Longitudinal pre-therapy sequencing of HIV-1 for use in phylogenetic characterisation of rebound virus

Lynn Tyers,[1]Matt Moeser A[2]nsie Yssel, M[1]elissa-Rose Abrahams, Sar[1]ah Beth Joseph, Nige[1] Garrett, Ronald[2] Swanstrom, Carolyn[2] Williamson1,3

[1]University of Cape Town Faculty of Health Sciences, Division of Medical Virology, Institute of Infectious Disease and Molecular Medicine, Cape Town, South Africa, 2University of North Carolina at Chapel Hill, Chapel Hill, United States, 3Centre for the AIDS Programme of Research in South Africa (CAPRISA), Durban, South Africa

Abstract

Although antiretroviral therapy (ART) has been successful in suppressing viral replication in people living with HIV (PLWH), the virus persists in latently infected cells and can re-emerge when therapy is interrupted. We previously observed, in women from the CAPRISA 002 acute infection cohort, Kwa-Zulu Natal, that the majority of replication-competent viruses induced from this latent reservoir were most closely related to viruses circulating in the blood within the year preceding ART initiation. While rebound and induced viruses originate from the same pool and overlap, they are not necessarily the same population. We aimed to determine when the rebound virus was seeded into the latent reservoir for nine women from the same cohort experiencing a viral rebound (>1000 RNA copies/ml) at least one-year post-ART initiation, using phylogenetic methods. Three gene regions, two in *env* and one in *nef* were amplified and sequenced from plasma-derived viral RNA on average every six months from acute infection to ART initiation, and at the time of rebound, using Illumina Miseq with Primer ID (PID). Relatedness of the rebound virus to pre-ART viruses was determined, using maximum-likelihood-like trees with patristic, clade and placement methods using a custom dating pipeline (https://primer-id.org/ogv). All three gene regions were successfully amplified and sequenced for the nine women. On average, 354 PID consensus sequences (range 18-4760) per time-point per region pre-ART and 293 PID consensus rebound virus sequences (range 3-2383) at rebound were generated for each participant. Ambiguity in the phylogenetic trees, likely resulting from multivariant infection and divergent lineages was identified as a challenge to our timing approach. Determining the best method/s for producing trees using optimal longitudinal alignments and for addressing the ambiguity that multivariant infection and divergent or recombinant lineages create is imperative.

Background

- The latent HIV-1 reservoir remains the greatest barrier to a cure in people living with HIV (PLWH), despite effective antiretroviral therapy (ART).
- Interruption of suppressive ART, typically results in rapid viral rebound from the latent reservoir.
- The viruses emerging during rebound only represent a fraction of this latent reservoir.
- Furthermore the viruses stimulated *in vitro* using methods such as quantitative viral outgrowth assays (qVOA) and those that emerge during rebound/treatment interruption overlap, but are distinct1,2.
- Previous studies have shown that the majority of the reservoir as measured by proviral or outgrowth virus sequencing was established near the time of ART initiation3,4,5. We aim to evaluate if this finding is consistent with rebound viruses.

Significance

- Analysis of the temporal origins of rebound virus will provide a better understanding of the fraction of the latent reservoir capable of re-emerging during treatment interruption.
- Improved understanding of the HIV-1 reservoir dynamics could inform strategies for restricting the formation and persistence thereof.

Methods

Study Participants

- 9 chronically infected PLWH from the CAPRISA002 cohort who were virologically suppressed for a minimum of 1 year, followed by a viral rebound were identified.

Sequencing of pre-ART and rebound virus

- Illumina MiSeq deep sequencing with Primer ID was used to generate sequences from viral RNA derived from plasma, sampled every 6 months pre-therapy as well as at rebound (Figure 1A).

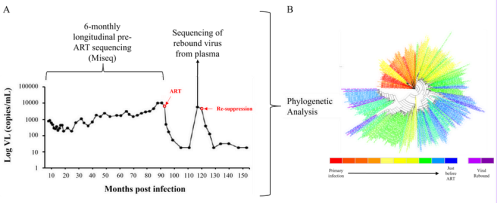

Figure 1: Overview of pre-ART and rebound virus sequencing approach.

- Timing of viral entry of rebound virus into the reservoir was done using maximum-likelihood-like trees with patristic, clade and placement methods using a custom dating pipeline (https://primer-id.org/ogv) (Figure 1B). Three genomic regions (V1V2, V4V5 and *nef,* HXB2:
- 8678-9156)

were amplified in a multiplexed reaction from plasma-derived RNA (Figure 2). Gene regions were selected based on previous work indicating that they produced strong phylogenetic signal5.

- Less than 18 consensus sequences per time-point pre-ART was considered a failed run.

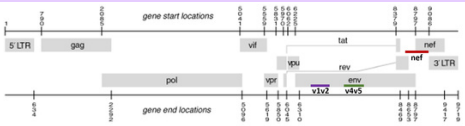

Figure 2: Regions amplified in the multiplexed approach for Illumina Miseq sequencing with Primer ID.

- Raw sequence reads were processed using a custom pipeline written in Python and R programming languages (https://github.com/HIVDiversity/NGS_processing_pipeline; DOI: 10.5281/zenodo.3372202).
- Quality filtering, merging of overlapping paired-end reads and Primer ID processing was performed using the MotifBinner2.R program (https://github.com/HIVDiversity/MotifBinner2; DOI:

541 10.5281/zenodo.3372204).

Data Analysis

Multiplexed sequence depth and timing of rebound viruses

- Participants were infected for an average of 5 years (range 2.6-8.2 years) and experienced a viral rebound after an average of 2.5 years on ART (range 1.3– 4 years).
- Of the participants 4 were identified as potential multivariant infections (44.4%), 4 as single variant infections (44.4%) and 2 were undetermined due to lack of early sampling (22.2%).
- Plasma viral load at rebound ranged from 1250 copies/ml to 640721 copies/ml.
- On average, 354 consensus sequences (range 18-4760) per time-point per region were generated pre-ART (Figure 3) and 293 rebound virus consensus sequences (range 3-2383) were generated per participant (Figure 4).

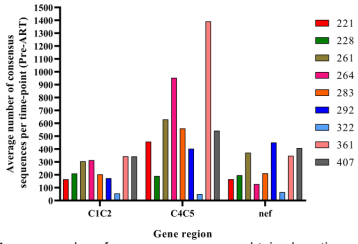

Figure 3: Average number of consensus sequences obtained per time-point for each participant.

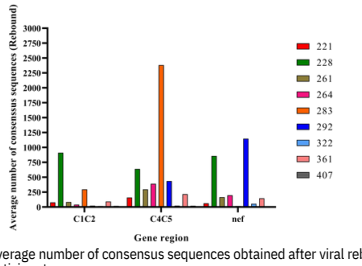

Figure 4: Average number of consensus sequences obtained after viral rebound for each participant.

- Tree R2 values ranged from 0.036 to 0.712. Trees with R2 values <0.4 were considered to have insufficient evolutionary signal/ladder-like structure for timing analyses.

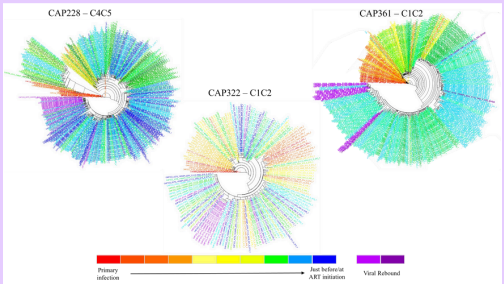

Figure 5: Representative maximum likelihood-like trees used to estimate rebound virus timing. Sampling time-points pre-ART are colour coded from red (study enrolment to blue (year preceding ART). Rebound virus sequences are shown in purple. The hypervariable loop regions of *env* were removed.

- Multivariant infection affected tree structure in various ways as depicted by the trees for CAP228-C4C5 and CAP322-C1C2 (Figure 5).
- CAP228-C4C5 illustrates an instance where two lineages diverge and CAP322-C1C2 illustrates an instance where sequences from longitudinal time-points are mixed with two distinct primary infection clusters present. Both of these situations negatively impact tree structure (R2=0.256 and 0.21, respectively), which may affect the timing estimates made by our pipeline.

Goals

- To determine the best approach to address the ambiguity introduced by multivariant infection and potential recombination.
- To derive a single timing estimate for the rebound virus using multiple gene regions where timing for each gene region are not concordant.

Acknowledgements

All participants from CAPRISA 002 acute infection cohort

National Institute of Allergy and Infectious Diseases

References

1.Pannus, P., Rutsaert, S., De Wit, S., Allard, S. D., Vanham, G., Cole, B., Nescoi, C., Aerts, J., De Spiegelaere, W., Tsoumanis, A., Couttenye, M. -M., Herssens, N., De Scheerder, M.-A., Vandekerckhove, L. and Florence, E. Rapid viral rebound after analytical treatment interruption in patients with very small HIV reservoir and minimal on-going viral transcription. *J Int AIDS Soc*. 2020; 23(2):e25453

2.Salantes DB, Zheng Y, Mampe F, Srivastava T, Beg S, Lai J, Li JZ, Tressler RL, Koup RA, Hoxie J, Abdel-Mohsen M, Sherrill-Mix S, McCormick K, Overton ET, Bushman FD, Learn GH, Siliciano RF, Siliciano JM, Tebas P, Bar KJ. HIV-1 latent reservoir size and diversity are stable following brief treatment interruption. J Clin Invest. 2018 Jul 2;128(7):3102-3115. doi: 10.1172/JCI120194. Epub 2018 Jun 18. PMID: 29911997; PMCID: PMC6026010.

3.Brodin J, Zanini F, Thebo L, et al. Establishment and stability of the latent HIV-1 DNA reservoir. *Elife*. 2016:1-15. doi:10.7554/eLife.18889

4.Abrahams M-R, Joseph SB, Garrett N, et al. The Replication-Competent HIV-1 Latent Reservoir is Primarily Established Near the Time of Therapy Initiation. *Sci Transl Med.* 2019:11(513). doi:10.1126/scitranslmed.aaw5589

5.Joseph SB, Abrahams M-R, Tyers L, et al Temporal dynamics of reservoir formation are not strongly influenced by replication competency or cellular tropism of HIV-1 proviruses [Manuscript in preparation].