Pervasive and non-random recombination in HIV genomes from Uganda

0093

Heather E Grant (1), Emma B Hodcroft (2), Deogratius Ssemwanga (3), John Kitayimbwa (4), Paul Kellam (5), Tulio de Oliveria (6), Rebecca Nsubuga (3), Sam Lycett (1) David L Robertson (7), Oliver Ratmann (8), Christophe Fraser (9), Deenan Pillay (10), Pontiano Kaleebu (3), **Andrew J Leigh Brown** (1) on behalf of the PANGEAHIV consortium

1) University of Edinburgh, 2) University of Basel, 3) MRC/UVRI/LSHTM Unit Uganda, 4) Makerere University 5) EMBL-EBI, 6) Africa Health Research Institute, 7) University of Glasgow 8) Imperial College London 9) Oxford University 10) University College London

BACKGROUND

- Subtypes A1 + D co-circulating in Uganda for 50+ years
- Recombination occurs during co-infection. Co-infection rates are hard to estimate, but is seen in many cohorts
- Study recombination of genomes in a population setting

Q1. Are there any similar recombination patterns?
Q2. Are there any hotspots of recombination?
Q3. Are there any circulating recombinant forms?

METHODS

- Next Generation sequencing used to obtain near full length genomes (n=465) from across southern Uganda between 2011 and 2017 by PANGEA-1 (and additional 1007 genomes between 5 and 8kb)
- SCUEAL used to find subtypes and breakpoints
- Generalised linear model used to test significance of breakpoint location by genome region
- Transmitted breakpoints detected using a window based linkage analysis

RESULTS

- 50% “pure” genomes, 50% inter-subtype recombinants (Fig 1)
- Broad-scale pattern whereby C2-TM, a section of envelope (C2 of gp120 to transmembrane region (TM) of gp41) is inherited intact in many recombinants, in both directions i.e. (A, D, A and D, A, D) (Fig 2)

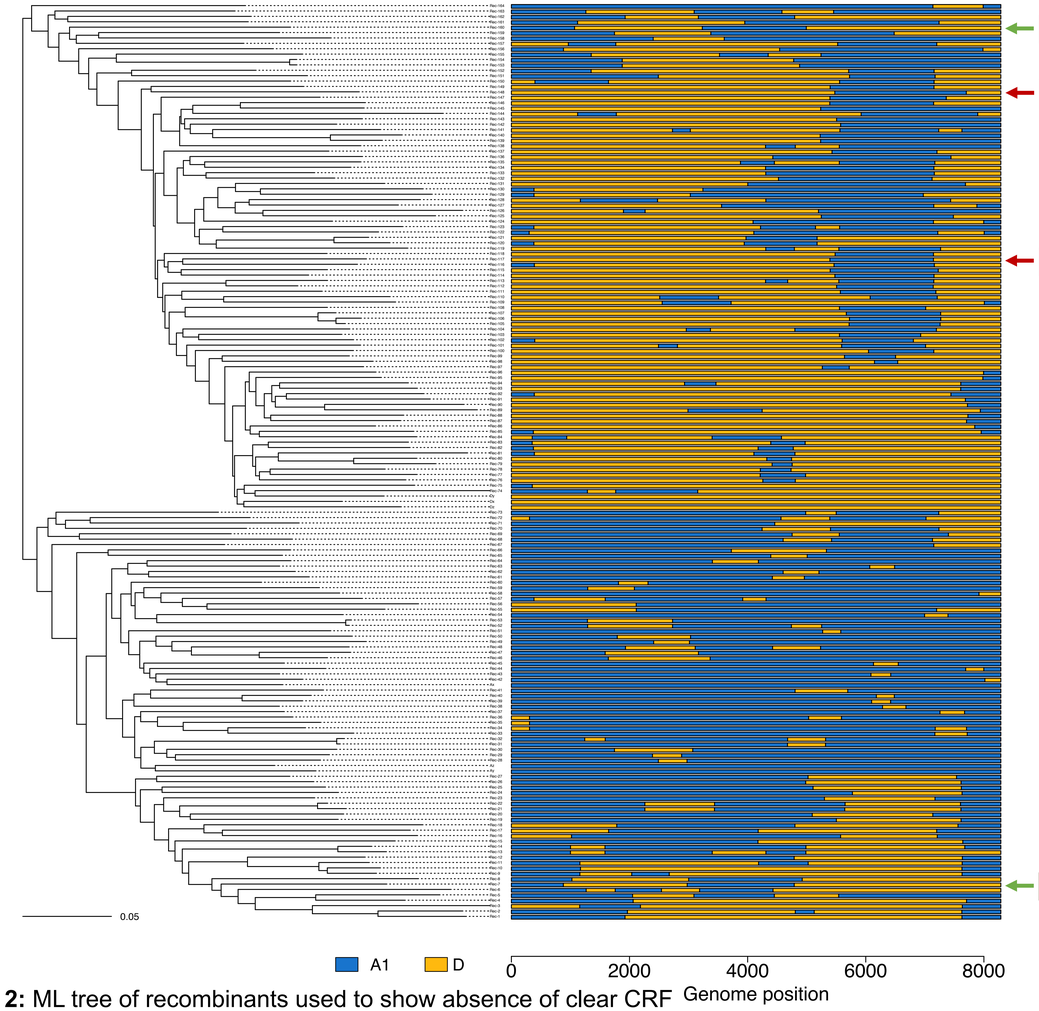
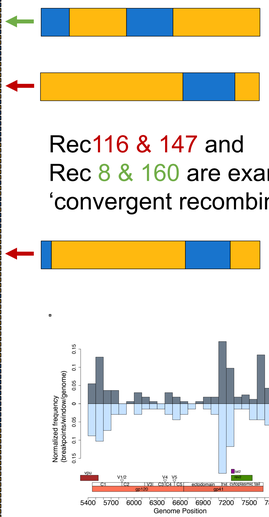


Fig 2: ML tree of recombinants used to show absence of clear CRF



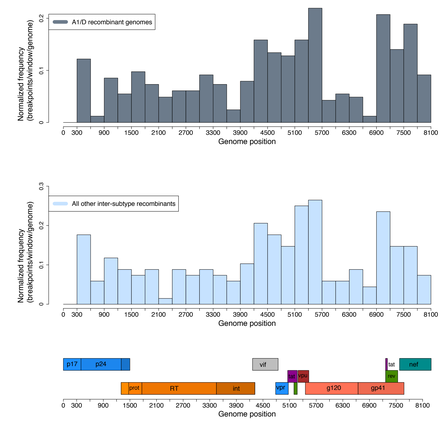


Fig 4: Recombination along the genome in 300 bp windows and (again close up in the envelope region in 100bp windows)

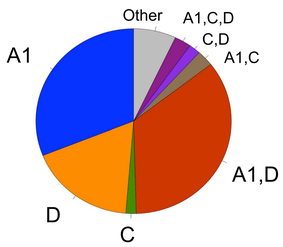


Fig 1: Subtype distribution

- A tree of recombinants violates a key assumption of phylogenetics, but any recently formed CRF should form a cluster amongst a background of other sequences, which they do not. In fact, similar recombination patterns are found in different parts of the tree suggesting ‘convergent recombination’ (Fig 2). A pairwise sliding window linkage analysis (Fig 3) showed only a small number of genomes had multiple linked windows (17 linked pairs / 164 recombinants) indicating few recombinants with recent common ancestry.
- Higher levels of recombination in the accessory genes than structural genes *gag-pol* and lower in C2-TM (GLM, p<0.001). Distribution largely mirrored in A1/D recombinants (n=164) and all “other” recombinants (n=68) (Fig 4)

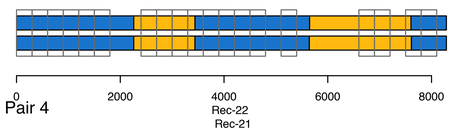


Fig 3: Example pair of genomes linked by a distance of less than 2% distance (TN93) in multiple 300 bp windows

CONCLUSIONS

- Misfolding of proteins or RNA cause functional constraints. Envelope is needed to gain entry into new host cell
- Higher co-infection and more complex transmission network in Uganda than generally thought?
- Traditional phylodynamic studies assume a lack of recombination, an assumption that may be violated in many cases
- Difficult to distinguish ancient CRFs from sequential recombination or convergent evolution

A1.Envelope region C2 –TM often inherited intact
A2. Higher recombination in accessory regions, and lower in C2-TM
A3. Little evidence for CRFs, instead we see examples of ‘convergent recombination’

Available to read in Virus Evolution

MRC/UVRI and LSHTM Uganda Research Unit

a.leigh-brown@ed.ac.uk @heather_eGrant

PRECISION MEDICINE DOCTORAL TRAINING PROGRAMME