I		gGagl			gGag3	
		↓		•		
#	GTTAAAAGAGACCATCA	ATGAGGAAGCTGCAGAA		AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	
1	GTTAAAAGAGACCATCA	ATGAGGAAGCTGCAGAA	WT	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
2	GCTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT	AGATGAGAGTA CCA AGG	AAGTGACATAGCAGG	-2
3	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
4	GTTAAAAGAGACCATCAC	ATG AGG AAGCTGCAGAA	+1	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
5	GTTAAAAGAGACCA	-TG AGG AAGCTGCAGAA	-4	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
6	GTTAAAAGAG	AAGCTGCAGAA	-13	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
7	GTTAAAAGAGACCATCA	ATG GGA AGCTGCAGAA	NS	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
8	GTTAAAAGAGAGGA	AGCTGCAGAA	-10	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
9	GTTAAAAGAGACCA	GGCAGAA	-13/NS	AGATGAGAGAACCAAGG	-GAAGTGACATAGCAGG	-1
10	GTTAAAAGAG	GAAGCTGCAGAA	-12	AGATGAGAGAACCAAGG	AAGTGACATAGCAGG	-2
11	GTTAAAAGAGACCATC-	-TG AGG AAGCTGCAGAA	-2	AGATGAGAGAA CCA AGG	GTGTGACATAGCAGG	-2/NS
12	GTTAAAAGAGACCA	GAG	-17/NS	AGATGAGAGAA CCA AGG	GAGTGACATAGCAGG	-2/NS
13	GTTAAAAGAGACCATC-		-50	AGATGAGAGAA CCA AGG	<u>c</u>	NS/-23
14	GTTAAAAGAGACCTA	GAGGAAGCTGCAGAA	NS/-4		AGTGACATAGCAGG	-24
15	GTTAAAAGAGACCAGGCC	ATG AGG AAGCTGCAGAA	NS/+1	AGATGAGAG		-44
16	GTTAAA	GAGGAAGCTGCAGAA	-13		GGAAGTGACATAGCAGG	-28
17	GTTAAAAGAGACCATCA	-TGAGGAAGCTGCAGAA	-1	AGATGAGAGAACCAAGG	GCGAAGTGACATAGCAGG	+1
18	GTTAAAAGAGACCACT-	GAGGAAGCTGCAGAA	-3	AGATGAGAGAACCAAGG	-GAAGTGACATAGCAGG	-1
19	GTTAAAAGAGACCATCA				GGAAGTGACATAGCAGG	EX
20	GTTAAAAGAGACCATCA				GGAAGTGACATAGCAGG	EX
21						EX -156
	00003333303030300					
22	GTTAAAAGAGACC				AGCAGG	EX -15
22 23	GTTAAAAGAGACC GTTAAAAGAGAGGAAGCTTAAC	<u>A</u>			GGAAGTGACATAGCAGG	EX -15 EX +4
		A				
23		<u>A</u> gGag1				
23 II	GTTAAAAGAG <mark>GAAGCTTAAC</mark>	gGag1		ļ	GGAAGTGACATAGCAGG gGag3	
23 II #	GTTAAAAGAGGAAGCTTAAC	gGag1 Atg AGG AAGCTGCAGAA		AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG	EX +4
23 II # 1	GTTAAAAGAG <mark>GAAGCTTAAC</mark> GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA	gGag1 ATG AGG AAGCTGCAGAA ATG AGG AAGCTGCAGAA	WT	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4
23 11 # 1 2	GTTAAAAGAGGAAGCTTAAC GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA	WT	AGATGAGAGAA CCA AGG AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4
23 II # 1 2 3	GTTAAAAGAGAAGCTTAAC GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGGGACCATCA GTTAAAAGGGACCATCA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA	WT WT	AGATGAGAGAACCAAGG AGATGAGAGAGACCAAGG AGATGAGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4
23 II # 1 2 3 4	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA	WT WT WT	AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGA	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31
23 II 1 2 3 4 5	GTTAAAAGAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GGTAA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA AAGCTGCAGAA	WT WT WT -18	AGATGAGAGAACCAAGG AGATGAGAGAGACCAAGG AGATGAGAGAGACCAAGG AGA AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31 WT
23 II # 1 2 3 4 5 6	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA	WT WT -18 -3	AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGA	GGAAGTGACATAGCAGG gGaaG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31 WT WT
23 II 1 2 3 4 5	GTTAAAAGAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GGTAA	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA AAGCTGCAGAA	WT WT -18 -3 -14	AGATGAGAGAACCAAGG AGATGAGAGAGACCAAGG AGATGAGAGAGACCAAGG AGA AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31 WT
23 II # 1 2 3 4 5 6	GTTAAAAGAGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA AAGCTGCAGAA ATCAGGAAGCTGCAGAA	WT WT -18 -3	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGA AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGaaG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31 WT WT
23 II # 1 2 3 4 5 6 7	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA	WT WT -18 -3 -14	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4 WT WT -31 WT WT WT
23 II 4 5 6 7 8	СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТ СТТАЛАЛСАСАТСА	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA AAGCTGCAGAA ATGAGAAGCTGCAGAA CAGAA CAGAA	WT WT -18 -3 -14 -10	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGA	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AAGTGACATAGCAGG	EX +4
23 II 2 3 4 5 6 7 8 9	GTTAAAAGAGAACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACC	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA AAGCTGCAGAA ATGAGAAGCTGCAGAA CAGAA CAGAA	WT WT -18 -3 -14 -10 -4	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAACCAAGG 	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AGTGACATAGCAGG CGGACATAGCAGG	EX +4 WT WT -31 WT WT -2 -36 -2
23 II 1 2 3 4 5 6 7 8 9 10	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACC GTTAAAAGAGACC GTTAAAAGAGACC GTTAAAAGAGACC	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA 	WT WT -18 -3 -14 -10 -4 -17	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGA- AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG 	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AGTGACATAGCAGG AGTGACATAGCAGG AGTGACATAGCAGG	EX +4 WT WT -31 WT WT -2 -36 -2
23 II 4 5 6 7 8 9 10 11	СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАССАТСА СТТАЛАЛСАСАССАТСА СТТАЛАЛСАСАССАТСА СТТАЛАЛСАСАССАТСА СТТАЛАЛСАСАССАТ СТТАЛАЛСАСАССАТ СТТАЛАЛСАССАТС СТТАЛАЛСАССАТСА СТТ	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA 	WT WT -18 -3 -14 -10 -4 -17 -16/NS	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG 	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG 	EX +4
23 II # 1 2 3 4 5 6 6 7 8 9 9 10 11 12	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA 	WT WT -18 -3 -14 -10 -4 -17 -16/NS NS/+7	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAAC	GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AAGTGACATAGCAGG GGAAGTGACATAGCAGG AAGTGACATAGCAGG GGAAGTGACATAGCAGG	EX +4
23 II # 1 2 3 4 4 5 6 7 8 9 10 11 12 13	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCATCA GTT GTTAAAAGAGACCATCA GTT-AAAAGAGACCATCA GTT-AAAAGAGACCATCA	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA CAGAAGCTGCAGAA ATCAGGAAGCTGCAGAA TAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA	WT WT 18 3 14 10 4 17 16/NS NS/+7 NS	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAAAC	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AAGTGACATAGCAGG AGTGACATAGCAGG GGAAGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG	EX +4
23 II # 1 2 3 4 4 5 6 7 8 9 10 11 12 13 14	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATC- GTTAAAAGAGACCAT GTTAAAAGAGACCATC GTTAAAAGAGACCATCA GTT-AAAGAGACCATCA GTT-AAAAGAGACCATCC GTTAAAAGAGACCATCC	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA 	WT WT -18 -3 -14 -10 -4 -17 -16/NS NS/+7 NS/-7 NS -31	AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG gGag3 GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AGTGACATAGCAGG AGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG AAGTGACATAGCAGG AGTGACATAGCAGG 	EX +4
23 II # 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15	GTTAAAAGAGAGACCATCA GTTAAAAGAGAGACCATCA GTTAAAAGAGAGACCATCA GTTAAAAGAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA GAGGAAGCTGCAGAA GAGGAAGCTGCAGAA TAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA	WT WT 18 3 14 10 4 17 16/NS NS/+7 NS 31 73	AGATGAGAAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG 	EX +4
23 II # 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGGGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCATCA GTT-AAAAGAGACCATCC GTTAAAAGAGACCATCC GTTAAAAGAGACCATCC GTTAAAAGAGACCATCC	gGag1 ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA ATCAGGAAGCTGCAGAA 	WT WT 	AGATGAGAAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG AGTGACATAGCAGG AGTGACATAGCAGG AGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG CGGCATAGCAGG 	EX +4 WT WT WT -31 WT WT -2 -36 -21 NS/-14 -2 -45 +4/NS -13 -6
23 II # 1 2 3 4 5 6 6 6 6 6 7 7 8 9 9 10 11 12 13 14 15 16 17 14 15 16 17 17 10 10 10 10 10 10 10 10 10 10	GTTAAAAGAGACATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCAT GTTAAAAGAGACCAT GTTAAAAGAGACCATCA GTT-AAAAGAGACCATCC GTTAAAAGAGACCATCC GTTAAAAGAGACCATCC GTTAAAAGAGACCATTAGG	gGag1 ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA ATGAGGAAGCTGCAGAA 	WT WT -18 -3 -14 -10 -4 -17 -16/NS NS/+7 NS -31 -73 -15 NS/+2	AGATGAGAAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG GGAAGTGACATAGCAGG 	EX +4 WT WT WT -31 WT WT -2 -36 -2 NS/-14 -2 -45 +4/NS -13 -6 +2

III	gGag1			gGag3	
#	GTTAAAAGAGACCATCA ATG AGG AAGCTGCAGAA		AGATGAGAGAACCA	GGAAGTGACATAGCAGG	
1	GTTAAAAGAGACCATGAGGAAGCTGCAGAA	-4	AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG	WT
2	GTTAAAAGAGACCCA ATGAGGAAGCTGCAGAA	NS/-2	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
3	GTTAGC	-28/NS	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
4	GTTAAAAGAGGAAGCTGCAGAA	-11	AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG	WT
5	CA ATGAGGAAGCTGCAGAA	NS/-15	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
6	GTTAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
7	GTTAAAAGAGACCCA ATGAGGAAGCTGCAGAA	NS/-2	AGATGAGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
8	GTTAAAAGAGACCGAA	-18	AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG	WT
9	GAGGAA	-28/NS	AGATGTGAGAACCAAGG	GGAAGTGACATAGCAGG	WT
10		-36	AGATGAGAGAA CCA AGG	GGAAGTGACATAGCAGG	WT
11	GTTAAAAGAGACC ATGAGGAAGCTGCAGAA	-4	AGATGAGAGAA CCA AGG	-AAAGTGACATAGCAGG	-1/NS
12	GTTAAAAGAGACCCCTGAGGAAGCTGCAGAA	NS/-3	AGATGAGAGAA CCA AGG	TGACATAGCAGG	-5
13	GTTAAAAGGCA ATGAGGAAGCTGCAGAA	-6		A T G	-32
14	GTTAAAAGAGACCATCAGCTGCAGAA	-8	AGATGAGAGAA CCA AGG	GAGCATAGCAGG	-5/NS
15	GTTAAAAGAGACCAT	-20	AGATGAGAGAACCAAGG	AGGAAGTGACATAGCAGG	+2
16	GTTAAAAGAGAGGAAGCTGCAGAA	-10	AGATGAGAGAAAAC	CATAGCAGG	NS/-11
17		-38	AGATGAGAGAA CCA AGG	-GAAGTGACATAGCAGG	-1
18	GTTAAAAGGTGCAGAA	-18	AGATGAGAGAA CCA AGG	AAGTGACATAGCAGG	-2
19	G			GAGCAG	EX -27/N
20	GTTAAAAGAGACCATC-			GAGCAGG	EX -11/N
21	GTTAAAAGAGACCATCA			GGAAGTGACATAGCAGG	EX
22	GTTAAAAGAGACCATCA			GGAAGTGACATAGCAGG	EX
23	GTTAAAAGAGACCATCA			GGAAGTGACATAGCAGG	EX
24	GTTAAAAGAGACCATC-			GGAAGTGACATAGCAGG	EX -1

-AGCAGG

GGAAGTGACATAGCAGG

EX -14

EX

GTTAAAAGAGACTA---GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA

Figure S1. Sequence analysis of gGag1+gGag3 targeted HIV-rtTA DNA. SupT1 cells stably expressing Cas9+gGag1+gGag3 were infected with HIV-rtTA and cellular DNA was isolated at 7 days post-infection as described in Figure 3 (n=3; **a-c**). The gRNA target regions were amplified by PCR, the PCR products were TA cloned and 23-24 cloned fragments were sequenced. Sequences were aligned to the wild-type reference HIV-LAI sequence (highlighted in yellow; PAM sequence in bold) and 17 nucleotides on either side of the Cas9 cleavage sites (indicated by black arrows) are shown. Wild-type (WT; green), mutation (white) and excision (EX; blue) products are indicated (mutations in red; NS, nucleotide substitution; -n, n nucleotides deleted; +n, n nucleotides inserted).

I		gGag1	
-		googe	
#	GTTAAAAGAGACCATCA	ATGAGGAAGCTGCAGAA	
1	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
3	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
4	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
5	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
6	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
7	GTTAAAA	GAGGAAGCTGCAGAA	-12
8	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1
9	GTTA		-57
10	GTTAAAAGAGACCATCAC	CTGAGGAAGCTGCAGAA	+1/NS
11	GTTAAAAGAGACAGAAT	GG G AGG AAGCTGCAGAA	NS
12	GTTAAAAGA	GAAGCTGCAGAA	-13
13	GT	CTGCAGAA	-24
14	GTTAAAAGAGA	CTGCAGAA	-15
15	GTTCCAAGAGACCGTT-	AT AAGG AAGCTGCAGAA	NS/-1
16	GTTAAAAGAGACCTCAT	CGAAGCTGCAGAA	NS/-4
17	GTTAA	ATG AGG AAGCTGCAGAA	-12
18	GTTAAAAGAGACCATC-	GCAGAA	-11
19	GTTA	AGAG	-26/NS
20	GTTAAAAGAGACCTTCT	ATG AGG AAGCTGCAGAA	NS
21	GTTAAAA	GAGGAAGCTGCAGAA	-12
22	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-14
23	GTTAAAAG		-44

I	gGag1					
-						
#	GTTAAAAGAGACCATCA	ATGAGGAAGCTGCAGAA				
1	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
3	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
4	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
5	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
6	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
7	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
8	GTTAGAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
9	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
10	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
11	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
12	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
13	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
14	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
15	GTTAAAAGAGACCATCA	ATGAGGAAGCTGCAGAA	WT			
16	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
17	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
18			-67			
19		GCTGCAGAA	-26			
20	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1			
21	GTTAAAAGAGACCATCAACC	ATG AGG AAGCTGCAGAA	+3			
22	GTTAAAAGAG		-34			
23	GTTAAAAGAGACCATCC	ATG AGG AAGCTGCAGGA	NS			
24	GTTAAAAGAGACCATCA	GTGAGGAAGCTGCAGAA	NS			

b			
II		gGag1	
#	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	
1	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCGGAA	WT
2	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1
3	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
4	GTTAAAAGAGAC	CTGCAGAA	-14
5	GTTAAA	AGGAAGCTGCAGAA	-14
6	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGTA	-1
7	GTTAAAAGAGACCATCAGA	ATG AGG AAGCTGCAGAA	+2
8	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1
9	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
10	GTTAAAAGAGACCATCA	CCAGCTGCAGAA	-5/NS
11			-40
12	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1
13		GGAAGCTGCAGAA	-22
14	GTTAAAAGAGACA	ATG AGG AAGCTGCAGAA	-4
15	GTTAAAAGAGACA	ATG AGG AAGCTGCAGAA	-4
16	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1

II		gGag1	
#	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	
1	GTTAAAAGAGAGCCATCA	ATGAGGAAGCTGCAGAA	WT
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
3	GTTAAAAG	AGCTGCAGAA	-16
4	GTTAAAAGAGACCATC-	GAGCTGCAGAA	-7/NS
5	GTTAAAAGAGACCAT	ATG AGG AAGCTGCAGAA	-2
6	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
7	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1
8	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
9	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1
10	GTTAAAAGAGACCA	GCAGAA	-14
11	G		-42
12	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1
13	GTTAAAAG	AGCTGCAGAA	-16

III		gGag1		III	gGag1	
#	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA		#	GTTAAAAGAGACCATCA	
1	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT	1	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT	2	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
3	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT	3	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
4	GTTAAAAGAGACGAA	ATG AGG AAGCTGCAGAA	NS/-2	4	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
5	GTTAAAAGAGACCA	ATG AGG AAGCTGCAGAA	-3	5	GTTAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1
6	GCTAAAAGAGACCATCACCCT	ATG AGG AAGCTGCAGAA	+4	6	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
7	GTTAAAAGAGACCATCTT	ATG AGG AAGCTGCAGAA	NS/+1	7	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
8	GTTA	CAGAA	-19	8	GTTAAAAGAGACCATCAC ATGAGGAAGCTGCAGAA	+1
9	GTTAAAAGAGACA	ATG AGG AAGCTGCAGAA	-4	9	GTTAAAAGAGACCATC- ATGGGGAAGCTGCAGAA	-1/NS
10			-79	10	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
11	GTTAAAAGAGAC	CTGCAGAA	-14	11	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
12	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1	12	GTTAAAAGAGACC ATGAGGAAGCTGCAGAA	-4
13	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1	13	GTTAAAAGAGACCAT CTGAGGAAGCTGCAGAA	-2/NS
				14		-149
				15		-76
				16	GAAGTTGCAGAA	-111/NS

е

Figure S2. Sequence analysis of gGag1+gTatRev targeted proviral DNA. SupT1 cells stably expressing Cas9+gGag1+gTatRev were infected with HIV-LAI (n=3; a-c) or HIV-rtTA (n=3; d-f) and cellular DNA was isolated at 7 days post-infection as described in Figure 5. The gGag1 target region was amplified through PCR, PCR products were TA cloned and 13-24 clones were sequenced. Sequences are shown as described in Supplementary Figure S1.

а	I	HIV-LAI	
I		gGag1	
#	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	
1	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
3	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
4	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
5	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
6	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
7	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
8	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT
9	GTTAAAAGAGACCATCA	GAAGCTGCAGAA	-5
10		GAGGAAGCTGCAGAA	-22
11	GTTAAAAGAGAC	TTAAGCTGCAGAA	-9/NS
12	GTTAAA	AGGAAGCTGCAGAA	-14
13	GTTAAAAGAGACA	ATG AGG AAGCTGCAGAA	NS/-4
14			-80
15	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1
16	GTTA	GCAGAA	-24
17	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
18	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1
19	GTTAAAAGAGACCAT	AGGAAGCTGCAGAA	-5
20	GTTAAAAGAGACCATCAGA	ATG AGG AAGCTGCAGAA	+2
21	GTTAAAAGAGACC	GGAAGCTGCAGAA	-8
22	GTTAAAAGAGACC	ATG AGG AAGCTGCAGAA	-4
23	GTTAAAAGAGAC	ATG AGG AAGCTGCAGAA	-5

d	HIV-	rtTA				
I	gGag1					
#	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA				
1	GTTAGAAGAGACCATCA	ATG AGG GAGCTGCAGAA	WT			
2	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
3	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
4	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
5	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
6	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
7	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
8	GTTAAAAGAGACCATCA	ATG AGG AAGCTGCAGAA	WT			
9	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1			
10	GTTAAAAGAGAC	AAGCTGCAGAA	-11			
11	GTTAAAAGAGACCATC TTTTAAC	ATG AGG AAGCTGCAGAA	NS/+6			
12	GTTAAAAGAGACCATCAA	ATG AGG AAGCTGCAGAA	+1			
13	GTTAAAAGAGACCATCAGA	ATG AGG AAGCTGCAGAA	+2			
14	GTTAAAAGAGAC	CTGCAGAA	-14			
15	GTTAAAAGAG	TG GCTGCAGAA	-13/NS			
16	GTTAAA	AGCGGGCTGCAGAA	-14/NS			
17	GTTAAAAGAGACCT	CTGAGGAAGCTGCAGAA	NS/-3			
18	GTTAAAAGAGACCATC-	ATG AGG AAGCTGCAGAA	-1			
19	GTTAAAA	GAGGGAGCTGCAGAA	-12			
20	GTTAAAAGAGAC	ATG AGG AAGCTGCAGAA	-5			
21		GCCGCAGAA	-25/NS			

b			е		
II	gGag1		II	gGag1	
#	GTTAAAAGAGACCATCA ATG AGG AAGCTGCAGAA		#	GTTAAAAGAGACCATCA ATG AGG AAGCTGCAGAA	
1	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT	1	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
2	GTAAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT	2	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
3	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT	3	GTTAAAAGAGACCATCATTTGT ATGAGGAAGCTGCAGAA	+5
4	GTTAAAAGAGACCATCAGAGGAAGCTGCAGAA	-2	4	GTTAAGA ATGAGGAAGCTGCAGAA	-10
5	GTTAAAAGAGACCA ATGAGGAAGCTGCAGAA	-3	5		-129
6	GTTAAAAGAGACA ATGAGGAAGCTGCAGAA	NS/-4	6	GTTAAAAGAGACC ATG AGG AAGCTGCAGAA	-4
7	GTTAAAAGAGACC ATGAGGAAGCTGCAGAA	-4	7	GGGAAGCTGCAGAA	-20
8	GTTAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1	8	GTTAAAAGAGAGGAAGCTGCAGAA	-10
9	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1	9	GTTAAAAGAGACCATCATTTGT ATGAGGAAGCTGCAGAA	+5
10	GTTAAAAGAGACCATCAAGCAGAA	-10/NS	10	GTTAAA AGGAAGCTGCAGAA	-14
11	GTTAAAAGAGACCATCA	-17	11	GTTAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1
12	GTTAAAAGAGACCATCATAAGCTGCAGAA	-5/NS	12	GTTAAAAGAGACCATCAAA ATGAGGAAGCTGCAGAA	+2
13	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1	13		-79
14	GTTAAAAGGAAGCTGCAGAA	-14	14	GTTAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1
15	GTCAAAAGAGACCATC- ATGAGGAAGCTGCAGAA	-1	15	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1
16 16	GTTAAAAGAGACCATCAA ATG AGG AAGCTGCAGAA	+1	16 f	GTTAAAAGGAAGCTGCAGAA	-14
16	GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	+1		GTTAAAAGGAAGCTGCAGAA	-14
16 C	_		16 f		-14
16 C III	gGag1	AA	16 f	gGag1	-14
16 C III #	gGag1 GTTAAAAGAGACCATCA ATG AGG AAGCTGCAG	AA -1	16 f 111 #	gGag1 GTTAAAAGAGACCATCA ATG AGG AAGCTGCAGAA	
16 C III # 1	GTTAAAAGAGACCATCA GTTAAAAGAGACCATC- ATG AGG AAGCTGCAG	AA -1 AA -9/NS	16 f 111 #	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT
16 C III # 1 2	gGag1 GTTAAAAGAGACCATCA GTTAAAAGAGACCATC- GTTAAAAGAGACCATC- GTTAAAAGAGAC TCAAGCTGCAG	AA -1 AA -9/NS AA NS/+4	16 f 111 <u>#</u> 1 2	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT WT
16 C III 1 2 3	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATC- GTTAAAAGAGACCATC- GTTAAAAGAGACCATC GTTAAAAGAGACCATC CTG AGG AAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4	16 f 111 1 2 3	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT WT WT
16 C III # 1 2 3 4	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- TCAAGCTGCAG GTTAAAAGAGACCATCCCATT CTGAGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4 35	16 f 111 <u></u> 1 2 3 4	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA	WT WT WT +1
16 C III 1 2 3 4 5	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGAAGCTGCAG GTTAAAAGAGACCATCC- TCAAGCTGCAG GTTAAAAGAGACCATCCCATT TTGAGAAGCTGCAG GTTAAAAGAGACCATCC- AGGAAGCTGCAG GTTAAAAGAGACCATCC- AGGAAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4 35 AA -8	16 f 111 2 3 4 5	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA	WT WT +1 +2
16 C III 1 2 3 4 5 6	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- TCAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG G GTTAAAAGA ATGAGGAAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4 35 AA -8 AA +2	16 f 111 <u></u> 1 2 3 4 5 6	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCACA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCACACA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCACACACAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCACACAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCACACAA ATGAGGAAGCTGCAGAA GTTAAACGACCATCACACAA ATGAGGAAGCTGCAGAA GTTAAACGACCATCACACAA ATGAGAAGCTGCAGAA GTTAAACGACCATCACACAA ATGAGAAGCTGCAGAA GTTAAACAGACCATCACACAA ATGAGAAGCTGCAGAA	WT WT +1 +2 -14
16 C III # 1 2 3 4 5 6 7	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- TCAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAAGACCATC- AGGAAGCTGCAG GTTAAAAGAAGACCATCAGG ATGAGGAAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4 35 AA -8 AA +2 AA -4	16 f <u>111</u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u> <u></u>	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAA-GAGACCATCAA ATGAGAAGCTGCAGAA GTTAAA	WT WT +1 +2 -14 +1
16 C III # 1 2 3 4 5 6 7 8	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC ATGAGGAAGCTGCAG GTTAAAAGAGACCATCCCCAT CTGAGAAGCTGCAG GTTAAAAGAGACCATCC AGGAAGCTGCAG GTTAAAAGAGACCATC AGGAAGCTGCAG GTTAAAAGAGACCATC AGGAAGCTGCAG GTTAAAAGAGACCATC AGGAAGCTGCAG GTTAAAAGAGACCATCAGG ATGAGGAGCTGCAG GTTAAAAGAGCCCATCAGG ATGAGGAGCTGCAG GTTAAAAGAGACCCATCAGG ATGAGGAGCTGCAG GTTAAAAGAGACCCATCA ATGAGGAAGCTGCAG	AA -1 AA -9/NS AA NS/+4 AA -4 -35 AA -8 AA +2 AA -4 -A -19	16 f <u>111</u> <u>#</u> 1 2 3 4 5 6 7 8	gGag1 GTTAAAAGAGACCATCA ATGAGGAASCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAASCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAGA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAACCTGCAGAA GTTAAAAGAGACCATCAA AGGAACCTGCAGAA GTTAAAAGAGACCATCAA AGGAACCTGCAGAA GTTAAAAGAGACCATCAA AGGAACCTGCAGAA	WT WT +1 +2 -14 +1 -7/NS
16 C III # 1 2 3 4 5 6 7 8 9	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATCCCATT CTAGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG	AA -1 AA -9/NS AA -9/NS AA -4 -35 AA -8 AA -4 -35 AA -8 AA -4 AA -12 AA -22	16 f <u>111</u> <u>#</u> 1 2 3 4 5 6 7 7 8 9	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGAAGACATCAA ATGAGAAGCTGCAGAA	WT WT +1 +2 14 +1 -7/NS -14
16 C III # 1 2 3 4 5 6 7 8 9 10	GTTAAAAGAGACCATCA GTTAAAAGAGACCATCA GTTAAAAGAGACCATC- ATG AGG AAGCTGCAG GTTAAAAGAGACCATC- GTTAAAAGAGACCATCCCATT GTAAAAGAGACCATCC GTTAAAAGAGACCATCA GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTAAAAGAGACCATCAG GTTACGACG	AA -1 AA -9/NS AA NS/+4 AA -4 -35 AA -8 AA -4 AA -8 AA -4 AA -22 AA -8/NS	16 f 11 2 3 4 5 6 6 7 8 9 10	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAASCTGCCAGAA GTTAAAAGAGACCATCAA ATGAGAASCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAASCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCCAGAA	WT WT +1 +2 -14 +1 -7/NS -14 +1
16 C III # 1 2 3 4 5 6 7 8 9 10 11	GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATCAGGAAGCTGCAG GTTAAAAGAGACCATCAGGAAGCTGCAG GTTAAAAGAGACCATCAGGAAGCTGCAG GTTAAAAGAGACCATCAG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCAG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCAG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCAG ATGAGGAAGCTGCAG GTTAAAAGAGACCA ATGAGGAAGCTGCAG GTTAAAAGAGACCA	AA -1 AA -9/NS AA NS/+4 AA -4 -35 AA -8 AA -4 AA -19 AA -19 AA -22 AA -8/NS AA -23	16 f iii ii i i i i i	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA	WT WT +1 +2 -14 +1 -7/NS -14 +1 -20
16 C III # 1 2 3 4 5 6 7 8 9 10 11 12	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- ATGAGGAAGCTGCAG GTTAAAAGAGACCATC- TCAAGCTGCAG GTTAAAAGAGACCATCCCATT CTGAGGAAGCTGCAG GTTAAAAGAGACCATC- AGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCATCACG ATGAGGAAGCTGCAG GTTAAAAGAGACCA- ATGAGGAAGCTGCAG GTTAAAAGAGACCA- ATGAGGAAGCTGCAG GTTA- CGACAG GTTA- CGACAG GTT-	AA -1 AA -9/NS AA NS/+4 AA -4 -35 AA -8 AA -4 AA -19 AA -19 AA -22 AA -8/NS AA -23	16 f iii * 1 2 3 4 5 6 7 7 8 9 10 11 12	gGag1 GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGGAAGCTGCAGAA GTTAAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAGAGACCATCAA ATGAGAAGCTGCAGAA GTTAAAAGAGACCATCAA </td <td>WT WT +1 +2 -14 +1 -7/NS -14 +1 -20 -15</td>	WT WT +1 +2 -14 +1 -7/NS -14 +1 -20 -15

Figure S3. Sequence analysis of gGag1+gEnv2 targeted proviral DNA. SupT1 cells stably expressing Cas9+gGag1+gEnv2 were infected with HIV-LAI (n=3; A-C) or HIV-rtTA (n=3; D-F) and cellular DNA was isolated at 7 days post-infection (as described in Figure 4). The gGag1 target region was amplified through PCR, PCR products were TA cloned and 15-23 clones were sequenced. Sequences are shown as described in Supplementary Figure S1.

Table S1. Primers and probe used for PCR and qPCR analysis

name	sequence
gGag1-forward (a)	ACCTAGAACTTTAAATGCATGG
gGag3-reverse (b)	CGGTCTACATAGTCTCTAAAGG
gTatRev-forward (c)	GAGCCAGTAGATCCTAGAC
gTatRev-reverse (d)	CTACTACTAATGCTGCTATTGC
gEnv2-forward (c)	GCAAAGAGAAGAGTGGTG
gEnv2-reverse (d)	AGCCAGGATTCTTGCC
gGag1-probe	CCCATGTTTTCAGCATTATCAGAAGGAGCC

Table S2. Selected gRNAs targeting HIV-1

Name	Target Sequence + PAM (underlined)	Orientation
gGag1	GTTAAAAGAGACCATCAATG <u>AGG</u>	Sense
gGag3	<u>CCA</u> AGGGGAAGTGACATAGCAGG	Antisense
gTatRev	<u>CCT</u> ATGGCAGGAAGAAGCGGAGA	Antisense
gEnv2	GGAGCAGCAGGAAGCACTAT <u>GGG</u>	Sense