4.93 vs. 4.73). There was no difference in presence of transmitted drug resistance. This study confirms the current epidemiological knowledge of local spread of HIV-1 in Belgium and provides a solid base for more in-depth characterization of transmission and for future real-time follow-up of cluster dynamics.

A6 Using phylodynamic modelling to estimate the population attributable fraction of HIV spread due to key populations in Dakar, Senegal

E. Geidelberg and E. Volz

Imperial College London, London, UK

Although it has long been believed that certain key populations contribute disproportionately to HIV infection, the proportion of transmission events attributable to them is poorly understood. Most existing methods for estimating the population attributable fraction (PAF) are derived from the proportion of prevalent infections found in each group, or from static modes of transmission studies. Although these methods are useful in obtaining a cross-sectional estimate of the fraction of incident infections acquired in each group, they do not take into account the chain of transmission, and thus may underestimate the contribution of key populations. Using a transmission dynamics model, we aim to estimate the PAF of female sex workers, clients, and men who have sex with men to the HIV epidemic of Dakar, Senegal. On top of behavioural and epidemiological data, we will have access to genetic data from these key populations from an ongoing study, as well as historical samples from the Los Alamos database. As genetic diversity is shaped by epidemiological history, population genetic modelling of our sequence data can be informative about epidemic size and the migration of lineages through space and between risk groups. Our model will be first parameterised and fitted to behavioural and epidemiological data. We will then perform a phylogenetic analysis on our sequence data, using known dates of sampling and a molecular clock model of sequence evolution. Using structured coalescent models, we can look at the balance of phylogenies and infer patterns of transmission (although we will not have a large enough sample to determine clusters). We can then refit the transmission model to the sequence data as well, and provide new estimates of the PAFs. The comparison of PAFs estimated with or without using sequence data will provide an insight into the added value of phylodynamic modelling, and may help reassess the role of key populations in this setting.

A7 The effect of the mechanism and amount of missingness on phylodynamic inference of heterosexual HIV transmission networks

David Niyukuri,[1] and Wim Delva[1,2,3,4]

[1]South Africa Centre of Epidemiological Modelling and Analysis, Stellenbosch University, South Africa, [2]Department of Microbiology and Immunology, Rega Institute for Medical Research, Clinical and Epidemiological Virology, KU Leuven, Belgium, [3]Center for Statistics, I-BioStat, Hasselt University, Belgium and [4]International Centre for Reproductive Health, Ghent University, Belgium

Successful HIV prevention requires a better understanding of the structure and the dynamics of the sexual network. The latter evolves over time and the transmission network emerges from that dynamic sexual network. At the same time, the genetic material of the virus evolves across the transmission network. Due to the fact that population-level HIV transmission dynamics and HIV evolutionary dynamics are on the same time scale, the molecular evolution of the virus becomes a footprint of the transmission network. Therefore, the analysis of HIV phylogenetic trees holds the promise to provide more objective methods to estimate transmission network characteristics. We aimed to investigate the effect of sub-optimal sequence coverage on the characteristics of the transmission network (e.g., degree distribution and link density) inferred from the phylogenetic tree. We simulated a small epidemic (seventy-two transmission events) using agent-based models. Across the transmission chain, the molecular evolution of the virus was simulated using appropriate substitution models and evolutionary rates for HIV-1. We considered one consensus sequence per individual, and we simulated five levels of sequence coverage. In addition, we simulated two sampling strategies: a cross-sectional sampling design at the end of the simulation time window, and longitudinal sampling design. For each level, we constructed a phylogenetic tree and the subsequent transmission network, which were compared to the true transmission network. We found that a reconstructed transmission network from a phylogenetic tree has characteristics close to those of the true transmission network when sequence coverage was at least 60 per cent of the infected individuals (forty-five sequences). The increase of taxa (sequence coverage) improves the inferred transmission network characteristics. In addition, transmission networks reconstructed with the cross-sectional sampling design had less overestimation of links (which are seen as potential transmission events).

A8 Improving the accuracy and precision of estimated temporal trends in HIV incidence among MSM populations by calibrating agent-based simulation models to phylogenetic tree data

W. Delva,[1,2,3,4] N. Hens,[3,5,6] and A.-M. Vandamme[2,7]

[1]South Africa Centre of Epidemiological Modelling and Analysis, Stellenbosch University, South Africa, [2]Department of Microbiology and Immunology, Rega Institute for Medical Research, Clinical and Epidemiological Virology, KU Leuven—University of Leuven, Belgium, [3]Center for Statistics, I-BioStat, Hasselt University, Belgium, [4]International Centre for Reproductive Health, Ghent University, Belgium, [5]Centre for Health Economic Research and Modelling Infectious Diseases, Vaxinfectio, University of Antwerp, Belgium, [6]Epidemiology and Social Medicine, University of Antwerp, Belgium and [7]Center for Global Health and Tropical Medicine, Unidade de Microbiologia, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisboa, Portugal

Effective prevention of sexually transmitted HIV infections requires knowledge of the sexual networks across which these infections are transmitted. Sexual behaviour surveys are routinely being conducted to ask people about their recent sexual partners, leading to emergence of star-shaped egocentric network data. At the same time, phylogenetic tree analysis has long been conducted to analyse HIV transmission clusters. In an ongoing project, we aim to show that (1) agent-based simulation models, embedded in a Bayesian framework, can provide a platform for combining these complementary data sources, and that (2) such an integrated approach can lead to more accurate and robust inferences in HIV epidemiology. Within this project, a first case study aims to provide evidence that phylogenetic tree data can increase the validity of agent-based model projections. Specifically, in a small-scale proof-of-concept study, we use synthetic data from a 'master model' to demonstrate that the accuracy and precision of estimated temporal trends in HIV incidence improves when agent-based simulation models of HIV transmission in MSM populations in Western Europe are not only calibrated to reported behavioural and epidemiological data, but also to phylogenetic tree data. Our model calibration

method employs a combination of Principal Component Analysis (to reduce the dimensionality of the space of output statistics), Latin Hypercube Sampling and sequential Approximate Bayesian Computation (to explore the parameter space in a time-efficient manner) and multivariate emulation (to interpolate output statistics for regions of the parameter space that have not been sampled explicitly).

A9 HIV subtype diversity across geography and transmission groups in Maryland

D. German,[1] S. Lamers,[2] R. Rose,[2] H. Boykin,[3] O. Laeyendecker,[4,5,6] and C. Flynn[3]

[1]Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA, [2]BioInfoExperts LLC, Thibodaux, LA, USA, [3]Maryland Department of Health and Mental Hygiene, Baltimore, MD, USA, [4]Department of Epidemiology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, USA, [5]Laboratory of Immunoregulation, Division of Intramural Research, National Institute of Allergy and Infectious Diseases, NIH, Baltimore, MD, USA and [6]Johns Hopkins University School of Medicine, Baltimore, MD, USA

The distribution of HIV subtypes within a local epidemic can provide insight into social dynamics underlying transmission patterns and inform treatment and prevention strategies. Subtype B predominates in most US regions, but non-B infections are increasingly evident in Maryland and large-scale examinations of subtype diversity are lacking. In 2014, Maryland HIV case surveillance began systematically collecting HIV genotypic sequence data to identify outbreaks and transmission patterns. We used sequences collected through 2015 to examine geographic and transmission group distribution associated with different HIV subtypes. HIV-1 pol sequences were aligned with Los Alamos HIV Database subtype reference sequences and subtyped using HIV-TRACE at a distance cutoff optimized for accurately identifying subtype B data (5 per cent). Non-B sequences were subsequently classified using the REGA HIV-1 Subtype Tool. All sequence subtypes were matched to data within the HIV case surveillance database for descriptive analysis. HIV sequences were available from 7,045 individuals between 2004 and 2015, ~20 per cent of all persons living with HIV in 2015. Subtype B accounted for 95.7 per cent of all sequences overall and 90.2 per cent of those from persons diagnosed in 2014. The most common non-B subtypes were subtype G, C, A, and D, and a variety of circulating recombinant forms. The proportion of non-B sub-types increased from 0 to 2 per cent of new annual diagnoses prior to 2000 to 7 to 9 per cent after 2010, particularly in the Central and Suburban regions. The highest frequency of non-B (4.3–25.5 per cent) was seen in five counties across the state. In contrast, <1 per cent of sequences from Baltimore City were non-B. Those with subtype B were 27.6 per cent heterosexual, 27.6 per cent people who inject drugs, and 25.2 per cent men who have sex with men (MSM). Non-B subtypes were primarily heterosexual (60.6 per cent), with increasing proportion among MSM after 2006. The large number of non-B subtypes likely reflects immigration patterns to the DC area. Non-B HIV is increasingly moving outside of DC and has been identified in some MSM and IDU cases. Increasing HIV viral diversity, including the potential for subtype recombination, across the state of Maryland poses challenges for clinical and preventive response and suggests geographic and social dimensions of HIV transmission patterns that require further examination.

A10 Using the molecular epidemiology of HIV transmission in New South Wales to inform public health response: Assessing the representativeness of linked phylogenetic data

Angie Pinto,[1] Francesca Di Giallonardo,[2] Phillip Keen,[1] Craig Cooper,[3] Barbara Telfer,[4] David Cooper,[1] Andrew Grulich,[1] and Anthony Kelleher[1]

[1]The Kirby Institute, UNSW, Kensington, Australia, [2]Sydney University, Sydney, Australia, [3]Positive Life NSW, Darlinghurst, Australia and [4]New South Wales Ministry of Health, Sydney, Australia

New South Wales has the largest number of people living with HIV in Australia, estimated at 11,500 and is undergoing a revolution in HIV prevention with the rollout of pre-exposure prophylaxis to 4,000 high-risk individuals. The recently established statewide HIV drug resistance database contains 9,982 sequences from 2004 to 2015 and has coverage of over 80 per cent of the HIV infected population. To allow enhanced analysis of these sequences, data linkage was performed on the HIV notifications database, which contains demographic data of all newly notified infections dating back to 1981. The aim of this study was to determine the representativeness of linked sequences compared to those that were unlinked using epidemiological and phylogenetic methods. Duplicate sequences per year were excluded. Deterministic linkage was performed using two by two namecode, date of birth, sex, and postcode. Differences in demographics were assessed between linked sequences and public health data for the state, using a chi-squared test for comparison of proportions. Sequences were aligned using MUSCLE, and phylogenetic inference using RaXML. Comparison of substitution rate, diversity, skyline plots, and cluster characteristics was made between linked and unlinked sequences. Only 2,843 (28 per cent) sequences were linked to HIV notification data. With the exception of 2011, the proportion of early infections in linked data was not significantly different to unlinked data from 2010, 2012 to 2015. Other demographics data were comparable. Between 2004 and 2007 there was a difference in the proportion of identical sites (23.5 vs. 8 per cent $P < 0.02$), while pairwise identity was not significantly different. For sequence data from 2008 to 2015, there was no significant difference in the proportion of identical sites, pairwise identity, or GC content. Despite the low proportion of sequences linked to notifications data, the epidemiological characteristics were representative of the HIV infected population in New South Wales. Phylogenetic analysis showed that the linked data were sufficiently representative to proceed with further cluster analysis to investigate the impact of public health interventions in this state.

A11 The 1988 Elista outbreak and the epidemiology of HIV subtype G in Russia

Ekaterina Serova,[1] Anastasia Murzakova,[2] Marina Nossik,[3] Andrey Shemshura,[4] Georgii Bazykin,[1] and Dmitry Kireev[2]

[1]Skolkovo Institute of Science and Technology, Moscow, Russia, [2]Central Research Institute of Epidemiology, Moscow, Russia, [3]I.I. Mechnikov Research Institute for Vaccines and Sera, Moscow, Russia and [4]'Clinical Center of HIV/AIDS' of the Ministry of Health of Krasnodar Region, Krasnodar, Russia

One of the first major outbreaks of HIV-1 in Russia happened in Elista in 1988. About 270 patients, mostly children, were nosocomially infected with HIV subtype G in seventeen hospitals in 1988–1989. The infection was mainly transmitted between