SUPPLEMENTARY METHODS

Ethics Statement

The study was approved by the Institutional Review Board at the University of California. All adult participants (age ≥ 18 years) provided written informed consent. No children were included in this study.

Study Population and Clinical Characteristics

We recruited study participants from the Owen Clinic at the University of California, San Diego. Criteria for inclusion to our study were: documented HIV-infection, at least 18 years of age, $CD4^+$ T cell counts >200 cells/mm^3, no recent changes to ART and sustained undetectable viral loads for at least 6 months. Enrolled participants provided informed consent and were given one 0.5mL dose of a standard clinical influenza vaccine (Fluarix®, GSK) after giving blood for a baseline comparison. Participants returned to the clinic on days 2, 4, 7, 14 and 28 following vaccination to provide blood samples. PBMC and serum were isolated from blood samples within 24 hours using Lymphoprep® (StemCell Technologies) per manufacturer's protocol, and were then frozen and stored at -150°C prior to analysis.

Digital Droplet PCR (ddPCR)

Copies of HIV gag DNA and RNA were evaluated by the extensively validated ddPCR assay, as previously described (BioRad) [15]. DNA and RNA were extracted from PBMC using Allprep or DNeasy and RNeasy spin columns (Qiagen). Total concentrations were measured with on a Nanodrop 1000 spectrophotometer (Thermo Scientific). DNA was

fragmented with a *BanII* restriction enzyme prior to ddPCR, and cDNA was prepared from RNA samples with Ambion Retroscript® (ThermoFisher Scientific). Next, droplets were generated from the digested DNA and cDNA samples, and transferred to a sealed PCR plate. Thermocycling was then used to amplify a region from the p24 region of HIV gag and total poly-adenylated mRNA using the primers listed in **Supplemental Digital Content Table S1**. Copy numbers were read in the ddPCR machine (BioRad) and normalized to total cell input, as in [15].

Flow Cytometry

PBMCs from study participants were washed in flow cytometry buffer consisting of phosphate buffered saline (Corning) with 2% fetal bovine serum (GemCell® cat# 100-500) and 0.1% sodium azide (Sigma-Aldrich). Cells were stained for the extracellular markers CD3 (clone: SK7), CD4 (SK3), HLA-DR (G46-6) and CD38 (HB7) (BD Biosciences). Analysis was performed on a BD Accuri CSampler™ (BD Biosciences, Cat# 653124), using the manufacturer's software. Tables and figures were generated and assembled using Microsoft Office®.

IgG titer

Influenza A-specific IgG was measured from serum samples by ELISA (AbCam, Cat# 108745), per manufacturer's instructions.

Sequencing and Analysis

DNA/RNA extraction and amplification. DNA and RNA were isolated from PBMCs using Qiagen DNA and RNA extraction kits (cat#s 51106 and 74104), per manufacturers

protocol, and cDNA was reverse transcribed from RNA samples using Ambion Retroscript® (ThermoFisher Scientific). HIV sequences were amplified from DNA or cDNA by nested PCR, using the primers listed in **Supplemental Digital Content Table S1.** PCR products were then cleaned prior to library preparation with Qiagen PCR purification kits (cat# 28106).

Library preparation. DNA Library Preparations were created per manufacturer specifications using the Illumina Nextera XT Index kit (FC-131-1001). Briefly, a minimum of 100ng of DNA template was combined with 25ul of 2x KAPA HiFi HotStart Ready Mix (KAPA Biosystems KK2601) and 5ul of each of the two Nextera index primers provided with the Nextera XT Index Kit V2 Set A (Illumina Cat. FC-131-2001). The total volume was adjusted to 50μl with the addition of PCR grade water. The preparations were then thermocycled, and >100ng DNA was purified using AMpure XP beads (ratio of 1.2μL beads:1μL DNA, Beckman A63881) and eluted with 10mM Tris pH 8.5.

Sequencing. The Illumina MiSeq instrument and MiSeq Reagent Kit V3 600-cycle Paired End sequencing kits (MS-102-2003/MS-102-3003) were used to sequence the DNA libraries. We performed 30 paired-end sequencing runs (5 time points, 3 individuals for HIV DNA and RNA). The median number of reads per amplicon was 320,452 (IQR: 198,053-1,494,661). All read files were uploaded to the NCBI sequence read archive, acquisition number XXX, and can also be obtained by contacting achaillon@ucsd.edu.

Read mapping and filtering. The reads were analyzed using a custom pipeline adapted from the publicly available pipeline https://github.com/iosonofabio/hivwholeseq [16]. Briefly, (i) reads were first mapped onto the HIV-1 reference HxB2 (ii) mapped reads

were classified into partial *gag*, *pol* and *env* regions used for RT-PCR (ambiguous reads were discarded), and trimmed for PHRED quality above or equal to 30 (iii) a consensus sequence was computed for each 3 regions in each sample from a subset of the reads, using a chain of overlapping local multiple sequence alignments; (iv) reads were re-mapped against their own consensus. (v) reads were trimmed for mapping errors at the edges (small indels) (vi) filtered reads were mapped a third time against a patient-specific consensus sequence from the initial time point (day 0); (vii) reads were re-filtered and checked again for cross-contamination. For each sample, we computed the mean of all pairwise Tamura-Nei 93 (TN93) distances between reads with at least 100 overlapping base pairs to quantify nucleotide diversity [17].

Phylogenetic analysis. HIV DNA and RNA haplotypes above a minimal frequency threshold of 0.01 were extracted from reads covering the gag, pol and env regions and were used to construct maximum likelihood phylogenies using FastTree [18, 19].

Viral Compartmentalization. Viral compartmentalization was first assessed by the Fst approach defined as $F_{ST} = 1 - \frac{\pi_I}{\pi_D}$, where π_I is the estimate of mean pairwise *intra-compartment* genetic distance (TN93) [20], and π_D is its *inter-compartment* counterpart [21]. Both quantities were computed by comparing all reads from DNA and RNA samples that had a minimum of 150 aligned nucleotide positions. Subsequently, to guard against inference of compartmentalization by skewing of allelic frequencies due to PCR amplification and other biases, we recomputed FST by discarding copy number counts for read clusters (i.e. each cluster was counted as having only one sequence), i.e. all haplotypes are assigned a relative weight of 1. Statistical significance of both tests was derived via 1,000 population-structure randomization/permutation test. We also

performed a tree-based Slatkin-Maddison (SM) test for compartmentalization [22]. HIV DNA and RNA sequenced populations were considered compartmentalized if all tests were congruent.

Table S1. Primers for ddPCR and Nested PCR

	Direction	Sequence	Location	Primer ID
ddPCR (gag)	Forward	AGTTGGAGGACATCAAGCAGCCATGCAAAT		1357 SK462 Gag-F
	Reverse	TGCTATGTCAGTTCCCCTTGGTTCTCT		1358 SK431 Gag-R
	Probe	AGACCATCAATGAGGAAGCTGCAGAATGGGAT		1359 SK102 Gag-Prb
ddPCR (PolyA)	Forward	CAGATGCTGCATATAAGCAGCTG		1369 VQA-F
	Reverse	TTTTTTTTTTTTTTTTTTTTTTTTGAAGCA		1368 VQA-R
	Probe	CCTGTACTGGGTCTCTCTGG		1370 VQA-Prb
Nested gag (1st Round)	Forward	TATCAGAAGGAGCCACCC		CI-p24gag1312_Fout
	Reverse	CTCCCTGACATGCTGTCATCA		CI-p24gag1846_Bout
Nested gag (2nd Round)	Forward	GGACATCAAGCAGCCATGCAAATG		CI-p24gag1366_Fin
	Reverse	TACATTCTTACTATTTTATT		CI-p24gag1619_Bin2
Nested pol (1st Round)	Forward	GGAAGAAATCTGTTGACTCAGATTGG		CI-Pol1
	Reverse	ACCCATCCAAAGGAATGGAGGTTCTTTC		3RT
Nested pol (2nd Round)	Forward	AAATCCATACAATACTCCAGTATTTGC		5RT
	Reverse	CATCCATGTATTGATAGATAACTATGTCTG		CI-185RT_R
Nested env (1st Round)	Forward	CAAAGGTATCCTTTGAGCCAAT		V3_Fout
	Reverse	ATTACAGTAGAAAAATTCCCCT		V3_Bout
Nested env (2nd Round)	Forward	GAACAGGACCAGGATCCAATGTCAGCACAGTACAAT		V3_Fin
	Reverse	GCGTTAAAGCTTCTGGGTCCCCTCCTGAG		V3_Bin
MiSeq	Forward	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG		
MiSeq	Reverse	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG		

Table S2. Demographics of study participants

PID	Age	Gender	Years since HIV diagnosis	CD4 count (cells/mm^3)	VL (copies/mL)	Race	ART Regimen
K1	68	F	30	856	<20	White	ABC/3TC/RAL
K2	64	M	19	277	<20	White	TDF/FTC/RAL/ DRV/r
K3	63	M	31	551	<20	White	ABC/3TC/RAL
K4	58	M	31	327	<20	White	TDF/FTC/ ETR/DTG
K5	59	M	17	721	<20	White	ABC/3TC/NVP
K6	31	M	4	661	<20	African American	TDF/FTC/RAL
K7	55	M	21	613	<20	African American	TDF/FTC/RAL/ DRV/r
Mean (all)	56.9	86% Male	21.9	572	<20	N/A	N/A
Mean (sequenced)	61.7	100% Male	27.0	385	<20	N/A	N/A

Supplementary Digital Content Figure S1. HIV transcription levels and immune activation of study participants following vaccination. A) Copies of cell-associated HIV gag (blue dotted lines) per million cells were measured by ddPCR. Averages from two aliquots are shown. B) Fold changes in *gag* copy number, relative to D0 C) Percentages of HLA-DR and CD38 dual-positive at the indicated days following vaccination. D) Influenza A-specific IgG was measured in serum samples by ELISA, in triplicate. Changes in absorbance, relative to D0, are shown for days 7, 14 and 28.

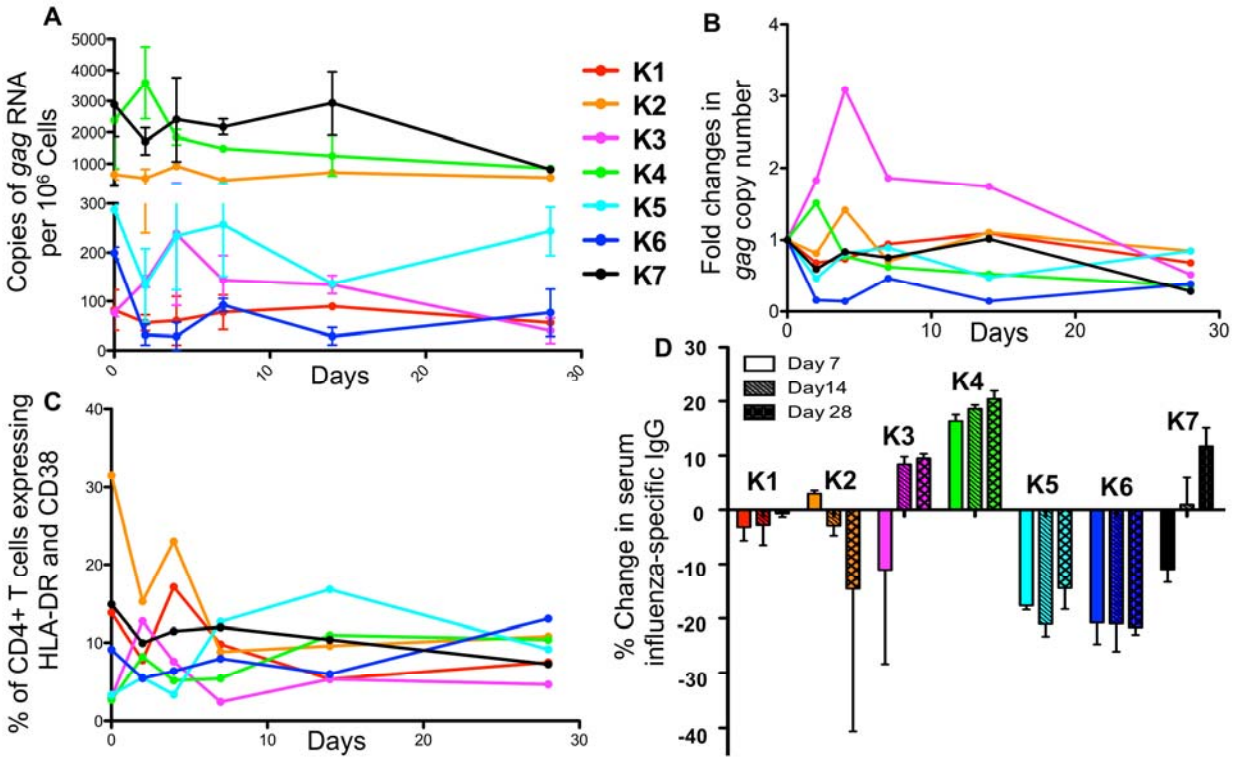
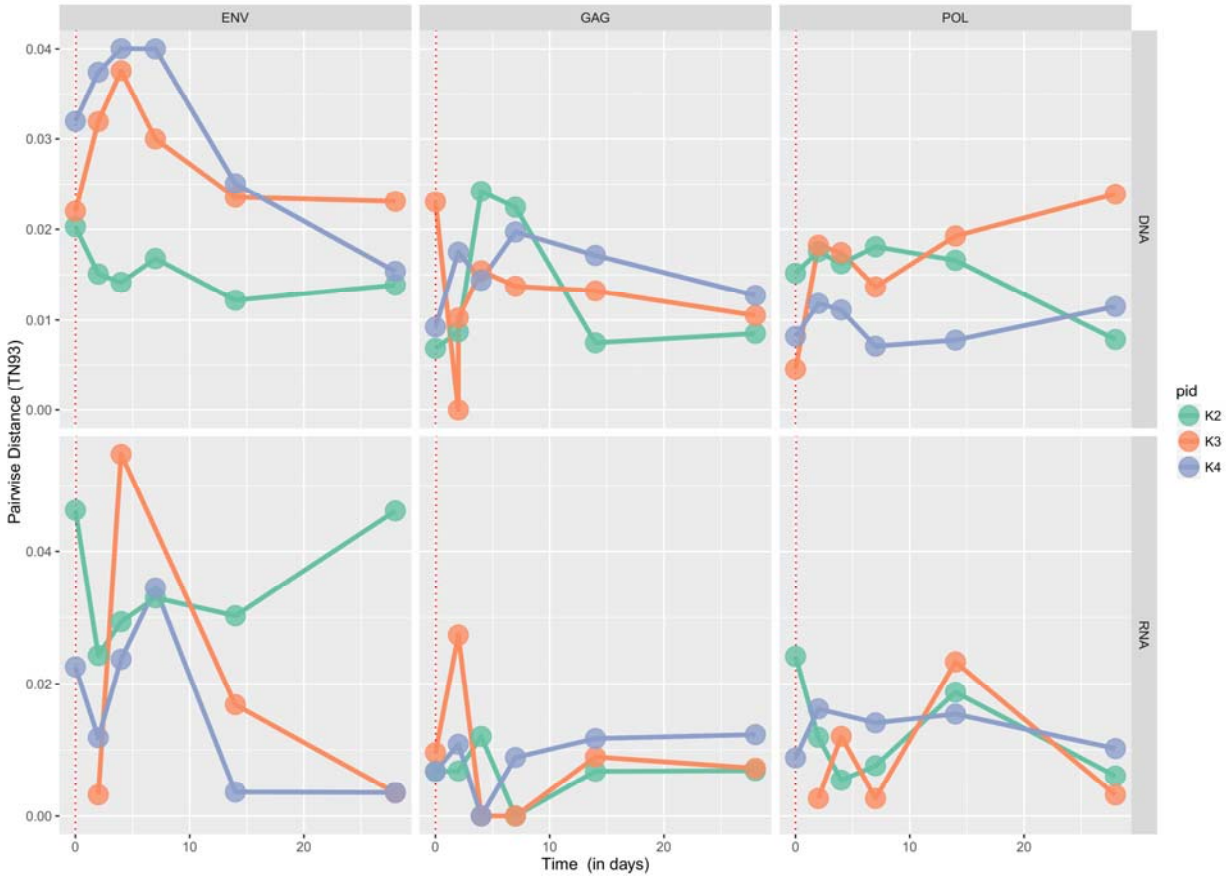


Table 1. Compartmentalization Analyses

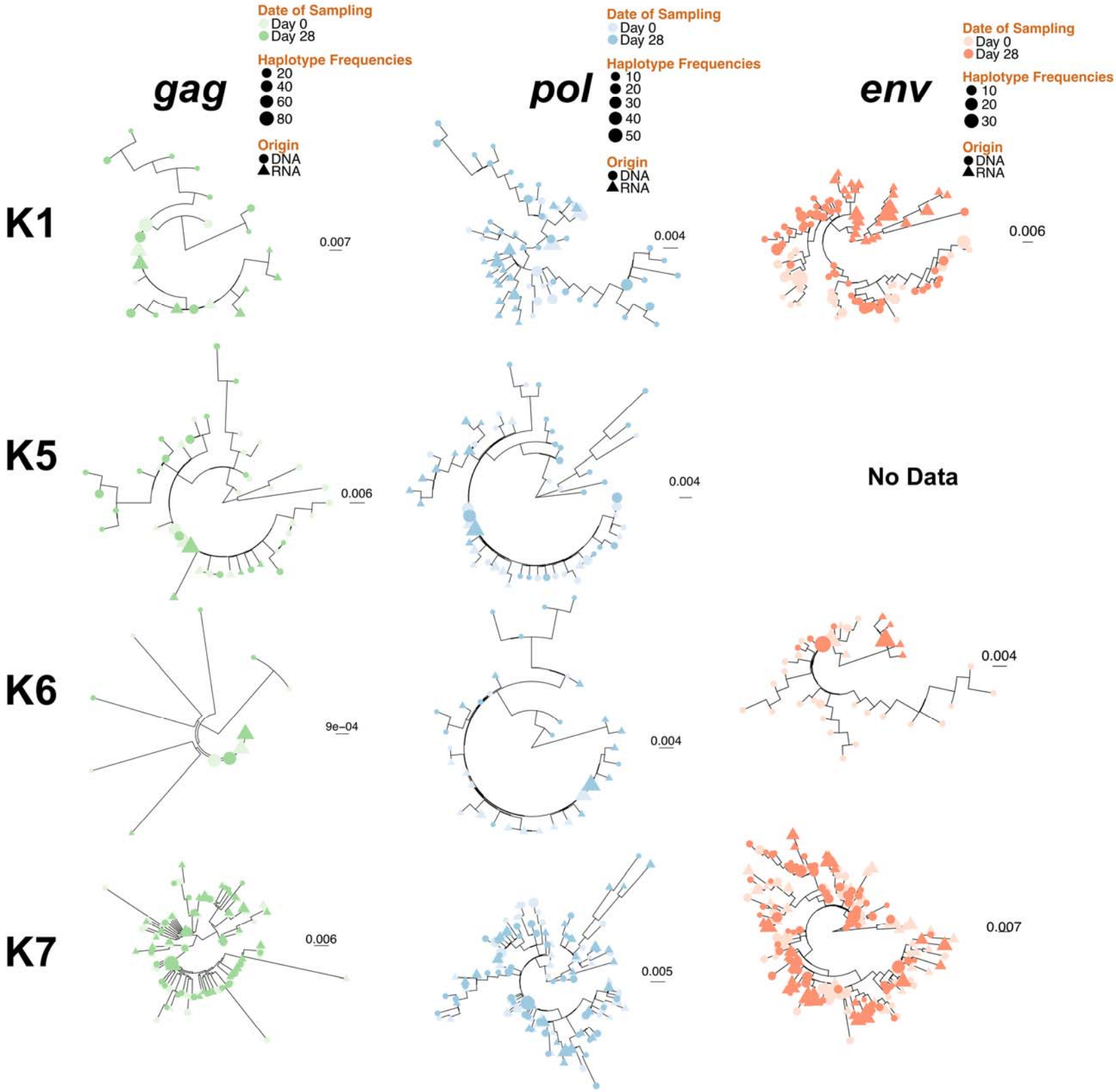
PID	Region	Fst	Slatkin-Maddison	p-value
K1	env	0.33848	0.20372	**<0.001**
	gag	-0.04858	-0.02371	<0.61
	pol	-0.01132	-0.00563	<0.58
K2	env	0.10896	0.05762	<0.07
	gag	-0.08600	-0.04123	<0.97
	pol	0.09275	0.04863	<0.06
K3	env	-0.07936	-0.03817	<0.93
	gag	0.15625	0.08475	<0.08
	pol	0.01368	0.00689	<0.39
K4	env	-0.01733	-0.00859	<0.80
	gag	-0.13949	-0.06520	<1.0
	pol	0.12644	0.06748	<0.07
K5	gag	-0.14788	-0.06885	<0.83
	pol	-0.10653	-0.05057	<0.89
K6	env	0.40914	0.25718	**<0.001**
	gag	-0.02848	-0.01404	<0.56
	pol	0.05127	0.02631	<0.19
K7	env	0.04487	0.02295	**<0.001**
	gag	-0.00001	-0.00001	<0.40
	pol	0.01566	0.00789	**<0.04**

Supplementary Digital Content Figure S2. Sequence diversity of HIV DNA and RNA at each time point for participants K2, K3 and K4. TN93 pairwise distances are shown at each time point for K2 (green), K3 (orange) and K4 (blue).



Supplemental Digital Content Figure S3. Maximum likelihood phylogenetic reconstruction of baseline and day 28 RNA and DNA HIV-1 populations in "non-kick" participants

HIV DNA and RNA haplotypes above a minimal frequency threshold of 0.01 were extracted from reads covering the *gag* (panel A), *pol* (panel B) and *env* (panel C) regions for individual K1, K5 K6 and K7 and were used to construct maximum likelihood phylogenies using FastTree (Price et al., 2009). HIV DNA and RNA haplotypes are depicted in circles and triangles respectively. Time point in days (0 and 28 days from vaccination) is indicated by color. Scale bars are in substitutions/site.


gag
pol
env
Date of Sampling
Day 0
Day 28
Haplotype Frequencies
20
40
60
80
10
20
30
40
50
10
20
30
Origin
DNA
RNA
K1
K5
K6
K7
0.007
0.004
0.006
0.006
0.004
No Data
0.004
9e−04
0.004
0.006
0.005
0.007

Supplemental Digital Content Figure S4. Model depicting standard of care vaccines as a tool to purge the reservoir

Purple dots: Vaccine components; Green dots: pro-inflammatory cytokines; Red squares: HIV. (1) Standard of care vaccines for people living with HIV, such as Influenza, Hepatitis, Pneumonococcus, Papillomavirus, Diphtheria, Pertussis, Tetanus. (2) Upon vaccination, antigen presenting cells are activated by adjuvant to process and present vaccine-specific antigens to T cells. (3) T cell receptor binding, co-stimulation, and cytokines activate antigen-specific T cells. (3) Activated T cells produce cytokines, which can promote (4) bystander T cells to become activated and (5) B cell maturation and vaccine-specific antibody production. (6) Bystander T cell activation could induce the expression of HIV from latent reservoirs.