

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