Whole-genome sequencing of Kaposi sarcoma-associated herpesvirus (KSHV/HHV8) reveals evidence for two African Lineages.

568

Razia Moorad[1], Angelica Juarez[1], Justin T. Landis[1], Linda J. Pluta[1], Megan Perkins[1], Avery Cheves[1], Dirk P. Dittmer[1],
[1]Lineberger Comprehensive Cancer Center, School of Medicine, Department of Immunology and Microbiology, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA

BACKGROUND

Kaposi sarcoma (KS)-associated herpesvirus (KSHV/ HHV-8) was first sequenced from a body cavity (BC) lymphoma cell line, BC-1, isolated from an AIDS patient in 1996. Fewer than 150 complete KSHV genomes are available as compared to several 1000 for Epstein-Barr Virus. This limits our ability to develop KSHV vaccines, therapies and cures. This study reports additional recent whole KSHV genomes and demonstrates their evolutionary relationship to those collected in the 1990s.

METHODS

Patient samples were obtained from the AIDS and Cancer Specimen Repository (ACSR), other, de-identified archival samples were from Malawi and the US (collected around 2010). Targeted sequencing libraries were prepared from 100 ng total DNA using the Ion AmpliSeq Library Preparation Kit 2.0 and sequenced on the Ion Torrent S5 system. Bioinformatic analysis was performed on CLC Genomics Workbench v20.0.3. High quality, trimmed reads were mapped to the KSHV reference genome (NC_009333). Duplicate mapped reads were removed and variants were called down to 60% frequency. Consensus sequences were built, and phylogenetic trees constructed using the MAFFT alignment tool implemented in Geneious v9.1.8. The maximum likelihood trees using the General Time Reversible protein model with the gamma distribution for site variation, implemented in RAxML with 1000 bootstrap replicates, were built. In addition, 117 full KSHV genomes were downloaded from Genbank to construct a phylogenetic tree based on the long unique region (67,021 bp of continuous sequence) of KSHV.

Supported by public health service grants RO1-CA239583, U54-CA254564, UM1-CA121947 to DPD, UM1-CA181255 to the AIDS Cancer Specimen Resource (ACSR), and an AIDS and Malignancy Consortium (AMC) fellowship to RM (UM1-CA121947).

This study would not have been possible without the prior work conducted by others in the field and their willingness to publicly share genomic sequences and de-identified clinical meta data without restrictions.

16 high coverage complete KSHV genomes, isolated from HIV positive KS patients in Sub-Saharan Africa and the USA were generated. A phylogenetic analysis of KSHV evolution based on this, the largest whole genome data set to date suggest the existence of one or more “African” lineages of KSHV. These represent KSHV strains that are responsible for 90% of the HIV-KS and KSHV-associated lymphoma burden in the world.

RESULTS

- Whole genome sequencing of KSHV from KS samples resulted in an average fold coverage of between X200-600. 13 full genomes of KSHV were added to Genbank repository.
- Phylogenies were based on KI and K15 genotyping as well as the entire LUR (67,021 bp); the KSHV strains from Malawi were of the K1 A5 subtype and those isolated from the USA of C3, consistent with previous data. The K15 genes were characterised as the P subtype.
- Shared single nucleotide variants were analysed between all groups, revealing 2 high quality, non-synonymous SNVs identified in all samples located in ORF11 and ORF61 at positions 15,906 and 99622, respectively (**Figure 1**). These should likely be included into an improved, 2022, reference sequence.
- Among all SSA genomes an additional 109 SNVs were uncovered. Four SNVs were found to be unique in genomes obtained from KS biopsies.
- Phylogenetic analysis performed on the largest KSHV dataset yet to be reported, supports two distinct subtypes of the genomes isolated from from Zambia, Malawi and Uganda.
- Genomes from the USA and Europe, including the BAC constructs clustered away from the Sub-Saharan African genomes (**Figure 2**).

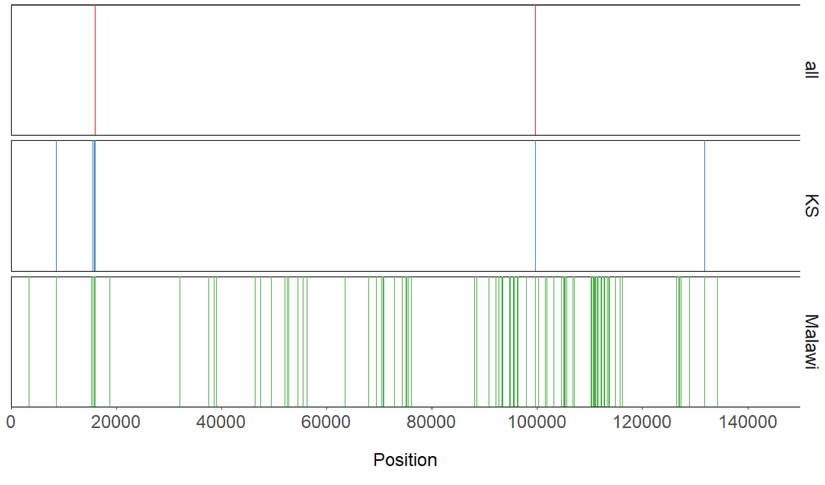

Figure 1: Distribution and Relative position of SNVs found in KSHV

CONCLUSION

This analysis indicates the existence of an ‘African’ lineage of KSHV representing viral strains that are currently circulated in HIV+ cancer patients. These largely resemble those which have been used as experimental tools for the past 20 years; however, individual viral proteins may differ significantly from the reference sequence. It is unclear, which of these novel SNVs are associated with novel phenotypes and how they may influence vaccine design.

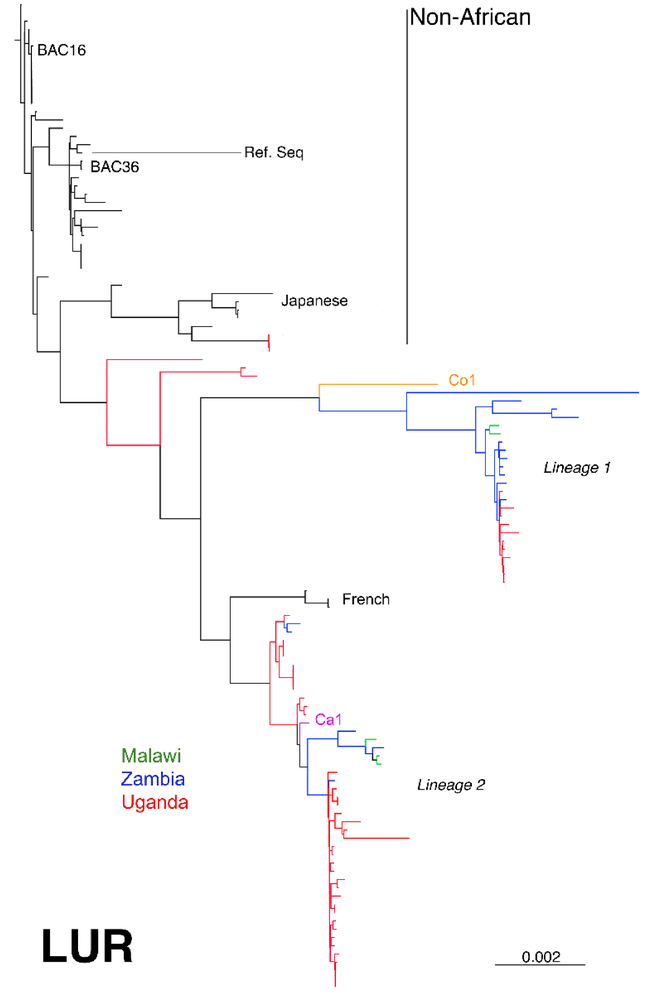

Figure 2: Maximum Likehood phylogenetic tree based on the LUR region of KSHV (from Moorad et al. *Virology*, in revision).