Biogeosciences

Discussions

Open Access

Interactive comment on "Phylogenetic support for the Tropical Niche Conservatism Hypothesis despite the absence of a clear latitudinal species richness gradient in Yunnan's woody flora" *by* G. Tang et al.

G. Tang et al.

ferryslik@hotmail.com

AUTHOR COMMENTS PLACED IN CAPS IN THE REVIEWERS COMMENTS

The authors tested for the Tropical Niche Conservatism Hypothesis using woody plants in Yunnan Province, China. There are major problems in this study. Here, I would mention some of them.

(1) The authors used herbarium collections to document species distributions and then related species compositions in geographic areas derived from the species distribu-

tions to environmental variables. As shown in several recent publications, including those for China's plants (e.g., Yang et al. 2013 Journal of Biogeography 40:1415-1426), studies using data derived from herbarium collections can result in substantially biased conclusions due to incompleteness of documentation of species richness and distributions. THIS IS INDEED TRUE IF THE ORIGINAL HERBARIUM COLLECTIONS WOULD HAVE BEEN USED. HOWEVER, WE CARRIED OUT A MODELLING EXERCISE THAT COUNTERS MOST OF THESE PROBLEMS. THIS MODELING IS BASED ON WHERE THE INDIVIDUAL COLLECTIONS OF EACH SPECIES WERE MADE AND THAN LINKING THESE LOCATIONS TO THE ENVIRONMENTAL CONDITIONS FOUND THERE TO PRODUCE A PREDICTION OF THE SUITABLE HABITAT FOR EACH INDIVIDUAL SPECIES. THESE SUITABLE HABITATS CAN SUBSEQUENTLY BE MAPPED. BECAUSE MOST PLANT COLLECTIONS ARE NOT MADE RANDOMLY, BUT ARE ENVIRONMENTALLY BIASED, WE ALSO TESTED WHETHER THERE WAS COLLECTION BIAS IN OUR DATA SET (WHICH THERE WAS) AND COMPENSATED FOR THIS BY APPLYING A VERY STRINGENT TEST SPECIFICALLY DESIGNED TO DETECT HABITAT ASSOCIATIONS AFTER ACCOUNTING FOR THIS COLLECTING BIAS. OVERALL OUR METHODS WERE SPECIFICALLY DESIGNED TO MINIMIZE THE PROBLEMS MENTIONED BY REVIEWER 2, AND ALTHOUGH THE PREDICTED SPECIES RANGES MAY NOT BE PERFECT (THEY IGNORE BIOTIC INTERACTIONS, FOR EXAMPLE), THEY DO REPRESENT THE STATE OF THE ART IN CURRENT SPECIES DISTRIBUTION MODELING, WHICH HAVE BEEN SHOWN TO PRODUCE RELATIVELY RELIABLE PREDICTIONS OF WHERE SPECIES MAY BE FOUND.

(2) Yunnan Province has about 13000 species of angiosperms, about 6000 of which are woody plants (according to Flora of Yunnan and Flora of China). The authors used only 2319 (or 40%) species of the entire woody flora of Yunnan. This incompleteness of data at the provincial level would have been translated into smaller scales used in the study. It is impossible that this low proportion of the woody flora can adequately represent the woody flora of Yunnan. Results for patterns of species richness (SR), phylogenetic

diversity (PD), and net relatedness index (NRI) can all be substantially biased when using such incomplete datasets. YES, THIS IS A VALID POINT. HOWEVER, MANY SPECIES HAD VERY FEW COLLECTIONS, AND BECAUSE OF STATISTICAL REASONS WE COULD ONLY USE SPECIES WHICH OCCURRED IN AT LEAST 5 SEPARATE LOCATIONS (GRID CELLS) IN OUR MODELLING EXERCISE. ADDITIONALLY, MANY COLLECTIONS DID NOT HAVE ADEQUATE LOCATION DESCRIPTIONS (I.E. IT WAS IMPOSSIBLE TO PINPOINT THESE COLLECTIONS TO FEW KILOMETERS ACCURACY). THIS AND SOME OTHER POINTS (NO PERMISSION TO USE FAGACEAE DATA AND EXCLUSION OF GYMNOSPERMS) EVENTUALLY RESULTED IN A DATASET THAT COMPRISED ABOUT 40% OF THE WOODY FLORA (OR ~2000 SPECIES). THIS IS STILL QUITE A LARGE AMOUNT, AND AT THE MOMENT THE MOST COMPLETE ANALYSIS EVER DONE FOR YUNNAN. MORE COLLECTION EFFORT IS NEEDED TO INCREASE THE NUMBER OF SPECIES THAT CAN BE INCLUDED IN THESE KIND OF ANALYSES. UNTIL THEN WE HAVE TO DO WITH WHAT IS AVAILABLE. SINCE THE ~2000 SPECIES USED PRESENT A VERY WIDE RANGE OF GENERA AND FAMILIES, WE THINK OUR RESULTS ARE RELIABLE ENOUGH TO ASSESS THE MAIN FLORISTIC DIVERSITY PATTERNS WITHIN YUNNAN.

(3) The authors used modeling methods to generate species distributions and then used the generated species distributions to related environmental variables. Because the species-presence data are very incomplete as mentioned above, it is very unlikely the modeled species distribution data can be reliable for their study. In addition, the current literature show that modeled species distributions can be substantially biased. It is incorrect to use modeled species distributions in the study. ACTUALLY THE BEST WAY TO DO DIVERSITY AND FLORISTIC STUDIES IN ANY REGION OF THE WORLD IS TO USE PLOT INVENTORY DATA WITH RELIABLE IDENTIFICATIONS. SUCH DATA IS HOWEVER, GENERALLY LACKING, OR THE NUMBER OF PLOTS ESTABLISHED TOO LOW TO BE USEFUL, ESPECIALLY IN THE TROPICS. THE ALTERNATIVE WAY TO STUDY DIVERSITY AND FLORISTICS IS

TO USE PRESENCE DATA ONLY (AS REPRESENTED BY HERBARIUM RECORDS) AND LINK THESE TO ENVIRONMENTAL VARIABLES AND THUS PREDICT SUITABLE HABITAT FOR EACH INDIVIDUAL SPECIES (SEE ALSO OUR COMMENT UNDER POINT (1)). BECAUSE OF THE LACK OF INVENTORY DATA IN YUNNAN, WE FOLLOWED THE SECOND APPROACH. MODELLED DISTRIBUTIONS OF INDIVIDUAL SPECIES CAN BE BIASED (AS MENTIONED BY THE REVIEWER), BUT MANY OF THESE MODELS COMBINED (AS DONE HERE) GENERALLY PRODUCE QUITE ACCURATE REGIONAL AND GLOBAL DIVERSITY AND COMPOSITION MAPS. SINCE WE CONCENTRATE ON THESE COMBINED (AS OPPOSED TO INDIVIDUAL SPECIES) DATA, WE ARE QUITE CONFIDENT IN THE PATTERNS THAT WE DETECTED FOR YUNNAN, ESPECIALLY AS THEY GENERALLY AGREE WITH WHAT IS ALREADY KNOWN ABOUT THE FLORA OF YUNNAN.

(4) Of the 4936 grid cells of 10 km by 10 km in Yunnan, the authors used only about 10% of the 4936 grid cells in their study. Such low coverage of the study area plus the low quality of the data as mentioned above would unlikely have resulted in unbiased results. According to the authors, the use of only 10% of their grid cells was to reduce computation. However, other studies have used much larger datasets in computing the same phylogenetic indices. For example, Brunbjerg et al. (2014, Journal of Plant Ecology 7:101-114) used 18463 samples. ACTUALLY THIS STATEMENT IS NOT TRUE. WE USED 10% OF GRID CELLS FOR EACH CALCULATION, BUT REPEATED THIS 50 TIMES, EACH TIME WITH A RANDOM SELECTION OF 10% OF YUNNAN'S GRID CELLS. OVERALL THIS ANALYSIS THUS COVERED YUNNAN 5 TIMES OVER, I.E. EVEN MORE THAN A SINGLE RUN WITH ALL GRID CELLS WOULD HAVE DONE. THE ADDITIONAL ADVANTAGE OF THIS APPROACH IS THAT IT ENABLES US TO COMPARE THE VARIABLES SELECTED IN EACH MODEL, I.E. PROVIDE A MEASURE OF IMPORTANCE FOR EACH INDIVIDUAL VARIABLE, WITH SOME VARIABLE BEING SELECTED IN ALL MODELS (IMPORTANCE OF 1), SOME VARIABLES ONLY IN PART OF THE MODELS (IMPORTANCE BETWEEN 0 AND 1), AND SOME VARIABLES NEVER BEING SELECTED IN ANY MODEL (IMPORTANCE 0). OUR

APPROACH THUS ADDED ADDITIONAL INFORMATION COMPARED TO A SINGLE RUN WITH ALL GRID CELLS.

In addition to major problems, there are some other potential problems. For example, at the end of section 2.4, the author indicated that their phylogeny was based on APG II (APG 2003). APG II is an outdated phylogeny. APG III (APG 2009) has been available for years and has been commonly used in phylogenetic studies. Why did you use APG II rather than APG III? THIS IS EXPLAINED IN OUR RESPONSE TO REVIEWER 1.

―――――――――――――――――――