

Additional file 1

Herpes simplex virus type 1 R-loops are targets for APOBEC-mediated mutagenesis

Márton Miskei, Dóra Varga, Lilla Hornyák, Éva Sipos, Éva Nagy, Qiuzhen Li, Zsolt Karányi, Zoltán Szabó, Rachel DeWeerd, Abby M Green, Dávid Szüts, Eszter Csoma, Lóránt Székvölgyi

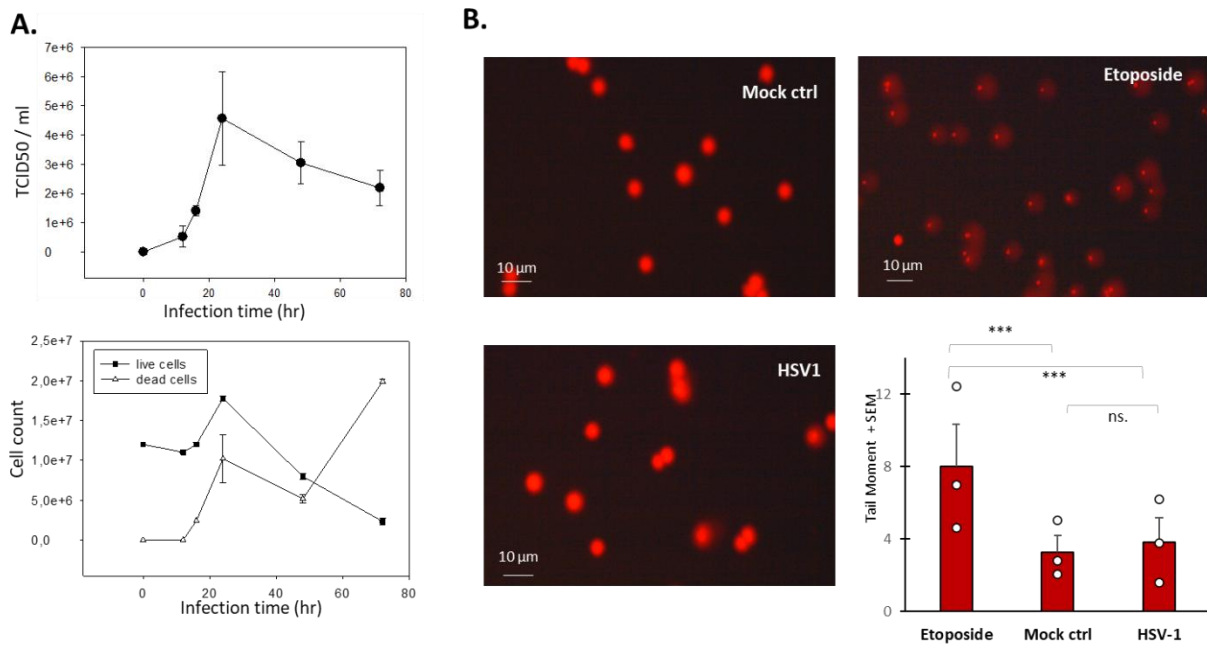


Fig. S1. A. Infective titration and cell viability of HSV-1 infected Jurkat cells. Top panel: each point on the resulting growth curve represents the average titer calculated from three independent experiments for each treatment or time point. Bottom panel: the number of live and dead cells as a function of infection time. Cell viability was determined by trypan blue exclusion method: cells taking up trypan blue were considered non-viable. B. Detection of DNA damage by alkaline comet assay in HSV-1 infected and Vero mock-treated cells. As a positive control, cells were treated with the topoisomerase II inhibitor etoposide (40 μ M for 1 hour). Representative images are shown. The DNA was stained with propidium iodide. Tail moments were quantified by the OpenComet software. *** means a statistically significant difference ($p < 0.0001$). ns.: non-significant. TCID₅₀: 50% tissue culture infectious dose.

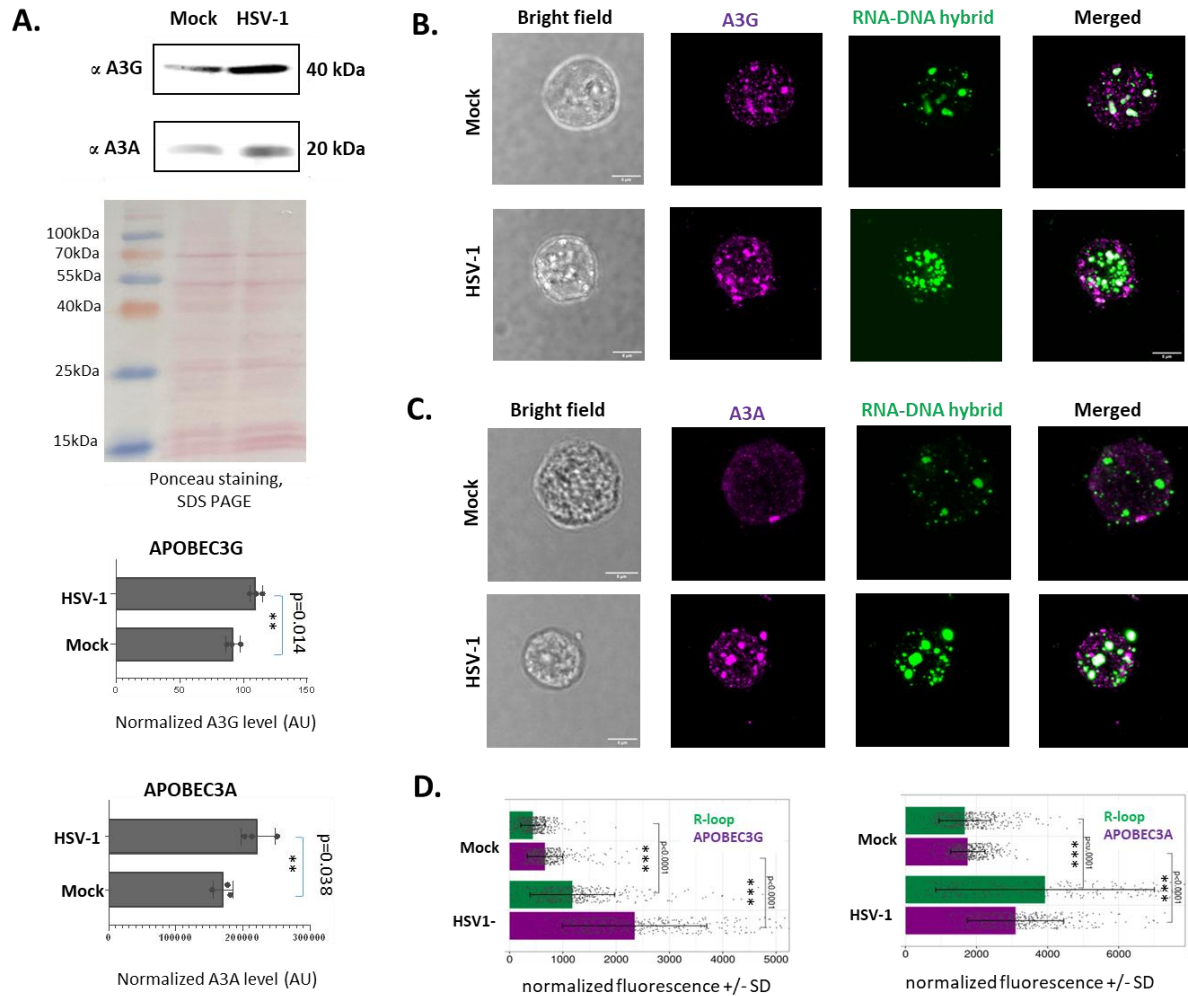


Fig. S2. APOBEC3A and APOBEC3G enzymes are upregulated during HSV-1 infection. (A) Western blot analysis of APOBEC3G and APOBEC3A expression. The antibodies detected a dominant band at the expected molecular weights of A3A (20 kDa) and A3G (40 kDa). Upregulation of A3G and A3A enzymes upon HSV-1 infection is significant (Mann-Whitney U test). Western blot band intensities were quantified and normalized to total protein levels, as determined by Ponceau red staining. The quantification results are displayed below the blot. (B-D) Immunofluorescence labeling of R-loop structures and A3A/A3G enzymes in HSV-1 infected Jurkat cells. (A) Representative confocal laser scanning microscopic image of HSV-1-infected and mock-treated Jurkat cells (with Vero cell extract). The cells were co-labeled using anti-APOBEC3G and anti-RNA-DNA hybrid (S9.6) antibodies to visualize the presence and overlap of A3G enzymes and R-loop structures. A total of 50 cells were analyzed. (C) This panel replicates the experiment described in panel A, substituting A3G with A3A. (D) Quantification of the fluorescent signals from the same samples using automated microscopy (LSC). Results demonstrate that both R-loop and A3A and A3G levels are significantly increased in HSV-1 infected cells compared to mock controls. A minimum of 5,000 cells were analyzed, and statistical significance is indicated (Student's t-test, significance level: 0.05).

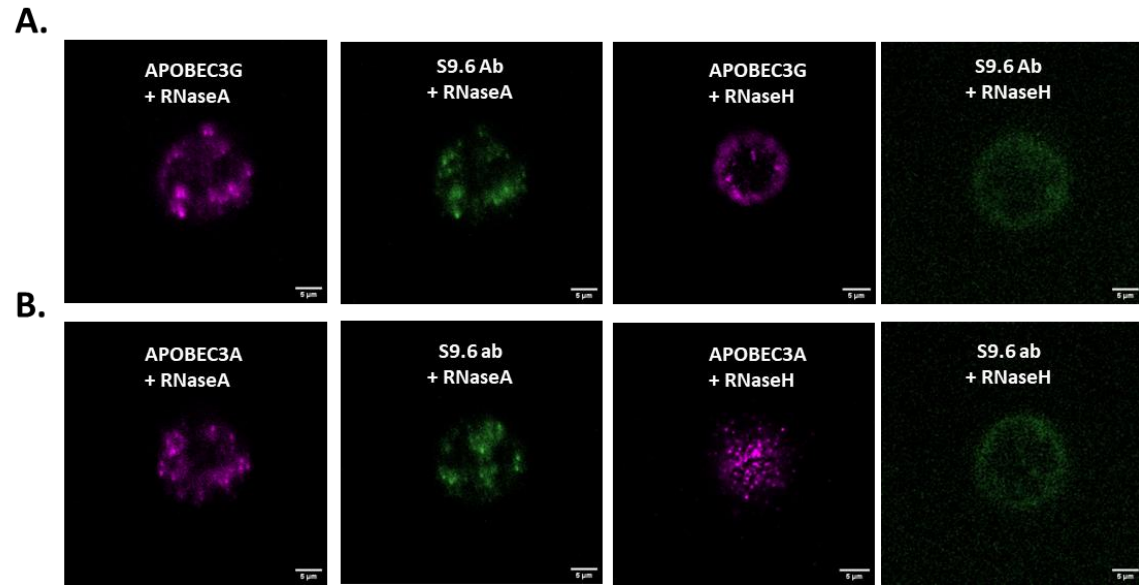


Fig. S3. Microscopic analysis of HSV-1-infected Jurkat cells treated with distinct ribonucleases. RNase A and RNase H were used to selectively degrade single-stranded RNA and RNA-DNA hybrids, respectively. RNA-DNA hybrids are visualized with the S9.6 antibody (green), while the (A) APOBEC3G and (B) APOBEC3A enzymes are labeled with specific antibodies (magenta). Representative images are displayed. Scale bar: 5 μ m.

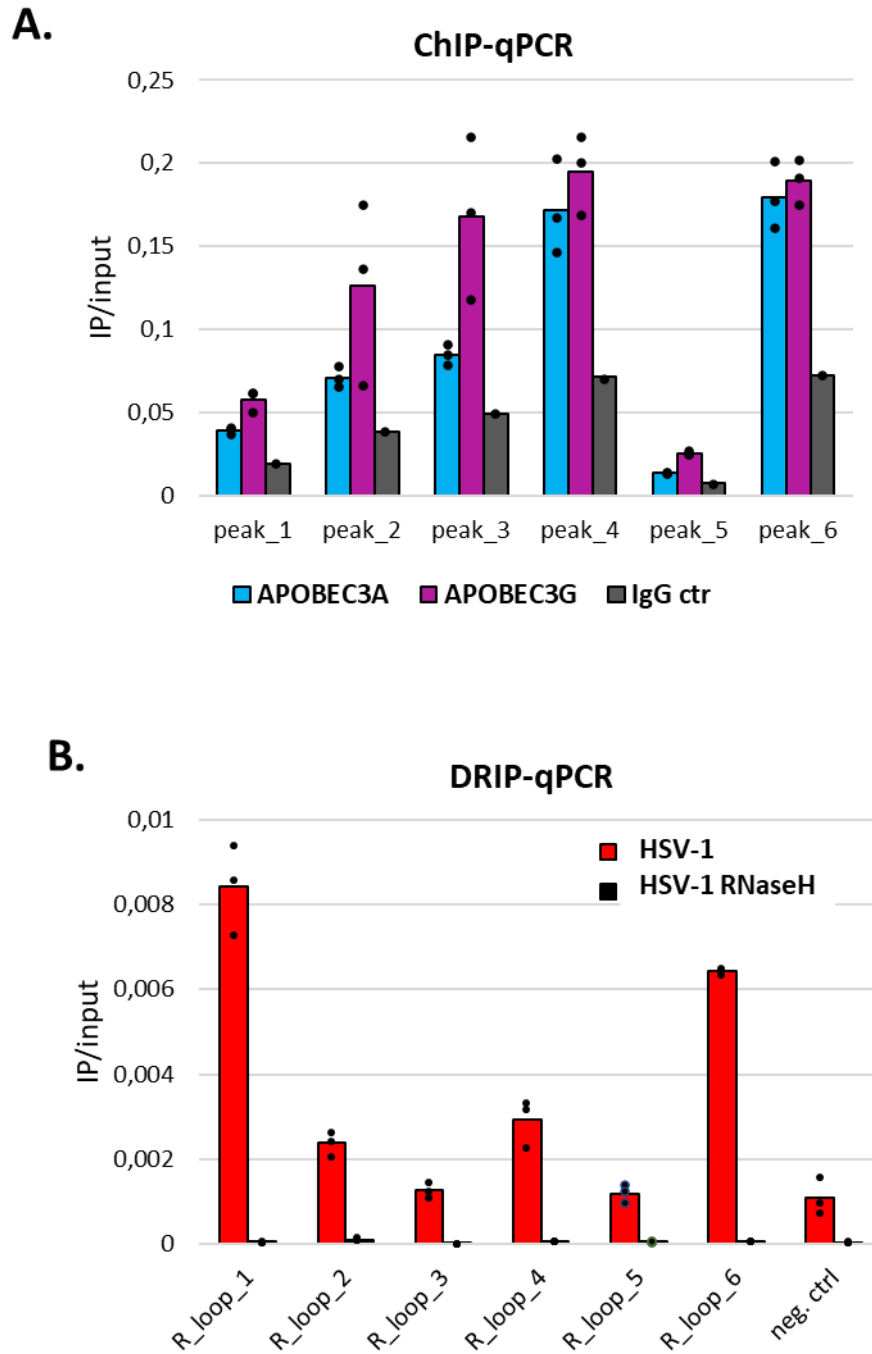


Fig. S4. qPCR validation of ChIP-seq and DRIP-seq data. Six peaks were selected for the analysis. For ChIP samples, IgG was included as a no-antibody control to account for background signal, with all peaks showing significant enrichment compared to IgG. For DRIP validation, RNase H-treated samples served as a negative control, exhibiting baseline S9.6 signal and confirming the specificity of S9.6 immunoprecipitation. Additionally, a control site from the genome (*ZNF554*) [1] was included.

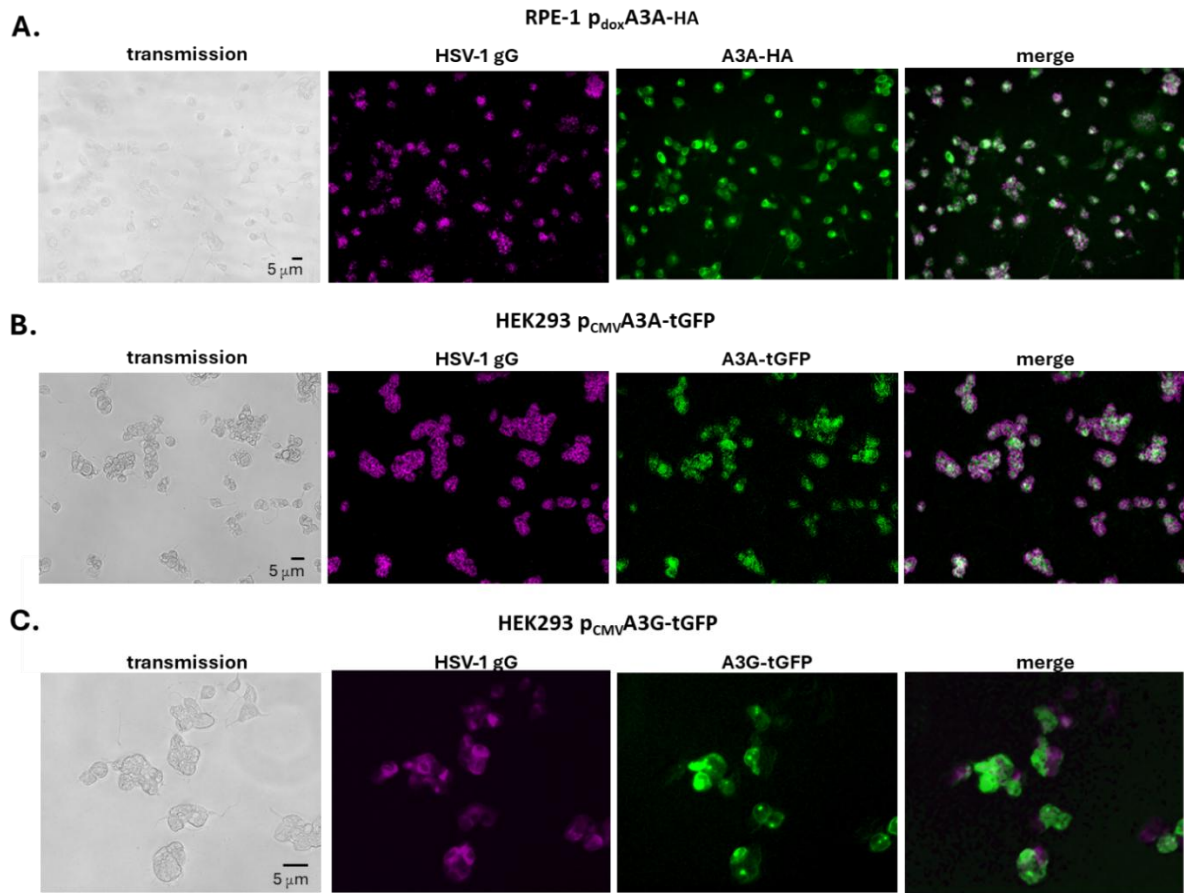


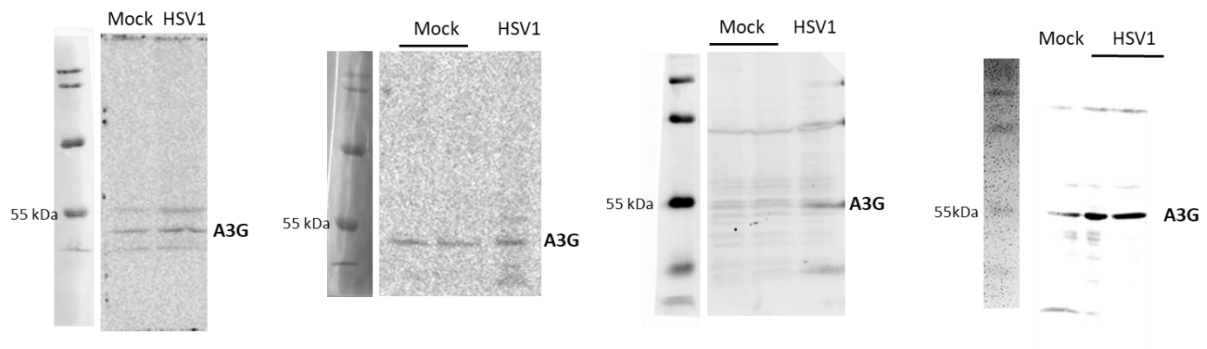
Fig. S6. Immunofluorescent detection of HSV-1 and APOBEC3A or APOBEC3G expression. (A) RPE-1 p_{dox}A3A-HA cells. (B) HEK293 p_{CMV}A3A-tGFP cells. (C) HEK293 p_{CMV}A3G-tGFP cells. HSV-1 was detected using anti-HSV-1 glycoprotein B antibody (magenta). A3A-HA was detected using anti-HA antibody (green), while A3A-tGFP and A3G-tGFP were visualized by turboGFP fluorescence (green).



Fig. S7. APOBEC-type C→T mutation hotspots shared between ChIP, DRIP, and input HSV-1 genomes. (A) Genome browser snapshot showing the ChIP and DRIP signals as well as the C>T mutations in RPE-1 p_{dox}A3A-HA, HEK293 p_{CMV}A3A-tGFP, and HEK293 p_{CMV}A3G-tGFP cells. C-to-T mutations are represented by vertical lines. Data was normalized to input. (B) High-resolution genome browser view of C→T mutation hotspots in the HSV-1 genome. Seven regions are highlighted and magnified to display local mutation patterns.

A.

Western blot anti-APOBEC3G (H-63) (Sc-48820, Santa Cruz Biotechnology, Inc)



B.

Western blot anti-APOBEC3A (D-23) (Sc-130688, Santa Cruz Biotechnology, Inc)

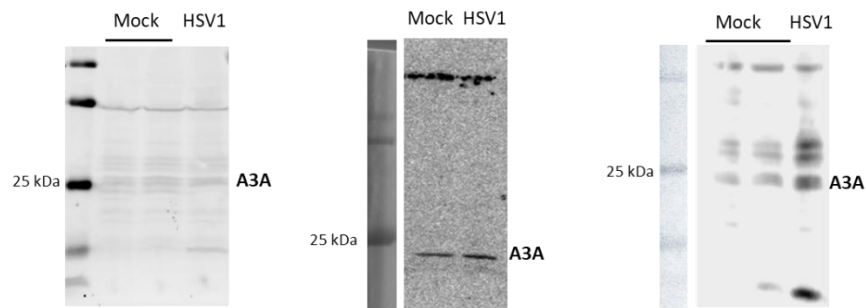
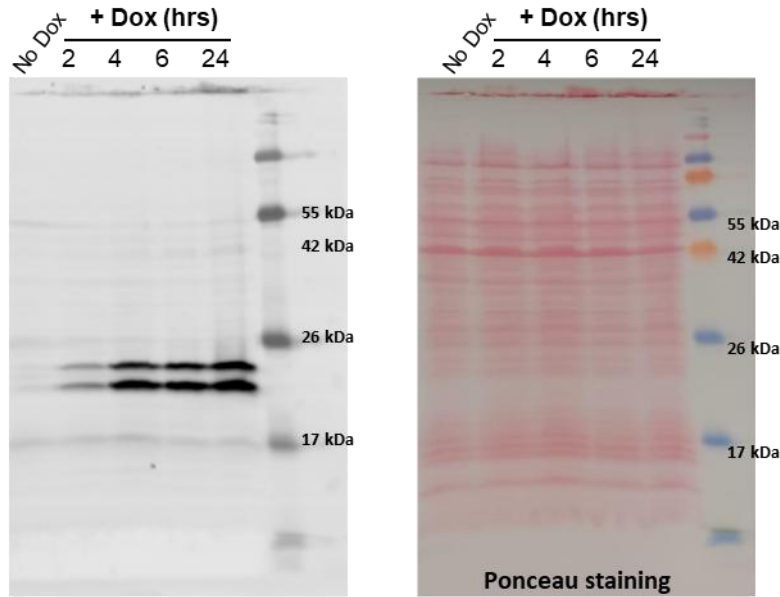


Fig. S8. Western blot analysis of APOBEC3A and APOBEC3G expression. Original images. The antibodies detected a dominant band at the expected molecular weights of A3A (20 kDa) and A3G (40 kDa). Upregulation of A3G and A3A enzymes upon HSV-1 infection is apparent.

Western blot anti-HA antibody



RPE-1 p_{dox}A3A-HA cells
APOBEC3A-HA size: ~25kDa

Fig. S9. Western blot analysis of APOBEC3A-HA expression in RPE-1 p_{dox}A3A-HA cells. Original image. The antibody detected a dominant band at the expected molecular weights of A3A-HA (25 kDa) that scaled with the time of doxycycline (dox) induction. The two bands observed in RPE-1 A3A-HA lysates both correspond to A3A-HA. The APOBEC3A transgene contains a second AUG at codon 13, which can serve as an alternative translation start site and produces a slightly smaller N-terminally truncated isoform; in our lentiviral constructs, this shorter isoform is often expressed more prominently, giving rise to the characteristic doublet.

Fig. S10. Amino acid substitutions in HSV-1 proteins arising from APOBEC-mediated mutations.

Aligned amino acid sequences of representative HSV-1 protein-coding genes that accumulate APOBEC-type C→T mutations are shown, highlighting the resulting amino acid changes.

UL36

UL36 (WT)
MIAGTPPHSTMERGGDRDIVVTGARNQFAPDLEPGGSVSCMRSSLSFLSLIFDVGPRDVL
UL36 (MUT)
MIAGTPPHSTMERGGDRDIVVTGARNQFAPDLEPGGSVSCMRSSLSFLSLIFDVGPRDVL

UL36 (WT)
SAEAIEGCLVEGGEWTRATAGPGPPRMCSIVELPNFLEYPGARGGLRCVFSRVYGEVGFF
UL36 (MUT)
SAEAIEGCLVEGGEWTRATAGPGPPRMCSIVELPNFLEYPGARGGLRCVFSRVYGEVGFF

UL36 (WT)
GEPAAAGLLETQCPAHTFFAGPWALRPLSYTLLTIGPLGMGLFRDGDYAYLFDPHGLPEGT
UL36 (MUT)
GEPAAAGLLETQCPAHTFFAGPWALRPLSYTLLTIGPLGMGLFRDGDYAYLFDPHGLPEGT

UL36 (WT)
PAFIAKVRAGDMYPYLTYYYTRDRPDVRWAGAMVFFVPSGPEPAAPADLTAAALHLYGASE
UL36 (MUT)
PAFIAKVRAGDMYPYLTYYYTRDRPDVRWAGAMVFFVPSGPEPAAPADLTAAALHLYGASE

UL36 (WT)
TYMQDEAFSERRVAITHPLRGEIAGLGEPVGVGPREGGGGPGHPPTAAQSPPTARR
UL36 (MUT)
TYMQDEAFSERRVAITHPLRGEIAGLGEPVGVGPREGGGGPGHPPTAAQSPPTARR

UL36 (WT)
DDRASETSRGTAGPSAKPEAKRPNRAPDDVWAVALKGTPPTDPPSADPPSAIPPPPSAP
UL36 (MUT)
DDRASETSRGTAGPSAKPEAKRPNRAPDDVWAVALKGTPPTDPPSADPPSAIPPPPSAP

UL36 (WT)
KTPAAEAAEEDDDDMRVLEMGVVPVGRHRARYSAGLPKRRRPTWTPPSSVEDLTSGEKTK
UL36 (MUT)
KTPAAEAAEEDDDDMRVLEMGVVPVGRHRARYSAGLPKRRRPTWTPPSSVEDLTSGEKTK

UL36 (WT)
RSAPPAKTKKKSTPKGKTPVGAAVPASVPEPVLASAPDPAGPPVAEAGEDDGPMVPASS
UL36 (MUT)
RSAPPAKTKKKSTPKGKTPVGAAVPASVPEPVLASAPDPAGPPVAEAGEDDGPMVPASS

UL36 (WT)
QALEALKTRRSPEPPGADLAQLFEAHPNVAATAVKFTACSATLAREVAACSRLTISALRS

UL36 (MUT)
QALEALKTRRSPEPPGADLAQLFEAHPNVAATAVKFTACSATLAREVAACSRILTISALRS

UL36 (WT)
PYPASPGLELVCVIFFFERVLAFLIENGARTHTQAGVAGPAAALLEFTLSMLPRKTAVGD
UL36 (MUT)
PYPASPGLELVCVIFFFERVLAFLIENGARTHTQAGVAGPAAALLEFTLSMLPRKTAVGD

UL36 (WT)
FLASTRLSLADVAAHLPLVQHVLDENSLIGRLALAKLILVARDVIRETDAFYGELADLEL
UL36 (MUT)
FLASTRLSLADVAAHLPLVQHVLDENSLIGRLALAKLILVARDVIRETDAFYGELADLEL

UL36 (WT)
QLRAAPPANLYTRLGEWLLERSQAHPDTLFAPATPTHPEPLLYRVQALAKFARGEIIRVE
UL36 (MUT)
QLRAAPPANLYTRLGEWLLERSQAHPDTLFAPATPTHPEPLLYRVQALAKFARGEIIRVE

UL36 (WT)
AEDRQMREALDALARGVDAVSQHAGPLGVMPAPAGAAPQGAPRPPPLGPEAVQVRLEEVR
UL36 (MUT)
AEDRQMREALDALARGVDAVSQHAGPLGVMPAPAGAAPQGAPRPPPLGPEAVQVRLEEVR

UL36 (WT)
TQARRAIEGAVKEYFYRGAVYSAKALQASDNNDRRFHVASA AVVPVQVQLLES LPVFDQHT
UL36 (MUT)
TQARRAIEGAVKEYFYRGAVYSAKALQASDNNDRRFHVASA AVVPVQVQLLES LPVFDQHT

UL36 (WT)
RDIAQRAAIPAPPIATSPTAILLRDLIQRGQTLDAPEDLAAWLSVLTDAANQGLIERKP
UL36 (MUT)
RDIAQRAAIPAPPIATSPTAILLRDLIQRGQTLDAPEDLAAWLSVLTDAANQGLIERKP

UL36 (WT)
LDELARSIRDINDQQARRSSGLAELRRFDALDAALGQQLSDA AFVPAPGASPY PDDGGL
UL36 (MUT)
LDELARSIRDINDQQARRSSGLAELRRFDALDAALGQQLSDA AFVPAPGASPY PDDGGL

UL36 (WT)
SPEATRMAEEALRQARAMDAAKLT AELAPDARARLRERARSLEAMLEGARERAKVARDAR
UL36 (MUT)
SPEATRMAEEALRQARAMDAAKLT AELAPDARARLRERARSLEAMLEGARERAKVARDAR

UL36 (WT)
EKFLHKLQGVLRPLPDFVGLKACPAVLATLRASLPAGWSDLPEAVRGAPPEVTAALRADM

UL36 (MUT)
EKFLHKLQGVLRPLPDFVGLKACPAVLATLRASLPAGWSDLPEAVRGAPPEVTAALRADM

UL36 (WT)
WGLLGQYRDALEHPTPDTATALSGLHPSFVVVLKNLFADAPETPFLLQFFADHAPIIAHA
UL36 (MUT)
WGLLGQYRDALEHPTPDTATALSGLHPSFVVVLKNLFADAPETPFLLQFFADHAPIIAHA

UL36 (WT)
VSNAINAGSAAVATADPASTVDAAVRAHRVLVDAVTALGAAASDPASPLAFLAAMADSAA
UL36 (MUT)
VSNAINAGSAAVATADPASTVDAAVRAHRVLVDAVTALGAAASDPASPLAFLAAMADSAA

UL36 (WT)
GYVKATRLALDARGAIAQLTTLGSAAADLVVQVRRANQPEGEHASLIQAATRATTGARE
UL36 (MUT)
GYVKATRLALDARGAIAQLTTLGSAAADLVVQVRRANQPEGEHASLIQAATRATTGARE

UL36 (WT)
SLAGHEGRFGLLHAEGTAGDHSPSGRALQELGKVIGATRRADELEAAIADLREKMAAQ
UL36 (MUT)
SLAGHEGRFGLLHAEGTAGDHSPSGRALQELGKVIGATRRADELEAAIADLREKMAAQ

UL36 (WT)
RARSSHERWAADVEAVLDRVESGAEFDVVELRRLQALAGTHGYNPRDFRKRAEQALGTNA
UL36 (MUT)
RARSSHERWAADVEAVLDRVESGAEFDVVELRRLQALAGTHGYNPRDFRKRAEQALGTNA

UL36 (WT)
KAVTLALETALAFNPYTPENQRHPMLPPLAAIHRIDWSAAFCAAADTYADMFRVDTEPLA
UL36 (MUT)
KAVTLALETALAFNPYTPENQRHPMLPPLAAIHRIDWSAAFCAAADTYADMFRVDTEPLA

UL36 (WT)
RLLRLAGGLLERAQANDGFIDYHEAVLHLSIDLGGVPALRQYVPFFQKGYAEYVDIRDRL
UL36 (MUT)
RLLRLAGGLLERAQANDGFIDYHEAVLHLSIDLGGVPALRQYVPFFQKGYAEYVDIRDRL

UL36 (WT)
DALRADARRAIGSVALDLAAAAEEISAVRNDPAAAAELVRAGVTLPSPEDALVACVAAL
UL36 (MUT)
DALRADARRAIGSVALDLAAAAEEISAVRNDPAAAAELVRAGVTLPSPEDALVACVAAL

UL36 (WT)
ERVDQSPVKDTAYADYVAFVTRQDLADTKDAVVRKQQRAEATERVTAGLREVLAARERR
UL36 (MUT)
ERVDQSPVKDTAYADYVAFVTRQDLADTKDAVVRKQQRAEATERVTAGLREVLAARERR

UL36 (WT)
AQLEAEGLANLKTLLKVVAVPATVAKTLDQARSAEEIADQVEILLDQTEKARELDVQAVA
UL36 (MUT)
AQLEAEGLANLKTLLKVVAVPATVAKTLDQARSAEEIADQVEILLDQTEKARELDVQAVA

UL36 (WT)
WLEHAQRTFETHPLNAASGDGPGLLTRQGARLQALFDTRRRVEALRRSLEEAEAEWDEVW
UL36 (MUT)
WLEHAQRTFETHPLNAASGDGPGLLTRQGARLQALFDTRRRVEALRRSLEEAEAEWDEVW

UL36 (WT)
GRFGRVRGGAWKSPEGFRAACEQLRALQD TTNTVSGLRAQRDYERLPAKYQGVLGAKSAE
UL36 (MUT)
GRFGRVRGGAWKSPEGFRAACEQLRALQD TTNTVSGLRAQRDYERLPAKYQGVLGAKSAE

UL36 (WT)
RAGAVEELGGRVAQHADLSARLRDEVVPRVAWEMNFDTLGGLLAEFDAVAGDLAPWAVEE
UL36 (MUT)
RAGAVEELGGRVAQHADLSARLRDEVVPRVAWEMNFDTLGGLLAEFDAVAGDLAPWAVEE

UL36 (WT)
FRGARELIQRRMGLYSAYAKATGQTGAGAAAAPAPLLVDLRALDARARASAPPGQEADPQ
UL36 (MUT)
FRGARELIQRRMGLYSAYAKATGQTGAGAAAAPAPLLVDLRALDARARASAPPGQEADPQ

UL36 (WT)
MLRRRGEAYLRVSGGPGPLVLREATSTLDRPFAPSFLVPDGTPLQYALCFPAVTDKLGAL
UL36 (MUT)
MLRRRGEAYLRVSGGPGPLVLREATSTLDRPFAPSFLVPDGTPLQYALCFPAVTDKLGAL

UL36 (WT)
LMCPEAACIRPPLPTDTLESASTVTAMYVITVINRLQLALSDAQAANFQLFGRFVRHRQA
UL36 (MUT)
LMCPEAACIRPPLPTDTLESASTVTAMYVITVINRLQLALSDAQAANFQLFGRFVRHRQA

UL36 (WT)
RWGASMDAAAELYVALVATTLTREFGCRWAQLEWGGDAAAPGPPLGPHSSTRH RVSFNEN
UL36 (MUT)
RWGASMDAAAELYVALVATTLTREFGCRWAQLEWGGDAAAPGPPLGPHSSTRH C VSFNEN

UL36 (WT)
DVLVALVASSPEHIYTFWRLDLVRQHEYMHLTLPRAFQNAADSMLFVQRLTPHPDARIRV
UL36 (MUT)
DVLVALVASSPEHIYTFWRLDLVRQHEYMHLTLPRAFQNAADSMLFVQRLTPHPDARIRV

UL36 (WT)
LPVFSTGGPPTRGLMFGTRLADWRRGKLSETDPLAPWRSVPPELGTERGAALGKLSPAQAL

UL36 (MUT)
LPVFSTGGPPTRGLMFGTRLADWRRGKLSETDPLAPWRSVPPELGTERGAALGKLSPAQAL

UL36 (WT)
AAVSVLGRMCLPSTALAALWTCMFPDDYTEYDSFDALLTARLESGQTLSPSGGREASPPA
UL36 (MUT)
AAVSVLGRMCLPSTALAALWTCMFPDDYTEYDSFDALLTARLESGQTLSPSGGREASPPA

UL36 (WT)
PPNALYRPTGQHVAVPAAATHRTPAARVTAMDLVLAAVLLGAPVVVALRNTTAFSRESEL
UL36 (MUT)
PPNALYRPTGQHVAVPAAATHRTPAARVTAMDLVLAAVLLGAPVVVALRNTTAFSRESEL

UL36 (WT)
ELCLTLFDSRARGPDAALRDAVSSDIETWAVRLLHADLNPIENACLAQPLRLSALIAER
UL36 (MUT)
ELCLTLFDSRARGPDAALRDAVSSDIETWAVRLLHADLNPIENACLAQPLRLSALIAER

UL36 (WT)
PLARGPPCLVLVDISMTPVAVLWENPDPPGPPDVR**F**VGSEATEELPFVAGGED**V**LAAS**S**AT
UL36 (MUT)
PLARGPPCLVLVDISMTPVAVLWENPDPPGPPDVR**V**VGSEATEELPFVAGGED**G**LAAR**R**AT

UL36 (WT)
DED**P**LARAILGRPFDASLLSGELFPGHPVYQRAPDDQSPSVPNPTPGPADLVGTEGSLG
UL36 (MUT)
DED**L**LARAILGRPFDASLLSGELFPGHPVYQRAPDDQSPSVPNPTPGPADLVGTEGSLG
____:_____

UL36 (WT)
PGSLAPTLFTDATPGEPVPPRMWAWIHGLEELASDDSGGPAPLLAPDPLS**T**ADQSVPTS
UL36 (MUT)
PGSLAPTLFTDATPGEPVPPRMWAWIHGLEELASDDSGGPAPLLAPDPLS**P**ADQSVPTS
____:_____

UL36 (WT)
QCAPRPPGPAVTAREARPGVPAESTRPAPVGPRDDFRRLPSPQSSPAPPDATAP**RPPASS**
UL36 (MUT)
QCAPRPPGPAVTAREARPGVPAESTRPAPVGPRDDFRRLPSPQSSPAPPDATAP**PPP****PPP**
____:_____

UL36 (WT)
RAS**AAS**SSGSRARRHRRARSLARATQASATTQGWRPPALPDTVAPVTDFA**RPP**APPKPPE
UL36 (MUT)
RAS**GTV**SSGSRARRHRRARSLARATQASATTQGWRPPALPDTVAPVTDFA**RPP**APPKPPE
____:_____

UL36 (WT)
PALHALVSGVPLPLGPQFAGQASPALPIDVPPPPVATGTVLPGGENRRPPLTSGPAPT**PP**
UL36 (MUT)
PALHALVSGVPLPLGPQFAGQASPALPIDVPPPPVATGTVLPGGENRRPPLTSGPAPT**PP**

UL15 (MUT)
RFKECAPADVVPQRNAYYSVLNTFQALHRSEAFRQLVHFVRDFAQLLKTSFRASSLTETT

UL15 (WT)
GPPKKRAKVDVATHGRTYGTLELFQKMILMHATYFLAAVLLGDHAEQVNTFLRLVFEIPL
UL15 (MUT)
GPPKKRAKVDVATHGRTYGTLELFQKMILMHATYFLAAVLLGDHAEQVNTFLRLVFEIPL

UL15 (WT)
FSDAAVRHFRQRATVFLVPRRHGKTWFLVPLIALSLASFRGIKIGYTAHIRKATEPVFEE
UL15 (MUT)
FSDAAVRHFRQRATVFLVPRRHGKTWFLVPLIALSLASFRGIKIGYTAHIRKATEPVFEE

UL15 (WT)
IDACLRGWFGSARVDHVKGETISFSFPDGSRSTIVFASSHNTNVSPLFFR^RMALPRGP^GST
UL15 (MUT)
IDACLRGWFGSARVDHVKGETISFSFPDGSRSTIVFASSHNTNVSPLFF^GMALPRGP^VST

UL15 (WT)
R^PT^TTH^PHTHT^TTRGRKVCVPGTDFYSGSLEEARATARATVGQLVANRRLMYEEKTHRRH
UL15 (MUT)
P^HT^HTH^THTHT^PPT^TGRKVCVPGTDFYSGSLEEARATARATVGQLVANRRLMYEEKTHRRH

UL15 (WT)
PALVGGMLSAPHRPPTTSCRRSVMQGRRGACSRVTVSTGWPKRSGHRLASPWTARVHWA
UL15 (MUT)
PALVGGMLSAPHRPPTTSCRRSVMQGRRGACSRVTVSTGWPKRSGHRLASPWTARVHWA

UL15 (WT)
LAHRHGATQGQDSVTTEESHSAESMWGGGAQDSQPPGWLVLARSHPRWRALLPGGQS
UL15 (MUT)
LAHRHGATQGQDSVTTEESHSAESMWGGGAQDSQPPGWLVLARSHPRWRALLPGGQS

UL15 (WT) DPRSLMAQQ
UL15 (MUT) DPRSLMAQQ

UL46

UL46 (WT)
MQRRTRGASSRLARCLTPANLIRGDNAGVPERRIFGGCLLPTEGLLSAAVGALRQRSD
UL46 (MUT)
MQRRTRGASSRLARCLTPANLIRGDNAGVPERRIFGGCLLPTEGLLSAAVGALRQRSD

UL46 (WT)
DAQPAFLTCTDRSVRLAARQHNTVPESLIVDGLASDPHYEYIRHYASAATQALGEVELPG
UL46 (MUT)
DAQPAFLTCTDRSVRLAARQHNTVPESLIVDGLASDPHYEYIRHYASAATQALGEVELPG

UL46 (WT)
GQLSRAILTQYWKYLQTVVPSGLDVPEDPVGDCDPSLHVLLRPTLAPKLLARTPFKSGAA
UL46 (MUT)
GQLSRAILTQYWKYLQTVVPSGLDVPEDPVGDCDPSLHVLLRPTLAPKLLARTPFKSGAA

UL46 (WT)
AAKYAATVAGLRDALHRIQQYMFFMRPADPSRPSTDTALRLNELLAYVSVLYRWASWMLW
UL46 (MUT)
AAKYAATVAGLRDALHRIQQYMFFMRPADPSRPSTDTALRLNELLAYVSVLYRWASWMLW

UL46 (WT)
TTDKHVCHRLSPSNRRFLPLGGSPEAPAETFARHLDRGPSGTTGSMQCMALRAAVSDVLG
UL46 (MUT)
TTDKHVCHRLSPSNRRFLPLGGSPEAPAETFARHLDRGPSGTTGSMQCMALRAAVSDVLG

UL46 (WT)
HLTCLANLWQTGKRSGGTYGTVDTVVSTVEVLSIVHHHAQYIINATLTGYGVWATDSLNN
UL46 (MUT)
HLTCLANLWQTGKRSGGTYGTVDTVVSTVEVLSIVHHHAQYIINATLTGYGVWATDSLNN

UL46 (WT)
EYLRAAVDSQERFCRTTAPLFPTMTAPSWARMELSIKAWFGAALAADLLRSGAPSLHYES
UL46 (MUT)
EYLRAAVDSQERFCRTTAPLFPTMTAPSWARMELSIKAWFGAALAADLLRSGAPSLHYES

UL46 (WT)
ILRLVASRRTTWSAGPPPDDMARGPGGHRAGGGTCREKIQRARRDNEPPPLPRPRLHSTP
UL46 (MUT)
ILRLVASRRTTWSAGPPPDDMARGPGGHRAGGGTCREKIQRARRDNEPPPLPRPRLHSTP

UL46 (WT)
APSTRFRRRRADGAGPPLPDADDPVAEPPAAAAQPATYYTHMGEVPPRLPARNVAGPDR
UL46 (MUT)
APSTRFRRRRADGAGPPLPDADDPVAEPPAAAAQPATYYTHMGEVPPRLPARNVAGPDR

UL46 (WT)
RPPAATCPLLVRASLGSLDRPRVWGPAPEGEPDQMEATYLTADD DDDDDARRKATHAASA
UL46 (MUT)
RPPAATCPLLVRASLGSLDRPRVWGPAPEGEPDQMEATYLTRDDDDDDGHGKAFHPAGA

UL46 (WT)
RERHAPYEDDESIYETVSEDGGRVYEEIPWMRVYENVCVNTANAAPASPYIEAENPLYDW
UL46 (MUT)
DERHAPYEDDESIYETVSEDGGRVYEEIPWMRVYENVCVNTANAAPASPYIEAENPLYDW

UL46 (WT)
GGSALFSPPGRTGPPPPPLSPSPVLARHRANALTNDGPTNVAALSALLTKLKREGRRSR
UL46 (MUT)
GGSALFSPPGRTGPPPPPLSPSPVLARHRANALTNDGPTNVAALSALLTKLKREGRRSR

US4

US4 (WT)
MSPGAMRAVVPIIPFLLVLVGVSgvPTNVsSTTQPQLQTTGRPSHEAPNMTQTGTtDSPT
US4 (MUT)
MSPGAMRAVVPIIPFLLVLVGVSgvPTNVsSTTQPQLQTTGRPSHEAPNMTQTGTtDSPT

US4 (WT)
AISLTTPDHTPPMPSIGLEEEEEEEEGAGDGEHLKGGDGTRDTLPQSPGPAVPLAGDDEK
US4 (MUT)
AISLTTPDHTPPMPSIGEEEEEEEEEEEEEEKGGDGTRDTLPQSPGPAVPLAGDDEK

US4 (WT)
DKPNRPVVPVPPGPNNSPARPETS RPKT PPTSIGPLATRP TTQLPSKGRPLVPTPQHTPLF
US4 (MUT)
DKPNRPVVPVPPGPNNSPARPETS RPKT PPTSIGPLATRP TTQLPSKGRPLVPTPQHTPLF

US4 (WT)
SFLTASPALDTL FVVSTVIHTLSFVCIVAMATHLCGGWSRRGRRTHPSVRYVCLPPERG
US4 (MUT)
SFLTASPALDTL FVVSTVIHTLSFVCIVAMATHLCGGWSRRGRRTHPSVRYVCLPPERG

Table S1: APOBEC mutational hotspots identified in RPE-1 pdoxA3A-HA, HEK293-AD pCMVA3A-tGFP, and HEK293-AD pCMVA3G-tGFP cells.

A. RPE-1 p_{dox}A3A-HA hotspots

Hotspots	# C>T	input	A3A-HA ChIP (rep1)	A3A-HA ChIP (rep2)	DRIP (rep1)	DRIP (rep2)
1. 30085-30087	1			+		
2. 50718-50735	7	+	+	+	+	+
3. 57204-57239	6		+	+		
4. 71589-71729	16	+	+	+	+	+
5. 77630-77648	3	+		+		
6. 136819-136928	19		+	+		
total: 52		24	20	38	10	10

B. HEK293 p_{CMV}A3A-tGFP hotspots

Hotspots	# C>T	input	A3A-tGFP ChIP	DRIP
1. 30033-30091	4		+	
2. 50718-50735	7	+	+	+
3. 57143-57260	19		+	
4. 71589-71729	16	+	+	+
5. 74144-74146	1		+	
6. 77630-77648	3	+		+
7. 136756-136950	46		+	+
total: 96		23	80	13

C. HEK293 p_{CMV}A3G-tGFP hotspots

Hotspots	# C>T	input	A3G-tGFP ChIP	DRIP
1. 30085-30087	3		+	
2. 50718-50735	7	+	+	+
3. 57204-57239	7		+	
4. 71589-71729	14	+	+	+
5. 77630-77648	4	+	+	+
6. 136819-136928	28	+	+	
total: 63		25	48	11

(A) RPE-1 pdoxA3A-HA hotspots. Mutational hotspots identified in doxycycline-inducible A3A-HA expressing RPE-1 cells. Each row represents a hotspot region with coordinates indicating HSV-1 genome positions. The #C→T column shows the number of C→T mutations within each hotspot. Plus signs (+) indicate enrichment above background in input DNA, A3A-HA ChIP-seq (two biological replicates), and DRIP-seq (two biological replicates). The total row shows cumulative values across all hotspots for each dataset (n = total mutations or number of hotspots with enrichment). (B) HEK293 pCMVA3A-tGFP hotspots. Mutational hotspots identified in HEK293-AD cells constitutively expressing A3A-tGFP. Table structure as in panel A. (C) HEK293 pCMVA3G-tGFP hotspots.

Mutational hotspots identified in HEK293-AD cells constitutively expressing A3G-tGFP. Table structure as in panels A and B.

Table S2: List of qPCR primers used in this study.

Oligo name	sequence
HSV-1_1_F	ATTGCCGTTGGGGTTTTGTG
HSV-1_1_R	CGTAACGCCAATCAAGATCGTC
HSV-1_2_F	AGGCGGTTGACGGTCAAATG
HSV-1_2_R	CCCACATACTTGACGCTGGAG
HSV-1_3_F	TTACGTGGCTGGGTTTTTGG
HSV-1_3_R	TTAATCGGGAGTGGGTTATCCG
HSV-1_4_F	TGTCACGGGAAAGGAAAGAGG
HSV-1_4_R	ACGCCTTTTGTGTGTGTGTG
HSV-1_5_F	TGCGTCTCTGGGTGTTTTTG
HSV-1_5_R	CCCGGGCATGTATTAACTGC
HSV-1_6_F	AACGCACCAAACAGATGCAG
HSV-1_6_R	TGTTGATGGCCTGTGCTTTC
ZNF554_FW	CGGGGAAAAGCCCTATAAAT
ZNF554_REV	TCCACATTCACCTGCATTCGT