Supplementary Table S1. Amplification primers for long-range HIV genotyping

Primer	Primer sequence, 5’ ➔ 3’	HXB2 location
First round PCR (RT-PCR):		
OFM19	GCA CTC AAG GCA AGC TTT GAG GCT TA	9,632–9,604
SK145	AGT GGG GGG ACA TCA AGC AGC CAT	1,359–1,382
Second round PCR, Amplicon 1:		
Pan-HIV-1_2F	GGG AAG TGA CAT AGC AGG AAC	1,485–1,505
Pan-HIV-1_2R	CTG CCA TCT GTT TTC CAT AGT C	5,037–5,058
Second round PCR, Amplicon 2:		
Pan-HIV-1_4F	CCT ATG GCA GGA AGA AGC G	5,967–5,985
Pan-HIV-1_4R	CTT ATA TGC AGC ATC TGA GGG	9,497–9,517

Supplementary Table S2. Sequencing primers for long-range HIV genotyping (direct sequencing of amplicon 1)

Primer	Primer sequence, 5’ ➔ 3’	HXB2 location
Major set of sequencing primers		
1340L	CTA TGT GCC CTT CCT TGC	1,976–1,993
981U-M	AAT GTA TAG CCC TGT **Y**AG	1,614–1,631
1870U-M	TTG GCT GAG GCA ATG AG	1,876–1,892
1959L-M	CAT TGT TTA ACC TTT GGG	2,611–2,628
2481U	GAA ATA GGG CAA CAT AGA	3,129–3,146
CR3268	GTC CAT TTR TCA GGA TGG AGY TCA T	3,244-3,265
SQ6.5FC	GCA GAG TTA GAA TTA GCA GAG AAC AG	3,444-3,469
3153L	TAC TAG GGG AGG GGT ATT	3,801–3,818
3658L	CTA TGG GTG GCA GAT TAA	4,306–4,323
CR10	ATT GTG GAT GAA TAC TGC C	4,763-4,781
Pan-HIV-1_2R-M	CTG CCA TCT GTT TTC CAT AGT C	5,037–5,058
SQ8FC	CCT GGT AGC AGT CCA TGT AGC	4,448-4,468
Alternate primers for re-sequencing problematic regions (if any):		
Pan-HIV-1_2F-M	GGG AAG TGA CAT AGC AGG AAC	1,485–1,505
984L	ATG CTG ACA GGG CTA TAC	1,617–1,634
1329U	TTT CAA CTG TGG CAA GGA	1,965–1,982
SQ14R(3)C	TCG AAT CTG AAG CTC TCT GCT GGT G	2,162-2,180
2546-M	CCA TTG AAA CTG TAC CAG	2,561–2,577
1954U	GCC CAA AGG TTA AAC AAT	2,602–2,619
2473L	GCC CTA TTT CTA AGT CAG	3,121-3,138
2973U	AGG ACT GCC CAC ACT AAT	3,621–3,638
2975L	TCA TTA GTG TGG GCA GTC	3,623–3,640
3110U	AGA CTA TTG GCA AGC CAC	3,758–3,775
3891L	TGG CCA TCT TCC TGC TAA	4,539–4,556
3969U	TGT TGG TGG GCA GGT ATC	4,617–4,634

Supplementary Table S3. Sequencing primers for long-range HIV genotyping (clone sequencing)

Primer	Primer sequence, 5’ ➔ 3’	HXB2 location
Amplicon 1		
pMini_F	ACC TGC CAA CCA AAG CGA GAA C	N/A
pMini_R	TCA GGG TTA TTG TCT CAT GAG CG	N/A
1870U-M	TTG GCT GAG GCA ATG AG	1,876–1,892
1959L-M	CAT TGT TTA ACC TTT GGG	2,611–2,628
CR3268	GTC CAT TTR TCA GGA TGG AGY TCA T	3,244-3,265
SQ6.5FC	GCA GAG TTA GAA TTA GCA GAG AAC AG	3,444-3,469
3153L	TAC TAG GGG AGG GGT ATT	3,801–3,818
CR10	ATT GTG GAT GAA TAC TGC C	4,763-4,781
Amplicon 2		
pMini_F	ACC TGC CAA CCA AAG CGA GAA C	N/A
pMini_R	TCA GGG TTA TTG TCT CAT GAG CG	N/A
6256L	TTA CAC TTT AGA ATC GCA	6,893–6,910
ED5	ATG GGA TCA AAG CCT AAA GCC ATG TG	6,557–6,582
ED12	AGT GCT TCC TGC TGC TCC CAA GAA CCC AAG	7,782–7,811
7182U	TGA CGG TAC AGG CCA GAC	7,834–7,851
7724L	GTT TGG GGT AAG GGT CTG	8,376–8,393
7847U	CGA CCT GCG GAG CCT GTG C	8,498–8,516

Supplementary Table S4. Distribution of selected drug-resistance mutations associated with APOBEC-induced hypermutations in MPP HIV-1C sequences (D30N, M46I and G73N are within Protease; D67, K103N and M184I are within Reverse Transcriptase; and E138K is within Integrase; drug-resistant mutations are highlighted).

	Selected drug-resistance mutations								
MPP sequences, n=36	D30N	M46I	G73S	D67N	K103N	M184I	E138K	Muts/ seq length	Hypermut ratio
mpp_00255_amp1	AAT	ATA	GGT	GAC	AAA	ATA	GAA	0.034	4.64
mpp_00267_amp1	GAT	ATG	GGT	GAC	AAC	ATG	GAA	0.003	0.63
mpp_00318_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.025	3.79
mpp_00018_amp1	AAT	ATA	AGT	GAC	AAA	ATA	AAA	0.082	3.61
mpp_00419_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.030	4.13
mpp_00450_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.021	2.3
mpp_00283_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.037	4.86
mpp_00136_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.018	3.58
mpp_00274_amp1	GAT	ATG	AGT	GAT	AAA	ATA	GAG	0.046	6.32
mpp_00142_amp1	AAT	ATA	AGT	AAC	AAA	ATA	AAA	0.075	3.54
mpp_00147_amp1	GAT	ATG	AGT	GAC	AAA	ATG	GAA	0.037	6.66
mpp_00378_amp1	GAT	ATA	AGT	GAC	AAA	ATA	GAA	0.058	8.77
mpp_00162_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.031	5.2
mpp_00124_amp1	AAT	ATG	AGT	GAC	AAA	ATA	AAA	0.057	6.17
mpp_00032_amp1	AAT	ATA	AGT	GAC	AAA	ATG	GAA	0.051	3.07
mpp_00470_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.023	4.23
mpp_00275_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.033	5.34
mpp_00134_amp1	AAT	ATA	AGT	GAC	AAC	ATA	GAA	0.044	5.91
mpp_00105_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.032	3.65
mpp_00328_amp1	AAT	ATA	GGT	GAC	AAA	ATA	AAA	0.061	4.32
mpp_00416_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.031	3.8
mpp_00248_amp1	GAT	ATG	GGT	GAC	AAG	ATA	GAA	0.028	3.89
mpp_00212_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.030	3.94
mpp_00137_amp1	AAT	ATA	GGT	AAC	AAA	ATG	AAA	0.054	3.73
mpp_00199_amp1	GAT	ATG	GGT	GAC	AAT	ATG	GAA	0.005	1.28
mpp_00411_amp1	GAT	ATG	AGT	GAC	AAA	ATG	GAA	0.027	5.2
mpp_00309_amp1	GAT	ATG	AGC	GAC	AAA	ATA	GAA	0.041	5.45
mpp_00476_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.028	4.66
mpp_00148_amp1	AAT	ATA	GGT	GAT	AAA	ATA	GAA	0.046	3.7
mpp_00214_amp1	AAT	ATG	AAG	GAC	AAA	ATA	GAA	0.040	6.92
mpp_00446_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.006	1.38
mpp_00234_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.051	5.17
mpp_00242_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	0.026	3.07
mpp_00206_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.043	5.87
mpp_00110_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.029	5.73
mpp_00491_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	0.031	3.23

Supplementary Table S5. Distribution of selected drug-resistance mutations associated with APOBEC-induced hypermutations in BCPP HIV-1C sequences (D30N, M46I and G73N are within Protease; D67, K103N and M184I are within Reverse Transcriptase; and E138K is within Integrase; drug-resistant mutations are highlighted).

	Selected drug-resistance mutations								
BCPP sequences, n=46	D30N	M46I	G73S	D67N	K103N	M184I	E138K	Hypermut ratio	Muts/ seq length
bcpp_00153_amp1	AAT	ATG	GGT	GAC	AAA	ATA	GAA	3.42	0.049
bcpp_00452_amp1	AAT	ATA	AGT	GAC	AAA	ATG	GAA	3.37	0.043
bcpp_00055_amp1	AAT	ATA	GGT	GAT	AAA	ATG	GAA	3.2	0.037
bcpp_00110_amp1	AAT	ATG	GGT	GAC	AAA	ATA	GAA	5.01	0.039
bcpp_00453_amp1	GAT	ATA	AGT	GAC	AAA	ATA	GAA	4.61	0.049
bcpp_00032_amp1	GAT	ATG	AGT	GAC	AAA	ATA	AAA	5.77	0.034
bcpp_00341_amp1	GAT	ATG	AGT	GAC	AAA	ATA	AAA	6.76	0.032
bcpp_00462_amp1	AAT	ATA	GGT	GAC	AAA	ATG	AAA	3.49	0.059
bcpp_00014_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	5.47	0.057
bcpp_00440_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	6.18	0.037
bcpp_00311_amp1	GAT	ATG	AGT	GAT	AAA	ATA	GAA	4.45	0.032
bcpp_00458_amp1	AAT	ATA	AGT	GAC	AAA	ATA	AAA	3.26	0.049
bcpp_00305_amp1	GAC	ATA	AGT	GAC	AAA	ATA	GAA	5.54	0.044
bcpp_00085_amp1	GAT	ATG	AGT	GAC	AGA	ATA	GAA	4.09	0.028
bcpp_00464_amp1	GAC	ATA	GGT	AAT	AAA	ATA	AAA	3.86	0.065
bcpp_00350_amp1	AAT	ATA	GGT	GAC	AAA	ATA	AAA	3.47	0.060
bcpp_00450_amp1	GAT	ATG	GGT	GAC	AAC	ATG	GAA	1.22	0.005
bcpp_00070_amp1	GAT	ATG	AGT	GAC	AAG	ATA	GAA	3.24	0.027
bcpp_00368_amp1	GAT	ATA	GGT	GAC	AAA	ATA	GAA	6.78	0.040
bcpp_00100_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	3.2	0.023
bcpp_00202_amp1	AAT	ATA	GGT	GAC	AAA	ATG	GAA	2.84	0.032
bcpp_00140_amp1	AAT	ATA	GGT	AAC	AAA	ATG	AAA	2.79	0.029
bcpp_00094_amp1	GAT	ATA	AGT	GAC	AAA	ATT	GAA	5.4	0.042
bcpp_00027_amp1	GAT	ATG	GGT	GAC	AAT	GTG	GAA	1.41	0.005
bcpp_00436_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	3.8	0.023
bcpp_00415_amp1	AAT	ATA	GGT	GAC	AAG	ATA	GAA	2.4	0.031
bcpp_00466_amp1	GAT	ATA	AGT	GAC	AAA	ATA	GAA	10.83	0.032
bcpp_00363_amp1	GAT	ATG	AGT	GAC	AAA	ATA	GAA	3.42	0.032
bcpp_00381_amp1	GAT	ATA	AGT	GAC	AAA	ATA	AAA	5.01	0.043
bcpp_00442_amp1	GAT	ATG	GGT	GAT	AAC	ATG	GAA	0.78	0.004
bcpp_00356_amp1	GAT	ATG	AGT	GAC	AAG	ATG	GAA	2.89	0.016
bcpp_00234_amp1	AAT	ATA	GGT	GAC	AAA	GTA	AAA	3.74	0.045
bcpp_00186_amp1	GAT	ATG	GGT	GAT	AAA	ATA	GAA	2.81	0.018
bcpp_00109_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	4.69	0.035
bcpp_00374_amp1	AAT	ATG	RGT	GAC	AAA	ATA	GAA	6.19	0.052
bcpp_00445_amp1	GAT	ATG	GGT	GAC	AAC	ATG	GAA	0.84	0.004
bcpp_00227_amp1	GAT	ATA	GGT	GAC	AAA	ATG	AAA	6.9	0.036

bcpp_00414_amp1	GAT	ATG	GGT	GAC	AAT	ATG	GAA	0.86	0.006
bcpp_00190_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	4.62	0.022
bcpp_00498_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	5.67	0.026
bcpp_00051_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	3.96	0.025
bcpp_00395_amp1	GAT	ATT	GGT	AAC	AAA	ATG	GAA	1.26	0.008
bcpp_00171_amp1	GAT	ATG	AGT	GAC	AAA	ATA	AAA	4.39	0.039
bcpp_00402_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	15.36	0.040
bcpp_00257_amp1	GAT	ATG	GGT	GAC	AAA	ATA	GAA	7.1	0.047
bcpp_00136_amp1	GAT	ATG	GGT	AAC	AAC	GTG	GAA	0.33	0.003