THE ROYAL SOCIETY
PUBLISHING

PROCEEDINGS B

Development and function explain the modular evolution of phalanges in gecko lizards

Priscila S. Rothier, Monique N. Simon, Gabriel Marroig, Anthony Herrel and Tiana Kohlsdorf

Article citation details
Proc. R. Soc. B **288**: 20212300.
http://dx.doi.org/10.1098/rspb.2021.2300

Review timeline

Original submission:	9 April 2021
1st revised submission:	21 October 2021
2nd revised submission:	6 December 2021
Final acceptance:	6 December 2021

Note: Reports are unedited and appear as submitted by the referee. The review history appears in chronological order.

Review History

RSPB-2021-0843.R0 (Original submission)

Review form: Reviewer 1

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
This MS presents an interesting study of the evolution of integration/modularity in gecko feet. Geckos represent an opportunity to study the evolution of novel structures because of the repeated origins of adhesive toepads in the group. Here the authors test multiple developmental and functional hypotheses about integration and ask whether patterns are similar in padless vs. padded geckos. They find that the two groups differ, but only for one model involving, as expected, the distal phalanges.

I think the science here is very good, and I am enthusiastic about seeing this published in PRSB. However, in many places, I found the MS somewhat confusing, which leads me to believe that a substantial revision is necessary before the MS is acceptable. The problems I had reading the MS are not really due to the writing style. In fact, I found the MS to be quite well written overall. I was able to follow the argument up through the middle of the methods, but upon the introduction of the various hypotheses, I found myself quite confused.

My suggestion is to revise the paper for clarity, focusing on two main goals. First, because they are central to understanding the paper, the various hypotheses should be made extremely clear both when they are first presented (see note for lines 190-210 below) and when they appear in the results. Second, the discussion should be streamlined, particularly the first paragraph (see note for lines 321-337), which should distill the main results of the paper into their essence for the reader.

Line comments:

45: Not sure "embrace" is the right word here

87: Missing word? "[do] not always"

190-210: I had a hard time keeping all these hypotheses straight. Would it be possible to provide some sort of visual guide for the reader, perhaps a figure or even a table, to help distinguish these hypotheses? I see that figures 1 and 2 are referenced in this paragraph, but it is hard for me to determine which aspects of that figure I am supposed to be looking at. I think a separate (pre-data) figure would be more helpful.

256-257: Wording of this sentence is a bit unclear

314-317: It took me a while to understand what this sentence was saying. I think the term "outgroups removal" was confusing me. I think the main point is that considering only geckos and not outgroups made the difference between padded/padless species more apparent (but still not significant) in the phalanges model. Is this correct?

321-337: I think this paragraph could do a better job summarizing the main results of the paper. In particular, the sentence on lines 331-334, which is really the biggest result of the paper, needs a bit of elaboration to make the result 100% clear. This sentence interprets the result before stating the result, and doing the opposite would help the reader.

352-357: I find these sentences confusing. It is difficult for me to link these to a specific result in the paper.

367-370: I had a hard time with this sentence on first read. I think the problem is that I didn't really understand the what the pattern of integration of the outgroups was when I got to this point in the paper. It took looking back at figure 2 (and reminding myself of the sentence on lines 304-305) to see that yes, the outgroups had modularity that resembled the padded, not the padless species. Would it be possible to show the outgroups in figure 3 to really drive this point home?

372: "it was observed" -> "as observed"

373: "which modular structure have" -> "the modular structure of which has"

Review form: Reviewer 2

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Acceptable

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Good

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
In this paper, Nothier and colleagues test the hypothesis that specialization of the distal phalanx of some geckos (evolution of toe pads) has impacted the modularity of phalanges within a digit. They use linear measurements of X-rays to test several models of modularity that are designed to consider developmental and functional hypotheses. They conclude that, broadly, species with and without toe pads exhibit similar types of modules but they can differ in the magnitude, or strength, of covariation among phalanges.

Major Comments:
(1) It is difficult to evaluate whether the results presented here can be considered generalizable across geckos. The clade in question contains approximately 1,500 species, and only ~1.5% of that diversity is measured. It might not be a problem, per se, to sample so selectively. However, it is stated that toe pads have evolved independently multiple times. The current sampling, thus, does not remotely provide the resolution required capture this history (e.g., parsimony reconstruction of toe-pad presence/absence on the tree shown and with taxa sampled would provide equivocal evidence for a single origin of toe pads, vs what is cited as true, multiple independent originations of this trait). I'll offer two suggestions to address this critique.

I understand that these analyses require large sample sizes and it is, indeed, commendable that they've been able to generate a data set as large as it is. My first suggestion is that they include an extended and critical examination of the limitations of sampling in the discussion. This would strengthen the manuscript and also make clear what future directions should be. Second, the authors should consider presenting sister-clade comparisons, for example testing whether differences between (a) Sphaerodactylus and Gonatodes, (b) Gymnodactylus and Phyllopezus and (c) Gekko and Nactus are consistent. These results, too, will require contextualization. It's not clear from the phylogeny in Figure 1 whether any of the lineages without toe pads reflect loss of the structure, which would impact scenarios of how modularity has evolved. An example of a paper that emphasizes pairwise comparisons to test for general trends following morphological transitions is https://www.pnas.org/content/114/17/4459.

(2) The model that is called the 'functional hypothesis' can also be explained by non-functional criteria. Namely, developmental models might predict that the terminal elements of a series should differ in their variability as compared to those in the middle. This hypothesis can be traced back at least to Bateson (1913). I believe that in the evodevo literature it referred to "the property of endedness," and that it has been discussed by Alec Panchen, among others. Note that this is a critique of the model itself. Only, it should not be presented as strictly a model based on functional criteria. This has implications for how authors might interpret the plesiomorphic condition of the autopod and developmental evolution in the clade.

(3) Figures 1 and 2 as well as their figure legends are identical the PDF that I received. One of them is missing.

Minor comments:
—Clarification: Do all species analyzed have the same phalangeal formula?

—Clarification: Does the terminal phalanx of D1 lack a claw or finger nail?
—Correction, Line 36: You did not study ancestral geckos. (To me, this reads as though you are claiming that you did.)
—Clarification, Line 45: It seems incorrect to say that 'phenotypes embrace' things. This implies agency. Maybe just state that "Phenotypes have" or "can show."
—Correction, Line 56: By my recollection, the Ruvinsky and Gibson-Brown paper does not claim that individual limb segments are derived from a single module (i.e., stylopod and zeugopod are from a common module). Rather, they discuss whether the fore- and hindlimb modules are derived from a single module.
—Clarification, Line 135: You state "anterior autopodia" were studied. Was it consistently the L or R side? Or whichever had less damage? Also, maybe state in the introduction that you studied forelimbs.
—Clarification, Line 142: You state that D1 was not considered because of its distinct modularity and cite other authors. Isn't it also true that not all of the models can be applied to a digit with only two phalanges (i.e., functional model)? That seems worth pointing out.
—Clarification, Line 153: Somewhere in the paper, either methods or SI, state how sex was determined.
— Clarification, Line 179: I might have missed it. Did you ever say how you measured size? (Was it snout-vent length, total length, length of an element like the femur, etc?)
—Clarification, Figure 1: Add units to the axis next to the phylogeny.
—Clarification, Supplementary Materials S1: I couldn't tell whether the list of specimens included those that were excluded as outliers. If outliers are still in this list, can you differentiate those from specimens included in downstream analyses, perhaps by an asterisk?
—Clarification: For species where sex was determined, can you report the sex of those specimens next to specimen number.

The following comments (listed as 'suggestions') are stylistic and intended to improve the MS. These are not required changes and do not require responses by the authors. I give them merely given for the authors' consideration.
—Suggestion: If there is space, Figure S1 could be moved to the main text. For this type of study, it can help to give readers an example of what the raw data look like before moving into the analyses. It's a beautiful X-ray. The image could be coupled with a visual representation of the models that you will be testing.
—Suggestion: I wonder if it the title might more appropriately refer to modular evolution of phalanges, rather than digits. It is not as though you're comparing the modularity of digits within the autopod (e.g., analyzing D1 vs D2-5).
—Suggestion: The phrase "diversification of locomotion" (line 55) feels awkward. I wouldn't refer to the 'diversification of smelling' or other behavioral categories. Maybe something like the 'diversification of locomotory modes/strategies/behaviors' would convey the same idea but with a more intuitive phrasing.
—Suggestion, line 74: I think it would clarify your meaning to state "within phalanges of a digit"
—Suggestion, Figure 2: You might add text to the top left of each panel ("all species" vs "geckos only") so that readers can more quickly understand what the data represent.
—Suggestion, Table S4: You might emphasize more strongly results of testing for sexual dimorphism in the correlation coefficients in the main text. I don't know how many times such tests have been done before; it might be useful for others in the field to find this result easily.
—Suggestion, Line 333: The phrase "morphologically disintegrated structures" could be written more clearly. Perhaps just "reduced integration"?

Review form: Reviewer 3 (P. Polly)

Recommendation
Accept with minor revision (please list in comments)

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Comments to the Author
This paper is thoughtful and interesting. It uses multivariate quantitative analysis to show how processes of embryonic development and functional integration interact in the evolution of tetrapod limbs. Specifically, the authors use tools for studying morphological integration and modularity to demonstrate that the evolution of a novel locomotor feature (toe pads in geckos) is associated with changes to the modular structure of the bones of the feet, causing it to depart from the pattern of inter-bone correlations found in lizards without toe pads. In other words, the authors show that a novel morphological module evolved in association with the evolution of the new functional structures. Interestingly, the authors argue that the "new" module can be viewed as the co-option of the structure of integration found in the non-squamate diapsid ancestors of lizards.

The paper is nearly publishable as it is. The authors have compiled an appropriate new data set to address the question, namely measurements of foot bones from nearly 500 individual lizards representing 20 extant species. These samples are about as good as one can get for a large comparative phenotypic analysis like this. Even though larger samples of each species would be ideal for estimating covariance matrices, the authors have adopted statistical bootstrapping and Monte Carlo methods that place appropriate standard errors on their statistics and prevent them from over-interpreting patterns. Many analytical strategies are available to measure and compare the strength of phenotypic modularity. Like all of them, the ones the authors chose (correlation matrices, random skewers, and Mantel tests) have their strengths and weaknesses, but they are more than adequate for the data in this paper. Any of the known biases in these particular methods ought not be a problem for correlation matrices of this size (15 x 15). The authors also performed appropriate statistical power tests (described in the Supplemental file). While I have two specific questions about methods below, I am convinced that the authors' analyses are both appropriate and convincingly support their conclusions.

The paper will be of broad interest, thus making it suitable for Proceedings B. It will, of course, be of interest to the very large community that studies phenotypic integration and modularity, but it will also be of interest to functional morphologists, developmental biologists, and large "integrative biology" audience that is interested in the adaptive radiation of Anolis lizards.

Specific comments:

Lines 56-57: I would not immediately have assumed that limb segments in tetrapods belong to the same developmental module, but I see that the paper you cite in this line does indeed make that assumption and in its own context I understand why. Nevertheless, I would have assumed something quite different. Maybe you could tweak the wording to, "Some have argued that limb segments in tetrapods are derived from a single ancestral developmental module…".

Line 153: "We performed linear models" > "We applied linear models". (the former is grammatically correct, but it is idiomatically unusual, even though I am struggling to come up with a logical reason why).

Lines 211-225: Regarding the test between residual correlation matrix and theoretical matrix, do you have any negative correlations in the residual matrix? If so, did you test using the absolute value? (otherwise a strong negative correlation in the residual matrix will be correlated with the lowest value in theoretical matrix, i.e., 0). As an aside, I have noticed that in geometric morphometric data sets with semilandmarks, odd things happen with the theoretical correlation matrix approach, including a bias toward significance because the large number of variables in the covariance matrix equates to larger sample size in the Mantel test, as well as non-random blocks of positive, negative, and near-zero correlations that relate to the x and y directions of variation at individual semilandmarks. I don't think your data set will be prone to either effect because the number of variables is comparatively small and because all your variables are the same biological type (proximo-distal lengths of bones).

Lines 231-240: I don't completely understand the resampling strategy or why it is employed. What precisely are you sampling from the multinormal distribution, a covariance matrix or a vector of trait values? If I assume the latter, am I correct that you are essentially sampling N sets of trait values, where N is the original sample size, from a distribution that has the same correlation structure at the original sample? Since each will be slightly different within the probability density of the multinormal, the N set of trait values represents sampling error on correlation matrix and AVG values calculated from them. Why do you use this strategy here instead of the bootstrapping approach you used in lines 168-172? The latter has the advantage of not assuming multivariate normal distribution of the traits.

lines 290-291: interesting that Sphenodon is the odd one out, given that it is also phylogenetically the most distant. Do you think this is because the pattern of integration changed in squamates, or could it somehow be an artifact of the PGLS and long branch? (to rephrase the question, does this result depend on the PGLS correction?)

lines 293-295: The second part of this sentence is confusing on the first reading because it starts with statement that the five geckos have negative AVG+ values but the value reported at the end of the sentence is positive. Maybe reword the entire phrase as "… five geckos had low values for within-module integration (AVG+) among the metapodials, indicating no evidence that they form their own module in these taxa (mean AVG difference = 0.21)."

Lines 308-311: Rewording the first two sentences would make this paragraph bolder: "We identified a divergence in modularity between padless and padded geckos, the latter evolving a new functional module involving the distal phalanges. This novel module was detected in both PGLS analyses (…) and was more pronounced after outgroup removal."

Figures 1 & 2: Can you add something that helps the reader orient themselves to the anatomy? That could be stating in the caption that the illustration shows a right foot, adding digit numbers to one or more of the anatomical drawings, or labeling radius and ulna (or tibia and fibula, as appropriate).

General disclaimer: By performing this review, I have accepted that it might be published along with the paper. Nevertheless I disagree with the policy of publishing reviews on general principles: I have not had the time to polish my text for publication, nor is my work intended for an audience other than the authors and editor. By the time authors have revised the paper, my suggestions will be moot (either because the authors changed the paper or successfully rebutted my comments) and will therefore be unintelligible to anyone who did not see their original manuscript. Lastly, reviewers should not be given the freedom to insert their own unreviewed opinions and interpretations (like this one) into the scientific literature. Despite my disagreement with the policy, reviewing is an essential part of the scientific process and anyone who publishes papers reviewed by their peers should reciprocate by reviewing the papers of others.

P. David Polly
Indiana University
pdpolly@indiana.edu

Decision letter (RSPB-2021-0843.R0)

03-Jun-2021

Dear Dr Kohlsdorf:

I am writing to inform you that your manuscript RSPB-2021-0843 entitled "Development and function explain the modular evolution of digits in gecko lizards" has, in its current form, been rejected for publication in Proceedings B.

This action has been taken on the advice of referees, who have recommended that substantial revisions are necessary. With this in mind we would be happy to consider a resubmission, provided the comments of the referees are fully addressed. However please note that this is not a provisional acceptance. Reviewers and the Associate Editor are supportive, although moderate revisions are needed.

The resubmission will be treated as a new manuscript. However, we will approach the same reviewers if they are available and it is deemed appropriate to do so by the Editor. Please note that resubmissions must be submitted within six months of the date of this email. In exceptional circumstances, extensions may be possible if agreed with the Editorial Office. Manuscripts submitted after this date will be automatically rejected.

Please find below the comments made by the referees, not including confidential reports to the Editor, which I hope you will find useful. If you do choose to resubmit your manuscript, please upload the following:

1) A ʻresponse to referees' document including details of how you have responded to the comments, and the adjustments you have made.
2) A clean copy of the manuscript and one with 'tracked changes' indicating your 'response to referees' comments document.
3) Line numbers in your main document.
4) Data - please see our policies on data sharing to ensure that you are complying (https://royalsociety.org/journals/authors/author-guidelines/#data).

To upload a resubmitted manuscript, log into http://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Resubmission." Please be sure to indicate in your cover letter that it is a resubmission, and supply the previous reference number.

Sincerely,
Dr John Hutchinson
mailto: proceedingsb@royalsociety.org

Associate Editor
Comments to Author:
We have now received three reviews of this manuscript, one of which is by a specialist in the areas of statistics employed in the manuscript. All of the reviewers are generally positive about the manuscript, in line with my own assessment. Nevertheless, the reviewers detail some concerns and issues with the manuscript that must be addressed before the manuscript would be suitable for publication in Proc B. I suggest that the authors carefully address all of the referee's comments. Of note, the first reviewer recommends some editing of the manuscript for clarity and the second reviewer provides some possible approaches to address the degree to which the results are generalizable across geckos.

Reviewer(s)' Comments to Author:
Referee: 1
Comments to the Author(s)
This MS presents an interesting study of the evolution of integration/modularity in gecko feet. Geckos represent an opportunity to study the evolution of novel structures because of the repeated origins of adhesive toepads in the group. Here the authors test multiple developmental and functional hypotheses about integration and ask whether patterns are similar in padless vs. padded geckos. They find that the two groups differ, but only for one model involving, as expected, the distal phalanges.

I think the science here is very good, and I am enthusiastic about seeing this published in PRSB. However, in many places, I found the MS somewhat confusing, which leads me to believe that a substantial revision is necessary before the MS is acceptable. The problems I had reading the MS are not really due to the writing style. In fact, I found the MS to be quite well written overall. I was able to follow the argument up through the middle of the methods, but upon the introduction of the various hypotheses, I found myself quite confused.

My suggestion is to revise the paper for clarity, focusing on two main goals. First, because they are central to understanding the paper, the various hypotheses should be made extremely clear both when they are first presented (see note for lines 190-210 below) and when they appear in the results. Second, the discussion should be streamlined, particularly the first paragraph (see note for lines 321-337), which should distill the main results of the paper into their essence for the reader.

Line comments:

45: Not sure "embrace" is the right word here

87: Missing word? "[do] not always"

190-210: I had a hard time keeping all these hypotheses straight. Would it be possible to provide some sort of visual guide for the reader, perhaps a figure or even a table, to help distinguish these hypotheses? I see that figures 1 and 2 are referenced in this paragraph, but it is hard for me to determine which aspects of that figure I am supposed to be looking at. I think a separate (pre-data) figure would be more helpful.

256-257: Wording of this sentence is a bit unclear

314-317: It took me a while to understand what this sentence was saying. I think the term "outgroups removal" was confusing me. I think the main point is that considering only geckos and not outgroups made the difference between padded/padless species more apparent (but still not significant) in the phalanges model. Is this correct?

321-337: I think this paragraph could do a better job summarizing the main results of the paper. In particular, the sentence on lines 331-334, which is really the biggest result of the paper, needs a bit of elaboration to make the result 100% clear. This sentence interprets the result before stating the result, and doing the opposite would help the reader.

352-357: I find these sentences confusing. It is difficult for me to link these to a specific result in the paper.

367-370: I had a hard time with this sentence on first read. I think the problem is that I didn't really understand the what the pattern of integration of the outgroups was when I got to this point in the paper. It took looking back at figure 2 (and reminding myself of the sentence on lines 304-305) to see that yes, the outgroups had modularity that resembled the padded, not the padless species. Would it be possible to show the outgroups in figure 3 to really drive this point home?

372: "it was observed" -> "as observed"

373: "which modular structure have" -> "the modular structure of which has"

Referee: 2
Comments to the Author(s)
In this paper, Nothier and colleagues test the hypothesis that specialization of the distal phalanx of some geckos (evolution of toe pads) has impacted the modularity of phalanges within a digit. They use linear measurements of X-rays to test several models of modularity that are designed to consider developmental and functional hypotheses. They conclude that, broadly, species with and without toe pads exhibit similar types of modules but they can differ in the magnitude, or strength, of covariation among phalanges.

Major Comments:
(1) It is difficult to evaluate whether the results presented here can be considered generalizable across geckos. The clade in question contains approximately 1,500 species, and only ~1.5% of that diversity is measured. It might not be a problem, per se, to sample so selectively. However, it is stated that toe pads have evolved independently multiple times. The current sampling, thus, does not remotely provide the resolution required capture this history (e.g., parsimony reconstruction of toe-pad presence/absence on the tree shown and with taxa sampled would provide equivocal evidence for a single origin of toe pads, vs what is cited as true, multiple independent originations of this trait). I'll offer two suggestions to address this critique.

I understand that these analyses require large sample sizes and it is, indeed, commendable that they've been able to generate a data set as large as it is. My first suggestion is that they include an extended and critical examination of the limitations of sampling in the discussion. This would strengthen the manuscript and also make clear what future directions should be. Second, the authors should consider presenting sister-clade comparisons, for example testing whether differences between (a) Sphaerodactylus and Gonatodes, (b) Gymnodactylus and Phyllopezus and (c) Gekko and Nactus are consistent. These results, too, will require contextualization. It's not clear from the phylogeny in Figure 1 whether any of the lineages without toe pads reflect loss of the structure, which would impact scenarios of how modularity has evolved. An example of a paper that emphasizes pairwise comparisons to test for general trends following morphological transitions is https://www.pnas.org/content/114/17/4459.

(2) The model that is called the 'functional hypothesis' can also be explained by non-functional criteria. Namely, developmental models might predict that the terminal elements of a series should differ in their variability as compared to those in the middle. This hypothesis can be traced back at least to Bateson (1913). I believe that in the evodevo literature it referred to "the property of endedness," and that it has been discussed by Alec Panchen, among others. Note that

this is a critique of the model itself. Only, it should not be presented as strictly a model based on functional criteria. This has implications for how authors might interpret the plesiomorphic condition of the autopod and developmental evolution in the clade.

(3) Figures 1 and 2 as well as their figure legends are identical the PDF that I received. One of them is missing.

Minor comments:
—Clarification: Do all species analyzed have the same phalangeal formula?
—Clarification: Does the terminal phalanx of D1 lack a claw or finger nail?
—Correction, Line 36: You did not study ancestral geckos. (To me, this reads as though you are claiming that you did.)
—Clarification, Line 45: It seems incorrect to say that 'phenotypes embrace' things. This implies agency. Maybe just state that "Phenotypes have" or "can show."
—Correction, Line 56: By my recollection, the Ruvinsky and Gibson-Brown paper does not claim that individual limb segments are derived from a single module (i.e., stylopod and zeugopod are from a common module). Rather, they discuss whether the fore- and hindlimb modules are derived from a single module.
—Clarification, Line 135: You state "anterior autopodia" were studied. Was it consistently the L or R side? Or whichever had less damage? Also, maybe state in the introduction that you studied forelimbs.
—Clarification, Line 142: You state that D1 was not considered because of its distinct modularity and cite other authors. Isn't it also true that not all of the models can be applied to a digit with only two phalanges (i.e., functional model)? That seems worth pointing out.
—Clarification, Line 153: Somewhere in the paper, either methods or SI, state how sex was determined.
— Clarification, Line 179: I might have missed it. Did you ever say how you measured size? (Was it snout-vent length, total length, length of an element like the femur, etc?)
—Clarification, Figure 1: Add units to the axis next to the phylogeny.
—Clarification, Supplementary Materials S1: I couldn't tell whether the list of specimens included those that were excluded as outliers. If outliers are still in this list, can you differentiate those from specimens included in downstream analyses, perhaps by an asterisk?
—Clarification: For species where sex was determined, can you report the sex of those specimens next to specimen number.

The following comments (listed as 'suggestions') are stylistic and intended to improve the MS. These are not required changes and do not require responses by the authors. I give them merely given for the authors' consideration.
—Suggestion: If there is space, Figure S1 could be moved to the main text. For this type of study, it can help to give readers an example of what the raw data look like before moving into the analyses. It's a beautiful X-ray. The image could be coupled with a visual representation of the models that you will be testing.
—Suggestion: I wonder if it the title might more appropriately refer to modular evolution of phalanges, rather than digits. It is not as though you're comparing the modularity of digits within the autopod (e.g., analyzing D1 vs D2-5).
—Suggestion: The phrase "diversification of locomotion" (line 55) feels awkward. I wouldn't refer to the 'diversification of smelling' or other behavioral categories. Maybe something like the 'diversification of locomotory modes/strategies/behaviors' would convey the same idea but with a more intuitive phrasing.
—Suggestion, line 74: I think it would clarify your meaning to state "within phalanges of a digit"
—Suggestion, Figure 2: You might add text to the top left of each panel ("all species" vs "geckos only") so that readers can more quickly understand what the data represent.
—Suggestion, Table S4: You might emphasize more strongly results of testing for sexual dimorphism in the correlation coefficients in the main text. I don't know how many times such tests have been done before; it might be useful for others in the field to find this result easily.

—Suggestion, Line 333: The phrase "morphologically disintegrated structures" could be written more clearly. Perhaps just "reduced integration"?

Referee: 3
Comments to the Author(s)
This paper is thoughtful and interesting. It uses multivariate quantitative analysis to show how processes of embryonic development and functional integration interact in the evolution of tetrapod limbs. Specifically, the authors use tools for studying morphological integration and modularity to demonstrate that the evolution of a novel locomotor feature (toe pads in geckos) is associated with changes to the modular structure of the bones of the feet, causing it to depart from the pattern of inter-bone correlations found in lizards without toe pads. In other words, the authors show that a novel morphological module evolved in association with the evolution of the new functional structures. Interestingly, the authors argue that the "new" module can be viewed as the co-option of the structure of integration found in the non-squamate diapsid ancestors of lizards.

The paper is nearly publishable as it is. The authors have compiled an appropriate new data set to address the question, namely measurements of foot bones from nearly 500 individual lizards representing 20 extant species. These samples are about as good as one can get for a large comparative phenotypic analysis like this. Even though larger samples of each species would be ideal for estimating covariance matrices, the authors have adopted statistical bootstrapping and Monte Carlo methods that place appropriate standard errors on their statistics and prevent them from over-interpreting patterns. Many analytical strategies are available to measure and compare the strength of phenotypic modularity. Like all of them, the ones the authors chose (correlation matrices, random skewers, and Mantel tests) have their strengths and weaknesses, but they are more than adequate for the data in this paper. Any of the known biases in these particular methods ought not be a problem for correlation matrices of this size (15 x 15). The authors also performed appropriate statistical power tests (described in the Supplemental file). While I have two specific questions about methods below, I am convinced that the authors' analyses are both appropriate and convincingly support their conclusions.

The paper will be of broad interest, thus making it suitable for Proceedings B. It will, of course, be of interest to the very large community that studies phenotypic integration and modularity, but it will also be of interest to functional morphologists, developmental biologists, and large "integrative biology" audience that is interested in the adaptive radiation of Anolis lizards.

Specific comments:

Lines 56-57: I would not immediately have assumed that limb segments in tetrapods belong to the same developmental module, but I see that the paper you cite in this line does indeed make that assumption and in its own context I understand why. Nevertheless, I would have assumed something quite different. Maybe you could tweak the wording to, "Some have argued that limb segments in tetrapods are derived from a single ancestral developmental module…".

Line 153: "We performed linear models" > "We applied linear models". (the former is grammatically correct, but it is idiomatically unusual, even though I am struggling to come up with a logical reason why).

Lines 211-225: Regarding the test between residual correlation matrix and theoretical matrix, do you have any negative correlations in the residual matrix? If so, did you test using the absolute value? (otherwise a strong negative correlation in the residual matrix will be correlated with the lowest value in theoretical matrix, i.e., 0). As an aside, I have noticed that in geometric morphometric data sets with semilandmarks, odd things happen with the theoretical correlation matrix approach, including a bias toward significance because the large number of variables in the covariance matrix equates to larger sample size in the Mantel test, as well as non-random blocks of positive, negative, and near-zero correlations that relate to the x and y directions of variation at individual semilandmarks. I don't think your data set will be prone to either effect because the number of variables is comparatively small and because all your variables are the same biological type (proximo-distal lengths of bones).

Lines 231-240: I don't completely understand the resampling strategy or why it is employed. What precisely are you sampling from the multinormal distribution, a covariance matrix or a vector of trait values? If I assume the latter, am I correct that you are essentially sampling N sets of trait values, where N is the original sample size, from a distribution that has the same correlation structure at the original sample? Since each will be slightly different within the probability density of the multinormal, the N set of trait values represents sampling error on correlation matrix and AVG values calculated from them. Why do you use this strategy here instead of the bootstrapping approach you used in lines 168-172? The latter has the advantage of not assuming multivariate normal distribution of the traits.

lines 290-291: interesting that Sphenodon is the odd one out, given that it is also phylogenetically the most distant. Do you think this is because the pattern of integration changed in squamates, or could it somehow be an artifact of the PGLS and long branch? (to rephrase the question, does this result depend on the PGLS correction?)

lines 293-295: The second part of this sentence is confusing on the first reading because it starts with statement that the five geckos have negative AVG+ values but the value reported at the end of the sentence is positive. Maybe reword the entire phrase as "… five geckos had low values for within-module integration (AVG+) among the metapodials, indicating no evidence that they form their own module in these taxa (mean AVG difference = 0.21)."

Lines 308-311: Rewording the first two sentences would make this paragraph bolder: "We identified a divergence in modularity between padless and padded geckos, the latter evolving a new functional module involving the distal phalanges. This novel module was detected in both PGLS analyses (…) and was more pronounced after outgroup removal."

Figures 1 & 2: Can you add something that helps the reader orient themselves to the anatomy? That could be stating in the caption that the illustration shows a right foot, adding digit numbers to one or more of the anatomical drawings, or labeling radius and ulna (or tibia and fibula, as appropriate).

General disclaimer: By performing this review, I have accepted that it might be published along with the paper. Nevertheless I disagree with the policy of publishing reviews on general principles: I have not had the time to polish my text for publication, nor is my work intended for an audience other than the authors and editor. By the time authors have revised the paper, my suggestions will be moot (either because the authors changed the paper or successfully rebutted my comments) and will therefore be unintelligible to anyone who did not see their original manuscript. Lastly, reviewers should not be given the freedom to insert their own unreviewed opinions and interpretations (like this one) into the scientific literature. Despite my disagreement with the policy, reviewing is an essential part of the scientific process and anyone who publishes papers reviewed by their peers should reciprocate by reviewing the papers of others.

P. David Polly
Indiana University
pdpolly@indiana.edu

Author's Response to Decision Letter for (RSPB-2021-0843.R0)

See Appendix A.

RSPB-2021-2300.R0

Review form: Reviewer 1

Recommendation
Accept as is

Scientific importance: Is the manuscript an original and important contribution to its field?
Excellent

General interest: Is the paper of sufficient general interest?
Good

Quality of the paper: Is the overall quality of the paper suitable?
Excellent

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

Is it accessible?
Yes

Is it clear?
Yes

Is it adequate?
Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
The authors have done an excellent job revising their MS in light of the reviewers' original comments. I am satisfied with their responses to my review and have no additional concerns. I believe that the MS should be accepted as is.

Review form: Reviewer 2

Recommendation
Major revision is needed (please make suggestions in comments)

Scientific importance: Is the manuscript an original and important contribution to its field?
Good

General interest: Is the paper of sufficient general interest?
Acceptable

Quality of the paper: Is the overall quality of the paper suitable?
Acceptable

Is the length of the paper justified?
Yes

Should the paper be seen by a specialist statistical reviewer?
No

Do you have any concerns about statistical analyses in this paper? If so, please specify them explicitly in your report.
No

It is a condition of publication that authors make their supporting data, code and materials available - either as supplementary material or hosted in an external repository. Please rate, if applicable, the supporting data on the following criteria.

> **Is it accessible?**
> Yes
>
> **Is it clear?**
> Yes
>
> **Is it adequate?**
> Yes

Do you have any ethical concerns with this paper?
No

Comments to the Author
In this study, the authors analyze the variational modularity of phalanges in geckos. Testing two hypothetical models, they describe the following major results: (1) Despite functional specialization, geckos exhibit a pattern of modularity that was previously described by Kavanagh et al (2013) and is plesiomorphic to tetrapods. (2) Non-toepad bearing geckos show a derived distal phalangeal module.

Major comment:
This second result is based upon phylogenetic regression (PGLS, shown in Fig. 3). My concern is that applying phylogenetic methods to a so highly restricted taxonomic sample may produce misleading results.

According to the authors (line 414, citing ref. 17), all of the geckos without toepads are inferred to never have had this character (i.e., toe pads have not been lost in any of these species). Therefore, the 10 genera with toe pads that they analyze represent at least 8 independent origins of the structure. It is the mismatch between this fact and what would be recovered by parsimony reconstructions about gains and losses for their sample (see tree in Fig. 2) that gives me pause. To be clear, it is not my intention to contest the evolutionary scenario of no loss. Rather, I think that when taxonomic sampling doesn't capture what we understand to be the major evolutionary history of a trait, it raises obvious questions about the robustness of results from phylogenetic methods.

That is why in my original review I suggested a non-phylogenetic method that might allow the authors to get around sampling limitations. Specifically, pairwise comparisons of sister clades with and without toe pads provide another means of testing for consistent differences between the groups. The authors now present this analysis in the supplementary materials. Pairwise comparisons (Fig. S3 D) to not clearly support the major conclusion that padless taxa have a derived condition of higher degree of modularity in the distal phalanges: Sphaerodactylus is overlapping the padless species; Phyllopezus has a lower value than the padless species; Gekko has a higher value than the padless species. How do the authors reconcile this observation? I would encourage the authors to consider whether there other ways to explore the robustness of their conclusion of a derived distal module in padless taxa. The paper will be improved if they can confront these questions analytically or via argumentation in the text.

Minor comments:
Line 40: Would it be more correct to say "Organisms have variable…"?
Line 55: Did you mean "which represent"?
Line 75: The line word "morphological" feels unnecessary.
Line 221: Did you mean "proximal" instead of "anterior"?
Line 279: To confirm, does Figure 2 reflect the resolved topology? It looks like a polytomy.
Line 305: You may want to make a new paragraph with the sentence that begins on this line.
Line 345: As a suggestion, don't just say "not significant." Give the p value. I believe it's 0.06, which makes your observation seem more reasonable.
Figure S2 and Table S8: Did you mean "metacarpal" instead of "metapodial"? That is how you've labeled the module in Fig. 2 and 3.
Figure S3: Functional is misspelled.

Decision letter (RSPB-2021-2300.R0)

29-Nov-2021

Dear Dr Kohlsdorf

I am pleased to inform you that your manuscript RSPB-2021-2300 entitled "Development and function explain the modular evolution of phalanges in gecko lizards." has been accepted for publication in Proceedings B. Congratulations!!

The referee(s) have recommended publication, but also suggest some minor revisions to your manuscript. Therefore, I invite you to respond to the referee(s)' comments and revise your manuscript. Because the schedule for publication is very tight, it is a condition of publication that you submit the revised version of your manuscript within 7 days. If you do not think you will be able to meet this date please let us know.

It is vital that the main recommendation, to tone down the interpretation (i.e. add caution) and/or new analyses to address the reviewer's concerns, is adequately addressed, but this may be easy enough.

To revise your manuscript, log into https://mc.manuscriptcentral.com/prsb and enter your Author Centre, where you will find your manuscript title listed under "Manuscripts with Decisions." Under "Actions," click on "Create a Revision." Your manuscript number has been appended to denote a revision. You will be unable to make your revisions on the originally submitted version of the manuscript. Instead, revise your manuscript and upload a new version through your Author Centre.

When submitting your revised manuscript, you will be able to respond to the comments made by the referee(s) and upload a file "Response to Referees". You can use this to document any changes you make to the original manuscript. We require a copy of the manuscript with revisions made since the previous version marked as 'tracked changes' to be included in the 'response to referees' document.

Before uploading your revised files please make sure that you have:

1) A text file of the manuscript (doc, txt, rtf or tex), including the references, tables (including captions) and figure captions. Please remove any tracked changes from the text before submission. PDF files are not an accepted format for the "Main Document".

2) A separate electronic file of each figure (tiff, EPS or print-quality PDF preferred). The format should be produced directly from original creation package, or original software format. PowerPoint files are not accepted.

3) Electronic supplementary material: this should be contained in a separate file and where possible, all ESM should be combined into a single file. All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

Online supplementary material will also carry the title and description provided during submission, so please ensure these are accurate and informative. Note that the Royal Society will not edit or typeset supplementary material and it will be hosted as provided. Please ensure that the supplementary material includes the paper details (authors, title, journal name, article DOI). Your article DOI will be 10.1098/rspb.[paper ID in form xxxx.xxxx e.g. 10.1098/rspb.2016.0049].

4) A media summary: a short non-technical summary (up to 100 words) of the key findings/importance of your manuscript.

5) Data accessibility section and data citation
It is a condition of publication that data supporting your paper are made available either in the electronic supplementary material or through an appropriate repository (https://royalsociety.org/journals/authors/author-guidelines/#data).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should be fully cited. To ensure archived data are available to readers, authors should include a 'data accessibility' section immediately after the acknowledgements section. This should list the database and accession number for all data from the article that has been made publicly available, for instance:

- DNA sequences: Genbank accessions F234391-F234402
- Phylogenetic data: TreeBASE accession number S9123
- Final DNA sequence assembly uploaded as online supplemental material
- Climate data and MaxEnt input files: Dryad doi:10.5521/dryad.12311

NB. From April 1 2013, peer reviewed articles based on research funded wholly or partly by RCUK must include, if applicable, a statement on how the underlying research materials – such as data, samples or models – can be accessed. This statement should be included in the data accessibility section.

If you wish to submit your data to Dryad (http://datadryad.org/) and have not already done so you can submit your data via this link http://datadryad.org/submit?journalID=RSPB&manu=(Document not available) which will take you to your unique entry in the Dryad repository. If you have already submitted your data to dryad you can make any necessary revisions to your dataset by following the above link. Please see https://royalsociety.org/journals/ethics-policies/data-sharing-mining/ for more details.

6) For more information on our Licence to Publish, Open Access, Cover images and Media summaries, please visit https://royalsociety.org/journals/authors/author-guidelines/.

Once again, thank you for submitting your manuscript to Proceedings B and I look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Sincerely,
Dr John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member
Comments to Author:
We have now received two reviews of this resubmitted manuscript. Reviewer #1 has no additional comments that need to be addressed. However, Reviewer #2 suggests that the authors temper their conclusions by discussing/mentioning the limitations of their approach. I suggest that the authors comply with Reviewer #2's suggestion. Otherwise, both reviewers, and this editor, are impressed by the revised manuscript and believe it would be an excellent contribution to Proc B. Well done!

Reviewer(s)' Comments to Author:
Referee: 1
Comments to the Author(s).
The authors have done an excellent job revising their MS in light of the reviewers' original comments. I am satisfied with their responses to my review and have no additional concerns. I believe that the MS should be accepted as is.

Referee: 2
Comments to the Author(s).
In this study, the authors analyze the variational modularity of phalanges in geckos. Testing two hypothetical models, they describe the following major results: (1) Despite functional specialization, geckos exhibit a pattern of modularity that was previously described by Kavanagh et al (2013) and is plesiomorphic to tetrapods. (2) Non-toepad bearing geckos show a derived distal phalangeal module.

Major comment:
This second result is based upon phylogenetic regression (PGLS, shown in Fig. 3). My concern is that applying phylogenetic methods to a so highly restricted taxonomic sample may produce misleading results.

According to the authors (line 414, citing ref. 17), all of the geckos without toepads are inferred to never have had this character (i.e., toe pads have not been lost in any of these species). Therefore, the 10 genera with toe pads that they analyze represent at least 8 independent origins of the structure. It is the mismatch between this fact and what would be recovered by parsimony reconstructions about gains and losses for their sample (see tree in Fig. 2) that gives me pause. To be clear, it is not my intention to contest the evolutionary scenario of no loss. Rather, I think that when taxonomic sampling doesn’t capture what we understand to be the major evolutionary history of a trait, it raises obvious questions about the robustness of results from phylogenetic methods.

That is why in my original review I suggested a non-phylogenetic method that might allow the authors to get around sampling limitations. Specifically, pairwise comparisons of sister clades with and without toe pads provide another means of testing for consistent differences between the groups. The authors now present this analysis in the supplementary materials. Pairwise comparisons (Fig. S3 D) to not clearly support the major conclusion that padless taxa have a derived condition of higher degree of modularity in the distal phalanges: Sphaerodactylus is overlapping the padless species; Phyllopezus has a lower value than the padless species; Gekko has a higher value than the padless species. How do the authors reconcile this observation? I would encourage the authors to consider whether there other ways to explore the robustness of their conclusion of a derived distal module in padless taxa. The paper will be improved if they can confront these questions analytically or via argumentation in the text.

Minor comments:
Line 40: Would it be more correct to say "Organisms have variable…"?
Line 55: Did you mean "which represent"?
Line 75: The line word "morphological" feels unnecessary.
Line 221: Did you mean "proximal" instead of "anterior"?
Line 279: To confirm, does Figure 2 reflect the resolved topology? It looks like a polytomy.
Line 305: You may want to make a new paragraph with the sentence that begins on this line.
Line 345: As a suggestion, don't just say "not significant." Give the p value. I believe it's 0.06, which makes your observation seem more reasonable.
Figure S2 and Table S8: Did you mean "metacarpal" instead of "metapodial"? That is how you've labeled the module in Fig. 2 and 3.
Figure S3: Functional is misspelled.

Author's Response to Decision Letter for (RSPB-2021-2300.R0)

See Appendix B.

Decision letter (RSPB-2021-2300.R1)

06-Dec-2021

Dear Dr Kohlsdorf

I am pleased to inform you that your manuscript entitled "Development and function explain the modular evolution of phalanges in gecko lizards." has been accepted for publication in Proceedings B.

You can expect to receive a proof of your article from our Production office in due course, please check your spam filter if you do not receive it. PLEASE NOTE: you will be given the exact page length of your paper which may be different from the estimation from Editorial and you may be asked to reduce your paper if it goes over the 10 page limit.

If you are likely to be away from e-mail contact please let us know. Due to rapid publication and an extremely tight schedule, if comments are not received, we may publish the paper as it stands.

If you have any queries regarding the production of your final article or the publication date please contact procb_proofs@royalsociety.org

Your article has been estimated as being 9 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Data Accessibility section
Please remember to make any data sets live prior to publication, and update any links as needed when you receive a proof to check. It is good practice to also add data sets to your reference list.

Open Access
You are invited to opt for Open Access, making your freely available to all as soon as it is ready for publication under a CCBY licence. Our article processing charge for Open Access is £1700. Corresponding authors from member institutions (http://royalsocietypublishing.org/site/librarians/allmembers.xhtml) receive a 25% discount to these charges. For more information please visit http://royalsocietypublishing.org/open-access.

Paper charges
An e-mail request for payment of any related charges will be sent out shortly. The preferred payment method is by credit card; however, other payment options are available.

Electronic supplementary material:
All supplementary materials accompanying an accepted article will be treated as in their final form. They will be published alongside the paper on the journal website and posted on the online figshare repository. Files on figshare will be made available approximately one week before the accompanying article so that the supplementary material can be attributed a unique DOI.

You are allowed to post any version of your manuscript on a personal website, repository or preprint server. However, the work remains under media embargo and you should not discuss it with the press until the date of publication. Please visit https://royalsociety.org/journals/ethics-policies/media-embargo for more information.

Thank you for your fine contribution. On behalf of the Editors of the Proceedings B, we look forward to your continued contributions to the Journal.

Sincerely,
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

FFCLRP

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO. DEPARTAMENTO DE BIOLOGIA.
AV. BANDEIRANTES, 3900. BAIRRO MONTE ALEGRE. RIBEIRÃO PRETO, SP. CEP 14040-901.
FAX: (16) 3602-4886

Editorial Office of Proceedings of the Royal Society B
Dr. Spencer Barrett, Editor-in-Chief

Response Letter: Resubmission of manuscript ID RSPB-2021-0843

Dear Dr. Spencer Barrett

I am submitting a new and fully revised version of our research article entitled "**Development and function explain the modular evolution of phalanges in gecko lizards**" (originally assigned as RSPB-2021-0843), authored by PS Rothier, MN Simon, G Marroig, A Herrel & T Kohlsdorf, to be considered for publication as a research article at the Proceedings of the Royal Society B. We acknowledge the positive feedback received from the reviewers and praise the very useful suggestions provided, which improved the clarity and precision of our text. We fully incorporated the suggestions in this revised version of the manuscript, and also performed new analyses. I am uploading a copy of the new version of the manuscript with all changes indicated in red, and here I also detail our answers to the reviewer's criticisms, describing how we incorporated the suggestions in the text. I acknowledge the attention given to our submission and the opportunity of having our manuscript being evaluated one more time at the Proceedings of the Royal Society B. I state that all data used is provided as supplementary files, and all authors agreed with this version of the manuscript. I appreciate your work as Editor of the journal Proceedings of the Royal Society B, and acknowledge in advance the attention given to our submission. Thank you for your time and consideration.

Sincerely,

Tiana Kohlsdorf

DETAILED ANSWERS TO THE REVIEWERS

Reviewer #1:

General comments to the authors:

This MS presents an interesting study of the evolution of integration/modularity in gecko feet. Geckos represent an opportunity to study the evolution of novel structures because of the repeated origins of adhesive toepads in the group. Here the authors test multiple developmental and functional hypotheses about integration and ask whether patterns are similar in padless vs. padded geckos. They find that the two groups differ, but only for one model involving, as expected, the distal phalanges.

I think the science here is very good, and I am enthusiastic about seeing this published in PRSB. However, in many places, I found the MS somewhat confusing, which leads me to believe that a substantial revision is necessary before the MS is acceptable. The problems I had reading the MS are not really due to the writing style. In fact, I found the MS to be quite well written overall. I was able to follow the argument up through the middle of the methods, but upon the introduction of the various hypotheses, I found myself quite confused. My suggestion is to revise the paper for clarity, focusing on two main goals. First, because they are central to understanding the paper, the various hypotheses should be made extremely clear both when they are first presented (see note for lines 190-210 below) and when they appear in the results. Second, the discussion should be streamlined, particularly the first paragraph (see note for lines 321-337), which should distill the main results of the paper into their essence for the reader.

Answer: We thank the thoughtful comments provided by the reviewer; we incorporated these suggestions to resubmit a new version of the manuscript that is easier to follow. Specifically, we clarify the two modularity hypotheses by including one additional figure (new Figure 1), which provides a visual aid on how we grouped the measured traits into the developmental and the functional modules. The main results are now more clearly stated, and we improved the discussion section for clarity and precision, as further detailed.

Specific comments:

1.1) Line 45: Not sure "embrace" is the right word here

Answer: We replaced the word "embrace" by "have" (line 40 of the revised version).

1.2) Line 87: Missing word? "[do] not always"

Answer: We corrected the sentence accordingly (line 82 of the revised version).

1.3) Lines 190-210: I had a hard time keeping all these hypotheses straight. Would it be possible to provide some sort of visual guide for the reader, perhaps a figure or even a table, to help distinguish these hypotheses? I see that figures 1 and 2 are referenced in this paragraph, but it is hard for me to determine which aspects of that figure I am supposed to be looking at. I think a separate (pre-data) figure would be more helpful.

Answer: We acknowledge the reviewer's suggestion, which certainly contributes for a proper understanding of our article. Accordingly, we included a new Figure 1, which corresponds to a 'pre-data image' that illustrates the two hypotheses of modularity using radiographs of a gecko autopodium. In this figure, each module is represented by a different colour. The figure also indicates the linear measurements obtained and the anatomical identification of key-bones.

1.4) Lines 256-257: Wording of this sentence is a bit unclear

Answer: We agree that the sentence was unclear, and therefore rephrased it as follows: "We used the R package ape [51] to calculate the structure of phylogenetic covariation between species assuming that traits evolved under a Brownian Motion model." (Lines 273-274 of the revised version).

1.5) Lines 314-317: It took me a while to understand what this sentence was saying. I think the term "outgroups removal" was confusing me. I think the main point is that considering only geckos and not outgroups made the difference between padded/padless species more apparent (but still not significant) in the phalanges model. Is this correct?

Answer: Yes, the interpretation is correct, and we agree that the original sentence required clarification. We rephrased the sentence as follows: "In the developmental hypothesis, the difference between pad-bearing and padless geckos was more apparent when we excluded the two outgroup species from the PGLS analysis (Figure 3, Table S8). Although the difference was not significant, analysis performed only with geckos suggests that padded species tend to have more integrated phalanges than the padless ones." (Lines 339-343 of the revised version).

1.6) Lines 321-337: I think this paragraph could do a better job summarizing the main results of the paper. In particular, the sentence on lines 331-334, which is really the biggest result of the paper, needs a bit of *elaboration to make the result 100% clear. This sentence interprets the result before stating the result, and doing the opposite would help the reader.*

Answer: Following the reviewer suggestion, we modified the sentences as follows (lines 352- 359 of the revised version): "Our results corroborate that the developmental model of digit formation [see 13] is a strong hypothesis to explain the modularity of hand bones for most species, regardless of the presence of adhesive toepads. Functional modules also exhibited strong modularity signal both in padded and padless geckos. However, although padded geckos exhibit a functional specialization that includes hyperextension of the distal phalanges during setal attachment and release, the padless species are the ones that show a new distal functional module, contradicting our expectation that this distal module would exhibit greater modular degree in padded geckos."

1.7) 352-357: I find these sentences confusing. It is difficult for me to link these to a specific result in the paper.

Answer: We thank the reviewer for such observation, and recognize that the main point of this passage, which focuses on the difference between evolution of trait covariation and mean trait evolution, was not clear. We clarified the sentences as follows (lines 376-384 of the revised version): "However, the conserved patterns of covariation and correlation among hand bones identified in geckos do not imply invariable mean morphology of phalangeal size and proportions in this lineage. Autopodial elements of geckos display a variety of shapes and sizes, as exemplified by the extremely reduced intermediate phalanx of *Hemidactylus* [23,29] and the overall shorter phalangeal proportions in pad bearing species [25]. Such mean morphological variation across species is possible even though species show conserved developmental modularity, and may occur because integration among autopodium elements is not complete, allowing some independent variation of single elements across species."

1.8) 367-370: I had a hard time with this sentence on first read. I think the problem is that I didn't really understand the what the pattern of integration of the outgroups was when I got to this point in the paper. It took looking back at figure 2 (and reminding myself of the sentence on lines 304-305) to see that yes, the outgroups had modularity that resembled the padded, not the padless species. Would it be possible to show the outgroups in figure 3 to really drive this point home?

Answer: We agree that changes in Figure 3 likely contribute for highlighting some visual patterns and provide a better support for discussions involving reference to the outgroup lineages. Therefore, we now indicate the two outgroups in each boxplot for the padless species using white dots (Figure 3A and B [all species] of the resubmitted version). Moreover, we have also rephrased the corresponding sentence in the discussion, in order to clarify the associated discussion (lines 395-397 of the revised version).

1.9) 372: "it was observed" -> "as observed"

Answer: We followed the suggestion and modified the sentence accordingly (line 399 of the revised version).

1.10) 373: "which modular structure have" -> "the modular structure of which has"

Answer: We modified the sentence according to the suggestion (line 400 of the revised version).

Reviewer # 2:

Comments to the Author(s)

In this paper, Rothier and colleagues test the hypothesis that specialization of the distal phalanx of some geckos (evolution of toe pads) has impacted the modularity of phalanges within a digit. They use linear measurements of X-rays to test several models of modularity that are designed to consider developmental and functional hypotheses. They conclude that, broadly, species with and without toe pads exhibit similar types of modules but they can differ in the magnitude, or strength, of covariation among phalanges.

Major Comments:

(1) It is difficult to evaluate whether the results presented here can be considered generalizable across geckos. The clade in question contains approximately 1,500 species, and only ~1.5% of that diversity is measured. It might not be a problem, per se, to sample so selectively. However, it is stated that toe pads have evolved independently multiple times. The current sampling, thus, does not remotely provide the resolution required capture this history (e.g., parsimony reconstruction of toe-pad presence/absence on the tree shown and with taxa sampled would provide equivocal evidence for a single origin of toe pads, vs what is cited as true, multiple independent originations of this trait). I'll offer two suggestions to address this critique.

I understand that these analyses require large sample sizes and it is, indeed, commendable that they've *been able to generate a data set as large as it is. My first suggestion is that they include an extended and critical examination of the limitations of sampling in the discussion. This would strengthen the manuscript and also make clear what future directions should be. Second, the authors should consider presenting sister-clade comparisons, for example testing whether differences between (a) Sphaerodactylus and Gonatodes, (b) Gymnodactylus and Phyllopezus and (c) Gekko and Nactus are consistent. These results, too, will require contextualization. It's not clear from the phylogeny in Figure 1 whether any of the lineages without toe pads reflect loss of the structure, which would impact scenarios of how modularity has evolved. An example of a paper that emphasizes pairwise comparisons to test for general trends following morphological transitions is https://www.pnas.org/content/114/17/4459.*

Answer: We understand the concern raised by the reviewer and recognize that our sample corresponds to one fraction of the overall diversity of Gekkota. However, and as also observed by the reviewer, the analyses implemented require a relatively large number of individuals for each species. Large samples were not always feasible due to limited sampling and damaged specimens (including digit stiffening and many broken bones) in several herpetological collections. While inferring ancestral state reconstructions of adhesive toepads in our limited sample might result on misinterpretation of trait evolution, current literature actually includes two formidable studies that already inferred pad evolution in gecko lizards using a broad taxonomic coverage (Gamble et al 2012; Russell and Gamble 2012). We agree with the reviewer that a critical examination of data limitation in our study certainly contributes for a stronger manuscript and likely points to future directions in the field, and therefore we included a new paragraph in the discussion section to address some of these aspects (lines 407-417 of the revised version). Results for some of the comparisons suggested by the reviewer were already present in the manuscript, although indirectly accessed. These were available with the confidence intervals of simulated distributions for each module shown in the Supplementary Material (Table S7). In the new paragraph included, we now highlight that the comparison of modularity between padded and padless sister-clades can be assessed in this Supplementary Table (lines 415-417 of the revised version), where we show confidence intervals of AVG differences calculated from the simulated matrices of each module per species. To synthesize, the differences between padded and padless geckos (mostly found in the functional distal module) are also consistent in comparisons between (a) *Sphaerodactylus argus* and *Gonatodes humeralis* and (b) *Gymnodactylus geckoides* and *G. amarali* and *Phyllopezus pollicaris*. In order to clarify such comparisons to the reader, we included two new figures in the Supplementary Material (Figures S2 and S3), which illustrate the distributions of AVG differences (i.e. modular degree) across species from sister-clades.

(2) The model that is called the 'functional hypothesis' can also be explained by non-functional criteria. Namely, developmental models might predict that the terminal elements of a series should differ in their variability as compared to those in the middle. This hypothesis can be traced back at least to Bateson (1913). I believe that in the evodevo literature it referred to "the property of endedness," and that it has been discussed by Alec Panchen, among others. Note that this is a critique of the model itself. Only, it should not be presented as strictly a model based on functional criteria. This has implications for how authors might interpret the plesiomorphic condition of the autopod and developmental evolution in the clade.

Answer: We agree with this comment, as in several morphological systems (such as the mammalian skull) the functional modules (e.g. nasal, oral, base, vault) are embedded within developmental modules (e.g. face and neurocranium). This observation indicates that modularity is hierarchical, and functional properties can be built within the developmental system. Indeed, both the developmental and the functional models of the gecko autopodium predict that terminal elements will vary differently from the more anterior elements, even though the number of terminal elements inserted in each hypothetical module differs between models (Figure 1 in the revised version). Following the reviewer's suggestion, we included the following sentence to highlight that the functional model can also be seen as a 'mixed development-functional model', which would account for both developmental and functional properties but differs from the strictly developmental model (lines 218-222 of the revised version): "Although we refer to this modularity hypothesis as a 'functional model', we clarify that this model might also be interpreted as a 'mixed development-functional model' because it also predicts that terminal elements will vary differently from more anterior elements, similarly to the developmental model (the number of terminal elements in each hypothetical module differs between models; see Figure 1)." We have also clarified this aspect in a previous sentence at lines 111-115 of the revised version, as follows: "To test these hypotheses, we established two hypothetical models based on different theoretical backgrounds to explain the modular pattern of the anterior autopodium, one strictly based on developmental processes, and the other incorporating developmental and functional relationships to explain autopodial modularization."

(3) Figures 1 and 2 as well as their figure legends are identical the PDF that I received. One of them is missing.

Answer: We recognize that including two figures in gray scale with very similar layouts can be confusing. Therefore, we colored the figures in order to express the developmental model and the functional model (Figure 2 in the revised version), so now the difference between the two models is more evident. We merged the two initial figures into a single one in order to respect size limitation defined by Proceedings B, and we assume that presenting these graphs side-by-side also contributes for identification of differenced between models. We thank the reviewer for noticing that, as the comment contributed for a better way of showing of our results.

Minor comments:

2.4) —Clarification: Do all species analyzed have the same phalangeal formula?

Answer: All species analyzed have the same phalangeal formula in the digits III, IV and V (4-5-3) in the *manus*, which are the digits we included in our study. Most of the species have the same phalangeal formula for all digits (2-3-4-5-3), except for *Tarentola mauritanica* and *Chondrodactylus bibronii*, which exhibit three phalanges in Digit I (3-3-4-5-3). We included the information about the consistency of phalangeal formula across the sampled species as follows (lines 139-139 of the revised version): "We measured three digits exhibiting the same phalangeal formula across all sampled species (Digit III = 4, Digit IV = 5, Digit V = 3)."

2.2) —Clarification: Does the terminal phalanx of D1 lack a claw or finger nail?

Answer: All padless species analyzed bear a claw in Digit I of the hand, but among pad-bearing geckos the morphology of Digit I is widely variable. In the hand, the claw of Digit I in padded species may be fully developed (as those sampled from the genus *Hemidactylus*), reduced (as in *Ptyodactylus hasselquistii*, *Phyllodactylus gerrhopygus* and *Phyllopezus pollicaris*) or completely absent (*Thecadactylus rapicauda*, *Tarentola mauritanica*, *Gekko gecko* and *Chondrodactylus bibronii*). We did not detect a "finger-nail" morphology in the hand digits of our sample. In case the reviewer is particularly interested about the ungual morphology of gecko digits, we suggest the article published in EvoDevo by Khannoon, Russell and Tucker (2015), where the authors review the subject of claw reduction and loss in geckos and discuss the developmental basis of such morphological variation in *Tarentola mauritanica* (Khannoon, ER; Russell, AP; Tucker, AS. 2015. *Developmental mechanisms underlying differential claw expression in the autopodia of geckos*. EvoDevo. 6:8. DOI 10.1186/s13227-015-0003-9)

2.3) —Correction, Line 36: You did not study ancestral geckos. (To me, this reads as though you are claiming that you did.)

Answer: We thank the reviewer for noticing that incongruence; in the revised version we replaced "ancestral" by "ancestrally" (i.e., lineages of geckos that evolved from padless ancestors, line 33 of the revised version).

2.4) —Clarification, Line 45: It seems incorrect to say that 'phenotypes embrace' things. This implies agency. Maybe just state that "Phenotypes have" or "can show."

Answer: We agree and therefore replaced "embrace" by "have".

2.5) —Correction, Line 56: By my recollection, the Ruvinsky and Gibson-Brown paper does not claim that individual limb segments are derived from a single module (i.e., stylopod and zeugopod are from a common module). Rather, they discuss whether the fore- and hindlimb modules are derived from a single module.

Answer: We thank the reviewer for this observation and acknowledge the oversimplification of our initial statement. We rephrased the sentence also considering the comment 3.1 from Reviewer #3 (further detailed), which now reads as follows: "Some authors have suggested that limb segments in tetrapods are derived from a single ancestral developmental module…" (lines 51-52 of the revised version).

2.6) —Clarification, Line 135: You state "anterior autopodia" were studied. Was it consistently the L or R side? Or whichever had less damage? Also, maybe state in the introduction that you studied forelimbs.

Answer: We agree that this important information should be provided in the Methods section. Therefore, we included the following sentence at Lines 133-135 of the revised version: "Measurements were preferably taken at the right side of each specimen. When the right autopodium was damaged, we assumed left-right symmetry and measured the left side." We also specified in the introduction that we studied the anterior autopodium (line 113).

2.7) —Clarification, Line 142: You state that D1 was not considered because of its distinct modularity and cite other authors. Isn't it also true that not all of the models can be applied to a digit with only two phalanges (i.e., functional model)? That seems worth pointing out.

Answer: Yes, this observation is true - it is not possible to test all the hypothetical modules of function in Digit I, but this aspect does not impose a caveat to the analyses. This is because we assessed the overall pattern of modularity of the autopodium and we did not intent to compare modularity between digits. For example, the functional module of the medial phalanx (which provides articulation to the distal phalanges during digital hyperextension in some species) is not present in the Digit V of the *manus*, which only has three phalanges – still, we were able to obtain the values of modularity for this module, present in Digits III and IV. If one aims to compare the integration of modules between digits, then we would indeed encourage testing for modules that are present in all digits.

2.8) —Clarification, Line 153: Somewhere in the paper, either methods or SI, state how sex was determined.

Answer: We included the information in the first paragraph of "Supplementary Methods S2 – Data preparation: control for sexual dimorphism", which reads as follows (Page 9 of SI): "Sexual dimorphism for *T. catalanensis* was determined by the colouration of ventral thighs, cloacal flap and abdomen: males present black-and-yellow ventral patches in these regions, which are absent in females [6]. Gecko species of the genus *Hemidactylus* are dimorphic in the cloacal and femoral regions, where only the males exhibit enlarged hemipenal swellings and femoral pores [7]. In the remaining nine gecko species, the hemipenis of male individuals is clearly visible at the x-ray images, a feature also observed in the three *Hemidactylus* species. Although this sexing method was efficient for 12 gecko species, we were not able to clearly identify the presence of hemipenis using x-rays for the remaining species. In the Supplementary List S1, we indicate the sexed specimens by an F (female) or M (male), near the voucher identification."

2.9) — Clarification, Line 179: I might have missed it. Did you ever say how you measured size? (Was it snout-vent length, total length, length of an element like the femur, etc?)

Answer: The reviewer's observation is correct. In fact, we did not include a linear measurement of body size in our sampling because we calculated size effect as the first eigenvector of the phenotypic covariance matrix for each species. The first eigenvector was considered size-related whenever all trait loadings showed the same sign (all positive or all negative), indicating that variation along this dimension was to increase or decrease all traits simultaneously. To disregard the size effect in our matrices and investigate the structure of modularity of bones in the autopodium, we removed the variation associated with the first eigenvector and obtained residual matrices without size effect. We modified the text at lines 181-187 of the revised version in order to include more details regarding these procedures and clarify the approach used for correcting size effects. Further explanation of these methods can also be found at:

- Marroig, G. & Cheverud, J. M. (2004). Did natural selection or genetic drift produce the cranial diversification of neotropical monkeys? The American Naturalist, 163(3), 417-428.
- Porto, A.; Shirai, L. T.; de Oliveira, F. B.; & Marroig, G. (2013). Size variation, growth strategies, and the evolution of modularity in the mammalian skull. Evolution, 67(11), 3305-3322. *(*cited in the list of references from the main text)*

2.10) —Clarification, Figure 1: Add units to the axis next to the phylogeny.

Answer: We included a scale of "millions years ago" (Mya) in the new version of these figures (Figure 2 in the revised version).

2.11) —Clarification, Supplementary Materials S1: I couldn't tell whether the list of specimens included those that were excluded as outliers. If outliers are still in this list, can you differentiate those from specimens included in downstream analyses, perhaps by an asterisk?

Answer: We acknowledge the observation. The original list at S1 did not include the outliers, so we decided to include the outliers to the list and identify specimens considered as outliers using an asterisk.

2.12)—Clarification: For species where sex was determined, can you report the sex of those specimens next to specimen number.

Answer: Specimens with sex determined are now indicated by an F (female) or M (male) alongside the voucher identification in the List S1 of the supplementary material.

The following comments (listed as 'suggestions') are stylistic and intended to improve the MS. These are not required changes and do not require responses by the authors. I give them merely given for the authors' consideration.

Answer: We appreciate the suggestions, which have been fully incorporated in the revised version of the manuscript. We provide answers to these comments because we consider these might help the Reviewers and Associate Editors to follow how we modified the manuscript.

2.13) —Suggestion: If there is space, Figure S1 could be moved to the main text. For this type of study, it can help to give readers an example of what the raw data look like before moving into the analyses. It's a beautiful X-ray. The image could be coupled with a visual representation of the models that you will be *testing.*

Answer: We thank the reviewer for this suggestion, which significantly contributed for a better understanding of our study. In fact, moving Figure S1 to the main text helps the readers to visualize the measurements we obtained, and also the modularity hypothesis we postulated. Combining this comment with suggestions provided by Reviewer #1 and Reviewer #3, we included a new Figure 1 to illustrate the linear measurements we obtained, as well as the two hypotheses of modularity tested.

2.14) —Suggestion: I wonder if it the title might more appropriately refer to modular evolution of phalanges, rather than digits. It is not as though you're comparing the modularity of digits within the autopod (e.g., analyzing D1 vs D2-5).

Answer: We acknowledge the suggestion and modified the title, which now reads: "Development and function explain the modular evolution of phalanges in gecko lizards".

2.15) —Suggestion: The phrase "diversification of locomotion" (line 55) feels awkward. I wouldn't refer to the 'diversification of smelling' or other behavioral categories. Maybe something like the 'diversification of locomotory modes/strategies/behaviors' would convey the same idea but with a more intuitive phrasing.

Answer: We agree with the reviewer, and rephrased the sentence as follows: "(...) the diversification of modes of locomotion" (line 50 of the revised version).

2.16) —Suggestion, line 74: I think it would clarify your meaning to state "within phalanges of a digit".

Answer: We agree and modified the sentence accordingly (line 69 of the revised version).

2.17) —Suggestion, Figure 2: You might add text to the top left of each panel ("all species" vs "geckos only") so that readers can more quickly understand what the data represent.

Answer: We thank the reviewer for this suggestion, which certainly contributes for a better understanding of our results. In the revised version we included the corresponding titles in this Figure.

2.18) —Suggestion, Table S4: You might emphasize more strongly results of testing for sexual dimorphism in the correlation coefficients in the main text. I don't know how many times such tests have been done before; it might be useful for others in the field to find this result easily.

Answer: We agree with the reviewer and therefore included one brief sentence in the beginning of the results section, in which we aim to direct the reader's attention to the tests for sexual dimorphism while we avoid missing the focus on the main results (lines 287-290 of the revised version).

2.19) —Suggestion, Line 333: The phrase "morphologically disintegrated structures" could be written more clearly. Perhaps just "reduced integration"?

Answer: We agree that the sentence was confusing. Considering also the 1.6 comment raised by Reviewer #1, we modified the sentence as follows (line 355-359 of the revised version): "However, although padded geckos exhibit a functional specialization that includes hyperextension of the distal phalanges during setal attachment and release, the padless species are the ones that show a new distal functional module, contradicting our expectation that this distal module would exhibit greater modular degree in padded geckos."

Reviewer # 3:

Comments to the Author(s)

This paper is thoughtful and interesting. It uses multivariate quantitative analysis to show how processes of embryonic development and functional integration interact in the evolution of tetrapod limbs. Specifically, the authors use tools for studying morphological integration and modularity to demonstrate that the evolution of a novel locomotor feature (toe pads in geckos) is associated with changes to the modular structure of the bones of the feet, causing it to depart from the pattern of inter-bone correlations found in lizards without toe pads. In other words, the authors show that a novel morphological module evolved in association with the evolution of the new functional structures. Interestingly, the authors argue that the "new" module can be viewed as the co-option of the structure of integration found in the non-squamate diapsid ancestors of lizards. The paper is nearly publishable as it is. The authors have compiled an appropriate new data set to address the question, namely measurements of foot bones from nearly 500 individual lizards representing 20 extant species. These samples are about as good as one can get for a large comparative phenotypic analysis like this. Even though larger samples of each species would be ideal for estimating covariance matrices, the authors have adopted statistical bootstrapping and Monte Carlo methods that place appropriate standard errors on their statistics and prevent them from over-interpreting patterns. Many analytical strategies are available to measure and compare the strength of phenotypic modularity. Like all of them, the ones the authors chose (correlation matrices, random skewers, and Mantel tests) have their strengths and weaknesses, but they are more than adequate for the data in this paper. Any of the known biases in these particular methods ought not be a problem for correlation *matrices of this size (15 x 15). The authors also performed appropriate statistical power tests (described in the Supplemental file). While I have two specific questions about methods below, I am convinced that the authors' analyses are both appropriate and convincingly support their conclusions. The paper will be of broad interest, thus making it suitable for Proceedings B. It will, of course, be of interest to the very large community that studies phenotypic integration and modularity, but it will also be of interest to functional morphologists, developmental biologists, and large "integrative biology" audience that is interested in the adaptive radiation of Anolis lizards.*

Answer: We thank the reviewer for the kind words and thoughtful comments. In fact, we found evidence for a new distal module in padless geckos, instead of in padded geckos as we originally hypothesized. We suggest that friction and subdigital structures in lineages with padless ancestors may promote the higher integration of the distal module, whereas in padded species the origin of toepads seems to disrupt the higher integration of distal elements. The reviewer provided very useful suggestions that were fully incorporated in the revised version of the manuscript, and raised questions that, after clarified, improved clarity and precision of the text.

Specific comments:

3.1) Lines 56-57: I would not immediately have assumed that limb segments in tetrapods belong to the same developmental module, but I see that the paper you cite in this line does indeed make that assumption and in its own context I understand why. Nevertheless, I would have assumed something quite different. Maybe you could tweak the wording to, "Some have argued that limb segments in tetrapods are derived from a single ancestral developmental module…".

Answer: We appreciate the suggestion, also addressed by Reviewer #2 (comment 2.5). We rephrased the sentence as follows: "Some authors have suggested that limb segments in tetrapods are derived from a single ancestral developmental module…" (lines 51-52 of the revised version).

3.2) Line 153: "We performed linear models" > "We applied linear models". (the former is grammatically correct, but it is idiomatically unusual, even though I am struggling to come up with a logical reason why).

Answer: We appreciate the reviewer's suggestion, and replaced "performed" by "applied" (line 152 of the revised version).

3.3) Lines 211-225: Regarding the test between residual correlation matrix and theoretical matrix, do you have any negative correlations in the residual matrix? If so, did you test using the absolute value?

(otherwise a strong negative correlation in the residual matrix will be correlated with the lowest value in theoretical matrix, i.e., 0). As an aside, I have noticed that in geometric morphometric data sets with semilandmarks, odd things happen with the theoretical correlation matrix approach, including a bias toward significance because the large number of variables in the covariance matrix equates to larger sample size in the Mantel test, as well as non-random blocks of positive, negative, and near-zero correlations that relate to the x and y directions of variation at individual semilandmarks. I don't think your data set will be prone to either effect because the number of variables is comparatively small and because all your variables are the same biological type (proximo-distal lengths of bones).

Answer: Yes, in the residual matrices correlations vary from -0.5 to 0.5, although most correlation values are around zero (see figure below). We interpret negative values as correlations that previously were positive only due to growth effects (global integration factors), having no contribution from local developmental or functional processes (see lines 179-181 of the revised version). Regarding the correlations that remained positive after removing the variation associated with allometric growth, we interpret that local processes explain the positive associations between traits. Consequently, we consider inappropriate to transform negative correlations into absolute values because effects of removing global integration factors associated with allometric growth would be erased in this approach. In the Mantel test, correlations close to zero (and not the strongly negative ones) likely increase the matrix correlation of empirical and theoretical matrices. Very negative correlations may actually be allocated to values of 1.0 in the theoretical matrix if that correlation is hypothesized to belong to a module, or may assume zero (0) values in the theoretical matrix when that correlation is hypothesized to not belong to a module. In fact, some average correlations within-modules (AVG+) are actually negative for some specific modules and species in figures 2A and 2B (revised version). The highest degrees of modularity were observed when AVG- was negative and AVG+ was positive after removing the variation associated with allometric growth.

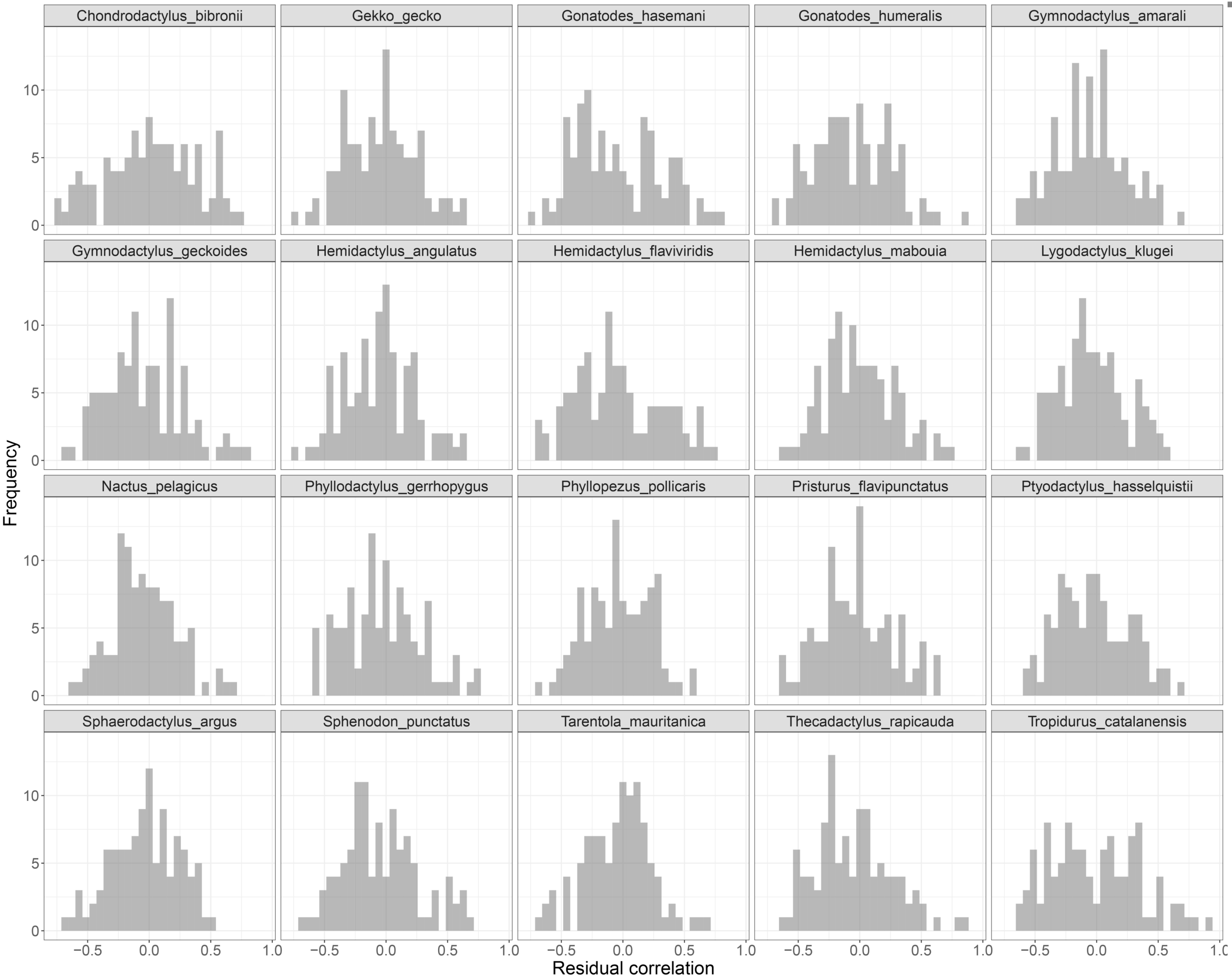

Figure: frequency of correlation values from residual matrices of each species.

To clarify this reasoning in the manuscript, we included the following sentences in the methods section (lines 19-196 of the revised version): "By removing variation associated with allometric size, several correlations previously positive in the original matrices became negative. Therefore, many positive associations between traits were originally related to growth effects. For the correlations that remained positive in residual matrices, we infer that local developmental and/or functional processes might be strong enough to contribute to the positive association between traits despite global integrating effects of allometric growth". We have also modified the results section as follows (lines 314-318 of the revised version): "The tip module exhibited a high modular signal (mean AVG difference = 0.52) for almost all species except for *Sphenodon punctatus*. This is the case because all average correlations between-modules (AVG-) are slightly negative whereas all average correlations within the tip module are highly positive, indicating that local processes related tip development are strong".

3.4) Lines 231-240: I don't completely understand the resampling strategy or why it is employed. What precisely are you sampling from the multinormal distribution, a covariance matrix or a vector of trait values? If I assume the latter, am I correct that you are essentially sampling N sets of trait values, where N is the original sample size, from a distribution that has the same correlation structure at the original sample? Since each will be slightly different within the probability density of the multinormal, the N set of trait values represents sampling error on correlation matrix and AVG values calculated from them. Why do you use this strategy here instead of the bootstrapping approach you used in lines 168-172? The latter has the advantage of not assuming multivariate normal distribution of the traits.

Answer: The reviewer is correct about the resampling procedure using a multivariate normal distribution. We sampled 1,000 sets of traits for each species, using the original sample size from a multivariate normal distribution that has the same correlation structure of the empirical matrices. We chose this approach because we aimed to incorporate sampling error in the degree of modularity, and we assume this method is suitable for this purpose by confirming that the matrices are reliable (i.e., have high repeatability using the bootstrap approach). Although deviations might occur due to low sample sizes within species, we expect that overall linear distances follow a multivariate normal distribution. We appreciate the concern raised by the reviewer, and we evaluated if results would change by adding the sampling error to the degree of modularity using bootstrapping (resampling 1,000 times the residuals of the linear models, with replacement, and using the original sample size). We summarize the results of AVG diff calculated using both Monte Carlo and Bootstrap resampling methods in the Table S7 of Supplementary Material and we detect that 72 out of the 160 modules exhibit different results depending on the method used. Specifically, all the modules that differ in significance depending on the method employed exhibit significant values when using Monte Carlo resampling (i.e., 95% CI range of positive values that do not contain 0) but are not significant when using Bootstrap. We therefore consider that the modules that are significant using both methods are more robust and with higher modular degree than those detected using only Monte Carlo resampling. We thank the reviewer for the suggestion and included this information in the methods (lines 251-255) and in the results (lines 303-305).

3.5) lines 290-291: interesting that Sphenodon is the odd one out, given that it is also phylogenetically the most distant. Do you think this is because the pattern of integration changed in squamates, or could it somehow be an artifact of the PGLS and long branch? (to rephrase the question, does this result depend on the PGLS correction?)

Answer: We agree with the reviewer this is indeed a very intriguing result. The results described in the section 'Tests of Modularity' actually derive from analyses that do not consider phylogenetic effects. The modular signal for the tip module, as well as for all the other modules tested, was individually calculated for each species through a matrix comparison, using the residual correlation matrix of each species and the theoretical modularity matrices (explained in the section 'Inferences on modularity', lines 198-258 of the revised version). Therefore, the relatively weaker modular degree in the tip module of *Sphenodon punctatus*, illustrated in Figure 2 (originally Figure 1 in the previous version submitted), is not a long-branch artifact derived from the PGLS, which was performed only after calculating the signal of each module for all species separately. The PGLS analyses were performed to investigate possible differences in the modularity signal between padded and padless species (results illustrated in Figure 3 of the revised version). We understand that representing the topology together with the average correlation plots in Figure 2 (revised version) might cause the impression that those results derive from phylogenetic comparative analyses. Inclusion of a topology in this figure aims to provide a graphical visualization of the phylogenetic relationships among padded and padless species. In order to clarify this eventual misunderstanding, we included the following sentence in the results section (lines 300-301 of the revised version): "Prior to the phylogenetic comparative analyses, we detected that the degree of modularity was variable among the modules tested, as well as among species (Table S7)".

The reviewer is correct to point out that the signal of modularity in the tips of *S. punctatus* is the most different among all species we studied, and to discuss that this is also the most phylogenetically distant lineage in our sampling. Given that the *Sphenodon* pattern does not correspond to an artifact from the comparative analysis, the hypothesis suggested by the reviewer – that the pattern of integration might have changed in Squamata – seems very interesting and deserves further investigation. We consider, however, that a different taxonomic coverage within Squamata is necessary to appropriately address this topic and, thus, provide more elusive conclusions. We hope our results stimulate future studies on the evolution of the modularity in the autopodia of Squamata.

3.6) lines 293-295: The second part of this sentence is confusing on the first reading because it starts with statement that the five geckos have negative AVG+ values but the value reported at the end of the sentence is positive. Maybe reword the entire phrase as "… five geckos had low values for within-module integration (AVG+) among the metapodials, indicating no evidence that they form their own module in these taxa (mean AVG difference = 0.21)."

Answer: We thank the reviewer for the suggestion, as indeed rephrasing the sentence clarified our results. Accordingly, we modified the sentence at lines 315-318 of the revised version.

3.7) Lines 308-311: Rewording the first two sentences would make this paragraph bolder: "We identified a divergence in modularity between padless and padded geckos, the latter evolving a new functional module involving the distal phalanges. This novel module was detected in both PGLS analyses (...) and was more pronounced after outgroup removal."

Answer: We appreciate the suggestion and modified the sentences accordingly, which now reads as follows: "We identified divergence in modularity between padded and padless species, the later involving a new functional module comprising the distal phalanges. This novel module was detected in geckos without toepads in both PGLS analyses, and divergence was more pronounced after removing the outgroups (*Sphenodon punctatus* and *Tropidurus catalanensis*; Figure 3, Tables S7 and S8)" (lines 335-339 of the revised version).

3.8) Figures 1 & 2: Can you add something that helps the reader orient themselves to the anatomy? That could be stating in the caption that the illustration shows a right foot, adding digit numbers to one or more of the anatomical drawings, or labeling radius and ulna (or tibia and fibula, as appropriate).

Answer: We considered the suggestions provided by the three reviewers and included a new Figure 1 in order to provide information on the anatomical orientation of digits and long bones (autopodium x-rays), indicating the linear measurements obtained, and also the two hypotheses of modularity tested (each module is now represented by a different colour).

Appendix B

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE FILOSOFIA, CIÊNCIAS E LETRAS DE RIBEIRÃO PRETO. DEPARTAMENTO DE BIOLOGIA.

AV. BANDEIRANTES, 3900. BAIRRO MONTE ALEGRE. RIBEIRÃO PRETO, SP. CEP 14040-901.

FAX: (16) 3602-4886

Editorial Office of Proceedings of the Royal Society B
Dr. John Hutchinson, Editor

Response to Referee: Revision of the manuscript ID RSPB-2021-2300

Dear Dr. John Hutchinson

We are pleased to submit the final revised version of manuscript RSPB-2021-2300 "*Development and function explain the modular evolution of phalanges in gecko lizards*", authored by PS Rothier, MN Simon, G Marroig, A Herrel & T Kohlsdorf, recently accepted for publication at the Proceedings of the Royal Society B. We acknowledge the final concerns raised by #Reviewer 2 and incorporated all suggestions in this revised version of the manuscript. In particular, we tempered our discussion on the results obtained to formally recognize the implications of restricted datasets to the analyses performed. I am uploading a response letter to the reviewer's suggestions, and also a clean copy of the revised version of the manuscript and a file with all changes indicated in red.

I would like to express my sincere gratitude for the careful evaluation of our study and for the opportunity given to contribute with the Proceedings of the Royal Society B.

Sincerely,

Tiana Kohlsdorf

DETAILED ANSWERS TO REVIEWER #2

Comments to the Author(s)

In this study, the authors analyze the variational modularity of phalanges in geckos. Testing two hypothetical models, they describe the following major results: (1) Despite functional specialization, geckos exhibit a pattern of modularity that was previously described by Kavanagh et al (2013) and is plesiomorphic to tetrapods. (2) Non-toepad bearing geckos show a derived distal phalangeal module.

Major comment:

This second result is based upon phylogenetic regression (PGLS, shown in Fig. 3). My concern is that applying phylogenetic methods to a so highly restricted taxonomic sample may produce misleading results.

According to the authors (line 414, citing ref. 17), all of the geckos without toepads are inferred to never have had this character (i.e., toe pads have not been lost in any of these species). Therefore, the 10 genera with toe pads that they analyze represent at least 8 independent origins of the structure. It is the mismatch between this fact and what would be recovered by parsimony reconstructions about gains and losses for their sample (see tree in Fig. 2) that gives me pause. To be clear, it is not my intention to contest the evolutionary scenario of no loss. Rather, I think that when taxonomic sampling doesn't capture what we understand to be the major evolutionary history of a trait, it raises obvious questions about the robustness of results from phylogenetic methods. That is why in my original review I suggested a non-phylogenetic method that might allow the authors to get around sampling limitations. Specifically, pairwise comparisons of sister clades with and without toe pads provide another means of testing for consistent differences between the groups. The authors now present this analysis in the supplementary materials. Pairwise comparisons (Fig. S3 D) to not clearly support the major conclusion that padless taxa have a derived condition of higher degree of modularity in the distal phalanges: Sphaerodactylus is overlapping the padless species; Phyllopezus has a lower value than the padless species; Gekko has a higher value than the padless species. How do the authors reconcile this observation? I would encourage the authors to *consider whether there other ways to explore the robustness of their conclusion of a derived distal module in padless taxa. The paper will be improved if they can confront these questions analytically or via argumentation in the text.*

Answer: We recognize that an extended discussion confronting these results might improve the article by enhancing criticism and transparency of our interpretations. We followed the reviewer suggestion to incorporate a critical argumentation to the main text that addresses the consequences of limited sampling for interpretations based on gains and losses of adhesive toe pads in geckos. Specifically, in the discussion section we attenuated the assumption of higher integrated distal modules in padless geckos as a novelty in Gekkota, and highlighted the conflicting results from *Sphaerodactylus argus* and *Gonatodes hasemani*, recognizing the limitation of our taxonomic sample for some generalizations (Lines 409-410 and 417-421 of the revised version). We thank the substantial attention given by the reviewer to improve our manuscript.

Minor comments:

Line 40: Would it be more correct to say "Organisms have variable…"?

Answer: We modified the sentence accordingly (Line 40 of the revised version).

Line 55: Did you mean "which represent"?

Answer: In this sentence, we aimed to state that fore and hind limbs of quadrupedal animals usually perform similar functions. In order to clarify the statement, we rephrased the sentence as follows: “in which fore and hind limbs perform similar functions” (Line 55 of the revised version).

Line 75: The line word "morphological" feels unnecessary.

Answer: We deleted the word ‘morphological’.

Line 221: Did you mean "proximal" instead of "anterior"?

Answer: We acknowledge the observation, and replaced “anterior” by “proximal”.

Line 279: To confirm, does Figure 2 reflect the resolved topology? It looks like a polytomy.

Answer: We confirm that Figure 2 is the resolved topology. The resemblance of *Hemidactylus* relationships with a polytomy derives from the fast clade divergence in relation to the time scale in millions of years (Carranza and Arnold, 2006, *Molecular Phylogenetics and Evolution,* 38: 531-545).

Line 305: You may want to make a new paragraph with the sentence that begins on this line.

Answer: We agree with the suggestion and modified the document accordingly.

Line 345: As a suggestion, don't just say "not significant." Give the p value. I believe it's 0.06, which makes your observation seem more reasonable.

Answer: We thank the observation and included the p-value alongside the statement (Line 342 of the revised version).

Figure S2 and Table S8: Did you mean "metacarpal" instead of "metapodial"? That is how you've labeled the module in Fig. 2 and 3.

Answer: We used "metacarpal" as a synonym for the hand "metapodials". However, we acknowledge that standardizing the nomenclature in all tables and figures is important to grant precision. We identified the same ambiguity when naming other modules, so we reviewed and modified all figures following the labels used in Table S8.

Figure S3: Functional is misspelled.

Answer: We acknowledge the observation and corrected the word.