Systematic Characterization of Viral-Human Chimeric mRNA Transcripts in Donors Living With HIV-1

215

Peyton K. Smith[1], Nancy Francoeur[2], Yanqin Ren[1], Gintaras Deikus[2], Colin Kovacs[3], Robert Sebra[2], R. Brad Jones[1], Marina Caskey[4], Melissa Smith[5], Guinevere Q. Lee[1]
[1]Weill Cornell Medicine, New York, NY, USA, [2]Icahn School of Medicine at Mount Sinai, New York, NY, USA, [3]Maple Leaf Medical Clinic, Toronto, OT, CA,
[4]Rockefeller University, New York, NY, USA, [5]University of Louisville, Louisville, KY, USA

Contact information
Peyton Smith, pes4008@med.cornell.edu
Guinevere Lee, gul4001@med.cornell.edu

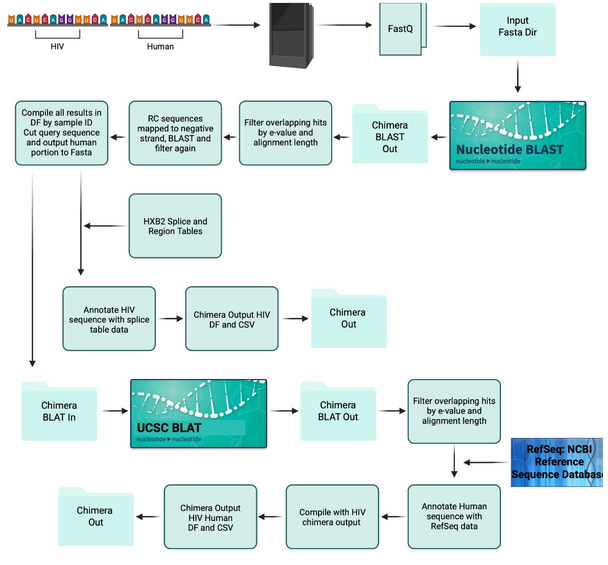
BACKGROUND

- HIV-1 transcription involves multiple splicing events. Previous studies (Cesana 2017, Liu 2020, Lee 2022 etc.) have reported chimeric transcripts containing both HIV and human mRNA portions.
- Chimeric transcripts are (i) joined by aberrant splicing event(s) or (ii) transcriptional readthrough.
- This may have implications in viral persistence, host gene dysregulation, and novel antigen production.
- Research question: What is the structural diversity, genomic distribution, and stimulation-dependence of HIV-host chimeric transcripts, and is their formation dependent on the proviral integration site?

METHODS

- mRNA was extracted from 8E5/LAV cell line (~70,000 cells; provirus integration in chr13) and primary CD4+ T cells (3 viremic donors, 3 aviremic donors; unstimulated, or stimulated with PMA/ionomycin, CD3/CD28, romidepsin).
- PacBio Iso-Seq sequencing (171 million (M) FLNC reads; 1.48M HIV-containing).
- Two-stage alignment (blastn → BLAT) with structural classification and GT/AG validation (±5 nucleotides).

Figure 1. Automated chimera identification pipeline. Full-length transcripts were aligned to HIV-1 (blastn/HXB2), human portions mapped to hg38 (BLAT), then structurally classified and validated for the presence of canonical GT/AG at splice junctions.



HIV-human chimeric RNAs are readily detected in CD4⁺ T cells from individuals living with HIV. They are rare in a clonal HIV cell line but are markedly enriched *in vivo* and strongly modulated by LRA stimulation. These transcripts reveal a previously underappreciated layer of host-virus interaction.

Figure 2. (A) HIV-human chimeric transcripts comprised 0.4% of HIV-containing reads in 8E5/LAV. (B) HIV-Human was the predominant configuration (72.9%). (C) HIV-Human configuration remained the predominant form among GT/AG-confirmed chimeras.

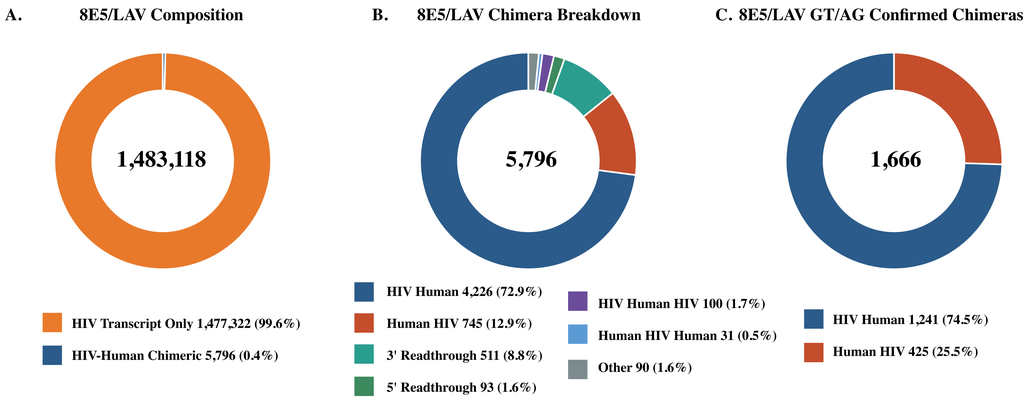


Figure 3. Genome-wide distribution of splice chimeras in 8E5/LAV. Despite confirmed single-site integration on chr13 (300/300 cells; 95% CI: 99.0–100%), 91.7% of splice chimeras mapped to non-chr13 chromosomes.

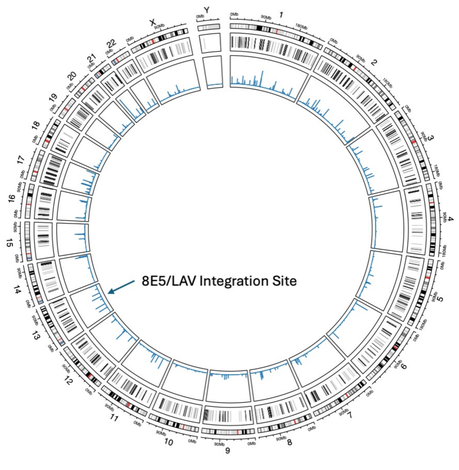
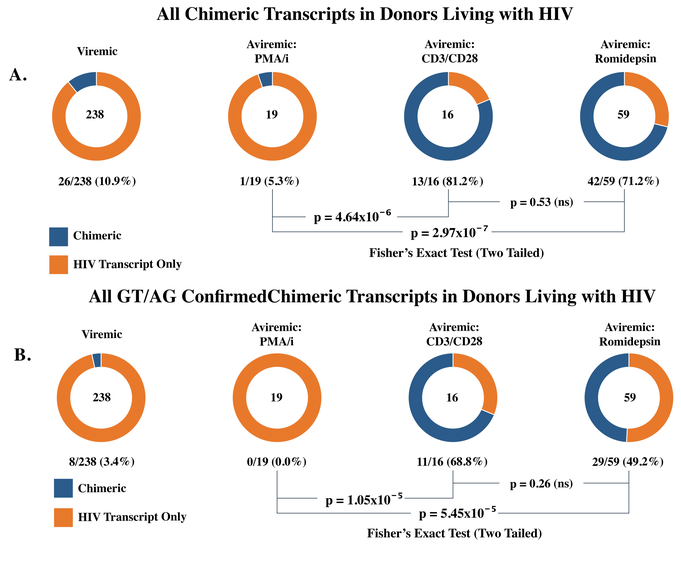


Figure 4. (A) Chimeras comprised 10.9% of HIV-containing reads in viremic donors. (B) Chimera frequencies varied by stimulation agents in aviremic donors: CD3/CD28 (81.3%) and romidepsin (71.2%) versus PMA/i (5.3%). (C) GT/AG-confirmed chimeras were most prevalent under romidepsin stimulation.



RESULTS

- **8E5/LAV cell line:** 5,796 chimeras in 1.48M HIV-containing reads (0.4%); 7 structural configurations; 1,666 GT/AG-confirmed (*Figure 2*). Splice chimeras distributed across all autosomes (91.7% not in chr13, *Figure 3*).
- **Clinical samples:** Chimera frequencies were higher in viremic donors than in 8E5/LAV ($p = 4.35 \times 10^{-29}$, Fisher's exact test). In aviremic donors, chimera frequencies varied by stimulation conditions (*Figure 4*). HIV-STAT5B chimeras were detected in two donors with protein coding potential through exon 2.

CONCLUSIONS

- We developed an automated pipeline enabling scalable chimera identification from long-read data.
- HIV-human chimeric RNAs are common in individuals living with HIV both with and without mitogen stimulation.
- Their occurrence highlights a novel mechanism of host-virus interaction that may shape viral persistence and immune response.
- Future work should define whether these transcripts encode proteins and/or alter host pathways.

ADDITIONAL KEY INFORMATION

Acknowledgments: This work is supported by NIH NIAID R01-AI162221 (GQL) and UM1-AI164565 (RBJ). Additionally, we would like to acknowledge our donors, without whom, this study would not be possible.

PLAIN LANGUAGE SUMMARY

When HIV becomes active, its genetic material can fuse with human genes to create hybrid RNA molecules. We found thousands of these hybrids with diverse structural configurations spread across the genome, and their frequencies depend on how infected cells are reactivated. These virus-human hybrid molecules may impact host biology and immune responses against the virus.

National Institutes of Health
REACH For the Cure — Research Enterprise to Advance a Cure for HIV
LEE LAB — Weill Cornell Medicine