Supplementary file of the manuscript entitled: 'HIV-1 subtype is an independent predictor of reverse transcriptase mutation K65R in HIV-1 patients treated with combination antiretroviral therapy including tenofovir'

Table S1: Pairwise subtype comparisons of K65R prevalence across HIV-1 subtypes

	Subtype A	Subtype B	Subtype C	Subtype F	Subtype G	CRF 02_AG
Subtype A	-	0.481	0.051	1.000	0.161	0.519
Subtype B	0.481	-	0.016	0.535	0.069	0.923
Subtype C	0.051	0.016	-	0.089	0.346	0.195
Subtype F	1.000	0.535	0.089	-	0.229	0.549
Subtype G	0.161	0.069	0.346	0. 229	-	0.569
CRF 02_AG	0.519	0.923	0.195	0. 549	0.569	-

This matrix shows the p-values for pairwise comparisons of K65R prevalence across subtypes. A chi-square test was used to compute the p-value based on a two-by-two contigency table of K65R presence for each pair of subtypes. Compared to the K65R prevalence in subtype B, only subtype C viruses displayed a significant different prevalence. Please see Table 2 in the manuscript text for the absolute patient count and the number of K65R cases for each subtype.

Table S2: Distribution of K65R prevalence across regression variables

Variables	Non-K65R	K65R
Prior exposure to TDF therapy (n)	93.3	6.7
Prior exposure to ABC or ddI therapy (n)	91.9	8.1
Second drug in combination with TDF (n)		
ABC	90.1	9.9
ddI	75.0	25.0
AZT	98.4	1.6
d4T	98.2	1.8
3TC	89.2	10.8
FTC	95.0	5.0
Third drug in combination with TDF (n)		
Boosted PI	97.8	2.2
Unboosted PI	96.4	3.6
NNRTIs	78.2	21.8
EFV	81.1	18.9
NVP	72.3	27.6
Co-occurrence of RT mutations (n)		
TAM 1	99.1	0.9
TAM 2	94.8	5.2
NAMs	66.9	33.1
M184V	81.9	18.1
NNRTI mutations	90.1	9.9

To illustrate varying K65R selection in our study population, this table shows K65R prevalence for discrete variables included in multivariate regression. These findings are in accordance with the factors identified by the regression model as significant predictors of K65R selection. Primarily the co-administration of didanosine or NNRTIs with tenofovir resulted in high selection of K65R, while the co-occurrence of K65R with TAMs is rarely observed. Abbreviations are as follows: ABC - Abacavir, ddI - didanosine, AZT - zidovudine, d4T - stavudine, 3TC - lamivudine, FTC - emtricitabine, PI - protease inhibitor, NNRTI - non-nucleoside reverse transcriptase inhibitor, EFV - efavirenz, NVP - nevirapine. Mutation patterns in RT were defined as thymidine analogue mutations (TAM) 1 (M41L, L210W or T215Y), TAM 2 (D67N, K70R, 215F or D219E/Q), M184V, other nucleoside analogue mutations (NAMs) (A62V, V75I, F77L, F116Y or Q151M) and non-nucleoside RT inhibitor (NNRTI) mutations (L100I, K103N or Y181C).

Table S3: Subtype distribution and K65R selection in the study by Gupta et al.

	Non-K65R	K65R
Subtype A	48	0
Subtype B	235	13
Subtype C	83	10

In the manuscript entitled 'K65R and Y181C are less prevalent in HAART-experienced HIV-1 subtype A patients' published by AIDS in 2005 by Gupta et al. (PMID 16227803), a lower prevalence of K65R in patients infected with subtype A was reported compared to those infected with subtypes B or C. In this table, the number of patients and the number of K65R detected are shown for subtypes A, B and C.

Table S4: A re-analysis of the reported subtype association with K65R selection in the study by Gupta et al.

Comparison of subtype combinations	P-value
A ~ B ~ C	0.022
A ~ B	0.137
A ~ C	0.016
B ~ C	0.089
A ~ (B + C)	0.094
(A + B) ~ C	0.040
B ~ (A+C)	0.505

Based on the information provided in Table S3, we compared the prevalence of K65R for all the different subtype combinations possible. The chi-square test of independence was used to compute statistical significance values, following the methodology used by Gupta et al. These results show that subtype A viruses do have a lower prevalence of K65R compared to subtype B or to the combined population of subtype B and subtype C viruses, as stated otherwise in the original manuscript by Gupta et al. Similar to the findings in our manuscript, subtype C viruses displayed a significant difference in K65R prevalence compared to subtype B viruses.