Exploring proviral evolutionary dynamics in the Women's Interagency HIV Study

Aniqa Shahid[1,2], Signe MacLennan[1], Natalie N. Kinloch[1,2], Hanwei Sudderuddin[2], Christina Ochsenbauer[3], Margaret Fischl[4], Igho Ofotokun[5], Adaora Adimora[6], Stephen Gange[7], Brett Williams[8], Bradley Aouizerat[9], Mark Kuniholm[10], Seble Kassaye[11], Kathryn Anastos[12], Zabrina L. Brumme[1,2]

[1]Faculty of Health Sciences, Simon Fraser University, Burnaby, British Columbia, Canada. [2]British Columbia Centre for Excellence in HIV/AIDS, Vancouver, British Columbia, Canada.
[3]Department of Medicine and CFAR, University of Alabama at Birmingham, Birmingham, Alabama, USA. [4]Department of Medicine, University of Miami School of Medicine, Miami, Florida, USA. [5]Division of Infectious Diseases, Emory University School of Medicine, Atlanta, Georgia, USA. [6]Departments of Medicine and Epidemiology, University of North Carolina School of Medicine, UNC Gillings School of Global Public Health.[7]Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, USA. [8]Department of Internal Medicine, Division of Infectious Disease, Rush University Medical Center, Chicago, Illinois; Ruth M. Rothstein Core Center of Cook County, Chicago, Illinois, USA. [9]Bluestone Center for Clinical Research, New York University, New York, NY, USA. [10]Department of Epidemiology and Biostatistics, University at Albany, State University of New York, Rensselaer, New York, USA. [11]Division of Infectious Diseases and Tropical Medicine, Georgetown University, Washington, DC, USA. [12]Albert Einstein College of Medicine, Bronx, New York, USA.

CONTACT:
Aniqa Shahid (ashahid@sfu.ca)
Zabrina Brumme (zbrumme@sfu.ca)

PRESENTED AT IAS 2021 – the 11th IAS Conference on HIV Science|18-21 JULY 2021
E-Poster supported by IAS 2021 Scholarship

BACKGROUND

Females have historically been underrepresented in HIV cure research, particularly in the studies of HIV reservoir evolutionary dynamics, though this is now changing[1,2]. Nevertheless, the majority of reservoir dynamics studies featuring females have focused on females with HIV subtype C. Females with HIV subtype B still remain underrepresented.

Women's Interagency HIV Study (WIHS), a multi-center, prospective, observational cohort established in 1993 across the USA, can help bridge this knowledge gap. WIHS features participants for whom longitudinal pre-cART plasma specimens dating back to transmission, along with PBMCs sampled following cART initiation, are available, thus offering us the opportunity to study HIV evolution and reservoir dynamics in females, and specifically, allow us **to infer HIV proviral ages and temporal stability during cART, and to compare ages of rebound viruses to the overall proviral pool.**

METHODS

Study participants: Three females with HIV subtype B were studied from the WIHS cohort.

HIV-1 amplification and sequencing: Pre-cART single-genome-amplified (SGA) plasma HIV RNA (env-*gp120*) sequences were previously published by Dapp *et al*[3]. In the present study, genomic DNA was extracted from 10 million PBMCs, using QIAamp DNA mini kit (Qiagen). SGA of HIV proviral DNA (env-*gp120*) was performed by limiting dilution approach. DNA extracts were then endpoint diluted such that ~25 to 30% of the PCR reactions would yield an amplicon. Amplicons were sequenced on an ABI 3730xl DNA sequencer. Chromatograms were base called using Sequencher v5.4.6 (Gene Codes).

Sequence alignment and QC: Sequences were aligned using MAFFT (v7.471), and manually inspected and edited using AliView (v1.26). Sequences that contained defects (e.g., large deletions/insertions), mixed bases, hypermutated sequences (identified using Hypermut, v2.0), or within-host recombinants (identified using RDP4 Beta 95) were discarded. Identical sequences were discarded to infer within-host phylogenies.

Phylogenetic inference of proviral ages and stability during cART: Each participant's plasma HIV RNA sequences collected during unsuppressed viremia, along with proviral sequences collected during suppressive cART, were used to infer a maximum-likelihood within-host phylogeny. This phylogeny was then rooted with root-to-tip regression function to maximize the correlation between the divergence from root and the sample collection dates of the pre-cART sequences, where this root represented the transmitted-founder virus. The regression was then used to convert root-to-tip distance of each proviral sequence to its integration date, with associated 95% confidence intervals (CI)[4].

Sampling timeline

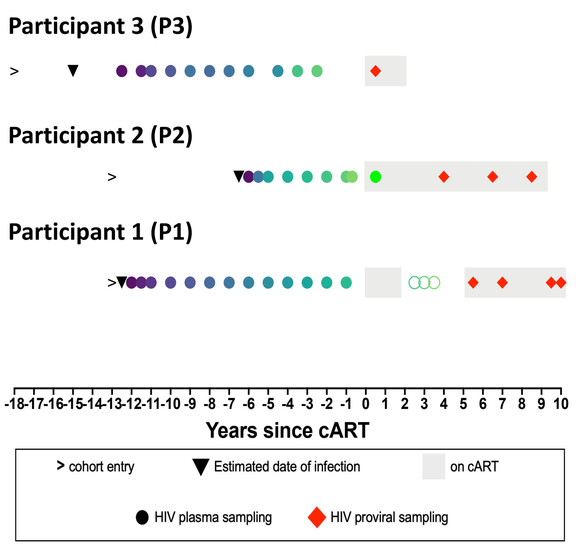

Figure 1: For each participant, the cohort entry is shown with right arrow, estimated date of infection as inverted triangle, plasma HIV RNA sampling in circles, PBMC-based proviral sampling in red diamonds, and suppression on cART in grey, with time in years since initiation of cART on x-axis. Participants initiated cART in chronic infection; total follow-up time since infection was 22 years for participant 1 (P1), 15 years for participant 2 (P2), and 16 years for participant 3 (P3). For P1 P2, and P3, pre-cART longitudinal plasma HIV RNA (env-*gp120)* were previously collected using SGA from 13, 9, and 11 timepoints as shown in colored circles, respectively[3]. For the present study, we used SGA to further characterize proviral DNA sampled on cART for the three participants, at the timepoints shown in red diamonds. Specifically, for P1, proviruses were sampled 4 times over 5 years, three times over 4 years for P2, and once for P3. P1 had a period of plasma viremia after their initial regimen was interrupted. During this time, the participant had three plasma samples taken, shown as colored open circles. HIV RNA sequences were collected using SGA from these three timepoints.

RESULTS

Participant 1

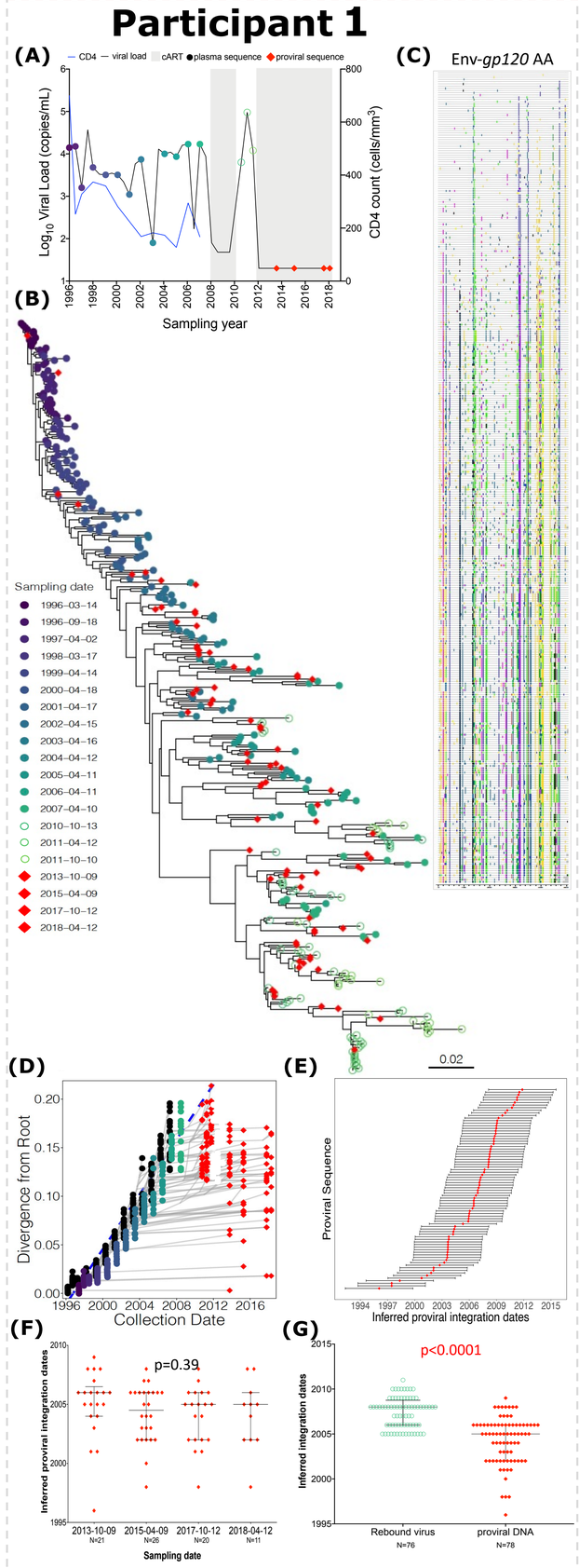

Figure 2: (A) P1 tested positive for HIV in 1996 and did not initiate cART until 2008, shown in grey, which was interrupted in 2010. A viremic episode was observed between October 2010 to October 2011, which was later suppressed to undetectable levels in 2012, until the last follow up in April 2018. **(B)** Maximum-likelihood within-host phylogeny (with it's legend on the left) was inferred from a total of 286 intact and unique pre-cART HIV RNA sequences from 16 plasma time points between March 1996 to October 2011, and 78 proviral sequences collected from four PBMC timepoints between October 2013 to April 2018. **(C)** A highlighter plot is shown based on amino acid sequences, which corresponds to the within-host phylogeny shown in B. The top sequence serves as the reference, where colored ticks in sequences beneath it denote non-synonymous substitutions with respect to the reference **(D)** Once the maximum-likelihood phylogeny was generated, these were rooted with root-to-tip regression function to maximize the correlation between the divergence from root and the sample collection dates of the pre-cART sequences, shown as blue dashed line. This line was then used to convert the root-to-tip distances of proviral sequences samples on cART to their integration dates. The light grey lines represent the underlying evolutionary relationships between HIV sequences sampled over time. **(E)** Point estimates and the associated 95% confidence intervals of the integration dates of sampled proviral sequences inferred from the regression function are shown in red and grey. **(F)** The inferred integration dates of proviral sequences sampled over 5 years on cART. **(G)** Comparison of the inferred integration dates of rebound viruses emerging after 1yr of suppressive cART (open green circles) to overall proviral pool (red diamonds).

Proviral sequences were archived throughout infection, recapitulating much of the viral diversity that arises during untreated HIV-1 infection. Unique proviruses initially sampled on cART were an estimated 8 years old on average with the oldest being 16 years of age, with most sequences dating to chronic phase. The proviral integration dates remained stable over 5 years on cART (*p*-value=0.39).

Rebound viruses that emerged after 1 year on cART were significantly younger than the persisting proviral pool (p-value<0.0001), suggesting that these viruses emerged from a recently established replication-competent reservoir, as compared to the older, persisting proviral pool.

Participant 2

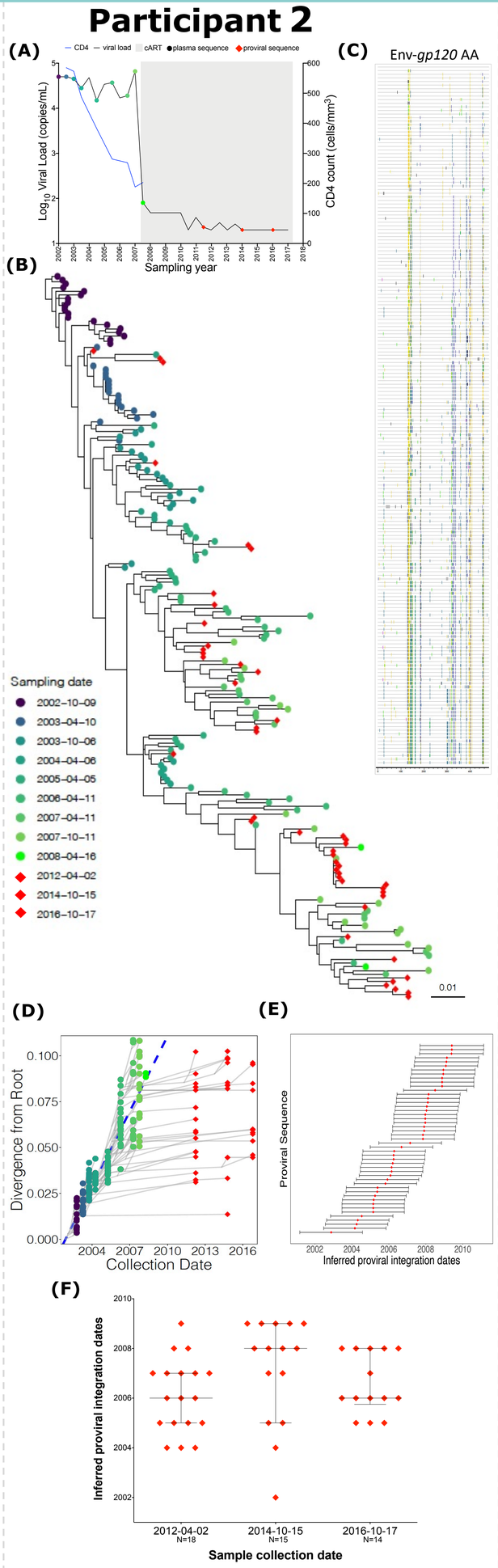

Figure 3: (A) P2 contracted HIV in late 2002, did not initiate cART until 2008, and viremia remain suppressed on cART until last follow-up. **(B)** Maximum-likelihood within-host phylogeny was inferred from a total of 147 intact and unique pre-cART sequences from 9 plasma time points and 47 proviral sequences collected from three PBMC timepoints between 2012 to 2016. **(C)** A highlighter plot is shown based on amino acid sequences, which corresponds to the within-host phylogeny shown in B. The top sequence serves as the reference, where colored ticks in sequences beneath it denote non-synonymous substitutions with respect to the reference **(D)** Once the maximum-likelihood phylogeny was generated, these were rooted with root-to-tip regression function to maximize the correlation between the divergence from root and the sample collection dates of the pre-cART sequences, shown as blue dashed line. This line was then used to convert the root-to-tip distances of proviral sequences samples on cART to their integration dates. The light grey lines represent the underlying evolutionary relationships between HIV sequences sampled over time. **(E)** Point estimates and the associated 95% confidence intervals of the integration dates of sampled proviral sequences inferred from the regression function are shown in red and grey. **(F)** The inferred integration dates of proviral sequences sampled over 4 years on cART.

P2's proviruses dated primarily to chronic infection, with relatively few dating to early infection. Unique proviruses were 5 years old on average with the oldest being 8 years of age at initial sampling, with most dating to chronic phase; proviral integration dates remained stable over 4 years.

Participant 3

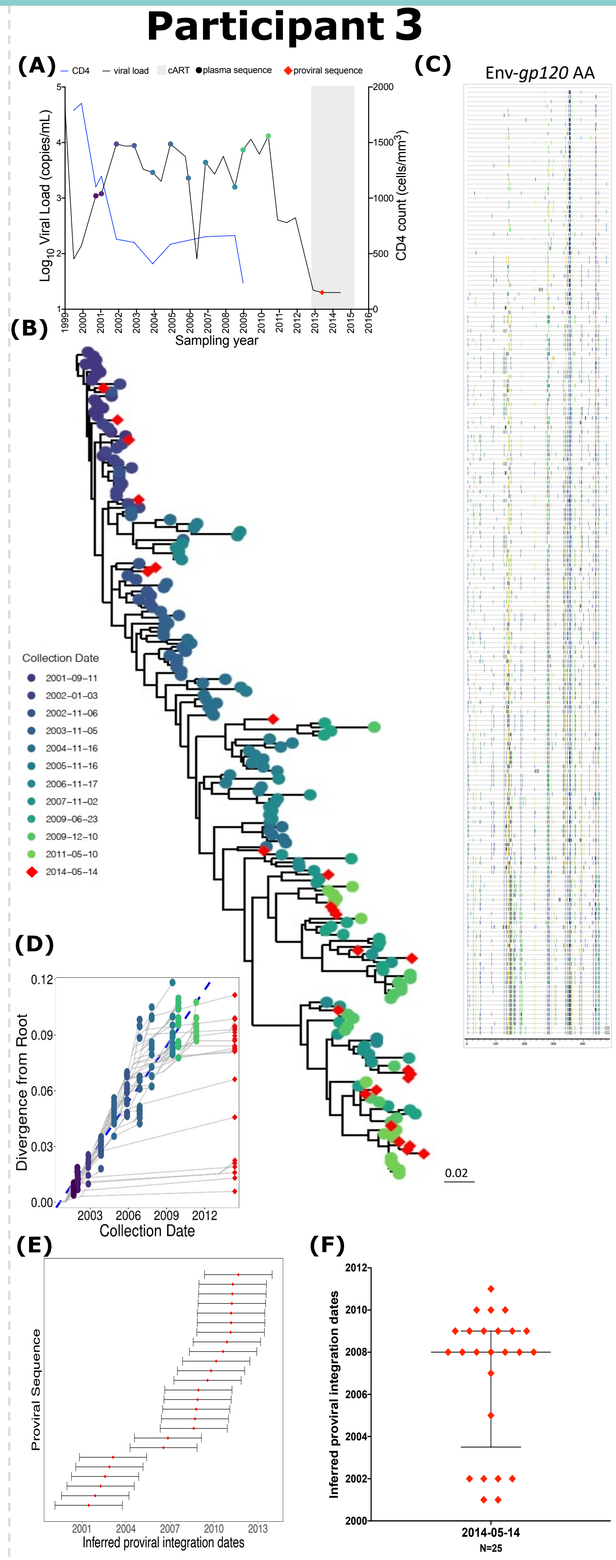

Figure 4: (A) P3 contracted HIV in late 1999 and did not initiate cART until much later in 2013. Maximum-likelihood within-host phylogeny was inferred from a total of 182 intact and unique pre-cART sequences collected from 11 plasma time points and 25 proviral sequences collected from one PBMC timepoint in 2014. **(C)** Similar to figure 3. **(D)** Similar to figure 3. **(E)** Similar to figure 3. **(F)** The inferred integration dates of proviral sequences sampled on cART.

P3's inferred proviral integration dates spanned their pre-cART history more evenly. Unique proviruses were an estimated 7 years old on average with the oldest being 13 years of age, with many dating to early infection.

CONCLUSIONS

- Consistent with reports in males and females with HIV subtype C, proviruses sampled on-cART spanned a wide age range, with some dating back to transmission. This is consistent with ongoing archiving of diverse proviral lineages pre-cART, and their subsequent persistence.
- In two participants, proviruses sampled on cART dated primarily to chronic infection, with relatively few dating to early infection. In the third, integration dates spanned their pre-cART history more evenly. We did not observe any participants whose proviral pools were dramatically skewed towards integration dates immediately before cART.
- In the one participant who interrupted cART, the viruses that emerged in plasma following this interruption were significantly younger than the total proviral pool. This is consistent with more rapid turnover of the replication-competent reservoir as compared to the total persisting proviral pool, which is largely defective[5].
- In the two participants with longitudinal proviral sampling of proviruses on-cART, the stability of these integration dates is consistent with very slow decay of the overall proviral pool during therapy[6,7].

REFERENCES

1. Abrahams *et al*, *Transl Med*. 2019, PMID: 31597754
2. Brooks *et al*, *PLoS Pathog*. 2020, PMID: 32492044
3. Dapp *et al*, *PloS One*, 2017, PMID: 29045410
4. Jones *et al*, *PNAS, 2018, PMID:* 30185556
5. Bruner *et al*, *Nat Med*. 2016, PMID: 27500724
6. Golob *et al*, *AIDS*. 2018, PMID: 30005008
7. Peluso *et al*, *CI Insight*. 2020, PMID: 32045386