

## **SUPPLEMENTARY INFORMATION**

### **HIITE: HIV Incidence and Infection Time Estimator**

Sung Yong Park<sup>1</sup>, Tanzy M. T. Love<sup>2</sup>, Shivankur Kapoor<sup>1</sup> and Ha Youn Lee<sup>1\*</sup>

<sup>1</sup> Department of Molecular Microbiology and Immunology, Keck School of Medicine, University of Southern California, Los Angeles, CA, United States

<sup>2</sup> Department of Biostatistics and Computational Biology, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States

Subject (Reference)	Minimum Duration of Infection (days)	Subtype	Risk Behavior	Sex	ART status	Viral Load (RNA copies/ml)	CD4 Count (cells/mm <sup>3</sup> )
<u>Bailey</u> <sup>1</sup>							
DQ410040-DQ410649 <sup>#</sup>							
C61	1,798	B	NS	F	Naïve	19,100	1,261
C62	2,071	B	NS	M	Naïve	33,300	481
C93	1,492	B	NS	M	Naïve	47,270	402
C94	1,886	B	NS	M	Naïve	22,898	351
C96	1,308	B	NS	M	Naïve	12,500	400
C109	882	B	NS	F	Naïve	61,000	222
<u>Yoshida</u> <sup>2</sup>							
AB588196-AB588337							
7060	1,460	B	MSM	M	Naïve	NS	NS
7015	1,460	B	Heterosexual	NS	Naïve	NS	NS
6826	1,460	B	MSM	M	Naïve	NS	NS
6739	1,460	B	Heterosexual	NS	Naïve	NS	NS
<u>Ren</u> <sup>3</sup>							
KC807912- KC807980							
GX752	1,095	C	NS	NS	Naïve	NS	NS
GX757	1,095	C	NS	NS	Naïve	NS	NS
0031	1,095	R-B/C	NS	NS	Naïve	NS	NS
0651	1,095	R-B/C	NS	NS	Naïve	NS	NS
1081	1,095	R-B/C	NS	NS	Naïve	NS	NS
CNXJ09277	1,095	C	NS	NS	Naïve	NS	NS
<u>Rong</u> <sup>4</sup>							
GQ485312- GQ485447							
ZM185F	787	C	Heterosexual	F	Naïve	NS	NS
	856	C	Heterosexual	F	Naïve	NS	NS
<u>Bunnik</u> <sup>5</sup>							
EU743973 -EU744175							
ACH18969	1,430	B	MSM	M	Naïve	NS	NS
	2,069	B	MSM	M	Experienced	NS	NS
	2,768	B	MSM	M	Experienced	NS	NS

ACH19542	1,308	B	MSM	M	Naïve	NS	NS
	1,916	B	MSM	M	Naïve	NS	NS
	2,616	B	MSM	M	Experienced	NS	NS
	3,559	B	MSM	M	Experienced	NS	NS
ACH19659	913	B	MSM	M	Naïve	NS	NS
	1,886	B	MSM	M	Naïve	NS	NS
	2,920	B	MSM	M	Experienced	NS	NS
ACH19768	1,095	B	MSM	M	Experienced	NS	NS
	2,038	B	MSM	M	Experienced	NS	NS
	2,829	B	MSM	M	Experienced	NS	NS
	3,863	B	MSM	M	Experienced	NS	NS
	4,472	B	MSM	M	Experienced	NS	NS
ACH19999	791	B	MSM	M	Naïve	NS	NS
	2,251	B	MSM	M	Naïve	NS	NS
	3,285	B	MSM	M	Experienced	NS	NS
<u>Edo-Matas</u> <sup>6</sup>							
EU743973.1-EU744175.1							
19383	1,886*	B	MSM	M	Naïve	NS	NS
	3,255*	B	MSM	M	Naïve	NS	NS
	4,045*	B	MSM	M	Naïve	NS	NS
19417	2,342*	B	MSM	M	Naïve	NS	NS
	3,072*	B	MSM	M	Naïve	NS	NS
19559	4,045*	B	MSM	M	Naïve	NS	NS
19576	882*	B	MSM	M	Naïve	NS	NS
	1,308*	B	MSM	M	Naïve	NS	NS
19828	1,430*	B	MSM	M	Naïve	NS	NS
19858	1,278*	B	MSM	M	Naïve	NS	NS
	3,437*	B	MSM	M	Naïve	NS	NS
19922	2,494*	B	MSM	M	Naïve	NS	NS
	4,106*	B	MSM	M	Naïve	NS	NS
<u>Van-Gils</u> <sup>7</sup>							
EU744055-EU744096							
GU455425-GU455525							
19298	1,034	B	MSM	M	Naïve	NS	NS
	1,460	B	MSM	M	Naïve	NS	NS
	2,190	B	MSM	M	Experienced	NS	NS
	2,647	B	MSM	M	Experienced	NS	NS

19554	1,430	B	MSM	M	Naïve	NS	NS
	2,069	B	MSM	M	Naïve	NS	NS
	2,525	B	MSM	M	Experienced	NS	NS
	3,255	B	MSM	M	Experienced	NS	NS
19663	1,430	B	MSM	M	Naïve	NS	NS
	2,768	B	MSM	M	Naïve	NS	NS
	3,377	B	MSM	M	Naïve	NS	NS
	4,259	B	MSM	M	Naïve	NS	NS
19642	882	B	MSM	M	Naïve	NS	NS
	1,491	B	MSM	M	Naïve	NS	NS
	2,251	B	MSM	M	Naïve	NS	NS
	3,285	B	MSM	M	Naïve	NS	NS
	3,985	B	MSM	M	Naïve	NS	NS
	4,320	B	MSM	M	Naïve	NS	NS

Gnanakaran <sup>8</sup>

HQ216367-HQ218052

HQ238279-HQ238288

MCST4474	3,468	B	MSM	M	Naïve	29,500	634
RHGA1581	3,468	B	MSM	M	Naïve	25,901	571
STCO5453	730	B	MSM	M	Naïve	67,964	796
WARO5662	913	B	Heterosexual	F	Naïve	16,758	598

Keele <sup>9</sup>

EU574937-EU579293

CRPE4571	730	B	Heterosexual	F	Naïve	21,917	67
FOJO4081	5,718	B	MSM	M	Naïve	73,677	408
LAHA4867	730	B	MSM	M	Naïve	53,500	610
MCRO3633	2,738	B	Heterosexual	F	Naïve	13,175	305
OLLA4645	760	B	Heterosexual	F	Naïve	382,000	150
SAMI4303	1,430	B	MSM	M	Naïve	116,000	1,012
TALA4022	2,525	B	MSM	M	Naïve	228,200	8
WICU4248	2,160	B	MSM	M	Naïve	8,424	564
YOMI4024	2,220	B	MSM	M	Naïve	14,178	486
UNC3405	2,890	B	MSM	M	Naïve	29,100	495
UNC5057	1,764	B	Heterosexual	F	Naïve	141,000	277
UNC5417	1,156	B	Heterosexual	M	Naïve	48,000	340
UNC5479	1,673	B	Heterosexual	F	Naïve	15,400	382
UNC5539	1,855	B	MSM	M	Naïve	46,600	355

UNC5769	973	B	MSM	M	Naïve	18,100	466
UNC5791	1,582	B	Heterosexual	F	Naïve	65,566	262
UNC7092	5,475	B	Heterosexual	F	Naïve	64,700	333
UNC2009	1,430	B	MSM	M	Naïve	56,000	343
UNC4295	2,707	B	MSM	M	Naïve	74,800	495
UNC4484	2,251	B	Heterosexual	F	Naïve	7,580	244
UNC4911	1,886	B	Heterosexual	F	Naïve	13,500	438
UNC5283	1,521	B	Heterosexual	F	Naïve	43,100	455
UNC5548	1,308	B	Heterosexual	F	Naïve	1,560	337
UNC5734	1,369	B	Heterosexual , IDU	F	Naïve	12,800	317
UNC5799	4,015	B	MSM	M	Naïve	81,400	211
UNC6064	1,491	B	MSM	M	Naïve	10,400	468
SC02	1,095	B	Heterosexual	M	Naïve	76,615	NS
SC03	2,525	B	Heterosexual	M	Naïve	528,140	121
SC05	1,977	B	Heterosexual	M	Naïve	19,514	NS
SC13	973	B	Heterosexual	M	Naïve	12,054	NS
SC24	1,399	B	Heterosexual	M	Naïve	178,753	NS
SC25	1,004	B	Heterosexual	F	Naïve	218,551	NS
TT31P	730	B	Heterosexual	F	Naïve	19,909	525
<u>Herbeck</u> <sup>10</sup>							
JN024459-JN024468							
PIC51550	3,285	B	MSM	M	Naïve	28,610	NS
<u>Herbeck</u> <sup>10</sup> <u>Liu</u> <sup>11</sup>							
DQ853426-DQ854622							
PIC1365	3,653*	B	MSM	M	Naïve	NS	NS
<u>Daniels</u> <sup>12</sup>							
AJ535590-AJ535607							
Pat-1	3,285	B	NS	M	Naïve	NS	220
Pat-2	5,476	B	NS	M	Naïve	NS	170
<u>Doria-Rose</u> <sup>13</sup>							
KF996701-KF996716							
CAP256	1,233	C	Heterosexual	F	Naïve	NS	NS
<u>Wu</u> <sup>14</sup>							
JQ609684-JQ610156							
1	3,650	B	NS	NS	Naïve	14,650	1,160
	5,840	B	NS	NS	Naïve	10,526	568

	7,665	B	NS	NS	Experienced	19,260	545
18	2,920	B	NS	NS	Naïve	24,648	462
45	4,015	B	NS	NS	Naïve	11,416	727
	5,840	B	NS	NS	Naïve	9,129	638
	6,935	B	NS	NS	Naïve	5,153	686
N26	7,665	B	NS	NS	Naïve	35,912	387
N90	8395	B	NS	NS	Naïve	8,216	912
<u>Blish<sup>15</sup>, Piantadosi<sup>16</sup></u>							
FJ396012-FJ396031							
EU163983-EU164399							
QA413	1,007	A	Heterosexual	F	Naïve	NS	296
	1,146*	A	Heterosexual	F	Naïve	NS	NS
	1,346*	A	Heterosexual	F	Naïve	NS	NS
<u>Geels<sup>17</sup></u>							
AY423381-AY423387							
H671	1,430	B	MSM	M	Naïve	NS	NS
<u>Sturdevant<sup>18</sup></u>							
KM353586-KM355197							
9018	873	B	NS	NS	Experienced	89,125	453
	980	B	NS	NS	Experienced	95,499	505
9040	753	B	NS	NS	Naïve	275,422	279
<u>SSemwanga<sup>19</sup></u>							
HM027653-HM027878							
PP1_F1	883	D	Heterosexual	F	NS	NS	NS
PP1_F3	2,312	D	Heterosexual	F	NS	NS	NS
PP1_F4	2,570	A/D	Heterosexual	F	NS	NS	NS
PP1_M	2,824	D	Heterosexual	M	NS	NS	NS
<u>Skar<sup>20</sup></u>							
JN251812-JN251888							
45755	2,555	B	NS	M	Naïve	832	NS
	2,556	B	NS	M	Naïve	794	NS
	2,557	B	NS	M	Naïve	1220	NS
	2,565	B	NS	M	Naïve	563	NS
	2,572	B	NS	M	Naïve	518	NS
	2,579	B	NS	M	Naïve	450	NS
	2,586	B	NS	M	Naïve	600	NS

	3,076	B	NS	M	Naïve	NS	NS
<u>Ping</u> <sup>21</sup>							
KC862610-KC863648							
HM638963-HM639197							
CAP61	835	C	Heterosexual	F	NS	NS	NS
DUI23	1,702	C	Heterosexual	F	NS	NS	NS
<u>Mukhopadhyay</u> <sup>22</sup>							
FJ515874-FJ515878							
JN400538-JN400545							
991566	2,920	B/C	NS	NS	Naïve	NS	720
<u>Malherbe</u> <sup>23</sup>							
KJ698244-KJ698348							
VC10014	1,037	B	Heterosexual		Naïve	NS	NS
	1,310	B	Heterosexual		Naïve	NS	NS
VC20013	774	B	NS		Naïve	NS	NS
<u>Evering</u> <sup>24</sup>							
KM258899 - KM259615							
NCN1	1,813	B	MSM	M	Experienced	65,600	591
NCN2	5,265	B	MSM	M	Naïve	749	802
NCN3	1,488	B	MSM	M	Experienced	8,620	687
NCN4	6,521	B	MSM	M	Naïve	13,900	790
NCN5	2,291	B	MSM	M	Naïve	19,700	420
NCN6	4,166	B	MSM	M	Experienced	18,900	418
ANI1	4,688	B	MSM	M	Naïve	5,320	743
ANI2	6,440	B	MSM	M	Experienced	29,800	220
ANI3	1,405	B	MSM	M	Naïve	1,890	819
ANI4	1,709	B	MSM	M	Naïve	38,600	237
ANI5	806	B	MSM	M	Naïve	9,210	244
ANI6	1,817	B	MSM	M	Experienced	13,000	329
ANI7	6,284	B	MSM	M	Experienced	50,100	466
MND1	2,282	B	MSM	M	Naïve	6,250	398
MND2	1,780	B	Heterosexual	F	Experienced	44,900	320
<u>Shankarappa</u> <sup>25</sup>							
<u>Jensen</u> <sup>26</sup>							
AF137629-AF138703							
AY348333-AY450257							
p1	1,044*	B	MSM	M	Naïve	13,243	845
	1,369*	B	MSM	M	Naïve	53,211	755

p2	1,544*	B	MSM	M	Naïve	13,614	463
	1,851*	B	MSM	M	Naïve	11,455	484
	1,993*	B	MSM	M	Naïve	17,258	534
	2,081*	B	MSM	M	Naïve	NS	524
	2,354*	B	MSM	M	Naïve	103,039	505
	2,435*	B	MSM	M	Naïve	206,063	147
	2,657*	B	MSM	M	Naïve	53,333	44
	2,851*	B	MSM	M	Naïve	107,152	59
	2,982*	B	MSM	M	Naïve	20,230	27
	3,194*	B	MSM	M	Naïve	41,020	42
	927*	B	MSM	M	Naïve	6,067	1,579
	1,205*	B	MSM	M	Naïve	3,532	1,556
	1,566*	B	MSM	M	Naïve	220,293	1,784
	1,843*	B	MSM	M	Naïve	46,774	1,667
	2,070*	B	MSM	M	Naïve	34,995	1,710
	2,230*	B	MSM	M	Naïve	48,753	1,654
	2,420*	B	MSM	M	Naïve	202,768	1,112
	2,595*	B	MSM	M	Naïve	345,939	693
	2,763*	B	MSM	M	Naïve	269,153	470
	p3	3,139*	B	MSM	M	Naïve	83,176
3,829*		B	MSM	M	Experienced	163,682	100
785*		B	MSM	M	Naïve	6,237	686
920*		B	MSM	M	Naïve	7,709	576
1,292*		B	MSM	M	Naïve	1,282	672
1,668*		B	MSM	M	Naïve	17,783	734
2,033*		B	MSM	M	Naïve	2,466	624
2,219*		B	MSM	M	Naïve	19,953	367
2,424*		B	MSM	M	Experienced	117,490	254
2,935*		B	MSM	M	Experienced	59,429	NS
p4	821*	B	MSM	M	NS	16,218	513
	989*	B	MSM	M	NS	3,715	529
	1,164*	B	MSM	M	NS	72,277	411
	1,278*	B	MSM	M	NS	NS	387
	1,453*	B	MSM	M	NS	91201	278
	1,599*	B	MSM	M	NS	62087	370
	1,796*	B	MSM	M	NS	4667	280
	1,971*	B	MSM	M	NS	209411	278



p5	2,154*	B	MSM	M	NS	254097	189
	2,321*	B	MSM	M	NS	130017	47
	2,712*	B	MSM	M	NS	232809	8
	843*	B	MSM	M	Naïve	66,834	742
	1,029*	B	MSM	M	Naïve	16,904	692
	1,223*	B	MSM	M	Naïve	101,625	500
	1,318*	B	MSM	M	Naïve	32,137	869
	1,497*	B	MSM	M	Naïve	21,330	615
	1,701*	B	MSM	M	Naïve	26,485	498
	1,894*	B	MSM	M	Naïve	77,268	334
p6	2,070*	B	MSM	M	Naïve	182,390	241
	2,456*	B	MSM	M	Experienced	41,591	18
	730*	B	MSM	M	Naïve	57,943	526
	902*	B	MSM	M	Naïve	53,211	404
	1,080*	B	MSM	M	Naïve	133,968	371
	1,263*	B	MSM	M	Naïve	228,034	291
	1,445*	B	MSM	M	Experienced	240,991	311
	1,628*	B	MSM	M	Experienced	240,991	276
	2,208*	B	MSM	M	Experienced	263,027	24
	763*	B	MSM	M	Naïve	6,823	739
p7	931*	B	MSM	M	Naïve	21,577	685
	1,343*	B	MSM	M	Naïve	21,979	460
	1,518*	B	MSM	M	Experienced	9,162	543
	1,690*	B	MSM	M	Experienced	124,738	494
	1,872*	B	MSM	M	Experienced	50,234	537
	2,252*	B	MSM	M	Experienced	42,855	373
	2,438*	B	MSM	M	Experienced	76,384	237
	876*	B	MSM	M	Naïve	9,441	545
	1,066*	B	MSM	M	Naïve	7,194	448
	1,234*	B	MSM	M	Naïve	35,075	591
p8	1,409*	B	MSM	M	Naïve	21,577	463
	1,617*	B	MSM	M	Naïve	22,594	359
	1,799*	B	MSM	M	Naïve	35,400	410
	1,989*	B	MSM	M	Naïve	113,763	278
	2,143*	B	MSM	M	Naïve	28,445	259
	2,456*	B	MSM	M	Experienced	17,783	55
	2,887*	B	MSM	M	Experienced	23,335	87

p9	1,172*	B	MSM	M	Naïve	8,204	625
	1,905*	B	MSM	M	Experienced	528	1,115
	2,621*	B	MSM	M	Experienced	4,477	1,295
	3,172*	B	MSM	M	Experienced	12,735	947
	3,362*	B	MSM	M	Experienced	29,717	NS
	3,544*	B	MSM	M	Experienced	73,451	549
	3,712*	B	MSM	M	Experienced	73,451	258
	3,909*	B	MSM	M	Experienced	201,837	612
	4,084*	B	MSM	M	Experienced	333,426	202
p10	785*	B	MSM	M	NS	802	1,588
	975*	B	MSM	M	NS	575	1,615
	1,336*	B	MSM	M	NS	3192	1,681
	1,511*	B	MSM	M	NS	NS	1,740
	1,891*	B	MSM	M	NS	NS	1,062
	2,051*	B	MSM	M	NS	NS	1,260
	2,241*	B	MSM	M	NS	NS	1,140
	2,424*	B	MSM	M	NS	NS	1,019
	2,602*	B	MSM	M	NS	11,041	1,095
	2,792*	B	MSM	M	NS	NS	953
	3,362*	B	MSM	M	NS	NS	1,030
	3,548*	B	MSM	M	NS	NS	774
	4,081*	B	MSM	M	NS	NS	885
	4,263*	B	MSM	M	NS	NS	805
	4,464*	B	MSM	M	NS	759	808
	4,668*	B	MSM	M	NS	437	789
	4,811*	B	MSM	M	NS	564	779
p11	883*	B	MSM	M	Naïve	741	831
	1,270*	B	MSM	M	Naïve	490	840
	1,763*	B	MSM	M	Naïve	468	1,230
	2,128*	B	MSM	M	Naïve	1,778	1,118
	3,030*	B	MSM	M	Naïve	NS	823
<u>Park</u> <sup>27</sup>		B					
KF945693-KF945994							
PL7408	749*	B	MSM	M	Naïve	507	840
IJ7234	2,984*	B	MSM	M	Naïve	5,599	364
QZ8149	3,216*	B	Heterosexual	F	Naïve	15,000	542
LE2707	2,407*	B	Heterosexual	F	Naïve	1,394	607

CX7332	5,464*	B	Heterosexual	M	Experienced	389,081	63
KI6633	2,428*	B	Heterosexual	M	Experienced	460,004	115
CF6610	3,710*	B	MSM	M	Experienced	17,001	634
<u>Liao</u> <sup>28</sup>							
KC247403-KC247483							
CH505	924	C	Heterosexual	M	Naive	51,554	294
	1,092	C	Heterosexual	M	Naive	110,230	339
<u>Park</u> <sup>29</sup>							
MF537774-MF538517							
TY2947	732*	B	NS	F	Naive	6400	767
	1,105*	B	NS	F	Naive	8500	631
RH7057	733*	B	NS	F	Experienced	4800	374
	1,126*	B	NS	F	Experienced	3200	359
IS5366	1,085*	B	NS	F	Naive	1100	125
GE6106	1,133*	B	IDU	F	Naive	160000	431
UN2011	1,079*	B	IDU	F	Experienced	3800	458
OY9080	1,087*	B	Heterosexual	F	Naive	34000	308
DA1342	789*	B	NS	F	Naive	1505	359
	1,070*	B	NS	F	Naive	1900	367
SI1392	739*	R-A/G	NS	F	Experienced	1148	357
	1,028*	R-A/G	NS	F	Experienced	4300	227
	1,182*	R-A/G	NS	F	Experienced	6100	245
CC2874	1,103*	B	NS	F	Naive	11000	252
PR4290	879*	B	NS	F	Naive	1530	368
XE9655	844*	B	NS	F	Naive	36000	444
FQ2419	756*	B	IDU	F	Experienced	47000	327
	1,134*	B	IDU	F	Experienced	32000	301
NM1689	775*	B	IDU	F	Experienced	12000	646
	1,197*	B	IDU	F	Experienced	1600000	658
SS0342	1,292*	B	IDU	F	Experienced	8700	173
BQ7042	1,035*	B	IDU	F	Experienced	9400	630
EJ7211	1,044*	B	Heterosexual	F	Naive	33000	683

**Table S1. 297 Chronic specimens.** NS, not stated in original publication; Naive, subject has not received ART; Experienced, patient currently or previously received ART; MSM, men who have sex with men; IDU, intravenous drug user. \*Only HXB2 7134-7499 segment is available. #Genbank accession numbers

Subject (Reference)	Estimated Days Post Infection	Fiebig Stage	Subtype	Risk Behavior	Sex	ART status	Viral Load (RNA copies/ml)	CD4 Count (cells/mm <sup>3</sup> )
<u>Baalwa</u> <sup>30</sup> , <u>Yue</u> <sup>31</sup>								
JX877476-JX877521								
KP223729-KP223854								
KJ190253-KJ190262								
190049	19.5 [13, 34]	I/II	D	Heterosexual	F	Naïve	38,266,666	NS
191084	31 [27, 43]	IV	A	Heterosexual	F	Naïve	34,600	NS
191727	31 [27, 43]	IV	D	Heterosexual	F	Naïve	15,200	NS
191845	31 [27, 43]	IV	A	Heterosexual	F	Naïve	31,200	NS
191859	17 [13, 28]	I	D	Heterosexual	F	Naïve	53,334	NS
191882	17 [13, 28]	I	D	Heterosexual	M	Naïve	15,378	NS
191947	17 [13, 28]	I	R-A/D	Heterosexual	M	Naïve	5,466	NS
191982	17 [13, 28]	I	A/D	Heterosexual	F	Naïve	7,156	NS
<u>Chen</u> <sup>32</sup>								
KM217584-KM218333								
BJOX23	31 [27, 43]	IV	R-A/E	MSM	M	Naïve	4,380,000	NS
BJOX35	19.5 [13, 34]	I/II	B	MSM	M	Naïve	11,600	NS
BJOX19	19.5 [13, 34]	I/II	R-B/C	MSM	M	Naïve	330,000	NS
BJOX20	19.5 [13, 34]	I/II	B	MSM	M	Naïve	64,700	NS
BJOX15	19.5 [13, 34]	I/II	R-A/E	MSM	M	Naïve	1,180,000	NS
BJOX03	19.5 [13, 34]	I/II	B	MSM	M	Naïve	2,150,000	NS
BJOX22	31 [27, 43]	IV	B	MSM	M	Naïve	322,000	NS
BJOX05	19.5 [13, 34]	I/II	R-A/E	MSM	M	Naïve	30,211	NS
BJOX06	31 [27, 43]	IV	B	MSM	M	Naïve	498,947	NS
BJOX25	19.5 [13, 34]	I/II	R-A/E	MSM	M	Naïve	1,740,000	NS
BJOX09	31 [27, 43]	IV	R-A/E	MSM	M	Naïve	77,600	NS
BJOX18	31 [27, 43]	IV	R-A/E	MSM	M	Naïve	16,200	NS
BJOX28	19.5 [13, 34]	I/II	R-A/E	MSM	M	Naïve	70,200	NS
BJOX10	19.5 [13, 34]	I/II	R-A/E	MSM	M	Naïve	19,300	NS
BJOX12	101 [71, 154]	V	R-A/E	MSM	M	Naïve	3,330	NS
BJOX29	101 [71, 154]	V	R-B/C	MSM	M	Naïve	7,440	NS
BJOX07	101 [71, 154]	V	B	MSM	M	Naïve	1,558	NS
BJOX11	101 [71, 154]	V	R-B/C	MSM	M	Naïve	11,900	NS
BJOX31	101 [71, 154]	V	R-A/E	MSM	M	Naïve	14,700	NS
BJOX47	101 [71, 154]	V	B	MSM	M	Naïve	4,480	NS

BJOX14	101 [71, 154]	V	B	MSM	M	Naïve	2,480,000	NS
BJOX41	101 [71, 154]	V	B	MSM	M	Naïve	6,080	NS
BJOX46	101 [71, 154]	V	B	MSM	M	Naïve	18,700	NS
<u>Masharsky</u> <sup>33</sup>								
GU481107-GU481683								
R163	19.5 [13, 34]	I/II	A	IDU	F	Naïve	316,228	NS
R392	31 [27, 43]	IV	A	IDU	M	Naïve	501,187	NS
R497	31 [27, 43]	IV	A	IDU	M	Naïve	398,107	NS
R575	31 [27, 43]	IV	A	IDU	M	Naïve	630,957	NS
<u>Manak</u> <sup>34</sup>								
JN687677-JN687847								
BP00002	22 [18, 34]	II	C	NS	M	Naïve	553,000	NS
BP00055	22 [18, 34]	II	B	NS	M	Naïve	13,200,000	NS
BP00008	101 [71, 154]	V	C	NS	M	Naïve	471,000	NS
BP00010	101 [71, 154]	V	C	NS	M	Naïve	112,000	NS
BP00067	101 [71, 154]	V	B	NS	M	Naïve	199,000	NS
<u>Keele</u> <sup>9</sup>								
EU574937-EU579293								
1006-11	25 [22, 37]	III	B	NS	M	NS	1,600,000	NS
1054-07	22 [18, 34]	II	B	NS	M	NS	320,000	NS
6240-08	22 [18, 34]	II	B	NS	M	NS	120,000	NS
6244-13	22 [18, 34]	II	B	NS	M	NS	274,000	NS
9010-09	22 [18, 34]	II	B	NS	F	NS	146,954	NS
9014-01	22 [18, 34]	II	B	NS	M	NS	25,680	NS
9015-07	22 [18, 34]	II	B	NS	M	NS	500,000	NS
9017-01	22 [18, 34]	II	B	NS	M	NS	32,390	NS
9020-20	22 [18, 34]	II	B	NS	F	NS	69,238	NS
9021-14	22 [18, 34]	II	B	NS	M	NS	143,379	NS
9023-21	22 [18, 34]	II	B	NS	M	NS	>500,000	NS
9024-12	22 [18, 34]	II	B	NS	NS	NS	>500,000	NS
9025-11	22 [18, 34]	II	B	NS	M	NS	>500,000	NS
9028-07	22 [18, 34]	II	B	NS	F	NS	>500,000	NS
9032-08	25 [22, 37]	III	B	NS	M	NS	40,815	NS
9033-16	25 [22, 37]	III	B	NS	M	NS	454,553	NS
9075-03	22 [18, 34]	II	B	NS	NS	NS	360,000	NS
9077-12	22 [18, 34]	II	B	NS	NS	NS	158,300	NS
9079-09	22 [18, 34]	II	B	NS	NS	NS	148,320	NS

61792-03	22 [18, 34]	II	B	NS	NS	NS	3,700,000	NS
62130-04	22 [18, 34]	II	B	NS	NS	NS	600,000	NS
62357-14	22 [18, 34]	II	B	NS	M	NS	302,000	NS
62995-05	17 [13, 28]	I	B	NS	NS	NS	75,000	NS
63054-04	22 [18, 34]	II	B	NS	NS	NS	1,153,260	NS
63396-07	22 [18, 34]	II	B	NS	NS	NS	420,000	NS
PRB926-04	22 [18, 34]	II	B	NS	NS	NS	756,060	NS
PRB931-06	25 [22, 37]	III	B	NS	NS	NS	2,000,000	NS
PRB956-04	22 [18, 34]	II	B	NS	NS	NS	600,000	NS
PRB958-06	25 [22, 37]	III	B	NS	NS	NS	678,900	NS
PRB959-02	22 [18, 34]	II	B	NS	NS	NS	>2,000,000	NS
SC05	22 [18, 34]	II	B	Heterosexual	M	NS	9,980,952	NS
SC11	22 [18, 34]	II	B	Heterosexual	F	NS	1,917,073	NS
SC20	31 [27, 43]	IV	B	Heterosexual	M	NS	2,789,313	NS
SC45	22 [18, 34]	II	B	Heterosexual	M	NS	6,318,529	NS
TRJO4551	22 [18, 34]	II	B	MSM	M	NS	8,121,951	NS
TT29P	22 [18, 34]	II	B	Heterosexual	M	NS	14,756,436	NS
WITO4160	22 [18, 34]	II	B	Heterosexual	M	NS	325,064	NS
Z05	22 [18, 34]	II	B	Heterosexual	M	NS	3,950,942	NS
Z20	25 [22, 37]	III	B	MSM	M	NS	13,466,851	NS
Z32	31 [27, 43]	IV	B	MSM	M	NS	5,788,591	NS
Z34	25 [22, 37]	III	B	Bisexual	M	NS	2,501,242	NS
1018-10	25 [22, 37]	III	B	NS	M	NS	270,000	NS
1053-07	25 [22, 37]	III	B	NS	M	NS	1,400,000	NS
6248-07	25 [22, 37]	III	B	NS	M	NS	87,700,000	NS
1001-07	25 [22, 37]	III	B	NS	F	NS	210,000	NS
9022-09	25 [22, 37]	III	B	NS	M	NS	>500,000	NS
TT35P	22 [18, 34]	II	B	Heterosexual	M	NS	1,849,301	NS
63358-04	22 [18, 34]	II	B	NS	NS	NS	260,000	NS
9031-19	31 [27, 43]	IV	B	NS	M	NS	237,938	NS
9029-12	22 [18, 34]	II	B	NS	M	NS	>500,000	NS
12007-04	22 [18, 34]	II	B	NS	NS	NS	75,815	NS
1012-11	25 [22, 37]	III	B	NS	M	NS	410,000	NS
SC22	22 [18, 34]	II	B	Heterosexual	F	NS	14,304,820	NS
9030-15	22 [18, 34]	II	B	NS	M	NS	>500,000	NS
SC31	31 [27, 43]	IV	B	Heterosexual	F	NS	4,449,612	NS
700010058	25 [22, 37]	III	B	NS	M	NS	394,649	NS

Z33	22 [18, 34]	II	B	Bisexual	M	NS	158,603	NS
Z31	22 [18, 34]	II	B	MSM	M	NS	6,507,075	NS
63068-05	22 [18, 34]	II	B	NS	NS	NS	100,000	NS
TT27P	31 [27, 43]	IV	B	Heterosexual	M	NS	1,489,202	NS
62615-03	22 [18, 34]	II	B	NS	NS	NS	2,200,000	NS
Z35	31 [27, 43]	IV	B	MSM	M	NS	2,281,415	NS
9026-07	25 [22, 37]	III	B	NS	NS	NS	211,703	NS
9076-08	25 [22, 37]	III	B	NS	NS	NS	410,000	NS
SC33	22 [18, 34]	II	B	Heterosexual	M	NS	3,277,559	NS
BORI0637	22 [18, 34]	II	B	MSM	M	NS	2,400,000	NS
PRB957-06	22 [18, 34]	II	B	NS	NS	NS	>2,000,000	NS
12008-09	22 [18, 34]	II	B	NS	NS	NS	>2,000,000	NS
SC42	31 [27, 43]	IV	B	Heterosexual	M	NS	1,380,000	NS
9019-03	101 [71, 154]	V	B	NS	F	NS	571,000	NS
REJO4541	101 [71, 154]	V	B	Heterosexual	M	NS	722,349	NS
THRO4156	101 [71, 154]	V	B	MSM	M	NS	5,413,140	NS
TT28P	101 [71, 154]	V	B	Heterosexual	M	NS	7,468,346	NS
TT34P	101 [71, 154]	V	B	Heterosexual	M	NS	3,050,125	NS
Z02	101 [71, 154]	V	B	Heterosexual	M	NS	296,111	NS
Z13	101 [71, 154]	V	B	Heterosexual	M	NS	53,941	NS
Z23	101 [71, 154]	V	B	Bisexual	M	NS	210,521	NS
Z27	101 [71, 154]	V	B	Heterosexual	M	NS	178,125	NS
RHPA4259	101 [71, 154]	V	B	Heterosexual	F	NS	1,458,354	NS
63215-03	101 [71, 154]	V	B	NS	NS	NS	550,000	NS
SC51	101 [71, 154]	V	B	Heterosexual	M	Naïve	4,229,703	NS
MEMI4948	101 [71, 154]	V	B	MSM	M	NS	22,686	NS
700010019	101 [71, 154]	V	B	IDU	F	NS	741,499	NS
Z18	101 [71, 154]	V	B	MSM	M	NS	2,845,823	NS
CAAN5342	101 [71, 154]	V	B	MSM	M	NS	608,560	NS
Z30	101 [71, 154]	V	B	Heterosexual	F	NS	1,408,114	NS
Z16	101 [71, 154]	V	B	MSM	M	NS	2,375,617	NS
Z03	101 [71, 154]	V	B	MSM	M	NS	726,859	NS
701010016	101 [71, 154]	V	B	NS	M	NS	1,802	NS
Z29	101 [71, 154]	V	B	Bisexual	M	NS	507,401	NS
<u>Abrahams</u> <sup>35</sup>								
FJ443128-FJ444362								
0114	31 [27, 43]	IV	C	Heterosexual	M	NS	99,000,000	NS

0334	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	22,000,000	NS
0393	31 [27, 43]	IV	C	Heterosexual	F	NS	12,000,000	NS
0478	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	2,166,667	NS
0595	31 [27, 43]	IV	C	Heterosexual	F	NS	1,006,864	NS
0626	31 [27, 43]	IV	C	Heterosexual	M	NS	711,061	NS
0665	31 [27, 43]	IV	C	Heterosexual	M	NS	67,728	NS
0682	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	17,960	NS
0985	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	2,894,737	NS
1086	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	1,409,208	NS
1172	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	4,805,921	NS
1176	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	200,000,000	NS
1196	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	30,500,000	NS
1335	31 [27, 43]	IV	C	Heterosexual	M	NS	576,282	NS
1373	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	3,221,979	NS
1394	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	7,341,772	NS
2010	31 [27, 43]	IV	C	Heterosexual	M	NS	1,203,333	NS
2052	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	339,475	NS
2060	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	1,530,364	NS
2103	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	152,364	NS
703010010	25 [22, 37]	III	C	Heterosexual	M	NS	408,727	NS
703010131	25 [22, 37]	III	C	Heterosexual	M	NS	437,369	NS
703010159	22 [18, 34]	II	C	Heterosexual	F	NS	73,453	NS
703010228	31 [27, 43]	IV	C	Heterosexual	M	NS	335,000	NS
704010083	25 [22, 37]	III	C	Heterosexual	F	NS	>750,000	NS
704809221	19.5 [13, 34]	I/II	C	Heterosexual	M	NS	>750,000	NS
CAP37	31 [27, 43]	IV	C	Heterosexual	F	NS	248,000	NS
CAP63	25 [22, 37]	III	C	Heterosexual	F	NS	3,210,000	NS
CAP69	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	15,300	NS
CAP129	31 [27, 43]	IV	C	Heterosexual	F	NS	1,800,000	NS
CAP177	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	359,000	NS
CAP188	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	13,800,000	NS
CAP200	31 [27, 43]	IV	C	Heterosexual	F	NS	398,000	NS
CAP217	31 [27, 43]	IV	C	Heterosexual	F	NS	3,260,000	NS
CAP221	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	24,300	NS
CAP222	19.5 [13, 34]	I/II	C	Heterosexual	F	NS	69,700	NS
CAP225	25 [22, 37]	III	C	Heterosexual	F	NS	63,600,000	NS
CAP237	25 [22, 37]	III	C	Heterosexual	F	NS	1,020,000	NS



089	101 [71, 154]	V	C	Heterosexual	M	NS	83,844	NS
703010054	101 [71, 154]	V	C	Heterosexual	M	NS	13,936	NS
703010193	101 [71, 154]	V	C	Heterosexual	M	NS	46,724	NS
704010042	101 [71, 154]	V	C	Heterosexual	M	NS	133,000	NS
704810053	101 [71, 154]	V	C	Heterosexual	F	NS	25,400	NS
705010015	101 [71, 154]	V	C	Heterosexual	F	NS	47,900	NS
705010026	101 [71, 154]	V	C	Heterosexual	F	NS	8,971	NS
705010078	101 [71, 154]	V	C	Heterosexual	M	NS	255,907	NS
CAP8	101 [71, 154]	V	C	Heterosexual	F	NS	207,000	NS
CAP84	101 [71, 154]	V	C	Heterosexual	F	NS	559,000	NS
CAP85	101 [71, 154]	V	C	Heterosexual	F	NS	621,000	NS
CAP136	101 [71, 154]	V	C	Heterosexual	F	NS	85,300	NS
CAP174	101 [71, 154]	V	C	Heterosexual	F	NS	40,000	NS
CAP206	101 [71, 154]	V	C	Heterosexual	F	NS	196,000	NS
CAP220	101 [71, 154]	V	C	Heterosexual	F	NS	2,070	NS
CAP224	101 [71, 154]	V	C	Heterosexual	F	NS	1,348,000	NS
CAP260	101 [71, 154]	V	C	Heterosexual	F	NS	17,600	NS
<u>Bar</u> <sup>36</sup>								
GU562001-GU562291								
GU938194-GU938295								
HDNDRPI029	31 [27, 43]	IV	B	IDU	F	NS	21,878	440
HDNDRPI034	25 [22, 37]	III	B	IDU	M	NS	75,857,758	240
CQLDR03	101 [71, 154]	V	B	IDU	M	NS	102,329	NS
HDNDRPI001	101 [71, 154]	V	B	IDU	M	NS	870,964	690
HDNDRPI032	101 [71, 154]	V	B	IDU	M	NS	3,388	1,040
HTM319	101 [71, 154]	V	B	IDU	M	NS	26,915	520
HTM385	101 [71, 154]	V	B	IDU	M	NS	234,423	406
<u>Heipertz</u> <sup>37</sup>								
JX863919-JX864047								
306159	19.5 [13, 34]	I/II	B	NS	NS	Naïve	NS	NS
306344	19.5 [13, 34]	I/II	B	NS	NS	Naïve	NS	NS
306376	19.5 [13, 34]	I/II	B	NS	NS	Naïve	NS	NS
306512	19.5 [13, 34]	I/II	B	NS	NS	Naïve	NS	NS
306517	19.5 [13, 34]	I/II	B	NS	NS	Naïve	NS	NS
<u>Li</u> <sup>38</sup>								
GU330247-GU331770								
INME0632	22 [18, 34]	II	B	MSM	M	Naïve	2,217,670	739

HOBRO961	22 [18, 34]	II	B	MSM	M	Naïve	599,238	794
4013171	31 [27, 43]	IV	B	MSM	M	Naïve	3,700,000	213
4013211	25 [22, 37]	III	B	MSM	M	Naïve	19,900,000	846
4013226	22 [18, 34]	II	B	MSM	M	Naïve	26,700,000	175
4013240	22 [18, 34]	II	B	MSM	M	Naïve	2,240,000	297
4013242	31 [27, 43]	IV	B	MSM	M	Naïve	5,790,000	251
4013291	101 [71, 154]	V	B	MSM	M	Naïve	1,490,000	179
4013296	22 [18, 34]	II	B	MSM	M	Naïve	8,050,000	395
4013321	22 [18, 34]	II	B	MSM	M	Naïve	6,250,000	407
4013327	31 [27, 43]	IV	B	MSM	M	Naïve	8,720,000	248
4013383	22 [18, 34]	II	B	MSM	M	Naïve	584,000	531
4013396	31 [27, 43]	IV	B	MSM	M	Naïve	1,600,000	581
4013419	22 [18, 34]	II	B	MSM	M	Naïve	21,200,000	226
4013440	22 [18, 34]	II	B	MSM	M	Naïve	>100,000	205
4013446	25 [22, 37]	III	B	MSM	M	Naïve	>100,000	438
4013448	22 [18, 34]	II	B	MSM	M	Naïve	28,600,000	536
701010055	22 [18, 34]	II	B	MSM	M	Naïve	31,513,812	432
701010068	31 [27, 43]	IV	B	MSM	M	Naïve	3,714,386	109
AD17	22 [18, 34]	II	B	MSM	M	Naïve	47,600,000	NS
AD75	22 [18, 34]	II	B	MSM	M	Naïve	21,400,000	NS
AD83	101 [71, 154]	V	B	MSM	M	Naïve	448,000	NS
700010106	22 [18, 34]	II	B	MSM	M	Naïve	84,545,454	277
701010027	101 [71, 154]	V	B	MSM	M	Naïve	194,744	542
700010246	31 [27, 43]	IV	B	MSM	M	Naïve	4,395,721	1,012
AD77	101 [71, 154]	V	B	MSM	M	Naïve	130,000	NS
701010108	101 [71, 154]	V	B	MSM	M	Naïve	14,711	592
700010238	101 [71, 154]	V	B	MSM	M	Naïve	596,908	587
<u>Parrish</u> <sup>39, 40</sup>								
JQ753730-JQ754304								
JX972238-JX973270								
KC149138-KC156401								
705010067	19.5 [13, 34]	I/II	C	Heterosexual	F	Naïve	639,000	NS
20258279	31 [27, 43]	IV	C	NS	F	Naïve	281,838	NS
1245045	19.5 [13, 34]	I/II	C	NS	M	Naïve	234,068	NS
2833264	19.5 [13, 34]	I/II	C	NS	M	Naïve	234,423	NS
2935054	19.5 [13, 34]	I/II	C	NS	M	Naïve	>10,000,000	NS

19157834	19.5 [13, 34]	I/II	C	NS	M	Naïve	275,423	NS
20927783	19.5 [13, 34]	I/II	C	NS	F	Naïve	1,886	NS
21197826	19.5 [13, 34]	I/II	C	NS	F	Naïve	343,923	NS
21283649	19.5 [13, 34]	I/II	C	NS	M	Naïve	3,180	NS
<u>Gnanakaran</u> <sup>8</sup>								
HQ216367-HQ218052								
HQ238279-HQ238288								
701010043	22 [18, 34]	II	B	NS	F	Naïve	5,532	NS
Z64	31 [27, 43]	IV	B	Heterosexual	M	Naïve	11,503,872	6
SC24	23.5 [18, 37]	II/III	B	Heterosexual	M	Naïve	523,183	NS
SC50	23.5 [18, 37]	II/III	B	Heterosexual	M	Naïve	259,747	NS
FASH1057	22 [18, 34]	II	B	Heterosexual	F	Naïve	21,783,636	NS
700010224	101 [71, 154]	V	B	NS	M	Naïve	300,000	NS
701010092	101 [71, 154]	V	B	NS	M	Naïve	207,143	NS
Z75	101 [71, 154]	V	B	MSM	M	Naïve	170,826	241
Z78	101 [71, 154]	V	B	MSM	M	Naïve	750,001	NS
Z91	101 [71, 154]	V	B	MSM	M	Naïve	37,706	541
Z92	101 [71, 154]	V	B	MSM	M	Naïve	7,633,876	NS
Z95	101 [71, 154]	V	B	Heterosexual	M	Naïve	59,603	808
<u>Nofemela</u> <sup>41</sup>								
HQ659584-HQ698053								
98	101 [71, 154]	V	C	Heterosexual	F	Naïve	5,330	NS
304	101 [71, 154]	V	C	Heterosexual	F	Naïve	>750,001	NS
556	101 [71, 154]	V	C	Heterosexual	F	Naïve	43,300	NS
49	25 [22, 37]	III	R- A/C/D	Heterosexual	F	Naïve	>750,001	NS
532	19.5 [13, 34]	I/II	R-A/D	Heterosexual	F	Naïve	>750,001	NS
390	101 [71, 154]	V	C	Heterosexual	F	Naïve	>750,001	NS
<u>Park</u> <sup>29</sup>								
MF537774-MF538517								
TY2947	[81 – 251]*	NS	B	NS	F	Naïve	920	967
RH7057	[0 - 146]*	NS	B	NS	F	Naïve	14000	700
	[188 – 334]*	NS	B	NS	F	Naïve	15000	252
	[384 – 530]*	NS	B	NS	F	Experienc ed	1592	328
GE6106	[0 – 183]*	NS	B	IDU	F	Naïve	NS	804
	[182 – 365]*	NS	B	IDU	F	Naïve	230	510

	[350 – 533]*	NS	B	IDU	F	Naive	190000	576
OY9080	[0 – 187]*	NS	B	Heterosexual	F	Naive	NS	NS
	[183 – 370]*	NS	B	Heterosexual	F	Naive	50000	585
	[357 – 544]*	NS	B	Heterosexual	F	Naive	45000	573
DA1342	[0 – 176]*	NS	B	NS	F	Naive	12000	584
	[160 – 336]*	NS	B	NS	F	Naive	20000	539
	[417 – 593]*	NS	B	NS	F	Naive	10000	416
SI1392	[0 – 134]*	NS	R-A/G	NS	F	Naive	1280	760
	[431 – 565]*	NS	R-A/G	NS	F	Naive	37000	232
JG8930	[0 – 237]*	NS	B	NS	F	Naive	11000	908
JY3080	[0 – 183]*	NS	B	NS	F	Naive	NS	NS
	[182 – 365]*	NS	B	NS	F	Naive	101925	940
XE9655	[0 - 176]*	NS	B	NS	F	Naive	NS	NS
DV6934	[0 - 245]*	NS	B	IDU	F	Naive	NS	377
NM1689	[0 – 173]*	NS	B	IDU	F	Naive	39000	836
SS0342	[0 - 177]*	NS	B	IDU	F	Naive	210000	352
	[160 – 337]*	NS	B	IDU	F	Naive	15000	487
	[354 – 531]*	NS	B	IDU	F	Experienced	3500	614
VE6375	[0 – 127]*	NS	B	IDU	F	Naive	16000	403
	[414 - 541]*	NS	B	IDU	F	Naive	140	260
EJ7211	[0 – 188]*	NS	B	Heterosexual	F	Naive	NI	909
	[358 – 546]*	NS	B	Heterosexual	F	Naive	2600	1220
TI9054	[0 – 176]*	NS	B	NS	F	Naive	5800	795
	[166 – 342]*	NS	B	NS	F	Naive	19000	493
	[349 – 525]*	NS	B	NS	F	Naive	17453	427

**Table S2. Incident Specimens.** 252 Incident specimens with Fiebig stages I, II, III, IV, and V and 31 incident specimens with documented infection of less than 1 year. \*Only HXB2 7134-7499 segment is available. \*The minimum infection duration was the time between the first HIV positive documentation and specimen collection, and the maximum infection duration was the time between the last HIV negative test and specimen collection.

Subject	Estimated Days Post Infection	Days from First Sample	Fiebig Stage	Subtype	Risk Behaviour	Sex	ART status	Viral Load (RNA copies/ml)	CD4 T cell Count (cells /mm <sup>3</sup> )
<u>Baalwa<sup>30</sup>, Yue<sup>31</sup></u>									
JX877476-JX877521									
KP223729-KP223854									
KJ190253-KJ190262									
191647	22 [18, 34]	0	II	D	Heterosexual	M	Naïve	242,222	NS
	149 [145, 161]	127	-	D	Heterosexual	M	Naïve	4,470	518
R463F	31 [27, 43]	0	IV	A	Heterosexual	F	Naïve	150,000,000	NS
	38 [34, 50]	7	-	A	Heterosexual	F	Naïve	3,980,000	324
	79 [75, 91]	48	-	A	Heterosexual	F	Naïve	332,000	427
	198 [194, 210]	167	-	A	Heterosexual	F	Naïve	398,000	368
	267 [263, 279]	236	-	A	Heterosexual	F	Naïve	246,000	344
	352 [348, 364]	321	-	A	Heterosexual	F	Naïve	167,054	375
R880F	25 [22, 37]	0	III	A	Heterosexual	F	Naïve	430,843	NS
	88 [85, 100]	63	-	A	Heterosexual	F	Naïve	2,410	483
	172 [169, 184]	147	-	A	Heterosexual	F	Naïve	49	742
	256 [253, 268]	231	-	A	Heterosexual	F	Naïve	49	NS
	338 [335, 350]	313	-	A	Heterosexual	F	Naïve	62	704
9004SS	31 [27, 43]	0	IV	A	Heterosexual	M	Naïve	29,921	NS
	60 [56, 72]	29	-	A	Heterosexual	M	Naïve	12,323	1,192
<u>Herbeck<sup>10</sup>, Liu<sup>11</sup></u>									
DQ853426-DQ854622									
PIC11286	101 [71, 154]	0	V	B	MSM	M	Naïve	2,804,000	NS
	106 [76,159]	5	-	B	MSM	M	Naïve	54,160	NS
	114 [84,167]	13	-	B	MSM	M	Naïve	29,180	NS
PIC38051	17 [13,28]	0	I	B	MSM	M	Naïve	620,820	NS
	21 [17,32]	4	-	B	MSM	M	Naïve	2,498,840	NS
PIC38417	101 [71,154]	0	V	B	MSM	M	Naïve	472,810	NS
	109 [79,162]	8	-	B	MSM	M	Naïve	169,580	NS
	116 [86,169]	15	-	B	MSM	M	Naïve	117,310	NS
	129 [99,182]	28	-	B	MSM	M	Naïve	71,230	NS
	159 [129,212]	58	-	B	MSM	M	Naïve	30,650	NS
	192 [162,245]	91	-	B	MSM	M	Naïve	32,430	NS
	220 [190, 273]	119	-	B	MSM	M	Naïve	24,740	NS

	276 [246,329]	175	-	B	MSM	M	Naïve	19,460	NS
	323 [293,376]	222	-	B	MSM	M	Naïve	14,610	NS
PIC55751	101 [71,154]	0	V	B	MSM	M	Naïve	128,814	NS
	115 [85, 168]	14	-	B	MSM	M	Naïve	8,469	NS
	3,426 [3396, 3479]	3325	-	B	MSM	M	Naïve	40,600	NS
	3,477 [3447, 3530]	3376	-	B	MSM	M	Naïve	22,190	NS
PIC71101	17 [13, 28]	0	I	B	MSM	M	Naïve	377,060	NS
	23 [19, 34]	6	-	B	MSM	M	Naïve	1,912,200	NS
	31 [27, 42]	14	-	B	MSM	M	Naïve	312,200	NS
	54 [50, 65]	37	-	B	MSM	M	Naïve	32,630	NS
	78 [74, 89]	61	-	B	MSM	M	Naïve	232,050	NS
	106 [102, 117]	89	-	B	MSM	M	Naïve	109,950	NS
	137 [133, 148]	120	-	B	MSM	M	Naïve	45,590	NS
	191 [187, 202]	174	-	B	MSM	M	Naïve	105,870	NS
PIC83747	17 [13, 28]	0	I	B	MSM	M	Naïve	923	NS
	23 [19, 34]	6	-	B	MSM	M	Naïve	1,049,870	NS
	31 [27, 42]	14	-	B	MSM	M	Naïve	26,785,000	NS
	53 [49, 64]	36	-	B	MSM	M	Naïve	665,680	NS
	59 [55, 70]	42	-	B	MSM	M	Naïve	668,660	NS
	67 [63, 79]	50	-	B	MSM	M	Naïve	124,000	NS
	80 [76, 91]	63	-	B	MSM	M	Naïve	254,300	NS
	109 [105, 120]	92	-	B	MSM	M	Naïve	151,000	NS
	162 [159, 173]	145	-	B	MSM	M	Naïve	242,940	NS
	221 [217, 232]	204	-	B	MSM	M	Naïve	60,950	NS
	284 [280, 295]	267	-	B	MSM	M	Naïve	51,120	NS
	361 [357, 372]	344	-	B	MSM	M	Naïve	49,600	NS
PIC90770	17 [13, 28]	0	I	B	MSM	M	Naïve	469,830	NS
	25 [21, 36]	8	-	B	MSM	M	Naïve	1600000	NS
	29 [25, 40]	12	-	B	MSM	M	Naïve	42,810	NS
	85 [81, 96]	68	-	B	MSM	M	Naïve	12,200	NS
	113 [109, 124]	96	-	B	MSM	M	Naïve	14,170	NS
	142 [138, 153]	125	-	B	MSM	M	Naïve	18,210	NS
	210 [206, 221]	193	-	B	MSM	M	Naïve	70,780	NS
PIC1362	17 [13, 28]	0	I	B	MSM	M	Naïve	8,240,000	876
	31 [27, 42]*	14	-	B	MSM	M	Naïve	NS	NS

	59 [55, 70]*	42	-	B	MSM	M	Naïve	NS	NS
	122 [118, 133]*	105	-	B	MSM	M	Naïve	NS	NS
	164 [160, 175]*	147	-	B	MSM	M	Naïve	NS	NS
	199 [195, 210]*	182	-	B	MSM	M	Naïve	NS	NS
	353 [349, 364]*	336	-	B	MSM	M	Naïve	NS	NS
	590 [586, 601]*	573	-	B	MSM	M	Naïve	NS	NS
	778 [774, 789]*	761	-	B	MSM	M	Naïve	NS	NS
	835 [831, 846]	818	-	B	MSM	M	Naïve	NS	NS
	1,044 [1040, 1055]*	1,027	-	B	MSM	M	Naïve	NS	NS
	1,256 [1252, 1267]	1,239	-	B	MSM	M	Naïve	NS	NS
	1,500 [1496, 1511]*	1,483	-	B	MSM	M	Naïve	NS	NS
<u>Parrish</u> <sup>40</sup>									
KC149138-KC149185									
703010200	19.5 [13, 34]	0	I/II	C	Heterosexual	M	NS	165,501	NS
	26.5 [20, 41]	7	IV	C	Heterosexual	M	NS	128,677	111
	417.5 [411, 432]	398	-	C	Heterosexual	M	NS	27,882	NS
	704.5 [698, 719]	685	-	C	Heterosexual	M	NS	18,300	NS
<u>Liu</u> <sup>11</sup> , <u>Keele</u> <sup>9</sup>									
<u>Salazar-Gonzalez</u> <sup>42</sup>									
<u>Abrahams</u> <sup>35</sup>									
DQ853426-DQ854622									
EU574937-EU579293									
FJ495804-FJ496214									
FJ919955-FJ919967									
EU578998-EU576521									
FJ443128-FJ444362									
CAP045	19.5 [13, 34]	0	I/II	C	Heterosexual	F	Naive	12,500	NS
	40.5 [34, 55]	21	-	C	Heterosexual	F	Naive	236,000	974
	369.5 [363, 384]	350	-	C	Heterosexual	F	Naive	556	1,030
CAP210	19.5 [13, 34]	0	I/II	C	Heterosexual	F	Naive	468,000	332
	89.5 [83,104]	70	-	C	Heterosexual	F	Naive	50,100	543
	159.5 [153,174]	140	-	C	Heterosexual	F	Naive	NS	NS
	187.5 [181, 202]	168	-	C	Heterosexual	F	Naive	20,400	475
CH040	22 [18, 34]	0	II	B	MSM	M	Naive	2,197,248	NS

	67 [63, 79]	45	-	B	MSM	M	Naive	298,026	NS
	133 [129, 145]	111	-	B	MSM	M	Naive	NS	NS
	203 [199, 215]	181	-	B	MSM	M	Naive	NS	NS
	434 [430, 446]	412	-	B	MSM	M	Naive	NS	NS
CH042	31 [27, 43]	0	IV	C	NS	M	Naive	181,000	NS
	52 [48, 64]	21	-	C	NS	M	Naive	201,000	318
	91 [87, 103]	60	-	C	NS	M	Naive	128,000	350
	203 [199, 215]	172	-	C	NS	M	Naive	102,000	357
	455 [451, 467]	424	-	C	NS	M	Naive	86,400	295
	707 [703, 719]	676	-	C	NS	M	Naive	267,000	349
CH058	22 [18, 34]	0	II	B	MSM	M	Naive	92,581	NS
	58 [54, 70]	36	-	B	MSM	M	Naive	394,649	NS
	98 [94, 110]	76	-	B	MSM	M	Naive	NS	NS
CH077	19.5 [13, 34]	0	I/II	B	MSM	M	Naive	NS	NS
	33.5 [27, 48]	14	-	B	MSM	M	Naive	NS	NS
	51.5 [45, 66]	32	-	B	MSM	M	Naive	NS	NS
	178.5 [172, 193]	159	-	B	MSM	M	Naive	NS	NS
CH131	19.5 [13, 34]	0	I/II	C	NS	M	Naive	411,873	NS
	40.5 [34, 55]	21	-	C	NS	M	Naive	7,764	NS
	47.5 [41, 62]	28	-	C	NS	M	Naive	10,655	NS
	53.5 [47, 68]	34	-	C	NS	M	Naive	53,169	NS
	82.5 [76, 97]	63	-	C	NS	M	Naive	45,697	NS
	110.5 [104, 125]	91	-	C	NS	M	Naive	15,424	NS
	194.5 [188, 209]	175	-	C	NS	M	Naive	50,709	310
	292.5 [286, 307]	273	-	C	NS	M	Naive	27,446	300
	352.5 [346, 367]	333	-	C	NS	M	Naive	44,488	298
	689.5 [683, 704]	670	-	C	NS	M	Naive	17,807	214
CH159	19.5 [13, 34]	0	I/II	C	Heterosexual	F	Naive	17,805	NS
	27.5 [21, 42]	8		C	Heterosexual	F	Naive	587,457	NS
	31.5 [25, 46]	12		C	Heterosexual	F	Naive	570,786	NS
	41.5 [35, 56]	22	-	C	Heterosexual	F	Naive	7,101	NS
	48.5 [42, 63]	29	-	C	Heterosexual	F	Naive	13,566	NS
	75.5 [69, 90]	56	-	C	Heterosexual	F	Naive	9,397	NS
	104.5 [98, 119]	85	-	C	Heterosexual	F	Naive	71,474	463
	321.5 [315, 336]	302	-	C	Heterosexual	F	Naive	3,722	448
	667.5 [661, 682]	648	-	C	Heterosexual	F	Naive	<400	432
	1151.5 [1144.5,	1132	-	C	Heterosexual	F	Naive	10,161	347



	1165]								
CH162	25 [22, 37]	0	III	C	NS	M	Naive	>10,000,000	NS
	46 [43, 58]	21	-	C	NS	M	Naive	18,260	NS
	102 [99, 114]	77	-	C	NS	M	Naive	3,625	NS
	204 [201, 216]	179	-	C	NS	M	Naive	102,793	533
	463 [460, 475]	438	-	C	NS	M	Naive	295,950	204
CH164	25 [22, 37]	0	III	C	NS	M	Naive	23,600	700
	39 [36, 51]	14	-	C	NS	M	Naive	NS	NS
	81 [78, 93]	56	-	C	NS	M	Naive	7,720	NS
	194 [191, 206]	169	-	C	NS	M	Naive	12,700	449
	445 [442, 456]	420	-	C	NS	M	Naive	750,000	NS
CH185	19.5 [13, 34]	0	I/II	C	Heterosexual	F	Naive	20,449	NS
	44.5 [38, 59]	25	-	C	Heterosexual	F	Naive	85,443	NS
	86.5 [80, 101]	67	-	C	Heterosexual	F	Naive	79,706	NS
	199.5 [193, 214]	180	-	C	Heterosexual	F	Naive	46,772	NS
	435.5 [429, 450]	416	-	C	Heterosexual	F	Naive	234,010	311
CH198	19.5 [13, 34]	0	I/II	C	Heterosexual	M	Naive	14,950,000	NS
	30.5 [24, 45]	11	-	C	Heterosexual	M	Naive	301,401	599
	79.5 [73, 94]	60	-	C	Heterosexual	M	Naive	2,433	NS
	466.5 [460, 481]	447	-	C	Heterosexual	M	Naive	3,476	NS
CH256	19.5 [13, 34]	0	I/II	C	Heterosexual	F	Naive	62,060	NS
	47.5 [41, 62]	28	-	C	Heterosexual	F	Naive	46,342	NS
	82.5 [76, 97]	63	-	C	Heterosexual	F	Naive	35,328	NS
	191.5 [185, 206]	172	-	C	Heterosexual	F	Naive	13,484	NS
	445.5 [439, 460]	426	-	C	Heterosexual	F	Naive	6,183	601
	703.5 [697, 718]	684	-	C	Heterosexual	F	Naive	28,066	531
CH470	31 [27, 43]	0	IV	B	NS	M	Naive	84,193	NS
	44 [40, 56]	13	-	B	NS	M	Naive	264,882	324
	72 [68, 84]	41	-	B	NS	M	Naive	27,991	522
	100 [96, 112]	69	-	B	NS	M	Naive	27,869	NS
	205 [201, 217]	174	-	B	NS	M	Naive	28,456	554
CAP239	101 [71, 154]	0	V	C	Heterosexual	F	Naive	95,800	845
	220 [190, 273]	119	-	C	Heterosexual	F	Naive	22,400	984
CH607	31 [27, 43]	0	IV	B	MSM	M	Experienced	8,009	NS
	45 [41, 57]	14		B	MSM	M	Experienced	NS	NS

	52 [48, 64]	21		B	MSM	M	Experienced	24,545	NS
<u>Masharsky</u> <sup>33</sup>									
GU481107-GU481683									
K84	31 [27, 43]	0	IV	A	IDU	M	NS	251,189	NS
	55 [51, 55]	24	V	A	IDU	M	NS	NS	NS
R053	31 [27, 43]	0	IV	A	IDU	F	NS	398,107	NS
	39 [35, 51]	8	V	A	IDU	F	NS	NS	NS
R526	19.5 [13, 34]	0	I/II	A	IDU	M	NS	501,187	NS
	84.5 [78, 99]	65	-	A	IDU	M	NS	NS	NS
H386	31 [27, 43]	0	IV	A	IDU	F	NS	630,957	NS
	40 [36, 52]	9	-	A	IDU	F	NS	NS	NS
H408	25 [22, 37]	0	III	A	IDU	M	NS	NS	NS
	38 [35, 50]	13	IV	A	IDU	M	NS	251,189	NS
H410	25 [22, 37]	0	III	A	IDU	M	NS	630,957	NS
	38 [35, 50]	13	IV	A	IDU	M	NS	NS	NS
<u>Keele</u> <sup>9</sup>									
EU574937-EU579293									
SUMA0874	22 [18, 34]	0	II	B	MSM	M	Naive	NS	NS
	23 [19, 35]	1	II	B	MSM	M	Naive	939,260	NS
	31 [27, 43]	9	III	B	MSM	M	Naive	701,640	NS
	38 [34, 50]	16	IV	B	MSM	M	Naive	106,543	NS
	52 [48, 64]	30	V	B	MSM	M	Naive	14,780	NS
WEAU0575	22 [18, 34]	0	II	B	MSM	M	Naive	216,415	NS
	30 [26, 42]	8	IV	B	MSM	M	Naive	355,360	NS
	37 [33, 49]	15	IV	B	MSM	M	Naive	100,877	NS
	51 [47, 63]	29	V	B	MSM	M	Naive	34,737	NS
	219 [215, 231]	197	VI	B	MSM	M	Naive	90,109	NS
1059	22 [18, 34]	0	II	B	NS	M	Naive	82,204	NS
	26 [22, 38]	4	II	B	NS	M	Naive	> 100,000	NS
	29 [25, 41]	7	III	B	NS	M	Naive	2,800,000	NS
1058	17 [13, 28]	0	I	B	NS	M	Naive	2,737	NS
	20 [16, 31]	3	II	B	NS	M	Naive	26,162	NS
	27 [23, 38]	10	IV	B	NS	M	Naive	550,000	NS
6247	17 [13, 28]	0	I	B	NS	F	NS	23,858	NS
	19 [15, 30]	2	II	B	NS	F	NS	124,000	NS
TT31P	22 [18, 34]	0	II	B	Heterosexual	F	NS	21,826,347	NS

	33 [29, 45]	11	IV	B	Heterosexual	F	NS	59,189	NS
1051	17 [13, 28]	0	I	B	NS	F	NS	13,000	NS
	27 [23, 38]	10	II	B	NS	F	NS	280,000	NS
1056	17 [13, 28]	0	I	B	NS	M	NS	7,000	NS
	20 [16, 31]	3	II	B	NS	M	NS	65,000	NS
	24 [20, 35]	7	II	B	NS	M	NS	140,000	NS

**Table S3. 194 longitudinal specimens from 43 subjects. \*Only HXB2 7134-7499 segment is available.**

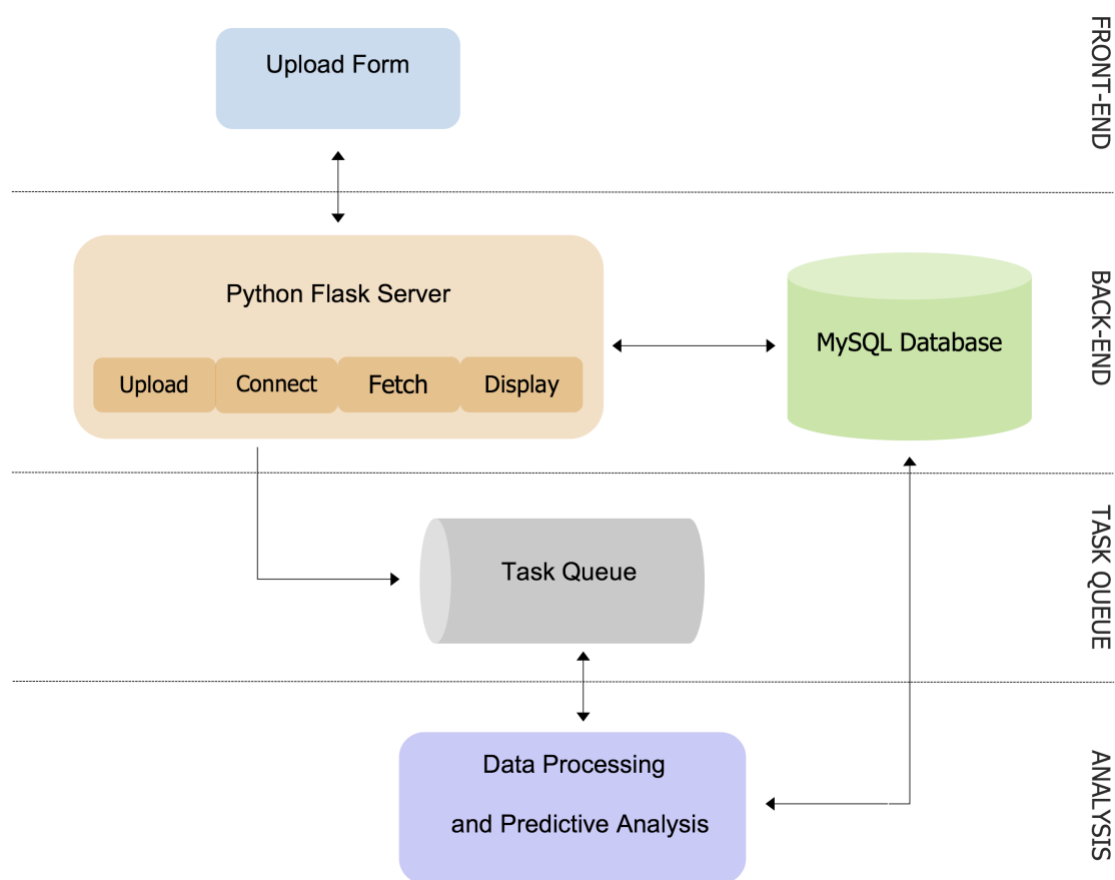
Subject	Minimum Duration of Infection (days)	Subtype	Sex	Vaccine status	Subject (Reference)	Minimum Duration of Infection (days)	Subtype	Sex	Vaccine status
AA001	618	CRF01_AE	M	placebo	AA065	152	CRF01_AE	M	vaccinee
AA002	178	CRF01_AE	M	placebo	AA066	143	CRF01_AE	F	vaccinee
AA003	513	CRF01_AE	M	placebo	AA067	191	CRF01_AE	M	placebo
AA004	168	CRF01_AE	F	placebo	AA068	170	CRF01_AE	M	placebo
AA005	203	CRF01_AE	F	placebo	AA070	461	CRF01_AE	F	vaccinee
AA006	511	CRF01_AE	F	vaccinee	AA072	184	CRF01_AE	M	vaccinee
AA007	184	CRF01_AE	M	placebo	AA073	182	CRF01_AE	M	vaccinee
AA008	433	CRF01_AE	F	placebo	AA074	163	CRF01_AE	M	vaccinee
AA009	182	CRF01_AE	M	vaccinee	AA075	168	CRF01_AE	M	placebo
AA010	135	B	M	vaccinee	AA076	157	CRF01_AE	F	placebo
AA011	223	B	M	placebo	AA077	189	CRF01_AE	F	vaccinee
AA012	168	CRF01_AE	M	vaccinee	AA078	185	CRF01_AE	F	placebo
AA013	170	CRF01_AE	M	vaccinee	AA079	181	CRF01_AE	M	placebo
AA014	353	CRF01_AE	M	placebo	AA081	20	CRF01_AE	F	placebo
AA015	196	CRF01_AE	M	vaccinee	AA082	183	CRF01_AE	F	placebo
AA016	191	CRF01_AE	F	placebo	AA083	176	CRF01_AE	F	placebo
AA017	167	CRF01_AE	M	placebo	AA084	175	CRF01_AE /B	M	vaccinee
AA018	173	CRF01_AE	F	vaccinee	AA085	218	CRF01_AE	M	vaccinee
AA019	189	CRF01_AE	M	placebo	AA086	192	CRF01_AE	F	placebo
AA020	201	CRF01_AE/B	F	vaccinee	AA088	168	CRF01_AE	F	placebo
AA021	169	CRF01_AE	M	vaccinee	AA089	180	CRF01_AE	M	vaccinee
AA022	52	CRF01_AE	M	placebo	AA090	336	CRF01_AE	M	vaccinee
AA023	475	CRF01_AE	M	placebo	AA091	209	CRF01_AE	F	placebo
AA024	181	CRF01_AE	M	vaccinee	AA092	173	CRF01_AE	F	vaccinee
AA025	182	CRF01_AE/B	M	placebo	AA093	157	B	M	vaccinee
AA026	176	CRF01_AE	M	placebo	AA094	169	CRF01_AE	M	placebo
AA027	173	CRF01_AE	F	placebo	AA095	186	CRF01_AE /B	M	placebo
AA028	222	CRF01_AE	M	vaccinee	AA096	177	CRF01_AE /B	F	vaccinee
AA029	168	CRF01_AE	M	placebo	AA097	105	CRF01_AE	M	vaccinee
AA031	183	CRF01_AE	F	placebo	AA098	167	CRF01_AE	M	placebo
AA032	187	CRF01_AE	F	placebo	AA099	126	CRF01_AE	F	placebo
AA033	161	CRF01_AE	F	placebo	AA100	182	CRF01_AE	F	vaccinee
AA034	185	CRF01_AE	F	vaccinee	AA101	154	CRF01_AE	F	placebo

AA035	203	CRF01_AE	M	placebo	AA102	126	CRF01_AE	F	placebo
AA036	174	CRF01_AE	M	vaccinee	AA103	168	CRF01_AE	M	vaccinee
AA037	272	CRF01_AE	M	placebo	AA104	154	CRF01_AE	F	placebo
AA038	205	CRF01_AE	M	placebo	AA105	192	CRF01_AE	M	placebo
AA039	500	CRF02_AG	F	vaccinee	AA107	204	CRF01_AE	M	placebo
AA040	285	B	F	placebo	AA108	179	CRF01_AE	F	vaccinee
AA041	191	CRF01_AE	F	placebo	AA109	196	CRF01_AE	M	placebo
AA042	399	CRF01_AE	M	placebo	AA110	283	CRF01_AE	F	placebo
AA044	191	CRF01_AE	M	placebo	AA111	178	CRF01_AE	F	vaccinee
AA045	183	CRF01_AE	M	placebo	AA112	175	CRF01_AE	M	placebo
AA048	512	CRF01_AE	M	vaccinee	AA113	147	CRF01_AE	F	vaccinee
AA049	179	CRF01_AE	M	placebo	AA115	183	B	M	placebo
AA050	208	CRF01_AE	F	vaccinee	AA116	164	CRF01_AE	M	placebo
AA051	179	CRF01_AE	M	placebo	AA117	178	CRF01_AE	M	placebo
AA052	228	CRF01_AE	M	vaccinee	AA118	181	CRF01_AE	M	vaccinee
AA054	168	CRF01_AE	M	placebo	AA119	161	CRF01_AE	F	placebo
AA055	182	CRF01_AE	F	vaccinee	AA121	154	CRF01_AE	F	vaccinee
AA057	156	CRF01_AE	F	placebo	AA122	168	CRF01_AE	M	placebo
AA058	161	CRF01_AE	M	vaccinee	AA123	182	CRF01_AE	M	vaccinee
AA059	164	CRF01_AE	F	placebo	AA125	162	CRF01_AE	F	vaccinee
AA060	170	CRF01_AE	F	placebo	AA126	78	CRF01_AE	M	placebo
AA061	190	CRF01_AE	M	vaccinee	AA127	156	CRF01_AE	M	vaccinee
AA062	158	CRF01_AE	M	vaccinee	AA129	199	CRF01_AE	F	placebo
AA063	182	CRF01_AE	M	vaccinee	AA130	186	CRF01_AE	M	placebo
AA064	167	CRF01_AE	M	vaccinee	AA131	245	CRF01_AE	M	vaccinee

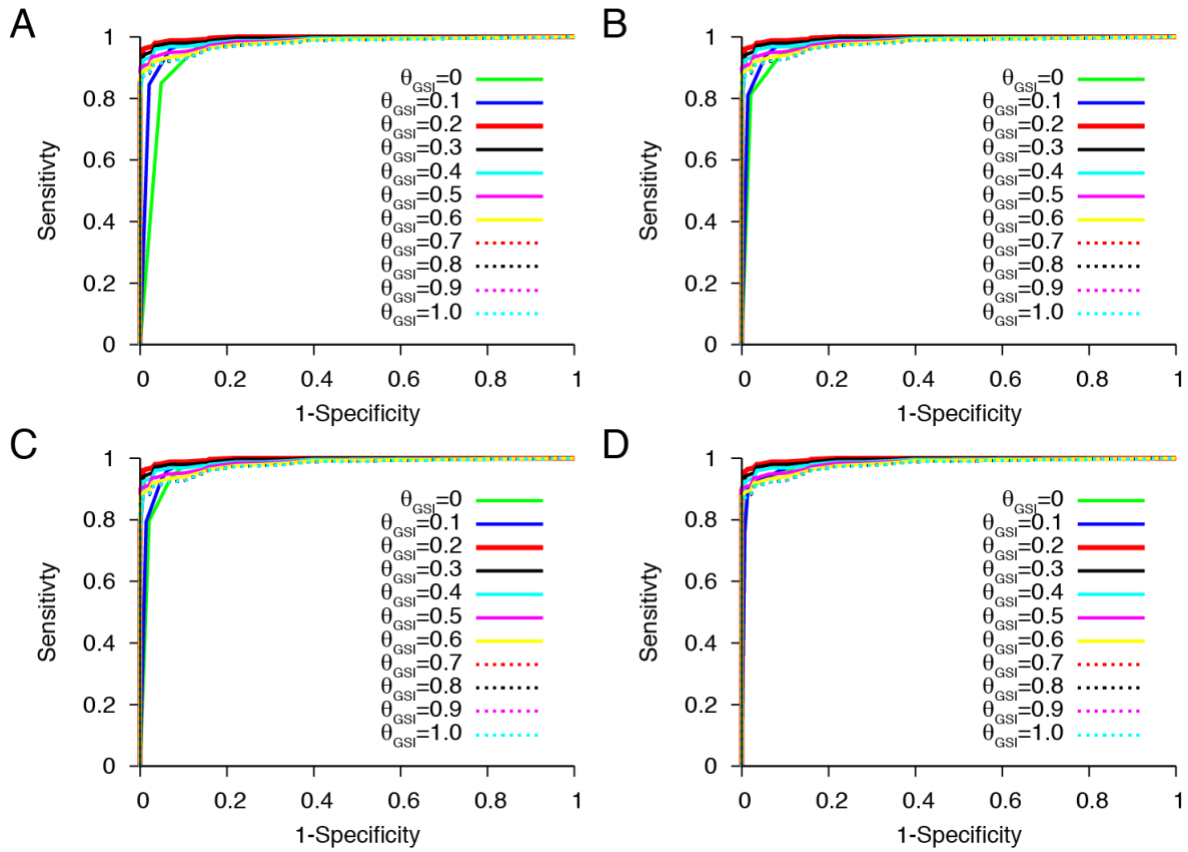
**Table S4. 116 incident specimens from RV144 vaccine trial from references, 43, 44, and 45.** Genbank accession numbers are JX446645–JX448316.

	Full Envelope (547 incident and 149 chronic)	HXB2 7134-7499 (585 incident and 305 chronic)
FRR	1.3% [0% - 4.0%]	1.3% [0% - 3.9%]
MDRI	468 [362 – 628] days	405 [310 – 531] days
Sensitivity	94.2% [89.8% – 97.8%]	85.0% [79.1% – 90.8%]
Prediction error for time since infection	13.6% [8.1% – 19.5%]	18.1% [12.1% – 24.6%]

**Table S5. HIITE performance from k-fold cross validation.**

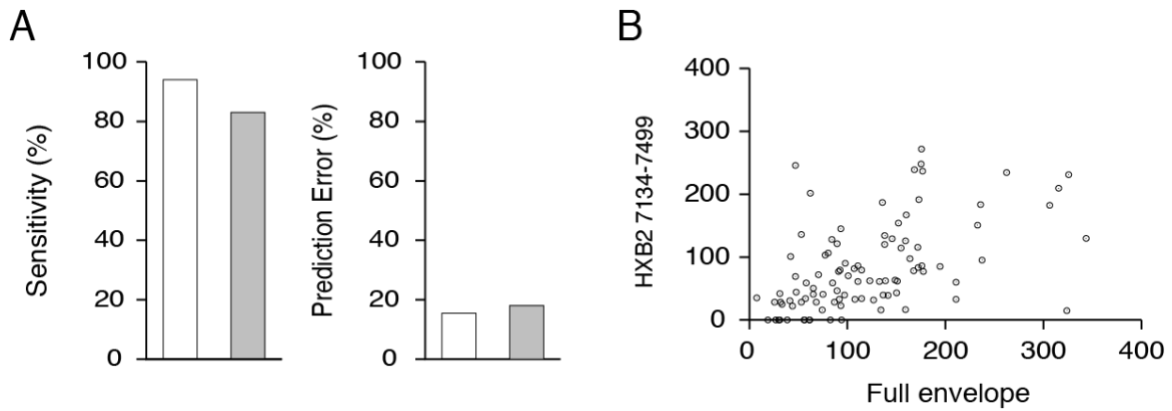


**Figure S1. HIITE’s frontend, backend and database architecture.** The frontend was developed using HTML5 and JavaScript, backend with Python using Flask framework, and database with MySQL. The frontend and backend are interlinked via a set of REST style APIs (Application Programming Interface) accessed over HTTP protocol. The upload API creates a bash file for each request, which is subjected to HIITE task queue and database. The task queue controls the execution of submitted requests and distributes each request into execution loads. The fetch API is invoked when users submit their request. This API polls the database and alerts the user that the request is being processed. Once processing is finished, the display API renders HIITE’s results including the stage of infection (chronic or incident), Hamming distance distribution plot, diversity, GSI, variance, and estimated days’ post infection with a 95% prediction interval for incident stage.

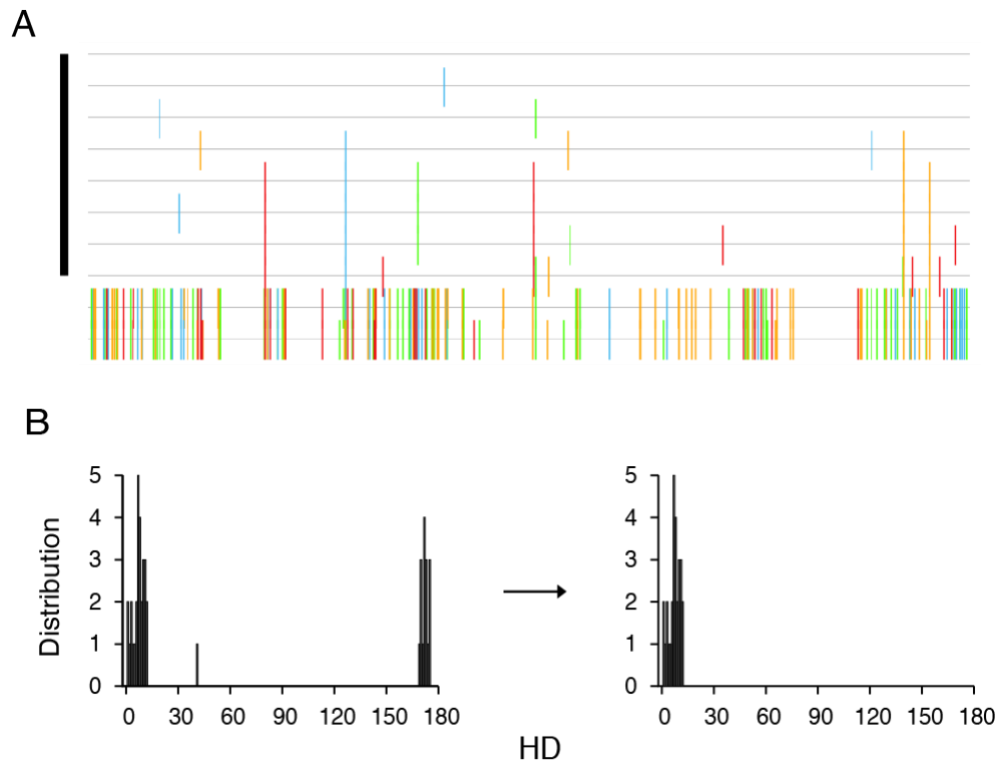


**Figure S2. The receiver operating characteristic (ROC) analysis.** **A.** The ROC curves for  $\theta_\beta = 1$  and  $\theta_{GSI}$  ranging from 0 to 1. The area under the ROC curve was 0.96232 for  $\theta_{GSI} = 0$ , 0.97937 for  $\theta_{GSI} = 0.1$ , 0.99673 for  $\theta_{GSI} = 0.2$ , 0.99414 for  $\theta_{GSI} = 0.3$ , 0.98950 for  $\theta_{GSI} = 0.4$ , 0.985174 for  $\theta_{GSI} = 0.5$ , 0.98185 for  $\theta_{GSI} = 0.6$ , 0.97799 for  $\theta_{GSI} = 0.7$ , 0.97798 for  $\theta_{GSI} = 0.8$ , 0.977954 for  $\theta_{GSI} = 0.9$ , and 0.977954 for  $\theta_{GSI} = 1.0$ . **B.** The ROC curves for  $\theta_\beta = 2$  and  $\theta_{GSI}$  ranging from 0 to 1. The area under the ROC curve was 0.97439 for  $\theta_{GSI} = 0$ , 0.98224 for  $\theta_{GSI} = 0.1$ , 0.99653 for  $\theta_{GSI} = 0.2$ , 0.99398 for  $\theta_{GSI} = 0.3$ , 0.98938 for  $\theta_{GSI} = 0.4$ , 0.98517 for  $\theta_{GSI} = 0.5$ , 0.98185 for  $\theta_{GSI} = 0.6$ , 0.97799 for  $\theta_{GSI} = 0.7$ , 0.97798 for  $\theta_{GSI} = 0.8$ , 0.977954 for  $\theta_{GSI} = 0.9$ , and 0.977954 for  $\theta_{GSI} = 1.0$ . **C.** The ROC curves for  $\theta_\beta = 4$  and  $\theta_{GSI}$  ranging from 0 to 1. The area under the ROC curve was 0.97591 for  $\theta_{GSI} = 0$ , 0.98250 for  $\theta_{GSI} = 0.1$ , 0.99647 for  $\theta_{GSI} = 0.2$ , 0.99395 for  $\theta_{GSI} = 0.3$ , 0.98933 for  $\theta_{GSI} = 0.4$ , 0.98517 for  $\theta_{GSI} = 0.5$ , 0.98185 for  $\theta_{GSI} = 0.6$ , 0.97799 for  $\theta_{GSI} = 0.7$ , 0.97798 for  $\theta_{GSI} = 0.8$ , 0.977954 for  $\theta_{GSI} = 0.9$ , and 0.977954 for  $\theta_{GSI} = 1.0$ . **D.** The ROC curves for  $\theta_\beta = 10$  and  $\theta_{GSI}$  ranging from 0 to 1. The area under the ROC curve was 0.98661 for  $\theta_{GSI} = 0$ , 0.98779 for  $\theta_{GSI} = 0.1$ , 0.99683 for  $\theta_{GSI} = 0.2$ , 0.99437 for  $\theta_{GSI} = 0.3$ , 0.98974 for  $\theta_{GSI} = 0.4$ , 0.98517 for  $\theta_{GSI} = 0.5$ , 0.98185 for  $\theta_{GSI} = 0.6$ , 0.97799 for  $\theta_{GSI} = 0.7$ , 0.97798 for  $\theta_{GSI} = 0.8$ , 0.977954 for  $\theta_{GSI} = 0.9$ , and 0.977954 for  $\theta_{GSI} = 1.0$ .

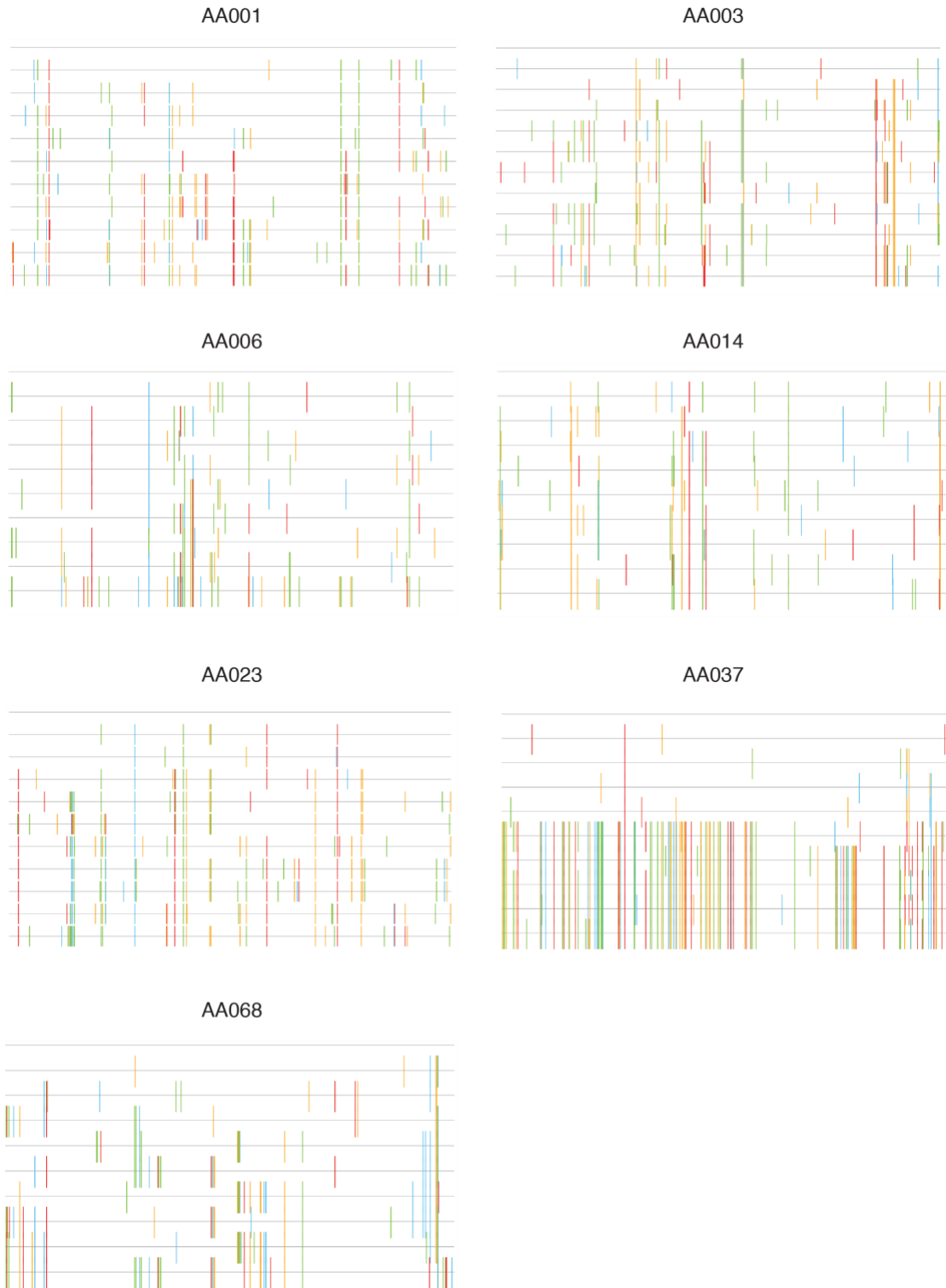




**Figure S3. HIITE’s performance on RV144 trial participants. A.** HIITE’s sensitivity and prediction error for time since infection are measured from 116 RV144 trial subjects’ full envelope gene sequences (white bar) and envelope gene segment HXB2 7134-7499 sequences (grey bar). The sensitivity is defined as the proportion of incident specimens correctly classified as incident. The prediction error denotes the proportion of specimens whose estimated infection time interval deviates from the HIV negative test record. **B.** Time since infection estimates for full envelope gene and those for segment HXB2 7134-7499 of the 96 RV144 subjects. The correlation between the two estimates was statistically significant ( $\rho=0.56$ ,  $p<0.001$ ).



**Figure S4. HIITE's performance on subject AA039 from RV144 vaccine trial. A.** Subject AA039's highlighter plot ([https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter\\_top.html](https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html)) shows the signature of multiple founder variants. HIITE successfully clustered subject AA039's sequences into two lineages (major lineage's sequences are marked by a black line), allowing incident classification. HIITE estimated an infection time of 141 [48 – 233] days, conforming with this subject's maximum duration of infection of 500 days. **B.** The distribution of subject AA039's intersequence nucleotide differences (Hamming distance) before clustering and the major lineage's Hamming distance distribution after clustering.



**Figure S5. Highlighter plot for 7 RV144 trial subjects classified as chronic infection.**

The Highlighter tool, [https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter\\_top.html](https://www.hiv.lanl.gov/content/sequence/HIGHLIGHT/highlighter_top.html), was used.

## REFERENCES

1. Bailey JR, Lassen KG, Yang HC, et al. Neutralizing antibodies do not mediate suppression of human immunodeficiency virus type 1 in elite suppressors or selection of plasma virus variants in patients on highly active antiretroviral therapy. *J Virol* 2006; **80**(10): 4758-70.
2. Yoshida I, Sugiura W, Shibata J, Ren F, Yang Z, Tanaka H. Change of positive selection pressure on HIV-1 envelope gene inferred by early and recent samples. *PLoS One* 2011; **6**(4): e18630.
3. Ren C, Liu S, Li Y, et al. Cross-neutralizing antibody profile of Chinese HIV-1-infected individuals and the viral envelope features from elite neutralizers. *J Acquir Immune Defic Syndr* 2014; **67**(5): 472-80.
4. Rong R, Li B, Lynch RM, et al. Escape from autologous neutralizing antibodies in acute/early subtype C HIV-1 infection requires multiple pathways. *PLoS Pathog* 2009; **5**(9): e1000594.
5. Bunnik EM, Pisas L, van Nuenen AC, Schuitemaker H. Autologous neutralizing humoral immunity and evolution of the viral envelope in the course of subtype B human immunodeficiency virus type 1 infection. *J Virol* 2008; **82**(16): 7932-41.
6. Edo-Matas D, Lemey P, Tom JA, et al. Impact of CCR5delta32 host genetic background and disease progression on HIV-1 intrahost evolutionary processes: efficient hypothesis testing through hierarchical phylogenetic models. *Mol Biol Evol* 2011; **28**(5): 1605-16.
7. van Gils MJ, Bunnik EM, Burger JA, et al. Rapid escape from preserved cross-reactive neutralizing humoral immunity without loss of viral fitness in HIV-1-infected progressors and long-term nonprogressors. *J Virol* 2010; **84**(7): 3576-85.
8. Gnanakaran S, Bhattacharya T, Daniels M, et al. Recurrent signature patterns in HIV-1 B clade envelope glycoproteins associated with either early or chronic infections. *PLoS Pathog* 2011; **7**(9): e1002209.
9. Keele BF, Giorgi EE, Salazar-Gonzalez JF, et al. Identification and characterization of transmitted and early founder virus envelopes in primary HIV-1 infection. *Proc Natl Acad Sci U S A* 2008; **105**(21): 7552-7.
10. Herbeck JT, Rolland M, Liu Y, et al. Demographic processes affect HIV-1 evolution in primary infection before the onset of selective processes. *J Virol* 2011; **85**(15): 7523-34.
11. Liu Y, McNevin J, Cao J, et al. Selection on the human immunodeficiency virus type 1 proteome following primary infection. *J Virol* 2006; **80**(19): 9519-29.
12. Daniels RS, Wilson P, Patel D, Longhurst H, Patterson S. Analysis of full-length HIV type 1 env genes indicates differences between the virus infecting T cells and dendritic cells in peripheral blood of infected patients. *AIDS Res Hum Retroviruses* 2004; **20**(4): 409-13.
13. Doria-Rose NA, Schramm CA, Gorman J, et al. Developmental pathway for potent V1V2-directed HIV-neutralizing antibodies. *Nature* 2014; **509**(7498): 55-62.
14. Wu X, Wang C, O'Dell S, et al. Selection pressure on HIV-1 envelope by broadly neutralizing antibodies to the conserved CD4-binding site. *J Virol* 2012; **86**(10): 5844-56.

15. Blish CA, Dogan OC, Derby NR, et al. Human immunodeficiency virus type 1 superinfection occurs despite relatively robust neutralizing antibody responses. *J Virol* 2008; **82**(24): 12094-103.
16. Piantadosi A, Chohan B, Chohan V, McClelland RS, Overbaugh J. Chronic HIV-1 infection frequently fails to protect against superinfection. *PLoS Pathog* 2007; **3**(11): e177.
17. Geels MJ, Cornelissen M, Schuitemaker H, et al. Identification of sequential viral escape mutants associated with altered T-cell responses in a human immunodeficiency virus type 1-infected individual. *J Virol* 2003; **77**(23): 12430-40.
18. Sturdevant CB, Joseph SB, Schnell G, Price RW, Swanstrom R, Spudich S. Compartmentalized replication of R5 T cell-tropic HIV-1 in the central nervous system early in the course of infection. *PLoS Pathog* 2015; **11**(3): e1004720.
19. Ssemwanga D, Lyagoba F, Ndemi N, et al. Multiple HIV-1 infections with evidence of recombination in heterosexual partnerships in a low risk Rural Clinical Cohort in Uganda. *Virology* 2011; **411**(1): 113-31.
20. Skar H, Gutenkunst RN, Wilbe Ramsay K, Alaeus A, Albert J, Leitner T. Daily sampling of an HIV-1 patient with slowly progressing disease displays persistence of multiple env subpopulations consistent with neutrality. *PLoS One* 2011; **6**(8): e21747.
21. Ping LH, Joseph SB, Anderson JA, et al. Comparison of viral Env proteins from acute and chronic infections with subtype C human immunodeficiency virus type 1 identifies differences in glycosylation and CCR5 utilization and suggests a new strategy for immunogen design. *J Virol* 2013; **87**(13): 7218-33.
22. Mukhopadhyay S, Ringe R, Patil A, Paranjape R, Bhattacharya J. Characterization of circulating HIV type 1 env genes in plasma of two antiretroviral-naive slow progressing patients with broad neutralizing antibody response with evidence of recombination. *AIDS Res Hum Retroviruses* 2012; **28**(7): 739-45.
23. Malherbe DC, Pissani F, Sather DN, et al. Envelope variants circulating as initial neutralization breadth developed in two HIV-infected subjects stimulate multiclade neutralizing antibodies in rabbits. *J Virol* 2014; **88**(22): 12949-67.
24. Evering TH, Kamau E, St Bernard L, Farmer CB, Kong XP, Markowitz M. Single genome analysis reveals genetic characteristics of Neuroadaptation across HIV-1 envelope. *Retrovirology* 2014; **11**: 65.
25. Shankarappa R, Margolick JB, Gange SJ, et al. Consistent viral evolutionary changes associated with the progression of human immunodeficiency virus type 1 infection. *J Virol* 1999; **73**(12): 10489-502.
26. Jensen MA, Li FS, van 't Wout AB, et al. Improved coreceptor usage prediction and genotypic monitoring of R5-to-X4 transition by motif analysis of human immunodeficiency virus type 1 env V3 loop sequences. *J Virol* 2003; **77**(24): 13376-88.
27. Park SY, Goeken N, Lee HJ, Bolan R, Dube MP, Lee HY. Developing high-throughput HIV incidence assay with pyrosequencing platform. *J Virol* 2014; **88**(5): 2977-90.
28. Liao HX, Lynch R, Zhou T, et al. Co-evolution of a broadly neutralizing HIV-1 antibody and founder virus. *Nature* 2013; **496**(7446): 469-76.
29. Park SY, Love TM, Reynell L, et al. The HIV Genomic Incidence Assay Meets False Recency Rate and Mean Duration of Recency Infection Performance Standards. *Scientific Reports* 2017; **7**(1): 7480.

30. Baalwa J, Wang S, Parrish NF, et al. Molecular identification, cloning and characterization of transmitted/founder HIV-1 subtype A, D and A/D infectious molecular clones. *Virology* 2013; **436**(1): 33-48.
31. Yue L, Pfafferoth KJ, Baalwa J, et al. Transmitted virus fitness and host T cell responses collectively define divergent infection outcomes in two HIV-1 recipients. *PLoS Pathog* 2015; **11**(1): e1004565.
32. Chen Y, Li N, Zhang T, et al. Comprehensive Characterization of the Transmitted/Founder env Genes From a Single MSM Cohort in China. *J Acquir Immune Defic Syndr* 2015; **69**(4): 403-12.
33. Masharsky AE, Dukhovlinova EN, Verevochkin SV, et al. A substantial transmission bottleneck among newly and recently HIV-1-infected injection drug users in St Petersburg, Russia. *J Infect Dis* 2010; **201**(11): 1697-702.
34. Manak M, Sina S, Anekella B, et al. Pilot studies for development of an HIV subtype panel for surveillance of global diversity. *AIDS Res Hum Retroviruses* 2012; **28**(6): 594-606.
35. Abrahams MR, Anderson JA, Giorgi EE, et al. Quantitating the multiplicity of infection with human immunodeficiency virus type 1 subtype C reveals a non-poisson distribution of transmitted variants. *J Virol* 2009; **83**(8): 3556-67.
36. Bar KJ, Li H, Chamberland A, et al. Wide variation in the multiplicity of HIV-1 infection among injection drug users. *J Virol* 2010; **84**(12): 6241-7.
37. Heipertz RA, Jr., Sanders-Buell E, Kijak G, et al. Molecular epidemiology of early and acute HIV type 1 infections in the United States Navy and Marine Corps, 2005-2010. *AIDS Res Hum Retroviruses* 2013; **29**(10): 1310-20.
38. Li H, Bar KJ, Wang S, et al. High Multiplicity Infection by HIV-1 in Men Who Have Sex with Men. *PLoS Pathog* 2010; **6**(5): e1000890.
39. Parrish NF, Wilen CB, Banks LB, et al. Transmitted/founder and chronic subtype C HIV-1 use CD4 and CCR5 receptors with equal efficiency and are not inhibited by blocking the integrin alpha4beta7. *PLoS Pathog* 2012; **8**(5): e1002686.
40. Parrish NF, Gao F, Li H, et al. Phenotypic properties of transmitted founder HIV-1. *Proc Natl Acad Sci U S A* 2013; **110**(17): 6626-33.
41. Nofemela A, Bandawe G, Thebus R, et al. Defining the human immunodeficiency virus type 1 transmission genetic bottleneck in a region with multiple circulating subtypes and recombinant forms. *Virology* 2011; **415**(2): 107-13.
42. Salazar-Gonzalez JF, Salazar MG, Keele BF, et al. Genetic identity, biological phenotype, and evolutionary pathways of transmitted/founder viruses in acute and early HIV-1 infection. *J Exp Med* 2009; **206**(6): 1273-89.