Dear Dr. Poisot,

We much appreciate your editorial review and the substantial and rigorous comments by the two reviewers. We feel that the reviews significantly improved our writing. We have addressed all comments to the best of our abilities as outlined in detail below. Please let us know if you still see room for improvement. We thank you and the reviewers for excellent and constructive feedback.

Sincerely,

Nik Grunwald

*Response to editor:*

*Both reviewers are extremely positive about this contribution, and I do share this feeling. Please make all suggested comments. I would appreciate that you simplify the figures as much as possible (use figshare if you want to refer to the complete versions). In particular, the color coding in Figure 3 is not adequate (not colorblind friendly). Please do make sure that sensible color codings are used.  
  
Response to reviewer's:  
Reviewer #1: Foster et al developed an R package for visualizing hierarchical data such as community taxonomic diversity data. Interestingly, they identify potential extensions of their package to other datasets, including gene ontology annotations, or even presidential election information! The plotting capabilities of the package are impressive and should prove useful for many scientists, justifying publication. Conversely, it should be noted that this manuscript describes a software package as opposed to conducting original research. I do not view this as a shortcoming of the manuscript, but it is something that should be considered by the editor.  
I found the manuscript to be clear, well written, and it included several examples that will simplify the process for others to implement the software. I have only optional minor comments below that should be considered:*

*1. Some comparison, perhaps even a figure, highlighting the differences between the capabilities of metacoder to those of GraPhlAn would be a nice addition to the manuscript.*

The following has been added to the heat tree plotting section:

“There is also a set of python scripts called GraPhlAn that can make similar kinds of trees. GraPhlAn has better annotation abilities than metacoder, supports edge length for phylogenetic trees, and can plot a variety of node shapes. However, metacoder’s heat tree fucntion can plot multiple trees per graph, use different layout algorithms, automatically transform raw data to color/size for quantitative display with a scale bar, and optimize the size range of nodes to avoid crowded or sparse graphs. Since metacoder is an R package, it is also easier to integrate into a growing collection of other bioinformatic R packages and use with R Markdown for digital research notebooks.“

*2. The resolution of figure 1 is not adequate.*

This has been fixed. *3. The color coding of Figure 2 is not clear to me. Are the colors a heat-scheme? If yes, hot and cold are not clear. Figure 1 of the cover-letter and most of the manuscript figures include a color key. I think Figure 2 of the manuscript needs one as well.*

Yes, I agree. I have added color keys and made the text more readable. *4. The authors repeatedly discuss digital PCR, but they don’t define it until page 9 (line 200). As a reader who isn’t familiar with digital PCR, I suggest that this brief description be moved earlier.*

Yes, good point. I have moved this to the first mention of digital PCR (besides the abstract) on line 48. *5. Refs for phyloseq, ape, seqinr, and taxize are missing from the manuscript (pg 12, line 281)*

This has been fixed, except for phyloseq, which was cited in the “heat tree plotting of taxonomic data” section. *Reviewer #2: The authors present their software package 'metacoder' for the manipulation and visualisation of hierarchical data (typically that resulting from metabarcoding). In doing so, they contribute valuable software to an emerging field of community ecologists. In particular, I agree with the authors that their visualisation approach far exceeds the current norms. The virtual PCR functionality they include is likewise a significant and useful contribution. Moreover, the software and its documentation is comprehensive. For the most part the manuscript, its figures, and the software itself are tidy and functional. In particular, I thought the authors framed the utility and need for this software nicely as well as highlighting the most important aspects of the package with analyses and figures. Having said that, the figures could use some work to increase their clarity given that a major part of this software is the visualisation. In particular, the authors should consider that many of the labels on their figures are illegible when the figures are the size of a full page. Therefore, labels will be distracting at best when the figures are shrunk to publication size. I also wondered throughout the MS whether the authors could explore (or, have already) applications of their software to other metabarcoding data sets. In particular, there are several groups currently using metabarcoding to examine diet and/or food webs. The software is obviously directly applicable to this data and the ms may benefit from this added perspective. I have a handful of comments on the text and software itself but all in all this manuscript is a good description of an important software contribution.*

I have gone through all the figures and increased the label size in most cases. Good idea in regards to examining food web / diet. That would indeed be very interesting, but we are already using 7 different data sets and we are concerned that adding more might overwhelm some readers.   
 *Specific comments:  
- ms is missing the Author Summary section required by PLoS CB*

I did not think we needed one since this is a software submission and the Author Summary is not listed in the document layout requirements. The “Software submissions” section in the submission quidlines (<http://journals.plos.org/ploscompbiol/s/submission-guidelines>) has its own list of sections which I assume replace the list for a normal submission, but we are happy to add one if I am wrong about this.

*- Figure 1: increase the resolution*

This has been done.

*- Figure 1: the middle section (feedback between manipulative functions and data is unintuitive at first glance. Can these be broken up to maintain the flow from box to box?*

Yes, i see what you mean. I rearranged the boxes into one vertical flow on the left and put the outputs of the analysis on the right.

*- Figure 1: the heat tree at the bottom is indistinguishable. Either, increase size and alter colour palette, or remove.*

I made it much bigger so it can be seen now.

*- L64: users should be user's*

Done.

*- L68--69: Sentence is confusing. I would suggest changing the middle part to "...a list of included functions...that create..."*

Done.

*- L83: "The structure and nomenclature of the taxonomy used" could be "The taxonomic structure and nomenclature used"*

Done.

*- L93: "figure 1" should be "Figure 1"*

*Done.*

*- L142: "and therefore" should be "and are therefore"*

Done.

*- Figure 2: Certainly needs a legend. Also I think it makes the reader look for things that are not there when node and edge size as well as colour are showing the same thing*

Yes, I agree after thinking about it more. I added a legend. In regards to the color and size showing the same statistic: without a legend it is indeed confusing, but adding a legend with only one side labeled makes it clear that it is representing one statistic. This is the default behavior of the package and although there could be some confusion upon first seeing it, I think that the benefit that the combination of effects has on quickly perceiving differences outweighs the risk of initial confusion. The color can often be used to differentiate small differences that size does not pick up, and setting size relative to abundance makes for better looking trees, since trees with nodes of all the same size have nodes that are either too small or overlap too much.

*- Figure 3: Too busy. Please use a subset of the data for part (a) so that the reader can discriminate what you are showing (i.e., that there are differences in node and edge size).*

Ok, I reduced the number of taxa shown by increasing the minimal number of reads needed for a taxon to be shown and lowered the maximum number of labels shown. I also changed the color scale in response to a comment from the editor. The label size has also been increased.

*- Figure 4: the figure itself is unclear to me. Does "All data" display all sequences in the database or all amplified sequences? Perhaps this figure would be better off with two or even one comparison and more in depth discussion of what the reader is seeing.*

“All data” refers to all sequences in the database. I hesitate to just show a single database since one of the features of the package I wanted to show with this is that the same analysis can be easily carried out on data from diverse sources. However, I did change “All data” to “Whole database” to try to avoid similar confusion from others.

*- Figure 5: the caption or figure itself needs to do a better job of explaining the colouring. Rather than giving examples of taxa that are enriched in certain areas, a better approach might be to say that green shows taxa enriched in the part of the body shown in the row and brown shows taxa enriched in the part of the body shown in the column.*

Good idea. This has been done.

*- Figure 5: the first sentence of the caption needs to tell the reader what the figure shows. Currently, this is the 3rd or 4th sentence. I think this will clear up a lot of the confusion around this figure (i.e., this figure shows pairwise comparisons of microbiome communities in different parts of the body)*

Done.

*- Figure 6: this is a really nice example of the greater applicability of the software.*

Thanks!

*- Figure 7: I would just reiterate here my earlier point about labeling.*

I have increased the label size and removed some of the labels.

*- CRAN documentation: example code for metacoder::extract\_taxonomy gives the error: "Error in validate\_regex\_key\_pair(regex, key, multiple\_allowed = c("taxon\_info", :Invalid key value "class\_name" given."*

Thanks for catching this! This was caused by upgrading `extract\_taxonomy` without updating the example code. It has been updated and now works and you can see the changes here:

<https://github.com/grunwaldlab/metacoder/commit/5d62afebcdbc5ac39c151dcffc97c8f2db53d423>.

You can install the development version to try it out using the following command:

devtools::install\_github("grunwaldlab/metacoder@dev")

*- CRAN documentation: typographical errors in the heading of the metacoder::remove\_redundant\_names function and in the details of metacoder::gen\_ex\_data*

Thank you for going through the documentation! I have fixed those errors. You can see the changes here:

https://github.com/grunwaldlab/metacoder/commit/216583ce5c6ae25e7a77b25a908c6ad6a68d898c