

Supplemental Text

Full-length sequence of HTLV-1 provirus integrated in TL-Om1 cells

Size : 8941 bps

TGACAATGACCATGAGCCCAATATCCCCCGGGGCTTAGAGCCTCTAGTGAAAAACATTTCCGCGAAACAGAAGTCTGAAAAGGTGAGGGCCAGACTA
AGGCTCTGACGTCTCCCCCGGAGGGACAGCTCAGCACCGGCTCAGGCTAGGCCCTGACGTGTCCCCCTGAAGACAAATCATAAGCTCAGACCTCCGGGA
AGCCACCGGAACCCCAATTTCCCTCCCATGTTTGTAAAGCCGCCCTCAGGCGTTGACGACAACCCCTCACCTCAAAAAAATTTTCATGGCAGCATATGCG
TGAATAAATACAGGAGTCTATAAAAGCGTGGAGACAGTTTACGAGGAGGCTCGCATCTCTCTTACGCGCCCGCCGCTTACCTGAGCGCCATCA
CGCCGGTTGAGTCGCTTCTGCCGCTCCCGCTGTGGTGCCTCTGAACCTCGCTCCGCGCTAGGTAAGTTTAGAGCTCAGGTCGAGACCGGGCTTT
GTCCGCGCTCCTTGGAGCTACCTAGACTCAGCCGCTCCACGCTTTGCCCTGACCTGTCTCAACTCTCGCTCTTGTTCGTTTTCTGTTCTGC
GTCCGTACAGATGGAAGTTCCACCCCTTCCCTTTCATTACAGACTGACTGCCGCTTGGCCACGGCCAAAGTACCGGACTCCGTTGGCTGGAGCCA
GGACAGCCCATCTATAGCACTCTCCAGGAGAGAACTTAGTACACAGTTGGGGGCTCGCCGGGATACGAGCGCCCTTTATCCCTAGGCAATGGGCCA
AACCTTTTCCCGTAGCGCTAGCCCTATTTCCGCGCCCGCCCGGGGCTGCGCGCTCATCACTGGCTTAACCTTCTCCAAGCGGCATATCGCTAGAACCCG
GTCCCTCCAGTTACGATTTCCACAGTTGAAAAAATTTCTAAAAATAGCTTTAGAAAACCGGCTCGGATCTGTCCCAATAACTCCCTCCTAGCCAGCCTAC
TCCCAAAGGATACCCCGGCGGTGAATGAAATTTACACATACCTCAAAACCCAGCCAGATCCCGTCCCGTCCCGCCACCCGCGCTCATCCC
CCACCCAGACCCCGGATTTCTGATCCAAATCCCCCTCCTATGTTGAGCTACGCGCCCAAGTCTTCCAGTCATGACCCACATGTTGCCCTC
CCAACCTAGCCCATGGAAATGAAAGCCTACAGCCATTAAGCAAGAGTCTCCAAGCAGCCCTGGAGCCCCAGTTTATCGACACCTACCGGCTT
CGGTCAGCAGTTTGGCCCACTCCAAAGACCTCCAAAGACCTCCGAGCTGACGTACCTTTGCTCCCTCCCTGCGCTTCCCTCCATCACCAGCAGTATGATGC
TTATATCGGAGGCCGAACCCGAGGTTATACAGTTATAACCCCTAGCCGCTCCCTCCGTTCCAAAGCCAAATCCCAACAACAAGGATTAAGCGGAGA
ATACAGCAACTCTGGCTCGCCGCTTCCGCGCTTCCGAGGAGTCCCAAGACCCCTTCCGCGCTCTATCCGAGGCTGGAGGAGCCTTACCC
GCCTCGTAGACCGCTCAACATAGCTTGTGACAATGGCTCCGAGAGGACCGCCAAAGACCCCATCTACGTTCTTACGCTACTCCATCAACAAAGT
GAATGCCAAAATTACTACAGCCCGGAGACACTAATAGCCCTTAGGAGATATGTCGGGCTTGTGACAGCTGGACCCCAAGACCAAAACAAAGTG
TTAGTTGTCCAGCTTAAAAACCCCGCCAAATAGCCGCTGCTCCGGTCCGGGAAAGCAGGCCACTGGAGTCGGGACTGCACTACGCTCCGCCCC
CCGGCCATGCCCTATGTCAGACCCCACTACTGGAAGCGAGACTGCCCGCTTAAAGCCACTATCCCAAGCAGGACGACGAGGAGGAGGCTCCCTC
CTATTAGACTTCCCGCTGACATCCCAACCCCAAAACCTCCATAGGCGGGGAGGTTTAACTCCCGCCACATACAGCAAGTCCCTTAAACAGAGCC
AGCATATTTTCCGAGTTATACCGTTAGATCCCGCCCGCTCGGCGGTAATTAAGCCCAAGTTGACACCCAGACCCAGCCCAAGACTTCAAGGCTTTA
CTAGATACAGGAGCAGACATGACAGTCTTCCGATAGCCTTGTCTCAAGTAATCTCCCTCAAAAAATACATCCGATATTAGGGCAGGGCCAAACCGA
ATCACTTTAAGCTACCTCCCTTCCGTGCTAATAGCCTCCCTTCCGACACAGCCTATTGTTTAAATATCTTCCGATGTTATACCAAAACAACTGGGCCA
TCATAGGTCGCGATGCTTCAACAAATGCCAGGCGCTCTGTACCTTCCGAGGCGCTTGTGAGGCTTGTGACAGCTGGACCCCAAGACCAAAAGTG
GGCTAGAACACCTCCCAAGCCCGCCGAATAGCCAGTTCCTTAAACCCAGAACGCTCCAGGCTTGAACACTTGGTCCGAAAGGCCCTGGAGGCA
GGCCATATGAACCTACACCCGACCCAGGAAATAACCCAGTATCCAGTAAAAAGGCCAATGAAACCTGGCGATTATCCACGACCTCGGGCCACTAAC
TCTTAACCATAGATCTCTCATCATCTTCCCGCGGCCCTGACTTGTCCAGCCTGCCAACTACACTAGCCCATTTGCAAACTATAGACTTAAAGAGCCCT
TTTTCAAAATCCCTTACCTAAACAGTTCCAGCCCTACTTTGCTTCACTGCTCCACAGCAGTGTAACTACGCGCCCGGCATAGATACGCTGGAAGGTA
CCCAAGGTTTAAAAATGCTCCACCTTGTTCGAAATGACAGTCCGCTTACCTGACGCTTCCGCAAGCTTCCGCAAGCTTCCCAAGCTTCCGATCT
GGATGACATTTCTCCGAGCCCTCCCATAGGACCTACTACTCTCAGAGGCCCAATGCTTCCCTAATCTCCCAAGGTTGCTGTGTCCGAAAA
CAAAACCCAGCAAGCTTGAACAAATTAAGTTCCTAGGGCAAAATAGTTCAACCAACCTGACTTATGATGACGCTCCCAAGCTTCCGATGCTCCG
TGGCGCTACCTGAACTTCAAGCCCTACTTGGCGAGATTGAGTGGTCTCAAGGGAACCTCTACCTTACGCCAGCCCTTCCAGTCTCTACTGTCCCTA
CAAAGGCATAGTCCCGAGACCAATATTAATCCTTCTCAAGTTCAATATTAGTGCAGCTCGGGCAGCCCTGTACAGAACTGCCGAGTAGACT
AGTCCAAACCTGCCCTTAGGGGCTATTATGCTGACCCCTACTTGTCCAGCCTGCCAACTACTGTAGTGTTCCAGTCCAAAGCAGCAGTGGCCACTTGT
TGCCCTACCCCACTACCCACACTGACAGTGCCTCGGGGAGCTACTTGGCTCAGCTGTGTTTACTCGACAATAACACTGCAATCCATGGCTACTCTGC
CAAACCCATCATAACACTTCCAGCCCAACCTTCAACCAATTTCAACCAACTGACCCACCTGACCCACCCAGTGTCCCTACTTACTCCACCACTACCCGAT
GAATGAGTGCACCAACTGGAGAATTTGGAACACTTTTCAAAACAGCTGCCCTTGGCTCCTGTGAAAGCCCTCATGCCAGTTTACTCTTCCCGGG
TGATCAAAAACCCGCCCTGCTGTTTTAGACGGATCTACCTCCGCGGACGCTATATTCTTGGGACAAAGCTATATTGTCAGAAAGTACTTCCCTCT
CCGCCACCCAGCAAGTCCGCGCAACGGCCGAATCTCGGACTTTGATGGCTTCCAGCCCTTCCAGCCCGCTTCCGTCGCTGTCTCAACATTTCTAG
CAAGGCTATCTTACTTATGACTTCCGACCTTGGCCCTAGGCACCTTCCAAGCGAGTCTCCTCAGGCCCTTCCAGGCCCTTCCGCGCTTACTACTG
AAGTCGCTTATTTGACCCAGTTCGACAGCATAACCTATCCCTGATCCCATCTCCAGGCTCAACGCTCACAGATGCCCTACTAATACCCCTGTCTCG
AGCTCTCTCCTGAGAACTACACAGTTTACCCTATGGGACAGAGGCCCTCACATGCAAGGGGCAACCAACTGAGGCTTCCAAATACCTCGCGCTT
GCCACGCTCCGCCAAAATAACCCCAACATCAGATGCTCGGGGACACATCCGCGTGGCCACTTCCCTAACCCACATCTGGCAGGGCGACATACCCAT
TCAATATAAAAATCACTGTATCGCTTCTATGATGGTAGACACTTAAACAGGCTTCCAGGACCATCTCAGCTACCCAAAAGAGAAAAGAAAGCTCAGAA
CCTTCTTTGCTTAGGCCATTTGCTATCTAGGCAAGCTTAGCTACATAAACACAGACAAAGCCCTGCCTATATTTCCCAAGACTTCCAAATGTAGTACT
CTTGCTATTGCGCATACTACCCATGTCCTTCAATCCAACAGCTCAGGACTGTAGAACGCTCTAATGGCATTCTTAAACCCCTATTATATAAGTCTTTACTG
ACAAACCCGACCTCCCATGGATAATGCTCTATCCATAGCCCTTAGGACAATCAACCCCTGAATGTGTTAAACCACTGCCACAAAACCCGATGGCAGCTTCA
CACTCCCGCCGACTCCAGCCGATCCAGACACTCCCTCAGCAATAAACCAACCCATTTGGTATTATTTCAAGTTTAAATAGCCCTGTTCTGACTGAC
AAGGACACAGGAGGCTTCCAAGAAGCTCGCGGCGCTGCTCTATCCGGTAAGCGCTAGTTCTGCCAGTGGATCCCGTGGAGACTCCTCAAGGGAGC
TGATGCCCAAGACCCGTCGGAGGCGCCCGGATCCCAAGAAAAGACCAACACCCATGGGTAAGTTTCTCGCCAATTTGATTTTCTTCCAGTTCT
GCCCTCTATCTCGGTGATTACAGCCCAAGCTGCTGACTCTCAAAATTTGAGGTCCTCTCATACCCTTAAACCCCTGAATCCTCGCCAGCCAGTTGCTT
GTGAGCCCTCGACCTCGCCGCTTTCAGCAGTACGCGCTTACACCCCTGCTTAACTTCTAGTAAAGTTACTCCAGCTACCTGACATTTCCCTAT
CTATCCCTCATTGGATAAAAAGCCAAACCGAAATGGCGGAGGCTATTATTAGTCTCTTATTAGACCTTGTCTTAAAGTGGCCATACCTGGGTTGCCA
ATCATGGACCTGTCCCTATAATGGGTTTTCCCTTCTCCCTTCTAGTGCAGCCTCCAGGATAGCCCATCTGGTTCTTAATACCGAACCAGCCCACTG
CTCCACCCGCTTCTTACTCCCACTTCAACCTAGACCACTCTCGAGCCCTTATACCATGGAATCAAAACTCTGACCCCTGACCCCTTCCAGTTAACCC
CAAAGCACTAATATACTGATTTGCTGTATCGATCGTGCAGCCTATCCACTTGGCAGCTCTATCCCAACGCTCTGTTCCATCTCTTCTTCTTACC
CCCTCCTTTTACCATCGTTAGCGCTTCCAGCCCGCCACTCAGCTTACCTTACCTGGACCCACTGCTTGGACCCCACTCAAGCTTATGCTCTCTCC
CCTGTATAAATCCCTCATCTCGCCCTTTTCCCTGTCACTTCCACCTTCCACCTAGGATCCCGCTCCCGCCGAGCGGTACCGGTTGGCGTGTGCT
CCGCTGGCCTGGGAGCGGGGTTGCTGGCGGATTACGGCTTCCCTCGCTCAGGGAAGGAGCCTTACATGAGTGGAAAGATTTCT
CCAATTAACCAAGCAATAGTCAAAAACCAAAAATCTACTCAAAATTTGCCAGTATGCTGCCAGAACGACGAGGCTTGTATCTCTGTCTGGGAGCAA
GGAGGATTATGCAAGCATTACAAGAACAGTGTCTTTTCTGAATATTACTAATTTCCCATGCTCAATACTACAAGAAAGACCCCGCTGGAGAATCGAGTCT
GACTGGCTGGGGCTTAACTGGGACCTTGGCTCTCACAGTGGGCTCGAAGGCTTACAAACTGGAATCACCCCTTGTCCGCTACTCCTTTCTGTTATCT
TGCAGGACCATGATCCTCGTACGCTACGACACCTCCCTCGCGCTCAGATACCCCACTTACTCTTATAAACCCCTGAGTATCCCTGTAAACCAAGCAC
ATAATTTGCAACCACTCGCTCCAGCCCTCCCTGCCAATAATTAACCTTCCCATTAATCCTCCTTCTCCTGCAGCACTTCCCTGCTCCGCTCCAA
GACTCCACTCGCTTCCAATCTGTAGTATAGCCATCAATCCCAACTCTGCTGATTTTCTTCTTAGCACTATGCTGTTTCCGCTTCTCAGCCCTTGTCT
CACTTGCCTCAGCGCTCCTGCTTCTCTTCTCCGCGCAGCTCAGCGCTTCTTCCGCGCGCTCCTGCGCGCTTCTCTCTCTCTCTCTCTCTCTCT
CCTTTTCAAACTCAGCAATCTGCTTTTCTCTCTCTTCTCCCGCTCTTTTTTTCGCTTCTCTCTCTCAGCCCGTGCCTGCGGATCAGATGCTTTCC
CGCGAGGTGGCGTTTTCTCCCTGGAGGGCCCGCTCGCAGCCGCGCGCGCTTCTCTTCTAAGGATAGCAAAACCGTCAAGCACAGCTTCTCTCTCT
CTTGTCTTAACTCTTCCCAAGGATAATAGCCCGTCCACCAATTTCTCCACAGCAGGCTCCTCGGGCATGGCACAGGCAAGCATCGAAACAGCCCTAC
AGATAAAAATTAACCTGCTTATATACGCCCACTTCCAGGGTTTGGACAGAGCTTCTTTTCCGATACCCAGTCTACGTTTGGAGACTGTGACAAAGC
GACTGGTCCCTCTCTGAGGAGGACTGTTCCGCGCGCTACATCGTACGCGCTTCCATTAATCCTCCTTCTCCTGCAGCACTTCCCTGCTCCGCTCCGA
TGGACGGTATACCGCTCAGCTTACAGTTCTTATCCTCGACTCCCTCTTCCCAACCCAGAGAACTTCCGAATGGATACATGGAACCCCACTTGGCAGC
ACTCATAACCCCACTTCCACTTCCCTCCAGCCATGCGCAAACTCCCTTCCGAATGGATACATGGAACCCCACTTGGCAGCAGCTTCC
CAACCTGTCTTTTCCAGACCCCGGACTCCGGCCCAAAAACCTGTACACCTCTGGGAGGCTCCGTTGTCTGCATGTACTTACCAGCTTACCAGCTTCC
TCACTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCT
AAATTTCCCTTACCAACAGGGCTTAATAATTTTCAACGAGACTGTTGCCAACCCCTTTCCAGCTTATAGGGACCCGATCGAAACAGCCCTGGCA
AAACGCTCTTCCGTTCCACTCAACCTCACCCTCCAGGCTTATTTGACATTTACCGATAGCAGCGCTATGATTTCCGGCCCTCCCTAAAGATGGC
CAGCCTTTTACTACAGCTCCTCCTCTTATATTTCAAAATTTCAAAACCCAGGCTTACCACCCCTCATTCTACTCTCAGACCGCTTCCATACGACTCT
CCTTTTCAATTTTACATCTCTGTTTGAAGAATACCAACATCCCAATTTCTACTTTTAAACGAAAAGAGGAGATGACATGACCATGAGCCCAATAT
CCCCGGGGCTTAGAGCTCCAGTGA AAAACATTTCCGCGAAACAGAAGTCTGAAAAGTCAAGGCCCAGACTAAGCTGTAGCTTCCCGGGAGG
GACAGCTAGCACCCGGCTCAGGCTAGGCCCTGACGTGTCCCTCGCAAGGACAAATATAAGCTCAGACTCCGGGAAAGCCAGCCAGCCACCCATTTCTCT
CCATGTTTGTCAAGCCCGCTCAGCGCTGACGACAAACCCCTCACTCAAAAACATTTTTCATGGCAGCATATGGCTGAATAAATCAACAGAGTCTATAAAA
GCGTGGAGACAGTTCAGGAGGAGGCTCGCATCTCTTCCAGCGCCCGCCCTTCCAGGCGGCTTCCAGGCGGCTTGAAGTGCAGCTTATGCTTCCGCG
CTCCGCGCTGTGGTCTCCTGAACTCGCTCCGCGCTTAGGTAAGTTTAGAGCTCAGGTCAGGACCGGGCTTGTCCGCGGCTCTCTTGGAGCCTACT
AGACTCAGCGGCTTCCAGCTTGTGCTGACCTGCTTGTCAACTGTGCTCAAGCTTTGTTTTCGTTTTCTGTTCTGCGTGCATGAGTGTCCACCC
CTTCCCTTTCATTACGACTGACTGCCGCTTGGCCACGGCAAGTACCGGCGACTCCGTTGGCTGGAGCCAGCGACAGCCATTCTATAGCACTCTC
CAGGAGAGAACTTAGTACAC

Supplemental Fig. 1

A

```

1' MGKFLATLIL FFQFCPLIFG DYSpscctLT IGVSSyHskP CNPAQpVCSW TLDLLALSAD
*****
1" MGKFLATLIL FFQFCPLILG DYSpscctLT IGVSSyHskP CNPAQpVCSW TLDLPALSAD

61' QALQPPcPNL VSYSSyHATY SLYLFPPhWTK KPNRNgGgYy SASySDpCSL KcPYLgCQSW
***,*****
61" QALHPPcPNL VSYSSyHATY SLYLFPPhWIK KPNRNgGgYy SVSySDpCSL KcPYLgCQSW

121' TCPyTGAVSS PYWKFQHDVN FTQEVsRLNI NLHfSKCGFP FSLlVDAPGY DpIWfLNTEP
****
121" TCPY-----KCGFP FSLlVDAPGY DpIWfFNTEP

181' SQLPPTAPPL LPHSNLDHIL EPsIPWkSKL LTLVQLTLQs TNYTCIVCID RASLSTWHVL
*****
150" SQLPPTAPPL LPHSNLDHIL EPsIPWkSKL LTLVQLTLQs TNYTCIVCID RASLSTWHVL

241' YSPNVsVPSS SSTPLLYPSL ALPAPHLTLp FNWTHCFDPQ IQAIVSSpCh NSLILPPFSL
*****
210" YSPNVsVPSS SSTPLLYPSL ALPAPHLTLp FNWTHCFDPQ IQAIVSSpCh NSLILPPFSL

301' SPVPTLGSRS RRAVPVAVWL VSALAMGAGV AGGITGSMsL ASGKsLLHEV DkDISQLTQA
*****
270" SPVPTLGSRS RRAVPVAVWL VSALAMGAGV AGGITGSMsL ASGKsLLHEV DkDISQLTQA

361' IVKNHKNLLK IAQYAAQNRR GLDLLFWEQg GLCKALQEQC RFPNITNSHV PILqERPPLE
*****
330" IVKNHKNLLK IAQYAAQNRR GLDLLFWEQg GLCKALQEQC CFLNITNSHV SILqERPPLE

421' NRVLtGWGLN WDLGLSQWAR EALQTGITLV ALLLLVILAG PCILRQLRHL PSRVRYPHYS
*****
390" NRVLtGWGLN WDLGLSQWAR EALQTGITLV ALLLLVILAG PCILRQLRHL PSRVRYPHYS

481' LIKPESSL
**,*
450" LINPESSL
450" LINPESSL
  
```

1st Amino Acid Sequence

File Name : ATK-1 Env_1AA
Sequence Size : 488

2nd Amino Acid Sequence

File Name : TL-Om1 Env_1AA
Sequence Size : 457

Unit Size to Compare : 2

Percent Similarity : 99.3(447/457)

Percent Identity : 97.8(454/457)

Env amino acid sequence domains

1-21 SP

21-313 SU

21-179 receptor binding domain (RBD)

180-215 proline-rich region (PRR)

215-313 SU C-term

313-488 TM 446-465 tm

Env sequence of HTLV-1 provirus isolated from TL-Om1 cells

Complete sequence of HTLV-1 from TL-Om1 cells showed 93 nt (31 aa) inframe deletion in the env sequence. A.

Comparison of the env sequences to ATK-1. The deleted region contained aa125-155 of ATK-1 Env and was located in the RBD.

Supplemental Table 1. Genomic long PCR primers

Forward Name	sequence	Reverse Name	Sequence	size (base)	regions
5' gPCR 1F	CAGTTCCCATGATCTGAGTCTAGGTCC	41R	GGGAGGTGAGCTTAAAGTGATCTT	3964	5' integration site, LTR, coding
18F	CCAAGTCCTTCCAGTCATGCA	188R	ATGGTCATTGTCATCTGCCTCTTT	6950	coding
174F	TGGACAGAGTCTTCTTTTCGGATAC	3' gPCR 2R	ACAAAAGCAGGAGGCGTCACACTAC	4065	LTR, 3' integration site

Supplemental Table 2. Primers for sequencing analysis

Name	Sequence	Name	Sequence
Determination of HTLV-1 integration site			
5'-seq 1F	GGGAACATGGATGGAGCTAG	5'-seq 2F	CTTATAAGTGGGAGCTAAATGATGAG
5'-seq 3F	ATCTTGCTGCTGCTCATTG	5'-seq 4F	ACTCCATTTGCCTTCCCTC
70F	ACAAATACACCTTGCAATCCTATGG	78F	CACCTCCCAGGCAGCCTATA
101F	TGCCTATATTTCCCAAGACTTCCT	5R	TCAGTCGTGAATGAAAGGGAAAG
18R	CGCTTGGGAGACTTCTTGCT		
Determination of HTLV-1 sequence			
18F	CCAAGTCCTTCCAGTCATGCA	18R	CGCTTGGGAGACTTCTTGCT
20F	CCCCAGTTTATGCAGACCAT	20R	CCACGAGGGAGGAGCAAA
33F	AATGCAAACAAAGAATGCCAAAA	33R	GGCTGGACAACAACTTTGGT
41F	ACAGTCCTTCCGATAGCCTTGTT	41R	GGGAGGTGAGCTTAAAGTGATCTT
47F	TCGTGATGCCTTACAACAATGC	47R	GGAGGTGTTCTAGCCCAAGGA
60F	CCGAAAACAAAACCCAGCAA	60R	CCAGCGGGACCGTATAGGTA
70F	ACAAATACACCTTGCAATCCTATGG	70R	GGGTGGTCAGATGTTTGAATGA
84F	CCTGCCCGCTTACTATCG	84R	GGCATCTGTGAGAGCGTTGA
101F	TGCCTATATTTCCCAAGACTTCCT	101R	TTCTACAAGTCCGGAGCTGGTT
111F	CCCCTCATCCTCGGTGATTAC	111R	CGAGGGTCCATGAACAAACTG
120F	CATTGGATCAAAAAGCCAAACC	120R	ATGATTGGCACCCCTAGGTATGG
139F	CACTGCTTTGACCCCGAGAT	139R	GGATCCTAGCGTGGGAACAG
156F	GCATCCTCCGTCAGCTACGA	156R	TTGCAATAATTGTGTGCTTGGTTT
161F	CCTCGCCTTCCAAGTGTCTAGTA	161R	TGAGCGCAAGTGGAGACAAG
174F	TGGACAGAGTCTTCTTTTCGGATAC	174R	TAGGCGGGCCGAACATAG
180F	GGCCTTCCCTACCAATGTTC	180R	GCAAACAGTCCTCGGGTAGAAT
187F	CAAGGCCTACCACCCTCAT	188F	GTACTCTTCCTTTCATAATTTGCATCTC
3'-1R	ACAAAAGCAGGAGGCGTCACACTAC		

Supplemental Table 3. Summary of gene copy numbers (N) determined in this report.

Method	Cells	Target genes								
		RPPH1	ALB	ACTB	CD81	HBB	RAG-1	HTLV-1	LTR	
FISH	TL-Om1	<u>3.95</u>	3	-	-	-	-	1.8	-	
	Jurkat	2	2	-	-	-	-	-	-	
Digital PCR	TL-Om1	<u>3.95</u>	2.93	-	-	-	-	2.00	-	
	Jurkat	<u>2</u>	1.72	-	-	-	-	-	-	
	PBMC1	2	2.17	-	-	-	-	-	-	
	PBMC2	2	1.99	-	-	-	-	-	-	
q-PCR Plasmid	TL-Om1	Mean	<u>3.95</u>	2.94	-	-	-	-	1.91	4.01
		SD	-	0.01	-	-	-	-	0.06	0.18
	Jurkat	Mean	<u>2</u>	1.84	-	-	-	-	-	-
		SD	-	0.04	-	-	-	-	-	-
q-PCR gDNA	TL-Om1	Mean	<u>3.95</u>	2.92	4.65	3.99	3.65	3.71	-	-
		SD	-	0.02	0.03	0.02	0.01	0.01	-	-
	Jurkat	Mean	<u>2</u>	1.90	2.14	1.98	1.79	2.16	-	-
		SD	-	0.01	0.02	0.01	0.01	0.02	-	-
	PBMC 1	Mean	2	1.98	1.99	1.96	1.98	2.00	-	-
		SD	-	0.01	0.04	0.03	0.02	0.01	-	-
	PBMC 2	Mean	2	2.02	2.02	1.98	2.00	2.02	-	-
		SD	-	0.04	0.05	0.04	0.05	0.04	-	-

Underlined number is the result from FISH analysis.

$N = (\text{FISH}) \times 2^{-\Delta\text{Ct}}$, $\Delta\text{Ct} = \text{Ct}(\text{Target gene}) - \text{Ct}(\text{RPPH1})$