

Table S1. Number of amino acid substitutions per subject in Gag during primary HIV-1 subtype C infection.

Subject	Fiebig stage	Time of follow-up, days p/s	ART (if any), days p/s initiation	Num of time points	Total amino acid substitutions			Dominant and fixed substitutions (subset)		
					Reverse	Forward	Poly-morphisms	Reverse	Forward	Poly-morphisms
NA	V-VI	169 – 448		4	10	26	2	7	3	0
A	II	6 – 356	181	10	9	25	2	4	1	2
OC	IV	27 – 482		6	18	35	4	8	1	1
OE	V	52 – 236	491	5	18	13	6	13	0	0
OG	V	104 – 317	450	7	20	43	8	14	1	2
OJ	IV	44 – 316		4	10	16	3	7	0	0
B	II	10 – 223	199	6	5	22	3	2	0	0
OM	V-VI	86 – 178		2	3	4	1	1	1	0
OQ	V	52 – 415	270	5	16	20	1	14	3	0
OS	V	62 – 494		6	16	39	3	15	4	0
OU	IV	13 – 471		7	6	42	9	1	2	0
OW	IV	13 – 469		6	14	38	7	11	3	4
OX	V	58 – 464		6	8	18	2	6	3	1
OY	V	51 – 423		7	23	21	8	20	0	3
OZ	V	78 – 395		6	20	22	1	14	2	0
C	II	4 – 445	362	7	9	22	0	6	1	0
O1	IV	6 – 312		5	4	23	0	1	0	0
PA	V	65 – 227		4	9	6	0	6	0	0
E	II	30 – 404		8	21	22	4	15	4	2
PC	V	54 – 418	388	6	11	15	6	9	0	2
F	II	7 – 447		7	7	14	2	6	4	0
PD	V	59 – 428		5	1	3	2	0	1	1
G	II	4 – 438		6	5	15	4	3	2	0
PK	V	47 – 195		4	5	7	3	4	3	0
D	II	6 – 483		7	26	15	5	7	4	1
PO	IV	6 – 277		5	3	14	1	0	3	0
PP	V	44 – 226		4	24	26	2	14	4	1
QA	IV	6 – 389	150	5	1	18	8	1	0	2
QC	V	80 – 415		3	4	25	1	1	2	1
QD	V-VI	79 – 420		4	5	12	1	2	0	0
QG	V	48 – 415		5	2	3	0	2	0	0
H	II	16 – 347		6	4	10	3	3	0	1
QI	IV	20 – 476	112	5	6	5	0	5	0	0
QJ	IV	8 – 471		7	1	16	1	0	4	0
QM	V	48 – 415		4	1	4	0	0	0	0
QP	V	48 – 414		4	7	25	3	4	3	0
QR	IV	7 – 485		5	16	26	0	16	5	0
QS	V	44 – 351		5	1	13	0	1	1	0
QT	V	44 – 349		5	2	15	2	0	1	0
QU	V-VI	84 – 266		4	10	14	0	8	0	0
RA	V	54 – 422		6	6	29	1	2	0	0
RB	V	52 – 332	424	4	6	21	5	5	2	3
Median (IQR):			316 (185-415)		7 (4-15.5)	18 (15.3-25)	2 (1-4)	5 (1.3-8.8)	1 (0-3)	0 (0-1)