Full-Length Integrated Proviral-Single Genome Sequencing (FLIP-SGS) to Determine Both the Integration Site and Sequence of Specific HIV-1 Proviruses in Infected Individuals

Sean C. Patro[1], Leah D. Brandt[2], Michael J. Bale[1], Elias K. Halvas[2], Jenna M. Hasson[1], Aurelie Niyongabo[1], Kevin Joseph[2], Xiaolin Wu[3], Shuang Guo[3], Michele Sobolewski[2], Brandon F. Keele[4], Steven G. Deeks[5], Stephen H. Hughes[1], John W. Mellors[2], John M. Coffin[6], Jason W. Rausch[7] and Mary F. Kearney[1]

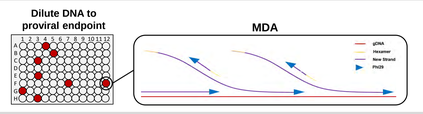
[1]HIV Dynamics and Replication Program, National Cancer Institute, Frederick, MD 21702; [2]Department of Medicine, University of Pittsburgh, Pittsburgh, PA 15213; [3]Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD 21702; [4]AIDS and Cancer Virus Program, Leidos Biomedical Research, Inc., Frederick National Laboratory, Frederick, MD 21702; [5]Department of Medicine, University of California, San Francisco, San Francisco, CA 94143; [6]Department of Molecular Biology and Microbiology, Tufts University, Boston, MA 02111; [7]Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick, MD 21702

Abstract

Most of the HIV-1 reservoir on ART is likely contained within clonally-expanded cells carrying intact proviruses. We describe **FLIP-SGS** (*full-length integrated proviral single-genome sequencing*), a method to sequence both the integration site and near full-length (NFL) HIV proviruses. FLIP-SGS was applied to identical P6-PR-RT sequences in PBMC or lymph nodes in 4 ART-suppressed HIV-infected individuals. We obtained the integration sites and NFL sequences from several clones that contained defective proviruses and one clone with an intact provirus that matched the NFL sequence of an infectious virus identified by VOA. In all 4 donors, identical P6-PR-RT were confirmed to be of clonal origin by FLIP-SGS (identical integration sites). We also found proviruses with identical P6-PR-RT sequences that had different integration sites, i.e. “false clones” that were more common in donors with low proviral diversity or with drug resistant variants. We identified a second clone *in vivo* that contains a replication-competent provirus, providing additional evidence that the HIV-1 reservoir is comprised, at least in part, by infectious proviruses in clonally expanded cells.

Methods

Multiple Displacement Amplification (MDA)



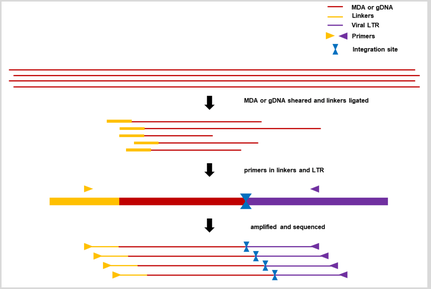
Random hexamers anneal to genomic DNA and are amplified with Phi29 for non-biased amplification. Hexamers anneal to growing strands leading to branching. New strands displace original strand.

Primary Screening and Phylogenetic Analysis

Positives are identified using nested P6-PR-RT amplification (1.5kb) and Sanger sequencing and analyzed with MEGA7. Neighbor-joining trees are built based on the number of nucleotide differences between sequences and are used to identify potential clones.

Integration Site Sequencing

Integration site analysis (ISA) sequences the host-proviral junction and maps the integration site within the human genome, identifying true expanded clones.



Near Full Length Sequencing

MDA wells of interest are subject to near full length (NFL) amplification using an outer NFL amplification (NFL1, 9kb) followed by nested near full length amplification (NFL2, 9kb) or four overlapping (3kb) fragments.

Results

Figure 1: P6-PR-RT Neighbor-joining trees analyzed by FLIP-SGS

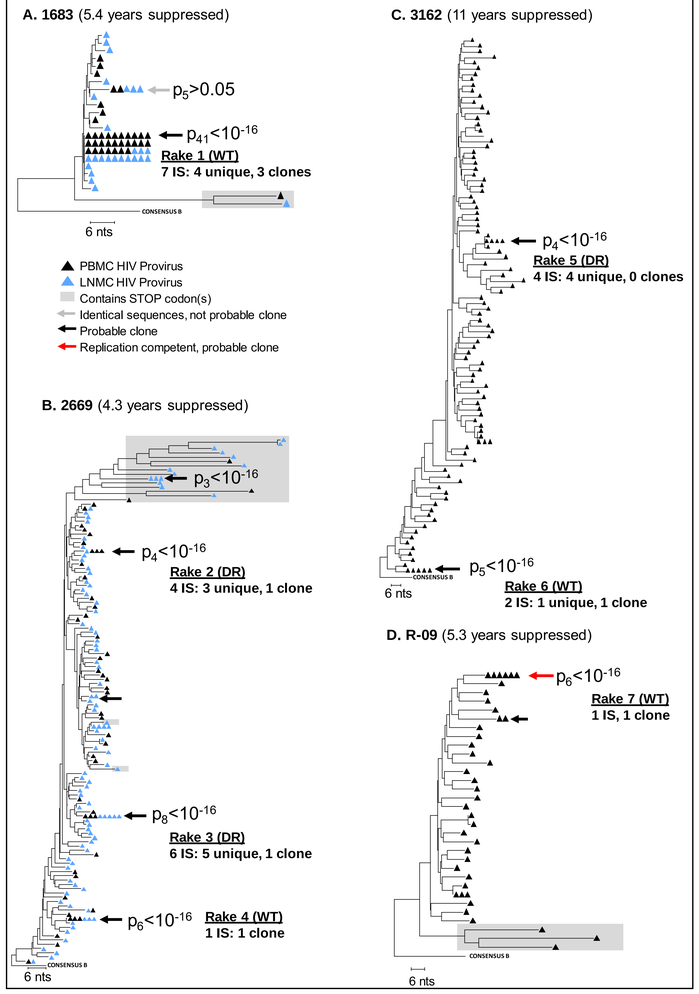
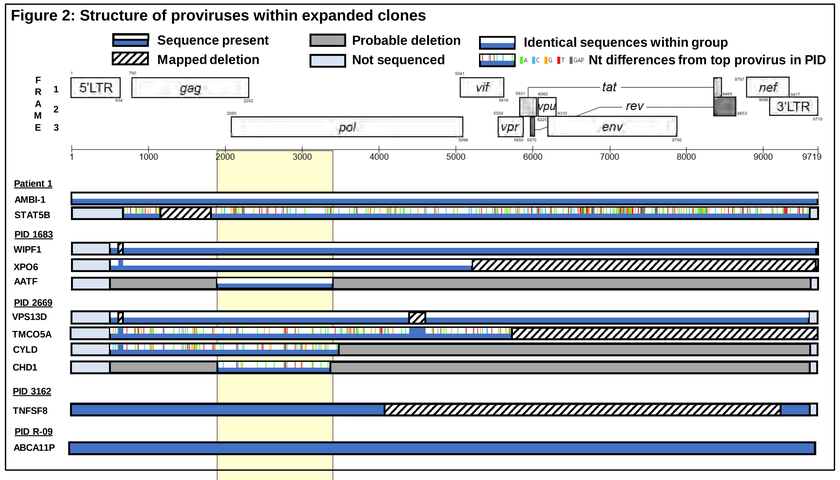


Table 1: Integration Sites in Rakes of Identical P6-PR-RT Sequences

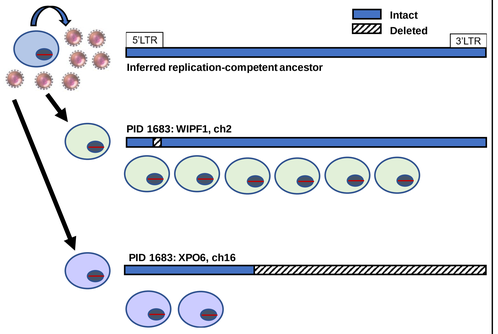
PID	Rake ID	Drug resistance mutations in rake	Chromosome	Gene (# of MDA wells with integration site)	Provirus orientation	Confirmed clone in rake
Patient 1	AMBI-1	none	Unmapped	Unmapped (12)	-	Yes
	STAT5B	none	17	STAT5B (2)	+	Yes
1683	1	none	2	WIPF1 (13)	+	Yes
			16	XPO6 (2)	+	
			17	AATF (1)	+	
			6	BACH2 (1)	-	
			9	TGFBR1 (1)	+	
			11	ATG2A (1)	+	
			17	CYTH1 (1)	-	
2669	CYLD	K103N, M184V, P225H, K238T	16	CYLD (2)	-	Yes
	2	K103N	15	TMCO5A (4)	+	Yes
			17	NPLOC4 (1)	-	
			2, 17	C9orf102 (1), C1GALT1 (1)	+, -	
			15, 16	TTBK2 (1), NFATC3 (1)	+, -	
	3	K103N	5	CHD1 (2)	-	Yes
			2	FAM117B (1)	-	
			4	LRBA (1)	-	
			7	GIMAP2 (1)	-	
			2, 17	USP34 (1), STAT3 (1)	-, +	
	4	None	1	VPS13D (2)	-	Yes
3162	5	I84V, L90M, D67N, K70R, M184V, K219Q	4	HTT (1)	-	No
			11	CHKA (1)	+	
			16	XPO6 (1)	+	
			22	ZBED4 (1)	-	
	6	None	9	TNFSF8 (3)	+	Yes
			17	B3GNTL1 (1)	+	
R-09	7	None	4	ABCA11P (2)	+	Yes

Figure 2: Structure of proviruses within expanded clones



Viral Paleontology

Figure 3: Intact viral ancestors can be inferred from proviruses archived in clones



Conclusions

- FLIP-SGS is a robust assay for paired ISA detection and NFL sequencing.
- In patients with low proviral diversity, identical P6-PR-RT may represent “true” clones and proviruses that were established early during infection.
- Characterization of highly-related/near-identical proviruses allows for reconstruction of historical, replication-competent viruses, such as transmitted-founder viruses.
- Proviruses with single and multiple drug-resistance mutations are archived in expanded clones.
- FLIP-SGS revealed a clone (like AMBI-1) that contains a replication-competent provirus in a donor fully suppressed on ART.

Acknowledgements: This project has been funded by intramural NCI funds to CCR, IATAP funding from OAR, and federal funds from the NCI under contract HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.