Response to David Rasmussen

Given that the manuscript has already gone through two rounds of review and that the major concerns of the reviewers' from the second round have largely been addressed, I have decided not to send the manuscript out for another round of review. However, there are still a few issues I hope the authors can quickly address before I write my recommendation.

It's stated that the origin of the epidemic in classic hosts is estimated to be in 1957 but it was not possible to estimate when the epidemic in 'new' host started. But doesn't the 1957 estimate reflect the MRCA of all samples, regardless of whether they are "classic" or "new"?

This is a good point. We now also infer a phylogeny using only ‘classical’ host sequences. We find that both the estimate of the date of the last common ancestor (1960) and the 95% HPD are identical to the target phylogeny. This is now specified in the Methods.

Line 119 goodness-of-fit [test]?

Yes thank you. We made the change in the manuscript.

Fig 3. do the summary statistics from the true phylogenies fall within the regions predicted by posterior simulations for individual statistics or just for the PCA on the summary statistics? Would it be more convincing to show the original summary statistics?

For 77 out of 101 summary statistics, the target value is in the 95% HPD of the summary statistics computed from the 10,000 simulated phylogenies used for the goodness-of-fit test. This is now shown in two supplementary figures and mentioned in the Results.

Lines 192-194: This is hard to understand, why would adding stages of infection make it "almost impossible to simulate phylogenies"?

When simulating phylogenies forward in time, we only have constraints on the tips (but not on their labels). However, when performing the coalescent (backward-in-time) phylogeny inference, we impose additional constraints about host classes (or stages). As a consequence, sometimes the coalescent process cannot be simulated because there are not enough hosts of a given stage of infection remaining to move to the next step.

Lines 201-203: "Although the multi-type birth-death model is unlikely to be directly applicable... because it links the two epidemics via mutation... whereas in our case the linking here the links is done via transmission events". This is not true. The multi-type birth-death model can handle type changes due to transmission or mutation/migration. Please remove!!

Yes we apologize for that. We made the change in the manuscript.

Lines 204-205: "We were unable to conclude anything from this analysis which rises the limitation of the likelihood-based approach for this dataset". In fairness to likelihood-based approaches, it is probably worth noting why the MTBD models implemented in BEAST did not work on this data set. In the response letter, the authors say that this is due to poor mixing. But is this due to difficulties in jointly estimating the phylogeny and evolutionary

parameters along with the epidemiological parameters? Does the MCMC converge if the phylogeny is fixed (as was done for ABC)?

Yes, you raise an important (and fair) point! We added this in the manuscript. We also tried to perform the MTBD estimation in Beast while fixing the phylogeny but we cannot get it to run... This may appear as strange because the same priors run when we try to infer the phylogeny but it is one of the difficulties associated with the Beast tool.