Bayesian phylogeography of HIV among MSM in the UK indicates a few source areas of widespread infection

Lucy Weinert*[1], Gareth Hughes[2], Esther Fearnhill[3], David Dunn[3], Andrew Rambaut[1], Andrew Leigh-Brown[1] on behalf of the UK HIV Drug Resistance Collaboration

*lucy.weinert@ed.ac.uk; [1]University of Edinburgh, Institute of Evolutionary Biology, King's Buildings, Edinburgh EH9 3JT UK; [2]Health Protection Agency, Huntingdon, Cambridgeshire, PE29 6FH; [3]Medical Research Council Clinical Trials Unit, 222 Euston Road, London, NW1 2DA UK

Summary

32,000 MSM in the UK are infected with HIV[1]. A rise in infection prevalence indicates a need for a better understanding of transmission dynamics.

Large epidemiologically-linked clusters were analysed using time-scaled phylogenies with geographic reconstruction.

There is a high rate of migration from London and a lower proportion of transmission than incidence. This may explain the decrease in infection over time.

In contrast, the west of England has a higher proportion of transmission indicating that infection in this area may increase in the future.

Introduction

Figure 1 show most people with HIV reside in London although infection is widespread.

Surveillance data helps to target HIV prevention strategies. However, it cannot inform us how HIV spreads geographically.

Understanding dynamics of spread needs a phylogenetic approach. Figure 2 shows how location state mapped on to topologies can inform geographic spread.

We used geographical and time-resolved phylogenies to investigate which areas have ongoing transmission, how transmission compares to eventual infection, and which areas are sources of infection.

Figure 1. Number of HIV diagnoses per area of the UK

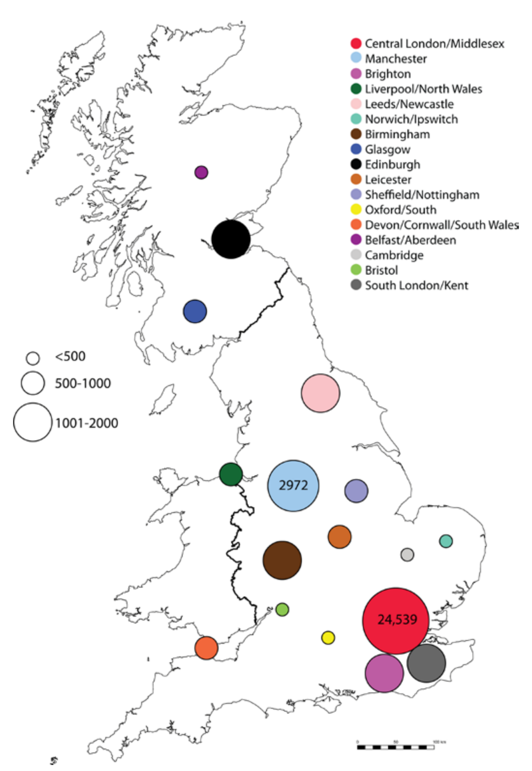
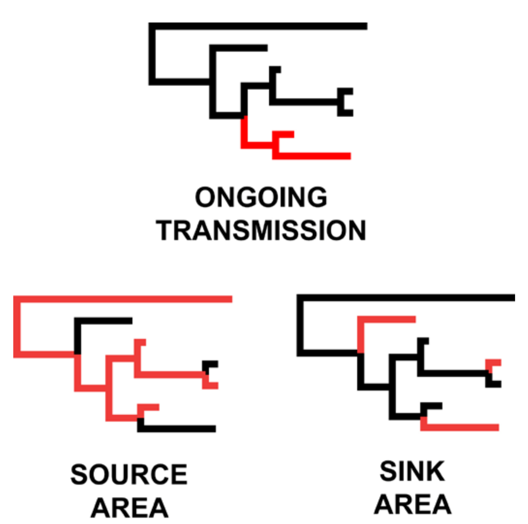


Figure 2. Signatures in phylogenetic trees and what they reveal about geographical spread (black and red represent two locations with the key in reference to the red location)



Methods

Epidemiologically-linked clusters were detected by pairwise comparisons between 14,560 partial *pol* subtype B sequences. A NJ tree was constructed with sequences of > 95.5% identity; the most basal node with ≥ 95% support represents one cluster[2].

Sequences associated with clinics were aggregated to 17 locations on the basis of proximity.

Table 1. Example of the model of evolution for geography that adopts three geographic areas

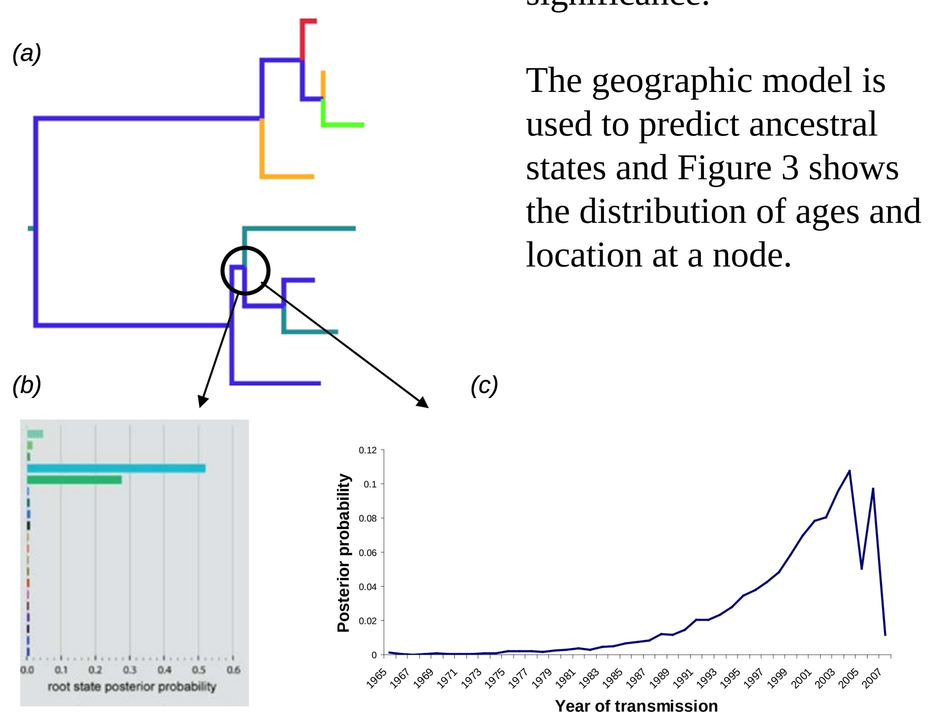
Geographic state	A	B	C
A	--	q_{AB}	q_{AC}
B	q_{BA}	--	q_{BC}
C	q_{CA}	q_{CB}	--

We used a Bayesian phylodynamic method that employs a geographical model of evolution[3] (Table 1) to analyse clusters with ≥ 10 individuals (*no. individuals* =1673).

Rate parameters shown in Table 1 were estimated between each area. Indicator variables which switch rates on or off were coupled to each rate parameter to test rate significance.

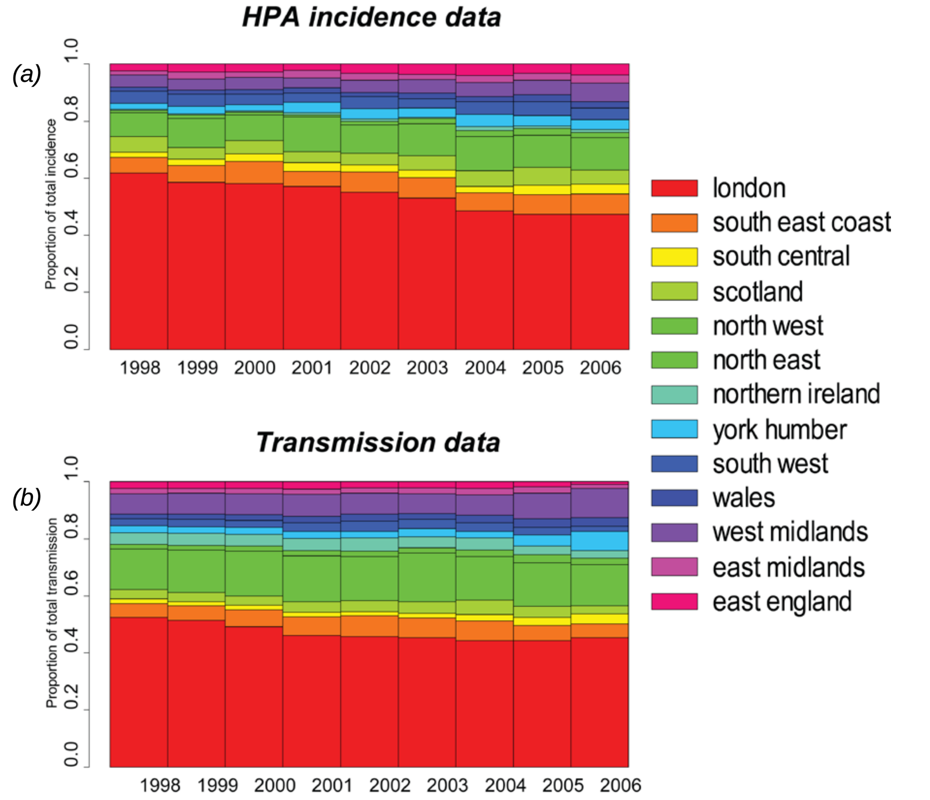
The geographic model is used to predict ancestral states and Figure 3 shows the distribution of ages and location at a node.

Figure 3. *(a)* A phylogenetic tree coloured by geographic location. Each node in the tree (where two branches collapse together) represents a transmission event and has *(b)* a posterior probability distribution of geographic state and *(c)* a posterior probability distribution of time when the transmission happened.



Results

Figure 4. Stack plots showing *(a)* the proportion of HIV in each area over time based on nation-wide incidence data from the HPA[1] and *(b)* the proportion of HIV transmission in each area over time based on posterior distributions of 1673 individuals.



>70% clusters had >50% sequences from one location signifying our method accurately detects epidemiologically-linked individuals.

Figure 4b shows that there is significant ongoing transmission in all areas of the UK.

Comparisons of incidence data with transmission data over time in Figure 4 shows that there is a lower prop^{n} of transmission in London than incidence. In contrast, the north west and the west midlands have a higher prop^{n} of transmission then incidence.

Figure 5 shows that London, Brighton, Manchester and Birmingham have the most outward spread.

Clusters are unique evolutionary events but may share geographical trends. We analysed clusters using one geographic model of evolution. Figure 6 demonstrate that clusters tend to always exhibit local spread.

Figure 5. Heat plot showing rates of HIV migration out of each particular area (source populations) for the largest 97 clusters. White blocks represent locations that were not represented in clusters. Blue blocks represent zero rates of migration. Blocks with any red tinge represent locations with significantly non-zero rates (using a a Bayes factor of 10). The legend represents the posterior probability of how significant the rate is, judged by indicator variables (see method).

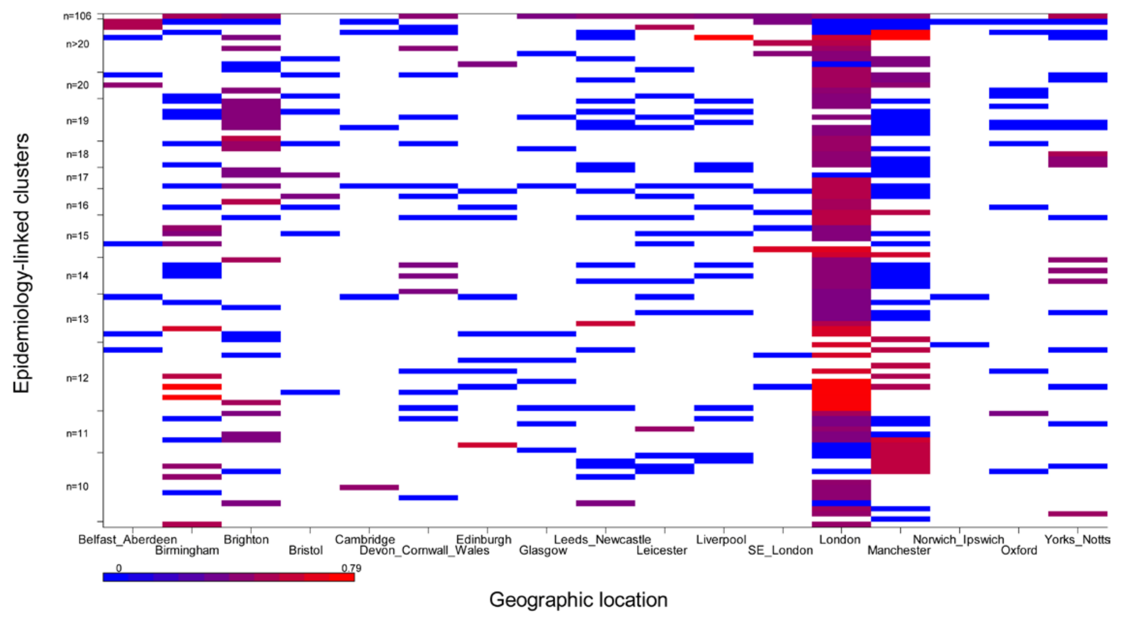
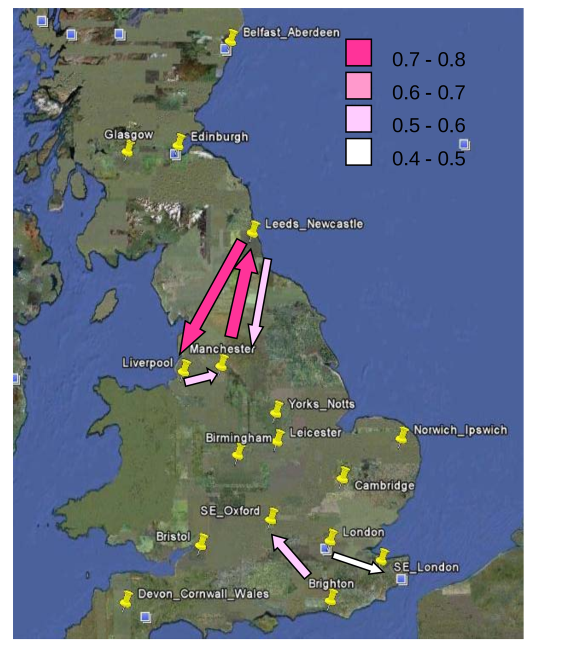


Figure 6. Significantly non-zero rates of HIV migration (using a Bayes factor of 10) using just one model of evolution to explain geographic spread in the 17 largest clusters



Conclusions

• Infection in the early UK epidemic was heavily London-based; our results suggest there is now ongoing infection in all other areas of the country

• Less transmission then incidence in London may contribute to the decreasing trend in incidence over time (Figure 4)

• The west of England shows the opposite trend and may be problem areas of infection in the future

• Areas with the most outward spread have the highest prevalence of HIV infection (Figure 5)

• Shared geographical trends between epidemic clusters indicate that transmission is primarily local (Figure 6)

• These results demonstrate the power of using Bayesian phylogenetics to inform transmission dynamics and have implications in more effective targeting of prevention strategies

References

[1]Heath Protection Agency data (www.hpa.org.uk)
[1]Lewis F, Hughes GJ, Rambaut A, Pozniak A, Leigh Brown AJ (2008) Episodic sexual transmission of HIV revealed by molecular phylodynamics. *PLoS Medicine* 5(3): e50
[2]Lemey P, Rambaut A, Drummond AJ, Suchard MA (2009) Bayesian Phylogeography Finds Its Roots. *PLoS Comput Biol* 5(9): e1000520

Acknowledgements

We would like to acknowledge the Medical Research Council and the Department of Health for funding and all other members of the UK HIV resistance collaboration, Philippe Lemey, Sam Lycett, Marc Suchard and John Welch for help.