1. A two-sentence editorial summary of the paper will appear on the journal homepage with the link to the paper. This is our proposed summry: ‘A re-analysis of virus diversity in mammals that now takes into account host sharing estimates that global diversity estimates have been previously overestimated by two orders of magnitude”. Please let us know of any factual inaccuracies.

We would propose a small change, to reduce re-use of the word “estimate”:

A re-analysis of viral diversity in mammals, which corrects for host sharing, suggests that global viral diversity has been previously overestimated by two orders of magnitude.

2. Please amend title to specify that it is about \*mammal\* viral diversity. We suggest "Global estimates of mammalian viral diversity accounting for host sharing". Titles should be fewer than 100 characters (including spaces) and contain no punctuation.

We have changed the title as such.

3. Please could you revise the colour scheme in Fig 2D so that it is readable by people who are red-green colourblind. See here for recommendations on Figure presentation: <https://www.nature.com/natecolevol/info/final-submission> . (for example, you could change Red to Magenta)

We have updated the figure as such.

4. Please read our policy about competing interests (<https://www.nature.com/authors/policies/competing.html>) and include a Competing Interests Statement in the article.

We have included a competing interests statement.

5. Please submit the main text, methods and figure legends as a single file (Word or TEX, not PDF), but the figures themselves as separate files.

We have made this change.

6. Please ensure that all figure axes are labelled as precisely as possible, as figures should be fully understandable without extensive reference to the main text. For example, please specify where axes refer to host or viral richness (to distinguish from abundance)

We have made these changes in Figures 1 and 2.

7. Please add a reference to the Olival dataset in the legend for Fig 2 (ie. “Node size is proportional to number of viruses sampled in the Olival dataset [14]”)

We have made this change.

8. Please include a citation in the main text to each Supplementary file (this guides readers to specific files of interest), and re-label each of these files “Supplementary Figure 1, Supplementary Figure 2 etc…” (not Figure S1, Figure S2 etc…)

We have made both of these changes.

9. Please specify in the Data Availability section where each data set can be found (e.g. “All data in this study is from previous studies and is available online for researchers to reproduce our results: global diversity of viruses in mammal hosts can be found in Carroll et al. [13]; plant-pollinator interactions can be found in Robertson [32, 33]….etc”

We have made this change.

10. Please include a separate Code Availability statement. We recommend including a citation to the R package in your reference list and including this in the statement e.g “The codependent R package [43] is available at [github.com/cjcarlson/codependent](http://github.com/cjcarlson/codependent)”

We have made this change.

11. Please follow the instructions below to link ORCID accounts of all corresponding authors to accounts in our system.

I have linked my ORCID account.

12. Please complete the Editorial policy checklist and the new version of the Reporting Summary (links below) and upload them with your revised manuscript. We will publish the latter along with the paper. Please note that these forms are dynamic ‘smart pdfs’ and must therefore be downloaded and completed in Adobe Reader. Please also ensure that “Final Submission” box is checked.  
a. Editorial policy checklist: <https://www.nature.com/authors/policies/Policy.pdf>  
b. Reporting summary: <https://www.nature.com/authors/policies/ReportingSummary.pdf>

I need to do this.

13. Please include full titles in each citation of the reference list.

We have made this change.

14. Please note that if you are not the copyright holder of the silhouettes used in the figures, you will need to obtain the permission from the copyright holders to reproduce them.

We have remade these figures entirely with public domain images from phylopic.org.

Reviewer #1 (Remarks to the Author):  
  
My major issue with this excellent paper related to sampling effort. The authors have responded satisfactorily to my concern, and given that there are no good enough representations of maximal sampling effort, what they do is the best that can be done. I have no other comments on the revised and extended version of the manuscript.

We appreciate this feedback.

\*\*Comments on response to Referee #2\*\*  
  
In my opinion, the responses are satisfactory overall. Also, the paper benefits from its current length, although under my opinion, it was clear enough for a brief communication. The reviewer raises an important concern: why a power-law fit and not other fits reported in other ecological networks. The authors provide a good argument, but it is only included as a supplement. I would recommend that they incorporate a paragraph in the main text explaining the rationale for using a power-law. It is key that they explain that the power law tends to overpredict at higher values, so this function should be interpreted as an upper bound of possible diversity. Other functions will predict even less diversity. And that's the whole point of this article: the common overstimation of viral diversity assuming extreme host-specificity (1:1 linear relationships) vs. host sharing, with a power-law fit as the most likely and conservative scenario.

We expanded in the results explaining how we arrived at the power law, and what we found in the supplement. The short paragraph about this, which previously just stated that we found power laws, is now substantially expanded into multiple paragraphs.  
  
Reviewer #3 (Remarks to the Author):  
  
The authors have done a nice job of responding to reviewer comments and tightening up the loose ends in this study. Its impact comes through even more clearly in the revised version, and they have framed the study in a broader and less adversarial way, which is a good change.  
  
I have a few suggestions for final tweaks to make the findings clear to readers:  
  
- I would advise a little bit more interpretive text setting up Table 2. From the current main text alone, it is easy to miss the rationale for including the estimates from the 'lower 50% of the subsampled curve' (as it is labeled in the table caption). I would add a strong sentence explaining that this is an approach to generating a working 'upper bound' estimate, which is a priority given the central message that the new estimates are much lower than those published by Carroll et al. I would also add a phrase to the table caption to make this clear, for those who scan papers by skimming display elements.

In the expanded two paragraphs at the start of the Results, we now do this. We include the following explanation of the upper bound message: “One important consequence of this error structure was that power laws fit to a smaller portion of the network produced higher estimates (see Figure 1E), with z values closer to 1. We took advantage of this property to create what we called ``upper bound'' estimates, using 50% of the network to generate estimates that reflect an upper bound on the overall size of the network.”

The table caption has also been updated: “Estimation of viral diversity using the lower 50% of the subsampled curve, an approach that produces an upper bound on extrapolated estimates due to the overestimation of the power law method at high and low values.”  
  
- Related, the 'only' on line 116 could use a bit more context.

We have added a clause here and separated it out as its own sentence: “Even using the 50% upper bound method, which produces a substantially increased estimate, we only calculated 55,784 possible total viruses, still less than 10% of previous estimates (Table 2).”

- The term 'host plasticity' arises just a few times, at high-leverage spots (e.g. the abstract,and line 111). I'm not sure its meaning will be clear to readers who aren't steeped in this literature (or even to some who are, e.g. I'm not sure if it's meant to differ from 'host breadth'), so a brief inline definition would be helpful.

We have changed to host breadth, which is a more commonly used term, in both locations. (There was no intended difference in meaning – we were merely using a term borrowed from Johnson *et al.* 2015) We also define it at line 131 as “(number of host species that can be infected)”.  
  
- The text references to Figure 2 are very cursory. Perhaps specify Fig 2A-C in the earlier references that are just about the rarefaction curves, so readers aren't trying to guess the connection between the text and network diagram.

We have added two references to these curves in the appropriate places (first when referencing the power law approach applied to viruses, and second when comparing the DNA and RNA curve exponents).

- line 295-296 - double 'ideally'

We appreciate this catch and have removed one.