HIV Type 1 Genetic Variation in Foreskin and Blood from Subjects in Rakai, Uganda

Ronald M. Galiwango,¹ Susanna L. Lamers,² Andrew D. Redd,³ Jordyn Manucci,⁴ Aaron A.R. Tobian,^{4,5} Nelson Sewankambo,^{1,6} Godfrey Kigozi,¹ Gertrude Nakigozi,¹ David Serwadda,^{1,7} Iga Boaz,¹ Fred Nalugoda,¹ David J. Sullivan,⁵ Xiangrong Kong,¹ Maria J. Wawer,^{1,5} Ronald H. Gray,^{1,5} Thomas C. Quinn,^{3,4} and Oliver Laeyendecker,^{3,4} on behalf of the Rakai Health Sciences Program

Abstract

The foreskin contains a subset of dendritic cells, macrophages, and CD4⁺ and CD8⁺ T cells that may be targets for initial HIV infection in female-to-male sexual transmission of HIV-1. We present analyses comparing HIV-1 sequences isolated from foreskin DNA and serum RNA in 12 heterosexual men enrolled in an adult male circumcision trial performed in Rakai, Uganda. Phylogenetic analysis demonstrated three topologies: (1) little divergence between foreskin and serum, (2) multiple genetic bottlenecks occurring in both foreskin and serum, and (3) complete separation of foreskin and serum populations. The latter tree topology provided evidence that foreskin may serve as a reservoir for distinct HIV-1 strains. Distance and recombination analysis also demonstrated that viral genotypes in the foreskin might segregate independently from the circulating pool of viruses.

HETEROSEXUAL TRANSMISSION of HIV-1 is the most common route of infection for men in sub-Saharan Africa, which can be influenced by factors such as the physiological state of the genital mucosa, genital trauma during intercourse, and the viral load of the infecting partner.¹ Compared to circumcised men, uncircumcised males are at an increased risk for HIV infection, likely due to the large number of Langerhans' cells (LC), CD4⁺ CD8⁺ T cells, and macrophages in the foreskin.²

Established HIV-1 infection is characterized by a large degree of viral genetic heterogeneity due to the high mutational error rate of the virus and the varied host immunologic pressures in different anatomic or cellular reservoirs and compartments. Viral compartments are cell types or tissues in which there is a restriction of virus flow to the primary replicating viral population, while viral reservoirs are cell types or tissues in which replication-competent viruses accumulate and persist with more stable kinetic properties than in the primary pool of actively replicating virus.³ Reservoir sequence populations often have higher levels of genetic diversity than circulating replicating viruses obtained at the same time point due to the low level generation of new viral species and the persistence of

historical viral species.⁴ Viral compartments and reservoirs have the potential to occasionally reseed the periphery with genetically distinct viral variants, thereby increasing the chance for immune escape, development of drug-resistant mutations, and reduced ability to control the infection.³ In this study, our goal was to employ a sequence analysis approach to determine if HIV-1 viral variants found in the male foreskin differ from those in circulation.

The study population consisted of HIV-1-positive control arm men enrolled into the male circumcision trial conducted in Rakai, Uganda, in 2003–2007. Control men were offered and accepted circumcision after 24 months of follow-up. Enrollment and study details have been previously described.⁵ Institutional Review Board approvals were obtained from the Uganda Virus Research Institute's Science and Ethics Committee, Uganda National Council for Science and Technology, Western IRB, and Johns Hopkins University, Bloomberg School of Public Health. All participants provided written informed consent prior to study enrollment. Prior to surgery, a medical assessment was performed, venous blood was drawn, and serum extracted and stored at –80°C. At surgery, foreskins were removed and preserved in 70% ethanol at

¹Rakai Health Sciences Program, Kalisizo, Uganda.

²BioInfoExperts Inc., Thibodaux, Louisiana.

³Laboratory of Immunoregulation, DIR, NIAID, NIH, Baltimore Maryland.

⁴Johns Hopkins Medical Institute, Johns Hopkins University, Baltimore Maryland.

⁵Bloomberg School of Public Health, Johns Hopkins University, Baltimore Maryland.

⁶School of Medicine, Makerere University, Kampala, Uganda.

⁷School of Public Health, Makerere University, Kampala, Uganda.

–80°C. In this study, 12 men with moderate to high plasma viremia (>10,000 copies/ml) were evaluated.

Serum samples were extracted and amplified as previously described⁶ with forward primer relative to HXB2 positions 5958–5977 (TAG GCA TYT CCT ATG GCA GG) and reverse primer relative to positions 7792–7822 (AGT GCT TCC TGC TGC TCC CAA GAA CCC AAG). Subsequently, nested primers were used relative to HXB2 positions 6197–6216 (ATW AGN GAA AGA GCA GAA GA) and reverse primer relative to positions 7667–7647 (CAC TTC TCC AAT TGT CCC TCA) using identical conditions. Total DNA was extracted from foreskin samples using the QIAGEN QIAamp DNA mini kit. PCR amplification was performed for foreskin as with the serum without using a reverse transcription step. The ~1500-bp PCR products were gel purified and cloned (Invitrogen Zero Blunt TOPO PCR kit). The transformed cells were then incubated on kanamycin selective media plates at 30°C for 48 h with controls. Bacterial culture plates were submitted for DNA nucleotide sequencing (Genewiz Inc., South Plainfield, NJ). Sequences from all individuals were aligned with the Clustal W algorithm and were further optimized by hand. An all-inclusive neighbor-joining phylogenetic tree was constructed from gap-stripped sequences to verify that each subject's sequence population segregated independently. Based on the NCBI genotyping tool, all sequences were subtype D with the exception of subject N17873, who was infected with HIV-1 subtype A1 and patient N15661, who was infected with HIV recombinant form A1/D. The sequences were deposited in GenBank (JF932509–JF933732).

Maximum-likelihood (ML) trees were inferred using the program PhyML and the best fitting nucleotide substitution

model determined by a hierarchical likelihood ratio test using the HKY85 substitution and gamma rate models. Statistical support for internal branches of each tree was obtained by bootstrapping (500 replicates). ML trees were mid-point rooted and color-coded using Figtree software (<http://tree.bio.ed.ac.uk/software/figtree/>). Evolutionary distances within and between sequence tissue populations from each subject $\pm$ standard errors were calculated using the Jukes–Cantor method with 1000 bootstrap replicates. Nonsynonymous/synonymous ratios (dN/dS) for foreskin and serum sequence populations were calculated using SNAP (<http://www.hiv.lanl.gov>). The number of recombinant sequences was determined for serum and foreskin sequence populations using SplitsTree as previously described.⁷ The Genetic Algorithm Recombination Detection (GARD) method⁸ was used to approximate the location of significant breakpoints found in alignments. To determine if there was an indication of recombination occurring between foreskin and serum, a smaller population of sequences was required for analysis; therefore, random subsets of 10 sequences were extracted from patient's sequence populations. First, a split decomposition network was constructed for each subset using SplitsTree. If the signal for recombination was significant (Phi test p -value < 0.05), the sequence population was checked for recombination breakpoints using a bootscanning approach with SimPlot software (<http://sray.med.som.jhmi.edu/SCSoftware/simplot/>). The envelope V3 loop sequences were isolated in order to evaluate potential coreceptor usage. We used the V3 loop length, the appearance of a V3 loop glycosylation site, V3 total charge, the charge of the amino acids at positions 11 and 25, and the Geno2pheno web server for tropism prediction (<http://coreceptor.bioinf.mpi-inf.mpg.de/>).

FIG. 1. Maximum-likelihood phylogenies for patient foreskin and serum sequence populations. Maximum-likelihood midpoint-rooted phylogenies of HIV gp120 isolated from serum (gray) and foreskin (black) samples for three patterns are shown. Each panel represents an example of an observed pattern of viral segregation in a single patient in each of the following groups: **(A)** subject N18674, **(B)** subjects L14622, N21656, N15661, M13181, M14720, M13806, M15343, and N21096, and **(C)** subjects M16827, M17505, and N17873. For comparison purposes, all phylogenies have been adjusted so that the distance along the branches represents an equal number of nucleotide substitutions per site. Asterisks indicate that the clade subtending the branch was supported by a bootstrap value of >70%.

Maximum-likelihood phylogenies demonstrated three patterns of viral segregation among subjects' sequences: (1) One subject displayed little significant separation at any point in the tree between serum and foreskin samples or significant branches within the overall sequence population (N18674) (Fig. 1A). This subject was infected less than 6 months prior to circumcision and the phylogeny suggests a close association between viruses in both tissues. (2) Sequence populations from eight subjects had significant clades containing either foreskin or serum sequences that were distributed throughout the tree and in these clades frequent segregation points were observed (L14622, N21656, N15661, M13181, M14720, M13806, M15342, and N21096) (Fig. 1B). (3) Sequence populations from three subjects exhibited clear separation of serum and foreskin sequences with little segregation in serum samples and multiple distinct strains in the foreskin sequence populations (M16827, M17505, and N17873) (Fig. 1C). The single subject infected with HIV-1 subtype A exhibited this third pattern of viral segregation.

Five subjects showed no significant differences between serum and foreskin distance calculations (N18674, N21096, M13181, M15345, and N15661). Seven subjects' sequence populations showed significantly higher distances between serum and foreskin tissues (M13806, L14622, M14720, N21656, N17873, M17505, and M16827). As expected, a large distance between foreskin and serum sequences was calculated for each of the three subjects with highly segregated serum and foreskin sequence populations. dN/dS ratios showed evidence for selective pressure acting on foreskin sequence populations from two subjects ($dN/dS > 1.0$) (N18674, M17505). All but two foreskin samples showed a higher dN/dS ratio than serum samples, although in most cases the difference was not significant (Table 1).

Recombination analysis revealed up to 59% of the isolates in foreskin and 44% of the isolates in serum were intrasubject recombinant viruses (Table 2). Again, the three sequence serum populations corresponding to sequence populations with the phylogenetic topology shown in Fig. 1C displayed little to no signal for recombination. Overall, 26 recombinant breakpoints were identified by GARD within foreskin samples as

TABLE 2. NUMBER OF RECOMBINANTS AND RECOMBINANT BREAKPOINTS CALCULATED USING SPLITSTREE AND GENETIC ALGORITHM RECOMBINATION

ID	Recombination			
	Within group (SplitsTree) ^a		Significant breakpoints (GARD) ^b	
	Foreskin (%)	Serum (%)	Foreskin	Serum
N18674	27	33	360	508
N21096	59	37	357, 511, 751	378, 734
M13181	58	44	252,562,672,863	355,915
M15343	50	35	167,417 736,961	294,655
M13806	25	0	674	None
M14720	20	43	563	251,559
L14622	10	33	None	525
N21656	7	39	535	438
N15661	0	25	453	576
N17873	30	0	120, 319, 791, 1038	None
M17505	35	2	273, 369,824	None
M16827	20	0	322,641,729	None

^aPercent recombinants in foreskin and serum using SplitsTree.

^bLocation of significant breakpoints in foreskin and serum using the Genetic Algorithm Recombination Detection (GARD).

compared to 12 in serum sequence populations. Figure 2 demonstrates a single example of recombination occurring between serum and foreskin. In Fig. 2A, a phylogenetic network is shown for sequences from patient N18674 that contains four serum and six foreskin sequences. The Phi test for recombination was significant at $p=0.005$. Removal of one of the serum sequences resulted in the phylogenetic network shown in Fig. 2B and an insignificant Phi test of $p=0.1$. Bootscanning of these sequences showed that the query sequence, which was derived from serum, shared a similar segment in V1-V2 with a serum sample and the V4-V5 region with a foreskin sequence. This patient was the most recently infected in our study set with infection occurring less than 6 months prior to circumcision. In sequence data sets corresponding to phylogenetic trees in Fig. 1A and B, recombination events

TABLE 1. CHARACTERIZATION AND DIVERSITY WITHIN SUBJECTS' SEQUENCE POPULATIONS

ID	Number of sequences		Mean DN/DS ^a		Within-group diversity ($\pm$ SE) ^b		Between-group diversity ($\pm$ SE) ^c
	Foreskin	Serum	Foreskin	Serum	Foreskin	Serum	
N18674	44	46	1.09	0.830	1.6 (0.2)	1.4 (0.2)	1.6 (0.2)
N21096	76	52	0.84	0.85	2.1 (0.3)	2.6 (0.3)	2.7 (0.3)
M13181	36	45	0.78	0.93	3.2 (0.3)	2.9 (0.3)	3.6 (0.3)
M15343	16	46	0.85	0.84	5.0 (0.4)	5.0 (0.4)	6.2 (0.5)
M13806	51	49	0.86	0.67	2.7 (0.3)	1.6 (0.2)	2.4 (0.2)
M14720	40	40	0.77	0.55	2.6 (0.3)	3.4 (0.3)	3.4 (0.3)
L14622	52	57	0.61	0.60	2.9 (0.3)	1.7 (0.2)	3.2 (0.4)
N21656	72	51	0.78	0.65	3.1 (0.3)	5.0 (0.4)	5.9 (0.5)
N15661	18	59	0.81	0.69	1.8 (0.2)	1.6 (0.20)	2.2 (0.3)
N17873	10	58	0.66	0.40	5.6 (0.5)	0.04 (0.1)	10.9 (0.9)
M17505	37	53	1.15	0.55	5.9 (0.5)	0.4 (0.01)	6.0 (0.5)
M16827	74	64	0.72	0.87	3.4 (0.3)	1.2 (0.1)	8.5 (0.8)

^aRatio of nonsynonymous substitutions per site to synonymous substitutions per site.

^bWithin group percent diversity $\pm$ standard error.

^cBetween group diversity $\pm$ standard error.

FIG. 2. Recombination between serum and foreskin. **(A)** A phylogenetic network is shown containing four serum and six foreskin sequences from patient N18674. The Phi test for recombination was significant at $p=0.005$. Removal of one of the serum sequences resulted in the phylogenetic network shown in **(B)** and an insignificant Phi test of $p=0.1$. **(C)** All sequences in **(A)** were subjected to bootscanning, with the suspected recombinant used as the query sequence. Colored circles in **(A)** correspond to the colors representing high bootscanning values in **(C)**. **(D)** A subset of the V4 domain alignment from a patient with highly segregated serum and foreskin sequences is shown (N17873). A consensus sequences for the entire alignment is followed by two serum and three foreskin sequences. Glycosylation motifs are highlighted in the consensus sequence. A potential case of recombination between serum and foreskin in V4 is highlighted in sequence P-N17873_3_10.

between serum and foreskin were identified in all populations except M15661. Recombination events between serum and foreskin samples were not detected using the methods in Splitstree or GARD in the sequences corresponding to phylogenetic trees displayed in Fig. 1C; however, in one of these cases (N17873) we found a small segment of sequence within the sequence alignment that appears to be exchanged in a hotspot for recombination: the highly glycosylated V4 domain. In Fig. 2D this potentially exchanged sequence segment between serum and prepuc is shown. Only one foreskin sequence contained the motif NSTWDKDG, whereas all serum sequences contained this motif.

Tropism prediction analysis showed that some subjects may harbor a majority of X4 viruses (N15661, M15343, and M17505). In the subject most recently infected (N18674), potential X4 variants appeared almost equally dispersed in the phylogenetic tree between both foreskin and serum sequences. There was no relationship between tropism prediction and viral load.

The three observed segregation patterns between foreskin and serum viruses in the 12 subjects (Fig. 1) demonstrate that foreskin is not a viral compartment; however, in some cases HIV-infected cells in the foreskin may act as a reservoir for HIV. This idea is supported by analysis of the three sequence

populations associated with the phylogenetic topology in Fig. 1C and corresponds to the distance and recombination calculations in the last three sequence populations listed in Tables 1 and 2. The foreskin sequences in these three subjects demonstrate the pattern of a viral reservoir as described by Nickle *et al.*,⁴ where increased sequence diversity is observed in a reservoir tissue or cell type when compared to non-reservoir sequence populations due to the presence of both recent and historical viral sequences. At the same time, the reduced degree of viral heterogeneity in the serum sequences in these three patients may suggest a highly stable circulating virus population under reduced immunological pressure.

Recombination is a major evolutionary force affecting viral genetic variation within an HIV-1-infected individual and is predicted to occur at the same order of magnitude as point mutational change⁹; therefore, it was not surprising to find a large number of intrasubject recombinants. However, it was interesting that 7 of the 12 patients displayed an increased number of recombinants in integrated DNA within foreskin when compared to replicating RNA in the serum. Higher numbers of recombinant sequences have previously been reported in macrophage-associated brain tissues in subjects with dementia⁷; therefore, this finding may be due to the infection of different cell types in foreskin than those infected in

serum. Alternatively, the finding could be due to the persistence of diverse viral species in a tissue reservoir. The presence of recombinant sequences containing sequence fragments from both serum and foreskin in most of the cases analyzed indicates mixing or migration of viruses to and from each reservoir is possible.

HIV-1 evolves primarily in CD4⁺ T cells and macrophages, although infection of other cell types has been reported.¹⁰ While the outer foreskin is similar to the epidermis of other regions of the penis, the inner foreskin is a distinct mucocutaneous tissue that is probably less keratinized and consists of stratified squamous epithelial cells interspersed with LCs.¹¹ LCs are antigen-presenting dendritic cells of monocyte origin that reside in the epidermis. Beneath this epidermal layer is the submucosa, which contains both tissue macrophages and T cells. LCs in the foreskin epithelium exist at a 12-fold higher density than macrophages or T cells and at a 3-fold higher density in the submucosa, thus suggesting that LCs may be the first cells that come into contact with HIV.¹² Unlike T cells, HIV-infected macrophages resist apoptosis once infected.¹³ Within 2 days after sexual transmission, HIV-1 can be found in fluid draining from iliac lymph nodes and a dynamic infection of the body is in progress.¹⁴ Recent studies have used the Trofile biological assay to evaluate tropism usage in subtype D virus populations and found that, different from subtype B viruses, CCR5, CXCR4, or dual tropic viruses are equally transmissible.¹⁵ In our attempt to predict coreceptor usage, we also found several subjects who appeared to harbor a primarily CXCR4 viral genotype in both foreskin DNA and serum RNA. This may be true; however, only biological tropism testing would be able to confirm this finding.

The reduced rate of heterosexual transmission of HIV-1 to circumcised men demonstrates that the foreskin is vulnerable to infection by the virus. In the current study our genetic analyses suggest that in some cases HIV-1 migrates between the foreskin and the primary replicating viral pool in the blood. Additionally, viruses in each tissue may segregate from one another once an established infection is in place. These findings have implications for the control of viral populations during antiretroviral treatment and vaccine studies.

Acknowledgments

This work was supported by the Bill and Melinda Gates Foundation Grants 22006.02 and 22006.03, the Fogarty International Center Grant 3D43TW000010-21S2, and the Division of Intramural Research, National Institute of Allergy and Infectious Diseases, National Institutes of Health.

Author Disclosure Statement

No competing financial interests exist.

References

- Quinn TC, Wawer MJ, Sewankambo N, *et al.*: Viral load and heterosexual transmission of human immunodeficiency virus type 1. Rakai Project Study Group. *N Engl J Med* 2000; 342(13):921–929.
- Johnson KE, Redd AD, Quinn TC, *et al.*: Effects of HIV-1 and herpes simplex virus type 2 infection on lymphocyte and dendritic cell density in adult foreskins from Rakai, Uganda. *J Infect Dis* 2011;203(5):602–609.
- Blankson JN, Persaud D, and Siliciano RF: The challenge of viral reservoirs in HIV-1 infection. *Annu Rev Med* 2002;53: 557–593.
- Nickle DC, Jensen MA, Shriner D, *et al.*: Evolutionary indicators of human immunodeficiency virus type 1 reservoirs and compartments. *J Virol* 2003;77(9):5540–5546.
- Tobian AA, Ssempijja V, Kigozi G, *et al.*: Incident HIV and herpes simplex virus type 2 infection among men in Rakai, Uganda. *AIDS* 2009;23(12):1589–1594.
- Conroy SA, Laeyendecker O, Redd AD, *et al.*: Changes in the distribution of HIV type 1 subtypes D and A in Rakai District, Uganda between 1994 and 2002. *AIDS Res Hum Retroviruses* 2010;26(10):1087–1091.
- Lamers SL, Salemi M, Galligan DC, *et al.*: Extensive HIV-1 intra-host recombination is common in tissues with abnormal histopathology. *PLoS One* 2009;4(3):e5065.
- Delpont W, Poon AF, Frost SD, and Kosakovsky Pond SL: Datamonkey 2010: A suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics* 2010;26(19):2455–2457.
- Shriner D, Rodrigo AG, Nickle DC, and Mullins JI: Pervasive genomic recombination of HIV-1 in vivo. *Genetics* 2004; 167(4):1573–1583.
- Coleman CM and Wu L: HIV interactions with monocytes and dendritic cells: Viral latency and reservoirs. *Retrovirology* 2009;6:51.
- Tobian AA, Gray RH, and Quinn TC: Male circumcision for the prevention of acquisition and transmission of sexually transmitted infections: The case for neonatal circumcision. *Arch Pediatr Adolesc Med* 2010;164(1):78–84.
- Donoval BA, Landay AL, Moses S, *et al.*: HIV-1 target cells in foreskins of African men with varying histories of sexually transmitted infections. *Am J Clin Pathol* 2006;125(3): 386–391.
- Swingler S, Mann AM, Zhou J, Swingler C, and Stevenson M: Apoptotic killing of HIV-1-infected macrophages is subverted by the viral envelope glycoprotein. *PLoS Pathog* 2007;3(9):1281–1290.
- Kahn JO and Walker BD: Acute human immunodeficiency virus type 1 infection. *N Engl J Med* 1998;339(1):33–39.
- Huang W, Eshleman SH, Toma J, *et al.*: Coreceptor tropism in human immunodeficiency virus type 1 subtype D: High prevalence of CXCR4 tropism and heterogeneous composition of viral populations. *J Virol* 2007;81(15):7885–7893.

Address correspondence to:

Thomas C. Quinn

Johns Hopkins University School of Medicine

Rangos Building, Room 530

855 North Wolfe Street

Baltimore, Maryland 21205

E-mail: tqinn2@jhmi.edu