

Dear Hervé, Thomas, managing board at PCI Evol Biol,

Thank you for the time invested in reviewing and providing constructive comments on our paper. We have taken the time to address each of the points made, resulting in what I consider to be a greatly improved text. To clarify what we have done, I am including below a copy of the entire decision letter with our responses interspersed in red type. I will also attempt to upload a version of the text with the changes tracked and your comments included in the relevant places. A clean version of the revised ms. has been uploaded as a new version of the bioRxiv preprint. To my embarrassment, I realised too late that I had omitted to include thanks to yourselves in the acknowledgements before submitting the latest version of the preprint. This will certainly go in at the next opportunity.

Best wishes,
Mike Pirie

Dear Michael David Pirie,

Your preprint, entitled **Parallel radiations of Crematosperma and Mosannona (Annonaceae), tropical rainforest trees tracking Neogene upheaval of the South American continent**, has now been reviewed. The referees' comments are enclosed. As you can see, we found your article very interesting and have suggested its recommendation, subject to certain revisions.

We shall, in principle, be happy to recommend your article as soon as it has been revised in response to the points raised by the referees. We hope to hear from you within two weeks. Please let us know if you are likely to need much longer than this to make your revisions. When your revised article is ready and you have responded to the reviewers' questions, please follow this link https://evolbiol.peercommunityin.org/user/my_articles. Once we have read the revised version, we may decide to recommend it directly, in which case, the recommendation will be published on the *PCI Evol Biol* website. Alternatively, a second round of reviews may be needed or it may be decided not to recommend your article. You will be notified by e-mail at each stage in the procedure.

Yours sincerely,

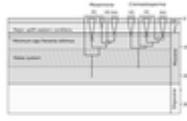
The Managing Board of *PCI Evol Biol*

PCI Evol Biol is the first community of the parent project Peer Community In. It is a community of researchers in Evolutionary Biology dedicated to both 1) the review and recommendation of preprints publicly available in preprint servers (such as bioRxiv) and 2) the recommendation of postprints published in traditional journals. This project was driven by a desire to establish a free, transparent and public recommendation system for reviewing and identifying remarkable articles. More information can be found on the website of *PCI Evol Biol* (<https://evolbiol.peercommunityin.org/>).

In case of any questions or queries, please use the following e-mail:

contact@evolbiol.peercommunityin.org.

If you wish to modify your profile or the fields and frequency of alerts, please follow this link: <https://evolbiol.peercommunityin.org/default/user/profile>.



Submitted by Michael David Pirie 2017-06-03 21:25

Michael D. Pirie, Paul J. M. Maas, Rutger A. Wilschut, Heleen Melchers-Sharrott & Lars W. Chatrou

Parallel radiations of *Crematosperma* and *Mosannona* (Annonaceae), tropical rainforest trees tracking Neogene upheaval of the South American continent

[10.1101/141127](https://doi.org/10.1101/141127)

Keywords: Andean orogeny; Pebas system; molecular dating; Neotropics; niche modelling; Panama isthmus; phylogeny reconstruction; tropical rainforest

Abstract

Much of the immense present day biological diversity of Neotropical rainforests originated from the Miocene onwards, a period of geological and ecological upheaval in South America. We assess the impact of the Andean orogeny, drainage of lake Pebas, and closure of the Panama Isthmus on two clades of trees (*Crematosperma*, c. 31 spp.; and *Mosannona*, c. 14 spp.; both Annonaceae) found in humid forest distributed across the transition zones between the Andes and Western (lowland) Amazonia and between Central and South America. We inferred phylogenies based on c. 80% of recognised species of each clade using plastid and nuclear encoded sequence markers, revealing similar patterns of geographically restricted clades. Using molecular dating we showed that the timing of diversifications in the different areas largely overlap and are consistent with Andean vicariance and Central American geodispersal. In apparent contradiction of high dispersal abilities of rainforest trees, *Crematosperma* clades within Amazonia are also geographically restricted, with a southern/montane clade that appears to have diversified along the foothills of the Andes sister to one of more northern/lowland species that diversified in a region once inundated by lake Pebas. Ecological niche modelling approaches show phylogenetically conserved niche differentiation, particularly within *Crematosperma*. Niche similarity and recent common ancestry of Amazon and Guianan *Mosannona* species contrasts with dissimilar niches and more distant ancestry of Amazon, Venezuelan and Guianan species of *Crematosperma* suggesting that this element of the similar patterns of disjunct distributions in the two genera had different biogeographic origins. The results provide further independent evidence for the importance of the Andean orogeny, the drainage of Lake Pebas, and the formation of links between South and Central America in the evolutionary history of Neotropical lowland rainforest trees.

Modifications requested

by Hervé Sauquet, 2017-07-07 09:26

Manuscript: [10.1101/141127](https://doi.org/10.1101/141127)

Decision & reviews

Dear authors,

Your preprint has now been reviewed by Thomas Couvreur and myself. You will find our comments attached. Although I was less critical than Thomas, who has more expertise in this field, I agree with most of his comments and believe that addressing them would make a stronger manuscript. Upon second thought, I also realize that this study did not use any of the recent parametric methods to reconstruct ancestral areas and biogeographic scenarios (such as Lagrange and BioGeoBears), yet most of the paper focuses on the biogeographic history of these two clades. It would be worthwhile explaining why in your text as some readers might also wonder about this.

Thanks for this suggestion – yes, I can imagine some people might expect such analysis.
Explanation added to methods:

“Given the phylogenetic results, we chose not to perform formal ancestral area analyses. The reasoning was first, that a realistic model for the biogeographic scenario would involve changing extents of areas through time, the definition of which would be likely to influence strongly the results. Second, the geographic structure in the phylogenetic trees (see Results) suggested a minimal number of range shifts, limiting the power of any parametric model (Pirie et al., 2012). We therefore adopt a parsimonious interpretation of the ancestral areas of the geographically restricted clades and use this to infer the timeframes for shifts in geographic range.”

I hope you will find these comments useful and would be very happy to recommend your preprint if you could address them. Although *PCI Evol Biol* is not a formal publication yet and we are all learning and testing the system here, I would suggest to treat this invitation to revise as you would for a traditional publication: please provide a point-by-point response to our comments and highlight your modifications in the revised manuscript.

I look forward to receiving your revised preprint.

Kind regards,

Hervé

Reviewed by Thomas Couvreur, 2017-07-07 09:26

This article sets out to test the timing and ecological dimensions of speciation within two moderate sized non sister Neotropical genera from the Annonaceae family (*Crematosperma* (31 sp.) and *Mosannona* (14 sp.)). Both genera belong however to the same tribe (*Malmeea* within the *Malmeoideae* sub family). First, they reconstruct the phylogenetic relationships separately of each of the two genera based on plastid and nuclear markers (6 plastid for both genera, and 2 additional nuclear markers for *Mosannona*). Then, using a combined plastid data matrix including both genera and several other *Malmeea* genera, they estimated the divergence times using two secondary fossil calibrations in BEAST. In addition, species distribution modeling was inferred for species in both genera and niche identity and similarity were tested. The authors show congruent biogeographic patterns with closely related species being close both in geography and ecology. The timing of diversification is inferred to be consistent with Andean vicariance and Central American dispersal. In general the methods are thorough and well applied.

General Comments: This is a good and solid new addition to studies around the origin and diversification of Neotropical biodiversity. This article provides further evidence of the importance of two major geological events in the diversification of Neotropical plant

diversification. The uplift of the Andes and the drainage of Lake Pebas in northwestern lowland Amazon. The congruence in the results of two well sampled non sister and co distributed genera adds support to the conclusions of the paper. I have however several comments. Title: It is curious that the two first words of your title never (almost) appear in the manuscript itself. "Parallel radiations". What are radiations (definition?) and what does it mean when they are parallel? Diversification instead of radiation? Moreover, the article addresses other points than just the Andean orogeny. Maybe the title needs a rethink given this.

Yes - I have switched "Radiation" with "diversification". I have introduced further explanation for the other elements of the title into the abstract, because even if their meaning might not be immediately obvious they strike me as rather apt: "Parallel" is in the sense both of "parallelism" (distributions that look similar but have, in part, different origins) and "in parallel" (happening at the same time).

"Using molecular dating we showed that diversifications in the different areas occurred in parallel, with timing consistent with Andean vicariance and Central American geodispersal."

"this element of the similar patterns of disjunct distributions in the two genera is instead a biogeographic parallelism, with differing origins"

"Upheaval" doesn't mean mountain building – this is a common, much more general, word for change or disruption.

I feel the article could do with a more focused writing. The introduction is very broad dealing very briefly with many different aspects of Neotropical biogeography. Some statements lack references both in the intro and the discussion (e.g. paragraph 2 "However, both species diversity ...").

Added references here and elsewhere as appropriate

The 5th paragraph ("In this paper..") appears too early, suggesting that it is the end of the intro, but then the authors add on an extra page and half, revisiting the major biogeographic hypotheses in the light of Annonaceae... making it confusing and hard to follow. It is a pity as this leads to a hard to read and follow intro, making the article confusing.

Moved that paragraph to the end of the intro and the one on molecular dating out of the intro (see below) – hopefully improving the flow.

The discussion does not refer to several important papers recently published.

Added references, particularly in response to specific comments.

Specific comments Intro: What is the point of the second paragraph page 4 about molecular dating? This seems out of place and more of a discussion point.

Removed and in part used to clarify the methods following Hervé's comment below.

In any case I think it could be shorter. Last paragraph of the intro: the term "lake Pebas" appears for the first time in the hypothesis, and is not directly introduced.

Added earlier in intro.

Methods: EMNTools: There are two different tests one can do: niche identity test and the background similarity test. Which one did you use and why? (see manual, and Couvreur et al 2011 for details) [Chatrou/Wilschut: We did neither of the two. Both the niche identity test and the background similarity test make pairwise comparisons of species distribution models that are generated by ENMtools. The phylogenetic trees that we inferred for *Cremastospema* and *Mosannona* are well resolved at deeper levels (i.e. relationships among clades occurring in the main biogeographical areas are well-supported), but the relationships among species are not fully resolved. Comparing species distribution models between almost any pair of sister species would incorporate this phylogenetic uncertainty, and reduce the meaningfulness of such analyses. Nevertheless, to exploit the information generated by ENMtools in a meaningful way, we plotted the niche overlap values, Hellinger's I, against phylogenetic distance.].

Added to text:

“This approach allowed a meaningful interpretation of niche overlap, despite phylogenetic uncertainty within the geographically restricted clades that in most cases prevented the identification of sister-species pairs. “

Page 7: species distribution modeling. This part is well done and the methods are well applied. Please indicate how many species have less than 10 unique occurrence points [New appendix added indicating this species by species]. Also please cite the following article where we showed that using remote sensing data in the tropics significantly improves models versus using worldclim. Why didn't the authors use that data? Deblauwe, V., Droissart, V., Bose, R., Sonké, B., Blach-Overgaard, A., Svenning, J.C., Wieringa, J.J., Ramesh, B.R., Stévant, T., & Couvreur, T.L.P. (2016) Remotely sensed temperature and precipitation data improve species distribution modelling in the tropics. *Global Ecology and Biogeography*, 25, 443–454. [Wilschut: We really weren't aware of this remote sensing data; therefore we didn't use it... I added a sentence with the reference to the discussion where I think it's most appropriate] Also it would be good to cite the following article which indicates that modeling is dependent on prevalence of the species more than on the strict number of occurrence. Statement 2 top of page 8 is thus not strictly correct. If you have a widespread species, 4 points will not be enough, but at least 13. In contrast a strict local endemic, 3 points could be ok. van Proosdij, A.S.J., Sosef, M.S.M., Wieringa, J.J., & Raes, N. (2016) Minimum required number of specimen records to develop accurate species distribution models. *Ecography*, 39, 542–552 [done]. How was independence between soil layers measured? Seems odd, but maybe I am missing something as I don't know what that type of data is made of. [Wilschut: These soil layers contain categorical data that refer to logical classes of numeric data; therefore you can test the correlations between them. Qualification added in the text “In a preliminary analysis we used Pearson correlation tests to remove correlated climate (containing continuous data) and soil (containing numeric categorical data) variables”.]

Page 11: 1st paragraph: I wouldn't call your sampling “dense” or “densely”. No doubt it is very respectable, but it isn't dense. In fact, that paragraph is a bit out of place and not very interesting. This again, leads to a confusing article. I would just directly go for the main result here.

Yes, this can go at no great loss – paragraph deleted.

3rd paragraph of discussion: It is said that Andean centered genera are the exception. Why not cite them here? If they are the exception then there aren't many, and it would make the rest of the sentence more comprehensible (multiple times, how many times?).

Done: *Malmea*, *Klarobelia*, *Cymbopetalum* are further examples.

Page 12 1st paragraph. Is “era” the correct term here? Seems too large and includes periods and ages. Check. [Changed to “period”] Several references missing for some of the statements (e.g. ref for the age of the Andean uplift etc). A ref to figure 3 would be nice after sentence 1. [Done] So in this paragraph you talk about the influence of the Andes on the origin of the species, yet no references are made to other studies that observe this same pattern... About Andean uplifts, I would recommend citing this new relevant paper: Pérez-Escobar, O.A., Chomicki, G., Condamine, F.L., Karremans, A.P., Bogarín, D., Matzke, N.J., Silvestro, D., & Antonelli, A. (2017) Recent origin and rapid speciation of Neotropical orchids in the world's richest plant biodiversity hotspot. *New Phytologist*, 215, 891–905. [Done] 3rd paragraph: Lake Pebas. Seems to me that most of this should be in the introduction. [Yes - moved part there] Here I would expect more to discuss your results to other findings (eg Roncal et al 2013).

Added refs; included more comparison

2nd paragraph: The use of “older estimates” versus younger “estimates” is confusing, especially as no results are cited (a table or a figure)... please improve the clarity of such statements. “Similar signatures” this sentence is vague I feel. Are the dates recovered by the cited studies the same? Another nice case of east west vicariance is Barfod, A.S., Trénel, P., & Borchsenius, F. (2010) Drivers of diversification in the vegetable ivory palms (Arecaceae: Ceroxyloideae, Phytelephea) – Vicariance or adaptive shifts in niche traits? *Diversity, phylogeny and evolution of Monocotyledons* (ed. by O. Seberg, A.T. Peterson, A. Barfod, and A.M. Davis), Aarhus University Press, 225-243.

Rephrased, refs added.

3rd paragraph: drainage of lake pebas. Please cite Roncal et 2013 as another study that presented this evidence in palms.

Yes - we already had.

Page 14: Closure of Isthmus. I would also add the article Cody, S., Richardson, J.E., Rull, V., Ellis, C., & Pennington, R.T. (2010) The Great American Biotic Interchange revisited. *Ecography*, 33, 326–332.

Added.

I am not convinced that plants are a good model to test hypotheses about closure of this Isthmus being more prone to long distance dispersal.

Fair point. This part added/modified in response: “and further data is warranted to further test the timing of dispersal events, potentially between separate land masses, in different groups”

The last paragraph about Annonaceae phylogenetics seems misplaced and not very interesting given the general theme of this paper. However, the discussion of the non monophyletic species is quite important, and should maybe be addressed in a separate paragraph somewhere. If the authors want to add a part about NGS then they will have to expand and better discuss the different options... such as sequence capture of nuclear regions etc. Also add references to past studies that have used these data to improve resolution (using full plastomes or nuclear genes).

Paragraph removed entirely.

Thomas Couvreur, Montpellier, 6th July 2017

Reviewed by Hervé Sauquet, 2017-07-07 09:26

This manuscript contributes to the growing number of studies aimed at understanding the drivers of diversification and biogeographic patterns in the Neotropics, considered to be the most biodiverse region in the World. As someone who is not too familiar with the specifics of this vast and rich field of research, I enjoyed very much reading the Introduction, which I thought set out the questions very clearly.

The study system here consists in two relatively small, closely related genera of the flowering plant family Annonaceae, *Crematosperma* and *Mosannonna*. The authors argue that this pair represents a good example of parallel radiations with both similarities and differences that might help us better understand the biogeographic history of the region. I was not entirely convinced by this and somehow wished that they had included more Neotropical clades (of Annonaceae and/or other plant families) in this study and applied the same methods to produce a more convincing contribution based on a larger sample of “radiations”. This being said, the strength of this paper is the near-complete sample of species they obtained in each clade and the high-quality data compiled for the niche models, and I certainly believe that careful, in-depth studies like this one are important to publish and equally useful to large meta-analyses.

Yes, a broader meta-analysis would certainly allow greater power to generalise, but indeed that is really not what this paper presents. We have done our best to address this by means of more comparison to other groups (as recommended by Thomas).

The results are interesting and not too surprising. Different clades experienced different biogeographic histories, as might be expected, but the detailed Discussion of how this may have happened is interesting. However, before recommending this paper, I believe a few minor points (below) are worth addressing to further strengthen the study.

Minor comments

Molecular dating: I agree that secondary calibration is a risky business (see also what we wrote about it in Sauquet et al. 2012) and that, given the known problems with evolutionary rates in Annonaceae (here you certainly want to cite the new paper by Paul Hoekstra et al. 2017 in *Mol Phyl Evol*), your solution to focus on *Malmeeae* only is a good idea to minimize error. However, given the paucity of fossil calibrations in Annonaceae, I would argue that

you are still likely to get the wrong age estimates here. First, I think it would be essential to clarify which analysis exactly you are relying on from Pirie & Doyle (2012) since that study involved various experiments with calibrations.

Clarified in text – we went for the full range of estimates, certainly not just the one.

Then, I would suggest to remind the reader here of which calibration(s) exactly were used in your reference analysis. I know that most of the Annonaceae community has more or less accepted the *Futabanthus* fossil, but this flower remains incomplete and I have reservations regarding its use as a calibration point until phylogenetic analyses of the fossil are actually conducted (this is why we excluded it from the review of magnoliid calibrations by Massoni et al. 2015).

Good suggestions – added to methods.

Last, I think that the normal prior used for the secondary calibration is not ideal. While it might attempt to convey the true shape of the uncertainty around the age estimate of the node in question (which in general looks more lognormal than normal), the reality of BEAST analyses is that such priors tend to be poorly informative and estimated node ages can shift largely from them, which would be problematic in the case of a node age assumed to be known (the secondary calibration). Therefore, I would recommend to either provide evidence that these normal priors were effective [I wasn't really aware of this issue, but have checked, and the posteriors are in fact a bit wider than they ought to have been – 37-21 Mya instead of 33-22 for the BEAST calibration; 69-47 instead of 52 +/- 3] or to rerun the analyses with uniform priors bounded by the 95% HPD from the original analysis [done – results added rather than replaced to illustrate this phenomenon. The conclusions are unaffected]. In summary, I believe this part of the manuscript could be improved by further justification of the choices made and discussion of how these choices may affect the conclusions presented here that are based on a direct comparison of estimated absolute ages and independently dated geological events.

Added to methods: “With these prior distributions we aimed to represent only uncertainty in the dating method (given the calibration), not the further uncertainty associated with fossil calibrations (i.e. that they represent minimum age constraints).”

Added to results: “The posterior age distributions for the root node given the normal prior were somewhat wider than the priors (37-21 Mya instead of 33-22 for the BEAST calibration; 69-47 instead of 52 +/- 3 for the PL calibration), reflected in somewhat wider/older age estimates for shallower nodes.”

I don't understand why patristic distances were calculated from PAUP* by re-estimating molecular branch lengths when the authors have RAxML, MrBayes, and BEAST trees at hand (all with branch lengths). Can't this be done easily in R without having to re-estimate branch lengths?

There would probably be simpler ways to have done this, but we just stuck with an established pipeline.

Figures 1 and 2: What is this “Unknown genus 0750”? Is it really useful to include in your analyses if you do not discuss it? The same question would apply to the various new species

included in the analyses. the focus of this paper is clearly not taxonomic so my suggestion would be to remove these taxa and sequences and keep them for a future paper.

I can see the argument, but would rather not remove these taxa at this stage. I don't think they have any negative impact on the results.

I personally do not like very much the attempt to show both branch lengths and nearly all support values: this might be too ambitious for a single figure and the curved lines might be confusing to some readers. I am not fond of the grey triangles either: would it be possible to denote areas with simple colored rectangles as you have done for Figure 3?

I personally am in favour of putting as much information as possible into single tree figures to avoid any need to compare between different ones. I have tried to tweak them for clarity, using colours to match Fig. 3.

Last, I am not sure the dual trees are so useful, since there does not seem to be so much incongruence when you include or exclude taxa for which ndhF was unavailable. I would suggest to show only the most complete trees (B) and move the other ones (A) to supplementary figures. Should you decide to keep both in the same figure, I would recommend to add a mini-caption (a text box) to quickly denote which is which, otherwise the reader might think that this is nuclear vs. plastid markers or some similar contrast.

I would be in favour of the latter – adding captions – for the same reason as above – avoids necessity for comparison.

In addition to these comments, I attach an annotated copy of the manuscript with a few additional minor suggestions.

Hervé Sauquet

Much appreciated!
On behalf of the authors,
Mike Pirie