NIH NATIONAL CANCER INSTITUTE

HIV Dynamics and Replication Program
NCI at Frederick

DIFFERENTIAL TRANSCRIPTIONAL LEVELS OF HIV-1 NEAR FULL-LENGTH AND HIGHLY DELETED PROVIRUSES

393

Thuy Nguyen[1], Lindsey Adams[1], Mary−Elizabeth Zipparo[1], Annie Glassey[1], Erin Madeen[1], Ulisses Santamaria[2], Catherine A Rehm[3], Jessica Earhart[1], Wei Shao[1], Chuen−Yen Lau[1], Frank Maldarelli[1]

1 HIV Dynamics and Replication Program, Frederick (United States), 2 Leidos Biomedical Research, Inc., Frederick National Laboratory For Cancer Research - Frederick (United States),
3 Clinical Research Section, National Institute Of Allergy And Infectious Diseases, Bethesda (United States)

BACKGROUND

HIV persistence is the principal obstacle to cure. The population of HIV-infected cells that harbor genetically intact or defective proviruses is dynamically changing during antiretroviral therapy (ART) [1,2]. Intact proviruses decline more rapidly than defective proviruses on suppressive therapy and the mechanisms for this are uncertain [1,2]. Proviral transcriptional activity might be one of the key factors contributing to the differential decay rate of proviral populations [3].

In this study, we aimed to compare and identify viral determinants affecting transcriptional levels of intact, near full-length (NFL) defective, and total proviruses after long-term ART.

METHODS

Peripheral blood mononuclear cells (PBMC) were collected from 11 HIV-suppressed patients at pretherapy and after long-term ART.

We performed single genome sequencing (SGS) of proviruses>7kb by Illumina MiSeq of pretherapy and long-term ART samples and characterized proviral genetic intactness by analyzing HIV open reading frames and regulatory genes [4,5].

The total proviral *gag* from pretherapy and long-term ART samples and cell-associated RNA (caRNA, unspliced transcript) populations from long-term ART samples were obtained by SGS of 1.3kb *gag* region.

We obtained 1434 sequences without stop codons on *gag*, ranging from 0-60 NFL (both intact and defective), 5-42 total *gag*, and 3-39 caRNA sequences/timepoint. We performed the analyses on *gag* sequences as following:

i) **Average pairwise distance** (APD) (MEGAX) to measure the genetic diversity of viral populations.
ii) **Approximate maximum-likelihood phylogenetic** analysis (FastTree 2.1) to visually evaluate the relationship of DNA and RNA populations.
iii) **Tests for population structure** (tree-based method Slatkin-Maddison) to statistically evaluate the structure of DNA and RNA populations.
iv) **Net average distance** of cell-associated *gag* RNA populations to proviral *gag* populations to calculate the divergence of caRNA from proviral populations.
v) **HLA-associated mutations** on *gag* RNA and DNA to evaluate the impact of HLA escaped mutations on the proviral transcriptional activity.

RESULTS

1. Participant characteristics

Characteristic	Median (IQR)
Age (n=11)	61 (59.0-68.5)
Sex, male	9
Log10 (Pretherapy viral load, copies/mL)	4.68 (4.37-4.96)
Nadir CD4 (cells/mm3)	222 (160-241)
CD4 count (cells/mm3)	
Pretherapy (n=11)	273 (165-490)
Long-term ART	530 (381-687)
Suppression time on ART (to the latest sample)	18.0 (13-19.0)
Time from pretherapy sample to ART initiation (days)	1 (1-21)

The length of proviruses is important for their transcriptional activities. Near-full length (NFL) proviruses, both genetically intact and defective, are equally transcriptionally silent. Our results suggest the advantage in transcription elongation of highly defective proviruses over NFL proviruses.

2. No change in the genetic diversity of proviral populations over long-term ART

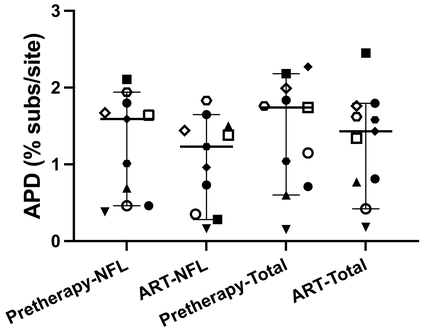

Fig1: Genetic diversity (measured as APD) of proviral populations prior to and following long-term ART

- NFL and total proviral *gag* populations at pretherapy had a median APD of 1.59% and 1.74%, respectively and did not change over long-term ART.

3. No significant difference in the genetic diversity of caRNA vs proviral populations

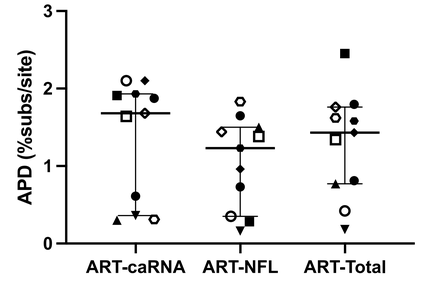

Fig2: Genetic diversity (measured as APD) of caRNA, NFL, and total proviral populations

- The median APD (IQR) of caRNA, NFL, and total *gag* populations after long-term ART were 1.68 (0.36- 1.93), 1.23 (0.35-1.50), 1.43 (0.77-1.76), respectively and were not significantly different.

4. CaRNA genetically closer to total *gag* than NFL populations

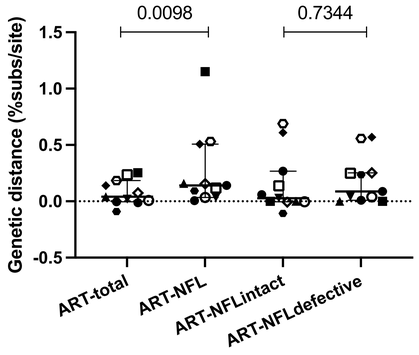

Fig4: Net genetic distance (% subs/site) of caRNA to total and vs NFL proviruses and to intact vs defective proviruses.

- CaRNA are **genetically closer** to total *gag* than NFL populations with the median (IQR) divergence of 0.04% (0.004-0.18) vs 0.14 (0.03-0.5) (p=0.0098).
- Interestingly, caRNA populations **are not differently divergent** from intact or NFL defective proviruses, 0.03 (0.005-0.27) vs 0.09 (0.009-0.25) (p=0.734).

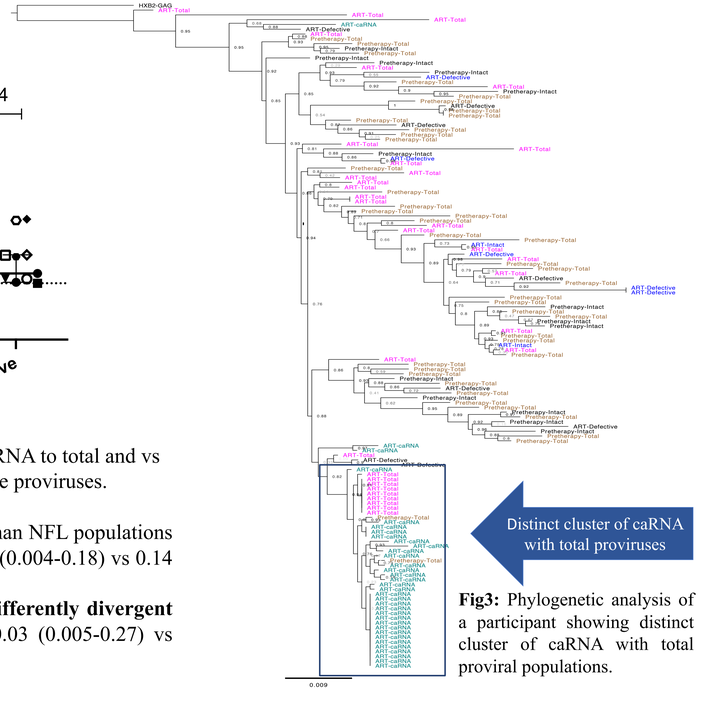

Fig3: Phylogenetic analysis of a participant showing distinct cluster of caRNA with total proviral populations.

5. Selection of HLA-associated escaped epitopes in proviruses vs caRNA *gag* populations in a small group of participants

Table 2: > 1.5-fold difference of HLA-associated escaped epitopes

Participant	HLA	HXB2 start on gag	Wildtype epitope	Variant epitope	Variant epitope frequency in total proviruses (%)	Variant epitope frequency NFL proviruses (%)	Variant epitope frequency in caRNA (%)
1	B*07:02	407	APRKKGCWK	APRKrGCWK	96.0	54.5	100
2	A*02:23	77	SLYNTVATL	SLfNTVATL	75.0	92.3	53.9
3	A*02:31	77	SLYNTVATL	SvYNTVATL	20.0	50.0	69.2
4	A*02:74	433	FLGKIWPSYK	FLGKIWPSrK	15.8	22.2	4.0
				FLGKIWPScK	26.3	11.1	52.0
5	A*30:02	76	RSLYNTVATLY	kSLYNTVATLY	46.7	79.4	33.3
6	A*02:01	77	SLYNTVATL	SLYNTiATL	27.3	57.1	0
				-----YKLK	100	0	90.9
7	B*08:01	24	GGKKKYKLK	-------LK	0	57.9	0
				---------	0	42.1	9.1

Overall, the frequencies of wild-type (capital) and escaped mutations (lowercase or "-" if deleted) on HLA-associated epitopes in *gag* were not significantly different in NFL vs total *gag* vs caRNA populations after long-term ART. However, 7 patients had evidence of selection for or against some epitopes in proviral *gag* compared to caRNA.

CONCLUSIONS

Transcriptional levels of intact and NFL defective proviruses are not significantly different.

However, total *gag* proviruses i.e highly deleted, are more transcriptionally active than NFL proviruses, and the HLA escape mechanism does not significantly contribute to this difference.

Our results suggest the advantage of highly defective proviruses in transcription elongation.

ACKNOWLEDGEMENT

Thank you to the study participants.

ADDITIONAL INFORMATION

Thuy Nguyen: thuy.nguyen2@nih.gov

REFERENCES

1. E M. Anderson et al, Viruses, 2020
2. MJ Peluso et al, JCI Insight, 2020
3. Kevin B Einkauf *et al*., Cell, 2022
4. Bonnie H *et al.,* Cell Rep, 2017
5. Sean C *et al*., PNAS, 2019