

1 Reviewer1:

2 This study addressed an interesting question of the effects of multi-level processes on
3 the geographic distribution of forest woody plant sexual diversity, which helps to
4 advance the understandings of the underlying drivers of geographical pattern of forest
5 plant diversity. The manuscript is overall clearly written. However, at the moment I
6 still have a couple of concerns and some minor comments, listed as below,

7 Major concerns:

8
9 **Comment 1:** According to my understanding, "region" in the text means continent-
10 level landmass (Lines 123-124) and its effects on shaping plot-level plant sexual
11 diversities were important (Lines 161-162), while latitude effects were unimportant
12 when other factors' effects were controlled (Lines 155-157). It would be better not to
13 group the regions as either tropical regions or temperate regions, which is confusing.

14 **Response 1:** Thank you for this comment. In the analysis, we grouped regions based
15 on continent-level landmass. In the abstract, we referred to tropical and temperate
16 regions in the context of previous debates (Lines 52-55). This might have led the
17 referee to believe we grouped the regions as either tropical regions or temperate
18 regions. In the revision, we rephrased the results in the abstract regarding regions as
19 follows:

20 *“Our results showed that plants were more likely to be dioecious than hermaphroditic*
21 *in Oceania and Tropical Asia, but were more likely to be monoecious than dioecious*
22 *in Europe and North America compared with Tropical Africa.”*

23 We have updated this result in the Abstract (Line 14-16).

24 **Comment 2:** As the authors stated in the Introduction, if (a) dioecious is more likely
25 associated with dry or poor resource habitats (Lines 61-56), and (b) areas with
26 younger species age is expected to have a higher incidence of dioecy (Lines 55-57),
27 then an interactive effect likely exists between (a) and (b) on uplifting the incidence of
28 dioecy. Have the authors considered of including this interactive term in the
29 regression models?

30 **Response 2:** Thanks for the suggestion. In the revision, we have added the interactive
31 term between mean species age and mean annual precipitation in the models (Table
32 1), which showed marginally significant effects on the proportion of plant sexual
33 systems. We added the results in the revisions (Line 177-179).

34 Minor comments:

35
36 **Comment 3:** Lines 14-16 I am not sure this finding (not significant when the
37 effects of climate etc. were controlled as shown in the Results at Lines 155-157)

38 supports Baker's law. Lack of dioecy in temperate regions can be a result of biotic
39 and/or abiotic filters, apart from the expected low colonization capacity of dioecious
40 plants.

41 **Response 3:** Agreed. We have rephrased the text to highlight the role of region
42 factors rather than Baker's Law. For example, in the abstract, we concluded that "Our
43 results showed that plants were more likely to be dioecious than hermaphroditic in
44 Oceania and Tropical Asia, but were more likely to be monoecious than dioecious in
45 Europe and North America compared with Tropical Africa." (Lines 14-16).

46 **Comment 4:** Lines 40-41 Mathematical inference? What is the underlying
47 biological/ecological meaning?

48 **Response 4:** Thanks a lot for this valuable comment. According to this suggestion, in
49 the revision, we have rephrased the sentence as:

50 *"The proportion of hermaphroditism might be higher in tropical forests than in
51 temperate forests as precipitation and the proportion of biotic pollination decrease
52 with latitude."*

53 **Comment 5:** Line 43 replace "showed" with "show"?

54 **Response 5:** Corrected.

55 **Comment 6:** Lines 53-55 I am not sure this expectation holds, if the effects of
56 abiotic and biotic filters have not been ruled out first. Besides, I think the authors
57 meant that dioecious plant species would have lower relative incidence in temperate
58 than tropical regions compared with other sexual systems plant species, due to the
59 expected low colonization rates of dioecious plants in temperate regions.

60 **Response 6:** We intended to mean dioecious species would have lower relative
61 incidence in temperate than other sexual systems supposing other conditions are the
62 same (Lines 53-57). We revised the sentence as:

63 *"Supposing abiotic and biotic conditions are equal, dioecious plant species would
64 have lower relative incidence in temperate regions, e.g., Europe and North America
65 than in tropical regions, e.g., Tropical Asia and Oceania, compared with other sexual
66 systems, due to the expected low colonization rates of dioecious plants in temperate
67 regions."*

68 **Comment 7:** Line 81 what does "PET" stand for?

69 **Response 7:** PET stands for annual potential evapotranspiration. We removed PET
70 here as it was not retained in the final model.

71 **Comment 8:** Lines 85 should be "forest dynamics plots"?

72 **Response 8:** Yes.

73 **Comment 9:** Lines 112, 132 replace "number of trees" with "number of stems"?

74 **Response 9:** We used number of trees.

75 **Comment 10:** Lines 127-134 A common format is that R package name be in italics
76 and R function name in quotation marks, please check and be consistent in the text.

77 **Response 10:** Addressed.

78 **Comment 11:** Line 180 add a "be" between "to" and "hermaphrodite". replace
79 "Fig. 5, S4" with "Fig. 5, S5"?

80 **Response 11:** Corrected.

81 **Comment 12:** Line 181 "Fig. S5" should be replaced with "Fig. S4"?

82 **Response 12:** We updated the figure in the appendix and revised it in the main text
83 accordingly.

84 **Comment 13:** Lines 196-199 These two findings contradict each other. Why low
85 colonization capacity is important in one case but not important in another?
86 Please elaborate more on this.

87 **Response 13:** In the revision, we found island effects were also important after
88 including the interactive term (between species age and precipitation) and plot
89 characteristics (number of species and number of trees). Plants were more likely to be
90 monoecious than dioecious in island communities (Table 1), which did not contradict
91 the effects of region (Lines 203-204).

92 **Comment 14:** Lines 221-222 How do these results support Baker's law?

93 **Response 14:** We appreciate your question. It was a mistake. We revised the sentence
94 as "These results suggested evolutionary processes diminish the effect of Baker's
95 law" (Lines 224-225).

96 **Comment 15:** Lines 224-226 It is not clear what regional processes or climate
97 factors, can the authors give an example?

98 **Response 15:** The effects of age could depend on precipitation as you suggested
99 (Lines 226-228). We revised the sentence as follows:

100 *"The effect of species age on plant sexual systems was different at regional and global*
101 *scales (Wang et al., 2020), which could be explained by the interactive effects*
102 *between species age and precipitation across different regions."*

103 **Comment 16:** Line 229 replace "less" with "more".

104 **Response 16:** Corrected.

105 Reviewer2:

106 I thank the authors for their interesting study. I have enjoyed reading the manuscript,
107 which is mostly presented clearly and concisely, and I find the emergence of
108 biogeographic patterns in plant reproductive traits quite exciting, personally. I have
109 nevertheless a few general points that I think need the author's attention:

110 **Comment 1:** I am a bit unsure about the chosen modelling strategy. At the very least,
111 some more details should be given (in the appendix), like the correlation between the
112 different variables, and the results of the VIF analyses. I also have some doubts that
113 fitting “regions” is the most meaningful variable to fit. I wonder if it wouldn't make
114 more sense, at least as an alternative approach, to model this variable as a random
115 factor in a mixed model.

116 **Response 1:** Thank you for the suggestions. We added tables for the pairwise
117 correlation among climate variables and VIF analyses in the appendix. We agree that
118 region could be treated as a random effect if we were not interested in the effect of
119 region. However, region is a major factor we are interested with as in many similar
120 global biogeographic studies. We need to include that variable. It was retained by
121 model selection procedure as the most important variable.

122 **Comment 2:** Another aspect that I think is somewhat problematic, is the apparent
123 overfitting of the models. As can be seen in all the graphs of the logistic models, the
124 blue line fits almost individual data points, which is also expressed in the relatively
125 high variances explained in the different models. I think that this “better” fit clearly
126 comes at the cost of interpretability. I suspect that also here the region variable plays a
127 major role, which again makes me wonder if it wouldn't be better to fit it as a random
128 factor. Another point in this context is that you present certain variables as not
129 significant in the models, but then still visualize their relationships with sexual system
130 in the presented graphs (e.g. latitude etc.).

131 **Response 2:** The region variable is of major interest in global biogeographic study
132 and it has been constantly shown to have significant effects on species richness and
133 functional traits (Ricklefs and He 2016, Zhang et al. 2016). That is also the case for
134 our study. Our model selection processes also retained region as the most important
135 variable in affecting the global distribution of plant sexual systems (Table 1, S2). We
136 moved the graph with latitude to the appendix because it has been used in previous
137 studies (Baker and Cox 1984). We further excluded other graphs with variables as not
138 significant in the models.

139 **Comment 3:** In some places in the discussion I have problems following what you
140 are trying to say, and you can find those in my detailed comments below.

141 **Response 3:** We revised the discussion to make it clearer. Please see our explanations
142 below.

143 Detailed comments:

144 **Comment 4:** Line 15: Throughout the manuscript, you are constantly confusing the
145 use of adjectives and nouns, when referring to the sexual systems. Please make sure
146 that you use dioecious/dioecy, monoecious/monoecy,
147 hermaphroditic/hermaphroditism appropriately.

148 **Response 4:** We have checked and corrected the usage throughout the manuscript.

149 **Comment 5:** Line 112, “plot characters”: These seem to me two important potentially
150 confounding variables, and I don't understand why they have not been kept in the
151 models (no matter their VIF). In addition, I would also think species richness in the
152 plots could be another important covariate.

153 **Response 5:** Thanks for this suggestion. In the revision, we kept the plot characters
154 (number of species, number of trees) in the final models (Table 1).

155 **Comment 6:** Line 128f: I don't understand how you can estimate lambda for a
156 discrete trait, as lambda is based on a Brownian motion model of trait evolution,
157 which is models continuous traits. Could you please comment and specify how that
158 works for discrete traits? Could that also be the reason that you get either 0 or very
159 high values for lambda?

160 **Response 6:** We used the “fitDiscrete” function in the R package *geiger* for discrete
161 traits, which could fit various likelihood models for discrete character evolution
162 (<https://cran.r-project.org/web/packages/geiger/index.html>). As the effects of lambda
163 were not significant, we did not include it as a variable in the model.

164 **Comment 7:** Line 136ff: This selection procedure based on variance inflation, while
165 maybe statistically valid, is somewhat elusive. First, it would be good to see the
166 results of this in the appendix, also to get a better overview of what variables were
167 used and which were excluded, and also how they were all correlated with each
168 other. I would also argue that it could be important to keep some potentially
169 confounding variables (i.e. the plot characteristics plus plot species richness) in the
170 model, no matter their VIF. Otherwise, what is the point of having those variables in
171 the first place?

172 **Response 7:** Thanks for this suggestion. We added the table of VIF in the appendix
173 and included the two plot characteristics in the model (Table 1, S2).

174 **Comment 8:** Line 156f: I would be more careful in how you express that here.
175 Latitude is clearly reflected in the "region" variable, so it is not that it is unimportant,

176 it is just that in the model all the explanatory power is "taken away" by the region
177 variable, but that region variable still expresses latitude as well. See also my earlier
178 general comment about the use of region as a variable.

179 **Response 8:** We agree that the region variable, to a good degree, expresses latitude,
180 but is more than the latitude. For example, temperate Asia, Europe, and North
181 America were coded as different regions though they have similar latitudes.
182 Neotropics and tropical Asia were also coded as two different regions. Anyway, we
183 rewrote the sentences describing the effects of latitude but kept the region variable
184 because it contains more information than latitude and it represents the major
185 hypothesis we aimed to test. We revised the sentences as follows:

186 *“However, when the effects of region, plot characteristics, and climate factors were*
187 *considered, the latitude variable was not retained in the final model, as it was closely*
188 *correlated with the region and environmental factors (Table 1, S2).”*

189 **Comment 9:** Line 162, “plot characteristics”: I can't see the plot characteristics in
190 these tables. In general, since you never show what the “full” model was that you
191 tested, it is not possible to understand from these table what variables were included
192 and which excluded, as you seem to present only the significant variables. As
193 mentioned earlier, it would be good if you could provide more details on that (see my
194 comments about VIF and correlation between variables).

195 **Response 9:** We added the tables in the appendix (Table S1, S2) and included the plot
196 characteristics in the model (Table 1).

197 **Comment 10:** Line 174: This table does not show anything about lambda. Again, it
198 would be helpful to see the correlations between variables presented somewhere.
199 Also, if it wasn't significant in the models, why did you decide to include plots in the
200 appendix? Also, I am sceptical about the calculation of lambda for these traits. First
201 because of the aforementioned doubt that lambda is meaningful for discrete variables.
202 Second, because as can be clearly seen in the graphs you provide in the appendix, the
203 lambda is either very high, or zero, which does not seem very meaningful, certainly
204 not for including lambda as an explanatory variable.

205 **Response 10:** To clarify, we did not include lambda as an explanatory variable and
206 removed the plots in the revision. See our above responses for other questions.

207 **Comment 11:** Line 194-195: I don't understand how that relates to the previous
208 sentence. You say that dioecy is disadvantageous for long distance colonisation, but
209 what has that to do with tropical regions? Why should they disfavour species with
210 long distance colonisation ability? Also, the end of the sentence doesn't make sense.

211 **Response 11:** We revised the sentence as: As temperate plant communities have been
212 assembled from tropical flora (Qian and Ricklefs 2016), tropical and the nearby

213 regions where dioecy originated are expected to have higher incidences of dioecy than
214 temperate regions (Bawa 1980; Renner and Ricklefs, 1995; Sakai and Weller, 1999;
215 Renner, 2014).

216 **Comment 12:** Line 197, “which supports...”: Again, I don't follow the logic here.
217 Please try to express yourself clearer.

218 **Response 12:** We revised the sentence as: Our study found a high proportion of
219 dioecy in Oceania and tropical Asia (Table 1, S3), which supports the deduction of
220 Baker's law on the dispersal limitation effect on dioecy.

221 **Comment 13:** Line 246: Unclear what processes you mean here, and it results in this
222 phrase sounding quite vague.

223 **Response 13:** We revised the sentences as: Regional processes such as long-distance
224 dispersal, evolution of sexual systems after colonization, and local climate, e.g.,
225 precipitation, together could shape the global distribution of plant sexual systems.