.

455

456   **6. Evaluating the Library**

457

458   6.1) Determine the library quality, concentration, and molarity.

459

460   6.1.1 Use a fluorometric quantification method. Measure 1 µL and 3 µL of the library with a high

461   sensitivity dsDNA assay kit as per manufacturer's instructions.

462

463   Note: Typical concentrations are between 1 and 10 ng/µL.

464

465   6.1.2 Measure the library DNA molecular weight spectrum by high sensitivity automated gel

466   electrophoresis as described above (step 5.2).

467

468   6.1.3 Use the automated gel electrophoresis analysis to determine the average molecular weight

469   of the library and calculate to dilute the library in nuclease-free H_2O to 4 nM. Several libraries can

470   be combined as long as all indices are unique.

471

472   6.2) Optional low throughput quality control

473

474   6.2.1 Subject the DNA library to TA cloning[19] to insert library molecules into vectors for

475   amplification. Follow the kit's instructions, grow out ~10-20 colonies and extract DNA *via*

476   miniprep protocols, as described here[20].

477

478   6.2.2 Sequence vector using local sequencing services and check that the inserts contain the

479   desired HIV-1 derived sequences and library-specific adaptors.

480

481   **7. High-Throughput Sequencing Run**

482

483   7.1) Create a sequencing sample sheet with the commercial software provided with the

484   sequencing platform.

485

486   7.1.1 Indicate the selected sequencing kit. Typically, choose a 150-cycle kit, but others are
487   suitable depending on the desired read length.

488

489   7.1.2 Select "Fastq only" as the application workflow. Choose one of the templates that contains
490   the 24 indices present in the multiplex oligonucleotide kits (indicated in kit manual).

491

492   7.1.2 Select "25 nt" for Read1 and "125 nt" for Read2. Keep 6 nt for single index read.

493

494   Note: In the in-house analysis only Read2 is used in the analysis. Keep Read1 at a minimum of 25
495   nt for sequencing platform algorithm purposes.

496

497   7.2) Follow manufacturer's instructions precisely for the pre-run library preparation and setup.
498   Opt for the maximum 20 pM concentration and use a 15% PhiX spike, as the library is of very low
499   complexity.

500

501   **8. Data Analysis**

502

503   8.1) Check if the pass filter percentage and average Q30 quality score are acceptable as per
504   sequencing platform manufacturer's guidelines.

505

506   Note: Pass filter is typically > 90% and Q30 scores are typically > 80%.

507

508   8.2) Download the .fastq.gz files from the manufacturer's sequencing hub.

509

510   8.3) Setting up the sequencing script

511

512   8.3.1 Create a new directory (folder) named "AnalysisXYZ" and go to
513   https://github.com/malimlab/seqparse to download all source code files (parse_sam.pl,
514   rc_extract.pl, parse.sh) into this directory.

515

516   8.3.2 Download the short-read aligner Bowtie, version 1.1.2, from http://bowtie-
517   bio.sourceforge.net/index.shtml into the same directory.

518

519   8.3.3 The download creates a subdirectory within "AnalysisXYZ" named "Bowtie-1.1.2". Within
520   this directory open subdirectory "Indexes" and download the provided template sequences
521   consisting of 6 files with .ebwt extensions.

522

523   8.3.4 Download the FASTQ/A short reads pre-processing toolkit fastx-0.0.13 from
524   http://hannonlab.cshl.edu/fastx_toolkit/download.html into the "AnalysisXYZ" directory.

525

526   8.3.5 Download both Samtools (https://sourceforge.net/projects/samtools/files/) and bam-
527   readcount (https://github.com/genome/bam-readcount) into the "Documents" directory.

528

529   8.4) Move the .fastq.gz files, downloaded in step 8.2, of all read 2s (ending in …_R2_001.fastq.gz)
530   into the "AnalysisXYZ" directory.

531

532   8.5) Open the command console/terminal. Move to the "AnalysisXYZ" as the current directory
533   using cd commands. Type "./parse.sh ." to run the scripts.

534

535   8.6) Find the .csv files with summaries for all samples on total read counts, length adjusted read
536   counts, and normalizes read counts, as well as files with the base variation, for each sample in a
537   directory named parse_results within the "Analysis XYZ" directory.

538

539   Note: See the discussion for further information about the analysis process. The script returns
540   csv files with total reads for each nucleotide along the HIV-1$_{NL4.3}$ strong-stop sequences and first
541   strand transfer until nucleotide 635. As a guidance, 50,000 to 100,000 unique reads are typically
542   observed in samples from infections with the indicated cell numbers and viral inocula and
543   without antiviral proteins or compounds. The oligonucleotide control sample usually produces
544   100,000 to 200,000 reads.

545

546   **REPRESENTATIVE RESULTS:**
547   The technique described in this article was applied to a wider study to address the mechanisms
548   underlying inhibition of HIV-1 reverse transcription by the antiretroviral human protein
549   APOBEC3G (A3G)[6]. **Figure 3** shows representative results obtained after employing the protocol
550   in samples from CEM-SS T-cells infected with *vif*-deficient HIV-1 in the absence or presence of
551   A3G. The total number of unique reads obtained from each sample after filtering out any PCR
552   duplicates that have the same 6 nt barcode and the same length (performed by the analysis
553   software provided) are plotted in **Figure 3a**. Increasing levels of A3G reduce the total read
554   number reflecting the inhibitory effect of A3G on RT mediated cDNA synthesis previously
555   described and measured by qPCR[6,13,21,22]. In **Figure 3b,** the fraction of molecules at each possible
556   length within the first 182 nt are shown. For HIV-1 infection in the absence of A3G, the most
557   abundant species is the main 180 nt strong-stop molecule itself, with some accumulation of reads
558   in the shorter range (23 to 40 nt) (top graph, blue histograms). The addition of A3G changes this
559   profile as a sharp increase of shorter, truncated cDNA molecules at a few very specific,
560   reproducible positions is detected (middle and lower graphs). Since A3G is a cytidine deaminase,
561   cytosine-to-uridine (identified as C-to-T) mutations in the cDNA occur when A3G is present in the
562   infecting virions[21,23,24]. Using the obtained sequencing information, the percentage of C-to-T
563   mutations was plotted on the same graph (red dotted line). It should be noted that the
564   mutational profile is derived from all unique reads combined and coverage of each nucleotide
565   will vary. However, if required sequence information can be related back to each molecule and
566   correlated with a specific 3'-terminus. The data provided were taken from Pollpeter *et al.*[6] and
567   the correlation between mutational and cDNA length profiles was demonstrated to be due to
568   detection and cleavage of deaminated cDNA by the cellular DNA repair machinery.

569

570   A positive control for the 3'-mapping approach can easily be produced by processing a pool of
571   synthetic oligonucleotides of known sequence, length, and concentration. This control is added

572   at the adaptor ligation in step 3.3.2 and advised to be included in all multiplexed libraries. Data
573   obtained from a control sample should have all the oligonucleotides at the expected input ratios,
574   with only very minor background reads. **Figure 4** shows results of a positive control set of 17
575   chemically synthesized oligonucleotides (for sequences, see **Table 2**), which were mixed at
576   equimolar ratios. As expected, all molecules appear in close to equal abundance with only small
577   variations (top graph). While most positions within the -sss DNA sequence that were not
578   represented by an oligonucleotide return zero read counts, we observed minor species that are
579   1 or 2 nt shorter than the actual control oligonucleotides. We have not further investigated these
580   minor species but assume that they represent degraded or incomplete products potentially
581   present in the provided oligonucleotide stocks at purchase (oligonucleotides were ordered as
582   HPLC purified, for which the manufacturer indicates > 80% purity). The bottom graph shows the
583   control sample from a different library run, where variation is slightly higher between the 17
584   oligonucleotides and correlates with overall length in that longer control molecules are detected
585   more efficiently then shorter ones. This may be due to a minor bias in PCR reactions or in
586   clustering during MiSeq sequencing, which has an insert size optimum and may occur with
587   libraries carrying particularly broad insert ranges. A basic way to address this bias is the
588   application of a normalization factor based on the slope that indicates the bias correlating to
589   molecule length (pink line). The required calculations are included in the analysis program (see
590   step 8.3 in the protocol).
591
592   **FIGURE AND TABLE LEGENDS:**
593   **Figure 1: Diagram showing the first steps of HIV-1 reverse transcription, starting with annealing**
594   **of tRNA(Lys,3) (orange) to the primer binding site (PBS) in the genomic viral RNA (step 1), which**
595   **allows the initiation and elongation of viral cDNA (blue, step 2)**. Concomitantly, the template
596   genomic RNA is degraded by RNaseH activity of RT (step 3). The first full intermediate in the
597   process of reverse transcription is the minus-strand strong-stop (-)sss cDNA, which is complete
598   when the RT catalyzed polymerization reaches the 5'-terminus of the gRNA repeat (R) region
599   (step 3). The (-)sss intermediate is transferred to the 3'-terminus of the genomic RNA template
600   by annealing to the complementary 3'-long terminal repeat (LTR) R region. From here,
601   polymerization continues (step 4). In the described method the reverse transcription progression
602   is determined by mapping the exact length of the nascent viral cDNA (blue). PPT, polypurine tract;
603   U5, unique 5'-sequence; U3, unique 3'-sequence. This figure is republished from a previous
604   publication[6].
605
606   **Figure 2: Workflow outline and schematics of the adapter ligation and PCR amplification**
607   **strategy.** (a) Workflow outlining main steps of the described technique to determine 3'-termini
608   of HIV-1 reverse transcripts in infected cells. The figure is adapted from a previous publication[6].
609   (b) Schematic of the adaptor ligation and PCR amplification strategy. Nascent cDNA molecules of
610   varying length that have been purified in previous steps are ligated to a single-stranded DNA
611   adaptor using T4 DNA ligase. The hairpin adaptor (named "full Kwok + MiSeq", see **Table 1**) design
612   was inspired by Kwok *et al*.[11]. The adaptor carries a random 6 nt barcode sequence, which allows
613   for base-pairing to facilitate ligation and simultaneously serves as an identifier for unique reads.
614   The 3'-termini of the adaptor carries a spacer (SpC3) to prevent self-ligation. Ligated products are
615   separated from excess adaptor by denaturing polyacrylamide gel electrophoresis (PAGE). Nucleic

616   acids in the gel are stained and cut into three separate, equal-size gel pieces in the area from
617   above the adaptor to the well as done in[25]. After elution, precipitation, and resuspension, the
618   products are PCR amplified with primers annealing to the known sequence of the adaptor (primer
619   1, multiplex oligonucleotide kit, see **Table of Materials**) and a primer carrying the first 22 nt of
620   the HIV-1 5'-LTR sequence immediately following the tRNA (primer 2, MP1.0 + 22HIV). The 5'-
621   termini of the chosen primers carry adaptors for the chosen sequencing platform (P5 and P7) as
622   well as an index sequence to distinguish individual samples run in the same library. Starting points
623   of the sequencing read primers are indicated. The blue box indicates the region of interest to
624   determine the original 3'-termini of the captured molecule. This figure is adapted from a previous
625   publication[6].
626
627   **Figure 3: Representative results.** (a) The total read count of representative samples processed
628   with the described protocol. This includes all sequences that were identified as unique reads of
629   HIV-1 molecules with their 3'-termini within the first 635 nt of the minus strand cDNA (up to the
630   PPT, see **Figure 1**). Infection with HIV-1 not carrying A3G yields the highest number of reads,
631   whereas A3G inhibits cDNA synthesis and thereby reduces the total read count. Uninfected cells
632   served as a negative control, while a set of synthetic oligonucleotides provides a positive control.
633   b) The relative abundance of cDNAs for each length between nt positions 23 and 182 (full-length
634   -sss cDNA is 180 to 182 nt) of the HIV-1$_{NL4.3}$ sequence (x-axis) is shown in blue histograms (scale
635   on the left y-axis). The relative abundance of cDNA was calculated from the absolute number of
636   sequences terminating at a given nucleotide within the -sss cDNA sequence divided by the sum
637   of all reads measuring 182nt or less. Shown in dashed red lines are the percentages of reads
638   carrying C-to-T/U mutations at the respective position (scale on the right y-axes). **Figure 3b** is
639   republished from a previous publication[6].
640
641   **Figure 4: Representative results of control samples.** Shown are two profiles for pools containing
642   equimolar amounts of 17 different length synthetic oligonucleotides. These oligonucleotides
643   have sequences from HIV-1$_{NL4.3}$ and were selected to cover various lengths and present all 4 bases
644   as a 3'-nucleotide (see **Table 2**). The top graph shows the positive control sample from **Figure 3a**.
645   No significant bias towards molecule length or the open 3'-termini is detected. The bottom graph
646   shows a different library run, which produced a minor length bias in sequencing. In this case, it is
647   advisable to apply a normalization factor, which is derived from the slope (shown in pink) that
648   represents the size bias. This figure is republished from a previous publication[6].
649
650   **Table 1: Table of oligonucleotides including length, sequences, and modifications that are**
651   **utilized in the described protocol**. The table is adapted from a previous publication[6].
652
653   **Table 2: Table of 17 synthetic control oligonucleotides used as a positive control sample.** The
654   top 13 oligonucleotides were chosen based on size [long (116 to 120 nt), mid (69 to 85 nt), short
655   (33 to 41 nt)] as well as their 3'-termini. The table is adapted from a previous publication[6].
656
657   **DISCUSSION:**
658   The availability of fast, reliable, and cost-effective deep sequencing has revolutionized many

659 aspects in the field of life sciences, allowing great depth in sequencing-based analyses. A
660 remaining challenge lies in the innovative design and creation of representative sequencing
661 libraries. Here we describe a protocol to capture nascent viral cDNA molecules, specifically the
662 intermediates of the HIV-1 reverse transcription process.

663

664 The most critical step in this strategy is the ligation of an adaptor to the open 3'-termini in a
665 quantitative and unbiased manner. Efficiencies of ligations between two ssDNA termini, both
666 inter- and intramolecular, have been investigated and optimized for various applications[11,26-29].
667 The choice of using a hairpin adaptor with T4 DNA ligase under the conditions described in step
668 3.3 is the result of empirical optimization in which we evaluated different ligases, adaptors and
669 reagents for the ligation of synthetic oligonucleotides representing HIV-1 sequences (**Table 2**)
670 (data not shown). In these *in vitro* test reactions, we confirmed that the T4 DNA ligase mediated
671 ligation of the hairpin adaptor, as described by Kwok *et al.*[11], has a very low bias, and achieves
672 near complete ligation of acceptor molecules when the adaptor is used in excess. The ligation
673 efficiency was unaffected by the addition of nucleotide sequence to render the adaptor
674 compatible for the multiplex primer system (see **Figure 4**). In comparison, we found that a
675 thermostable 5'DNA/RNA ligase ("Ligase A", see **Table of Materials** for exact ligases compared
676 here), which is an engineered RNA ligase that was developed in part to improve on ligation
677 efficiency with ssDNA as the acceptor[27], was indeed more effective at ligating two ssDNA
678 molecules than RNA ligase ("Ligase B") but had a significant bias, with strong differences in
679 ligation efficiency even between oligonucleotides with single base length differences [**Table 2**;
680 HTP con mid G (a) and (b)]. Furthermore, we found only a minimal bias in reactions with "Ligase
681 C" combined with an adaptor carrying a randomized 5'-termini (a strategy used to offset known
682 nucleotide bias of "Ligase C"; see for example Ding *et al.*[30]). However, the "Ligase C"-mediated
683 intermolecular ligations were incomplete, rendering the T4 DNA ligase system the superior
684 choice.

685

686 Several quality control steps over the course of the protocol and the inclusion of positive and
687 negative controls allow for the detection of potential problems before assay continuation and
688 provide guidance for troubleshooting efforts. The qPCR quantifications in steps 2.2.2 and 2.3.12
689 ensure that the quantity of the input material is sufficient. Typical cDNA copy numbers in the 200
690 µL elution (from step 2.1) range from around 10,000 to 300,000 per µL. The hybrid capture step
691 can result in some loss of overall HIV-1 cDNA quantity but should result in a strong enrichment
692 of specific HIV-1 cDNA over cellular DNA, which can be determined by using appropriate primers
693 to quantify genomic DNA before and after enrichment by qPCR or by measuring total DNA
694 concentration. The loss at the hybrid capture steps should be no higher than 50%. Low starting
695 material may otherwise explain a successful oligonucleotide positive control (see step 3.3.2) but
696 only limited reads achieved in the samples. Low read numbers overall could also be explained by
697 overestimation of the library concentration due to the presence of irrelevant DNA species
698 without MiSeq adaptors. This would result in low cluster density and can be improved by
699 determining the concentration of HIV-1 sequences in the library by qPCR in addition to the total
700 DNA amount by fluorometric assays. Due to the highly sensitive nature of the method, special
701 care should be taken to avoid even low-level contamination, both from other samples (in
702 particular, from the high concentration control oligonucleotide stocks) as well as from laboratory

703 equipment. Working in a UV sterilizing PCR workstation is beneficial in this regard. The
704 automated gel electrophoresis of the final library (step 6.1.2) is a further quality control measure.
705 The nucleic acid size range typically observed is between 150 to 500 nt. Primers that can be
706 detected in the optional control after the PCR and before purification (see note in step 5.2) should
707 now be absent. In a representative result, the sample intensity curve has a peak around 160 to
708 170 nt and a second sharper peak around 320 to 350 nt. This likely reflects the often-seen higher
709 abundance in both relatively short (1 to 20 nt insert length) reverse transcripts and full-length
710 strong-stop (180 to 182 nt insert length) (**Figure 3b**).

711

712 While the presented protocol and selected primers are specific for early HIV-1 reverse
713 transcription constructs, the method is generally applicable to any study aiming to determine
714 open 3'-termini of DNA. The main modifications required in other contexts will be the method
715 for hybrid capture and the primer design strategy. For example, if the target is to be adapted to
716 late HIV-1 transcripts, a larger number of different capturing biotinylated oligonucleotides
717 annealing across the length of the cDNA would be advisable and will likely decrease the loss in
718 the hybrid capture step. As mentioned in the introduction, it is important to consider limitations
719 when designing the range over which 3'-termini are to be detected to avoid different sources of
720 bias. First, there may be a bias in the PCR reactions if the templates with the adaptor are of vastly
721 varying length. Second, the sequencing platform used here (*e.g.,* MiSeq) has a preferred insert
722 length range for optimal clustering, and significantly shorter and longer products may not be
723 sequenced with the same efficiency. In part, this can be addressed computationally, as was done
724 by calculating a correction factor for linear length bias (see **Figure 4**, bottom graph). However, if
725 the region of where 3'-termini mapping is desired is long (> 1000 nt), it is more advisable to split
726 the reactions with the ligated transcripts and use multiple upstream primers to assess 3'-termini
727 in sections.

728

729 The analysis program was written in-house for the specific purpose of analyzing both the last
730 nucleotide of the HIV-1 sequence adjacent to the fixed adaptor sequence as well as the base
731 variation of all bases to identify any mutations. The individual steps comprise the following: first,
732 the adaptor sequences are trimmed using the fastx-0.0.13 toolkit; then, any sequences that are
733 duplicated (meaning identical sequences including the barcode) are removed. All remaining
734 unique reads are then aligned to the HIV-1 sequence using Bowtie (http://bowtie-
735 bio.sourceforge.net/index.shtml) with the maximum mismatch set at three bases. The template
736 sequence is comprised of the first 635 nt of HIV-1 cDNA (NL4.3 strain), which includes the -sss
737 sequence and the first strand transfer product up to the polypurine track (U5-R-U3-PPT; see
738 **Figure 1**). Thereby, the provided software and templates are only directly suitable if the method
739 is used for the same application (detection of early reverse transcripts of the HIV-1$_{NL4.3}$).
740 Adjustments will have to be made for other target sequences. The positions of the 3'-termini for
741 each read were determined by the position in the alignment. Base calls for each position are
742 recorded and mutation rates are calculated from the total coverage of each base, which varies,
743 as reads are of different lengths and long inserts may not be entirely covered by the 125-base
744 sequencing in Read2.

745

746 To conclude, we believe the described method to be a valuable tool for many types of studies.

Obvious applications include investigations of the mechanisms underlying reverse transcription inhibition through antiretroviral drugs or cellular restriction factors. However, only relatively minor adjustments should be necessary to adapt the system to 3'-termini mapping within other single-stranded DNA viral intermediates, which are present, for example, in parvovirus replication. Furthermore, the principle of the method, particularly its optimized ligation step, can provide a core part of library preparation design for the characterization of any 3'-DNA extensions, including elongations catalyzed by cellular double-stranded DNA polymerases.

DISCLOSURES:
The authors declare that they have nothing to disclose.

REFERENCES:

1   Herschhorn, A., Hizi, A. Retroviral reverse transcriptases. *Cellular and Molecular Life Sciences.* **67** (16), 2717-2747 (2010).

2   Hu, W. S., Hughes, S. H. HIV-1 reverse transcription. *Cold Spring Harbor Perspectives in Medicine.* **2** (10), (2012).

3   Levin, J. G., Mitra, M., Mascarenhas, A., Musier-Forsyth, K. Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription. *RNA Biology.* **7** (6), 754-774 (2010).

4   Menendez-Arias, L., Sebastian-Martin, A., Alvarez, M. Viral reverse transcriptases. *Virus Research.* (2016).

5   Telesnitsky, A., Goff, S. P. *Retroviruses.* Coffin, J. M., Hughes, S. H., Varmus, H. E., editors. (1997).

6   Pollpeter, D., *et al.* Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G. *Nature Microbiology.* **3** (2), 220-233 (2018).

7   Frohman, M. A., Dush, M. K., Martin, G. R. Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proceedings of the National Academy of Sciences of the United States of America.* **85** (23), 8998-9002 (1988).

8   Liu, X., Gorovsky, M. A. Mapping the 5' and 3' ends of Tetrahymena thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Research.* **21** (21), 4954-4960 (1993).

9   Ince, I. A., Ozcan, K., Vlak, J. M., van Oers, M. M. Temporal classification and mapping of

791   non-polyadenylated transcripts of an invertebrate iridovirus. *Journal of General Virology.*
792   **94** (Pt 1), 187-192 (2013).

793   10   Hafner, M., *et al.* RNA-ligase-dependent biases in miRNA representation in deep-
794   sequenced small RNA cDNA libraries. *RNA.* **17** (9), 1697-1712 (2011).

795   11   Kwok, C. K., Ding, Y., Sherlock, M. E., Assmann, S. M., Bevilacqua, P. C. A hybridization-
796   based approach for quantitative and low-bias single-stranded DNA ligation. *Analytical*
797   *Biochemistry.* **435** (2), 181-186 (2013).

798   12   Abram, M. E., Tsiang, M., White, K. L., Callebaut, C., Miller, M. D. A cell-based strategy to
799   assess intrinsic inhibition efficiencies of HIV-1 reverse transcriptase inhibitors.
800   *Antimicrobial Agents and Chemotherapy.* **59** (2), 838-848 (2015).

801   13   Bishop, K. N., Verma, M., Kim, E. Y., Wolinsky, S. M., Malim, M. H. APOBEC3G inhibits
802   elongation of HIV-1 reverse transcripts. *PLoS Pathogens.* **4** (12), e1000231 (2008).

803   14   Zack, J. A., Haislip, A. M., Krogstad, P., Chen, I. S. Incompletely reverse-transcribed human
804   immunodeficiency virus type 1 genomes in quiescent cells can function as intermediates
805   in the retroviral life cycle. *Journal of Virology.* **66** (3), 1717-1725 (1992).

806   15   Adachi, A., *et al.* Production of acquired immunodeficiency syndrome-associated
807   retrovirus in human and nonhuman cells transfected with an infectious molecular clone.
808   *Journal of Virology.* **59** (2), 284-291 (1986).

809   16   Shah, V. B., Aiken, C. In vitro uncoating of HIV-1 cores. *Journal of Visualized Experiments.*
810   (57), (2011).

811   17   JoVE Science Education Database: Basic Methods in Cellular and Molecular Biology:
812   Passaging Cells. *Journal of Visualized Experiments.* (2018).

813   18   JoVE Science Education Database: Basic Methods in Cellular and Molecular Biology: Using
814   a Hemocytometer to Count Cells. *Journal of Visualized Experiments.* (2018).

815   19   Zhou, M. Y., Gomez-Sanchez, C. E. Universal TA cloning. *Current Issues in*
816   *Molecular Biology.* **2** (1), 1-7 (2000).

817   20   Zhang, S., Cahalan, M. D. Purifying plasmid DNA from bacterial colonies using the QIAGEN
818   Miniprep Kit. *Journal of Visualized Experiments.* (6), 247 (2007).

819   21   Mangeat, B., *et al.* Broad antiretroviral defence by human APOBEC3G through lethal
820   editing of nascent reverse transcripts. *Nature.* **424** (6944), 99-103 (2003).

821   22   Gillick, K., *et al.* Suppression of HIV-1 infection by APOBEC3 proteins in primary human
822   CD4(+) T cells is associated with inhibition of processive reverse transcription as well as
823   excessive cytidine deamination. *Journal of Virology.* **87** (3), 1508-1517 (2013).

824   23   Harris, R. S., *et al.* DNA deamination mediates innate immunity to retroviral infection. *Cell.*
825   **113** (6), 803-809 (2003).

826   24   Zhang, H., *et al.* The cytidine deaminase CEM15 induces hypermutation in newly
827   synthesized HIV-1 DNA. *Nature.* **424** (6944), 94-98 (2003).

828   25   Konig, J., *et al.* iCLIP--transcriptome-wide mapping of protein-RNA interactions with
829   individual nucleotide resolution. *Journal of Visualized Experiments.* (50), (2011).

830   26   Troutt, A. B., McHeyzer-Williams, M. G., Pulendran, B., Nossal, G. J. Ligation-anchored
831   PCR: a simple amplification technique with single-sided specificity. *Proceedings of the*
832   *National Academy of Sciences of the United States of America.* **89** (20), 9823-9825 (1992).

833   27   Zhelkovsky, A. M., McReynolds, L. A. Structure-function analysis of Methanobacterium
834   thermoautotrophicum RNA ligase - engineering a thermostable ATP independent

835 enzyme. *BMC Molecular Biology.* **13,** (24), (2012).

836 28 Li, T. W., Weeks, K. M. Structure-independent and quantitative ligation of single-stranded
837 DNA. *Analytical Biochemistry.* **349** (2), 242-246 (2006).

838 29 Gansauge, M. T., *et al.* Single-stranded DNA library preparation from highly degraded DNA
839 using T4 DNA ligase. *Nucleic Acids Research.* **45** (10), e79 (2017).

840 30 Ding, Y., Kwok, C. K., Tang, Y., Bevilacqua, P. C., Assmann, S. M. Genome-wide profiling of
841 in vivo RNA structure at single-nucleotide resolution using structure-seq. *Nature*
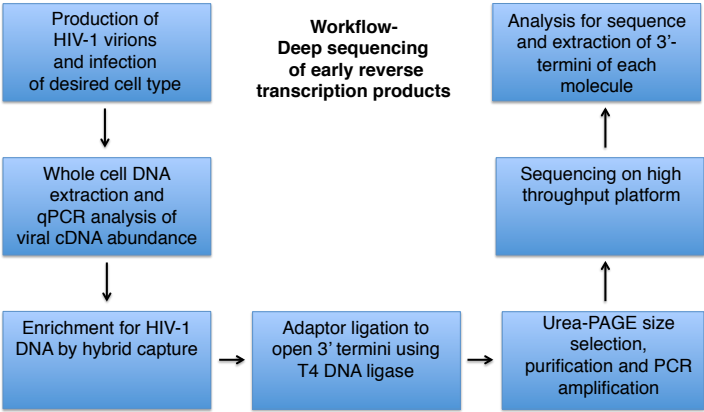842 *Protocols.* **10** (7), 1050-1066 (2015).
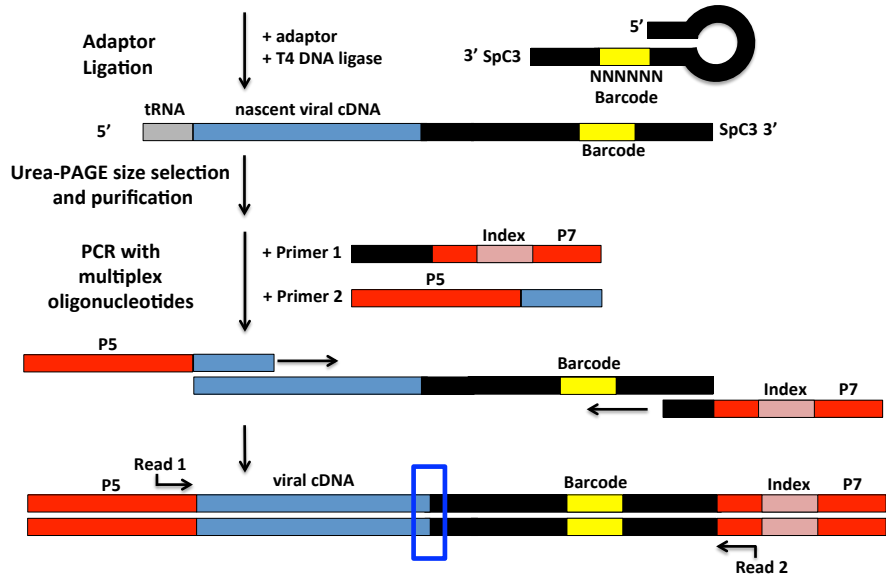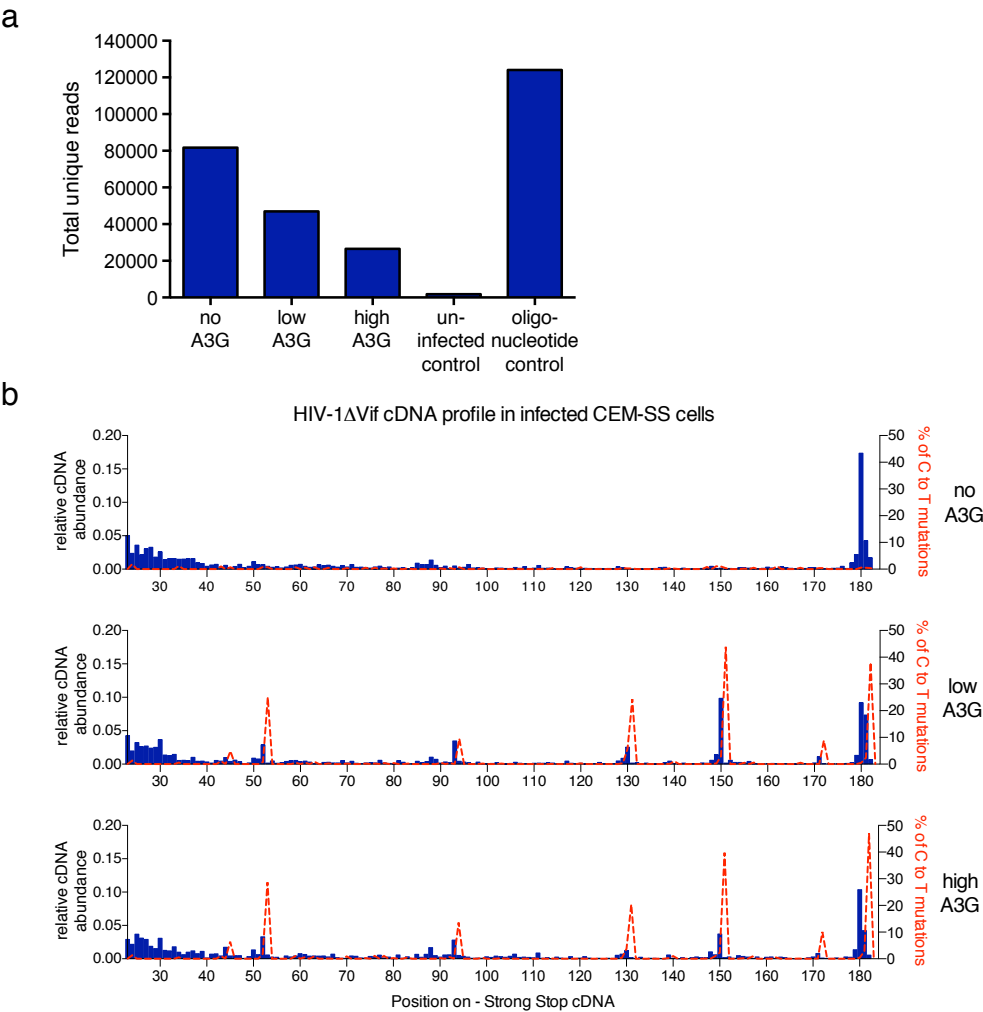
843

Figure 1

Figure 1

First steps of HIV-1 reverse transcription:
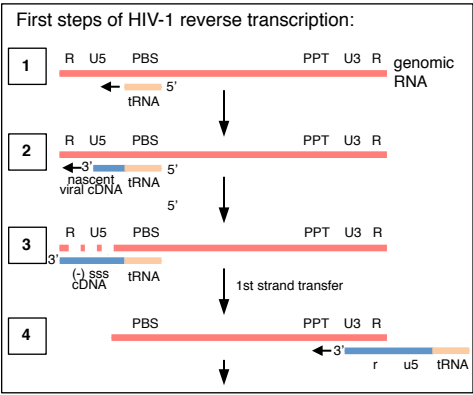
Figure 2

a

b

Figure 3

Figure 3

a

b

HIV-1ΔVif cDNA profile in infected CEM-SS cells

Position on - Strong Stop cDNA

Figure 4

Figure 4

Table 1 revised
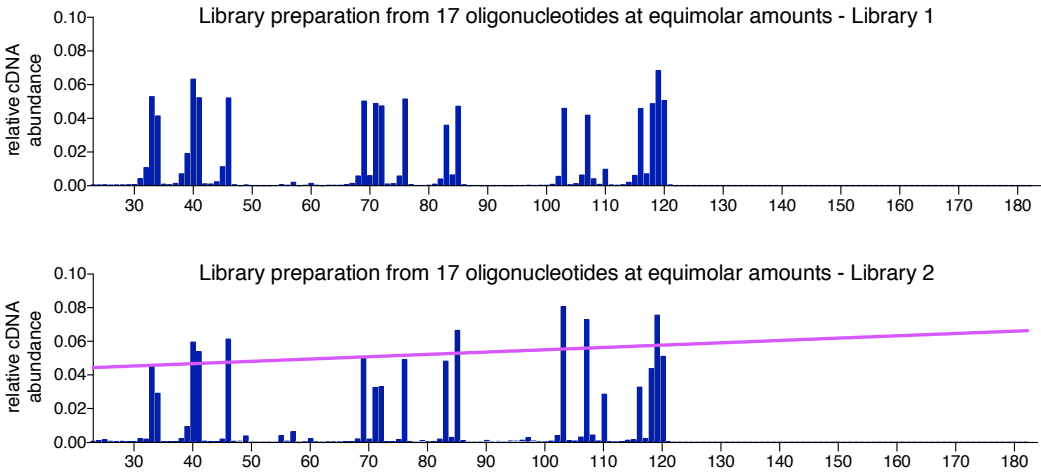
Oligo name	Length in nt	Sequence	Purpose	Manufacturer (Purification)
full Kwok + MiSeq	61	5'-PHO-tgaagagcctagtcgctgttcannnnnnctgcccatagagagatcggaagagcacacgtct-SpC3-3'	Adaptor	IDT DNA technologies (HPLC)
2xBiotin SS bait	40	5'-biotin-cagtgtggaaaatctctagcagtggcgcccgaacagggac-biotin-3'	Hybrid Capture	MWG Eurofins (HPLC)
Biotin 1-16 ss	22	5'-cagtgtggaaaatctctagcag-BiTEG-3'	Hybrid Capture	MWG Eurofins (HPLC
Biotin tRNA + CTG	16	5'-cagtggcgcccgaaca-BITEG-3'	Hybrid Capture	MWG Eurofins (HPLC)
MP1.0 + 22HIV	82	5'-aatgatacggcgaccaccgagatctacactctttccctacacgacgctcttccgatctcactgctagagattttccacactg-3'	PCR amplification	MWG Eurofins (HPLC

Table 2 revised

Oligo name	Length in nt	Sequence	Manufacturer (Purification)
HTP con long C	120	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagctttattgaggcttaagc-3'	MWG Eurofins (HPLC)
HTP con long G	119	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagctttattgaggcttaag-3'	MWG Eurofins (HPLC)
HTP con long T	116	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagctttattgaggctt-3'	MWG Eurofins (HPLC)
HTP con long A	118	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagctttattgaggcttaa-3'	MWG Eurofins (HPLC)
HTP con mid C	76	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcac-3'	MWG Eurofins (HPLC)
HTP con mid G (a)	71	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacg-3'	MWG Eurofins (HPLC)
HTP con mid G (b)	72	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgg-3'	MWG Eurofins (HPLC)
HTP con mid A	69	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacaga-3'	MWG Eurofins (HPLC)
HTP con mid T	85	5'-ctgctagagattttccacactgactaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactt-3'	MWG Eurofins (HPLC)
HTP con short A	40	5'-ctgctagagattttccacactgactaaaagggtctgaggga-3'	MWG Eurofins (HPLC)
HTP con short T	33	5'-ctgctagagattttccacactgactaaaagggt-3'	MWG Eurofins (HPLC)
HTP con short G	41	5'-ctgctagagattttccacactgactaaaagggtctgaggg-3'	MWG Eurofins (HPLC)
HTP con short C	34	5'-ctgctagagattttccacactgactaaaagggtc-3'	MWG Eurofins (HPLC)
HTP Con 46 (T)	46	5'-ctgctagagattttccacactg actaaaagggtctgagggatctct-3'	MWG Eurofins (HPLC)
HTP Con 83 (C)	83	5'-ctgctagagattttccacactg actaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactac-3'	MWG Eurofins (HPLC)
HTP Con 103 (C)	103	5'-ctgctagagattttccacactg actaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagc-3'	MWG Eurofins (HPLC)
HTP Con 107 (A)	107	5'-ctgctagagattttccacactg actaaaagggtctgagggatctctagttaccagagtcacacaacagacgggcacacactactttgagcactcaaggcaagcttta-3'	MWG Eurofins (HPLC)

Table of Materials - revised

Name of Material/ Equipment	Company	Catalog Number	Comments/Description
293T cells	ATCC	CRL-3216	
Dulbecco's Modified Eagle's Medium	Gibco	31966-021	
Penicillin/Streptomycin	Gibco	15150-122	
Fetal Bovine Serum	Gibco	10270-106	
HeraCell Vios 250i CO2 Incubator	Thermo Scientific	51030966	
Laminar flow hood - CAS BioMAT2	Wolflabs	CAS001-C2R-1800	
10mm TC-treated culture dish	Corning	430167	
TrypLE™ Express (1x), Stable Trypsin Replacement Enzyme	Gibco	12605-010	
OptiMEM® (Minimal Essential Medium)	Gibco	31985-047	
HIV-1 NL4-3 Infectious Molecular Clone (pNL4	NIH Aids reagent program	114	
Polyethylenimine (PEI) - MW:25000	PolySciences Inc	23966-2	dissolved at 1mg/ml and adjusted to pH7
RQ1- Rnase free Dnase	Promega	M6101	
Filter 0.22 μm	Triple Red Limited	FPE404025	
15 mL polypropylene tubes	Corning	CLS430791	
Sucrose	Calbiochem	573113	
Phosphate Buffered Saline (1x)	Gibco	14190-094	
Ultracentrifuge tubes	Beckman Coulter	344060	
Ultracentrifuge	Sorval	WX Ultra Series	Th-641 Rotor
Alliance HIV-1 p24 antigen ELISA kit	Perkin Elmer	NEK050001KT	
CEM-SS cells	NIH Aids reagent program	776	
Roswell Park Memorial Institute Medium	Gibco	31870-025	
CoStar® TC treated multiple well plates	Corning	CLS3513-50EA	
Benchtop centrifuge: Heraus™ Multifuge™ X3 FR	Thermo Scientific	75004536	
TX-1000 Swinging Bucket Rotor	Thermo Scientific	75003017	

Microcentrifuge: 5424R	Eppendorf	5404000060	
Total DNA extraction kit (DNeasy Blood and Tissue kit)	Qiagen	69504	
Nuclease free H2O	Ambion	AM9937	
Cutsmart buffer	New England Biolabs (part of DpnI enzyme)	R0176S	
DpnI restriction enzyme	New England Biolabs	R0176S	
Oligonucleotides for qPCR	MWG Eurofins	N/A	HPSF purification
TaqMan PCR Universal Mastermix	Thermo	4304437	
LoBind Eppendorf® tubes	Eppendorf	30108078	
Axygen™ aerosol filter pipette tips, 1000 µL	Fisher Scientific	TF-000-R-S	
Axygen™ aerosol filter pipette tips, 200 µL	Fisher Scientific	TF-200-R-S	
Axygen™ aerosol filter pipette tips, 20 µL	Fisher Scientific	TF-20-R-S	
Axygen™ aerosol filter pipette tips, 10 µL	Fisher Scientific	TF-10-R-S	
PCR clean hood	LabCaire	Model PCR-62	
DynaMag™2-magnet	Thermo	12321D	
Streptavidin MagneSphere® paramagnetic particles	Promega	Z5481	
Casein	Thermo Scientific	37582	
End over end rotator, Revolver™ 360°	Labnet	H5600	
Tris-Base	Fisher Scientific	BP152-5	
Hydrochloric Acid	Sigma	H1758-100ML	
EDTA disodium salt dihydrate	Electran (VWR)	443885J	
Sodium Chloride	Sigma	S3014	
Dri-Block® Analog Block Heater	Techne	UY-36620-13	
PCR tubes and domed caps	Thermo Scientific	AB0266	
PCR machine	Eppendorf	Mastercycler® series	
T4 DNA ligase	New England Biolabs	M0202M	

40% Polyethylene glycol solution (PEG) in H2O, MW: 8000	Sigma	P1458-25ML	
Betaine solution, 5M	Sigma	B0300-1VL	
Gel loading buffer II (formamide buffer)	Thermo Scientific	AM8546G	
Precast 6% TBE urea gels	Invitrogen	EC6865BOX	
Mini cell electrophoresis system	Invitrogen, Novex	XCell SureLock™	
Tris/Borate/EDTA solution (10x)	Fisher Scientific	10031223	
Needle 21 G x1 1/2	VWR	613-2022	
SYBR Gold nucleic acid stain (10000x)	Life Technologies	S11494	
Dark Reader DR46B transilluminator	Fisher Scientific	NC9800797	
Ammonium acetate	Merck	101116	
SDS solution 20% (w/v)	Biorad	161-0418	
Centrifuge tube filter	Appleton Woods	BC591	
Filter Glass Fibre Gf/D 10mm	Whatman (VWR)	512-0427	
polyadenylic acid (polyA) RNA	Sigma	10108626001	
Glycogen, molecular biology grade	Thermo Scientific	R0561	
Isopropanol (2-propanol)	Fisher Scientific	15809665	
Ethanol, molecular biology grade	Fisher Scientific	10041814	
Accuprime™ Supermix I (DNA polymerase premix)	Life Technologies	12342-010	
NEBNext® Multiplex Oligo for Illumina (Index Primer Set 1 and 2)	New England Biolabs	E7335S; E7500S	
Tapestation D1000 Screentape High sensitivity	Agilent Technologies	5067- 5584	
Tapestation D1000 Reagents	Agilent Technologies	5067- 5585	
2200 Tapestation - automated gel electrophoresis system	Agilent Technologies	G2965AA	
Agencourt® AMPure® beads XP	Beckman Coulter	A63880	
Qubit™ dsDNA HS Assay Kit	Invitrogen	Q32851	
Qubit™ 2.0 Fluorometer	Invitrogen	Q32866	
Topo™ TA cloning Kit	Invitrogen	450071	
Sequencing platform: MiSeq System	Illumina		

Experiment Manager (Sample sheet software)	Illumina	Note: Use TruSeq LT as a template	
Miseq™ Reagent kit V3 (150 cycle)	Illumina	MS-102-3001	
Sequencing hub: Basespace	Illumina	https://basespace.illumina.com	
Ligase A: Thermostable 5' App DNA/RNA l	*NEB*	*M0319S*	*Not used in this protocol, but*
Ligase B: T4 RNA ligase 1	*NEB*	*M0204*	*tested in optimization process*
Ligase C: CircLigase	*Epicentre*	*CL4111K*	*with results described in the*

jove
JOURNAL OF VISUALIZED EXPERIMENTS

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

Title of Article:

Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells

Author(s):

Darja Pollpeter, Andrew E. Sobala, Michael H. Malim

Item 1: The Author elects to have the Materials be made available (as described at http://www.jove.com/publish) via:

☐ Standard Access   ☑ Open Access

Item 2: Please select one of the following items:

☑ The Author is **NOT** a United States government employee.

☐ The Author is a United States government employee and the Materials were prepared in the course of his or her duties as a United States government employee.

☐ The Author is a United States government employee but the Materials were NOT prepared in the course of his or her duties as a United States government employee.

ARTICLE AND VIDEO LICENSE AGREEMENT

1.   **Defined Terms.** As used in this Article and Video License Agreement, the following terms shall have the following meanings: "**Agreement**" means this Article and Video License Agreement; "**Article**" means the article specified on the last page of this Agreement, including any associated materials such as texts, figures, tables, artwork, abstracts, or summaries contained therein; "**Author**" means the author who is a signatory to this Agreement; "**Collective Work**" means a work, such as a periodical issue, anthology or encyclopedia, in which the Materials in their entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole; "**CRC License**" means the Creative Commons Attribution 3.0 Agreement (also known as CC-BY), the terms and conditions of which can be found at: http://creativecommons.org/licenses/by/3.0/us/legalcode ; "**Derivative Work**" means a work based upon the Materials or upon the Materials and other pre-existing works, such as a translation, musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Materials may be recast, transformed, or adapted; "**Institution**" means the institution, listed on the last page of this Agreement, by which the Author was employed at the time of the creation of the Materials; "**JoVE**" means MyJove Corporation, a Massachusetts corporation and the publisher of The Journal of Visualized Experiments; "**Materials**" means the Article and / or the Video; "**Parties**" means the Author and JoVE; "**Video**" means any video(s) made by the Author, alone or in conjunction with any other parties, or by JoVE or its affiliates or agents, individually or in collaboration with the Author or any other parties, incorporating all or any portion of the Article, and in which the Author may or may not appear.

2.   **Background.** The Author, who is the author of the Article, in order to ensure the dissemination and protection of the Article, desires to have the JoVE publish the Article and create and transmit videos based on the Article. In furtherance of such goals, the Parties desire to memorialize in this Agreement the respective rights of each Party in and to the Article and the Video.

3.   **Grant of Rights in Article.** In consideration of JoVE agreeing to publish the Article, the Author hereby grants to JoVE, subject to **Sections 4 and 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Article in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Article into other languages, create adaptations, summaries or extracts of the Article or other Derivative Works (including, without limitation, the Video) or Collective Works based on all or any portion of the Article and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. If the "Open Access" box has been checked in **Item 1** above, JoVE and the Author hereby grant to the public all such rights in the Article as provided in, but subject to all limitations and requirements set forth in, the CRC License.

612542.6   For questions, please contact us at submissions@jove.com or +1.617.945.9051.

4. **Retention of Rights in Article.** Notwithstanding the exclusive license granted to JoVE in **Section 3** above, the Author shall, with respect to the Article, retain the non-exclusive right to use all or part of the Article for the non-commercial purpose of giving lectures, presentations or teaching classes, and to post a copy of the Article on the Institution's website or the Author's personal website, in each case provided that a link to the Article on the JoVE website is provided and notice of JoVE's copyright in the Article is included. All non-copyright intellectual property rights in and to the Article, such as patent rights, shall remain with the Author.

5. **Grant of Rights in Video - Standard Access.** This **Section 5** applies if the "Standard Access" box has been checked in **Item 1** above or if no box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby acknowledges and agrees that, subject to **Section 7** below, JoVE is and shall be the sole and exclusive owner of all rights of any nature, including, without limitation, all copyrights, in and to the Video. To the extent that, by law, the Author is deemed, now or at any time in the future, to have any rights of any nature in or to the Video, the Author hereby disclaims all such rights and transfers all such rights to JoVE.

6. **Grant of Rights in Video - Open Access.** This **Section 6** applies only if the "Open Access" box has been checked in **Item 1** above. In consideration of JoVE agreeing to produce, display or otherwise assist with the Video, the Author hereby grants to JoVE, subject to **Section 7** below, the exclusive, royalty-free, perpetual (for the full term of copyright in the Article, including any extensions thereto) license (a) to publish, reproduce, distribute, display and store the Video in all forms, formats and media whether now known or hereafter developed (including without limitation in print, digital and electronic form) throughout the world, (b) to translate the Video into other languages, create adaptations, summaries or extracts of the Video or other Derivative Works or Collective Works based on all or any portion of the Video and exercise all of the rights set forth in (a) above in such translations, adaptations, summaries, extracts, Derivative Works or Collective Works and (c) to license others to do any or all of the above. The foregoing rights may be exercised in all media and formats, whether now known or hereafter devised, and include the right to make such modifications as are technically necessary to exercise the rights in other media and formats.

7. **Government Employees.** If the Author is a United States government employee and the Article was prepared in the course of his or her duties as a United States government employee, as indicated in **Item 2** above, and any of the licenses or grants granted by the Author hereunder exceed the scope of the 17 U.S.C. 403, then the rights granted hereunder shall be limited to the maximum rights permitted under such statute. In such case, all provisions contained herein that are not in conflict with such statute shall remain in full force and effect, and all provisions contained herein that do so conflict shall be

deemed to be amended so as to provide to JoVE the maximum rights permissible within such statute.

8. **Protection of the work.** The Author(s) authorize JoVE to take steps in the Author(s) name and on their behalf if JoVE believes some third party could be infringing or might infringe the copyright of either the Author's Article and/or Video.

9. **Likeness, Privacy, Personality.** The Author hereby grants JoVE the right to use the Author's name, voice, likeness, picture, photograph, image, biography and performance in any way, commercial or otherwise, in connection with the Materials and the sale, promotion and distribution thereof. The Author hereby waives any and all rights he or she may have, relating to his or her appearance in the Video or otherwise relating to the Materials, under all applicable privacy, likeness, personality or similar laws.

10. **Author Warranties.** The Author represents and warrants that the Article is original, that it has not been published, that the copyright interest is owned by the Author (or, if more than one author is listed at the beginning of this Agreement, by such authors collectively) and has not been assigned, licensed, or otherwise transferred to any other party. The Author represents and warrants that the author(s) listed at the top of this Agreement are the only authors of the Materials. If more than one author is listed at the top of this Agreement and if any such author has not entered into a separate Article and Video License Agreement with JoVE relating to the Materials, the Author represents and warrants that the Author has been authorized by each of the other such authors to execute this Agreement on his or her behalf and to bind him or her with respect to the terms of this Agreement as if each of them had been a party hereto as an Author. The Author warrants that the use, reproduction, distribution, public or private performance or display, and/or modification of all or any portion of the Materials does not and will not violate, infringe and/or misappropriate the patent, trademark, intellectual property or other rights of any third party. The Author represents and warrants that it has and will continue to comply with all government, institutional and other regulations, including, without limitation all institutional, laboratory, hospital, ethical, human and animal treatment, privacy, and all other rules, regulations, laws, procedures or guidelines, applicable to the Materials, and that all research involving human and animal subjects has been approved by the Author's relevant institutional review board.

11. **JoVE Discretion.** If the Author requests the assistance of JoVE in producing the Video in the Author's facility, the Author shall ensure that the presence of JoVE employees, agents or independent contractors is in accordance with the relevant regulations of the Author's institution. If more than one author is listed at the beginning of this Agreement, JoVE may, in its sole discretion, elect not take any action with respect to the Article until such time as it has received complete, executed Article and Video License Agreements from each such author. JoVE reserves the right, in its absolute and sole

![jove logo] JOURNAL OF VISUALIZED EXPERIMENTS

1 Alewife Center #200
Cambridge, MA 02140
tel. 617.945.9051
www.jove.com

ARTICLE AND VIDEO LICENSE AGREEMENT - UK

discretion andwithout giving any reason therefore, to accept or decline any work submitted to JoVE. JoVE and its employees, agents and independent contractors shall have full, unfettered access to the facilities of the Author or of the Author's institution as necessary to make the Video, whether actually published or not. JoVE has sole discretion as to the method of making and publishing the Materials, including, without limitation, to all decisions regarding editing, lighting, filming, timing of publication, if any, length, quality, content and the like.

12.   **Indemnification.** The Author agrees to indemnify JoVE and/or its successors and assigns from and against any and all claims, costs, and expenses, including attorney's fees, arising out of any breach of any warranty or other representations contained herein. The Author further agrees to indemnify and hold harmless JoVE from and against any and all claims, costs, and expenses, including attorney's fees, resulting from the breach by the Author of any representation or warranty contained herein or from allegations or instances of violation of intellectual property rights, damage to the Author's or the Author's institution's facilities, fraud, libel, defamation, research, equipment, experiments, property damage, personal injury, violations of institutional, laboratory, hospital, ethical, human and animal treatment, privacy or other rules, regulations, laws, procedures or guidelines, liabilities and other losses or damages related in any way to the submission of work to JoVE, making of videos by JoVE, or publication in JoVE or elsewhere by JoVE. The Author shall be responsible for, and shall hold JoVE harmless from, damages caused by lack of sterilization, lack of cleanliness or by contaminationdue to the making of a video by JoVE its employees, agents or independent contractors. All sterilization, cleanliness or

decontamination procedures shall be solely the responsibility of the Author and shall be undertaken at the Author's expense. All indemnifications provided herein shall include JoVE's attorney's fees and costs related to said losses or damages. Such indemnification and holding harmless shall include such losses or damages incurred by, or in connection with, acts or omissions of JoVE, its employees, agents or independent contractors.

13.   **Fees.** To cover the cost incurred for publication, Jo VE must receive payment before production and publication the Materials. Payment is due in 21 days of invoice. Should the Materials not be published due to an editorial or production decision, these funds will be returned to the Author. Withdrawal by the Author of any submitted Materials after final peer review approval will result in a US$1,200 fee to cover pre-production expenses incurred by JoVE. If payment is not received by the completion of filming, production and publication of the Materials will be suspended until payment is received.

14.   **Transfer, Governing Law.** This Agreement may be assigned by JoVE and shall inure to the benefits of any of JoVE's successors and assignees. This Agreement shall be governed and construed by the internal laws of the Commonwealth of Massachusetts without giving effect to any conflict of law provision thereunder. This Agreement may be executed in counterparts, each of which shall be deemed an original, but all of which together shall be deemed to me one and the same agreement. A signed copy of this Agreement delivered by facsimile, e-mail or other means of electronic transmission shall be deemed to have the same legal effect as delivery of an original signed copy of this Agreement.

A signed copy of this document must be sent with all new submissions. Only one Agreement is required per submission.

CORRESPONDING AUTHOR

Name: *Darja Pollpeter*

Department: *Infectious Diseases*

Institution: *King's College London*

Title: *Postdoctoral Research Associate*

Signature: *Darje Pollpbr*   Date: *02-July-2018*

Please submit a **signed** and **dated** copy of this license by one of the following three methods:
1. Upload an electronic version on the JoVE submission site
2. Fax the document to +1.866.381.2236
3. Mail the document to JoVE / Attn: JoVE Editorial / 1 Alewife Center #200 / Cambridge, MA 02140

Rebuttal letter for submitted manuscript JoVE58715 "Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells" by Pollpeter et al.

We thank the editor and reviewers for their careful reading of our manuscript. We have made all the requested changes and believe this improved the manuscript for the reader. Please find the actions taken in bold and italics.

Editorial comments:

Changes to be made by the Author(s):

1. Please take this opportunity to thoroughly proofread the manuscript to ensure that there are no spelling or grammar issues. The JoVE editor will not copy-edit your manuscript and any errors in the submitted revision may be present in the published version. *- Done*

2. Please obtain explicit copyright permission to reuse any figures from a previous publication. Explicit permission can be expressed in the form of a letter from the editor or a link to the editorial policy that allows re-prints. Please upload this information as a .doc or .docx file to your Editorial Manager account. The Figure must be cited appropriately in the Figure Legend, i.e. "This figure has been modified from [citation]." *– See submitted document, lease advice if any additional information is required.*

3. Please use SI abbreviations for all units: L, mL, µL, h, min, s, etc. – *Changed where applicable*

4. JoVE cannot publish manuscripts containing commercial language. This includes trademark symbols (™), registered symbols (®), and company names before an instrument or reagent. Please remove all commercial language from your manuscript and use generic terms instead. All commercial products should be sufficiently referenced in the Table of Materials and Reagents. For example: OptiMEM, Alliance, Qiagen, Cutsmart, LoBind, Eppendorf, etc. *- Done*

5. Please revise the protocol text to avoid the use of any personal pronouns (e.g., "we", "you", "our" etc.). *- Done*

6. Please revise the protocol to contain only action items that direct the reader to do something (e.g., "Do this," "Ensure that," etc.). The actions should be described in the imperative tense in complete sentences wherever possible. Avoid usage of phrases such as "could be," "should be," and "would be" throughout the Protocol. Any text that cannot be written in the imperative tense may be added as a "Note." Please include all safety procedures and use of hoods, etc. Please move the discussion about the protocol to the Discussion. *– Done. This is the main change in the resubmission. Several notes were moved to the discussion, other, shorter notes were incorporated. Some steps were split into sub steps and hence step numbering has now changed.*

7. Please add more details to your protocol steps. There should be enough detail in each step to supplement the actions seen in the video so that viewers can easily replicate the protocol. Please ensure you answer the "how" question, i.e., how is the step performed? Alternatively, add references to published material specifying how to perform the protocol action. *- Done*

Some examples:

1.1.2: Please specify the incubation temperature. **- *Done***

1.1.3: What happens after replacing with new medium? Is the cultured incubated? Please specify. **- *Done***

1.1.4: What happens after filtration? **- *Done***

1.1.5: Please specify ultracentrifugation parameters (force and time). What container is used in this step? Please break up into sub-steps. **- *Force and time was already indicated, changed others.***

1.2.3: Please specify the incubation conditions. Please specify throughout. **- *Done***

8. Please ensure that conditions and primers are listed all PCR procedures. **– *Double checked, though not sure where any were missing.***

9. The Protocol should be made up almost entirely of discrete steps without large paragraphs of text between sections. Please simplify the Protocol so that individual steps contain only 2-3 actions per step and a maximum of 4 sentences per step. Use sub-steps as necessary. Please move the discussion about the protocol to the Discussion. **– *Done, see comment 6 above.***

10. References: Please do not abbreviate journal titles. **– *Changed to full journal titles.***

Reviewers' comments:

Reviewer #1:
Manuscript Summary:
This manuscript describing the method to determine 3'-termini and sequences of HIV viral DNA in infected cells will significantly help researchers in the field. The manuscript is very well written, detailed and very comprehensive. Being a HIV researcher myself (although not specialized in the RT step), I had no difficulty whatsoever in understanding this protocol. I strongly recommend this for further consideration.

Major Concerns:
None

Minor Concerns:
None

Reviewer #2:
Manuscript Summary:
The method addresses a gap in research enabling detailed analysis of reverse transcripts in HIV infected cells. Previously, methods would detect if reverse transcription had reached a specific stage or not, but were unable to determine the sequence of the reverse transcripts and were based on population level analyses. This method enables capture and sequencing of individual nascent reverse transcribed (-) DNA molecules. There is also a method for analyzing sequences to determine the length of the transcript and deamination events (if APOBEC3 enzymes are being studied).

Major Concerns:
The analysis of sequences is the major culmination of a very involved method. However, the reference to the website alone does not seem to be sufficient;

especially if a researcher has not done their own sequence analysis previously. Some more detail of how to use the GitHub and SeqParse within the manuscript is warranted. *– **We have now added a step-by-step description of how to install, set up and run our in-house script, see section 8.***

Minor Concerns:

Section 1.1.4: It is not clear if the 48 h later is after the media change or the transfection. *– **We have now specified that it is after the transfection***.

Section 1.1.5: It is not clear that the virus will be in a pellet below both the sucrose and supernatant, needs to be directly stated. *- **Done***

Section 4.2.5: It is not clear until later in the protocol that each high, med, and low gel fragment should be kept separate. *– **We have now made this clear at several steps (see step 4.2.1 and in a note after step 4.2.6)***

Figure 2a: Has an uncompiled reference. *- **Changed***

Representative Results Section: When discussing the oligos and possibly having incomplete products; what was the purity of the oligos that were ordered (Desalted, HPLC, gel purified?); would your recommend purification to avoid this or it still occurs? *– **We have now added the information that we had ordered oligos as HPLC purified. We don't know, whether further purification would avoid the presence of incomplete products altogether.***

Reviewer #3:

The manuscript entitled 'Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells' by Pollpeter et al. describes a detailed protocol to analyze nucleotide sequences of the early reverse-transcribed DNA products in HIV-1-infected cells. The authors' group recently reported a possible molecular mechanism of APOBEC3G-mediated reverse transcription inhibition (ref #6), which was conducted by using the protocol. To date, many retrovirus researchers have attempted to determine the 3'-termini of the early reverse-transcribed DNA intermediates in HIV-1-infected cells. But there were no feasible and sensitive assay systems to analyze them, because of some technical difficulties. In this manuscript, the authors optimized the best method of the adaptor ligation to prepare a DNA library for the next generation sequencing tool, Illumina MiSeq and elaborated this protocol. Overall, this manuscript is well written, important and very informative. I have only a few minor comments as described below.

Minor comments:

1: line 85. A typo. The "RML-RACE" should be "RLM-RACE". *- **Changed***

2: lines 191-207. Specify the temperature(s) in these processes. *– **Added, all at room temperature.***

3: lines 239 and 245. Should 40% PEG be used here? In the ref #6 that the authors previously reported, it says 50% PEG. *– **We have used both 50% PEG and 40% and did not find a difference. We currently use 40%, simply because the 50% product was discontinued from out chosen company***.

4: lines 542. The "long (116-134 nt)" is correct? It seems "long (116-120)". –
134 is indeed incorrect, we have now changed this to 120.

Explicit copyright permission for article "Determining 3'-termini and sequences of nascent single stranded viral DNA molecules during HIV-1 reverse transcription in infected cells" by Pollpeter et al., submitted to JoVE on July 2nd 2018.

As the authors of the original publication in Nature Microbiology, from which the figures are taken, the copyrights remain with us. All re-used material is appropriately labeled as such. Please find below a link to Nature Microbiology's editorial policies on re-used material. In addition we provide the outcome of an official author request through RightsLink®, the copyrights clearance center that is partnered with Nature Microbiology. Further we provide the email response from the Nature Microbiology editorial office confirming the copyright remains with the authors.

Link to the editorial policies:

https://www.nature.com/reprints/permission-requests.html

SPRINGER NATURE

Title: Deep sequencing of HIV-1 reverse transcripts reveals the multifaceted antiviral functions of APOBEC3G

Author: Darja Pollpeter et al

Publication: Nature Microbiology

Publisher: Springer Nature

Date: Nov 20, 2017

Copyright © 2017, Springer Nature

Author Request

If you are the author of this content (or his/her designated agent) please read the following. If you are not the author of this content, please click the Back button and select no to the question "Are you the Author of this Springer Nature content?".

Ownership of copyright in original research articles remains with the Author, and provided that, when reproducing the contribution or extracts from it or from the Supplementary Information, the Author acknowledges first and reference publication in the Journal, the Author retains the following non-exclusive rights:

To reproduce the contribution in whole or in part in any printed volume (book or thesis) of which they are the author(s).

The author and any academic institution, where they work, at the time may reproduce the contribution for the purpose of course teaching.

To reuse figures or tables created by the Author and contained in the Contribution in oral presentations and other works created by them.

To post a copy of the contribution as accepted for publication after peer review (in locked Word processing file, of a PDF version thereof) on the Author's own web site, or the Author's institutional repository, or the Author's funding body's archive, six months after publication of the printed or online edition of the Journal, provided that they also link to the contribution on the publisher's website. Authors wishing to use the published version of their article for promotional use or on a web site must request in the normal way.

If you require further assistance please read Springer Nature's online author reuse guidelines.

For full paper portion: Authors of original research papers published by Springer Nature are encouraged to submit the author's version of the accepted, peer-reviewed manuscript to their relevant funding body's archive, for release six months after publication. In addition, authors are encouraged to archive their version of the manuscript in their institution's repositories (as well as their personal Web sites), also six months after original publication.

v1.0

BACK   CLOSE WINDOW

Dear Darja,

Many thanks for reaching out to us, I can confirm that this usage is absolutely fine!

Do not hesitate to contact me directly if you require anything further

Best,

Ana

Anastasia Broadbridge
Senior Editorial Assistant
Nature Microbiology and Nature Plants
Nature Research

-----Original Message-----
From: Pollpeter, Darja [mailto:darja.pollpeter@kcl.ac.uk]
Sent: 06 June 2018 22:26
To: Nonia Pariente Penamaria
Subject: Copyright question

Dear Nonia,
I hope you are doing well!
Mike Malim and I have been invited to submit a manuscript to the Video Journal JoVE (Journal of Visualised Experiments) about the sequencing method we have first described in our recent publication with Nature Microbiology (https://emea01.safelinks.protection.outlook.com/?url=https%3A%2F%2Fwww.nature.com%2Farticles%2Fs41564-017-0063-9&data=01%7C01%7Cdarja.pollpeter%40kcl.ac.uk%7Cddc01b56a9214a7e771608d5cc7b4930%7C8370cf1416f34c16b83c724071654356%7C0&sdata=JytXJcbCOhaskJL6rd9gJOu3ZgwNvvf4DD2Gi8xVoO8%3D&reserved=0). We think this is a good idea and it would make reproduction of the technique easier for other laboratories. Since the Nature Microbiology article has quite a detailed description in the supplementary data I would like to re-use a couple of the schematics. The same goes for one main figure as representative data, which are typically part of JoVE. I have been to your website and read through the RightsLink instructions. From that it actually sounded like for us as authors re-using this, with appropriate referencing obviously, would be ok, but I want to make absolutely sure I am going through all the right channels.
Could you please let me know, who best to contact about this?

Thank you very much,
Darja