Editorial correspondence about the preprint:

Simon Dellicour, Guy Baele, Gytis Dudas, Nuno R. Faria, Oliver G. Pybus, Marc A. Suchard, Andrew Rambaut, Philippe Lemey. **Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak.** *bioRxiv, 163691. https://doi.org/10.1101/163691*

Revision round #2

2018-02-08

The reviewers and I are very happy with this revised version.

I will be happy to write a recommendation.

In the meantime, one of the reviewers suggested some minor edits that could be worth adding.

Preprint DOI: https://doi.org/10.1101/163691

Reviewed by Chris Wymant, 2018-01-18 14:58

See next page

The authors addressed all of my previously concerns well. Further minor issues are detailed below; these should be easy to address.

In the authors' response letter:

>> Page 6: the authors provide no information about how the wavefront
>> was calculated.
> Answer: We have now added a sentence to describe this as well as a
> reference to the Methods section.
I cannot find what the authors are referring to. My best guess is following:
"the evolution of... the maximal wavefront distance through time. For the latter, we plot the maximal distance of the epidemic wavefront from the spatial origin as a function of time."
The second, new sentence doesn't really add anything. What defines the wavefront? Is it simply the single lineage that at a given moment in time is the furthest from the location of the ancestral lineage? (Which would mean that the use of 'maximal' in describing the wavefront distance is redundant?) The term 'wavefront' suggests to me an expanding wave-like model, fit to the locations of the most distant handful of lineages at each given time, but unless I've misunderstood I don't think that's the case here.

In the revised manuscript:

In the abstract, in the space of six sentences the authors describe three different things as 'critical'. The repetition aside, I think the hype doesn't help: most readers will appreciate the importance of these issues.

Much of what the authors say about the use of viral genomes in understanding virus spread is also applicable to bacteria; they might consider broadening the terminology from viruses to pathogens.

Page 1: *"metapopulation dynamics were critical for connecting rural and urban areas"*
I find this statement odd. I would not say that the metapopulation dynamics caused this connectedness; they arose because of the (time- and space-varying) connectedness of rural and urban areas.

page 1: *"interventions within the region such as border closures, lockdowns, and restricting travel may more challenging to investigate"*.
Missing "be" after "may".

Page 1: *"...this outbreak occurred in a highly connected region of Africa with large population centres. This connectivity is also important to consider in local management strategies because interacting populations do not necessarily implement policies that are coordinated, as was the case for EBOV in Guinea, Sierra Leone and Liberia."*
Personally I would find this clearer if the second sentence were fused with the first, without expanding on what "this outbreak" refers to (the reader knows). e.g.

"...this outbreak occurred in a highly connected region of Africa with large population centres, spread over multiple countries, without fully coordinated intervention policies."

Page 1: *"The increasing individual-level spatio-temporal disease data... may offer"*
Consider removing the opening 'The'.

Page 2: *"In that case, percentages of reduction in epidemic size and duration are estimated for the period of time during which the intervention strategy is effective"*
I suggest adding "(i.e. considering only that part of the tree dated after June 2014)". It is important to make this point as clear as possible (perhaps also repeating it in the part of the caption referring to the brown histograms of Fig 1C). I misunderstood this in the original version; anyone who still fails to understand it is likely to conclude that certain interventions can be applied quite late in the epidemic with the same final impact as if they were applied earlier. If such a person decides policy, this would be a fatal misunderstanding.

Page 2: *"Figure S2 provides summarises"*
Correction needed.

Page 3: *"Further, the epidemic was generally less likely to spread across international borders, but did so specifically both early and late in the epidemic, between locations that share such an international border."*
Suggest replacing 'locations' by a more specific term such as administrative district, otherwise this reads as a tautology. (Continuous motion over an international border must proceed via a location just before the border and one just after; I think that what the authors are trying to say is that a single lineage has been inferred to start in a district touching one side of the border and end in a region touching the other side of the border, rather than leap-frogging a border district.)

Page 3: *"we assess how critical such long-distance events were"*
"How critical" implies that the events were definitely important, and you assess whether they're merely quite important, or very important. "How important" is the more appropriate neutral phrasing that allows for the possibility that the events were not important.

Page 4: *"Weighted dispersal velocities are both reported in"*
-> are reported in both

Page 4: *"events recorded in the MCC (maximum clade credibility),"*
"tree" is missing between the closing bracket and the comma

Page 5: *"These estimates provide strong evidence for a significantly reduced frequency in international border crossing from September 2014 (p-value <0.05)"*
Apologies for not noticing this in the original draft. p-values should be stated, not provided as inequalities: we are interested in the level of evidence, not the false dichotomy of 'statistically significant' or not.

Page 5: *"with more cases and fatalities than all previous outbreaks combined"*

Would "reported outbreaks" be more accurate?

Page 5: *"the region has been declared Ebola free"*
Ebola-free

Page 5: *"we investigate the two key elements of viral spread in a gravity model"*
I think the authors do themselves a disservice here: they have investigated viral spread tout cours, not within the framework of (and subject to the applicability of) a gravity model.

Page 5: *"short-distance dispersal realised may be more important"*
'realised' should be removed (left over from 'by human mobility').

Page 5: *"This reflects the important role of highly populated locations in maintaining gravity-model transmission, as previously identified"*
See my previous comment on 'identification' of a gravity model. Also, a gravity model is a description of the situation in which highly populated locations are important for transmission. Saying that highly populated locations are important to maintain the model seems either false or tautological, I can't decide which. I suggest simplifying the statement.

Page 5: *"If viral lineage movement to a single capital could have been prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. In contrast, preventing lineage movement to all the capitals reduced epidemic size to about one-third, while their sample size percentage and case count percentage are 28% and 39%, respectively"*
In this passage, the three fractional reductions are presented as i) by 15-37%, ii) to one-third, iii) by 28%. This asymmetry initially confused me - I slipped into reading the second "to" as "by", and lost the point being made. It would be easier to read if the second reduction was described as by ~67%.

Page 6: *"may be a reasonable approximated by"*
Correction needed.

Page 6: *"In our study we use posterior predictive simulation based on continuous phylogeographic reconstruction to identify a significant decline in international border crossing. Our findings confirm a significant decline following the announcements of border closures between Guinea, Sierra Leone and Liberia, which was also observed using a discrete phylogeographic approach"*
The second sentence more or less just repeats the first – consider merging and simplifying.

Page 7: *"allows calculating"*
-> "allows the calculation of" (this was correct first time round - 'allow' is transitive)

Supplementary Figure S3: I suggest removing the legend ("No intervention before July 2014"), as the fact that it describes only one of the two things plotted is confusing. The two things can be described in the caption as for figures 1C and 1D.

Appendix S1: *"or dispersal velocity estimates"*

-> our

Appendix S1: *"each sequence were"*
-> was

Appendix S1: *"the first 5% samples"*
-> of samples

Appendix S1: I suggest quoting the velocities here in km/day, or the velocities in the main text in km/year, for easier comparison: the main result of this appendix is that these velocities are similar.

Best,
Chris Wymant

Reviewed by Christian Althaus, 2018-01-18 15:01

I would like to thank the authors for revising their manuscript and adequately addressing my previous comments.

Author's reply:

See next page

Dr Simon Dellicour
Rega Institute, KULeuven
Minderbroedersstaat 10
3000 Leuven, Belgium

02.02.2018

Dear Managing Board of ***PCI Evol Biol***,

We want to thank again the Recommender and the two Referees for all their constructive comments on our preprint manuscript entitled "Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak".

We addressed the new comments of the first Reviewer in a second revised version of our preprint manuscript (see below for a point-by-point reply). We are very grateful to the Referees and the Recommender for the time and effort they have put into this.

Best regards,

Simon Dellicour

Reviewer 1:

The authors addressed all of my previously concerns well. Further minor issues are detailed below; these should be easy to address.

In the authors' response letter:
>> Page 6: the authors provide no information about how the wavefront was calculated.
> Answer: We have now added a sentence to describe this as well as a reference to the Methods section.
I cannot find what the authors are referring to. My best guess is following: "the evolution of... the maximal wavefront distance through time. For the latter, we plot the maximal distance of the epidemic wavefront from the spatial origin as a function of time." The second, new sentence doesn't really add anything. What defines the wavefront? Is it simply the single lineage that at a given moment in time is the furthest from the location of the ancestral lineage? (Which would mean that the use of 'maximal' in describing the wavefront distance is redundant?) The term 'wavefront' suggests to me an expanding wavelike model, fit to the locations of the most distant handful of lineages at each given time, but unless I've misunderstood I don't think that's the case here.
Answer: We have modified the sentence to make more explicit what exactly the phylogeographic estimate of wavefront distance entails.
New sentence: "***In addition, we also summarize, for each data set, the evolution of the mean dispersal velocity and of the wavefront distance through time. For the latter, we plot the distance between the estimated location at the root and the lineage that is estimated to be the furthest from the root location, summarized for a series of time-slices of the posterior tree distribution.***"

In the revised manuscript:
In the abstract, in the space of six sentences the authors describe three different things as 'critical'. The repetition aside, I think the hype doesn't help: most readers will appreciate the importance of these issues.
Answer: We edited the abstract and removed the terms "critical"/"important" in the first sentences.

Much of what the authors say about the use of viral genomes in understanding virus spread is also applicable to bacteria; they might consider broadening the terminology from viruses to pathogens.
Answer: Thank you for pointing this out. We now use the term "pathogen" in the beginning of our introduction.

Page 1: "metapopulation dynamics were critical for connecting rural and urban areas".
I find this statement odd. I would not say that the metapopulation dynamics caused this connectedness; they arose because of the (time- and space-varying) connectedness of rural and urban areas.
Answer: We have changed this sentence to ***"…metapopulation dynamics were critical for EBOV dispersal between rural and urban areas…"***.

Page 1: "interventions within the region such as border closures, lockdowns, and restricting travel may more challenging to investigate". Missing "be" after "may".
Answer: Corrected.

Page 1: "...this outbreak occurred in a highly connected region of Africa with large population centres. This connectivity is also important to consider in local management strategies because interacting populations do not necessarily implement policies that are coordinated, as was the case for EBOV in Guinea, Sierra Leone and Liberia."
Personally I would find this clearer if the second sentence were fused with the first, without expanding on what "this outbreak" refers to (the reader knows). e.g. "...this outbreak occurred in a highly connected region of Africa with large population centres, spread over multiple countries, without fully coordinated intervention policies."
Answer: Thank you for this suggestion, we have modified the two sentences accordingly.

Page 1: "The increasing individual-level spatio-temporal disease data... may offer"
Consider removing the opening 'The'.
Answer: We have now made this modification.

Page 2: "In that case, percentages of reduction in epidemic size and duration are estimated for the period of time during which the intervention strategy is effective"
I suggest adding "(i.e. considering only that part of the tree dated after June 2014)". It is important to make this point as clear as possible (perhaps also repeating it in the part of the caption referring to the brown histograms of Fig 1C). I misunderstood this in the original version; anyone who still fails to understand it is likely to conclude that certain interventions can be applied quite late in the epidemic with the same final impact as if they were applied earlier. If such a person decides policy, this would be a fatal misunderstanding.
Answer: We have now made this modification.

Page 2: "Figure S2 provides summarises"
Correction needed.
Answer: Corrected.

Page 3: "Further, the epidemic was generally less likely to spread across international borders, but did so specifically both early and late in the epidemic, between locations that share such an international border."
Suggest replacing 'locations' by a more specific term such as administrative district, otherwise this reads as a tautology. (Continuous motion over an international border must proceed via a location just before the border and one just after; I think that what the authors are trying to say is that a single lineage has been inferred to start in a district touching one side of the border and end in a region touching the other side of the border, rather than leap-frogging a border district.)
Answer: Agreed, we have now replaced “locations” by administrative areas.

Page 3: "we assess how critical such long-distance events were"
"How critical" implies that the events were definitely important, and you assess whether they're merely quite important, or very important. "How important" is the more appropriate neutral phrasing that allows for the possibility that the events were not important.
Answer: We have replaced “critical” by “important”.

Page 4: "Weighted dispersal velocities are both reported in"
-> are reported in both
Answer: Corrected.

Page 4: "events recorded in the MCC (maximum clade credibility),"
"tree" is missing between the closing bracket and the comma
Answer: Corrected.

Page 5: "These estimates provide strong evidence for a significantly reduced frequency in international border crossing from September 2014 (p-value <0.05)"
Apologies for not noticing this in the original draft. p-values should be stated, not provided as inequalities: we are interested in the level of evidence, not the false dichotomy of 'statistically significant' or not.
Answer: We agree but in this particular case, this does not refer to a single p-value (we thus changed this to p-values < 0.05).

Page 5: "with more cases and fatalities than all previous outbreaks combined"
Would "reported outbreaks" be more accurate?
Answer: It should not make much difference, but we now use “reported outbreaks”.

Page 5: "the region has been declared Ebola free"
Ebola-free
Answer: Corrected.

Page 5: "we investigate the two key elements of viral spread in a gravity model"
I think the authors do themselves a disservice here: they have investigated viral spread tout cours, not within the framework of (and subject to the applicability of) a gravity model.
Answer: We realize that this sentence could be misinterpreted. We have now changed it to: ***"In the first part of this study, we investigate the two key elements of a gravity model of spread …"***.

Page 5: "short-distance dispersal realised may be more important"
'realised' should be removed (left over from 'by human mobility').
Answer: Corrected.

Page 5: "This reflects the important role of highly populated locations in maintaining gravity model transmission, as previously identified"
See my previous comment on 'identification' of a gravity model. Also, a gravity model is a description of the situation in which highly populated locations are important for transmission. Saying that highly populated locations are important to maintain the model seems either false or tautological, I can't decide which. I suggest simplifying the statement.
Answer: We have modified this sentence as follows: ***"This reflects the important role of highly populated locations in fuelling EBOV transmission, as previously highlighted"***.

Page 5: "If viral lineage movement to a single capital could have been prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. In contrast, preventing lineage movement to all the capitals reduced epidemic size to about one-third, while their sample size percentage and case count percentage are 28% and 39%, respectively"
In this passage, the three fractional reductions are presented as i) by 15-37%, ii) to onethird, iii) by 28%. This asymmetry initially confused me - I slipped into reading the second "to" as "by", and lost the point being made. It would be easier to read if the second reduction was described as by ~67%.
Answer: We opt not to reduce the information reported here but for clarity, we have now modified these two sentences as follows: ***"If viral lineage movement to a single capital could have been prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. Preventing lineage movement to all three capitals would have reduced epidemic size by two-thirds, while their sample size percentage and case count percentage are 28% and 39%, respectively"***.

Page 6: "may be a reasonable approximated by"
Correction needed.
Answer: Corrected.

Page 6: "In our study we use posterior predictive simulation based on continuous phylogeographic reconstruction to identify a significant decline in international border crossing. Our findings confirm a significant decline following the announcements of border closures between Guinea, Sierra Leone and Liberia, which was also observed using a discrete phylogeographic approach"
The second sentence more or less just repeats the first – consider merging and simplifying.
Answer: We have modified the first sentence as follows: ***"In our study we use posterior predictive simulation in a continuous phylogeographic framework to assess changes in international border crossing through time"***.

Page 7: "allows calculating"
-> "allows the calculation of" (this was correct first time round - 'allow' is transitive)
Answer: Corrected.

Supplementary Figure S3: I suggest removing the legend ("No intervention before July 2014"), as the fact that it describes only one of the two things plotted is confusing. The two things can be described in the caption as for figures 1C and 1D.
Answer: This has been modified.

Appendix S1: "or dispersal velocity estimates"

-> our
Answer: Corrected.

Appendix S1: "each sequence were"
-> was
Answer: Corrected.

Appendix S1: "the first 5% samples"
-> of samples
Answer: Corrected.

Appendix S1: I suggest quoting the velocities here in km/day, or the velocities in the main text in km/year, for easier comparison: the main result of this appendix is that these velocities are similar.
Answer: We now report velocities in km/day in the Appendix.

Revision round #1

2018-01-18
This work offer another very nice illustration of the power of recent advances in phylodynamics when applied to a dataset with dense sampling and rich meta-data (here the location of the infections). It focuses on the recent devastating ebola virus (EBOV) outbreak in West Africa and extends an earlier enormous analysis by Dudas et al. of the by adding a continuous phylogeography approach. It also refines the interpretation of the results by pinpointing the importance of the three capital cities in the magnitude of the outbreak.

Reviewer #1 made some very detailed suggestions and raised a general question about the interpretation of the tree prunning. He/she and Reviewer #2 also made suggestions to broaden the perspective of the article, for instance by discussing epidemiological studies that did not involve phylodynamics to estimate the spread of the epidemics.

In addition to the comments made by the reviewers, I have a couple of my own.

1) Would it be possible to provide confidence intervals for Figure 1 (there are some for panels E, F and G but only for the unprunned tree). The reason why I ask this is because it could help assess the magnitude of the effect. It could also explain why the curves in panels A and B increase at first (I was expecting a steady decrease).

2) Figure 1E is really beautiful! I was wondering if there is an explanation to the fact that recent case counts are below the inferred population size.

3) Figure 2 is also very nice but I expected to be able to find similarities because the sampled data should be the same. However, even the recent timepoints (in blue), which should all I guess be sampled, did not seem to be in the same place.

4) About the model choice (HKY+GAMMA and skygrid) the authors refer to Dudas et al. but it seems that the model choice is not really justified over there, e.g. testing for the most appropriate substitution model. If there is actual support, it would be worth mentionning it. Regarding the details about the priors, I guess the xml files will be made available?

Reviewed by anonymous reviewer, 2017-11-28 15:59
See next page

As a preface to my comments, I note that am not familiar with phylogeographic analyses.

In the abstract the authors summarise their work as showing that
1. "long-distance dispersal events were not crucial for epidemic expansion",
2. "preventing viral lineage movement to single locations would, in most cases, have had little impact",
3. "urban areas – specifically those encompassing the three capital cities and their suburbs – represented major ‘transit centers’ for transmission chains, but preventing viral lineage movement to all three simultaneously would have only contained epidemic size to about one third"
4. there was "considerable heterogeneity in dispersal velocity through time", and
5. "announcements of border closures were followed by a significant but transient effect on international virus dispersal".
These conclusions are drawn by performing state reconstruction of the location of ancestral viruses in a large posterior set of timed phylogenies, allowing inferences about the movements of different viral lineages over time, and exploration of counterfactual interventions that would have stopped certain viral movements. I found that authors' work supported all five of these points, and was well described.

However, the major concern I have is with the relevance of these conclusions for public health interventions in practise. The authors' branch cutting model of preventing transmissions means that when one transmission is blocked, all people who would ultimately have their infection as a result of that transmission are assumed to be perfectly immune / not re-exposed. (The authors do "acknowledge that in the interpretation of intervention scenarios, we assume that all other efforts would have remained unchanged when an introduction to a specific location is hypothetically prevented", which is not the same as the no-second-exposure assumption.) This biases all estimates of intervention effect size downwards, by an unknown amount.

The effect of this bias is relevant to the authors only considering interventions that block transmission with 100% effectiveness. Modelling by Hollingsworth et al (Nature Medicine 2006) showed that for flu, long distance travel bans had to be implemented with >99% efficiency in order to slow epidemic growth from a time scale of days to weeks. How much worse would less-than-perfect interventions have been in the current study? I don't know how this question can be addressed with the authors' current method of cutting branches and removing all descendant lineages from the tree. Removing only x% of the descendants, or cutting the branch with only x% probability, would not capture the relevant effect - these would both trivially reduce the effectiveness to x% of its previous value. The relevant effect is that in the counterfactual where most but not all transmissions to a particular area are blocked, those transmissions that did get through can result in the same people as before becoming infected. If this effect cannot be modelled in the current work, it should at least be added as a strong caveat that as only 100% effective interactions have been considered, and impact depends strongly non-linearly on effectiveness, the effect of preventing long-distance transmission in reality is uncertain.

Modest suggestions & concerns

The authors consider the impact of interventions on reducing epidemic duration. I think that most readers' will be intuitively think that reducing duration is a good thing, because of an assumed link between duration and the final size. However the authors also consider the impact of interventions on the epidemic's final size directly - which is what we really care about - and for a fixed size, reducing epidemic duration is a bad thing. The same sized epidemic concentrated into a shorter time period gives any additional interventions outside of those being modelled less time in which to act. This should be clarified; ideally statements about duration should only be made at the same time as statements about size, less the conclusion be misinterpreted. For example, the authors comment that halting introductions into any one of the three capitals would not have reduced the epidemic duration by much (lower part of Figure S1). The reader will naturally interpret this as "there would have been little point in intervening in only one of the capitals". The authors do not point out that halting introductions into Freetown alone would have reduced the epidemic size by around 40% (upper part of Figure S1). Reducing epidemic size is what we care about - why comment on duration but not size?

Could the authors mention whether a sampling fraction of much less than 100% is expected to affect the inferred dispersal velocity, and if so how? It's not clear to me but I could imagine it results in an overestimate: the serial interval is defined specifically for one infected individual, yet many (95%) of individuals are missing from the tree, so viral lineages are spending a lot of time outside of sampled hosts.

page 5: "The picture that emerges from our phylogeographic analyses is one of multiple moving targets"
Evidence for this (i.e. a metapopulation) would be showing that locations of high burden appear and disappear at different times in different places. It may be that the geographically annotated phylogenies show this, but I don't think the authors have shown that here, aside from the example in Fig 2 suggesting a tendency of the epicentre to drift over time. This picture could be shown by plotting the number of reported cases in different locations as a function of time (without need for phylogeographic analysis).

page 6: "we remove sequences such that monophyletic clusters of sequences sampled from the same administrative region are only represented by a single sequence. Such clusters would largely represent dispersal within administrative regions, which will be determined by the ‘noise’ assigned to their location within an administrative region"
My concern with this procedure is not the uncertainty in the location of the single sequence representing the cluster (which is what I understand by the 'noise' comment), it is the fact that by preferentially removing sequences that seem to be connected by short-distance transmission before fitting a model of diffusion, the authors will upwardly bias the estimated rate of diffusion. I suggest testing for the presence of this bias. If present, and if including all such sequences renders the problem computationally unfeasible, could the rate of diffusion within such close

clusters can be estimated separately, and merged somehow with the estimation where these clusters have been collapsed?

page 6: "We prune each of the posterior trees by removing the same extant taxa as identified in the MCC tree by the pruning selection process."
Why not identify branches with viral lineage movement over distances >d in each posterior tree separately, and then prune the descendant subtree? (Why only identify these branches in the MCC tree and then prune the same taxa in all posterior trees?) This would seem to make better use of the fact that each posterior tree represents a slightly different evolutionary history, and we want to integrate over all possibilities.

page 6/7: the authors describe a procedure for comparing the number of border crossings inferred with the number expected by chance (in a border-unaware simulation), and plot the level of evidence for a discrepancy between the two in Figure 4. Some kind of direct plot of the number of border crossings over time (or the fraction of movements that cross a border, to normalise to the growing epidemic size) would provide a more intuitive visualisation of the effect the authors are trying to test for - a transient decrease after border closures - than Figure 4. As there, the rate of crossing within-country borders could be plotted for comparison. Quantifying the level of evidence is clearly important, but the plot I suggest could be included in addition. This would also show the magnitude of the effect, about which no information is currently given (only the frequency with which N_inferred < N_simulated is presented).

Minor points

The authors could consider citing Ratmann et al. (Science Translational Medicine 2016), who also identified transmission patterns using viral phylogenies and quantified the impact of removing transmissions between certain groups of individuals.

The authors refer to timed phylogenies (with internal annotation of geographic states) as transmission trees. Transmission trees are not the same as phylogenies.

I was confused in a few places whether the velocity being discussed was the velocity of the epidemic wavefront or the velocity of an individual viral lineage.

I read the first two sentences of the abstract as stating that viral genomic data is critical for viral molecular epidemiology. This is essentially tautological, analogous to saying that measurements of human height are critical for studies of human height. A statement of interest would be about how important the study conclusions are. If the intended point is the importance of being able to do viral molecular analyses rapidly, the statement should be reworded to clarify.

page 1: "but preventing viral lineage movement to all three simultaneously would have only contained epidemic size to about one third".

A three-fold reduction in the total size of a large epidemic is substantial; the "but... only" sounds odd.

page 1: "the impact that specific intervention strategies made, had or could have made".
The first two items in this three-item list are identical.

Figure 1 C&D: "% of locations in each range" - does this mean that all samples from the same location only contribute 1 to each bin? If so, this doesn't seem very informative. Or should this be "% of samples in each range"?

page 2: "this GLM approach identified a gravity model of transmission".
A modelling analysis does not 'identify' the correct model, it tests multiple models and may find that one is a better description of reality than the others. All models are wrong, but some are useful.

Figure 2: the legend states that "nodes are coloured according to a colour scale" for which the minimum and maximum are stated, but what the colour itself actually represents is left unsaid. It's fairly obvious but this could easily be clarified.

page 3: "Although about 27% of the genome samples were from these administrative areas"
And what fraction of the total reported cases? It is mentioned in the introduction that sampling intensity correlates well with the infection burden, but it would be helpful to clarify for this example.

Figure 3: the x axis and the legend both state that panel C shows dispersal velocity; I think it shows distance. The legend also contains "1,000 trees sampled sampled" (duplicated "sampled").

page 4: "This result shows that preventing viral lineage movement to these locations halts the dynamic spread of lineages, which in turn continue to generate numerous clusters of cases in other locations, even in different capitals. This stands in contrast to a model of separate, independent and local chains of transmission in each capital city."
A simpler summary would be "This result shows that there was transmission between the capitals." Transmission between capitals (including via external locations as intermediates) is necessary and sufficient for the observation, given the branch-cutting approach for modelling blocked transmissions.

page 4: "The same extent of spatial spread is not achieved for the data set restricted to dispersal events <250 km, indicating that relatively long-distance dispersal events contributed to the maximum epidemic wavefront distance."
How much smaller is the spatial spread? (How big is the contribution of these long-distance events?)

page 5: "short-distance dispersal realised by human mobility"
All mobility here is human. Do the authors mean mobility on foot? If so this seems unwarranted - there is no examination of the method of transport. Cars, bikes etc. could have been used to cover short distances. Best just to leave this statement at "short-distance dispersal" I think.

page 5: "If viral lineage movement to each of the capitals was prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. This emphasises the importance of these cities"
The size of the resulting reduction needs to be balanced against the total number of cases that occurred in the cities. If the latter is 15% to 37%, the observation would be compatible with the cities exporting no cases at all.

page 5: "justify the use of a continuous diffusion process... at least for relatively restricted geographic scales"
Preferable to complete the latter sentence with "such as...". A few hundred kilometres?

page 6: "When unique sampling coordinates are not available for every sequence... we associate a random coordinate within the administrative area of sampling to each sequence."
Was this random coordinate drawn once only for each such sequence, fixing one realisation of a stochastic effect across many replicates, or re-drawn many times (for each realisation of the diffusion model)? That latter would seem to be better, as the advantage of the kind of Bayesian analysis used by the authors is to integrate over uncertainties in the model.

page 6: the authors describe the form of the distributions used for priors, but not the parameters of the priors. Could all relevant information be captured by providing the BEAUti file as supplementary information? If so this would greatly facilitate replication of the analysis.

page 6: the authors provide no information about how the wavefront was calculated.

page 7: "in continental African" -> in continental Africa

page 7: "As illustrated in Figure 3, posterior and posterior predictive diffusion histories are roughly similar except for the position and orientation of branches in the West African study area."
Consider clarifying what it is about these histories that is similar, given that the position and orientation of branches seems to be the dominant feature. Consider removing the final "in the West African study area", which slightly confuses the sentence, hinting that branches are different inside West Africa but similar outside of it.

In supplementary Figure S1, some of the filled bars reach greater values that the corresponding open bars, e.g. for epidemic size for Freetown. How is possible that halting only those introductions into Freetown that occur after July 2014 has greater reduction on the epidemic size than halting all introductions into Freetown at any time?

Reviewed by Christian Althaus, 2017-11-28 15:59

The study by Dellicour and colleagues makes use of phylodynamic analyses for studying the spatial spread of Ebola during the 2013-2016 outbreak in West Africa. The authors extended their previously published phylogeographic framework to examine 1) the potential effect of intervention strategies - such as border closures - and 2) the process of spatial spread by introducing a continuous diffusion process (as opposed to the discrete approach in their earlier analysis). The methods are state-of-the-art and described in sufficient detail. The main findings of the study suggest that the Ebola epidemic was mainly driven by short- rather than long-distance dispersal. Furthermore, the study corroborates the notion that urban transmission was a major contributor to the characteristic spatial transmission dynamics that was observed in West Africa. I found the study rather technical and applying its findings in public health practice is maybe somewhat limited. However, the study is certainly a valuable contribution to the field of phylodynamics and provides an excellent example how genomic analyses can be used to infer the spatial spread of epidemics.

What I was missing a bit was a deeper discussion and comparison of the results to other studies outside the field of phylodynamics that investigated the spatial spread of Ebola and the impact of control interventions (e.g., border closures). The authors briefly mention two key papers using gravity-type models by Backer & Wallinga (ref. 27) and Kramer et al. (ref. 28). Others have estimated the velocity of Ebola spread at 1004 km per year (Zinszer et al., 2015, Lancet Infect Dis, PMID: 26333328) which seems to be in rough agreement with Fig. 3D. I also have a question related to how the authors call this velocity (mean dispersal distance per infection). Shouldn't it be per generation? In my view, it is not a single infection that spreads, but an epidemic that expands over subsequent generations.

Minor comments:
- Methods: The authors associate a random coordinate within the entire administrative area for sampled sequences that have the same geographic coordinates. I was wondering whether this assumption could introduce any sort of bias. For example, if all sequences came from exactly the same place in an otherwise large area, wouldn't associating random coordinates suggest wider spatial spread than what effectively happened?
- Fig. 1A: What is the dashed line on the peak of the distribution of lineage dispersal distances supposed to show?
- Fig. 1C: I could not find the dashed line that is described in the figure caption.
- Fig. 3: The word "sampled" appears twice in the second sentence.
- Fig. 4: I could not find a reference to this figure in the main text of the paper.

Author's reply:

Dear Recommender, Please find uploaded our reply to the comments as well as a tracked changes version. The latter was produced using latex diff, which does not do a perfect job, but we hope it does reflect the major changes to the manuscript. Thank for handling our submission for recommendation and please do not hesitate to let us know if there are any remaining issues. Kind regards, Philippe

See next page

Dr Simon Dellicour
Rega Institute, KULeuven
Minderbroedersstaat 10
3000 Leuven, Belgium

17.01.2018

Dear Managing Board of ***PCI Evol Biol***,

We are grateful for the interesting and constructive comments from the Recommender and the two Referees on our preprint manuscript entitled "Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak".

We addressed these comments in a revised version of our preprint manuscript. We include below a point-by-point reply to the issues raised by the Referees. We believe that the revision has significantly improved our manuscript and we hope it can now be found suitable for recommendation.

Sincerely,

Simon Dellicour,
on behalf of all authors

Recommender:

This work offer another very nice illustration of the power of recent advances in phylodynamics when applied to a dataset with dense sampling and rich meta-data (here the location of the infections). It focuses on the recent devastating ebola virus (EBOV) outbreak in West Africa and extends an earlier enormous analysis by Dudas et al. of the by adding a continuous phylogeography approach. It also refines the interpretation of the results by pinpointing the importance of the three capital cities in the magnitude of the outbreak.

Reviewer #1 made some very detailed suggestions and raised a general question about the interpretation of the tree pruning. He/she and Reviewer #2 also made suggestions to broaden the perspective of the article, for instance by discussing epidemiological studies that did not involve phylodynamics to estimate the spread of the epidemics.

In addition to the comments made by the reviewers, I have a couple of my own.

1) Would it be possible to provide confidence intervals for Figure 1 (there are some for panels E, F and G but only for the unprunned tree). The reason why I ask this is because it could help assess the magnitude of the effect. It could also explain why the curves in panels A and B increase at first (I was expecting a steady decrease).
Answer: We understand the request to represent the uncertainty of the estimates. Concerning the density plots in panels A and B, these are based on the posterior distribution of Markov jumps, so already marginalizing over the evolutionary histories. The reason why these densities increase initially is because we consider jumps between administrative areas and the associated distances are determined between the major population centres in these areas. So, this will not include smaller transmission distances within administrative areas.
We attempted to include credible intervals for the different percentages reported on panels A to D, but the different options we explored overload the overall picture that already contains a lot of information. For clarity reasons, we thus finally decided to not include this additional information in the main figures. However, because one of the estimates was sensitive to conditioning on the MCC tree, we now include the alternative versions of Figures 1A-1D as Supplementary Information, and include a note on this sensitivity issue.
Concerning panels 1E-1G, adding credible intervals for all the effective population size dynamic curves (other than the one reported for the un-pruned tree) would lead to too much overlap. In addition, since we used empirical tree distributions for the pruned trees, the uncertainty would also not be adequately accommodated anyway.

2) Figure 1E is really beautiful! I was wondering if there is an explanation to the fact that recent case counts are below the inferred population size.
Answer: Thank you for appreciating our efforts in making these figures. Bearing in mind that different transformations were used for case counts and effective population sizes, there are indeed some differences for the recent estimates. We can only speculate about potential reasons. First, there could be an underreporting of cases in that recent time period, but it is difficult to understand why that would be the case for the period during which the epidemic is largely contained. Alternatively, it could be a period for which the coalescent has difficulties of inferring appropriate population sizes. In this period, prevalence declined such that only a handful of lineages persisted, but these are relatively divergent lineages. So the coalescent could be misled by the maintenance of a relatively large diversity through these remaining lineages, and a smoothed estimate of population size change may simply reflect the loss over time of these lineages. Because we cannot substantiate these explanations and since we only draw on the large-scale correspondence of the coalescent estimates, we did not discuss this further detail in the revised manuscript.

3) Figure 2 is also very nice but I expected to be able to find similarities because the sampled data should be the same. However, even the recent time points (in blue), which should all I guess be sampled, did not seem to be in the same place.
Answer: Because we made forward-in-time simulations and not backward-in-time simulations we cannot expect to reproduce the same coordinates as the sampling distribution. The point of this

posterior predictive procedure is to simulate a diffusion process using the same parameters as those estimated from the data on the same genealogies as estimated from the data. The simulations are then parameterised according to the estimates from the real data, but the outcome of these simulations does not need to match the realised pattern of spread. It serves as a 'landscape-unaware' model of spread against which we can contrast the realized pattern of spread.

4) About the model choice (HKY+GAMMA and skygrid) the authors refer to Dudas et al. but it seems that the model choice is not really justified over there, e.g. testing for the most appropriate substitution model. If there is actual support, it would be worth mentioning it. Regarding the details about the priors, I guess the xml files will be made available?
Answer: We indeed adopted of nucleotide substitution parameterisation from Dudas ***et al***. that is customised for this Ebola genomic data. The customisation lies in the partitioning of data, with 3 partitions for the three codon positions and an additional partition for the intergenic regions. For each of the 4 partitions, we specify an independent HKY model, an independent gamma distribution for among-site rate heterogeneity (ASRH), and we allow for relative rate differences among the partitions. We did not assess the model fit of this parameterization because:

1) There is a gigantic number of parameterisations to test against in terms of how the data can be partitioned and how these partitions are crossed with specific nucleotide substitution models including or not ASRH. Common substitution model testing procedures do not consider partitioning although adequately modelling rate heterogeneity is far more important than modelling the differences in the type of substitutions.
2) We strongly believe that it is more important to ask whether the estimates of interest are sensitive to the model choice than to ask which model from a pre-specified limited collection of models may fit the data best. Substitution model parameters are nuisance parameters in the inference of the diffusion process that we study. Nevertheless, the HKY model is likely to capture the most pronounced signal about differences in types of substitutions because the kappa estimates all range between 8 and 14 for the partitions.

All the xml files are now available online (https://github.com/ebov/space-time/tree/master/Analyses/continuousDiffusion).

Reviewer 1:

As a preface to my comments, I note that am not familiar with phylogeographic analyses. In the abstract the authors summarise their work as showing that
1. "long-distance dispersal events were not crucial for epidemic expansion",
2. "preventing viral lineage movement to single locations would, in most cases, have had little impact",
3. "urban areas – specifically those encompassing the three capital cities and their suburbs - represented major 'transit centers' for transmission chains, but preventing viral lineage movement to all three simultaneously would have only contained epidemic size to about one third"
4. there was "considerable heterogeneity in dispersal velocity through time", and 5. "announcements of border closures were followed by a significant but transient effect on international virus dispersal".
These conclusions are drawn by performing state reconstruction of the location of ancestral viruses in a large posterior set of timed phylogenies, allowing inferences about the movements of different viral lineages over time, and exploration of counterfactual interventions that would have stopped certain viral movements. I found that authors' work supported all five of these points, and was well described.

However, the major concern I have is with the relevance of these conclusions for public health interventions in practise. The authors' branch cutting model of preventing transmissions means that when one transmission is blocked, all people who would ultimately have their infection as a result of that transmission are assumed to be perfectly immune / not re-exposed. (The authors do "acknowledge that in the interpretation of intervention scenarios, we assume that all other efforts would have remained unchanged when an introduction to a specific location is hypothetically prevented", which is not the same as the no-second-exposure assumption.) This biases all estimates of intervention effect size downwards, by an unknown amount. The effect of this bias is relevant to the authors only considering interventions that block transmission with 100% effectiveness. Modelling by Hollingsworth et al (Nature Medicine 2006) showed that for flu, long distance travel bans had to be

implemented with >99% efficiency in order to slow epidemic growth from a time scale of days to weeks. How much worse would less-than-perfect interventions have been in the current study? I don't know how this question can be addressed with the authors' current method of cutting branches and removing all descendant lineages from the tree. Removing only x% of the descendants, or cutting the branch with only x% probability, would not capture the relevant effect - these would both trivially reduce the effectiveness to x% of its previous value. The relevant effect is that in the counterfactual where most but not all transmissions to a particular area are blocked, those transmissions that did get through can result in the same people as before becoming infected. If this effect cannot be modelled in the current work, it should at least be added as a strong caveat that as only 100% effective interactions have been considered, and impact depends strongly non-linearly on effectiveness, the effect of preventing long-distance transmission in reality is uncertain.

Answer: We fully agree that our approach of preventing transmission through pruning phylogenetic lineages assumes that relevant people were not exposed to the virus in other ways. Our approach conditions on the reconstruction of a realised pattern, so it is not a model that can take on a different course. We also considered cutting branches with a particular probability, but as the reviewer indicates, this would not adequately capture the effect of less-than-perfect interventions. We follow the reviewer's recommendation to better discuss the limitations of our approach, specifically that we consider no re-exposure and 100% effective interactions:

"***Our phylogenetic approach of assessing hypothetical containment strategies rests on a number of assumptions, with a 100% effectiveness of their implementation being an important one. While it would be straightforward to introduce a probability on the effectiveness of preventing the movement events we target, quantifying the corresponding impact using our phylogenetic measures may not be so relevant. Even if only a fraction of movements is allowed to escape prevention, the resulting transmission chains in the relevant area may have put everyone at risk of infection. In other words, our approach needs to assume that persons that were not infected by a particular lineage, because its transmission was halted, were not exposed to other transmission chains that were not contained. Our phylodynamic approach therefore offers a best-case scenario as starting point, and different degrees of effectiveness and its potential non-linear impact on outcome may be further examined using computational models. Further investigations will be important to assess whether interventions such as travel restrictions can in practice be implemented with reasonable success. In the case of air travel and influenza spread for example, travel restrictions were shown to be practically unfeasible to effectively contain the international spread of a pandemic (Hollingsworth et al. 2006)***".

Modest suggestions & concerns

The authors consider the impact of interventions on reducing epidemic duration. I think that most readers' will be intuitively think that reducing duration is a good thing, because of an assumed link between duration and the final size. However the authors also consider the impact of interventions on the epidemic's final size directly - which is what we really care about - and for a fixed size, reducing epidemic duration is a bad thing. The same sized epidemic concentrated into a shorter time period gives any additional interventions outside of those being modelled less time in which to act. This should be clarified; ideally statements about duration should only be made at the same time as statements about size, less the conclusion be misinterpreted. For example, the authors comment that halting introductions into any one of the three capitals would not have reduced the epidemic duration by much (lower part of Figure S1). The reader will naturally interpret this as "there would have been little point in intervening in only one of the capitals". The authors do not point out that halting introductions into Freetown alone would have reduced the epidemic size by around 40% (upper part of Figure S1). Reducing epidemic size is what we care about - why comment on duration but not size?

Answer: We agree and we now note that epidemic size, and not duration, is the most relevant measure to evaluate the impact of containment strategies before reporting the results and at the end of the results.

Could the authors mention whether a sampling fraction of much less than 100% is expected to affect the inferred dispersal velocity, and if so how? It's not clear to me but I could imagine it results in an overestimate: the serial interval is defined specifically for one infected individual, yet many (95%) of

individuals are missing from the tree, so viral lineages are spending a lot of time outside of sampled hosts.
Answer: With a sampling fraction of about 5% of the known cases, viral lineages are indeed spending time outside of the sampled hosts, but we believe the relevant question is how much time they have spent outside of the reconstructed history. Because our sampling shows a very good correlation with case counts through time in administrative areas, it is likely to adequately capture the large-scale transmission history. So, many additional samples will have come from local transmission chains and would increase the density of tips in our phylogenetic reconstructions, but that would not add considerable much ancestral history to our reconstruction. Because of this and the fact we only took a single representative sequence for a cluster of admin-specific sequences, our dispersal velocity estimate will reflect the rate of lineages that largely spread across different admin regions (and not within admin regions). We further address this in reply to the comment about pruning short-distance transmission.

Page 5: "The picture that emerges from our phylogeographic analyses is one of multiple moving targets" Evidence for this (i.e. a metapopulation) would be showing that locations of high burden appear and disappear at different times in different places. It may be that the geographically annotated phylogenies show this, but I don't think the authors have shown that here, aside from the example in Fig 2 suggesting a tendency of the epicentre to drift over time. This picture could be shown by plotting the number of reported cases in different locations as a function of time (without need for phylogeographic analysis).
Answer: The demonstration of metapopulation dynamics was indeed more explicit in our previous work, and as part of this, we have already created an animation of both the phylogeographic reconstruction and the case count evolution (the latter is shown by the colour intensity for the administrative regions):
https://github.com/ebov/space-time/blob/master/Visualizations/EBOV_animation.HD.264.mp4
We now explicit cite our previous study when we refer to the metapopulation dynamics.

Page 6: "we remove sequences such that monophyletic clusters of sequences sampled from the same administrative region are only represented by a single sequence. Such clusters would largely represent dispersal within administrative regions, which will be determined by the noise assigned to their location within an administrative region" My concern with this procedure is not the uncertainty in the location of the single sequence representing the cluster (which is what I understand by the 'noise' comment), it is the fact that by preferentially removing sequences that seem to be connected by short-distance transmission before fitting a model of diffusion, the authors will upwardly bias the estimated rate of diffusion. I suggest testing for the presence of this bias. If present, and if including all such sequences renders the problem computationally unfeasible, could the rate of diffusion within such close clusters can be estimated separately, and merged somehow with the estimation where these clusters have been collapsed?
Answer: We believe the Reviewer raises an important point. First, we would like to clarify that for most sequences, there were no specific geographic coordinates available within the administrative area. Therefore, estimating diffusion rates based on clusters of sequences that only constitute samples within an area would not be possible. It is because of the absence of this information as well as the computational challenges of integrating over the administrative areas for unknown locations that we decided to reduce those clusters to one representative sequence. As a result, our procedure ignores a lot of short-distance transmission within administrative areas. To investigate to what extent this would affect our dispersal rate estimate, we now analysed a data set of sequences for which more precise geographic coordinates were available and for which we did not restrict monophyletic clusters of sequences from the same administrative area to a single representative sequence. While this data set therefore accommodates transmission within administrative areas, and also differs in the time interval and total area of sampling (Sierra Leone), the dispersal velocity estimates are remarkable consistent with our previous estimates. This offers reasonable reassurance that our procedure does not result in strong biases. This additional analysis and related results are now described in the Methods, reported in the Results and included as Supplementary Information.

Page 6: "We prune each of the posterior trees by removing the same extant taxa as identified in the MCC tree by the pruning selection process." Why not identify branches with viral lineage movement over distances >d in each posterior tree separately, and then prune the descendant subtree? (Why only

identify these branches in the MCC tree and then prune the same taxa in all posterior trees?) This would seem to make better use of the fact that each posterior tree represents a slightly different evolutionary history, and we want to integrate over all possibilities.
Answer: We indeed had to remove exactly the same taxa across all trees in the posterior so that we could also use the posterior genealogies to infer the demographic trajectories (our proxy of viral effective size through time). We would not be able to do this coalescent inference while averaging over empirical trees that contain different taxa. This is now explicitly mentioned in the Methods section. In addition, we have now examined how sensitive the tree height and tree length summaries are to conditioning on the MCC tree. We report the results in the Supplementary Information in the form of two new figures: one that includes credible intervals for reductions in epidemic size and duration and one that is based on pruning that does not condition on the MCC tree. For the less important epidemic duration measure, we noticed that the reduction associated with preventing spread to administrative locations with >1,000k people is highly uncertain and not represented well by the MCC tree. We also highlight this in the main manuscript and thank the Reviewer for encouraging us to examine this.

Page 6/7: the authors describe a procedure for comparing the number of border crossings inferred with the number expected by chance (in a border-unaware simulation), and plot the level of evidence for a discrepancy between the two in Figure 4. Some kind of direct plot of the number of border crossings over time (or the fraction of movements that cross a border, to normalise to the growing epidemic size) would provide a more intuitive visualisation of the effect the authors are trying to test for - a transient decrease after border closures - than Figure 4. As there, the rate of crossing within-country borders could be plotted for comparison. Quantifying the level of evidence is clearly important, but the plot I suggest could be included in addition. This would also show the magnitude of the effect, about which no information is currently given (only the frequency with which N_inferred < N_simulated is presented).
Answer: We agree and plots with the frequency of border crossing for both the estimated and simulated distributions are included as Figure S2.

Minor points

The authors could consider citing Ratmann et al. (Science Translational Medicine 2016), who also identified transmission patterns using viral phylogenies and quantified the impact of removing transmissions between certain groups of individuals.
Answer: Thank you for the suggestion, we now refer to this study when starting our exposé on the phylogenetic pruning procedure.

The authors refer to timed phylogenies (with internal annotation of geographic states) as transmission trees. Transmission trees are not the same as phylogenies.
Answer: We have now replaced the expression “transmission tree” by “phylogenetic tree”.

I was confused in a few places whether the velocity being discussed was the velocity of the epidemic wavefront or the velocity of an individual viral lineage.
Answer: We have edited the text to avoid the confusion between these two concepts (“velocity of the epidemic wavefront” is now clearly distinguished from “the weighted/mean dispersal velocity”).

I read the first two sentences of the abstract as stating that viral genomic data is critical for viral molecular epidemiology. This is essentially tautological, analogous to saying that measurements of human height are critical for studies of human height. A statement of interest would be about how important the study conclusions are. If the intended point is the importance of being able to do viral molecular analyses rapidly, the statement should be reworded to clarify.
Answer: The second sentence has now been modified.

Page 1: "but preventing viral lineage movement to all three simultaneously would have only contained epidemic size to about one third". A three-fold reduction in the total size of a large epidemic is substantial; the "but... only" sounds odd.
Answer: We corrected the text.

Page 1: "the impact that specific intervention strategies made, had or could have made". The first two items in this three-item list are identical.
Answer: Corrected.

Figure 1 C&D: "% of locations in each range" - does this mean that all samples from the same location only contribute 1 to each bin? If so, this doesn't seem very informative. Or should this be "% of samples in each range"?
Answer: Indeed, it was actually the percentage of samples. This mistake has been corrected.

Page 2: "this GLM approach identified a gravity model of transmission". A modelling analysis does not 'identify' the correct model, it tests multiple models and may find that one is a better description of reality than the others. All models are wrong, but some are useful.
Answer: This has been reworded.

Figure 2: the legend states that "nodes are coloured according to a colour scale" for which the minimum and maximum are stated, but what the colour itself actually represents is left unsaid. It's fairly obvious but this could easily be clarified.
Answer: This has been clarified.

Page 3: "Although about 27% of the genome samples were from these administrative areas" And what fraction of the total reported cases? It is mentioned in the introduction that sampling intensity correlates well with the infection burden, but it would be helpful to clarify for this example.
Answer: This specific information has been added.

Figure 3: the x axis and the legend both state that panel C shows dispersal velocity; I think it shows distance. The legend also contains "1,000 trees sampled sampled" (duplicated "sampled").
Answer: Panel C shows the density for velocity (in km per infection). The duplicate "sampled" has been removed.

Page 4: "This result shows that preventing viral lineage movement to these locations halts the dynamic spread of lineages, which in turn continue to generate numerous clusters of cases in other locations, even in different capitals. This stands in contrast to a model of separate, independent and local chains of transmission in each capital city." A simpler summary would be "This result shows that there was transmission between the capitals." Transmission between capitals (including via external locations as intermediates) is necessary and sufficient for the observation, given the branch-cutting approach for modelling blocked transmissions.
Answer: We agree, and have modified the text accordingly.

Page 4: "The same extent of spatial spread is not achieved for the data set restricted to dispersal events <250 km, indicating that relatively long-distance dispersal events contributed to the maximum epidemic wavefront distance." How much smaller is the spatial spread? (How big is the contribution of these long-distance events?)
Answer: This information has been added in the text: "***Our phylogeographic estimates of the epidemic wavefront through time indicate that EBOV spread up to ~500 km from its location of origin in about 8 to 9 months (Fig. 3A). With a maximum wavefront distance of ~400 km, the same extent of spatial spread is not achieved for the data set restricted to dispersal events <250 km, indicating that relatively long-distance dispersal events contributed to the maximum epidemic wavefront distance***".

Page 5: "short-distance dispersal realised by human mobility" All mobility here is human. Do the authors mean mobility on foot? If so this seems unwarranted - there is no examination of the method of transport. Cars, bikes etc. could have been used to cover short distances. Best just to leave this statement at "short-distance dispersal" I think.
Answer: Corrected.

Page 5: "If viral lineage movement to each of the capitals was prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. This emphasises the importance of these cities". The size of the resulting reduction needs to be balanced against the total number of cases that occurred in the cities. If the latter is 15% to 37%, the observation would be compatible with the cities exporting no cases at all.
Answer: For clarity, we now contrast this result with the percentages of reduction obtained when preventing lineage movement to all the capitals together: "***If viral lineage movement to a single capital could have been prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. In contrast, preventing lineage movement to all the capitals reduced epidemic size to about one-third, while their sample size percentage and case count percentage are 28% and 39%, respectively. This theoretical result emphasises the importance of urban transmission, but at the same time, it indicates that no single capital was critical for the maintenance of all co-circulating lineages***".

Page 5: "justify the use of a continuous diffusion process... at least for relatively restricted geographic scales" Preferable to complete the latter sentence with "such as...". A few hundred kilometres?
Answer: We prefer not to associate a particular distance to this, as it remains to be determined how large the scale can be. Moreover, the appropriate scale will also depend on the mobility dynamics in the area of interest. We refer to a 'restricted scale' to remain cautious.

Page 6: "When unique sampling coordinates are not available for every sequence... we associate a random coordinate within the administrative area of sampling to each sequence." Was this random coordinate drawn once only for each such sequence, fixing one realisation of a stochastic effect across many replicates, or re-drawn many times (for each realisation of the diffusion model)? That latter would seem to be better, as the advantage of the kind of Bayesian analysis used by the authors is to integrate over uncertainties in the model.
Answer: In a first attempt, we indeed tried to integrate out every location over the relevant administrative area. While this approach works well for a restricted number of tips, we were confronted with significant mixing issues when trying to do this for all tips. We are currently working on new integration techniques for unobserved locations that we will hopefully be able to use for such purposes in the future.

Page 6: the authors describe the form of the distributions used for priors, but not the parameters of the priors. Could all relevant information be captured by providing the BEAUti file as supplementary information? If so this would greatly facilitate replication of the analysis.
Answer: Yes, the xml files will be made available.

Page 6: the authors provide no information about how the wavefront was calculated.
Answer: We have now added a sentence to describe this as well as a reference to the Methods section.

Page 7: "in continental African" -> in continental Africa.
Answer: Corrected.

Page 7: "As illustrated in Figure 3, posterior and posterior predictive diffusion histories are roughly similar except for the position and orientation of branches in the West African study area." Consider clarifying what it is about these histories that is similar, given that the position and orientation of branches seems to be the dominant feature. Consider removing the final "in the West African study area", which slightly confuses the sentence, hinting that branches are different inside West Africa but similar outside of it.
Answer: We have now clarified this sentence (and this should actually refer to Figure 2): "***Figure 2 illustrates the difference in position and orientation of branches between a diffusion history reconstructed from the data and a diffusion history simulated using our posterior predictive simulation procedure***".

In supplementary Figure S1, some of the filled bars reach greater values that the corresponding open bars, e.g. for epidemic size for Freetown. How is possible that halting only those introductions into

Freetown that occur after July 2014 has greater reduction on the epidemic size than halting all introductions into Freetown at any time?
Answer: We apologize for the lack of clarity about this. In the case of the delayed intervention strategies, percentages of reduction were computed relative to period impacted by the intervention strategy (after June 2014). This explains why we can observe higher reductions in this case. This has now been clarified (in the Results section as well as in the legend of Figure 1).

Review 2:

The study by Dellicour and colleagues makes use of phylodynamic analyses for studying the spatial spread of Ebola during the 2013-2016 outbreak in West Africa. The authors extended their previously published phylogeographic framework to examine 1) the potential effect of intervention strategies - such as border closures - and 2) the process of spatial spread by introducing a continuous diffusion process (as opposed to the discrete approach in their earlier analysis). The methods are state-of-the-art and described in sufficient detail. The main findings of the study suggest that the Ebola epidemic was mainly driven by short- rather than long-distance dispersal. Furthermore, the study corroborates the notion that urban transmission was a major contributor to the characteristic spatial transmission dynamics that was observed in West Africa. I found the study rather technical and applying its findings in public health practice is maybe somewhat limited. However, the study is certainly a valuable contribution to the field of phylodynamics and provides an excellent example how genomic analyses can be used to infer the spatial spread of epidemics.

What I was missing a bit was a deeper discussion and comparison of the results to other studies outside the field of phylodynamics that investigated the spatial spread of Ebola and the impact of control interventions (e.g., border closures). The authors briefly mention two key papers using gravity-type models by Backer & Wallinga (ref. 27) and Kramer et al. (ref. 28). Others have estimated the velocity of Ebola spread at 1004 km per year (Zinszer et al., 2015, Lancet Infect Dis, PMID: 26333328) which seems to be in rough agreement with Fig. 3D.
Answer: We agree and added a comparison to specific modelling results to the discussion. In addition, we now also better integrate modelling studies in the first paragraph of the introduction. Specifically, we refer to containment efforts, how they have been investigated using mathematical modelling, and how sequence data and phylodynamic approaches may complement these to evaluate long-range interventions.

I also have a question related to how the authors call this velocity (mean dispersal distance per infection). Shouldn't it be per generation? In my view, it is not a single infection that spreads, but an epidemic that expands over subsequent generations.
Answer: We are afraid that the word 'generation' may be too vague, as there is host generation, viral generation, transmission generation, etc. If we say 'per infection', we do not imply that it is a single infection that spreads just as we would not imply a single generation if we would use 'per generation'.

Minor comments:

- Methods: The authors associate a random coordinate within the entire administrative area for sampled sequences that have the same geographic coordinates. I was wondering whether this assumption could introduce any sort of bias. For example, if all sequences came from exactly the same place in an otherwise large area, wouldn't associating random coordinates suggest wider spatial spread than what effectively happened?
Answer: For a detailed answer, we refer to our reply to the related comment by Reviewer 1. In short, we have now performed an additional analysis to show that our rate estimate is robust to this (as reported in the new Appendix in Supplementary Information).

- Fig. 1A: What is the dashed line on the peak of the distribution of lineage dispersal distances supposed to show?

Answer: It corresponds to the distribution of distances associated with lineage dispersal occurring after June 2014. We apologize for the fact that this information was lacking in the previous version of the legend.

- Fig. 1C: I could not find the dashed line that is described in the figure caption.
Answer: We have now solved this issue.

- Fig. 3: The word "sampled" appears twice in the second sentence.
Answer: Corrected.

- Fig. 4: I could not find a reference to this figure in the main text of the paper.
Answer: Corrected.

Phylodynamic assessment of intervention strategies for the West African Ebola virus outbreak

Simon Dellicour[1,*], Guy Baele[1], Gytis Dudas[2], Nuno R. Faria[3], Oliver G. Pybus[3], Marc A. Suchard[4,5,6], Andrew Rambaut[7,8], and Philippe Lemey[1]

[1] Department of Microbiology and Immunology, Rega Institute, KU Leuven – University of Leuven, Herestraat 49, 3000 Leuven, Belgium

[2] Fred Hutchinson Cancer Research Center, 1100 Fairview Ave N, 98109, Seattle, WA, USA

[3] Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom

[4] Department of Biostatistics, UCLA Fielding School of Public Health, University of California, Los Angeles, CA, USA

[5] Department of Biomathematics David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA

[6] Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, CA, USA

[7] Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh, EH9 3FL, UK

[8] Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

∗Corresponding author (simon.dellicour@kuleuven.be)

The recent Ebola virus (EBOV) outbreak in West Africa witnessed considerable efforts to obtain viral genomic data as the epidemic was unfolding. ~~**Such data are critical for the investigation of viral molecular epidemiology and can complement contact tracing**~~ **If such data can be deployed in real-time, molecular epidemiological investigations can take up a critical role in complementing contact tracing efforts by public health agencies. Analysing the accumulated EBOV genetic data can also deliver important insights into epidemic dynamics, as demonstrated by a recent viral genome study that revealed a metapopulation pattern of spread. Although metapopulation dynamics were critical for connecting rural and urban areas during the epidemic, the implications for specific intervention scenarios remain unclear. Here, we address this question using a collection of phylodynamic approaches. We show that long-distance dispersal events (between administrative areas >250 km apart) were not crucial for epidemic expansion and that preventing viral lineage movement to** ~~**single locations**~~ **a specific administrative area would, in most cases, have had little impact. In addition, urban areas – specifically those encompassing the three capital cities and their suburbs –** ~~**represented major 'transit centers' for transmission chains, but**~~ **were critical in attracting and further disseminating the virus, and preventing viral lineage movement to all three simultaneously would have** ~~**only**~~ **contained epidemic size to about one third. Using continuous phylogeographic reconstructions we estimate a distance kernel for EBOV spread and reveal considerable heterogeneity in dispersal velocity through time. We also show that announcements of border closures were followed by a significant but transient effect on international virus dispersal.** ~~**Our study illustrates how phylodynamic analyses can answer specific epidemiological and epidemic control questions and can be used to quantify**~~ **By quantifying the hypothetical impact of different intervention strategies as well as the impact of barriers on dispersal frequency**~~**.**~~**, our study illustrates how phylodynamic analyses can help to address specific epidemiological and outbreak control questions.**

Keywords: Ebola virus, West Africa, phylogeography, spatial epidemiology, intervention strategies, ~~**borders impact**~~**border closures.**

The recent Ebola virus (EBOV) epidemic in West Africa emerged around end of 2013 in the prefecture of Guéckédou in Guinea[1] and caused at least 11,310 deaths among 28,616 recorded cases in Guinea, Sierra Leone and Liberia[2]. It has been argued that the West African EBOV epidemic ~~illustrates~~ illustrated problems in the early detection of, and rapid response to, infectious disease outbreaks of public health importance[3]. ~~In general, inadequate surveillance and/or reporting systems can cause delayed detection and underreporting. In addition, misdiagnosis and a lack of real-time data sharing impedes understanding of the magnitude of an epidemic[3]. The use of rapid point-of-care (POC) diagnostic tests to improve the speed of diagnosis can improve outbreak control, and the promise of effective assistance can help to overcome reluctance to report disease events[3].~~ Various reasons may explain the slow initial response to the West African EBOV epidemic, including poor public health infrastructure and local unfamiliarity with Ebola virus disease (EVD), as well as a lack of preparedness by the international community. ~~This highlights the need for effective systems to detect and stop infectious disease threats and for rapidly expanding international responsecapacity when confronted with such threats[4]. Improving infection prevention and control in health care settingsis critical in these efforts~~

Because efforts to control the epidemic could not rely on vaccination or effective antiviral drugs, the outbreak response focused on standard medical practices (e.g. case identification and isolation) as well as community practices (e.g. sanitary funeral practices)[5]. The dynamics of EBOV transmission (e.g. superspreading events[6]) and the performance of local containment measures that were applied as part of the international response[7–9], as well as the potential impact of other hypothetical strategies (e.g. the use of rapid diagnostic test that were not available yet[10]), have been studied extensively through mathematical modelling. Concerning spatial transmission dynamics, the impact of air travel restrictions out of the affected region has also been assessed in detail (e.g. Poletto *et al.*[11]), but long-range interventions within the region such as border closures, lockdowns, and restricting travel may more challenging to investigate. They are however important to consider because, unlike previous outbreaks that were confined to remote villages, this outbreak occurred in a highly connected region of Africa with large population centres[5,12]. This connectivity is also important to consider in local management strategies because interacting populations do not necessarily implement policies that are coordinated, as was the case for EBOV in Guinea, Sierra Leone and Liberia[5]. The increasing individual-level spatio-temporal disease data, e.g. made available through mobile phone data[12], may offer invaluable opportunities to accommodate connectivity in modelling efforts (e.g. Lau *et al.*[13]). Viral genetic data represents an interesting alternative or complementary data source because it contains information about spatio-temporal spread that can be extracted using phylodynamic approaches. While routinely used for epidemiological reconstructions, opportunities to harness the power of evolutionary approaches to inform intervention strategies have remained scarce. Now that genomic surveillance systems can be deployed for real-time viral genome sequencing in resource-limited settings[14], it is critical to examine what relevant information for control strategies can be gleaned from viral genome data.

Viral genome sequencing is increasingly being used to assist with the identification of unknown sources and transmission chains, as viral genomes contain valuable information that complements contact tracing efforts. In the case of Ebola, Arias *et al.*[15] demonstrated that rapid outbreak sequencing in locally-established sequencing facilities can identify transmission chains linked to sporadic cases. Consequently, it is unsurprising that there have been many calls for making sequence data openly available in outbreak situations[3,16–18]. In addition to identifying specific transmission pathways, viral genome analyses can also shed light on the origins, evolution and transmission dynamics of a virus during an epidemic[19]. Early in the EBOV epidemic, analyses such as those by Gire and colleagues[20] demonstrated that the virus entered the human population in late 2013 and crossed from Guinea to Sierra Leone in May 2014 through sustained human-to-human transmission. The EBOV genome data that was generated also stimulated phylodynamic efforts to characterise transmission dynamics early in the epidemic (e.g. superspreading[21]) and to estimate critical epidemiological parameters such as the basic reproductive number[22]. Various molecular epidemiological studies subsequently attempted to trace Ebola spread[14,15,23–27] (see Holmes *et al.*[19] for a detailed overview), marking the beginning of large-scale real-time molecular epidemiology[19]. All these efforts culminated in an impressive collection of over 1,600 EBOV genome sequences corresponding to more than 5% of known cases[28]. These data represent a unique opportunity to learn important lessons about the evolutionary and

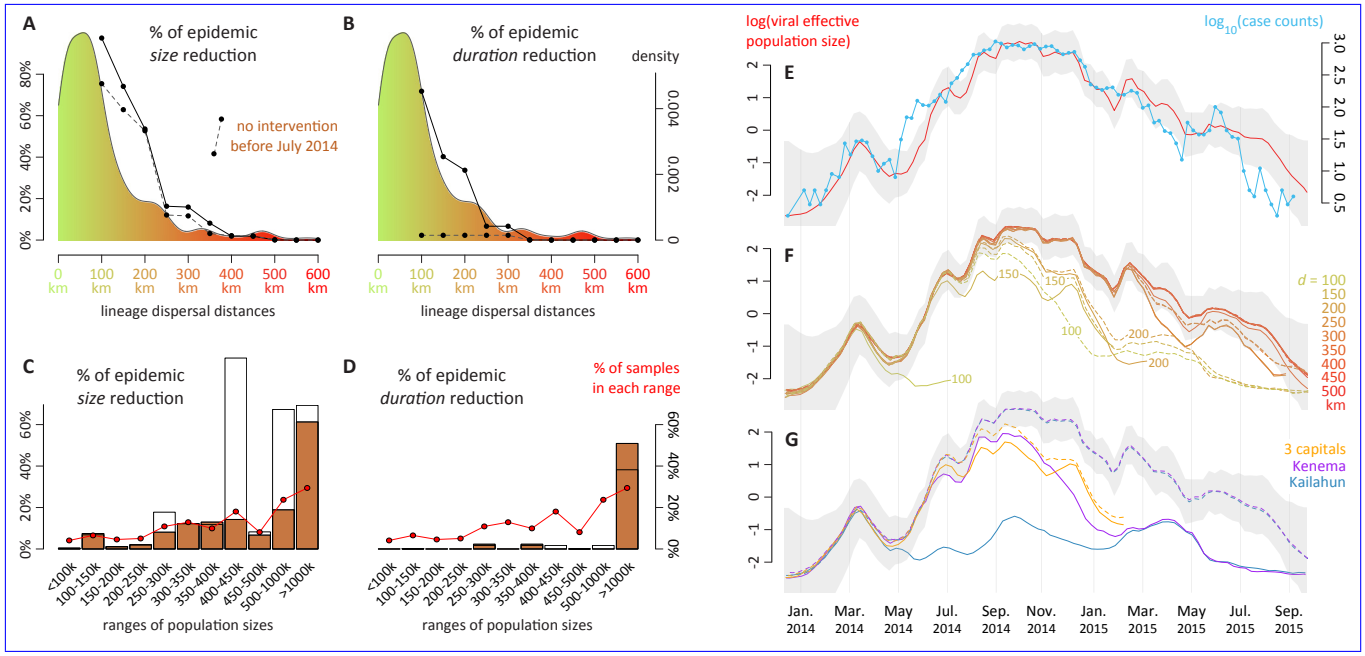

Figure 1. Hypothetical impact of intervention strategies. (**A**) Impact of halting transmission following long-distance dispersal events on *epidemic size*. The plot depicts the epidemic size reduction when removing transmission following dispersal events over various distances, i.e. dispersal events between administrative areas separated by a least 100, 150, 200, 250, 300, 350, 400, 450 and 500 km. The effect on epidemic size is quantified by comparing (in %) the resulting tree length to the original phylogenetic tree. In addition, we also report corresponding epidemic size reductions when transmission is prevented only after June 2014 (dashed line). In that case, percentages of reduction in epidemic size and duration are estimated for the period of time during which the intervention strategy is effective. Both epidemic size reduction curves are superimposed on the distribution of lineage dispersal distances summarised based on the posterior Markov jump history (coloured from green to red and as inferred by discrete phylogeographic analysis of the outbreak in West Africa). (**B**) Impact of removing transmission following long-distance dispersal events on *epidemic duration*. This plot corresponds to Figure 1A but focuses on the impact on epidemic duration. In this case, the effect on epidemic duration is quantified by comparing (in %) the resulting tree height to the original phylogenetic tree. (**C**) Impact of preventing dispersal events to specific administrative areas on *epidemic size*. The plot reports the epidemic size reduction when removing transmission following dispersal events in administrative areas belonging to one of the following population sizes ranges: <100k, 100-150k, 150-200k, 200-250k, 250-300k, 300-350k, 350-400k, 400-450k, 450-500k, 500-1000k and >1000k people. As for Figure 1A, the effect on epidemic size is quantified by comparing (in %) the resulting tree length to the original phylogenetic tree, and we also report corresponding epidemic size reductions when transmission is prevented only after June 2014 (brown histogram). (**D**) Impact of removing transmission following dispersal events in specific administrative areas on the *epidemic duration*. This plot corresponds to 1C but focuses on the impact on epidemic duration. As for Figure 1B, the effect on epidemic duration is quantified by comparing (in %) the resulting tree height to the original phylogenetic tree. We refer to Supplementary Figure S1 for credible intervals associated with percentages of epidemic size/duration reductions reported in Figures 1A-D. All these percentages of reduction have been computed by conditioning the pruning on transition events recorded in the MCC (maximum clade credibility) tree summary of the discrete phylogeographic reconstruction. Supplementary Figure S2 provides summarises the equivalent results for pruning trees based on Markov jump histories associated with each posterior tree. (**E**) Estimates of viral population size (in red; 95% HPD in grey) and the evolution of the case counts (in green). (**F**) Impact of removing transmission following long-distance dispersal events on viral effective population size through time. As in Figure 1E, 95% HPD of viral population size based on the entire dataset (no intervention strategy) is displayed in grey. On this graph, dashed lines correspond to viral population size evolution when transmission is prevented only after June 2014. (**G**) This plot corresponds to Figure 1F but focuses on the impact of removing transmission following dispersal events to specific locations on viral effective population size through time.

epidemiological dynamics of an Ebola outbreak. Here, we aim to contribute to this progress by investigating how viral genome analyses can assess the impact that specific intervention strategies made ~~, had~~ or could have made if implemented.

~~**Hypothetical impact of intervention strategies.** (A) Impact of halting transmission following long-distance dispersal events on the *epidemic size*. The plot depicts the epidemic size reduction when removing transmission following dispersal events over various distances, i.e. dispersal events between administrative areas separated by a least 100, 150, 200, 250, 300, 350, 400, 450 and 500 km. The effect on epidemic size is quantified by comparing (in %) the resulting tree length to the original transmission tree. In addition, we also report corresponding epidemic size reductions when transmission is prevented only after June 2014 (dashed line). These epidemic size reduction curves are superimposed on the distribution of lineage dispersal distances within the transmission tree (as inferred by discrete phylogeographic analysis of the outbreak in West Africa). (B) Impact of removing transmission following long-distance dispersal events on the *epidemic duration*. This plot corresponds to Figure 1A, but focuses on the impact on epidemic duration. In this case, the effect on epidemic duration is quantified by comparing (in %) the resulting tree height to the original transmission tree. (C) Impact of preventing dispersal events to specific administrative areas on the *epidemic size*. The plot reports the epidemic size reduction when removing transmission following dispersal events in administrative areas belonging to one of the following population sizes ranges: <100k, 100-150k, 150-200k, 200-250k, 250-300k, 300-350k, 350-400k, 400-450k, 450-500k, 500-1000k and >1000k people. As for Figure 1A, the effect on epidemic size is quantified by comparing (in %) the resulting tree length to the original transmission tree, and we also report corresponding epidemic size reductions when transmission is prevented only after June 2014 (dashed line). (D) Impact of removing transmission following dispersal events in specific administrative areas on the *epidemic duration*. This plot corresponds to 1C but focuses on the impact on epidemic duration. As for Figure 1B, the effect on epidemic duration is quantified by comparing (in %) the resulting tree height to the original transmission tree. (E) Estimation of viral population size (in red; 95% HPD in grey) and the evolution of the case counts (in green). (F) Impact of removing transmission following long-distance dispersal events on the viral effective population size through time. As in Figure 1E, 95% HPD of viral population size based on the entire dataset (no intervention strategy) is displayed in grey. On this graph, dashed lines correspond to viral population size evolution when transmission is prevented only after June 2014 (dashed line). (G) This plot corresponds to Figure 1F but focuses on the impact of removing transmission following dispersal events to specific locations on the viral effective population size through time.~~

Although Ebola viral genomes were reported across numerous studies, the collected genetic data cover the epidemic exceptionally well, and sampling intensity correlates well with the infection burden in different locations throughout the course of the outbreak[28]. This data set motivated a detailed phylogeographic study that identified the patterns and drivers of spatial spread[28]. Specifically, a generalised linear model (GLM) of ~~transitions~~ transition rates between discrete locations in a Bayesian statistical framework was used to test which

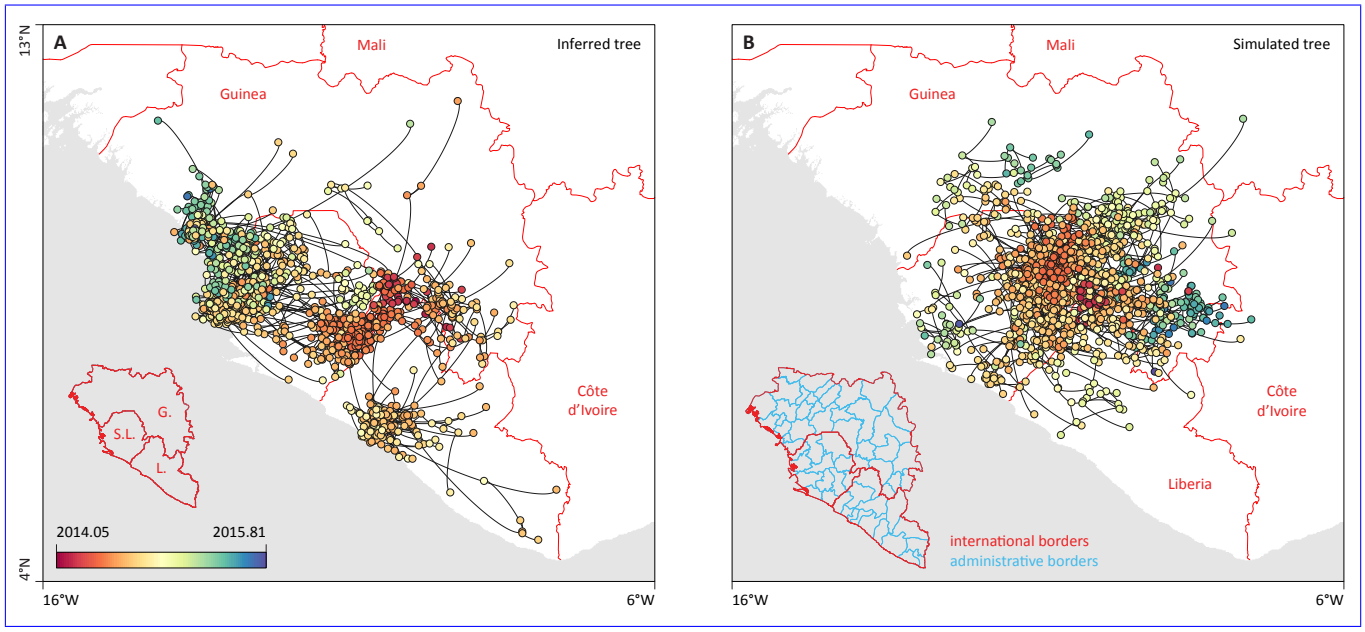

Figure 2. **Example of a continuous phylogeographic estimate (A) and the corresponding posterior predictive simulation unaware of international borders (B).** In both cases, phylogenetic nodes are coloured according to their time of occurrence.

causal factors might have influenced the spread of the virus at sub-national administrative levels (termed districts in Sierra Leone, prefectures in Guinea, and counties in Liberia). By considering a range of geographic, administrative, economic, climatic, infrastructural and demographic predictors, this GLM approach ~~identified~~ provided support for a gravity model of transmission, albeit one that was attenuated by international borders[28]. The gravity model emphasises the impact of population size on viral dispersal and implies that large urban populations acted as sources, re-seeding smaller limited epidemics in more outlying locations. Further, the epidemic was generally less likely to spread across international borders, but did so specifically both early and late in the epidemic, between locations that share such an international border. More detailed temporal analyses suggested that border attenuation may have resulted from border closures between Guinea, Sierra Leone and Liberia, although their containment effects were limited. Within the three affected countries, viral spread was not always maintained by continuous transmission in each location, but often by repeated introductions into a location, generating small, well-connected, clusters of cases. This dynamical pattern of connectivity characterises a metapopulation, highlighting the need for responsive, mobile and measured interventions.

Here, we extend the phylogeographic analyses of the West African Ebola epidemic in two different ways. First, we examine the implications of EBOV metapopulation dynamics on particular intervention strategies. Specifically, we assess to what extent limiting long-distance spread, or preventing spread to highly populated locations, might have impacted the epidemic. Second, we introduce continuous diffusion models as an alternative phylogeographic framework, which can characterise aspects of the process of Ebola spread that were not captured by the discrete approach employed by Dudas *et al*[28]. We quantify important parameters of spatial spread and demonstrate how a posterior predictive simulation procedure can be used to evaluate potential barriers to transmission, specifically, the impact of border closures. These new evolutionary approaches deepen our understanding of the public health implications of EBOV epidemic dynamics and the extent to which viral spread could be curbed by particular intervention strategies.

RESULTS

~~**Example of a continuous phylogeographic estimate (A) and the corresponding posterior predictive simulation unaware of international borders (B).** In both cases, phylogenetic nodes are coloured according to a colour scale ranging from red (most recent common ancestor) to blue (most recent sampling time).~~

To understand the implications of EBOV metapopulation dynamics during the 2013-2016 epidemic, we first investigate the impact of hypothetical intervention strategies on epidemic size and duration. To model the effects of reducing long-distance dispersal (e.g. through travel restrictions), we prune the EBOV ~~transmission~~ phylogenetic trees when a long-distance lineage translocation between administrative areas was inferred, by effectively removing transmission following such dispersal events. We then quantify the effect of this intervention on the ~~size (tree length) and duration (tree height) of the transmission trees~~ epidemic size as reflected in the reduction in tree length in the phylogenetic reconstructions (Fig. 1 A). ~~This procedure~~therefore assesses We also report reductions in epidemic duration (based on tree height), but note that epidemic size is the most relevant measure to evaluate the impact of containment strategies. Using this procedure, we assess how critical such long-distance events were to the expansion and maintenance of EBOV transmission. For comparison, we also perform ~~this~~ the analysis by preventing transmission only after a particular point in time, i.e. by only removing viral lineage movements that occurred after June 2014. The latter analysis reflects a scenario in which hypothetical intervention strategies are implemented only some time after the onset of the outbreak, in this case after 6 months, at which time all three countries had already been seeded. This also avoids the effect of pruning early transmission bottlenecks that are specific to this epidemic, which could affect the generality of our conclusions. For the delayed intervention strategies, percentages of reduction in epidemic size and duration are estimated relative to the period of time during which the intervention strategy is effective.

Figures 1A and 1B depict the impact of preventing long-distance viral lineage movements on relative epidemic size and duration, respectively, under both the immediate and delayed intervention strategies. Transitions between administrative areas involving distances greater than 300 km are rare and the transmission chains they generate do not contribute substantially to the total epidemic size (Fig. 1A). Only if viral lineage movement is impeded over shorter distances ($<$250 km) do we start to observe a stronger impact on relative epidemic size. This is the case if lineage movements are prevented ~~throughout the entire duration of the epidemic~~from the initial stages of the outbreak. If viral lineage movement is restricted only after June 2014, then the epidemic size effect diverges ~~(97%/75% of reduction) at smaller~~ at small distances (e.g. a 75% epidemic size reduction at 100 km instead of a 97% reduction without this restriction, Fig. 1A). Epidemic duration is similarly affected if short-distance transitions are prevented from the start of the epidemic, but not if the restriction is implemented only after June 2014 (Fig. 1B). This implies that specific transitions between administrative areas with distances larger than $>$100 km early in the epidemic were critical for generating long-term transmission chains, and that after June 2014, epidemic duration could have been largely maintained by viral lineage movements over shorter distances, albeit at smaller epidemic size.

To further investigate the effects of potential interventions on reducing epidemic size through time, we also undertook coalescent ~~estimation~~ inference of viral effective population size through time, which ~~appears to be~~ results in estimates that are remarkably proportional to case counts (Fig. 1E). This relationship can be statistically tested by using ~~an~~ a GLM-based extension of the coalescent approach that considers case counts as a potential covariate of viral effective population size[28,29]. The GLM coefficient for the association between case counts and estimated effective population size is high (0.55) and its credible intervals exclude zero (95% highest posterior density interval [HPD]: 0.18-0.90), indicating a significant association. The effective population size

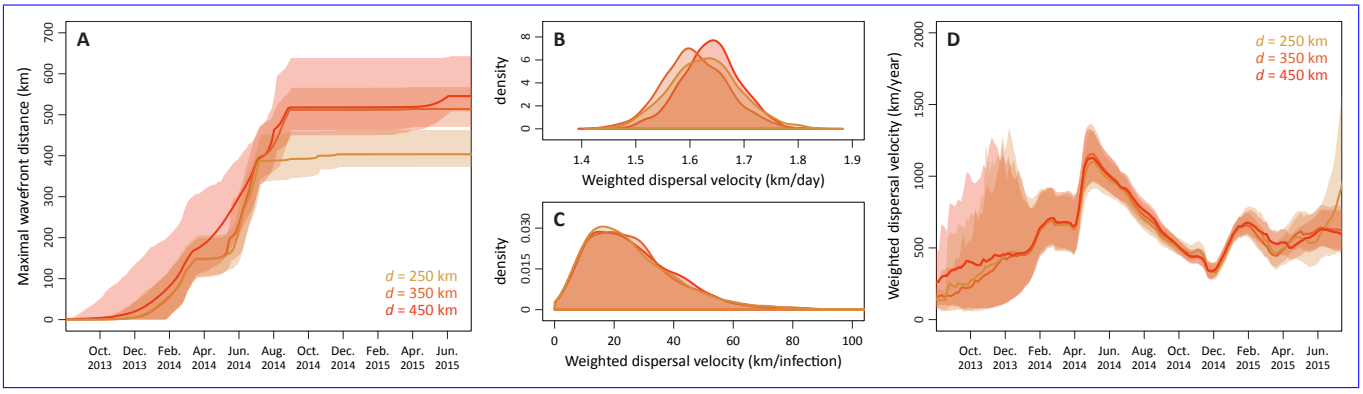

Figure 3. **Evolution of maximal wavefront distance (A), mean dispersal velocity (B-C) and dispersal velocity through time (D) for each data set (*d* = 250, 350 and 450 km).** These statistics were summarized from 1,000 trees sampled from the posterior distribution of trees. Weighted dispersal velocities are both reported in km per day (**B**) and in km per infection (**C**). Weighted dispersal velocities in km per infection are obtained by multiplying mean branch velocities (in km/day) by serial interval values (in days between successive infections) randomly drawn from the generation time distribution estimated by the WHO Ebola Response Team (2014).

estimates indicate that the impact of pruning dispersal events between administrative areas ≥200 km apart from the ~~transmission~~ phylogenetic trees has a notable effect on epidemic size only after the time of the epidemic peak (Fig. 1F). These estimates also illustrate the pronounced effect on epidemic size and duration of applying the delay on preventing viral lineage movement over smaller distances. If viral lineage movement had been prevented over distances >100 km between administrative areas at the onset of the outbreak, then the epidemic would have been restricted to an initial small peak in epidemic size (Fig. 1C), which represents the emergence in the Guéckédou prefecture and neighbouring areas[28].

We next apply a similar procedure, but this time restricting viral lineage movement according to the population sizes of the 'destination' (administrative area) of viral lineage movement. Specifically, we bin areas according to their population size and remove all descendent transmission (subtrees) that follow lineage dispersal events to corresponding areas. We then examine the effect of this restriction on epidemic size and duration assuming, as in the distance-based pruning, that the restriction results from a 100% effective intervention. Figures 1C and 1D summarise the relative reduction in epidemic size and duration, with and without a delay on transmission prevention (filled and open bars respectively), as well as the sample sizes from all areas within each population size range. For two population size ranges (~~350-400k~~ 400-450k and 500-1000k), the presence or absence of a delay on transmission prevention has a large effect on the observed reduction in epidemic size. By examining the impact of each administrative area separately (Supplementary Figure ~~S1~~S3), we can attribute this difference to the impact of preventing viral lineage movement to the Kailahun and Kenema districts in Sierra Leone prior to June 2014. These administrative areas represent early transmission bottlenecks that are specific to this EBOV epidemic. The virus spread extremely rapidly from Kailahun district to several counties of Liberia[25] and Guinea[23,27]. However, preventing early viral lineage movement to Kailahun district would not have noticeably reduced the duration of the epidemic (Fig. 1G), because a basal phylogenetic lineage specific to Guinea would have remained unaffected by this restriction and would have continued to circulate, albeit with a limited epidemic burden. We acknowledge that in the interpretation of intervention scenarios, we ~~assume that~~ not only assume that they are 100% effective, but also that people for which infection was prevented in particular areas would not have been infected through other introductions, and that all other efforts would have remained unchanged~~when an introduction to a specific location is hypothetically prevented. However~~. With respect to the latter however, a localised lineage causing limited cases in Guinea as a consequence of ~~such intervention~~ interventions would have been more easily contained. Preventing early spread to Kenema district would ~~have also~~ also have halted much of the westward spread from Kailahun district, but its impact on epidemic size is smaller (Fig. 1G), because Kenema district was much less important than Kailahun district for viral spread to Liberia. From these two districts, EBOV spread over relatively long distances to other administrative areas, which is why preventing early lineage movements over such distances generates a strong reduction in epidemic size in the distance-based examination.

~~**Evolution of maximal wavefront distance (A), mean dispersal velocity (B-C) and dispersal velocity through time (D) for each data set (*d* = 250, 350 and 450 km).** These statistics were estimated from 1,000 trees sampled sampled from the posterior distribution of trees. Weighted dispersal velocities are both reported in km per day (**B**) and in km per infection (**C**). Weighted dispersal velocities in km per infection are obtained by multiplying mean branch velocities (in km/day) by serial interval values (in days between successive infections) randomly drawn from the generation time distribution estimated by the WHO Ebola Response Team (2014).~~

When applying the delay to the intervention strategies, the effect of preventing viral lineage movement to sets of administrative areas with different population sizes generally had limited impact on predicted epidemic size (Fig. 1C) and virtually no impact on epidemic duration, except for the case where viral lineage movement was prevented to areas with population sizes >1,000,000. This category corresponds precisely to the areas encompassed by the three capitals, i.e. Montserrado, Freetown (and suburbs) and Conakry. Although about ~~27~~28% of the genome samples were from these administrative areas (and about 39% of reported cases), removal of viral lineages that moved into these areas lead to a disproportionate reduction in epidemic size of about 60%, which starts to take effect before the epidemic peak (Fig. 1G). This reflects the important role of highly populated locations in maintaining gravity-model transmission, as previously identified[28]. We note that the epidemic size reduction for the three capitals together is less than the sum of the reduction obtained by removing viral lineage movement to each capital individually (Supplementary Figure ~~S1). This result shows that preventing viral lineage movement to these locations halts the dynamic spread of lineages, which in turn continue to generate numerous clusters of cases in other locations, even in different capitals. This stands in contrast to a model of separate, independent and local chains of transmission in each capital city. We also~~ S3), indicating that there was transmission between the capitals. Furthermore, we observe a substantial reduction in epidemic duration ~~only~~ when preventing the viral lineage movement to all the three capitals (i.e. administrative areas with >1,000k people; Fig. 1D), but not ~~preventing the~~ when preventing viral lineage movement to a single capital (Supplementary Figure ~~S1). This result indicates that they were jointly~~ S3). However, this result appears to be sensitive to conditioning the tree pruning on transition events recorded in the MCC (maximum clade credibility), as this is one of the few trees in the posterior distribution for which a capital is important for the ~~spread and maintenance of a handful of lineages that persisted over some time in 2015.~~ maintenance of all residual lineages in the late stage of the epidemic (see the comparison between Supplementary Figures S1 and S2). So, epidemic duration is not only less relevant for assessing intervention strategies, it is also less robust to the phylogenetic procedure we employ.

In the second part of this study, we use a new continuous phylogeographic approach to ~~address further~~ further address key aspects of EBOV spatial spread. ~~Although viral spread driven by human mobility~~ Because human mobility can make pathogen diffusion highly irregular when measured against geographic distance, viral spread may not always be adequately modelled by a ~~Brownian diffusion process~~Brownian-like diffusion process[30,31]. However, the strongly distance-dependent diffusion of the Ebola virus (Fig. 1A) justifies its use on this relatively restricted geographic scale. Here we employ a relaxed random walk (RRW) model that accommodates diffusion rate heterogeneity among lineages. We assess the sensitivity of our continuous phylogeographic reconstructions to long-distance dispersal by analysing the data sets that were pruned based on discrete transitions larger than 450 km, 350 km and 250 km (~~cfr~~ see above). Figure 2A illustrates an EBOV ~~transmission~~ phylogenetic tree estimated under this model of continuous phylogeographic ~~inference~~diffusion, which is spatially mapped onto the study area.

Our phylogeographic estimates of the epidemic wavefront through time indicate that EBOV spread up to ~500 km from its location of origin in about

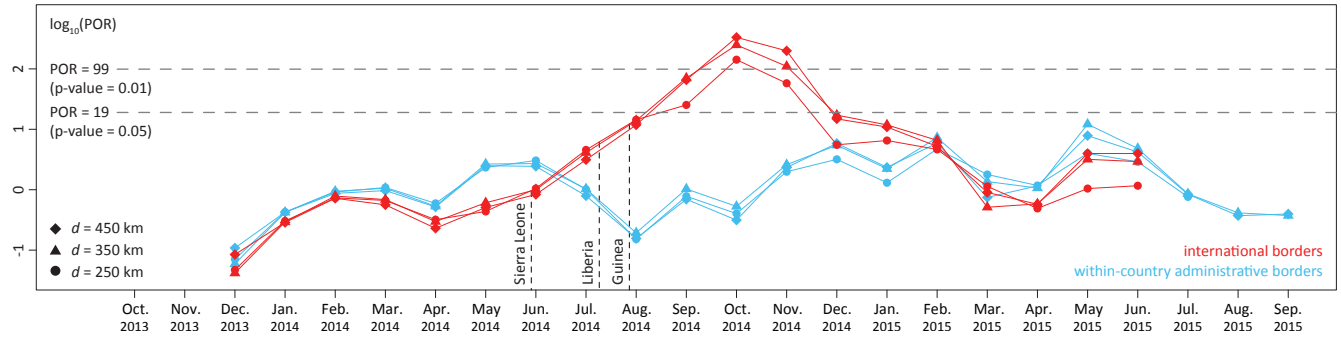

Figure 4. Analysis of the monthly impact of borders on the EBOV dispersal frequency. The plot depicts the predictive odds ratio (POR) estimates per month and per data set (d = 250, 350 and 450 km). PORs estimates for these borders are also included as a negative control as we do not expect any significant impact of within-country administrative borders on dispersal frequency. PORs >3 and >20 can be considered as 'positive' and 'strong' evidence for the impact of borders on the dispersal frequency. Vertical hashed lines indicate the time at which Sierra Leone, Liberia and Guinea announced their border closure. We refer to Supplementary Figure ~~S2~~ S4 for a more detailed representation of the monthly differences in ~~terms of~~ crossing border events ~~among~~ between inferred and simulated diffusion processes.

8 to 9 months (Fig. 3A). ~~The~~ With a maximum wavefront distance of ~400 km, the same extent of spatial spread is not achieved for the data set restricted to dispersal events <250 km, indicating that relatively long-distance dispersal events contributed to the maximum epidemic wavefront distance. The ~~mean~~ velocity of the epidemic wavefront from the beginning of 2014 to early September 2014 is ~1.9 km/day~~.~~, which is consistent with, but still smaller than the estimated velocity of spread of 2.8 km/day based on weekly counts of confirmed cases per administrative area[32]. Our continuous phylogeographic approach also estimates a mean dispersal velocity of 1.64 km/day (95% HPD [1.52, 1.74]), with little variability across data sets (Fig. 3B). Based on a mean serial interval of transmission of 15.3 days and its uncertainty (SD = 9.3 days[33]), this translates to a mean dispersal distance of 25.4 km per infection (95% HPD [5.33, 61.21]; Fig. 3C) confirming that ~~viral infections may spread because of~~ transmission was fuelled by relatively high mobility of infected individuals in this region of Africa[12]. Our estimates are mostly informed by dispersal between districts, prefectures and counties (see Methods), and therefore ignores a lot of local transmission within these administrative areas. However, similar analysis of an alternative data set of genomes from Sierra Leone with more precise coordinates and multiple samples per administrative area provided largely consistent estimates (see Appendix S1). The mean dispersal velocity shows a remarkable variability during the course of the epidemic (Fig. 3D). It steeply increases from January 2014 to a peak around May/June, coinciding with the time of spread across the region from Kailahun. The peak in mean dispersal velocity is followed by a marked drop until the end of the year 2014. This drop appears to begin before the announced border closures (Sierra Leone on 11 June 2014, Liberia on 27 July 2014, and Guinea on 9 August 2014). In addition, we observe the same pattern even if we summarise dispersal velocity only for branches that do not cross borders (as determined by their inferred node locations), suggesting that the decrease in dispersal velocity may be attributed to the impact of a more general awareness of the outbreak, and of the emerging response against it.

Although the mean dispersal velocity may be affected by factors other than border closures, phylogeographic testing based on the discrete phylogeographic approach showed that the frequency at which national borders were crossed by viral lineages significantly decreased following the announced border closures between the countries[28]. In order to evaluate this using our continuous phylogeographic approach, we developed a test procedure that compares the estimated posterior frequency of border crossing events to the same frequency in posterior predictive simulations that are unaware of borders (~~cfr. Methods; see also~~ see the Methods section as well as Figure 2B for an example of such simulation). We quantify the deviation in estimated border frequency crossing from the expected frequency obtained by posterior simulation as a predictive odds ratio, and calculate posterior predictive p-values to assess statistical significance. We also perform this analysis using the data sets for d = 250, 350 and 450 km and include the frequency of within-country administrative border crossing as a negative control (because we do not expect any significant changes in this frequency as a result of border ~~closure~~closures). In Figure ~~3~~4, we plot the predictive odds ratio estimates per month (see also Supplementary Figure ~~S2~~ S4 for a more detailed representation of the monthly differences in crossing border events among inferred and simulated diffusion processes). These estimates provide strong evidence for a significantly reduced frequency in international border crossing from September 2014 (p-value <0.05), i.e. starting shortly after the announced border closures (Fig. ~~2~~4), to about the end of the year. As expected, the frequency of within-country administrative border crossing does not depart significantly from the border-unaware expectation.

DISCUSSION

West Africa has experienced the largest outbreak of Ebola virus in history, with more cases and fatalities than all previous outbreaks combined since 1976. Although the region has been declared Ebola free since 2016, it remains critically important to learn as much as possible from this devastating epidemic. Our study builds on a recent viral genome analysis of EBOV in West Africa by investigating further the implications of the metapopulation dynamics revealed using phylogeographic inference[28]. Specifically, we investigate the two key elements of viral spread in a gravity model, distance and population size, by measuring the predicted reduction in epidemic size ~~and duration~~ (and duration) that results from restricting viral lineage movement, either by preventing viral lineage movements ~~of~~ over different distances, or by preventing movement ~~events~~ to areas of different population size. The latter can be used to evaluate the impact of hypothetical intervention strategies. We found that long-distance dispersal events were not critical for epidemic expansion, and only when dispersal events are restricted to 200 km or less did we observe significant reductions in epidemic size. While this result does not immediately translate into practical intervention strategies, it suggests that frequent short-distance dispersal realised ~~by human mobility~~ may be more important than rarer long-distance dispersal events in driving the EBOV epidemic spread.

We demonstrate that the contribution of population size to the previously identified gravity model of EBOV spread[28] is almost entirely driven by viral lineage movement to and from the areas encompassing the three capital cities, which are the most highly populated areas in the outbreak region. The fact that the West African EBOV epidemic also affected urban areas, in addition to rural areas, ~~represents a critical difference between it and~~ makes it stand apart from all previous EBOV outbreaks. If viral lineage movement to ~~each of the capitals was~~ a single capital could have been prevented, beginning from the onset of the epidemic, then epidemic size could have been reduced by 15% to 37%. ~~This~~ Preventing lineage movement to all the capitals would have reduced epidemic size to about one-third, while their sample size percentage and case count percentage are 28% and 39%, respectively. This theoretical result emphasises the importance of ~~these cities~~urban transmission, but at the same time, it indicates that no single capital was critical for the maintenance of all co-circulating lineages. ~~Only if viral lineage movement to all three capitals was prevented, would the epidemic size have been reduced to about one-third. Such an intervention would also have a considerable impact on epidemic duration. We therefore conclude that, collectively, urban transmission was a major contributor to the metapopulation dynamics underlying this epidemic.~~ The inability to strongly reduce epidemic size by preventing viral lineage movement to other collections of locations further underscores the highly pervasive and distributed nature of the metapopulation ~~dynamic~~ dynamics underlying the epidemic. ~~The~~ Together with the metapopulation dynamics highlighted by Dudas *et al.*[28], the picture that emerges from our phylogeographic analyses is one of multiple moving targets: potential intervention strategies that are piecemeal, reactive and geographically restricted are predicted to have a limited impact on epidemic size and duration. ~~This result argues~~ These dynamics argue for coordinated intervention strategies across the whole outbreak region. By applying a delay to the hypothetical intervention strategies, which will be realistic for most outbreaks, we avoided the impact of early transmission bottlenecks that were identified in Kailahun district and Kenema district in Sierra Leone. The importance of these transmission bottlenecks to the 2014-2016 EBOV outbreak, and their role in spreading the virus to other areas – in particular for Kailahun district – has been highlighted before[23,25,27].

Our phylogenetic approach of assessing hypothetical containment strategies rests on a number of assumptions, with a 100% effectiveness of their implementation being an important one. While it would be straightforward to introduce a probability on the effectiveness of preventing the movement events we target, quantifying the corresponding impact using our phylogenetic measures may not be so relevant. Even if only a fraction of movements is allowed to escape prevention, the resulting transmission chains in the relevant area may have put everyone at risk of infection. In other words, our approach needs to assume that persons that were not infected by a particular lineage, because its transmission was halted, were not exposed to other transmission chains that were not contained. Our phylodynamic approach therefore offers a best-case scenario as starting point, and different degrees of effectiveness and its potential non-linear impact on outcome may be further examined using computational models. Further investigations will be important to assess whether interventions such as travel restrictions can in practice be implemented with reasonable success. In the case of air travel and influenza spread for example, travel restrictions were shown to be practically unfeasible to effectively contain the international spread of a pandemic[34].

Phylogeographic reconstructions in continuous space have been used primarily for animal viruses because ~~relaxed random walk diffusion may be a reasonable approximation for~~ the dispersal of their ~~host~~hosts may be a reasonable approximated by a relaxed random walk process[35–39]. However, the distance-dependent transmission dynamics of EBOV also justify the use of a continuous diffusion process for phylogeographic reconstruction of this outbreak. We present this as an alternative and complementary approach for the study of spatial epidemiological dynamics in human populations, at least for relatively restricted geographic scales. Continuous phylogeographic reconstructions enable us to quantify several aspects of dispersal dynamics, such as mean dispersal velocity per infection (25.4 km per infection [5.33, 61.21]). We observe a strong heterogeneity in this mean dispersal velocity over time, with a significant decrease from May/June 2014 until the end of that year. This likely reflects the general impact of control strategies and awareness on human behaviour.

~~In our study we use posterior predictive simulation to identify a significant decline in international border crossing. Our findings confirm a significant decline following the announcements of border closures between Guinea, Sierra Leone and Liberia, which was also observed using a discrete phylogeographic approach[28]~~Several of our findings are in line with or are complementary to the results of modelling studies. Based on case occurrence data, Backer *et al*.[40] ~~previously~~ estimated that only 4-10% of newly infected EBOV cases migrated to another district and that, among these migrants, only 0-23% left their country of origin. ~~Similarly,~~ Kramer *et al*.[41] ~~have~~ used a spatial network approach to demonstrate that the probability of viral lineage movement between locations depended on international border closures. Phylogeographic approaches can contribute important insights by directly inferring the historical connections underlying viral spread. In our study we use posterior predictive simulation based on continuous phylogeographic reconstruction to identify a significant decline in international border crossing. Our findings confirm a significant decline following the announcements of border closures between Guinea, Sierra Leone and Liberia, which was also observed using a discrete phylogeographic approach[28]. The procedure we used here has a number of advantages relative to the discrete phylogeographic approach that uses epoch modelling to incorporate time-varying predictors[28,42]. A continuous phylogeographic reconstruction does not require a prior specification of the number of change-points, which, in the previous EBOV analysis, was restricted to a single change-point[28]. Instead, we flexibly identify the relevant time period by using a statistic that deviates from the null expectation. In addition, we avoid a discrete approach with large state spaces ~~in the discrete approach~~ that arise from large numbers of locations and ~~which~~ that are computationally expensive, despite the ability to employ multicore GPU architecture[43].

Our analysis expands on the findings from discrete phylogeographic reconstructions and underlines how border closure decisions may spatially structure an epidemic without necessarily having a strong containment effect. The methodology presented in this study could be used to study the impact of potential barriers on the epidemic spread of other important pathogens. Assessing the impact of both hypothesised intervention strategies, as well as the impact of border closures is of interest for public health agencies and policy makers and may provide a better general understanding of outbreak dynamics.

METHODS

Assessing the impact of preventing long-distance or location-specific viral lineage movements. Similar to Ratmann *et al*.[44], we use a phylogenetic pruning approach to investigate prevention strategies, but now specifically relying on the embedded estimates of spatial spread. We build on the phylogeographic reconstruction performed by Dudas *et al*.[28], who used a ~~GLM-parameterization~~ GLM-parameterisation of discrete phylogenetic diffusion[45]. Based on a data set of 1,610 viral genomes sampled between 17 March 2014 and 24 October 2015 (available at https://github.com/ebov/space-time~~/data/~~), Dudas *et al*.[28] used this approach to reconstruct a history of lineage movements between 56 administrative regions in Guinea (prefectures), Sierra Leone (districts) and Liberia (counties). Their Bayesian inference resulted in a posterior distribution of time-measured trees, each annotated with inferred ancestral locations, which was summarised as a maximum clade credibility (MCC) tree. In order to assess the impact of preventing viral lineage movement over specific geographic distances, or to specific locations, we condition on the full transition history in the MCC tree recorded using Markov jumps. Markov jump estimation provides a stochastic mapping of the realisations of the continuous-time Markov process throughout evolutionary history[46,47]. Because each transition between a pair of locations is associated with a geographic distance (the great-circle distance between the locations' population centroids), we are able to assess the impact of preventing viral lineage movement over distances $>d$ by pruning from the complete tree all subtrees that represent the transmission history following branches that accommodate such transition. As possible values for d, we test a series of decreasing distances: 500, 450, 400, 350, 300, 250, 200, 150 and 100 km; this yields pruned trees with 1607, 1567, 1567, 1498, 1368, 1242, 875, 383 and 53 sequences, respectively. ~~We prune each of the posterior trees by removing the same extant taxa as identified in~~ In order to prune the ~~MCC tree by the pruning selection process~~same taxa from all posterior trees (using PAUP*[48]), such that pruned trees can also be used as empirical tree distributions in subsequent coalescent inference (cfr. below), we condition on the Markov jump history in the MCC tree to determine which taxa need to be pruned. However, for the measures of epidemic size and duration, we examine the sensitivity to conditioning on the MCC tree by also pruning each tree of the posterior distribution based on its specific Markov jump history (Supplementary Figure S2). From the resulting set of pruned posterior trees, we ~~record the~~ use the program TreeStats[49] to compute the tree length (the sum of all branch lengths) and tree height (the time to the most recent common ancestor), which we interpret as measures of relative epidemic size and epidemic duration. In order to also obtain an estimate of the effect on relative epidemic size through time, we make use of coalescent estimates under a flexible Bayesian skygrid model[50]. For the complete data set, the effective population size estimates are highly correlated with case counts through time (Fig. 1A). Conditioning on a posterior subset of 1,000 pruned ~~empirical genealogies from the posterior tree distribution~~genealogies, we re-estimate effective population sizes through time using the Bayesian skygrid model and compare these estimates to the original coalescent estimates. All the effective population size plots are based on summaries from the program Tracer 1.7.

We follow a similar procedure to assess the impact of preventing viral lineage movement to a specific category of administrative areas or to individual administrative areas. Different categories of administrative areas were defined on the basis of their population size: $<$100k, 100-150k, 150-200k, 200-250k, 250-300k, 300-350k, 350-400k, 400-450k, 450-500k, 500-1000k and $>$1,000k people. In these analyses, we pruned subtrees from the complete tree that were the result of transitions from any location to the location(s) within the category under consideration. As before, we obtain the estimates of epidemic duration, relative size, and relative size through time.

Bayesian skygrid estimation with covariates. To assess the strength of association between case counts and effective population size through time, we use a recent extension of the ~~non-parameteric~~ non-parametric Bayesian skygrid model that incorporates potential covariates[29,29]. This approach allows us to include external time series as covariates in a GLM framework while accounting for demographic uncertainty. By applying this GLM framework to the complete genome data set[28], we model the Ebola outbreak effective population size as a log-linear function of case counts and estimate the effect sizes for the latter as a GLM coefficient.

Continuous phylogeographic inference. As an alternative to discrete phylogeographic inference, we estimate the spatiotemporal dynamics of Ebola virus by inferring viral lineage movements in continuous space with a multivariate diffusion approach implemented in BEAST[51]. This approach infers ~~the~~ ancestral locations (in geographic coordinates) of internal nodes using a (relaxed) random walk diffusion process. In order to assess the impact of long-distance dispersal on the continuous diffusion estimates, we analyse the pruned data sets obtained using the procedure described in the previous section. Specifically, we remove the same taxa as in the subtrees that need to be pruned based on transitions with distances $> d$ on particular branches. For computational convenience, we restrict ourselves to ~~fewer~~ the values of d, specifically $d = 450$ km, 350 km and 250 km. ~~Additionally, we~~ For most sequences, only admin-2

level locations of sampling are known (administrative areas that correspond to prefectures in Guinea, districts in Sierra Leone and counties in Liberia). We therefore remove sequences without a known administrative region of sampling (e.g. when only the country of sampling is known), and we remove sequences such that monophyletic clusters of sequences sampled from the same administrative region are only represented by a single sequence. Such clusters would largely represent dispersal within administrative regions, which will be determined by the 'noise' ~~assigned to their location within an~~ because of their randomly drawn geographic coordinates within the administrative region (see below). The final data sets include 722 (d = 450 km), 676 (d = 350 km) and 527 (d = 250 km) sequences, respectively.

Inference under the multivariate diffusion models is problematic when different sequences are associated with identical trait values[51], or in this case, with the same geographic coordinates. When unique sampling coordinates are not available for every sequence, a common practice is to add a restricted amount of noise to duplicated traits. In BEAST[49], this is facilitated by a jitter option, that uniformly draws such noise from a user-defined window size for each dimension of the trait (i.e. each coordinate). Using a jitter may be problematic for two reasons in our case. Many administrative areas are along the coast, so the added noise may lead to sampling coordinates in the sea. Due to the different sizes and shapes of the administrative areas, the noise may also move coordinates to areas neighbouring their actual sampling area. To avoid these issues, we associate a random coordinate within the administrative area of sampling to each sequence. Because this approach ignores a lot of short-distance transmission within administrative areas, we show that our estimates are consistent with those based on an analysis of a data set that is composed of sequences with more precise sampling locations (the admin-3 chiefdom level) and that does not restrict monophyletic clusters of sequences sampled from the same administrative region to a single representative. The details of this additional analysis and the associated results are reported in Appendix S1.

We analyse each data set using a relaxed random walk (RRW) model with an underlying Cauchy distribution to represent among-branch heterogeneity in branch velocity[51], and with a multivariate Wishart distribution as a prior on the precision matrix[52]. We follow Dudas *et al.*[28] in choosing substitution, molecular clock and coalescent models, and their prior specifications, and reiterate those choices here. We model molecular evolution according to a HKY+Γ_4[53,54] substitution model independently across four partitions (codon positions 1, 2, 3 and non-coding intergenic regions), allowing for partition-specific relative rates. We use a non-parametric coalescent Bayesian skygrid model as a prior density over the tree[29,50] and model branch-specific evolutionary rates according to a relaxed molecular clock with an underlying log-normal distribution[55]. We specify a continuous-time Markov chain (CTMC) reference prior[56] for the overall evolutionary rate, while the rate multipliers for each partition were given an uninformative uniform prior over their bounds. All other priors in the phylogenetic inference are left at their default values. For each of the continuous phylogeographic inferences, we ran an MCMC chain using BEAST 1.8.4[49] for 130 (d = 250 km), 470 (d = 350 km) and 740 (d = 450 km) million iterations, removing the first 5% samples in each chain as burn-in. All the BEAST xml files are available at https://github.com/ebov/space-time. Based on these continuous phylogeographic inferences, we estimate, for each data set, the mean or weighted dispersal velocity v_{weighted} defined as follows:

$$v_{\text{weighted}} = \frac{\sum_{i=1}^{n} d_i}{\sum_{i=1}^{n} t_i}$$

where n is the number of branches in the phylogeny, d_i the geographic distance travelled and t_i the time elapsed on each phylogeny branch. We estimate this statistic for 1,000 samples of the posterior distribution of trees and report a posterior distribution of estimated values. In addition, we also ~~reconstruct~~summarize, for each data set, the evolution of the mean dispersal velocity and of the maximal wavefront distance through time. For the latter, we plot the maximal distance of the epidemic wavefront from the spatial origin as a function of time. We obtain all these statistics using the R package 'seraphim'[57,58].

Impact of borders on dispersal frequency. To test if administrative/international borders act as barriers to dispersal frequency, we adopt a Bayesian posterior predictive simulation procedure. Such a procedure allows ~~the calculation of~~ calculating a Bayesian counterpart of the classical p-value, using posterior predictive replications of the data and employing a test statistic that depends on both data and unknown (nuisance) parameters[59,60]. In our setup, we record the number of times borders are crossed by tree branches in the posterior set of trees, as determined by the location at the parent and child node of the branches, and compare this to posterior predictive values for the same statistic. We obtain the posterior predictive values by simulating a forward-in-time RRW process along each posterior tree using the sampled precision matrix parameters and location at the root node. We also condition on the branch-specific posterior rate scalars to generate the mixture of normals that characterises the RRW. In addition, we constrain the RRW simulations such that the simulated node locations ~~are in continental African and cannot~~ remain within continental Africa (and do not fall in the ocean. ~~As illustrated in Figure 3, posterior and posterior predictive diffusion histories are roughly similar except for the~~). Figure 2 illustrates the difference in position and orientation of branches ~~in the West African study area~~between a diffusion history reconstructed from the data and a diffusion history simulated using our posterior predictive simulation procedure. We integrate over all possible realisations, weighted by their posterior probabilities based on 1000 samples, to generate a test based on the border crossing frequency. In the absence of a border impact on Ebola movement in the West African epidemic, we expect a similar number of border crossing events between the posterior and posterior predictive diffusion histories, as the latter is deliberately unaware of borders. For each data set, we again use a sample of 1,000 trees from the post burn-in posterior distribution of phylogenies in order to accommodate phylogenetic uncertainty. Upon simulation of RRW diffusion along these sampled trees, we count and compare the number of crossing border events for each pair of inferred and simulated diffusion processes associated with a particular tree. Each 'inferred' N value (N_{inferred}) is thus compared to its corresponding 'simulated' value ($N_{\text{simulated}}$) to compute a predictive odds ratio (POR) as follows:

$$\text{POR} = \frac{p_e}{1-p_e} \Big/ \frac{0.5}{1-0.5}$$

where p_e is the posterior probability that $N_{\text{inferred}} < N_{\text{simulated}}$, i.e. the frequency at which $N_{\text{inferred}} < N_{\text{simulated}}$ in the sampled posterior distribution. In interpreting POR estimates, we adopt the Bayes factor scale defined by Kass & Raftery[61]: values higher than 3, 20 and 150 are considered as 'positive', 'strong' and 'very strong' evidence respectively for the impact of administrative/international borders on dispersal frequency. Both types of within-country administrative borders were obtained from the Global Administrative Areas database (GADM, www.gadm.org). The human population density and inaccessibility rasters were obtained from the Global Rural-Urban Mapping Project (GRUMP, sedac.ciesin.columbia.edu) and the Joint Research Centre of the European Commission (forobs.jrc.ec.europa.eu/products/gam) respectively. We implement the posterior predictive simulation procedure, along with a related tutorial, in the R package 'seraphim'[57,58].

~~REFERENCES~~

1. Baize, S. *et al.* Emergence of Zaire Ebola virus disease in Guinea. *N. Engl. J. Med.* **371**, 1418–1425 (2014).
2. World Health Organization. Ebola situation report - 10 june 2016 (2016).
3. Woolhouse, M. E. J., Rambaut, A. & Kellam, P. Lessons from Ebola: improving infectious disease surveillance to inform outbreak management. *Sci. Transl. Med.* **7**, 307rv5 (2015).
4. Bell, B. P. *et al.* Overview, control strategies, and lessons learned in the CDC response to the 2014-2016 Ebola epidemic. *MMWR Suppl.* **65**, 4–11 (2016).
5. Blackwood, J. C. & Childs, L. M. The role of interconnectivity in control of an Ebola epidemic. *Sci. Rep.* **6**, 29262 (2016).
6. Lau, M. S. Y. *et al.* Spatial and temporal dynamics of superspreading events in the 2014–2015 west africa ebola epidemic spatial and temporal dynamics of superspreading events in the 2014–2015 west africa ebola epidemic. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 2337–2342 (2017).
7. Fang, L.-Q. *et al.* Transmission dynamics of Ebola virus disease and intervention effectiveness in Sierra Leone. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4488–4493 (2016).
8. Kucharski, A. J. *et al.* Measuring the impact of Ebola control measures in Sierra Leone. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14366–71 (2015).
9. Pandey, A. *et al.* Strategies for containing Ebola in West Africa. *Science* **346**, 991–995 (2014).
10. Nouvellet, P. *et al.* The role of rapid diagnostics in managing Ebola epidemics. *Nature* **528**, S109–16 (2015).
11. Poletto, C. *et al.* Assessing the impact of travel restrictions on international spread of the 2014 West African Ebola epidemic. *Euro Surveill.* **19**, 20936 (2014).
12. Wesolowski, A. *et al.* Commentary: containing the Ebola outbreak - the potential and challenge of mobile network data. *PLoS Curr.* **6** (2014).
13. Lau, M. S. Y. *et al.* A mechanistic spatio-temporal framework for modelling individual-to-individual transmission—with an application to the 2014-2015 west africa ebola outbreak. *PLoS Comput. Biol.* **13**, e1005798 (2017).

14. Quick, J. *et al.* Real-time, portable genome sequencing for Ebola surveillance. *Nature* **530**, 228–232 (2016).

15. Arias, A. *et al.* Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies transmission chains linked to sporadic cases. *Virus Evol.* **2**, vew016 (2016).

16. Gardy, J., Loman, N. J. & Rambaut, A. Real-time digital pathogen surveillance — the time is now. *Genome Biol.* **16**, 155 (2015).

17. Modjarrad, K. *et al.* Developing global norms for sharing data and results during public health emergencies. *PLoS Med.* **13**, e1001935 (2016).

18. Nature Microbiology editorial. Don't wait to share data on Zika. *Nat. Microbiol.* **1**, 16028 (2016).

19. Holmes, E. C., Dudas, G., Rambaut, A. & Andersen, K. G. The evolution of Ebola virus: insights from the 2013-2016 epidemic. *Nature* **538**, 193–200 (2016).

20. Gire, S. K. *et al.* Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* **345**, 1369–72 (2014).

21. Volz, E. & Pond, S. Phylodynamic analysis of Ebola virus in the 2014 Sierra Leone epidemic. *PLoS Curr.* **6** (2014).

22. Stadler, T., Kühnert, D., Rasmussen, D. A. & du Plessis, L. Insights into the early epidemic spread of Ebola in Sierra Leone provided by viral sequence data. *PLoS Curr.* **6** (2014).

23. Carroll, M. W. *et al.* Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature* **524**, 97–101 (2015).

24. Kugelman, J. R. *et al.* Monitoring of Ebola virus Makona evolution through establishment of advanced genomic capability in Liberia. *Emerg. Infect. Dis.* **21**, 1135–1143 (2015).

25. Ladner, J. T. *et al.* Evolution and spread of Ebola virus in Liberia, 2014-2015. *Cell Host Microbe* **18**, 659–669 (2015).

26. Park, D. J. *et al.* Ebola virus epidemiology, transmission, and evolution during seven months in Sierra Leone. *Cell* **161**, 1516–1526 (2015).

27. Simon-Loriere, E. *et al.* Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. *Nature* **524**, 102–104 (2015).

28. Dudas, G. *et al.* Virus genomes reveal factors that spread and sustained the Ebola epidemic. *Nature* **544**, 309–315 (2017).

29. Gill, M. S., Lemey, P., Bennett, S. N., Biek, R. & Suchard, M. A. Understanding past population dynamics: Bayesian coalescent-based modeling with covariates. *Syst. Biol.* **65**, 1041–1056 (2016).

30. Faria, N. R., Suchard, M. A., Rambaut, A. & Lemey, P. Toward a quantitative understanding of viral phylogeography. *Curr. Opin. Virol.* **1**, 423–429 (2011).

31. Pybus, O. G., Tatem, A. J. & Lemey, P. Virus evolution and transmission in an ever more connected world. *Proc. Royal Soc. B* **282** (2015).

32. Zinszer, K., Morrison, K., Anema, A., Majumder, M. S. & Brownstein, J. S. The velocity of Ebola spread in parts of West Africa. *Lancet Infect. Dis.* **15**, 1005–1007 (2015).

33. WHO Ebola Response Team. Ebola virus disease in West Africa — The first 9 months of the epidemic and forward projections. *N. Engl. J. Med.* **371**, 1481–1495 (2014).

34. Hollingsworth, T. D., Ferguson, N. M. & Anderson, R. M. Will travel restrictions control the international spread of pandemic influenza? *Nat. Med.* **12**, 497 (2006).

35. Pybus, O. G. *et al.* Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15066–15071 (2012).

36. Trovo, N. S., Suchard, M. A., Baele, G., Gilbert, M. & Lemey, P. Bayesian inference reveals host-specific contributions to the epidemic expansion of Influenza A H5N1. *Mol. Biol. Evol.* **32**, 3264–3275 (2015).

37. Streicker, D. G. *et al.* Host-pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 10926–10931 (2016).

38. The Global Consortium for H5N8 and Related Influenza Viruses. Role for migratory wild birds in the global spread of avian influenza H5N8. *Science* **354**, 213–217 (2016).

39. Dellicour, S. *et al.* Using viral gene sequences to compare and explain the heterogeneous spatial dynamics of virus epidemics. *Mol. Biol. Evol.* **34**, 2563–2571 (2017).

40. Backer, J. A. & Wallinga, J. Spatiotemporal analysis of the 2014 Ebola epidemic in West Africa. *PLoS Comput. Biol.* **12**, e1005210 (2016).

41. Kramer, A. M. *et al.* Spatial spread of the West Africa Ebola epidemic. *R. Soc. Open Sci.* **3**, 160294 (2016).

42. Bielejec, F., Lemey, P., Baele, G., Rambaut, A. & Suchard, M. A. Inferring heterogeneous evolutionary processes through time: from sequence substitution to phylogeography. *Syst. Biol.* **63**, 493–504 (2014).

43. Ayres, D. L. *et al.* BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).

44. Ratmann, O. *et al.* Sources of HIV infection among men having sex with men and implications for prevention. *Sci. Transl. Med.* **8**, 320ra2–320ra2 (2016).

45. Lemey, P. *et al.* Unifying viral genetics and humant transportation data to predict the global transmission dynamics of human influenza H3N2. *PLoS Path.* **10**, e1003932 (2014).

46. Minin, V. M. & Suchard, M. A. Fast, accurate and simulation-free stochastic mapping. *Phil. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3985–3995 (2008).

47. Minin, V. M. & Suchard, M. A. Counting labeled transitions in continuous-time Markov models of evolution. *J. Math. Biol.* **56**, 391–412 (2008).

48. Swofford, D. L. PAUP*: Phylogenetic Analysis Using Parsimony (and other methods), version 4 (2002).

49. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–1973 (2012).

50. Gill, M. *et al.* Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. *Mol. Biol. Evol.* **30**, 713–724 (2013).

51. Lemey, P., Rambaut, A., Welch, J. J. & Suchard, M. A. Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–1885 (2010).

52. Wishart, J. The generalised product moment distribution in samples from a normal multivariate population. *Biometrika* **20A**, 32–52 (1928).

53. Hasegawa, M., Kishino, H. & Yano, T.-a. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).

54. Yang, Z. Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: Approximate methods. *J. Mol. Evol.* **39**, 306–314 (1994).

55. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).

56. Ferreira, M. A. R. & Suchard, M. A. Bayesian analysis of elapsed times in continuous-time markov chains. *Can. J. Stat.* **36**, 355–368 (2008).

57. Dellicour, S., Rose, R. & Pybus, O. G. Explaining the geographic spread of emerging epidemics: A framework for comparing viral phylogenies and environmental landscape data. *BMC Bioinfo.* **17**, 1–12 (2016).

58. Dellicour, S., Rose, R., Faria, N. R., Lemey, P. & Pybus, O. G. SERAPHIM: studying environmental rasters and phylogenetically informed movements. *Bioinformatics* **32**, 3204–3206 (2016).

59. Meng, X. L. Posterior predictive p-values. *Ann. Statist.* **22**, 1142–1160 (1994).

60. Gelman, A., Meng, X. & Stern, H. Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Stat. Sin.* **6**, 733–807 (1996).

61. Kass, R. E. & Raftery, A. E. Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).