EARLY EVOLUTION OF HIV-1 FROM TRANSMITTED FOUNDERS DURING ACUTE INFECTION

HIV- Dynamics and Replication Program
NCI-Frederick

Valerie F Boltz[1], Wei Shao[2], Clarissa Halpern[1], Adam A Capoferri[1], Morgane Rolland[3,4], Nittaya Phanuphak[7], Denise Hsu[3,4], Carlo Sacdalan[7], John W Mellors[5], John M Coffin[6], Mary F Kearney[1], on behalf of the RV254/SEARCH 010 Study Team

[1]HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702, [2]Leidos Biomedical Research, Inc., Frederick National Laboratories for Cancer Research, Frederick, MD 21702; [3]US Military HIV Research Program, Walter Reed Army Institute of Research, Sliver Spring MD, 20901, [4]Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., Bethesda, MD, USA [5]Department of Medicine, University of Pittsburgh, Pittsburgh, PA 15261, [6]Department of Molecular Biology and Microbiology, Tufts University, Boston, MA 02111, [7]SEARCH, Institute of HIV Research and Innovation, Bangkok, Thailand

BACKGROUND/AIMS

To better define the number of transmitted founder variants and sequence divergence from founders in the first few weeks following HIV-1 transmission, we performed ultra-sensitive single genome sequencing (uSGS) of *pol* and *env* HIV-1 RNA in plasma samples from individuals with acute HIV-1 infection.

METHODS

HIV-1 RNA was extracted from plasma of donors with acute CRF01_AE infection (Fiebig II-IV) enrolled in the RV254/SEARCH010 Cohort (NCT00796146). Five donors had been identified by standard SGS as having multiple transmitted founders (TFs) and 10 as single TFs, consistent with a prior report showing 70% having single TFs in RV254. Our ultrasensitive SGS method with primer IDs and paired-end Illumina sequencing was applied to identify >10,000 independent *pol* and *env* sequences per sample. We used radial trees (Figure 1) and star phylogeny to identify TFs, calculated transition/transversion (Ti/Tv), non-synonymous/synonymous mutations, identified drug resistance mutations and analyzed the phylogenetic relationships among emergent variants.

Fig 1.

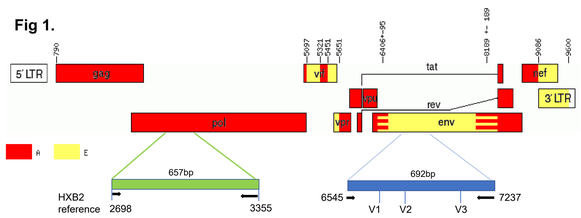

Figure 1. HIV genome map showing the 2 sub-genomic regions amplified and sequenced. *pol* uSGS amplicons resulted in 500 bp of RT from amino acid 59-146 and 184-260. *env* amplicons resulted in 500 bp of gp120 from amino acids 95-183 encompassing all of V1-V2 and amino acids 225-302 encompassing all of v3

Table 1. Donor demographics of 15 SEARCH trial plasma samples

Donor code (ID)	Donor sex	Donor age	Sample date	Fiebig Stage	Viral load (copies/ml)	CD4 count (cells/ul)	Subtype	T/F	No. of *pol* seq	No. of NFL seq
6609	M	29	2010	III	4,112,500	206	CRF01_AE	Multiple	10	10
6340	M	29	2011	III	23,317,000	338	CRF01_AE	Multiple	10	10
7905	M	29	2011	II	7,263,860	213	CRF01_AE	Multiple	9	9
9813	M	22	2015	II	5,163,823	165	CRF01_AE	Multiple	10	10
7268	M	23	2015	III	8,914,961	334	CRF01_AE	Multiple	10	10
3928	M	46	2009	III	5,517,440	525	CRF01_AE	Single	11	ND
3832	M	36	2010	II	36,694,000	269	CRF01_AE	Single	11	11
5436	M	32	2010	III	30,811,000	621	CRF01_AE	Single	10	10
3698	M	30	2010	III	4,939,160	352	CRF01_AE	Single	10	ND
8123	F	45	2011	III	25,579,700	132	CRF01_AE	Single	10	ND
3513	M	28	2011	III	13,557,900	359	CRF01_AE	Single	10	10
9114	M	22	2011	III	2,656,900	350	CRF01_AE	Single	9	9
3193	M	24	2012	II	2,412,840	289	CRF01_AE	Single	9	9
7279	M	27	2014	IV	7,249,767	234	CRF01_AE	Single	11	11
8522	M	29	2015	III	56,127,150	265	CRF01_AE	Single	10	10

Fig 2.

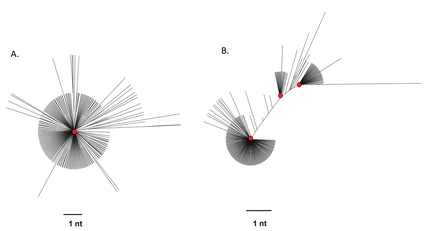

Figure 2. Star phylogeny in donor in Fiebig III after transmission of a single founder (A). Phylogeny in donor in Fiebig II after transmission of multiple founders (B). The red circles show the consensus sequences of the transmitted founders. Identical sequences have been collapsed to one.

RESULTS

Table 2. Total numbers of single genomes resulting from uSGS of 500 bp from 2 sub-genomic regions: *pol* and *env*

		pol sub-genomic region			*env* sub-genomic region		
ID	Fiebig Stage	Total # Genomes	# Possible Founders	# Identical Genomes in TF	Total # Genomes	# Possible Founders	# Identical Genomes in TF
6340	III	12,524	3	9,431, 656, 239	10,105	3	7,229, 693, 459
7905	II	20,924	4	10,536, 2,062, 3,492, 3,223	11,203	3	5,222, 3,273, 1129
7279	IV	11,265	2	9,107, 798	9,755	1	8,428
9813	II	7,841	3	9,107, 798, 831	8,890	3	6,924, 1,079, 358
7268	III	2,265	3	1,524, 190, 133	2,769	2	1,282, 295
3698	III	8,156	1	4,895, 1,309	10,127	1	9,296
3928	III	19,526	1	15,782	17,245	1	16,129
3832	II	12,520	1	11,757	12,994	1	12,411
5436	III	7,762	1	6,182	10,543	1	9,730
6609	III	10,400	1	8,981	13,483	1	11,911
8123	III	11,508	1	9,773	13,870	1	12,813
3513	III	12,918	1	12,102	12,459	1	11,268
9114	III	27,912	1	23,496	10,957	1	10,233
3193	II	3,705	1	3,357	6,914	1	6,631
8522	III	22,675	1	20,520	12,616	1	11,921
Total		191,901			163,930		
Avg		12,793			10,929		

Fig 3.

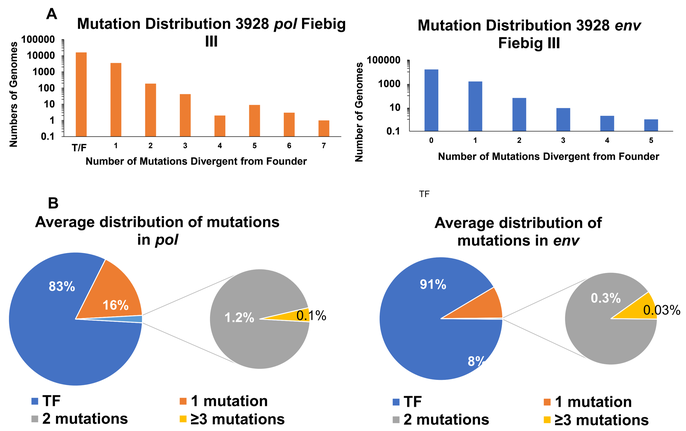

Figure 3. Histogram showing distributions of numbers of genomes that have mutations divergent from the TF in both *pol* and *env* (A). Average percent of mutations divergent from founder in all single founder infections in both *pol* and *env* (B)

Table 3. Total numbers of transitions and transversions in uSGS of 500 bp from 2 sub-genomic regions: *pol* and *env*

	pol										
Fiebig Stage	III	II	III	III	III	III	III	III	II	III	**Mean**
ID	3928	3832	5436	3698	6609	8123	3513	9114	3193	8522	
x 10^6 copies/ml	5.5	36.7	30.8	4.9	25.5	13.5	2.6	2.4	5.1	56	**18**
Total # Sequences	19,526	12,520	7,762	8,156	10,400	11,508	12,918	27,912	3,705	22,675	**13,705**
Transitions (Ti)	3,303	729	1,574	3,597	968	1,444	653	3,620	329	1,495	**1,775**
Transversions (Tv)	774	106	187	221	575	399	164	1,302	61	781	**457**
Ti/Tv ratio	4.3	6.9	8.4	16.3	1.7	3.6	4.2	2.8	5.4	1.9	**5.5**
	env										
Total # Sequences	17,245	12,994	10,543	10,127	13,483	13,870	12,459	10,957	6,914	12,616	**12,120**
Transitions (Ti)	851	456	517	638	1,320	944	1,008	674	223	438	**706**
Transversions (Tv)	353	146	355	244	297	136	235	162	75	260	**226**
Ti/Tv ratio	2.4	3.1	1.5	2.6	4.4	6.9	4.3	4.2	3.0	1.7	**3.4**

Table 4. Drug resistance mutations found in *pol* genomes in 10 single TF IDs

ID	NRTI mutation present (# of genomes)	NNRTI mutation present (# of genomes)	Linkage (# of genomes)	Total genomes
3928	K65R (10), L74V (3), Y115F(1), M184I (51), L210W (1), K219E (22)	K101E (2), Y188C (1), Y188H (1), G190E (28), M230L (1)	M184I/G190E (2)	19,530
3832	M184I (4)	G190E (10)		12,520
5436	D67N (1), M184I (3)	G190E (14), G109A (2)		7,762
3698	K65R (1), D67N (4), K70E (3), M184I (1)	Y188H (4), G190E (5)		8,156
6609	M184I (7), K219Q (7)	G190E (10)		10,400
8123	K65R (4), L70R (9), M184I (6), M184V (1)	Y188C (1), Y188H (37), G190E (12), M230L (1)		11,510
3513	M184I (9), K219E (1)	K101E (1), Y188C (1), G190A (2), G190E (5)	M184I/G190E (1)	12,918
9114	K65R (21), D67N (6), L74V (1), Y115F (1), M184I (46)	K101E (1), V106A (1), Y188H (1), G190A (1), G190E (120)	M184I/G190E (1) D67N/G190E (2)	27,884
3193	M184V (1)	Y188H (1), G190E (3)		3,607
8522	M184I (4)	Y188C (1), G190E (17), M230L (4)		22,675

Fig 4.

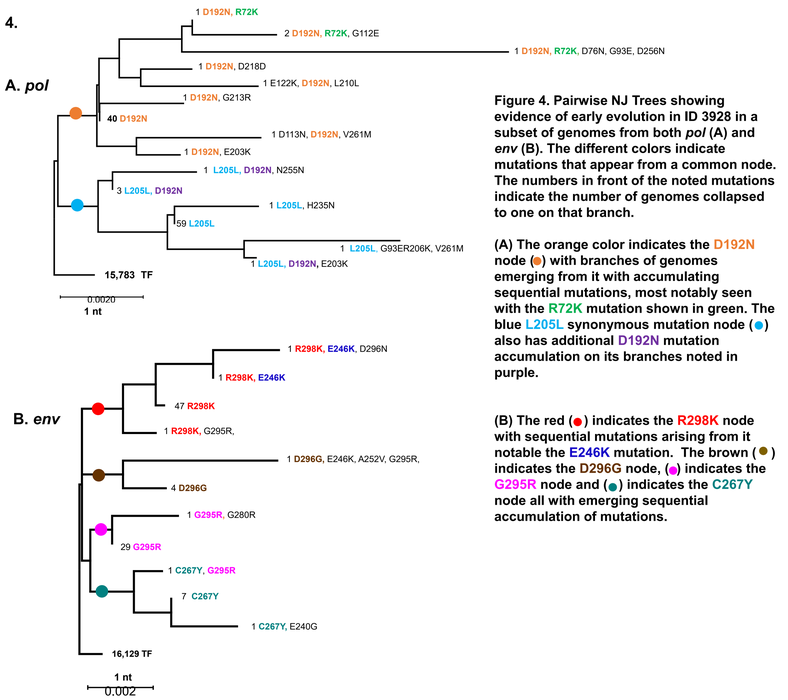

Figure 4. Pairwise NJ Trees showing evidence of early evolution in ID 3928 in a subset of genomes from both *pol* (A) and *env* (B). The different colors indicate mutations that appear from a common node. The numbers in front of the noted mutations indicate the number of genomes collapsed to one on that branch.

(A) The orange color indicates the D192N node (●) with branches of genomes emerging from it with accumulating sequential mutations, most notably seen with the R72K mutation shown in green. The blue L205L synonymous mutation node (●) also has additional D192N mutation accumulation on its branches noted in purple.

(B) The red (●) indicates the R298K node with sequential mutations arising from it notable the E246K mutation. The brown (●) indicates the D296G node, (●) indicates the G295R node and (●) indicates the C267Y node all with emerging sequential accumulation of mutations.

CONCLUSION

Ultrasensitive SGS (>10,000 genomes per sample) identified similar proportions of single vs multiple founder variants than low depth (10 genomes) SGS, verifying that HIV-1 infection is most commonly established by a single variant. By contrast, ultrasensitive SGS revealed viral evolution from the founder variants in all samples even those from as early as Fiebig stage II, indicating rapid evolution of HIV-1, including appearance of mutations at sites known for DRMs, soon after transmission.

Acknowledgements:
We would like to thank the study participants who committed so much of their time for this study. The participants were from the RV254/SEARCH 010, which is supported by cooperative agreements (WW81XWH-18-2-0040) between the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc., and the United States Army Medical Research and Development Command (USAMRDC) and by an intramural grant from the Thai Red Cross AIDS Research Centre and, in part, by the Division of AIDS, the National Institute of Allergy and Infectious Diseases, National Institutes of Health (AAI21058-001-01000). Antiretroviral therapy for RV254/SEARCH 010 participants was supported by the Thai Government Pharmaceutical Organization, Gilead Sciences, Merck and ViiV Healthcare.

Disclaimer:
The views expressed are those of the authors and should not be construed to represent the positions of the U.S. Army, the Department of Defense, the National Institutes of Health, the Department of Health and Human Services, or the Henry M. Jackson Foundation for the Advancement of Military Medicine, Inc. The investigators have adhered to the policies for protection of human subjects as prescribed in AR-70-25.