

Using phylogenetics and modeling to understand HIV-1 transmission dynamics in Ugandan fishing communities



Nicholas Bbosa, PhD Fellow

MRC/UVRI and LSHTM Uganda Research Unit

Makerere University, College of Health Sciences (CoH), Department of Medical Microbiology

INTRODUCTION

- In Uganda, HIV is at high prevalence (29% [1]) and incidence (6/100 PYAR [2]) in the fishing communities (FCs) of Lake Victoria relative to the general population (GP) and other high risk groups [3]
- Transmission dynamics in FCs are still not well understood.
 - Understanding the transmission dynamics is critical for effective epidemic control.
- Due to high HIV-1 incidence rates in FCs [2, 4-5] they have been thought to act as reservoirs of viral transmission [4, 6]
- But there was lack of enough evidence to prove this (source of HIV transmission is from the FCs to the GP).
- Phylogenetic techniques provide a powerful tool to examine transmission dynamics

Study 1

Qn. Are fishing communities of Lake Victoria sources of or sinks for HIV-1 transmission from the General population



HYPOTHESIS

- Fishing communities of Lake Victoria are viral reservoirs or sources of HIV-1 transmission.

OBJECTIVE

1) To identify HIV-1 transmission networks in the FCs and GP and reconstruct the temporal/geospatial viral transmission patterns.

Methodology

- A cross-sectional study of 8 FCs and 2 GP cohorts
- 606 HIV-1 partial *pol* sequences from FCs (255 sequences) and GP (351 sequences) cohorts were analyzed by phylogenetic methods
- Socio-demographic data from study participants was captured using structured questionnaires
- HIV-1 transmission networks on RAxML phylogenetic trees were identified using Cluster Picker
- Phylogeography was used to reconstruct the viral migration patterns between the FCs and the GP
- A phylodynamic analysis in BEAST was used to determine the temporal dynamics of HIV transmission in identified networks (time-resolved phylogenies)

Bayesian Evolutionary Analysis Sampling Trees



- A Bayesian Markov Chain Monte Carlo (MCMC) method was implemented in BEAST v1.8.4 for a chain length of 300 million generations sampling after every 10,000th iteration.
- An uncorrelated lognormally-distributed relaxed molecular clock with the SRD06 model of nucleotide substitution and a coalescent skygrid tree prior were used as determined by path sampling/stepping-stone method
- Phylogeographic analysis was performed using an asymmetrical discrete traits substitution model
- Phylogeography was based on the reconstruction of ancestral states (Markov model) and the count of the number of location changes that occurred in phylogenies
- A Bayesian Stochastic Search Variable selection (BSSVS) method was used to limit the number of significant rates that describe the diffusion process and generate a Bayes factor (BF) test
- The viral migration patterns were reconstructed in the SPREAD program

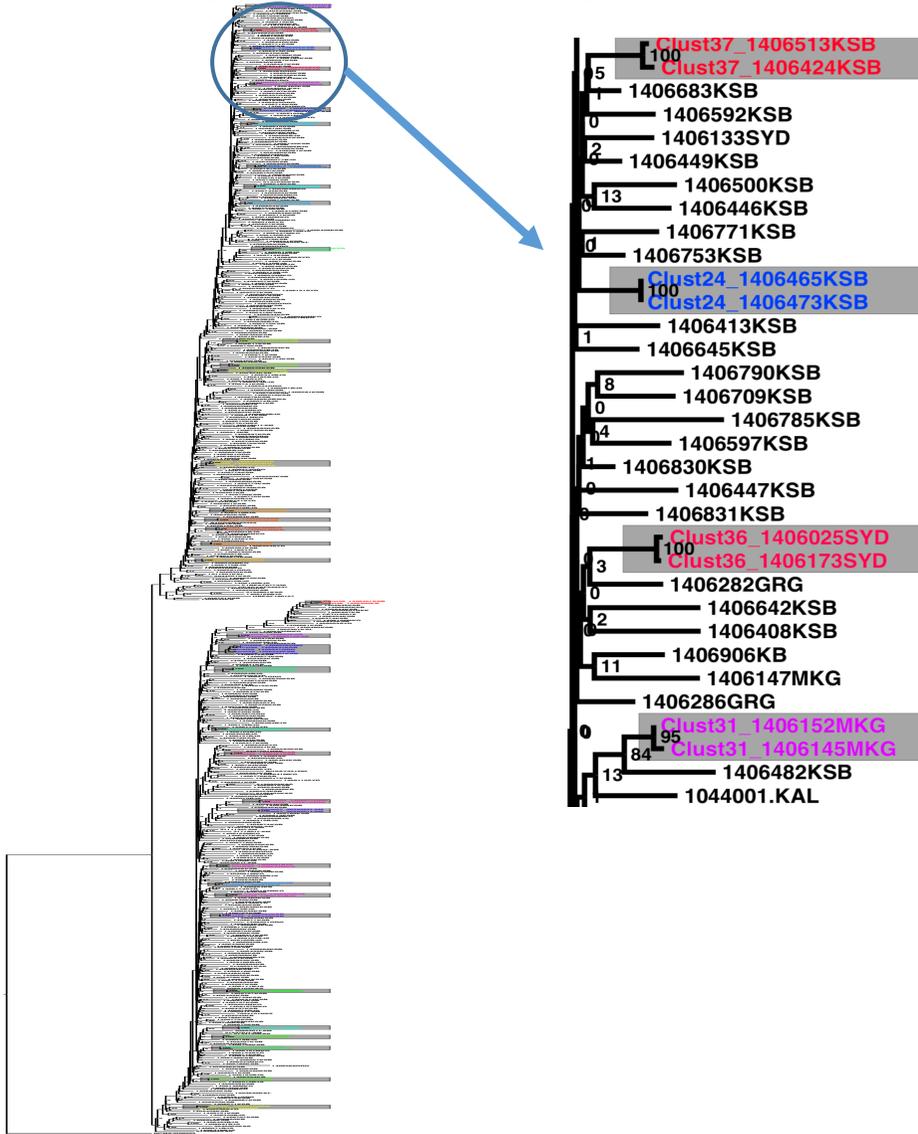
Results

- Phylogeography showed strong support (Bayes Factor >50) for viral migration from the GP to FCs

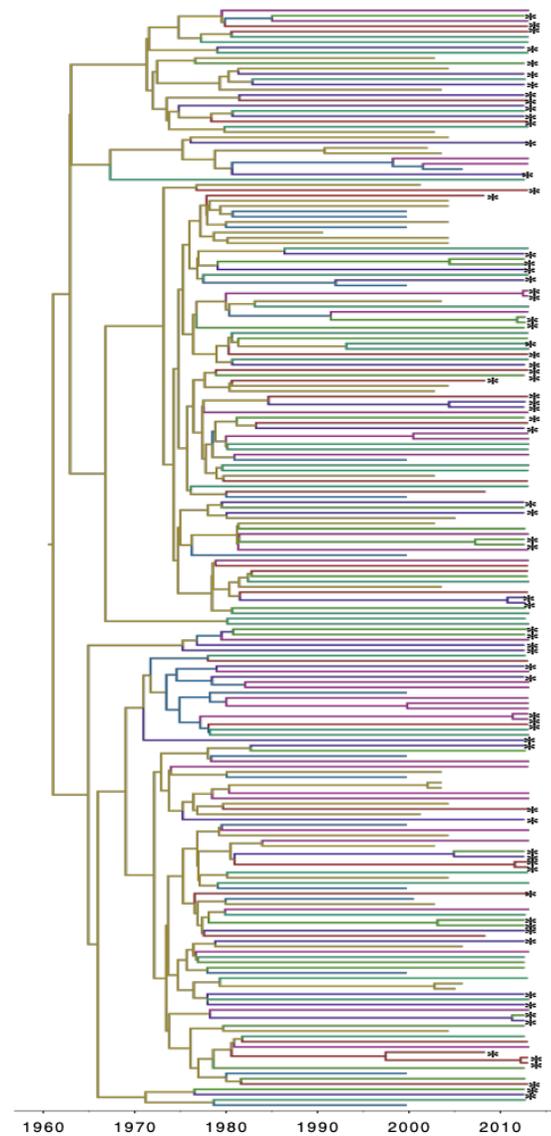
Table showing significant viral migration with BF>3

Viral migration FCs and GP	Subtype A1	Subtype D
GP-FC	178	51
GP-FC	177	246
GP-FC	110	50
GP-FC	94	98
GP-FC	52	61

RAxML Phylogenetic tree showing transmission networks



Simplified location-annotated MCC tree



Results

- More recent HIV transmission pairs from FCs were found at a genetic distance (GD) <1.5% than in the GP (Fisher's exact test, $p=0.001$)

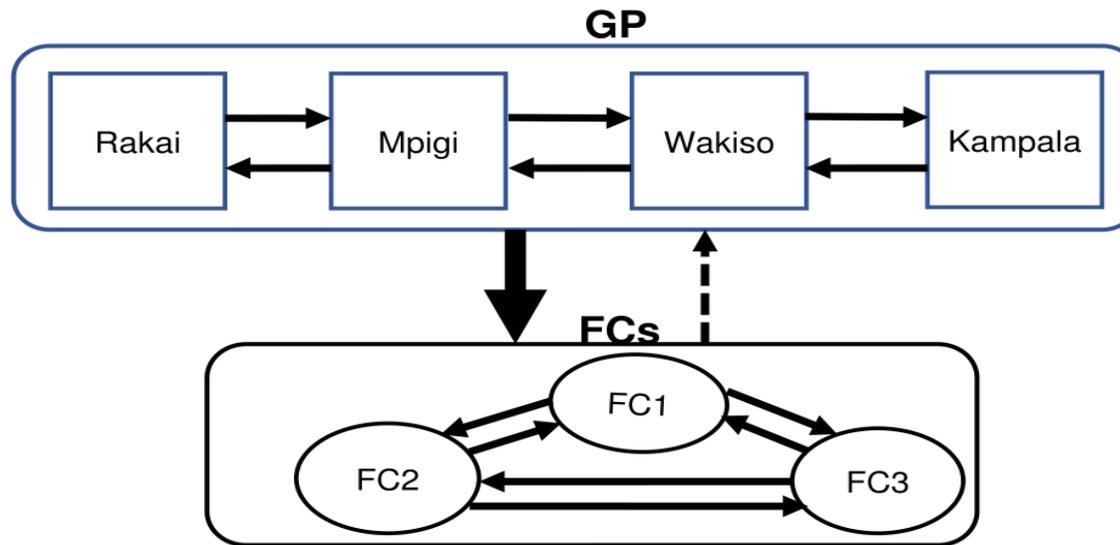
Contingency table showing pure subtype pairs identified at GD thresholds of 1.5%-4.5% and <1.5%

Group	Number of pairs 1.5-4.5%	Number of pairs <1.5%	Total
FCs	3	10	13
GP	15	3	18
Total	18	13	31

Any linkage found among individuals in the FCs is more likely to be recent (Fisher's exact test, $p=0.001$)

- A phylodynamic analysis: The mean time depth (TD) for pairs was shorter in FCs (5 months) than in the GP (4 years)

Major finding



- Phylogeographic analysis showed strong support for viral migration from the GP to FCs without evidence of substantial viral dissemination to the GP
- Targeted interventions in FCs should be extended to include the neighbouring GP for effective epidemic control

Phylogeography of HIV-1 suggests that Ugandan fishing communities are a sink for, not a source of, virus from general populations (Accepted Scientific Reports 2018)

Study 2: Combining Phylogenetics with Modeling

FCs are not sources for HIV transmission but act as transmission sinks **but what is the underlying network structure in this population**



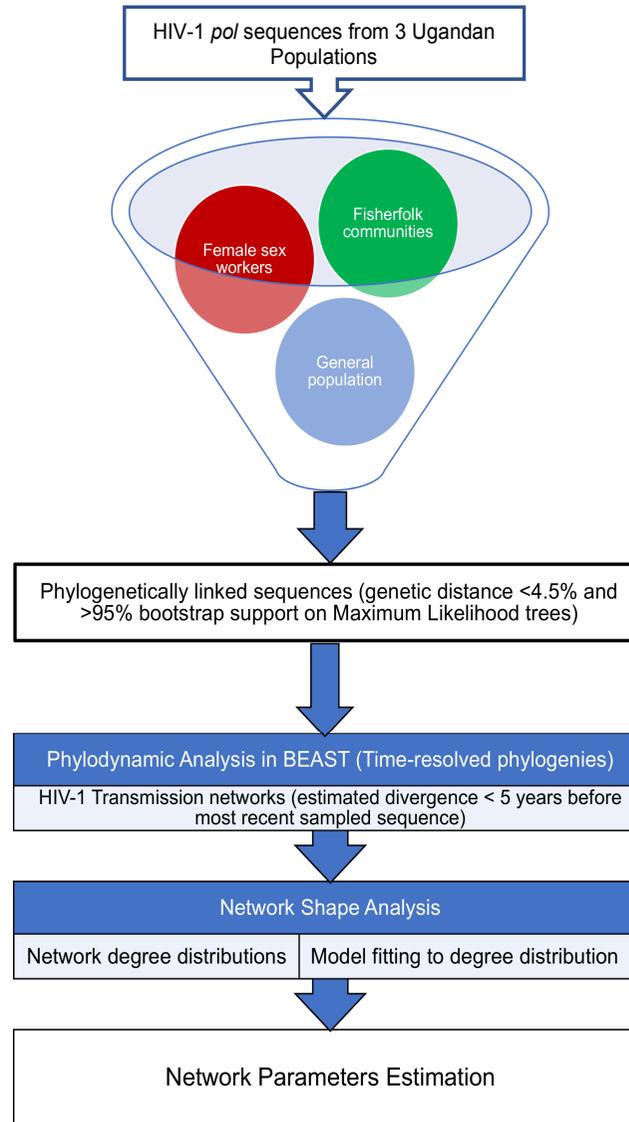
Transmission network parameters and dynamics inferred from HIV-1 sequences of fisherfolk communities, female sex workers and general populations in Uganda” submitted to PLOS ONE

Background

- Transmission network studies have provided useful insight in understanding HIV-1 transmission in high-risk and general populations in Uganda
- However, estimating parameters that influence network formation and evaluating the underlying network structure has not been done
- This is important in assessing determinants of disease spread or persistence in sexual networks that is critical in predicting and making informed decisions about effective epidemic control.
- We aimed to estimate transmission network parameters and characterize the network shape from HIV-1 nucleotide sequences (n=2,017) of Lake Victoria fisherfolk communities (FFCs, n=728), female sex workers (FSWs, n=592) and general population (GP, n=697) in Uganda for the first time in these populations.

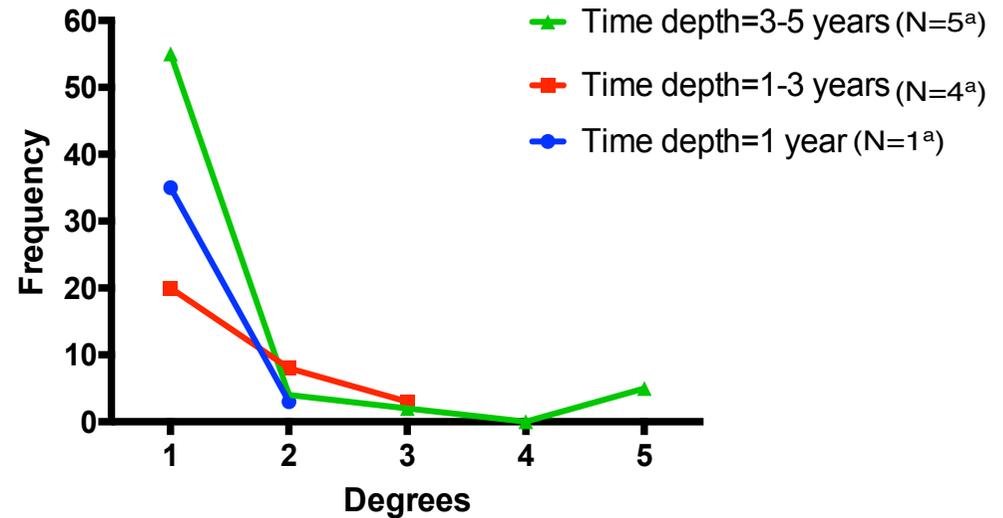
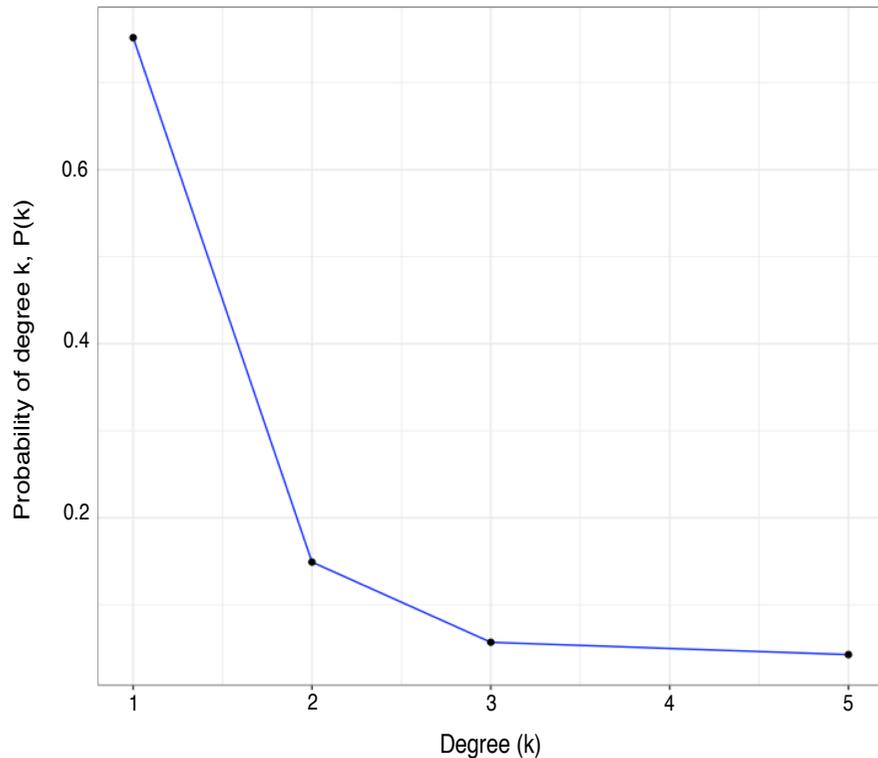
 In the GP, we focused on cohorts that were sampled more densely

Summary of Methodology



Results

- The degree distribution for all sequences was approximated by a power law with an increase of higher degree nodes at increasing time depth (from 1 to 5 years).
- An increase in cluster formation at longer time depths.



^a Includes a cluster of 3 linked sequences,

^b four clusters that include 3 clusters of 3 linked sequences and 1 cluster of 4 linked sequences,

^c five clusters that include 3 clusters of 3 linked sequences, 1 cluster of 4 linked sequences and 1 cluster of 6 linked sequences.

Results

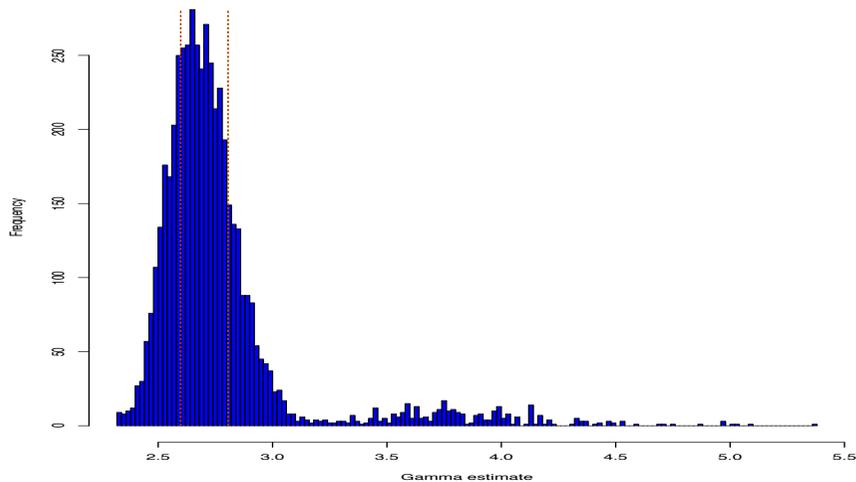
- Older nodes in the network have more connections or edges



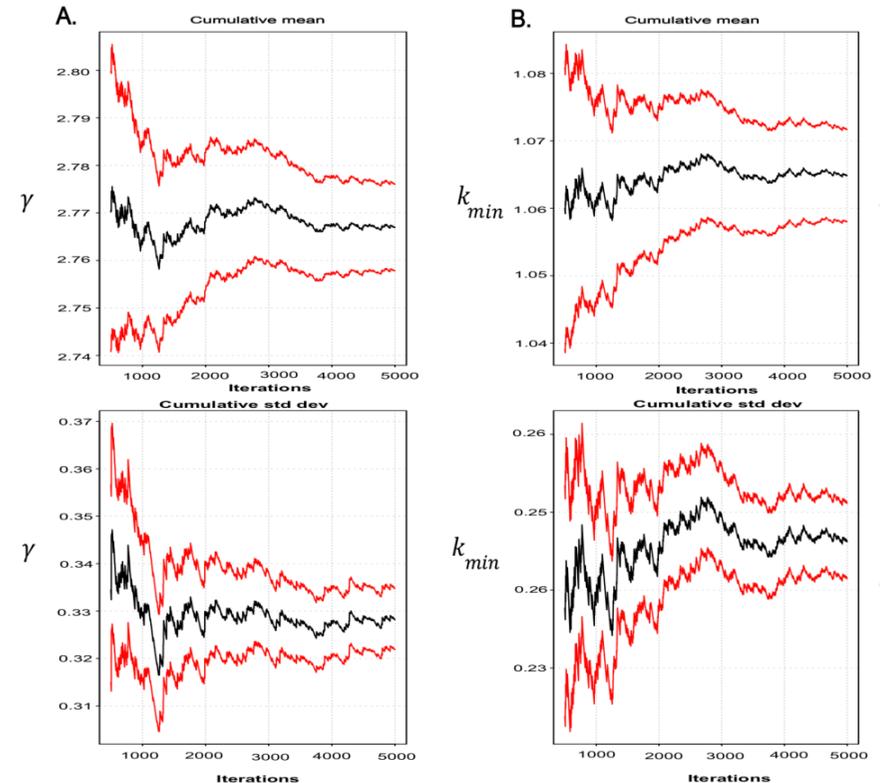
Results

- Networks of fisherfolk communities were scale-free with exponent γ between 2 and 3, suggesting an infinite variance of the degree distribution.
- This implies that HIV transmission involving a few but highly connected individuals in a network could result in the significant spread and persistence of the disease.

Graph showing the distribution of bootstrap estimates for γ with 95% confidence intervals



A graph showing the bootstrap resampling of parameter estimates



- The gamma (γ) parameter is an estimation of the exponent for the power law distribution

Results

- The Waring, Yule and negative binomial models fitted best networks of FFCs, FSWs and the GP respectively.
- The Yule and Waring models arise from a preferential attachment process.
- The negative binomial makes the assumption that individuals have a fixed rate of partner acquisition over time .

Major findings

- This study showed that the underlying network structure in the fisherfolk population was scale-free and shaped by preferential attachment in contrast to the GP.
- Preferential attachment network: Nodes prefer to attach to others with many connections



- This suggests that while generalised random interventions could be effective in preventing disease spread in the GP, epidemic control in high-risk populations would necessitate early and focussed interventions targeted at groups with high network degrees.

Acknowledgement

MRC/UVRI and LSHTM Uganda Research Unit



Uganda
Virus
Research
Institute

LONDON
SCHOOL of
HYGIENE
& TROPICAL
MEDICINE



Prof. Pontiano Kaleebu

Dr. Deogratius Ssemwanga

Dr. Rebecca Nsubuga



Prof. Andrew Leigh-Brown

Dr. Gonzalo Yebra



Dr. John Kitayibwa

Mr. Alfred Ssekagiri

THANK YOU