THE ROYAL SOCIETY PUBLISHING

PROCEEDINGS B

Evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards

Gen Morinaga and Philip J. Bergmann

Article citation details
Proc. R. Soc. B **287**: 20200192.
http://dx.doi.org/10.1098/rspb.2020.0192

Review timeline

Original submission:	6 August 2019
1st revised submission:	28 January 2020
2nd revised submission:	21 February 2020
Final acceptance:	25 February 2020

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

Review History

RSPB-2019-1841.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?
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
I have now read the paper by Morinaga and Bergmann on 'the evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards'. I enjoyed the paper and think the authors have a nice data set. I do find that they have a tendency to oversell their data a bit by forcing it into the 'origin of snakes hypothesis'. In itself it is already an interesting question and I don't think it needs that. I do have a number of comments and suggestions that I would like to see addressed:

line 15: it is important to not that there need not be one all encompassing reason for limb loss. Different taxa may have lost limbs for different reasons (very likely even) and the two hypotheses are not mutually exclusive. Here you present it as a real dichotomy.

line 47: the anatomy of the snake eye

line 57: this is more of a conceptual issue, but is sand swimming really burrowing ? I would argue not really as the constraints and mechanics are very different for sand swimming (sand behaving almost like a fluid) versus burrowing in a solid, compact substrate. I would like to see this addressed and discussed.

Line 93: I was really confused as how you were able to film burrowing with a regular camera ... by definition burrowing and even moving in leaf litter happens in the substrate and as such the animals are not visible. I think it is important to make very clear that what you studied is initial soil penetration from the surface. Even when this is what you appear to have studied (based on the supplementary methods) I still wonder how much you are able to quantify as you will loose the point you track right away. I would really like to see more information on the filming protocol to have an ideal about what it is you really studied.

line 99: this is a uni-axial force transducer. Especially in skinks forces are intrinsically 3D with much of the forces being not in the Z direction (see the paper by Vanhooydonck in BJLS). As such your force measures are likely underestimates. This in itself is not worrisome but if species different in burrowing style then your measures will be biased by that and will not represent true performance. I would definitely like to see this discussed in the paper.

line 169: is a leaf litter mat really a structurally complex habitat? I really wonder about this as leaves are soft and compressible and as such the animal does not really have to move through

tight interstitial spaces but can and will simply create space by compacting the soft dead leaves. This makes the system not so great for testing your initial question.

line 180: there are many eels that burrow head-first into the substrate; see papers by Choi and colleagues (J. Exp. Biol.) and De Schepper and colleagues, for example.

line 192: this is where you need to make a difference between burrowing and sand swimming. In sand swimming the tail can effectively be used to produce propulsive forces as in in any anguilliform swimmer. Consequently long tails are not a disadvantage for sand swimming.

The figures of my review copy weren't great but that's probably just a conversion issue. I liked some of the figures in the supplementary as well (S1 & S2) and would like to have seen them in the main manuscript.

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?
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.
Yes

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?
No

Do you have any ethical concerns with this paper?
No

Comments to the Author
GENERAL COMMENTS:
I believe the manuscript by Morinaga and Bergmann addresses very general issues regarding evolution, structure/function relationships, evolutionary and functional trade-offs, and niche partitioning that have great potential to appeal to a very wide readership for the Proceedings of the Royal Society. I wholeheartedly agree with the authors that a lack of extant intermediate forms has hindered our ability to understand the adaptive benefits involved in the early stages of some profound changes in animal body plans. However, the study system of the authors (lizards with varying degrees of limb reduction and elongation) has a wealth of intermediate forms that the authors have put to very good use. Hence, in my opinion, overall the authors address very interesting questions and have made good choices of study taxa and experimental methods required to address the primary hypotheses that they pose.

Despite my overall enthusiasm for this interesting and potentially important manuscript, I believe that it could be improved greatly by addressing the specific concerns that I detail below. I appreciate that the authors have an abundance of data, and they must deal with significant constraints on manuscript length. Hence, the supplemental material is vitally important for reader to fully understand this work, and I read it in detail. I am also aware of the fact that some of the species studied by authors are quite difficult to obtain; hence, making some allowances for that also seems appropriate. Nonetheless, my approach below is to list all concerns that come to mind and largely leave it to the authors to sort out what is and is not practical to include or revise. Some items of greater concern are indicated by **, and most importantly they include a critical lack of a more logical quantitative measure of elongation (key aspect of snake-like morphology) such as the some relation between mass, cross sectional area or width to total length (rather than residual values based on head length). Additionally, many potential readers in a general audience will not be familiar with these study taxa. Consequently, some combination of images of these lizards along with some anatomical variables that are easier to interpret than the rather abstract PCAs, could go a long way for painting a more concrete picture for the readers regarding the relevant structural variation in this system.

Despite some of my concerns, I believe that the authors have performed experiments sufficient to provide very compelling insights for how: 1) a more elongate and snake-like body plan facilitates sub-surface locomotion, 2) no tradeoffs occur between subsurface and surface locomotion despite a reasonable a-priori expectation for this, and 3) the data for these lizards with a continuum of variation in morphology and locomotion strongly support the suggestion that the ancestors of snakes were elongate as a result of adaptation for subsurface locomotion. I just believe that an extensively revised manuscript with some rethinking and revision of methods and illustrative information can be even more compelling and accessible to an even wider audience. I found this work and the questions it addressed extremely interesting and insightful.

SPECIFIC COMMENTS:
1) I had difficulty following the second part of the last sentence in the abstract. The authors did not find performance trade-offs between surface and subsurface locomotion (lines 143-144). Consequently, I was confused at how the authors could then suggest here that habitat partitioning mitigates the effects of performance trade-offs.

2) Background.
a) (line 33) Perhaps “cluttered” be a better choice of words than “structurally complex”. Consider a flat sand surface versus one that has ripples from the wind. The latter is arguably more complex that the former, but this sort of increased complexity does not seem to be likely to have much consequence for impeding the locomotion of limbed animals.
b) (line 39) Perhaps substitute “several clades of ectothermic tetrapod vertebrates” (or some similar wording) for “most major vertebrate clades”.
c) (lines 46-48) To me it would make more sense to combine the information on visual pigments and eye anatomy into one issue regarding the near loss of a functional eye as subsequent re-

elaboration of an eye in snakes arising from a burrowing ancestor.
d) (lines 51-54) This has the potential to cause confusion. After first referring to 2 ecomorphs, the authors proceed to state that the genus Lerista contains 12 morphs. To me the fact that "Lerista do not fit neatly into either ecomorph" (lines 57-58) diminishes the likelihood that reduced limbed lizards actually have 2 distinct ecomorphs rather than forming a continuum of variation
e) (last paragraph) Perhaps the suggestion that more lizard-like lizards are faster than snake-like lizards could be strengthened by including the fact that a wealth of studies have found very rapid speeds of lizards using limbed locomotion (very commonly more than 10 body lengths per second) compared to the surface locomotion of snakes (very rarely > 3 body lengths per second).

3) Methods.
a**) To me it seemed that the authors lacked a good metric of overall size. I realize that they cite some previous studies of elongate lizards that have used head length for this purpose (supplement p. 3), but based on the biomechanics of this system focused on locomotion, this did not seem like a good choice to me. Either mass or total length would be much better metric of overall size. I realize that the length of the tail compared to snout-vent length can vary substantially among species of reduced-limbed lizards, and so I also understand the desire to include both SVL and tail length in the PCA. Since the authors want to calculate some residuals from a regression using a variable related to overall size, total length would have a drawback because it is the sum of SVL and tail length rather than being independent of these two quantities. By contrast, using mass as an independent variable to calculate residual values of other variables would not have such a drawback. Furthermore, the residuals of lengths predicted from mass to would get right at the heart of what is a more elongate (snake-like) body plan (greater residual value = greater elongation) much better than residuals predicted from head length. In my opinion the authors' analyses are presently deficient because of this lack of some more logical quantitative variable that relates quantities such as mass, maximal cross-sectional area, or maximal width to total length and hence better reflects elongation.
b) (lines 106-115) I favor clarifying the distance over which velocity was calculated. Here authors indicated that they fit a cubic spline to a plot of displacement versus time and then calculated velocity as the slope of that spine. I believe that this was done between successive frames, which were less than 5 ms apart (line 97). Their field of view was at most 518 pixels for ~ a 660 mm long racetrack (~1.3 mm per pixel). Based on Fig. 2, many of the maximal speeds were only approximately 100 mm/s, which would correspond less than 0. 4 mm between successive frames, which is less than 1 pixel. Consequently, I suspect that the displacement between successive images was often less than or equal to the digitizing error, and hence calculating quantities frame by frame does not seem like a good choice. The authors subsequently state that they calculated an average velocity for frame by frame values (line 111), but they did not specify over how many images or over what distance. If they are taking an average of the spline slopes for several successive time intervals, they just as easily could have taken the difference between first and last displacement and divided by the time interval for which the animal was moving, and the result would be simpler and perhaps even more accurate. Perhaps I am just misunderstanding some methods that could be explained more clearly? Maximal velocities decease as the distance over which they are measured increases. Consequently, one good method for quantifying maximal velocity would be to use a fixed distance for calculating an average maximal velocity. Unfortunately, neither the supplement nor the main manuscript provide the ranges of lengths of their study species. If there is a large range in size, then another option would be to calculate average velocities per cycle and then select the fastest cycle so that the relative distance traveled by different sized animals would be reasonably similar.
c**) In part, in light of some of the issues above, I strongly favor including in the supplement a tabular summary with, at a minimum, means and ranges of SVL, tail lengths and mass for each species. I also believe that some outline figures or good photographs of the study taxa would be extremely helpful and much more useful than the current Fig. 1.
d) (lines 113-115) The biomechanical relevance of a lateral head angle for penetration was not especially clear to me. Wouldn't the amount of ventral flexion would be as important or perhaps even more relevant for understanding how the lizards penetrated the surface of the sand?
e) It was not clear to me if the velocities of burrowing included both penetration of the surface

and sub-surface locomotion or just one or the other. If the speed of penetration was analyzed, am I correcting in assuming that it was the time taken until the entire animal was buried? The authors should explicitly state what the case is.
f) For subsurface locomotion, am I correct in assuming that velocity was estimated by tracking the disturbance of the sand rather than a paint mark on the animals? If so this should be clarified as the currently the methods imply that only paint marks were used to track the displacement of the lizards
g) (lines 119-121) Are the authors trying to convey that they only used one maximal speed per individual per substrate, and then they took the mean value (where N = number of individuals for each species. If so, this could be stated more clearly.
h**) I believe leg length could be dealt more effectively. One key benchmark is whether or not a leg is long enough to lift the belly of an animal off the ground, but this is not at all obvious for values of absolute leg length. Although absolute leg lengths were included in the PCA, the supplement lacked any summaries of quantitative data for this important aspect of morphology. I believe a maximally useful quantitative measure of this would be the length of the leg divided by the vertical distance from the ventral surface of the body to the point of attachment of the limb. The current lack of information along these lines obscures the simple but very important issue of whether or not it is even possible for the limbs to be used in a conventional manner rather than just sticking out to the sides of the animal. The authors mentioned that in some circumstances the lizards used their legs when they move, but the wording was not precise enough to know whether or not the legs wiggled ineffectually or if they were contributing to propulsion. Once again, understanding such an issue would be enhanced greatly, in part, by good images of the study taxa.
h) (lines 125-126) I realize it is common practice, but I do not favor asserting that a particular PC is one certain thing such as "overall size" or how "snake-like" an animal is. Each PC is simply a mathematical construct for which the first PC explains the maximal amount of variance of the cloud of points, and by their very construction PCs are not one thing but rather a composite measure of many things. It is perfectly reasonable mention factors that load high for a particular PC, but that is quite different from asserting that one PC is "shape" and another PC is "size". For example, see Gould (1981, The Mismeasure of Man) for one line of argument against reification of factors (that admittedly has not been without controversy).
i**) Except in very special circumstances, I do not favor performing statistical analysis of multiple species with only one individual per species, but the authors have 3 species for which this is the case. This is potentially problematic for a couple of reasons. First, measurements of performance are often highly variable as a result of being strongly influenced by the behavioral motivation of the test subjects. Consequently, a single individual could have an extreme value of performance as a result of behavioral variation rather than variation in its morphology and physiology. Second, the within species variability of morphology in the study taxa is huge, and large variation within species is best dealt with by having sample sizes larger than 1 per species. Note that in Fig. S1 the values of PC2 range from approximately -2.5 to 3, and the range of values for just one (medium green) species (but hard to tell some shades of colors apart) range from approximately -2.5 to 2, which is roughly 80% of the range of PC2 for all species! Imagine potentially what could have happened if just one of these extreme points was sampled for this particular species.
j) The large ranges in PC scores among some of the individuals belonging to the same species also made me wonder if the authors had restricted their analysis to only adults to reduce the variability of size within each of their study species. They should explicitly address the issue and provide compelling reasons for their sampling strategy.

4) Results.
a) (line 139-143) The authors performed several statistical tests, but made no effort to account for this by adjusting P-values. Clearly, this is a thorny issue with no easy answer as using extremely small P-values arising from adjusting for multiple comparisons just introduces a different sort of error. However, three important results highlighted as significant (Tables S3 and S5) had P-values ranging from 0.026 to 0.046. Hence, some of these conclusions might be on rather shaky ground.
b) (line 144) As indicated above, I do not favor using head size and body size interchangeably.

c**) Tables S3, S4, S5 S6 included many non-significant independent variables in (forced-fit) multiple regression regressions. Including these non-significant independent variables can lead to spurious conclusions regarding the significance level of the other independent variable in the multiple regression. Hence, I favor redoing the all of the multiple regressions and only reporting the models in which all of coefficients of the independent variables are statistically significant.

5) Discussion.
a**) The authors repeatedly acknowledge the importance of head shape for affecting burrowing performance. Consequently, I was surprised by the absence of any variable quantifying head shape for the study taxa, especially since the biomechanical relevance of this seems much more obvious than that of the lateral angle of the head, which was quantified.
b) (lines 186-196) The authors mention that the relative length of the tail varies substantially among elongate lizard taxa. A substantial difference between many elongate lizards and snakes is also the greater relative length tail in lizards. Furthermore, most burrowing snakes have quite short and blunt tails relative to their terrestrial counterparts. However, despite having snout vent length and tail length, no data were reported for relative tail length.
c) By the way, do any of these study taxa have tail autotomy? If so, did all of the individuals used in the study have original rather than regenerated tails? This should be stated explicitly. I also strongly favor including only intact individuals.

6) (Throughout the main manuscript and supplement) Given that some taxa did not use their limbs for surface locomotion, "running" seems to be a poor choice of terminology compared to "surface locomotion".

7) Supplement.
a**) I strongly favor including a conventional rooted cladogram of the study taxa as a figure in the supplement.
b) A better color scheme should be possible for Fig S1. I was not confident I could distinguish between the two of the shades green (L. kendricki and L.connivens) and two of the shades of blue (L. macropishopus and L. dorsalis).
c) The authors should explicitly state if a value of zero was used for limb length when limbs were absent.
d) I could not find anywhere close to an N =13 for the light blue symbols and dark blue squares in Fig. S3. Perhaps the author are trying to cram too much information in each panel of this figure. Since it is in the supplement anyway, I favor clarity over small size.
e) What is the definition of a "stump" limb?
f) (Fig. S3) It seemed somewhat odd to me to include the seven species that that lacked front limbs (Table S1) in the analysis of the proportion of individuals that used the fronts limbs since front limb use is not possible in these species.
g) The heading for Table S5 refers to "penetration during burrowing". However, two entries in the body of this table use the term "sand-swimming". It is vitally important that the authors use consistent terminology for the different types of locomotion, which I believe include the initial penetration of loose substrates, sub-surface locomotion and surface locomotion.
h) A few supplementary videos might be nice for illustrating the presence or absence of limb use or variation in how the limbs are used.
i**) The multiple regression analyses in Tables S3-7 included both pPC1 and head length as independent variables. However, Table S2 indicates that head length was a factor in the PCA. Consequently, it seemed inappropriate to perform a multiple regression analysis that in some sense used the same variable (head length) twice.

Decision letter (RSPB-2019-1841.R0)

19-Sep-2019

Dear Dr Morinaga:

I am writing to inform you that your manuscript RSPB-2019-1841 entitled "The evolution of fossorial locomotion in the transition from tetrapod to snake-like in 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. The 2 reviewers and Associate Editor all agree that the paper has potential but is marginal in its current form. Substantial improvements will be required; not simply stylistic ones.

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 'response 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.

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,
Professor John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor
Board Member: 1
Comments to Author:
Thank you for the opportunity to review this paper. Herein the authors attempt to test existing but competing hypotheses for the evolution of body elongation and limb loss in vertebrates: that these traits evolved as adaptations for fossorial locomotion, or as a means to better locomote through "structurally complex habitats." To test these competing hypotheses, the authors collect experimental data on lizards locomoting through sand and leaf litter. Based on specific metrics (e.g. velocity) they conclude that their experiments support the hypothesis that elongation and limb loss is better explained as an adaptation for fossoriality.

I personally enjoyed reading the paper and believe questions driving it are of wide biological

significance. Both reviewers were also enthusiastic about the what the paper attempts to achieve, but do raise some important and quite fundamental concerns about the methods and the depth of interpretation (or extrapolation) from the data collected. Notable issues that need to be confronted one way or the other include (but are not limited to) justification for sand and leaf litter as a appropriate substrate choices, how burrowing locomotion is being measured given submersion in sand, body size quantification and a number of issues with the statistics raised by reviewer 2. If the authors feel they can address these issues constructively then I would encourage them to do so.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
I have now read the paper by Morinaga and Bergmann on 'the evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards'. I enjoyed the paper and think the authors have a nice data set. I do find that they have a tendency to oversell their data a bit by forcing it into the 'origin of snakes hypothesis'. In itself it is already an interesting question and I don't think it needs that. I do have a number of comments and suggestions that I would like to see addressed:

lien 15: it is important to not that there need not be one all encompassing reason for limb loss. Different taxa may have lost limbs for different reasons (very likely even) and the two hypotheses are not mutually exclusive. Here you present it as a real dichotomy.

line 47: the anatomy of the snake eye

line 57: this is more of a conceptual issue, but is sand swimming really burrowing ? I would argue not really as the constraints and mechanics are very different for sand swimming (sand behaving almost like a fluid) versus burrowing in a solid, compact substrate. I would like to see this addressed and discussed.

Line 93: I was really confused as how you were able to film burrowing with a regular camera ... by definition burrowing and even moving in leaf litter happens in the substrate and as such the animals are not visible. I think it is important to make very clear that what you studied is initial soil penetration from the surface. Even when this is what you appear to have studied (based on the supplementary methods) I still wonder how much you are able to quantify as you will loose the point you track right away. I would really like to see more information on the filming protocol to have an ideal about what it is you really studied.

line 99: this is a uni-axial force transducer. Especially in skinks forces are intrinsically 3D with much of the forces being not in the Z direction (see the paper by Vanhooydonck in BJLS). As such your force measures are likely underestimates. This in itself is not worrisome but if species different in burrowing style then your measures will be biased by that and will not represent true performance. I would definitely like to see this discussed in the paper.

line 169: is a leaf litter mat really a structurally complex habitat? I really wonder about this as leaves are soft and compressible and as such the animal does not really have to move through tight interstitial spaces but can and will simply create space by compacting the soft dead leaves. This makes the system not so great for testing your initial question.

line 180: there are many eels that burrow head-first into the substrate; see papers by Choi and colleagues (J. Exp. Biol.) and De Schepper and colleagues, for example.

line 192: this is where you need to make a difference between burrowing and sand swimming. In

sand swimming the tail can effectively be used to produce propulsive forces as in in any anguilliform swimmer. Consequently long tails are not a disadvantage for sand swimming.

The figures of my review copy weren't great but that's probably just a conversion issue. I liked some of the figures in the supplementary as well (S1 & S2) and would like to have seen them in the main manuscript.

Referee: 2

Comments to the Author(s)
GENERAL COMMENTS:
I believe the manuscript by Morinaga and Bergmann addresses very general issues regarding evolution, structure/function relationships, evolutionary and functional trade-offs, and niche partitioning that have great potential to appeal to a very wide readership for the Proceedings of the Royal Society. I wholeheartedly agree with the authors that a lack of extant intermediate forms has hindered our ability to understand the adaptive benefits involved in the early stages of some profound changes in animal body plans. However, the study system of the authors (lizards with varying degrees of limb reduction and elongation) has a wealth of intermediate forms that the authors have put to very good use. Hence, in my opinion, overall the authors address very interesting questions and have made good choices of study taxa and experimental methods required to address the primary hypotheses that they pose.

Despite my overall enthusiasm for this interesting and potentially important manuscript, I believe that it could be improved greatly by addressing the specific concerns that I detail below. I appreciate that the authors have an abundance of data, and they must deal with significant constraints on manuscript length. Hence, the supplemental material is vitally important for reader to fully understand this work, and I read it in detail. I am also aware of the fact that some of the species studied by authors are quite difficult to obtain; hence, making some allowances for that also seems appropriate. Nonetheless, my approach below is to list all concerns that come to mind and largely leave it to the authors to sort out what is and is not practical to include or revise. Some items of greater concern are indicated by **, and most importantly they include a critical lack of a more logical quantitative measure of elongation (key aspect of snake-like morphology) such as the some relation between mass, cross sectional area or width to total length (rather than residual values based on head length). Additionally, many potential readers in a general audience will not be familiar with these study taxa. Consequently, some combination of images of these lizards along with some anatomical variables that are easier to interpret than the rather abstract PCAs, could go a long way for painting a more concrete picture for the readers regarding the relevant structural variation in this system.

Despite some of my concerns, I believe that the authors have performed experiments sufficient to provide very compelling insights for how: 1) a more elongate and snake-like body plan facilitates sub-surface locomotion, 2) no tradeoffs occur between subsurface and surface locomotion despite a reasonable a-priori expectation for this, and 3) the data for these lizards with a continuum of variation in morphology and locomotion strongly support the suggestion that the ancestors of snakes were elongate as a result of adaptation for subsurface locomotion. I just believe that an extensively revised manuscript with some rethinking and revision of methods and illustrative information can be even more compelling and accessible to an even wider audience. I found this work and the questions it addressed extremely interesting and insightful.

SPECIFIC COMMENTS:

1) I had difficulty following the second part of the last sentence in the abstract. The authors did not find performance trade-offs between surface and subsurface locomotion (lines 143-144). Consequently, I was confused at how the authors could then suggest here that habitat partitioning mitigates the effects of performance trade-offs.

2) Background.

a) (line 33) Perhaps "cluttered" be a better choice of words than "structurally complex". Consider a flat sand surface versus one that has ripples from the wind. The latter is arguably more complex that the former, but this sort of increased complexity does not seem to be likely to have much consequence for impeding the locomotion of limbed animals.

b) (line 39) Perhaps substitute "several clades of ectothermic tetrapod vertebrates" (or some similar wording) for "most major vertebrate clades".

c) (lines 46-48) To me it would make more sense to combine the information on visual pigments and eye anatomy into one issue regarding the near loss of a functional eye as subsequent re-elaboration of an eye in snakes arising from a burrowing ancestor.

d) (lines 51-54) This has the potential to cause confusion. After first referring to 2 ecomorphs, the authors proceed to state that the genus Lerista contains 12 morphs. To me the fact that "Lerista do not fit neatly into either ecomorph" (lines 57-58) diminishes the likelihood that reduced limbed lizards actually have 2 distinct ecomorphs rather than forming a continuum of variation

e) (last paragraph) Perhaps the suggestion that more lizard-like lizards are faster than snake-like lizards could be strengthened by including the fact that a wealth of studies have found very rapid speeds of lizards using limbed locomotion (very commonly more than 10 body lengths per second) compared to the surface locomotion of snakes (very rarely > 3 body lengths per second).

3) Methods.

a**) To me it seemed that the authors lacked a good metric of overall size. I realize that they cite some previous studies of elongate lizards that have used head length for this purpose (supplement p. 3), but based on the biomechanics of this system focused on locomotion, this did not seem like a good choice to me. Either mass or total length would be much better metric of overall size. I realize that the length of the tail compared to snout-vent length can vary substantially among species of reduced-limbed lizards, and so I also understand the desire to include both SVL and tail length in the PCA. Since the authors want to calculate some residuals from a regression using a variable related to overall size, total length would have a drawback because it is the sum of SVL and tail length rather than being independent of these two quantities. By contrast, using mass as an independent variable to calculate residual values of other variables would not have such a drawback. Furthermore, the residuals of lengths predicted from mass to would get right at the heart of what is a more elongate (snake-like) body plan (greater residual value = greater elongation) much better than residuals predicted from head length. In my opinion the authors' analyses are presently deficient because of this lack of some more logical quantitative variable that relates quantities such as mass, maximal cross-sectional area, or maximal width to total length and hence better reflects elongation.

b) (lines 106-115) I favor clarifying the distance over which velocity was calculated. Here authors indicated that they fit a cubic spline to a plot of displacement versus time and then calculated velocity as the slope of that spine. I believe that this was done between successive frames, which were less than 5 ms apart (line 97). Their field of view was at most 518 pixels for ~ a 660 mm long racetrack (~1.3 mm per pixel). Based on Fig. 2, many of the maximal speeds were only approximately 100 mm/s, which would correspond less than 0. 4 mm between successive frames, which is less than 1 pixel. Consequently, I suspect that the displacement between successive images was often less than or equal to the digitizing error, and hence calculating quantities frame by frame does not seem like a good choice. The authors subsequently state that they calculated an average velocity for frame by frame values (line 111), but they did not specify over how many images or over what distance. If they are taking an average of the spline slopes for several successive time intervals, they just as easily could have taken the difference between first and last displacement and divided by the time interval for which the animal was moving, and the result would be simpler and perhaps even more accurate. Perhaps I am just misunderstanding some methods that could be explained more clearly? Maximal velocities decease as the distance over which they are measured increases. Consequently, one good method for quantifying maximal velocity would be to use a fixed distance for calculating an average maximal velocity. Unfortunately, neither the supplement nor the main manuscript provide the

ranges of lengths of their study species. If there is a large range in size, then another option would be to calculate average velocities per cycle and then select the fastest cycle so that the relative distance traveled by different sized animals would be reasonably similar.

c**) In part, in light of some of the issues above, I strongly favor including in the supplement a tabular summary with, at a minimum, means and ranges of SVL, tail lengths and mass for each species. I also believe that some outline figures or good photographs of the study taxa would be extremely helpful and much more useful than the current Fig. 1.

d) (lines 113-115) The biomechanical relevance of a lateral head angle for penetration was not especially clear to me. Wouldn't the amount of ventral flexion would be as important or perhaps even more relevant for understanding how the lizards penetrated the surface of the sand?

e) It was not clear to me if the velocities of burrowing included both penetration of the surface and sub-surface locomotion or just one or the other. If the speed of penetration was analyzed, am I correcting in assuming that it was the time taken until the entire animal was buried? The authors should explicitly state what the case is.

f) For subsurface locomotion, am I correct in assuming that velocity was estimated by tracking the disturbance of the sand rather than a paint mark on the animals? If so this should be clarified as the currently the methods imply that only paint marks were used to track the displacement of the lizards

g) (lines 119-121) Are the authors trying to convey that they only used one maximal speed per individual per substrate, and then they took the mean value (where N = number of individuals for each species. If so, this could be stated more clearly.

h**) I believe leg length could be dealt more effectively. One key benchmark is whether or not a leg is long enough to lift the belly of an animal off the ground, but this is not at all obvious for values of absolute leg length. Although absolute leg lengths were included in the PCA, the supplement lacked any summaries of quantitative data for this important aspect of morphology. I believe a maximally useful quantitative measure of this would be the length of the leg divided by the vertical distance from the ventral surface of the body to the point of attachment of the limb. The current lack of information along these lines obscures the simple but very important issue of whether or not it is even possible for the limbs to be used in a conventional manner rather than just sticking out to the sides of the animal. The authors mentioned that in some circumstances the lizards used their legs when they move, but the wording was not precise enough to know whether or not the legs wiggled ineffectually or if they were contributing to propulsion. Once again, understanding such an issue would be enhanced greatly, in part, by good images of the study taxa.

h) (lines 125-126) I realize it is common practice, but I do not favor asserting that a particular PC is one certain thing such as "overall size" or how "snake-like" an animal is. Each PC is simply a mathematical construct for which the first PC explains the maximal amount of variance of the cloud of points, and by their very construction PCs are not one thing but rather a composite measure of many things. It is perfectly reasonable mention factors that load high for a particular PC, but that is quite different from asserting that one PC is "shape" and another PC is "size". For example, see Gould (1981, The Mismeasure of Man) for one line of argument against reification of factors (that admittedly has not been without controversy).

i**) Except in very special circumstances, I do not favor performing statistical analysis of multiple species with only one individual per species, but the authors have 3 species for which this is the case. This is potentially problematic for a couple of reasons. First, measurements of performance are often highly variable as a result of being strongly influenced by the behavioral motivation of the test subjects. Consequently, a single individual could have an extreme value of performance as a result of behavioral variation rather than variation in its morphology and physiology. Second, the within species variability of morphology in the study taxa is huge, and large variation within species is best dealt with by having sample sizes larger than 1 per species. Note that in Fig. S1 the values of PC2 range from approximately -2.5 to 3, and the range of values for just one (medium green) species (but hard to tell some shades of colors apart) range from approximately -2.5 to 2, which is roughly 80% of the range of PC2 for all species! Imagine potentially what could have happened if just one of these extreme points was sampled for this particular species.

j) The large ranges in PC scores among some of the individuals belonging to the same species also made me wonder if the authors had restricted their analysis to only adults to reduce the variability of size within each of their study species. They should explicitly address the issue and provide compelling reasons for their sampling strategy.

4) Results.

a) (line 139-143) The authors performed several statistical tests, but made no effort to account for this by adjusting P-values. Clearly, this is a thorny issue with no easy answer as using extremely small P-values arising from adjusting for multiple comparisons just introduces a different sort of error. However, three important results highlighted as significant (Tables S3 and S5) had P-values ranging from 0.026 to 0.046. Hence, some of these conclusions might be on rather shaky ground.

b) (line 144) As indicated above, I do not favor using head size and body size interchangeably.

c**) Tables S3, S4, S5 S6 included many non-significant independent variables in (forced-fit) multiple regression regressions. Including these non-significant independent variables can lead to spurious conclusions regarding the significance level of the other independent variable in the multiple regression. Hence, I favor redoing the all of the multiple regressions and only reporting the models in which all of coefficients of the independent variables are statistically significant.

5) Discussion.

a**) The authors repeatedly acknowledge the importance of head shape for affecting burrowing performance. Consequently, I was surprised by the absence of any variable quantifying head shape for the study taxa, especially since the biomechanical relevance of this seems much more obvious than that of the lateral angle of the head, which was quantified.

b) (lines 186-196) The authors mention that the relative length of the tail varies substantially among elongate lizard taxa. A substantial difference between many elongate lizards and snakes is also the greater relative length tail in lizards. Furthermore, most burrowing snakes have quite short and blunt tails relative to their terrestrial counterparts. However, despite having snout vent length and tail length, no data were reported for relative tail length.

c) By the way, do any of these study taxa have tail autotomy? If so, did all of the individuals used in the study have original rather than regenerated tails? This should be stated explicitly. I also strongly favor including only intact individuals.

6) (Throughout the main manuscript and supplement) Given that some taxa did not use their limbs for surface locomotion, "running" seems to be a poor choice of terminology compared to "surface locomotion".

7) Supplement.

a**) I strongly favor including a conventional rooted cladogram of the study taxa as a figure in the supplement.

b) A better color scheme should be possible for Fig S1. I was not confident I could distinguish between the two of the shades green (L. kendricki and L.connivens) and two of the shades of blue (L. macropishopus and L. dorsalis).

c) The authors should explicitly state if a value of zero was used for limb length when limbs were absent.

d) I could not find anywhere close to an N =13 for the light blue symbols and dark blue squares in Fig. S3. Perhaps the author are trying to cram too much information in each panel of this figure. Since it is in the supplement anyway, I favor clarity over small size.

e) What is the definition of a "stump" limb?

f) (Fig. S3) It seemed somewhat odd to me to include the seven species that that lacked front limbs (Table S1) in the analysis of the proportion of individuals that used the fronts limbs since front limb use is not possible in these species.

g) The heading for Table S5 refers to "penetration during burrowing". However, two entries in the body of this table use the term "sand-swimming". It is vitally important that the authors use consistent terminology for the different types of locomotion, which I believe include the initial penetration of loose substrates, sub-surface locomotion and surface locomotion.

h) A few supplementary videos might be nice for illustrating the presence or absence of limb use or variation in how the limbs are used.

i**) The multiple regression analyses in Tables S3-7 included both pPC1 and head length as independent variables. However, Table S2 indicates that head length was a factor in the PCA. Consequently, it seemed inappropriate to perform a multiple regression analysis that in some sense used the same variable (head length) twice.

Author's Response to Decision Letter for (RSPB-2019-1841.R0)

See Appendix A.

RSPB-2020-0192.R0

Review form: Reviewer 1

Recommendation
Accept with minor revision (please list in comments)

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?
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
I have no read the revised version of the paper entitled 'The evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. I found the revision improved and the replies to the comments of both reviewers adequate and careful. I only have two minor remaining remarks that I'd like to see addressed.

1) on line 21 the statement is ambiguous ... as you did not quantify subsurface locomotion but rather soil penetration velocity it would be good to be very clear about this in the abstract.

2) on line 98, please state that this is a uni-axial force transducer. I agree that you data are likely comparable across taxa but for non-experts it is important that they are aware of the fact that you measured only one of the force components. Related it would be nice to add a sentence stating that it would be nice to validate these measures with 3D force plate recordings and/or Xray video.

Review form: Reviewer 2

Recommendation
Accept with minor revision (please list in comments)

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

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

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
I liked a lot that the first version of the manuscript had to offer, and the revised version is even better still. I would like to compliment the authors on their thoroughness in responding to my suggestions, virtually all of which have been dealt effectively. I found the new supplementary tables very helpful and I especially liked the new supplemental figures that showed the lizards penetrating sand and moving through the cluttered substrate. I have only a few very minor points detailed below that I believe merit attention.

1) Locomotion in sand.
a) It is very clear in the response to reviews that the authors studied the initial penetration of sand (time taken to bury the body up to the pelvic girdle) rather than undulatory locomotion beneath the surface of the sand when the body is completely submerged (sand-swimming). In my opinion some very significant differences exist between initial penetration of sand and "swimming" beneath the sand. Hence, wouldn't it clearest and most intuitive simply to use "sand penetration" rather than "sand swimming"?
b) Part of my logic for the comment above is that some species of lizards with well-developed legs and impressive sprinting speeds on the surface of sand are also proficient at penetrating the surface of the sand. For example, fringe-toed lizards (SVL=7-10 cm) can bury themselves up to their pelvic girdle in approximately 1.6 s (Jayne & Daggy, 2000, J. Exp. Biol.), and I believe this compares quite favorably to some of the faster penetration times for the more snake-like species studied by the authors (Fig. 3). The fringe-toed lizards also have a rather stereotyped behavior in which both of the hind limbs are simultaneously retracted, and this appears important for providing a purchase from which the lizards can forcibly push forward and penetrate the sand. I do not dispute that within the clade of lizards studied by the authors more snake-like lizards had superior sand penetration. However, at least for fringe-toed lizards, well developed limbs and other aspect of morphology and behavior can result in superior penetration performance than many of the reduced limb species studied by the authors. I was uncertain whether or not the authors were aware of this, and I leave it to them to decide if this is worth a very brief mention in their manuscript.

2) (lines 47-48) Squamates = lizards + snakes + amphisbaenians not just lizards and snakes.

3) Line 150 cites non-existent Table 4S. After this citation, many of the other citations appear to be incorrect for how the supplementary Tables are presently numbed. Hence, the authors should update the table numbers and carefully double check all of the citations.

4) Because I found this paper so interesting, I really liked having many of the additional details provided in the new supplementary materials which allowed me to relate some data to a particular species. Unfortunately, in the current scheme I realized that I could not find sand penetration velocities for the species that had hind limbs but no forelimbs. Try as I might, I could not envision some system of unique symbols for each species in all plots of the functional data (Figs. 3, S6, S7) because of the admirably large number of study species. Hence, I believe the easiest fix for this would behave a supplementary table with the all of the mean and values of functional data for each of the species (Figs. 3, S6, S7). I apologize for not catching this in the first round of comments.

Once, congratulations on a fine study! I wish more of the other papers I have been asked to review were this interesting and well done.

Decision letter (RSPB-2020-0192.R0)

11-Feb-2020

Dear Dr Morinaga:

Your manuscript has now been peer reviewed and the reviews have been assessed by an Associate Editor. The reviewers' comments (not including confidential comments to the Editor) and the comments from the Associate Editor are included at the end of this email for your reference. As you will see, the reviewers and the Editors have raised some concerns with your manuscript and we would like to invite you to revise your manuscript to address them.

We do not allow multiple rounds of revision so we urge you to make every effort to fully address all of the comments at this stage. If deemed necessary by the Associate Editor, your manuscript will be sent back to one or more of the original reviewers for assessment. If the original reviewers are not available we may invite new reviewers. Please note that we cannot guarantee eventual acceptance of your manuscript at this stage.

To submit your revision please 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 Revision". Your manuscript number has been appended to denote a revision.

When submitting your revision please upload a file under "Response to Referees" in the "File Upload" section. This should document, point by point, how you have responded to the reviewers' and Editors' comments, and the adjustments you have made to the 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.

Your main manuscript should be submitted as a text file (doc, txt, rtf or tex), not a PDF. Your figures should be submitted as separate files and not included within the main manuscript file.

When revising your manuscript you should also ensure that it adheres to our editorial policies (https://royalsociety.org/journals/ethics-policies/). You should pay particular attention to the following:

Research ethics:
If your study contains research on humans please ensure that you detail in the methods section whether you obtained ethical approval from your local research ethics committee and gained informed consent to participate from each of the participants.

Use of animals and field studies:
If your study uses animals please include details in the methods section of any approval and licences given to carry out the study and include full details of how animal welfare standards were ensured. Field studies should be conducted in accordance with local legislation; please include details of the appropriate permission and licences that you obtained to carry out the field work.

Data accessibility and data citation:
It is a condition of publication that you make available the data and research materials supporting the results in the article. Datasets should be deposited in an appropriate publicly available repository and details of the associated accession number, link or DOI to the datasets must be included in the Data Accessibility section of the article (https://royalsociety.org/journals/ethics-policies/data-sharing-mining/). Reference(s) to datasets should also be included in the reference list of the article with DOIs (where available).

In order to ensure effective and robust dissemination and appropriate credit to authors the dataset(s) used should also be fully cited and listed in the references.

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.

For more information please see our open data policy http://royalsocietypublishing.org/data-sharing.

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. Please try to submit all supplementary material as a single file.

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].

Please submit a copy of your revised paper within three weeks. If we do not hear from you within this time your manuscript will be rejected. If you are unable to meet this deadline please let us know as soon as possible, as we may be able to grant a short extension.

Thank you for submitting your manuscript to Proceedings B; we look forward to receiving your revision. If you have any questions at all, please do not hesitate to get in touch.

Best wishes,
Professor John Hutchinson, Editor
mailto: proceedingsb@royalsociety.org

Associate Editor Board Member
Comments to Author:
I thank the authors for responding constructively and exhaustively to the initial reviews of their manuscript and carrying out the additional analyses suggested by the expert reviewers. I agree with the reviewers that these modified analyses, and the other editorial and aesthetic changes, to the manuscript have greatly improved the paper. Both reviewers have suggested some very minor but reasonable additional modification to improve clarity and smooth out the rough edges on the manuscript. It is my recommendation that the authors be given the opportunity to resubmit a modified version to incorporate those changes.

Reviewer(s)' Comments to Author:
Referee: 1

Comments to the Author(s).
I have no read the revised version of the paper entitled 'The evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. I found the revision improved and the replies to the comments of both reviewers adequate and careful. I only have two minor remaining remarks that I'd like to see addressed.

1) on line 21 the statement is ambiguous ... as you did not quantify subsurface locomotion but rather soil penetration velocity it would be good to be very clear about this in the abstract.

2) on line 98, please state that this is a uni-axial force transducer. I agree that you data are likely comparable across taxa but for non-experts it is important that they are aware of the fact that you measured only one of the force components. Related it would be nice to add a sentence stating that it would be nice to validate these measures with 3D force plate recordings and/or Xray video.

Referee: 2

Comments to the Author(s).
I liked a lot that the first version of the manuscript had to offer, and the revised version is even better still. I would like to compliment the authors on their thoroughness in responding to my suggestions, virtually all of which have been dealt effectively. I found the new supplementary tables very helpful and I especially liked the new supplemental figures that showed the lizards penetrating sand and moving through the cluttered substrate. I have only a few very minor points detailed below that I believe merit attention.

1) Locomotion in sand.
a) It is very clear in the response to reviews that the authors studied the initial penetration of sand (time taken to bury the body up to the pelvic girdle) rather than undulatory locomotion beneath the surface of the sand when the body is completely submerged (sand-swimming). In my opinion some very significant differences exist between initial penetration of sand and "swimming" beneath the sand. Hence, wouldn't it clearest and most intuitive simply to use "sand penetration" rather than "sand swimming"?
b) Part of my logic for the comment above is that some species of lizards with well-developed legs and impressive sprinting speeds on the surface of sand are also proficient at penetrating the surface of the sand. For example, fringe-toed lizards (SVL=7-10 cm) can bury themselves up to their pelvic girdle in approximately 1.6 s (Jayne & Daggy, 2000, J. Exp. Biol.), and I believe this compares quite favorably to some of the faster penetration times for the more snake-like species studied by the authors (Fig. 3). The fringe-toed lizards also have a rather stereotyped behavior in which both of the hind limbs are simultaneously retracted, and this appears important for providing a purchase from which the lizards can forcibly push forward and penetrate the sand. I do not dispute that within the clade of lizards studied by the authors more snake-like lizards had superior sand penetration. However, at least for fringe-toed lizards, well developed limbs and other aspect of morphology and behavior can result in superior penetration performance than many of the reduced limb species studied by the authors. I was uncertain whether or not the authors were aware of this, and I leave it to them to decide if this is worth a very brief mention in their manuscript.

2) (lines 47-48) Squamates = lizards + snakes + amphisbaenians not just lizards and snakes.

3) Line 150 cites non-existent Table 4S. After this citation, many of the other citations appear to be

incorrect for how the supplementary Tables are presently numbed. Hence, the authors should update the table numbers and carefully double check all of the citations.

4) Because I found this paper so interesting, I really liked having many of the additional details provided in the new supplementary materials which allowed me to relate some data to a particular species. Unfortunately, in the current scheme I realized that I could not find sand penetration velocities for the species that had hind limbs but no forelimbs. Try as I might, I could not envision some system of unique symbols for each species in all plots of the functional data (Figs. 3, S6, S7) because of the admirably large number of study species. Hence, I believe the easiest fix for this would behave a supplementary table with the all of the mean and values of functional data for each of the species (Figs. 3, S6, S7). I apologize for not catching this in the first round of comments.

Once, congratulations on a fine study! I wish more of the other papers I have been asked to review were this interesting and well done.

Author's Response to Decision Letter for (RSPB-2020-0192.R0)

See Appendix B.

Decision letter (RSPB-2020-0192.R1)

25-Feb-2020

Dear Dr Morinaga

I am pleased to inform you that your manuscript entitled "The evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards" has been accepted for publication in Proceedings B. Congratulations!!

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

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.

Your article has been estimated as being 7 pages long. Our Production Office will be able to confirm the exact length at proof stage.

Paper charges
An e-mail request for payment of any related charges will be sent out after proof stage (within approximately 2-6 weeks). 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.

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,
Professor John Hutchinson
Editor, Proceedings B
mailto: proceedingsb@royalsociety.org

Appendix A

Associate Editor
Board Member: 1
Comments to Author:
Thank you for the opportunity to review this paper. Herein the authors attempt to test existing but competing hypotheses for the evolution of body elongation and limb loss in vertebrates: that these traits evolved as adaptations for fossorial locomotion, or as a means to better locomote through "structurally complex habitats." To test these competing hypotheses, the authors collect experimental data on lizards locomoting through sand and leaf litter. Based on specific metrics (e.g. velocity) they conclude that their experiments support the hypothesis that elongation and limb loss is better explained as an adaptation for fossoriality.

I personally enjoyed reading the paper and believe questions driving it are of wide biological significance. Both reviewers were also enthusiastic about the what the paper attempts to achieve, but do raise some important and quite fundamental concerns about the methods and the depth of interpretation (or extrapolation) from the data collected. Notable issues that need to be confronted one way or the other include (but are not limited to) justification for sand and leaf litter as a appropriate substrate choices, how burrowing locomotion is being measured given submersion in sand, body size quantification and a number of issues with the statistics raised by reviewer 2. If the authors feel they can address these issues constructively then I would encourage them to do so.

---We are happy that the Associate Editor enjoyed the paper and considers our work significant to warrant publication in Proceedings of the Royal Society. Our original submission did not provide some details about locomotor trials done in sand and leaf litter, as pointed out by both reviewers. We now better describe *Acacia* leaf litter in the introduction [lines 53–54] and better discuss their significance in the discussion [lines 218–220]. We have also added Supplementary Figures (S3–S5) which show sequences of still-frames from a sample video of sand-swimming, leaf litter, and surface locomotion trials.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s)
I have now read the paper by Morinaga and Bergmann on 'the evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards'. I enjoyed the paper and think the authors have a nice data set. I do find that they have a tendency to oversell their data a bit by forcing it into the 'origin of snakes hypothesis'. In itself it is already an interesting question and I don't think it needs that. I do have a number of comments and suggestions that I would like to see addressed:

---We appreciate that the reviewer enjoyed the paper, given the work we put into the project. We agree that our original submission was heavy-handed on the 'origin of snakes' aspect and now focus more on the evolution of locomotion in fossorial and cluttered habitats.

lien 15: it is important to not that there need not be one all encompassing reason for limb loss. Different taxa may have lost limbs for different reasons (very likely even) and the two hypotheses are not mutually exclusive. Here you present it as a real dichotomy.

---We agree and have edited the entire manuscript to not imply that the evolution of snake-like body shapes evolved only for cluttered or fossorial habitats, and that these two hypotheses are not mutually exclusive.

line 47: the anatomy of the snake eye

---Done.

line 57: this is more of a conceptual issue, but is sand swimming really burrowing ? I would argue not really as the constraints and mechanics are very different for sand swimming (sand behaving almost like a fluid) versus burrowing in a solid, compact substrate. I would like to see this addressed and discussed.

---We thank the reviewer for bringing this point up. We agree that there is a clear distinction between burrowing and sand-swimming, notably in how the substrate behaves as the animals move through it. We originally felt that for non-specialist readers burrowing would be more intuitive than sand-swimming. However, given that there is an important distinction between these modes of locomotion, we have changed all appropriate instances of 'burrowing' to 'sand-swimming'. Prior treatments of fossorial squamates have lumped burrowers and sand-swimmers together and have considered sand-swimmers as burrowers.

Line 93: I was really confused as how you were able to film burrowing with a regular camera ... by definition burrowing and even moving in leaf litter happens in the substrate and as such the animals are not visible. I think it is important to make very clear that what you studied is initial soil penetration from the surface. Even when this is what you appear to have studied (based on the supplementary methods) I still wonder how much you are able to quantify as you will loose the point you track right away. I would really like to see more information on the filming protocol to have an ideal about what it is you really studied.

---This is an important point that we have now clarified in both the main manuscript [lines 92–96] and in the Supplementary Methods. We also now provide new Supplementary Figures S3–S5 which are sequences of frames from representative videos of sand-swimming, burrowing into leaf litter, and surface locomotion, respectively and reference them in the main manuscript. We note that by following a point on the animals' pelvic girdle we were able to track progress of submerging into the substrate until that point disappeared. This approach allowed us to accurately quantify velocity as the animal disappeared into the substrate.

line 99: this is a uni-axial force transducer. Especially in skinks forces are intrinsically 3D with much of the forces being not in the Z direction (see the paper by Vanhooydonck in BJLS). As such your force measures are likely underestimates. This in itself is not worrisome but if species different in burrowing style then your measures will be biased by that and will not represent true performance. I would definitely like to see this discussed in the paper.

--- All of the Lerista species we examined engaged in sand-swimming, and it is likely all other Lerista do as well. Thus we are confident that our data should not have different bias from one species to the next. We do agree that our measurements are underestimates and now acknowledge this in the Supplementary Methods. However, we believe that the majority of the force exerted by the animals during our force measurements were aligned with the force transducer. We also state that all species we examined only engaged in sand-swimming when placed in the sand substrates and provide references that describe Lerista in general as sand-swimmers.

line 169: is a leaf litter mat really a structurally complex habitat? I really wonder about this as leaves are soft and compressible and as such the animal does not really have to move through tight interstitial spaces but can and will simply create space by compacting the soft dead leaves. This makes the system not so great for testing your initial question.

---The fallen Acacia leaves that comprise the litter mats tightly interdigitate and are stiff and slender, producing a habitat with many small crawl spaces. The Lerista that crawl through them are also quite small, meaning that the stiff leaves are certainly not experienced as soft and compressible by the animals, although they are still somewhat compliant. Other species of lizards that inhabit these complex habitats often live in dense brush or grass, which are also compliant. Hence, we would argue that the leaf litter actually makes an ideal substrate for testing this question. We have worked to better describe this habitat in the manuscript to address these concerns, and now address it in the Introduction [59–63], Discussion [203–207; 235–245], and the Supplementary Methods. We also note that, following suggestions of the second reviewer, we now refer to this habitat as “cluttered” as opposed to “structurally complex” because it more accurately describes it.

line 180: there are many eels that burrow head-first into the substrate; see papers by Choi and colleagues (J. Exp. Biol.) and De Schepper and colleagues, for example.

--- We now cite De Schepper et al. (2005) as it explicitly collected functional data.

line 192: this is where you need to make a difference between burrowing and sand swimming. In sand swimming the tail can effectively be used to produce propulsive forces as in in any anguilliform swimmer. Consequently long tails are not a disadvantage for sand swimming.

---While we agree with the reviewer that anguilliform swimmers generate propulsive forces with their tails, it is still unknown whether sand-swimmers move their bodies in ways that resemble anguilliform swimming under the surface, or that the tail is capable of exerting propulsive forces during sand-swimming bouts. We now state this explicitly [198–199], and qualify that increases in total length has diminishing returns on cost of transport while also incurring greater drag [200–202].

The figures of my review copy weren't great but that's probably just a conversion issue. I liked some of the figures in the supplementary as well (S1 & S2) and would like to have seen them in the main manuscript.

---We appreciate the positive feedback on our figures. Adding additional figures is challenging because space is at a premium, but considering the reviewer comments, we have decided to move Figure S1 into the main manuscript as part of a panel figure (now Figure 2).

Referee: 2

Comments to the Author(s)
GENERAL COMMENTS:

I believe the manuscript by Morinaga and Bergmann addresses very general issues regarding evolution, structure/function relationships, evolutionary and functional trade-offs, and niche partitioning that have great potential to appeal to a very wide readership for the Proceedings of the Royal Society. I wholeheartedly agree with the authors that a lack of extant intermediate forms has hindered our ability to understand the adaptive benefits involved in the early stages of some profound changes in animal body plans. However, the study system of the authors (lizards with varying degrees of limb reduction and elongation) has a wealth of intermediate forms that the authors have put to very good use. Hence, in my opinion, overall the authors address very interesting questions and have made good choices of study taxa and experimental methods required to address the primary hypotheses that they pose.

Despite my overall enthusiasm for this interesting and potentially important manuscript, I believe that it could be improved greatly by addressing the specific concerns that I detail below. I appreciate that the authors have an abundance of data, and they must deal with significant constraints on manuscript length. Hence, the supplemental material is vitally important for reader to fully understand this work, and I read it in detail. I am also aware of the fact that some of the species studied by authors are quite difficult to obtain; hence, making some allowances for that also seems appropriate. Nonetheless, my approach below is to list all concerns that come to mind and largely leave it to the authors to sort out what is and is not practical to include or revise. Some items of greater concern are indicated by **, and most importantly they include a critical lack of a more logical quantitative measure of elongation (key aspect of snake-like morphology) such as the some relation between mass, cross sectional area or width to total length (rather than residual values based on head length). Additionally, many potential readers in a general audience will not be familiar with these study taxa. Consequently, some combination of images of these lizards along with some anatomical variables that are easier to interpret than the rather abstract PCAs, could go a long way for painting a more concrete picture for the readers regarding the relevant structural variation in this system.

Despite some of my concerns, I believe that the authors have performed experiments sufficient to provide very compelling insights for how: 1) a more elongate and snake-like body plan facilitates sub-surface locomotion, 2) no tradeoffs occur between subsurface and surface locomotion despite a reasonable a-priori expectation for this, and 3) the data for these lizards with a continuum of variation in morphology and locomotion strongly support the suggestion that the ancestors of snakes were elongate as a result of adaptation for subsurface locomotion. I just believe that an extensively revised manuscript with some rethinking and revision of methods and illustrative information can be even more compelling and accessible to an even wider audience. I found this work and the questions it addressed extremely interesting and insightful.

SPECIFIC COMMENTS:

1) I had difficulty following the second part of the last sentence in the abstract. The authors did not find performance trade-offs between surface and subsurface locomotion (lines 143-144). Consequently, I was confused at how the authors could then suggest here that habitat partitioning mitigates the effects of performance trade-offs.

---We have changed the wording for these lines to better reflect our findings.

2) Background.

a) (line 33) Perhaps "cluttered" be a better choice of words than "structurally complex". Consider a flat sand surface versus one that has ripples from the wind. The latter is arguably more complex that the former, but this sort of increased complexity does not seem to be likely to have much consequence for impeding the locomotion of limbed animals.

---We agree with the reviewer's suggestion and have replaced 'structurally complex' with 'cluttered' throughout.

b) (line 39) Perhaps substitute "several clades of ectothermic tetrapod vertebrates" (or some similar wording) for "most major vertebrate clades".

---We elected to keep the wording from our original submission because we also consider fishes in our statement, which are not tetrapods, and mammals, which are not ectotherms. We had included a reference about fishes and now also cite Law et al. (2019), which is about elongate mammals (mustelids) to better support our statement [now lines 41–43].

c) (lines 46-48) To me it would make more sense to combine the information on visual pigments and eye anatomy into one issue regarding the near loss of a functional eye as subsequent re-elaboration of an eye in snakes arising from a burrowing ancestor.

---In addressing Reviewer 1's opening comment about overstating the "origin of snakes" aspect, we have elected to remove these lines and shorten the paragraph overall to better focus on the evolution of locomotion in fossorial and cluttered habitats.

d) (lines 51-54) This has the potential to cause confusion. After first referring to 2 ecomorphs, the authors proceed to state that the genus Lerista contains 12 morphs. To me the fact that "Lerista do not fit neatly into either ecomorph" (lines 57-58) diminishes the likelihood that reduced limbed lizards actually have 2 distinct ecomorphs rather than forming a continuum of variation

---We have changed the wording here from 'morph' to 'forms' [line 55] and deleted "Lerista do not fit neatly into either ecomorph…" to avoid confusion. To clarify here, the two ecomorphs we referred to were the long-tailed surface-dwelling and short-tailed fossorial that Brandley et al. refer to. The twelve forms of Lerista refer to species with different number of digits on front and hind limbs.

e) (last paragraph) Perhaps the suggestion that more lizard-like lizards are faster than snake-like lizards could be strengthened by including the fact that a wealth of studies have found very rapid speeds of lizards using limbed locomotion (very commonly more than 10 body lengths per second) compared to the surface locomotion of snakes (very rarely > 3 body lengths per second).

---We now cite two references to support this statement [line 70], and thank the reviewer for this suggestion.

3) Methods.

a**) To me it seemed that the authors lacked a good metric of overall size. I realize that they cite some previous studies of elongate lizards that have used head length for this purpose (supplement p. 3), but based on the biomechanics of this system focused on locomotion, this did not seem like a good choice to me. Either mass or total length would be much better metric of overall size. I realize that the length of the tail compared to snout-vent length can vary substantially among species of reduced-limbed lizards, and so I also understand the desire to include both SVL and tail length in the PCA. Since the authors want to calculate some residuals from a regression using a variable related to overall size, total length would have a drawback because it is the sum of SVL and tail length rather than being independent of these two quantities. By contrast, using mass as an independent variable to calculate residual values of other variables would not have such a drawback. Furthermore, the residuals of lengths predicted from mass to would get right at the heart of what is a more elongate (snake-like) body plan (greater residual value = greater elongation) much better than residuals predicted from head length. In my opinion the authors' analyses are presently deficient because of this lack of some more logical quantitative variable that relates quantities such as mass, maximal cross-sectional area, or maximal width to total length and hence better reflects elongation.

---We wholeheartedly agree with the reviewer and appreciate the critical appraisal. We have redone all analyses using mass as the proxy for size. As the reviewer points out, mass has none of the drawbacks of the other options.

b) (lines 106-115) I favor clarifying the distance over which velocity was calculated. Here authors indicated that they fit a cubic spline to a plot of displacement versus time and then calculated velocity as the slope of that spine. I believe that this was done between successive frames, which were less than 5 ms apart (line 97). Their field of view was at most 518 pixels for ~ a 660 mm long racetrack (~1.3 mm per pixel). Based on Fig. 2, many of the maximal speeds were only approximately 100 mm/s, which would correspond less than 0. 4 mm between successive frames, which is less than 1 pixel. Consequently, I suspect that the displacement between successive images was often less than or equal to the digitizing error, and hence calculating quantities frame by frame does not seem like a good choice. The authors subsequently state that they calculated an average velocity for frame by frame values (line 111), but they did not specify over how many images or over what distance. If they are taking an average of the spline slopes for several successive time intervals, they just as easily could have taken the difference between first and last displacement and divided by the time interval for which the animal was moving, and the result would be simpler and perhaps even more accurate. Perhaps I am just misunderstanding some methods that could be explained more clearly?

Maximal velocities decease as the distance over which they are measured increases. Consequently, one good method for quantifying maximal velocity would be to use a fixed distance for calculating an average maximal velocity. Unfortunately, neither the supplement nor the main manuscript provide the ranges of lengths of their study species. If there is a large range in size, then another option would be to calculate average velocities per cycle and then select the fastest cycle so that the relative distance traveled by different sized animals would be reasonably similar.

---To address this, we assembled a new dataset of the velocity from the fastest 5 cm interval for each individual on each substrate (i.e., coarse, fine, leaf litter, and mixed). These new data for showed slightly higher averages for each species (as expected), but much higher variances for sand-swimming and leaf litter trials, but not for running trails. We believe that this is due to quite unsteady velocity during burying behavior. Given these findings, we felt that the average velocity during burying (whether in sand or leaf litter) provided the most accurate and lowest noise data. Then we argue that it is important to collect the running velocity data in the same way. The pattern of variation among species was consistent between our original data and the 5 cm interval data. Therefore, we decided to present our original dataset. Considering that there was some confusion over how many frames we calculated average velocity, we have clarified in the methods that we estimated average velocity from all of the frame-by-frame values during which the animal was moving [lines 124–125].

c**) In part, in light of some of the issues above, I strongly favor including in the supplement a tabular summary with, at a minimum, means and ranges of SVL, tail lengths and mass for each species. I also believe that some outline figures or good photographs of the study taxa would be extremely helpful and much more useful than the current Fig. 1.

---We have added Supplementary Table S2 that summarizes average SVL, hindlimb length, tail length, and mass for each species and their standard deviations. We have also added silhouettes of L. micra, L. elegans, L. praepedita, and L. lineopunctulata (So) to the phylomorphospace (now Figure 2).

d) (lines 113-115) The biomechanical relevance of a lateral head angle for penetration was not especially clear to me. Wouldn't the amount of ventral flexion would be as important or perhaps even more relevant for understanding how the lizards penetrated the surface of the sand?

---We have removed this analysis from the current submission. In addressing the reviewer's later comment (Results a), the effect of body shape on angle was no longer significant. Furthermore, we believe that in removing this, we are better able to focus on how differences in body shape affect locomotion in different substrates, while also shortening the manuscript overall.

e) It was not clear to me if the velocities of burrowing included both penetration of the surface and sub-surface locomotion or just one or the other. If the speed of penetration was analyzed, am I correcting in assuming that it was the time taken until the entire animal was buried? The authors should explicitly state what the case is.

---We tracked the pelvic point starting from the moment the tip of the snout penetrated the substrate until the pelvic point disappeared into the substrate. We now state this in the methods [lines 118–119]. The velocity we estimated was the frame-by-frame average as described in the methods [lines 123–124].

f) For subsurface locomotion, am I correct in assuming that velocity was estimated by tracking the disturbance of the sand rather than a paint mark on the animals? If so this should be clarified as the currently the methods imply that only paint marks were used to track the displacement of the lizards

---We clarify that we tracked the pelvic point during sand-swimming and leaf litter trials [lines 117–118]. We also now provide a sequence of frames for both leaf litter and sand-swimming trials in Figures S3 and S4.

g) (lines 119-121) Are the authors trying to convey that they only used one maximal speed per individual per substrate, and then they took the mean value (where N = number of individuals for each species. If so, this could be stated more clearly.

---This is what was done. We now state that more clearly in the manuscript [line 133] as well as in the Supplementary Methods.

h**) I believe leg length could be dealt more effectively. One key benchmark is whether or not a leg is long enough to lift the belly of an animal off the ground, but this is not at all obvious for values of absolute leg length. Although absolute leg lengths were included in the PCA, the supplement lacked any summaries of quantitative data for this important aspect of morphology. I believe a maximally useful quantitative measure of this would be the length of the leg divided by the vertical distance from the ventral surface of the body to the point of attachment of the limb. The current lack of information along these lines obscures the simple but very important issue of whether or not it is even possible for the limbs to be used in a conventional manner rather than just sticking out to the sides of the animal. The authors mentioned that in some circumstances the lizards used their legs when they move, but the wording was not precise enough to know whether or not the legs wiggled ineffectually or if they were contributing to propulsion. Once again, understanding such an issue would be enhanced greatly, in part, by good images of the study taxa.

---This is an important point that we have tried to clarify in our revision. Although we did not take measurements of the distance from where the limb exits the body to the ventral edge of the body, the limbs in Lerista are positioned quite ventrolaterally, such that even in species with extremely reduced limbs, like L. praepedita, they do reach the ground. Also, in all of the species that we studied, the belly remained in contact with the ground during locomotion (the limbs were never used to elevate the body from the ground). Hence, when animals used their limbs during locomotion, the limbs always gained purchase on the substrate. However, it is unknown how much they actually contribute to propulsion. We now provide a better description of how we determined whether the hindlimb was used in the methods [lines 124–127].

h) (lines 125-126) I realize it is common practice, but I do not favor asserting that a particular PC is one certain thing such as “overall size” or how “snake-like” an animal is. Each PC is simply a mathematical construct for which the first PC explains the maximal amount of variance of the cloud of points, and by their very construction PCs are not one thing but rather a composite measure of many things. It is perfectly reasonable mention factors that load high for a particular PC, but that is quite different from asserting that one PC is “shape” and another PC is “size”. For example, see Gould (1981, The Mismeasure of Man) for one line of argument against reification of factors (that admittedly has not been without controversy).

---We have re-worded our interpretations of the pPCs and state which variables load strongly and their direction (i.e., positive or negative) [line137–140], as well as in figure captions and table headings.

i**) Except in very special circumstances, I do not favor performing statistical analysis of multiple species with only one individual per species, but the authors have 3 species for which this is the case. This is potentially problematic for a couple of reasons. First, measurements of performance are often highly variable as a result of being strongly influenced by the behavioral motivation of the test subjects. Consequently, a single individual could have an extreme value of performance as a result of behavioral variation rather than variation in its morphology and physiology. Second, the within species variability of morphology in the study taxa is huge, and large variation within species is best dealt with by having sample sizes larger than 1 per species. Note that in Fig. S1 the values of PC2 range from approximately -2.5 to 3, and the range of values for just one (medium green) species (but hard to tell some shades of colors apart) range from approximately -2.5 to 2, which is roughly 80% of the range of PC2 for all species! Imagine potentially what could have happened if just one of these extreme points was sampled for this particular species.

---This is a valid concern that we now address. Working with wild cryptic animals in remote locations presents numerous challenges, and chief among them is reliably finding species of interest in sizeable numbers. We agree that sampling single individuals for some species is not ideal, but removing these species from our analysis would exclude potentially critical data, and lower sample size for the number of species included.

To address this, we supplemented our morphometric dataset with morphometric measurements of eight museum specimens for each species we had n=1 (L. connivens, L. labialis, L. praepedita). We then performed a pPCA on this supplemented dataset and compared the pPC scores of the specimens we collected to those of the museum specimens. We found that the pPC score for the L. labialis and L. praepedita specimens we collected in the field were similar to those of the museum specimens. However, the scores for the L. connivens specimen we collected in the field differed substantially from the museum specimens. Therefore, we re-analyzed the dataset without L. connivens, but retained L. labialis and L. praepedita. We now justify the n=1 in the methods for the species we retained [lines 141–145], and present the results from the pPCA we describe in this response in the Supplementary Materials (Table S4; Figure S5). The main manuscript now excludes L. connivens.

j) The large ranges in PC scores among some of the individuals belonging to the same species also made me wonder if the authors had restricted their analysis to only adults to reduce the variability of size within each of their study species. They should explicitly address the issue and provide compelling reasons for their sampling strategy.

---We only included adults and sub-adults in our analyses, and now explicitly state this on line [82]. It is also worth noting that all of our analyses accounted for size by including mass as one of the explanatory variables.

4) Results.

a) (line 139-143) The authors performed several statistical tests, but made no effort to account for this by adjusting P-values. Clearly, this is a thorny issue with no easy answer as using extremely small P-values arising from adjusting for multiple comparisons just introduces a different sort of error. However, three important results highlighted as significant (Tables S3 and S5) had P-values ranging from 0.026 to 0.046. Hence, some of these conclusions might be on rather shaky ground.

---We agree that how to best account for multiple comparisons is a complicated issue. Under different circumstances (e.g., correlation), we would have adjusted for false discovery rate using the Benjamini-Hochberg correction. However, it is unclear how correction should take place in a multiple regression framework—should P-values be corrected within each regression (i.e., P-values for three parameters in our case), within similar sets of analyses (i.e., within a table), or globally (i.e., pool all P-values for all parameters estimated)? It is also unclear how multiple comparisons should be handled when reducing models to the minimal significant model, as requested in comment "Results c**".

We have taken two steps to address this, albeit indirectly. First, we followed the reviewer's suggestion for comment "Results c**" and reduced each model to include only significant model parameters. This had the effect of reducing P-values for parameters with strong effects, while parameters that were on the cusp of significance became non-significant after model reduction. Second, we have re-worded parts of the results to not emphasize significance, and now report adjusted R^2 to bring to attention the importance of effect size.

b) (line 144) As indicated above, I do not favor using head size and body size interchangeably.

---We now use ln-mass as our measure of size in all multiple regression analyses.

c**) Tables S3, S4, S5 S6 included many non-significant independent variables in (forced-fit) multiple regression regressions. Including these non-significant independent variables can lead to spurious conclusions regarding the significance level of the other independent variable in the multiple regression. Hence, I favor redoing the all of the multiple regressions and only reporting the models in which all of coefficients of the independent variables are statistically significant.

---Done.

5) Discussion.

a**) The authors repeatedly acknowledge the importance of head shape for affecting burrowing performance. Consequently, I was surprised by the absence of any variable quantifying head shape for the study taxa, especially since the biomechanical relevance of this seems much more obvious than that of the lateral angle of the head, which was quantified.

---We thank the reviewer for this suggestion. We had collected head shape data, but we originally opted to focus our message on body shape. We have taken the reviewer's suggestion and analyzed these data but found no effect of head shape on burying performance and kinematics. We feel that including these new analyses not only increases the length of the manuscript, but also distracts from our main focus. Thus, we have decided not to include these data, and avoid any discussion of the head.

b) (lines 186-196) The authors mention that the relative length of the tail varies substantially among elongate lizard taxa. A substantial difference between many elongate lizards and snakes is also the greater relative length tail in lizards. Furthermore, most burrowing snakes have quite short and blunt tails relative to their terrestrial counterparts. However, despite having snout vent length and tail length, no data were reported for relative tail length.

---We have added Table S2 which now includes the means and standard deviations of tail length each species. We also have written an interpretation of our pPC results that addresses tail length [lines 125–128].

c) By the way, do any of these study taxa have tail autotomy? If so, did all of the individuals used in the study have original rather than regenerated tails? This should be stated explicitly. I also strongly favor including only intact individuals.

---All Lerista species are capable of tail autotomy. We now state this, and clarify that we only sampled individuals that had original tails or long, regenerated tails [line 82].

6) (Throughout the main manuscript and supplement) Given that some taxa did not use their limbs for surface locomotion, "running" seems to be a poor choice of terminology compared to "surface locomotion".

---We agree with the reviewer and have changed all appropriate instances of "running" to "surface locomotion".

7) Supplement.

a**) I strongly favor including a conventional rooted cladogram of the study taxa as a figure in the supplement.

---Done. Please see Figure S1.

b) A better color scheme should be possible for Fig S1. I was not confident I could distinguish between the two of the shades green (L. kendricki and L.connivens) and two of the shades of blue (L. macropishopus and L. dorsalis).

---We have changed the figure to better distinguish species using both colors and symbols. We also include species names. This figure is now Figure 2.

c) The authors should explicitly state if a value of zero was used for limb length when limbs were absent.

---Done.

d) I could not find anywhere close to an N =13 for the light blue symbols and dark blue squares in Fig. S3. Perhaps the author are trying to cram too much information in each panel of this figure. Since it is in the supplement anyway, I favor clarity over small size.

---We agree with the reviewer that the overlapping symbols made the figure more difficult to interpret. We have remade the figure (now Figure S7) such that each substrate is plotted in separate panels.

e) What is the definition of a “stump” limb?

---A ‘stump’ limb is one that lacks an autopodium. We now state this in the header for Table S1.

f) (Fig. S3) It seemed somewhat odd to me to include the seven species that that lacked front limbs (Table S1) in the analysis of the proportion of individuals that used the fronts limbs since front limb use is not possible in these species.

---We agree with reviewer, and in our revising decided to remove this analysis.

g) The heading for Table S5 refers to “penetration during burrowing”. However, two entries in the body of this table use the term “sand-swimming”. It is vitally important that the authors use consistent terminology for the different types of locomotion, which I believe include the initial penetration of loose substrates, sub-surface locomotion and surface locomotion.

---Done.

h) A few supplementary videos might be nice for illustrating the presence or absence of limb use or variation in how the limbs are used.

---Done. We have now included Supplementary figures S3–S5, which include a series of stills from representative videos included in the dataset.

i**) The multiple regression analyses in Tables S3-7 included both pPC1 and head length as independent variables. However, Table S2 indicates that head length was a factor in the PCA. Consequently, it seemed inappropriate to perform a multiple regression analysis that in some sense used the same variable (head length) twice.

---In our revision we have addressed this issue by re-doing all analyses using ln-mass as a measure of size.

Appendix B

Associate Editor Board Member
Comments to Author:
I thank the authors for responding constructively and exhaustively to the initial reviews of their manuscript and carrying out the additional analyses suggested by the expert reviewers. I agree with the reviewers that these modified analyses, and the other editorial and aesthetic changes, to the manuscript have greatly improved the paper. Both reviewers have suggested some very minor but reasonable additional modification to improve clarity and smooth out the rough edges on the manuscript. It is my recommendation that the authors be given the opportunity to resubmit a modified version to incorporate those changes.

Reviewer(s)' Comments to Author:

Referee: 1

Comments to the Author(s).
I have no read the revised version of the paper entitled 'The evolution of fossorial locomotion in the transition from tetrapod to snake-like in lizards. I found the revision improved and the replies to the comments of both reviewers adequate and careful. I only have two minor remaining remarks that I'd like to see addressed.

1) on line 21 the statement is ambiguous ... as you did not quantify subsurface locomotion but rather soil penetration velocity it would be good to be very clear about this in the abstract.

---This is a valid point raised by both reviewers. We have changed this line to read "…snake-like species penetrated sand substrates faster than lizard-like species…". In addition, we judiciously re-worded appropriate instances of "sand-swimming" (or variant) to "sand-penetration" (or similar) [e.g., lines 93, 99, 160, 164, 225, 241, Supplementary Materials].

2) on line 98, please state that this is a uni-axial force transducer. I agree that you data are likely comparable across taxa but for non-experts it is important that they are aware of the fact that you measured only one of the force components. Related it would be nice to add a sentence stating that it would be nice to validate these measures with 3D force plate recordings and/or Xray video.

---We have added "uniaxial" to line 100. We have also changed the terminology (from "1-component" to "uniaxial") in the supplementary methods for consistency. We also now state on line 219–221 that amphisbaenians exhibit three dimensional force exertion during burrowing bouts, and that it would interesting to investigate if sand-swimming is similarly three-dimensional.

Referee: 2

Comments to the Author(s).
I liked a lot that the first version of the manuscript had to offer, and the revised version is even better still. I would like to compliment the authors on their thoroughness in responding to my suggestions, virtually all of which have been dealt effectively. I found the new supplementary tables very helpful and I especially liked the new supplemental figures that showed the lizards penetrating sand and moving through the cluttered substrate. I have only a few very minor points detailed below that I believe merit attention.

1) Locomotion in sand.

a) It is very clear in the response to reviews that the authors studied the initial penetration of sand (time taken to bury the body up to the pelvic girdle) rather than undulatory locomotion beneath the surface of the sand when the body is completely submerged (sand-swimming). In my opinion some very significant differences exist between initial penetration of sand and "swimming" beneath the sand. Hence, wouldn't it clearest and most intuitive simply to use "sand penetration" rather than "sand swimming"?

---Reviewer 1 had a similar concern, and we agree that "sand-penetration" is a more accurate description of the data we collected. Therefore, we have judiciously changed "sand-swimming" to "sand-penetration" throughout the manuscript [e.g., lines 93, 99, 160, 164, 225, 241, Supplementary Materials].

b) Part of my logic for the comment above is that some species of lizards with well-developed legs and impressive sprinting speeds on the surface of sand are also proficient at penetrating the surface of the sand. For example, fringe-toed lizards (SVL=7-10 cm) can bury themselves up to their pelvic girdle in approximately 1.6 s (Jayne & Daggy, 2000, J. Exp. Biol.), and I believe this compares quite favorably to some of the faster penetration times for the more snake-like species studied by the authors (Fig. 3). The fringe-toed lizards also have a rather stereotyped behavior in which both of the hind limbs are simultaneously retracted, and this appears important for providing a purchase from which the lizards can forcibly push forward and penetrate the sand. I do not dispute that within the clade of lizards studied by the authors more snake-like lizards had superior sand penetration. However, at least for fringe-toed lizards, well developed limbs and other aspect of morphology and behavior can result in superior penetration performance than many of the reduced limb species studied by the authors. I was uncertain whether or not the authors were aware of this, and I leave it to them to decide if this is worth a very brief mention in their manuscript.

---We agree with the reviewer that this is an interesting consideration. We now state at the end of the discussion [lines 242–244] that the evolution of sand-swimming behavior itself is likely a result of historical contingency because short-bodied, robustly limbed taxa such as *Uma scoparia* and *Scincus scincus* also engage in sand-swimming and perform at similar levels.

2) (lines 47-48) Squamates = lizards + snakes + amphisbaenians not just lizards and snakes.

---Done.

3) Line 150 cites non-existent Table 4S. After this citation, many of the other citations appear to be incorrect for how the supplementary Tables are presently numbed. Hence, the authors should update the table numbers and carefully double check all of the citations.

---Sorry for the error; we neglected to insert "Supplementary" in a few cases in the main manuscript and have now corrected and double checked this.

4) Because I found this paper so interesting, I really liked having many of the additional details provided in the new supplementary materials which allowed me to relate some data to a particular species. Unfortunately, in the current scheme I realized that I could not find sand penetration velocities for the species that had hind limbs but no forelimbs. Try as I might, I could not envision some system of unique symbols for each species in all plots of the functional data (Figs. 3, S6, S7) because of the admirably large

number of study species. Hence, I believe the easiest fix for this would behave a supplementary table with the all of the mean and values of functional data for each of the species (Figs. 3, S6, S7). I apologize for not catching this in the first round of comments.

---We thank the reviewer for taking such keen interest in our work. We agree that this information could be useful to readers and have now added mean±standard deviations of performance variables as well as the proportions of individuals that used hindlimbs as Supplementary Table 3, and reference it in the methods [line 118].

Once, congratulations on a fine study! I wish more of the other papers I have been asked to review were this interesting and well done.