Supplemental Material Table S1. Reproducibility of RT-PCR amplification of the C2V3 region of the HIV-1 *env* gene (480 nt)

(A) Viral load 1,001 – 5,000 copies of viral RNA/ml

Replicate 2	n = 5	+	-
	+	5	0
	-	0	0
		Replicate 1	

(B) Viral load 5,001 – 10,000 copies of viral RNA/ml

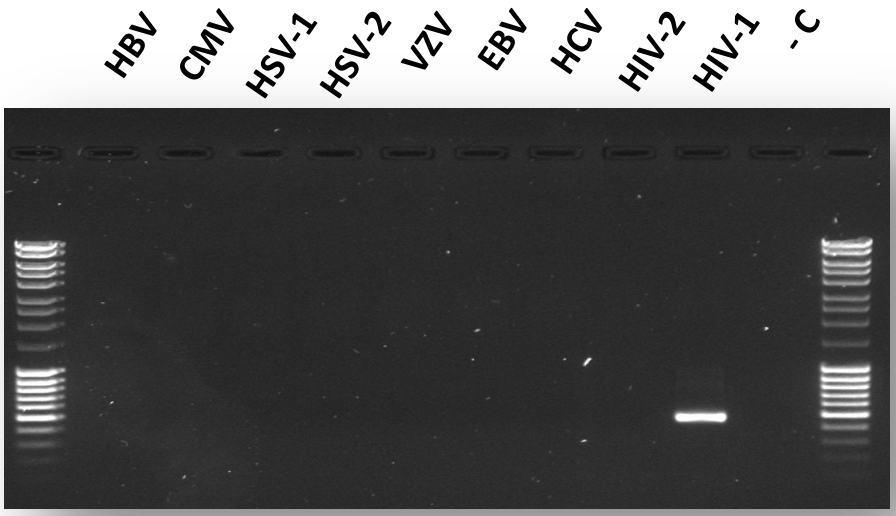
Replicate 2	n = 5	+	-
	+	5	0
	-	0	0
		Replicate 1	

(C) Viral load >10,000 copies of viral RNA/ml

Replicate 2	n = 5	+	-
	+	5	0
	-	0	0
		Replicate 1	

RT-PCR amplification reproducibility involved the analysis of 15 plasma samples from HIV-infected individuals with different viral loads: (A) 1,001 – 5,000 copies of viral RNA/ml (n = 5), (B) 5,001 – 10,000 copies of viral RNA/ml (n = 5), and (C) >10,000 copies of viral RNA/ml (n = 5). RT-PCR amplification of the patient-derived C2V3 fragment was performed by two different operators, using different lots of critical reagents over a seven-day period. A perfect (100%) RT-PCR amplification reproducibility was observed using plasma samples with viral loads >1,000 copies/ml.

Supplemental Material Figure S1. RT-PCR amplification cross-reactivity of the C2V3 region of the HIV-1 *env* gene (480 nt)



RT-PCR amplification specificity involved the analysis of nucleic material from 9 DNA or RNA viruses (i.e., HBV, HCV, HIV-2, BKV, EBV, CMV, HSV- 1, HSV-2, and VZV). No cross-reactivity was observed with any of these viruses as all RT-PCR reactions failed to generate any detectable amplicons.

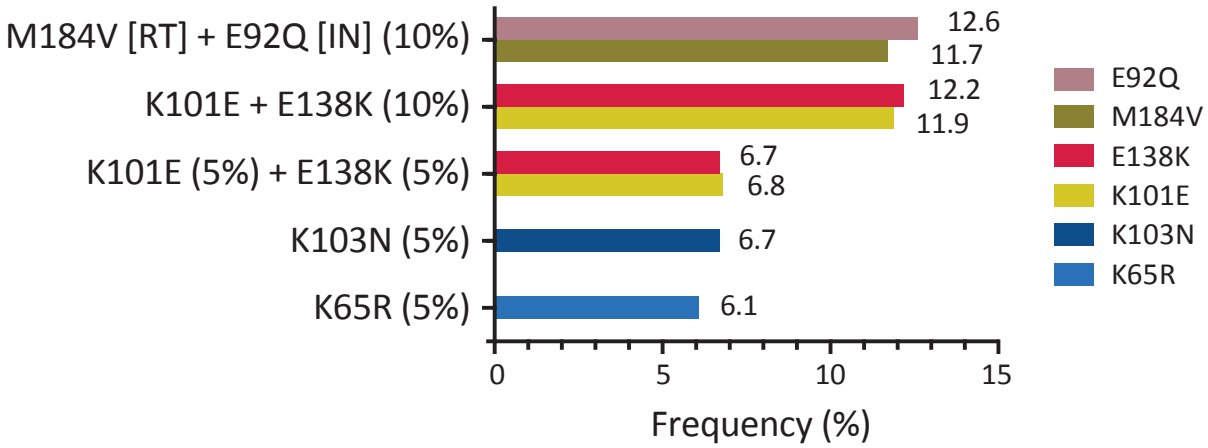
Supplemental Material Table S2. HIV-1 isolates used to RT-PCR amplify the C2V3 region of the HIV-1 *env* gene (480 nt)

Virus ID [a]	*env* subtype [b]
92RW009	A
93RW020	A
V115	A
V120	A
92UG029	A
92BR014	B
92TH593	B
92US714	B
92US727	B
92US076	B
C18	C
C20	C
C21	C
C22	C
92BR025	C
V89	D
V122	D
V126	D
94UG108	D
92UG038	D
93UG065	D
93BR029	F
VI820	F
V164	F
CA16	F
CA20	F
93BR020	F
RU570	G
RU132	G
CMU02	AE
CMU06	AE
92TH021	AE
93BR019	BF

[a] All viruses were obtained from the AIDS Research and Reference Reagent Program (ARRRP), Division of AIDS, NIAID, NIH or as a gift from Dr. Eric J. Arts' laboratory at Case Western Reserve University (CWRU), Cleveland, OH as described in Materials and Methods.

[b] *env* subtype determined by in-house population (Sanger) sequencing to corroborate published (NIH ARRRP) results.

Supplemental Material Figure S2. Analytical sensitivity determined in mixtures of plasmid DNA



Five plasmid mixtures containing amino acid substitutions in the HIV-1 *pol* gene associated with drug resistance, i.e., K65R (5%) + wild type (95%), K103N (5%) + wild type (95%), K101E (5%) + E138K (5%) + wild type (90%), K101E + E138K (10%) + wild type (90%), and M184V (RT) + E92Q (IN) (10%) + wild type (90%) were kindly provided by Gilead Sciences, Inc. (Foster City, CA). Plasmid DNA was used as template for the amplification of the *gag*-p2/NCp7/p1/p6/*pol*-PR/RT/IN fragment and deep sequenced as described in Materials and Methods. All amino acid changes were detected and quantified at approximately the right proportion.

Supplemental Material Figure S3. Analytical sensitivity determined mixing DNA from two plasmids containing the HIV-1 genome from a multidrug resistant (08-180) and a wild-type (NL4-3) virus.

A *gag*-p2/NCp7/p1/p6/*pol*-PR/RT/IN PCR product was obtained from an antiretroviral-experienced patient (08-180) and used to construct p2-INT recombinant viruses (Weber *et al* 2013 PLoS ONE e65631). The plasmid generated by the yeast cloning method contains a representation of the *in vivo* HIV-1 quasispecies (Weber et al 2011 Antimicrob. Agents Chemother. 55:3729). This plasmid preparation contained the *pol* gene from the patient and the *env* gene from the CXCR4-tropic HIV-1_{NL4-3} strain, i.e., **08-180 *pol* / NL43- (X4) *env*.** Plasmid **NL4-3 *pol* / YU2 (R5) *env*** contains the genome of the wild-type HIV-1_{NL4-3} virus carrying the *env* gene from the R5 HIV-1_{YU2} virus (Weber et al 2007 Antivir. Ther. 12:S155). A series of plasmid mixtures were created by mixing 0.1%, 1%, 2%, 3%, 5% and 10% of the 08-180 *pol* / NL43- (X4) *env* plasmid with the corresponding amount of the NL4-3 *pol* / YU2 (R5) *env* plasmid at a final concentration of 0.1 ng/ml. DNA from the entire plasmid mixtures, together with the two individual plasmids as controls (100%), was purified and deep sequenced as described in Materials and Methods. Frequency of amino acids detected in positions associated with drug resistance in the protease, reverse transcriptase, and integrase coding regions, as well as the proportion of X4 sequences, were determined. Amino acid changes detected at ≥1% of the population (threshold calculated based on the intrinsic error rate of the assay) in each mixture of plasmid DNA are indicated in red. Most drug resistance mutations present at ~100% in the original **08-180 *pol* / NL43- (X4) *env*** plasmid were detected when the plasmid was diluted to approximately 1% of the population in the plasmid mixture. Two amino acid changes in the protease gene were detected as minority members of the quasispecies in the original **08-180 *pol* / NL43- (X4) *env*** plasmid (L33F at 21.9% and F53Y 1.7%). As expected, detection of the L33F failed after the second dilution while F53Y was not detected when the original plasmid was diluted to 10% of the population.

Protease - Mutation Frequency

% Plasmid in the mixture		L	K	V	L	M	K	M	I	F	I	I	L	A	G	V	I	L	L
NL4-3 *pol*	08-180 *pol*	10I	20T	32I	33F	36L	43T	46I	47V	53Y	54L	62V	63P	71V	73S	82I	84V	89V	90M
YU2 (R5) *env*	NL4-3 (X4) *env*																		
100	0	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
99.9	0.1	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
99	1	I (1.38%)	T (1.13%)	I (1.65%)	0%	L (1.79%)	T (1.59%)	I (1.66%)	V (1.6%)	0%	L (1.5%)	V (1.74%)	P (1.62%)	V (1.74%)	S (1.61%)	I (1.54%)	V (1.64%)	V (1.54%)	M (1.59%)
98	2	I (1.26%)	T (1.21%)	I (1.67%)	0%	L (1.7%)	T (1.71%)	I (1.65%)	V (1.56%)	0%	L (1.41%)	V (1.65%)	P (1.44%)	V (1.52%)	S (1.45%)	I (1.51%)	V (1.43%)	V (1.28%)	M (1.41%)
97	3	I (2.43%)	T (2.37%)	I (3.24%)	0%	L (3.33%)	T (3.14%)	I (2.9%)	V (2.76%)	0%	L (2.57%)	V (3.46%)	P (3.5%)	V (3.54%)	S (3.47%)	I (3.58%)	V (3.66%)	V (3.71%)	M (3.79%)
95	5	I (11.99%)	T (12.81%)	I (16.73%)	F (3.65%)	L (16.84%)	T (15.69%)	I (15.46%)	V (14.24%)	0%	L (14.72%)	V (15.9%)	P (15.49%)	V (16.8%)	S (17.17%)	I (18.75%)	V (18.78%)	V (20.18%)	M (20.28%)
90	10	I (10.78%)	T (11.49%)	I (15.1%)	F (3.49%)	L (14.98%)	T (14.15%)	I (13.86%)	V (12.7%)	0%	L (13.18%)	V (14.62%)	P (14.14%)	V (15.27%)	S (15.56%)	I (16.9%)	V (16.9%)	V (17.27%)	M (17.38%)
0	100	I (99.82%)	T (99.43%)	I (99.82%)	F (21.91%)	L (99.97%)	T (98.19%)	I (99.49%)	V (99.76%)	Y (1.73%)	L (99.91%)	V (99.83%)	P (98.88%)	V (99.72%)	S (99.9%)	I (99.66%)	V (99.71%)	V (99.9%)	M (99.56%)

Reverse Transcriptase - Mutation Frequency

% Plasmid in the mixture		M	D	V	F	A	M	L	T	K
NL4-3 *pol*	08-180 *pol*	41L	67N	75M	77L	98S	184V	210F	215Y	219N
YU2 (R5) *env*	NL4-3 (X4) *env*									
100	0	0%	0%	0%	0%	0%	0%	0%	0%	0%
99.9	0.1	0%	0%	0%	0%	0%	0%	0%	0%	0%
99	1	L (1.1%)	N (1.59%)	M (1.38%)	L (1.49%)	S (1.28%)	V (1.61%)	F (1.43%)	Y (0.95%)	N (1.46%)
98	2	L (1.21%)	N (1.58%)	M (1.45%)	L (1.8%)	S (1.56%)	V (1.46%)	F (1.57%)	Y (1.12%)	N (1.47%)
97	3	L (2.85%)	N (3.05%)	M (2.86%)	L (2.97%)	S (2.62%)	V (2.83%)	F (3.12%)	Y (2.33%)	N (3.29%)
95	5	L (15.47%)	N (16.15%)	M (17.19%)	L (17.48%)	S (16.4%)	V (16.04%)	F (17.63%)	Y (13.99%)	N (17.84%)
90	10	L (14.82%)	N (15.12%)	M (15.87%)	L (16.2%)	S (14.84%)	V (14.81%)	F (15.34%)	Y (12.13%)	N (15.9%)
0	100	L (99.69%)	N (99.15%)	M (99.59%)	L (99.68%)	S (99.72%)	V (99.55%)	F (99.49%)	Y (99.48%)	N (99.67%)

Integrase - Mutation Frequency

% Plasmid in the mixture		E	L
NL4-3 *pol*	08-180 *pol*	92Q	101I
YU2 (R5) *env*	NL4-3 (X4) *env*		
100	0	0%	0%
99.9	0.1	0%	0%
99	1	Q (1.86%)	I (1.94%)
98	2	Q (2.6%)	I (2.04%)
97	3	Q (9.62%)	I (9.39%)
95	5	Q (13.86%)	I (13.71%)
90	10	Q (20.08%)	I (19.38%)
0	100	Q (99.77%)	I (99.63%)

Coreceptor Tropism

% Plasmid in the mixture		X4 (%)
NL4-3 *pol*	08-180 *pol*	
YU2 (R5) *env*	NL4-3 (X4) *env*	
100	0	0%
99.9	0.1	0%
99	1	2.4%
98	2	0%
97	3	13.9%
95	5	4.3%
90	10	5.2%
0	100	100%

Supplemental Material Figure S4. Frequency of amino acids detected in positions associated with drug resistance in the protease, reverse transcriptase, and integrase coding regions using Sanger (population) and deep sequencing (DEEPGEN™HIV).

A. Protease

	L	V	G	K	L	D	V	L	E	M	K	M	I	G	I	F	I	Q	D	I	L	I	H	A	G	T	L	V	V	N	I	I	N	L	L	I
	10	11	16	20	24	30	32	33	34	36	43	46	47	48	50	53	54	58	60	62	63	64	69	71	73	74	76	77	82	83	84	85	88	89	90	93
Population	F	.	E	T	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	.	H	.	.	T	.	.	.	.	T	.	.	.	.	.	.	.
Deep Sequencing	F (97.8)	. (99.6)	E (99.2)	T (94.9) . (4.2)	. (99.0)	. (99.3)	. (95.9) I (3.7)	. (99.2)	. (99.2)	. (99.3)	. (96.6) N (2.0)	. (91.3)	. (92.3) V (3.9) N (2.1)	. (95.1)	. (90.8) L (2.7) M (2.7)	. (90.0) L (8.0)	V (89.3) . (5.9) F (3.0)	. (91.9) K (6.1)	. (99.3)	. (99.7)	H (99.6)	. (99.7)	. (96.9)	T (90.4) . (4.9) H (4.0)	. (99.2)	. (99.6)	. (99.7)	. (99.6)	T (94.4) A (3.5)	. (98.9)	. (99.2)	. (99.4)	. (99.4)	. (99.8)	. (99.3)	. (98.7)

B. RT

	M	A	K	D	T	K	L	V	F	V	A	L	K	K	V	V	Y	F	E	Q	V	Y	M	Y	G	L	T	K	H	P	F	M
	41	62	65	67	69	70	74	75	77	90	98	100	101	103	106	108	115	116	138	151	179	181	184	188	190	210	215	219	221	225	227	230
Population	L	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	V	.	.	L/W	Y	.	.	.	.	.
Deep Sequencing	L (84.9) . (7.9)	. (90.9) D (4.0) V (3.1)	. (98.0)	. (78.1) N (16.4) E (3.1)	. (97.8)	. (98.0)	. (96.5)	. (94.2) I (5.0)	. (96.6) L (3.4)	. (98.6)	. (96.4)	. (96.7)	. (99.5)	. (96.6)	. (99.8)	. (98.5)	. (97.2)	. (99.9)	. (98.0)	. (99.4)	. (98.3)	. (99.2)	V (95.6) . (3.4)	. (99.2)	. (95.1) A (2.7)	. (75.3) W (17.8) M (4.9)	Y (87.5) . (6.3) H (2.5)	. (93.5) Q (3.4) N (2.0)	. (86.3) N (5.7) P (3.3) Q (2.4)	. (91.4) T (7.8)	. (97.4)	. (96.5)

C. Integrase

	H	T	L	E	T	L	F	G	Y	S	Q	S	N	G
	51	66	74	92	97	101	121	140	143	147	148	153	155	193
Population	.	.	.	.	.	.	.	.	.	S/G	R	.	.	.
Deep Sequencing	. (99.4)	. (94.1) I (2.7)	. (98.6)	. (93.1) Q (4.0)	. (95.3) Q (3.3)	. (98.4)	. (98.2)	. (98.6)	. (97.8)	. (66.3) G (33.2)	R (73.5) . (25.9)	. (94.7) Y (4.0)	. (95.7) K (2.1)	. (99.4)

HIV-1 RNA was purified from an antiretroviral-experienced patient and the protease, reverse transcriptase (RT), and integrase coding regions sequenced using Sanger (population) and deep sequencing (DEEPGEN™HIV). Deep sequencing identified all amino acids detected by population sequencing. In addition, a series of amino acids, at a frequency below the detection level of Sanger sequencing (<20%), were only detected using DEEPGEN™HIV. Amino acid substitutions associated with drug resistance are indicated in red.