Detecting Changes in Incidence Using Phylogenetic Tools:
Simulation-Based Studies within the PANGEA_HIV Initiative

Emma Hodcroft[1], Oliver Ratmann[2], Anne Cori[2], Mike Pickles[2], Samantha Lycett[1], Manon Ragonnet[1], Matthew Hall[1], Andrew Leigh Brown[1], and Christophe Fraser[2] on behalf of the PANGEA_HIV Consortium

[1]University of Edinburgh, UK; [2]Imperial College, London, UK

summary

- Extensive next-generation sequencing of sub-Saharan African HIV epidemics will allow better characterization and evaluation of prevention efforts
- Simulated HIV sequences from detailed epidemiological models allow evaluation of the accuracy of phylogenetic tools at detecting changes in HIV epidemics
- Independent analysis of the simulated data shows the power of phylogenetic methods to identify changes in incidence and prevalence

contact

Emma Hodcroft
emma.hodcroft@ed.ac.uk
emmahodcroft.com

Oliver Ratmann
oliver.ratmann@imperial.ac.uk

www.pangea-hiv.org

introduction

- PANGEA_HIV (Phylogenetics and Networks for Generalized HIV Epidemics in Africa) is generating large volumes of next generation HIV sequences from epidemics in sub-Saharan Africa
- Two models have been developed to simulate HIV transmission paths and viral phylogenies
- To estimate the accuracy and reliability of phylogenetic methods to better characterize HIV epidemics and evaluate prevention methods
- Participants independently validated methods' ability to detect changing epidemic dynamics

methods

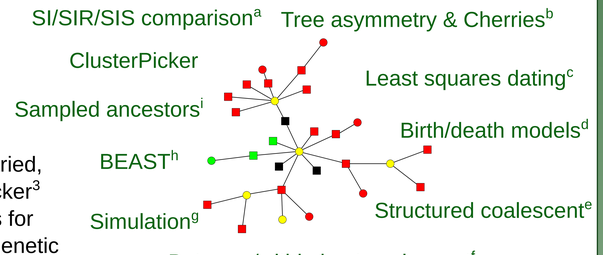
- Epidemics were simulated with treatment (see Models)
- Samples were taken at different epidemic dynamics
- Subtype C-like sequences with gag, pol, and env were generated
- 9 groups from Europe, New Zealand, and America analyzed simulated data from both models
- Results were presented at a PANGEA meeting in December 2014

analyses: first round

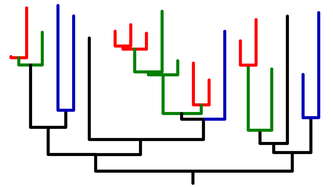
Participating groups used a variety of methods to analyze the simulated data including:

- Phylogenetic trees were generated e.g. using RAxML[2] or Bayesian methods
- Subsequent steps varied, for example ClusterPicker[3] (CP) identified clusters for each sample using a genetic distance threshold

Variety of Methods Used

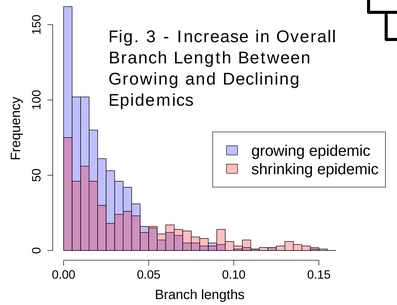


Phylogeny Component in All



- CP detected a highly significant difference between the growing and declining epidemics (Fig. 3)

Fig. 3 - Increase in Overall Branch Length Between Growing and Declining Epidemics



conclusions

- We have developed two **highly-detailed and flexible models** to simulate HIV epidemics and generate realistic sequence data
- Initial analysis using simulated data highlights the **power of phylogenetic tools to detect changes** in epidemic dynamics
- The flexibility of the models will allow important **assessment of the performance of phylogenetic tools** in a generalised African epidemic scenario
- The **latest release of simulated data** from both models is now available at: http://bit.ly/PANGEAHIVsim Results expected in May '15

models

Both simulations are detailed, stochastic, individual-based models that incorporate disease stage and HIV-dependant and -independent death.

Village

"Village" simulations
- based on ~8,000 individuals in 2-person households and sex workers (Fig. 1)
- Grow for 30-40 years, plateau, and decline after treatment is introduced

Regional

"Regional" simulations based on PopART trial[1]
- model 80,000 individuals and more detailed clinical information over 40 years (Fig. 2)
- Incorporate initial treatment under WHO guidelines with later intensification

Fig. 1 - Deme and Contact Structure

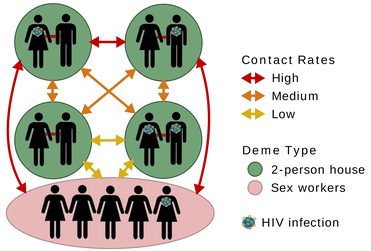
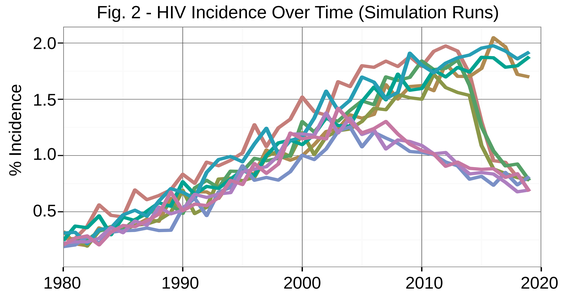


Fig. 2 - HIV Incidence Over Time (Simulation Runs)



This work was supported through the Pangea_HIV Consortium with support provided by the Bill & Melinda Gates Foundation and by NIH GM110749

Round One Participants:
Gabriel Leventhal[c,d]
Giacomo Plazotta & Caroline Colijn[b]
Daniela Bezemer[h]
Erik Volz & Simon Frost[c,e]
Michelle Kendall & Caroline Colijn[f]
David Welch, Alexei Drummond, Sasha Gavryushkina[d,i]
Marc Suchard, Michael Karchner, Guy Baele[c,e]
Bethany Dearlove[a,b,h]
Art Poon & Jeffrey Joy[g]
Tanja Stadler, David Rasmussen, Denise Kuehnert[c,d,h]
Andrew Rambaut

references

1. Cori A, et al. PLoS One. (2014) 9 (1):e84511.
2. Stamatakis A, Bioinformatics (2006) 22 (21):2688–90.
3. Ragonnet-Cronin M, et al. BMC Bioinf (2013) 14(1):317.