Biogeosciences
Discussions

Open Access

EGU

Interactive comment on "Phylogeny of the Stipa and implications for grassland evolution in China: based on biogeographic evidence" *by* Qing Zhang et al.

Anonymous Referee #1

Received and published: 24 May 2018

This manuscript focuses on an interesting topic - the biogeography and diversification of an ecologically important grass lineage; however, the data presented are insufficient to justify the conclusions that are drawn, and several additional analyses should be done before this manuscript is published in Biogeosciences. I am not convinced that the basic premise of this paper makes sense given that Stipa is a clade with about 350 species that occur on six continents (GrassBase) and only a small number of species from a limited geographic area are represented here. The biogeographic analysis that is presented relies on the assumption that the Chinese Stipa species form a clade and assumes that these lineages have not arrived in China from other regions e.g. Southeast Asia or Europe. For instance it seems possible that some of the Chinese Stipa

might actually have sister species or other close relatives outside of China. There are sequences for more than 75 different species of Stipa available on GenBank and as many of these as possible should be included. With a more complete (or global) phylogeny of Stipa, inferences about dispersal and vicariance events would be much more reliable. I also do not think the data in this manuscript are sufficient to test anything about the diversification rates of Stipa. Again a broader phylogeny that includes Stipa from outside of China would be necessary for such an analysis. Aside from sampling issues, I think that the results of this paper are stated too strongly. There is uncertainty in the phylogentic reconstruction, the inferred divergence dates, and in the biogeographic reconstruction, and very little of this uncertainty is presented or discussed in the current manuscript. Finally, there are grammatical mistakes throughout the manuscript and overall this paper needs to be thoroughly edited for clarity. One of the most common errors is "grasslands" should be plural rather than singular.

Introduction: There should be a section in the introduction with background information about Stipa. How many species are there? What is its distribution outside of China? Maybe also something about the ecology of the clade could be added.

Line 49. See also Liu et al. 2018 "Phylogeny of Orinus (Poaceae), a dominant grass genus on the Qinghai-Tibet Plateau"

Lines 51-57. "Late Tertiary" seems like a misleading description of when Chinese grasslands emerged. "Mid tertiary" would be more appropriate for 33MaBP. Also see Edwards et al. 2010.

Line 73. What are "constructive species?"

Methods: Line 99. Chloroplast not chlorophyll

105. Why is Ramirez et al. 2017 cited related to diluting DNA samples?

Line 134 where is this rate from? Is it possible to get a rate of chloroplast evolution that is specific to grasses?

Line 138-139. The sentence structure makes it sound as though FigTree was used to infer divergence times. Instead "Finally the tree and divergence times were visualized in FigTree."

Lines 145-147. I'm confused whether 100 or 2500 posterior trees were used for this analysis.

Line 148. Setting the maximum number of areas a species can occupy to 2 seems justifiable but needs more context. How many areas do the most widespread Stipa species today occupy?

More details are needed on the biogeographic analysis. How were results summarized across the posterior trees?

What analysis was used to generate Figure 4?

Results: Lines 153-155. This isn't hugely important, but these numbers don't really add up in a way that makes sense. With 2988 conserved and 1102 variable sites there would only be 4090 sites total - what happened to the other three? I'm guessing it has to do with idels or gaps in the alignment? I'm not sure this information is necessary to report, but if you're going to include it, it would make sense to include something about indels as well.

Did the BEAST analysis converge? Were all ESS values >200?

Was the tree well supported? The posterior probability for each node should be added to Figure 2.

Line 160. What are the confidence intervals for these ages?

Line 162. The tree alone does not provide adequate evidence for an "explosive radiation" of Stipa. This could be tested using BAMM (Rabosky et al. 2013) or other diversification methods.

Lines 165-179. This scenario could be plausible if these species form a clade, but there

C3

are other possibilities. Moderate support for other scenarios seem to be indicated by multiple colors in the backbone nodes in Figure 3.

Line 182. What are "larger expansion and isolation events"? How are these different from the expansion and isolation events described in the previous section?

Discussion: How do the age estimates from this study compare to estimates from Poaceae-wide analyses?

Line 200. See Xing and Ree 2017 for a nuanced perspective on this. "Uplift-driven diversification in the Hengduan Mountains, a temperate biodiversity hotspot"

Line 213-4. These lines are confusing because the entire tree is dichotomous. For the second part maybe it should read "the inner branches are short compared to the tip branches?"

Line 216-225. The discussion of rapid radiation in Stipa should either be qualified or removed since it is not explicitly tested in this paper.

Lines 296-301. Because the Chinese grasslands pre-dated Stipa (last section of Discussion), it doesn't seem possible to infer the order of Chinese grassland formation from biogeographic patterns of Stipa.

Figure 3. What are the colors in 3B? I'm confused by what the "largest possible distribution area of Stipa species" means? Largest compared to what? I think it would be clearer to color the map in part A to match the phylogeny in part B. Then the letters and numbers might not be necessary at all? I'm also confused by the fact that part A and the letters on part B seem to tell a very clear story, but the colors of the nodes suggest that there is a lot of uncertainty in the deeper nodes of the tree. That amount of uncertainty would be expected for this kind of an analysis and it seems like it should be reflected more clearly here and in the text of the paper.

Figure 4. How was this graph generated? The colors on this graph aren't very clearâĂŤred/pink/purple all look similar to me. Maybe the lines could be thicker or a

C4

green/light blue color could be used instead of one of the other colors.

C5