<div align="center">**Peer Review File**</div>

Manuscript Title: Subaqueous foraging among carnivorous dinosaurs

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

This paper examines the ecology of dinosaurs and critically tests the hypothesis that some dinosaurs have aquatic (or at least semi aquatic) habits. The authors present a clear hypothesis, novel analyses with a novel dataset, integrate previous arguments and finding about aquatic dinosaurs, and present the best argument to date that at least some dinosaurs were partially aquatic, the spinosaurids. They also test other dinosaurs that have been thought to be aquatic, but their methods suggest that they were not.

Overall, the paper accomplishes its goals focused on testing the aquatic spinosaurid hypothesis and the approach is novel; this dataset is important and the analyses can be easily repeated. The discussion and conclusion follow from the analyses with one exception (diversification of spinosaurids and timing of aquatic-correlated changes; see comments). I think their arguments are solid and can be repeated, added to, etc. The paper includes all classic and the most-recent publications hence is 'state-of-the-art'. The phylogenetic analysis and interpretations are well constructed and clear; others may disagree but this hypothesis is repeatable and transparent.

Most of my comments are minor and can be addressed easily. All are directly in comments in the provided word doc. In sum, the authors could be specific about 'aquatic life' categories up front, the term bone-compactness needs some revision for clarity of what was being measured, the correlation with Karroo-Ferrar needs some more careful consideration, and adult vs skeletally mature needs to be defined better in the supplement when discussing ontogenetic age through histology. Finally, the data availability needs to be clearer, especially concerning the images of histology slides and where the CT datasets are available from.

The paper only needs minor revision after my points are considered.

Minor
Line 26 - evolved multiple times - Give an estimate. Too vague.
Line 40-42 – 'acquisition of adaptation to aquatic environments followed the faunal turnover and diversification of tetanuran theropods initiated by the emplacement of the Karoo-Ferrar large igneous province during the late Early Jurassic' - The divergence estimate and correlation with the Karoo-Ferrar is OK, but the acquisition of aquatic environments correlation is not. These could have evolved from the divergence up to the Cretaceous.
Line 46 - life in water - I see that you are keeping this vague here, but I think it would be much more useful to show the range of aquatic lifestyles upfront. You will have people arguing for years (likely

will anyway) about the aquatic modes if they are not better defined upfront. I think this is a service for the community and can be followed from this work.

Line 53 – believed ◊ hypothesized

Line 71 - aquatic habits - Similar comment as above. Be a bit more specific here or refer to your categories earlier

Line 73 - body plan - I would suggest adding 'over millions or tens of millions of years' to give a temporal context.

Line 75 - gradually - Add a range of time here.

Line 76-78 - As such, many aquatic taxa possess few anatomical indicators of water-related ecology, and instead share numerous traits with land animals - You might want to add something about the power of fossil record records these changes, but the ecological – anatomical transformations are not necessarily clear, only the end members.

Line 79 - considered to have been terrestrial - Something to think about or mention – phylogenetic baggage of previous habits.

Line 80-82 - early stages of an evolutionary transition towards aquatic ecologies, or amphibious animals that evolved relatively limited anatomical transformations despite spending much of their lives in water. - This sentence assume that the lineage will hit 'full aquatic habits'. Please modify to include a statement that some may 'end' at certain stages.

Line 85 - bone compactness - I think it is a bit misleading to sweep your measures into the term 'bone compactness'. You are including density of the cortex and the amount of cortex (approaching the center) into bone compactness here. This is important because both have been used to infer aquatic ecologies for amniotes. Please be careful. I would make sure these measures are clearly delineated throughout the paper.

Line 101 - bone compactness - See comment above.

Line 104-111 - Disparate ecological niches are herein recognized among spinosaurids: whereas Baryonyx and Spinosaurus are strongly predicted as subaqueous foraging species, Suchomimus is inferred as a more terrestrial, potentially wading animal, although sharing morphological similarities with Baryonyx. Our findings greatly expand the ecological disparity of non-avian dinosaurs. Furthermore, the appearance of aquatic adaptations among spinosaurids is linked to a major burst in evolutionary diversification of the here-recovered clade Carnosauria that followed the emplacement of the Karoo-Ferrar large igneous province during the Jurassic - This is all 'abstract' or discussion and con be deleted here. I know this is a common practice, but it is just repetition.

Line 128 - modern crocodilians and diving birds ◊ living crocodylians and various clades of

Line 187 - adult specimen - Specimen # please, please define adult or use skeletally mature (see long review by Griffin et al. in Biological Reviews on growth)

Line 188 - subadult – minor change here and in the supplement. Please define subadult or use skeletally immature or where the animal is on the growth curve (see long review by Griffin et al. in Biological Reviews on growth).

Line 199 – an unexpectedly ◊ a

Line 234-241 - Spinosaurids were part of the rapid radiation of Tetanurae following the emplacement of the Karoo-Ferrar large igneous province during the late stages of the Early Jurassic16,17. Increased diversification appeared in concert with morphological innovation and high

rates of homoplasy across vertebrate clades, including tetanurans17. Exploration of ecological regimes may have facilitated ecomorphological radiation linked to the invasion of previously uninhabited environments. Among these major ecomorphological radiations, multiple dinosaur clades independently evolved aerial capabilities32. - I like this section, but you must be a bit more conservative here. The aquatic adaptations may have come well after diversification given that the spinosaurid taxa that you use are Cretaceous. If there are other spinosaurids with similar aquatic adaptations that are much more fragmentary, specifically cite them here or say that the onset of ecological shift(s) are in a specific time interval, not as a consequence of Karoo.

Line 234-235 - following the emplacement of the Karoo-Ferrar large igneous - This is correlated with the timing, so please be careful here; it is not a causation, or at least the evidence is pretty weak.

Line 248 – density - Refer to supplement here.

Line 250 - To quantify bone compactness, - I would really like to see and 'element and evaluation method' table to summarize this information specifically for the spinosaurids given that this is new information. Please include element, side, specimen # and method.

Line 259 – while ◊ whereas

Line 296 - thin sectioning - Were these all imaged by your team? If so, it would be important to provide the image as part of your data availability statement.

Line 296 - micro-CT scanning - Please provide where EACH dataset can be obtained in your excel file. These should all be publicly available or made available.

Baryonyx – should be a NHM number not BMNH

Referee #2 (Remarks to the Author):

Review of Fabbri et al, 2021:
Subaqueous foraging and wading behavior among carnivorous dinosaurs

Thank you for the opportunity to review the manuscript submitted by Fabbri and others, with the title above.

This manuscript investigates subaqeous foraging behavior among carnivorous dinosaurs using a novel approach combining measurements of osteosclerosis analyzed by statistical means of ecomorphological inference. With this approach, the authors build a broadly sampled dataset of bone cortical thicknesses across living amniotes with diverse foraging strategies and use this to make robust inferences as to the habits of extinct dinosaurs (both non-avian and avian). From this set of inferences, the authors focus most of their paper on their finding that two spinosaurid taxa show clear, significant adaptations for sub-aqueous foraging.

The broadest impact of the manuscript is the comparative dataset and subsequent analysis. The authors are indeed correct when they point out that while bone compactness has been used on a taxon-by-taxon basis as a means of assessing foraging habits, that a comparative framework for understanding how osteosclerotic changes are related to activity pattern changes is lacking. Therefore, I think that their research in this particular area is pioneering and will be widely cited by

researchers using the comparative method in palaeontology, by those seeking to derive more information from osteohistology, and by those studying aquatic adaptations in a range of extinct organisms (in fact, dinosaurs may be the least important of these!).

The more discipline-specific impact of this paper is its contribution to a rampant debate about the habits of spinosaurids. More than a dozen papers over the last decade have investigated some aspect of spinosaurid palaeobiology, with some studies reaching diametrically opposed conclusions about their aquatic habits, based on different means of assessment (e.g., buoyancy vs isotopes vs cranial geometry vs axial biomechanics). Fabbri et al.'s finding presents a statistically defensible means of answering this question from a completely novel perspective – that is, while the debate about spinosaur habits is not in itself novel, the finding presented here involves a novel source of data, setting this manuscript apart from others on the topic. My impression is that this issue has a broad public audience as well as a more focused scientific audience.

Based on these two merits, I think that the manuscript is worth publication in your journal. However, I have a number of concerns about the manuscript which I lay out below, in no particular order.

1. Allometry. One of my immediate concerns about bone compactness was its relationship with organismal size. I think that many of the scientific readers of this paper will have similar concerns in the back of their minds. However, my concerns were alleviated when I found abundant evidence (in the form of the statistical model tests, and the more qualitative assessment of supplemental figure 4) that allometric changes are a poor explainer of bone compactness differences among the sampled taxa. However, I found little treatment of this in the main text, and I would urge the authors to make more explicit, reasoned arguments against allometry.
a. Also on this topic: while the main figures look nice, I found the black and white femoral slices compared between foraging groups to be a more convincing explanation of the differences between subaqueous foragers and other categories. The first in-text figure does not provide that comparative sample set, but instead focuses only on spinosaurids. Researchers interested in citing the paper for non-dinosaurian foraging strategy research will undoubtedly get more out of a main-text figure with a broader array of comparisons, in the style of several of the convincing figures in the supplement.
b. More on allometry: Another observation I think clearly shows that allometry isn't a good explanatory variable is the stark cross-sectional differences when looking in comparison between Spinosaurus/Baryonyx and the comparable-massed Tyrannosaurus rex. That's just an observation and the authors can make of it what they want.
2. Deep diving versus shallow diving. A casual reader or a person poorly versed in diving categories might be confused or remain unconvinced by the massive differences between the cross-sectional bone geometries of deep diving animals (e.g., many cetaceans, ichthyosaurs that are full of spongiosa) and shallow, subaqueous foragers (e.g., manatees and crocodilians, where the bones have thick cortices plus spongiosa). I found these striking, and only discovered an explanation of these differences in the supplement. A pass through the dataset shows that the global compactness values of these types of taxa are more in line with those of non-aquatic animals (although stem members like Basilosaurus show heightened bone compaction), further confusing me. It might be worth explaining these differences in the main text – for example, you can certainly state that the trabecular morphology of spinosaurids resembles shallow divers such as crocodilians rather than obligate aquatic deep divers such as ichthyosaurs and whales. In fact, perhaps deep diving could be

its own category in some of your statistical tests?

3. Graviportalism as a locomotor category. I was struck by how a graviportal taxon like Antetonitrus fell neatly in between the cross sectional geometry of T rex and Spinosaurus. And graviportal postures certainly must represent a unique loading regime that clearly results in stereotyped osteosclerosis, as you indicate in your citation of relevant mammalian literature. Did you ever consider graviportality as a categorical variable in your various analyses? Its signal seems quite strong from a qualitative viewpoint, and it might be important to show that you can differentiate it from aquatic taxa.

4. Utility of quadratic DFA. I am not familiar with the use of quadratic DFA, but I'm wondering as to the utility of this approach. It doesn't seem to incorporate phylogenetic information explicitly (especially when compared to pFDA) and given the high phylogenetic signal in the dataset, I'd be worried interpreting its results without necessary caveats about shared ancestry. It's also difficult to fully understand the output image you show in Supplemental Figure 8 (more explanation would help). Moreover, the pFDA analysis gives similar results, incorporates phylogeny, and involves a straightforward output table that is easily interpretable (and which has been used in several recent papers where the comparative method is applied to palaeontological questions). But see my comment next about the sample set of the pFDA.

5. Inclusion of samples in pFDA: I found your statement that "We excluded non-archosaurian taxa from the training set for this analysis as many of them are subaqueous-foragers and as such they may unbalance the phylogenetic signal" (main text, pg. 16, line 358 and on) to be curious. Do you mean that, e.g., you had different sample sizes of sub-aqueous vs non-diving animals in archosaurs versus mammals? Different sample sizes of categories have long been realized to cause issues for DFA and similar analyses, however these can either be addressed by subsampling (e.g., method employed by Chapelle et al., 2020 for bipedal vs quadrupedal locomotor categories) or by the insertion of prior category membership proportions. If you didn't mean that, could you explain why you made this statement? It seems to act to undermine the general applicability of your results as it stands…

6. Please correct typos in taxon names in supplemental worksheet

a. E.g., "Ichtyosaurus"

7. Lack of informative clade labels on supplemental trees. While the phylogenetic work of this paper is secondary to the diving hypothesis, some researchers will undoubtedly cite it for its systematic findings. Can you please provide trees with explicit labels showing the positions of major clades that you recover? And, if possible, typical support values?

8. Unsupported, vague, or inefficient statements in body text: The main text contains several passages that seem out of place with the comparative biological approach of this paper. These make unsubstantiated claims or read as hyperbolic. They can be amended or deleted without making any change in the significance of your findings. For example:

a. "This finding reveals spinosaurids to be aquatic specialists with surprising ecological disparity, including subaqueous foraging behavior in Spinosaurus and Baryonyx and potentially wading habits in Suchomimus" (main text, page 2, line starting at 37).

i. You examine three spinosaurids, and find that two of them do the same thing and one does not. Is that "surprising ecological disparity" in the group? I find that hard to assess.

ii. "Potentially wading habits" is really not supported by your analysis at all. Your novel data consists of global bone compactness indices – and on those indices Suchomimus looks identical to many non-waders, for example Rhamphorhynchus. So the inference of wading is based on prior data that has

been well-published. I would just indicate that you firmly predict Baryonyx and Spinosaurus as shallow divers, and surprisingly that Suchomimus lacks such osteosclerotic adapations, suggesting it has secondarily reverted to non-diving life habits (while acknowledging, as you do later on, that other parts of the skeleton suggest at least some affinity to water).

b. "The divergence of spinosaurids and their acquisition of adaptation to aquatic environments followed the faunal turnover and diversification of tetanuran theropods initiated by the emplacement of the Karoo-Ferrar large igneous province during the late Early Jurassic16,17, which led to a remarkable increase in ecomorphological disparity among predatory dinosaurs." (page 2, lines starting at 39).

i. The K-F LIP and the divergence of the spinosaurids you mention and include on your tree are separated by at least 40mya. I find it hard to make a compelling argument that these form a causal linkage. This argument is perpetuated elsewhere in the paper, with relatively little rhetorical support. I'd suggest deleting it altogether, or instead providing definitive statements upon which it can be evaluated as a hypothesis.

c. "Our findings greatly expand the ecological disparity of non-avian dinosaurs. Furthermore, the appearance of aquatic adaptations among spinosaurids is linked to a major burst in evolutionary diversification of the here-recovered clade Carnosauria that followed the emplacement of the Karoo-Ferrar large igneous province during the Jurassic16,17. (page 5, lines starting 106).

i. Your findings expand the ecological niche space of non-avian dinosaurs by exactly one category, which is already in the literature as a fairly well-developed hypothesis. So I would avoid making this statement, it seems hyperbolic. The importance of your finding is rather that a novel method, that has statistical power, has resolutely supported the notion of diving in some spinosaurids, resolving a major controversy in favor of an expanded niche.

ii. What do you mean by a "major burst in evolutionary diversification"? I don't understand if this is in terms of lineages/per time or in terms of ecospace of some sort? Perhaps I missed this in the text and the supplement, but if not then I think this should be explicitly stated and if necessary quantified. Or you could just leave this sort of statement out – it would not detract from the study in any way.

d. "In conclusion, our data reveal that Spinosauridae, a geographically widespread clade of predatory dinosaurs, was ecomorphologically adapted to life in water. Our findings greatly increase the ecological disparity of non-avian dinosaurs, further challenging the longstanding hypothesis that this group was restricted to terrestrial environments. Spinosaurids were part of the rapid radiation of Tetanurae following the emplacement of the Karoo-Ferrar large igneous province during the late stages of the Early Jurassic16,17. Increased diversification appeared in concert with morphological innovation and high rates of homoplasy across vertebrate clades, including tetanurans17. Exploration of ecological regimes may have facilitated ecomorphological radiation linked to the invasion of previously uninhabited environments. Among these major ecomorphological radiations, multiple dinosaur clades independently evolved aerial capabilities32. Our study demonstrates that ecomorphological radiations among non-avian dinosaurs also included the invasion of aquatic systems." (page 10, beginning line 226)

i. Please see previous comments about ecological disparity. You might even consider that since you show that ornithomimosaurs and haszkaraptorines are not divers, your analysis has reduced the disparity!

ii. You present no data supporting the rapid tetanuran radiation post K-F LIP emplacement. So why are you concluding that you did?

iii. Where is the data supporting high rates of homoplasy across vertebrate clades? What does this mean in the context of this paper and why is it important?

iv. What is an ecomorphological radiation in the context of exploration of ecological regimes? You're describing three taxa and a single theropod aquatic "invasion". While I do understand that sampled fossils really represent a much greater amount of evolutionary innovation and cladogenesis, these statements are so vague as to be nonsensical. I'd just leave them out entirely. Again, this will not diminish your work.

Sincerely,

Jonah Choiniere

Referee #3 (Remarks to the Author):

The study uses a broad survey of bone compactness metrics among archosaurs to predict aquatic behavior in non-avian dinosaurs. The key results of the work are 1) that subaqueous foraging behavior in extant archosaurs can be predicted by bone compactness with a high degree of accuracy, and 2) strong support is found for aquatic ecology in spinosaurids, a subject of much recent interest and controversy. The analysis is rigorously performed and well documented, and the writing and figures are excellent. The conclusions are original and will be of great interest to people who study dinosaur ecology. More broadly, the refinement of methods to predict behavior from bone compactness will be useful for people who study paleoecology and evolutionary biomechanics. My only major suggestion is that the limitations of the approach and how the data should be interpreted could be made clearer. This would not only avoid misinterpretation but also increase the manuscript's impact by making the implications obvious to a broader audience.

Introduction and literature: The literature cited seems appropriate, but previous studies have used different definitions or terminology, and some clarification would be helpful for those outside the field of ecology. For example, Quemeneur et al. 2013 used aquatic, semi-aquatic, and terrestrial as categories. How do these categories relate to probability of subaqueous foraging? Why is probability of subaqueous foraging used instead in this study – because it is more specific and/or more closely linked with bone compactness, or because it is a behavior rather than an ecological category? It seems that in this study anatomical variables other than bone histology are used to infer aquatic habits. Please clarify the relationship between aquatic/semi-aquatic, diving, and subaqueous foraging, ideally with examples from extant taxa.

Data & methodology: The approach has been validated in this and previous studies. Although almost all the extant taxa in this study are birds, previous studies have shown similar results for lepidosaurs and mammals. The analysis has been carefully chosen to separate the phylogenetic signal in the data from a functional signal. The data are comprehensive and very well presented and the methods are sufficient for replication. One suggestion: I don't quite get the connection between Liem's paradox and the inclusion of flying as a variable - I thought Liem's paradox was that highly specialized taxa are also "jacks of all trades." Clarification would be appreciated.

The results support the conclusions. I'm a bit concerned that the limitations of interpretation are not clear enough and that readers may misinterpret or over-interpret the results. As I understand it, this

method cannot be used to distinguish between animals that feed aquatically but do not submerge (e.g., storks) and completely terrestrial animals, nor between deep-diving fully aquatic animals (e.g., cetaceans) and semiaquatic animals that forage subaqueously (e.g., ducks). Presumably this is because the relationship between bone compactness and aquatic lifestyle/behavior beyond two or three broad categories is not well understood (as pointed out by previous studies). However, in the supplementary information a distinction is made between the histology of "deep divers" such as ichthyosaurs and cetaceans and swimmers in shallow water, which include Spinosaurus (supp. Fig. 2), and other studies have discussed qualitative differences in bone histology between taxonomic groups and types or degrees of aquatic adaptation (Alexandra Houssaye, P. Martin Sander, Nicole Klein, Adaptive Patterns in Aquatic Amniote Bone Microanatomy—More Complex than Previously Thought, Integrative and Comparative Biology, Volume 56, Issue 6, December 2016, Pages 1349–1369, https://doi.org/10.1093/icb/icw120 and references therein). Could the writers comment on this, and perhaps how future studies might go about distinguishing between different degrees of aquatic adaptation and/or different types of aquatic locomotor or feeding behaviors (e.g., looking at differences in compactness index of vertebrae, ribs, and humerus where available)? This is set up by paragraph 1 of the introduction which discusses various types and degrees of aquatic adaptation found in extant animals, but it is not revisited in the light of the results (other than Suchomimus). Also, how should one interpret a result like the 34% probability of subaqueous foraging in Megalosaurus: as an inconclusive result or an indication of semi-aquatic adaptation? I understand the limitations of interpreting fossils with unique anatomical characteristics (223-224), but I think it could be valuable to hear the authors' educated opinions on possible ecological interpretations of spinosaurids.

Referee #4 (Remarks to the Author):

It's not entirely clear to me what the conclusions of this paper are, because the fundamental language used is unclear – what exactly _is_ subaqueous foraging? Is it completely submerged, swimming pursuit? Being an ambush predator lying in wait under the water? Does it extend as far as wading in deep water, and the subaqueous applies only to that being foraged? Are crocodiles and sharks both "subaqueous foragers"? The reason this matters is because it directly affects the novelty of the study – it is well established that Spinosaurus is a piscivore (e.g. 1,2,3,4,5), and therefore probably spent time in the water, and certainly foraged from under the water's surface. What is contentious (2,3,4,6), is whether it actively swam in pursuit of prey, or was an ambush predator, either wading or partially/fully submerged. This paper does not answer that, even though its primary conclusion is that "Spinosauridae… [were] ecomorphologically adapted to life in water". I don't know if the authors want to hedge their bets, or if their data simply cannot discriminate between types of aquatic lifestyles, but in this regard, there is little novel about the conclusions.

The second major part of the study is in showing that bone compactness is increased in aquatic (sub-aqueously-foraging) animals. Again, this is not particularly novel (7, 8, 9, 10), though admittedly most previous work has been on mammals.

That being said, the methods and data are an excellent resource, and clearly contribute to the literature. My biggest concern is that only a single bone, the femur, is used in the study. Is this

always going to be the most representative of lifestyle? A number of waders are listed (ln 134-137), but all of these are relatively small, and also fly. Yes, subsequent models incorporate flight into the relationship with bone compactness, but if Spinosaurus were a wader, would we perhaps expect dense hind-limb bones, but typical theropod pneumatisation elsewhere? This could presumably be tested through histology/CT scanning of other bones in the skeleton, and seeing if they are as compact. The authors could also look at other bones (e.g. vertebrae) from their sample of 141 taxa and explore if other parts of the skeleton are more or less indicative of a subaqueous lifestyle. This of course is difficult to apply to Spinosaurus, because no matter how large the sample size of extant taxa (or extinct taxa of conclusive aquatic affinity), there is nothing closely analogous to Spinosaurus (very large, potentially wader/shoreline predator, potentially bipedal on land) – which is presumably why the study was carried out with that animal as the focus in the first place.

In summary: I think the methods are great, but the conclusions are not particularly novel or interesting.

Minor comments:

Line 119: use of (1) and (2) for two variables that have 0, 1 and 2 might be confusing, perhaps use letters and numbers: "taxa were scored using two categorical explanatory variables that encode the presence of (a) subaqueous foraging (0, unable, 1, able…"
Line 121: Would "two independent variables" not be simpler than "two independently varying variables"?
Line 193: do these results point to all Spinosaurids as being aquatic? If so, how can only some of them be subaqueous foragers?
Data availability: Please provide direct links to all CT data used, in accordance with standard open data practice (11)

References:
1. Buffetaut, E., 1989. New remains of the enigmatic dinosaur Spinosaurus from the Cretaceous of Morocco and the affinities between Spinosaurus and Baryonyx. Neues Jahrbuch für Geologie und Paläontologie-Monatshefte, pp.79-87
2. Ibrahim, N., Sereno, P.C., Dal Sasso, C., Maganuco, S., Fabbri, M., Martill, D.M., Zouhri, S., Myhrvold, N. and Iurino, D.A., 2014. Semiaquatic adaptations in a giant predatory dinosaur. Science, 345(6204), pp.1613-1616.
3. Gimsa, J., Sleigh, R. and Gimsa, U., 2016. The riddle of Spinosaurus aegyptiacus' dorsal sail. Geological Magazine, 153(3), pp.544-547.
4. Hone, D.W. and Holtz Jr, T.R., 2021. Evaluating the ecology of Spinosaurus: Shoreline generalist or aquatic pursuit specialist?. Palaeontologia Electronica, 24(1), p.a03.
5. Cuff, A.R. and Rayfield, E.J., 2013. Feeding mechanics in spinosaurid theropods and extant crocodilians. PLoS One, 8(5), p.e65295.
6. Henderson, D.M., 2018. A buoyancy, balance and stability challenge to the hypothesis of a semi-aquatic Spinosaurus Stromer, 1915 (Dinosauria: Theropoda). PeerJ, 6, p.e5409.
7. Wall, W.P., 1983. The correlation between high limb-bone density and aquatic habits in recent mammals. Journal of Paleontology, pp.197-207.
8. Fish, F.E. and Stein, B.R., 1991. Functional correlates of differences in bone density among

terrestrial and aquatic genera in the family Mustelidae (Mammalia). Zoomorphology, 110(6), pp.339-345.

9. Gray, N.M., Kainec, K., Madar, S., Tomko, L. and Wolfe, S., 2007. Sink or swim? Bone density as a mechanism for buoyancy control in early cetaceans. The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology: Advances in Integrative Anatomy and Evolutionary Biology, 290(6), pp.638-653.

10. Taylor, M.A., 2000. Functional significance of bone ballastin in the evolution of buoyancy control strategies by aquatic tetrapods. Historical Biology, 14(1-2), pp.15-31.

11. Davies, T. G. et al. (2017) 'Open data and digital morphology', Proceedings of the Royal Society B: Biological Sciences, 284.

Author Rebuttals to Initial Comments:

We are grateful for the insightful and thorough revisions received by the reviewers, which improved and strengthened the results and impact of our manuscript.

We followed the majority of reviewers' comments. We significantly expanded the femur dataset (which now includes 206 entries and 200 taxa *contra* the original version of the dataset that had only 141) and assembled a *de novo* dataset using dorsal ribs (174 taxa, including the spinosaurids *Baryonx* and *Spinosaurus*). The taxonomic overlap between the two dataset is equal to 83 (39.8% and 46.85% of the femur and dorsal rib dataset, respectively). Unfortunately, restrictions to travel, facilities and collections did not allow us to expand the taxonomic overlap between the two datasets. On the other hand, our results show that the patterns recovered from our analyses are not affected by taxonomic sampling bias. The choice of dorsal ribs as an additional skeletal element to test for subaqueous foraging among dinosaurs was based on the availability of these elements in the taxa of interest (*Baryonyx* and *Spinosaurus*) and on the fact that this skeletal element was previously observed to be informative to infer ecological adaptations in other clades (such as mammals). Our results are now based on 380 data points and 292 taxa, almost triplicating the original dataset we submitted with the first version of the manuscript.

The femur and dorsal rib datasets were analyzed separately to infer if 1) these elements are valid proxies for inferring ecologies among extant amniotes and 2) to then extrapolate ecological adaptations among non-avian dinosaurs. As suggested by the reviewers, we only focused on analyses taking into account phylogenetic correction. Additionally, the informal supertrees used for the analyses were time calibrated not only for the tips of the branches, but also for the deeper nodes. Our results reveal that the best model is the correlation between bone compactness and subaqueous foraging in both datasets. *Spinosaurus* and *Baryonyx* are found to be subaqueous foragers, while *Suchomimus* and all other non-avian dinosaurs were recovered as non-divers, in agreement with our original results. We added new figures and tables to illustrate the novel results obtained by analyzing the new datasets. Figures previously included in the supplementary information were moved to the main text as extended data figures. The main text was changed accordingly to reviewer comments, as well as new results and analyses.

We hope that the reviewers will be satisfied with our efforts to improve the manuscript and further strengthen our results and interpretations.

Referees' comments:

Referee #1 (Remarks to the Author):

This paper examines the ecology of dinosaurs and critically tests the hypothesis that some dinosaurs have aquatic (or at least semi aquatic) habits. The authors present a clear hypothesis, novel analyses with a novel dataset, integrate previous arguments and finding about aquatic dinosaurs, and present the best argument to date that at least some dinosaurs were partially aquatic, the spinosaurids. They also test other dinosaurs that have been thought to be aquatic, but their methods suggest that they were not.

Overall, the paper accomplishes its goals focused on testing the aquatic spinosaurid hypothesis and the approach is novel; this dataset is important and the analyses can be easily repeated. The discussion and conclusion follow from the analyses with one exception (diversification of spinosaurids and timing of aquatic-correlated changes; see comments). I think their arguments are solid and can be repeated, added to, etc. The paper includes all classic and the most-recent publications hence is 'state-of-the-art'. The phylogenetic analysis and interpretations are well constructed and clear; others may disagree but this hypothesis is repeatable and transparent.

Most of my comments are minor and can be addressed easily. All are directly in comments in the provided word doc. In sum, the authors could be specific about 'aquatic life' categories up front, the term bone-compactness needs some revision for clarity of what was being measured, the correlation with Karroo-Ferrar needs some more careful consideration, and adult vs skeletally mature needs to be defined better in the supplement when discussing ontogenetic age through histology. Finally, the data availability needs to be clearer, especially concerning the images of histology slides and where the CT datasets are available from.

The paper only needs minor revision after my points are considered.

Thank you for your thorough revision. We incorporated all requested changes.

Minor
Line 26 - evolved multiple times - Give an estimate. Too vague.

We specified that this happened more than 30 times through time. Uncertainty in a proper quantification is related to phylogenetic uncertainty of deep divergences among reptilian clades.

Line 40-42 – 'acquisition of adaptation to aquatic environments followed the faunal turnover and diversification of tetanuran theropods initiated by the emplacement of the Karoo-Ferrar large igneous province during the late Early Jurassic' - The divergence estimate and correlation with the Karoo-Ferrar is OK, but the acquisition of aquatic environments correlation is not. These could have evolved from the divergence up to the Cretaceous.

We specified that, based on the available data, aquatic adaptations among spinosaurs probably appeared in the Early Cretaceous. We additionally deleted our suggestion of a link with the emplacement of the Karoo-Ferrar large igneous province in order to satisfy the requests by reviewer 2.

Line 46 - life in water - I see that you are keeping this vague here, but I think it would be much more useful to show the range of aquatic lifestyles upfront. You will have people arguing for years (likely will anyway) about the aquatic modes if they are not better defined upfront. I think this is a service for the community and can be followed from this work.

We introduced the three main habits that will be discussed in the manuscript, namely wading behavior (meant as shoreline specialists and/or only partially submerged habit), subaqueous foraging (defined as fully submerged active behavior), and deep diving

Line 53 – believed ◊ hypothesized

Changed

Line 71 - aquatic habits - Similar comment as above. Be a bit more specific here or refer to your categories earlier

We specified subaqueous foraging and deep diving, as these are the major categories that require alteration of the body plan to achieve buoyancy control and active swimming

Line 73 - body plan - I would suggest adding 'over millions or tens of millions of years' to give a temporal context.
We added "over millions or tens of millions of years" after line 75 to better link "gradually" with a time range

Line 75 - gradually - Add a range of time here.

We added "over millions or tens of millions of years"

Line 76-78 - As such, many aquatic taxa possess few anatomical indicators of water-related ecology, and instead share numerous traits with land animals - You might want to add something about the power of fossil record records these changes, but the ecological – anatomical transformations are not necessarily clear, only the end members.

We added "available through the fossil record" to suggest that such patterns are generally observable because of preservation of early diverging taxa.

Line 79 - considered to have been terrestrial - Something to think about or mention – phylogenetic baggage of previous habits.

We added "based on anatomical proxies and phylogenetic bracketing" to explain that such inferences might be related to anatomical studies or assumptions based on phylogenetic bracketing. A clear example of this are spinosaurids: previous studies have assumed similar ecologies among these taxa based on similar cranial morphology and phylogenetic bracketing based on their tetanuran affinities.

Line 80-82 - early stages of an evolutionary transition towards aquatic ecologies, or amphibious animals that evolved relatively limited anatomical transformations despite spending much of their lives in water. - This sentence assume that the lineage will hit 'full aquatic habits'. Please modify to include a statement that some may 'end' at certain stages.

Line 85 - bone compactness - I think it is a bit misleading to sweep your measures into the term 'bone compactness'. You are including density of the cortex and the amount of cortex (approaching the center) into bone compactness here. This is important because both have been used to infer aquatic ecologies for amniotes. Please be careful. I would make sure these measures are clearly delineated throughout the paper.

Line 101 - bone compactness - See comment above.

Line 104-111 - Disparate ecological niches are herein recognized among spinosaurids: whereas Baryonyx and Spinosaurus are strongly predicted as subaqueous foraging species, Suchomimus is inferred as a more terrestrial, potentially wading animal, although sharing morphological similarities with Baryonyx. Our findings greatly expand the ecological disparity of non-avian dinosaurs. Furthermore, the appearance of aquatic adaptations among spinosaurids is linked to a major burst in evolutionary diversification of the here-recovered clade Carnosauria that followed the emplacement of the Karoo-Ferrar large igneous province during the Jurassic - This is all 'abstract' or discussion and con be deleted here. I know this is a common practice, but it is just repetition.

Line 128 - modern crocodilians and diving birds ◊ living crocodylians and various clades of

Line 187 - adult specimen - Specimen # please, please define adult or use skeletally mature (see long review by Griffin et al. in Biological Reviews on growth)
Line 188 - subadult – minor change here and in the supplement. Please define subadult or use skeletally immature or where the animal is on the growth curve (see long review by Griffin et al. in Biological Reviews on growth).

Line 199 – an unexpectedly ◊ a

Line 234-241 - Spinosaurids were part of the rapid radiation of Tetanurae following the emplacement of the Karoo-Ferrar large igneous province during the late stages of the Early Jurassic16,17. Increased diversification appeared in concert with morphological innovation and high rates of homoplasy across vertebrate clades, including tetanurans17. Exploration of

ecological regimes may have facilitated ecomorphological radiation linked to the invasion of previously uninhabited environments. Among these major ecomorphological radiations, multiple dinosaur clades independently evolved aerial capabilities32. - I like this section, but you must be a bit more conservative here. The aquatic adaptations may have come well after diversification given that the spinosaurid taxa that you use are Cretaceous. If there are other spinosaurids with similar aquatic adaptations that are much more fragmentary, specifically cite them here or say that the onset of ecological shift(s) are in a specific time interval, not as a consequence of Karoo.

We specified that, although the origin of the spinosaur branch falls in the Jurassic, aquatic adaptations evolved in the Cretaceous for this group. This is a more conservative statement that better reflects our results and available data. Additionally, we deleted any connection to the Karoo-Ferrar province.
Line 234-235 - following the emplacement of the Karoo-Ferrar large igneous - This is correlated with the timing, so please be careful here; it is not a causation, or at least the evidence is pretty weak.

We deleted "following the emplacement of the Karoo-Ferrar large igneous province"
Line 248 – density - Refer to supplement here.

Done
Line 250 - To quantify bone compactness, - I would really like to see and 'element and evaluation method' table to summarize this information specifically for the spinosaurids given that this is new information. Please include element, side, specimen # and method.

We added Supplementary Table 1 to satisfy the reviewer's request.
Line 259 – while ◊ whereas

Changed
Line 296 - thin sectioning - Were these all imaged by your team? If so, it would be important to provide the image as part of your data availability statement.

The thin sections of *Spinosaurus* and *Suchomimus* were already imaged in Ibrahim et al. 2014-2020. The thin sections for *Condorraptor* (femur) and *Baryonyx* (dorsal ribs) were studied for this manuscript. Supplementary Figure 2 shows osteohistological characterization of these thin sections. Extended Data Figures 1-7 were prepared to make all the novel femoral and dorsal ribsections collected for bone density in this study freely available (these include black and white pictures of the sections). Note that these do not include taxa for which bone density was previously quantified in the literature (please, see the supplementary datasets to see which data were collected by our team and which ones were found in the literature). Additionally, all the CT scans collected for this study will be made freely available in Morphosource upon publication of the manuscript.

Line 296 - micro-CT scanning - Please provide where EACH dataset can be obtained in your excel file. These should all be publicly available or made available.
Baryonyx – should be a NHM number not BMNH

We will make all novel data freely available: the CT scan datasets (see list below) of modern and extinct amniotes (mammals and birds) will be uploaded in Morphosource (see our comment above),

We changed the acronym for *Baryonyx* throughout the whole manuscript.

List of the CT scans collected for this project (n=69):

Taxa

Alca_torda	Porzana_carolina
Anas_aucklandica	Psophia_crepitans_napensis
Anas_discors	Rhynchotus_rufescens_
Anas_erythroincha	Rynochetus_jubatus
Anhinga_anhinga	Sarothrura_rufa
Anser_fabalis	Spheniscus
Anseranas_semipalmata	Strigiops_habroptilus
Apteryx_owenii_	Sula
Aramus_guarauna	Tachybaptus
Atlantisia_rogersi	Tribonyx_mortieri
Baryonyx	Trybonix_ventralis
Caloenas_nicobarica	Turdus_olivator
Chauna_chavaria	Uria
Cinclus_cinclus	Xenicus_giliventris
Crypturellus_tataupa	Xenicus_longipes
Cygnus_olor	Chironectes_minimus
Eudromia_elegans	Dasyurus_maculatus
Eustreptospondylus	Desmana_moschata
Fratercula_arctica	Diplomesodon_pulchellum
Fregata_aquila	Euroscaptor_micrura
Gavia_immer	Hemiechinus_azurites
Glareola_pratincola	Lutra_vulgaris
Goura_cristata	Micropotamogale_euwenzorii
Megalosaurus	Monodelphis_domestica
Nestor_notabilis	Mustela_nivalis
Oxyura_jamaicensis	Noemys_fodies
Pelagodroma_marina	Ornithorhynchus_anatinus
Pelecanoides_urinatryx	Procavia_capensis
Phaeton_aethereus	Rhynchocyon_petersi
Phalacrocorax_carbo	Taxidea_taxus
Phalacrocorax_harrisi	Tenrec_ecaudatus
Phoenicopterus_roseus	Rattus_rattus
Podica_senegalensis	Microgale_thomasi
Porphyrio_Poliocephalus	Hydromys_chrysogaster
Porphyrio_hochstetteri	

Referee #2 (Remarks to the Author):

Review of Fabbri et al, 2021:
Subaqueous foraging and wading behavior among carnivorous dinosaurs

Thank you for the opportunity to review the manuscript submitted by Fabbri and others, with

nature portfolio

the title above.

This manuscript investigates subaqeous foraging behavior among carnivorous dinosaurs using a novel approach combining measurements of osteosclerosis analyzed by statistical means of ecomorphological inference. With this approach, the authors build a broadly sampled dataset of bone cortical thicknesses across living amniotes with diverse foraging strategies and use this to make robust inferences as to the habits of extinct dinosaurs (both non-avian and avian). From this set of inferences, the authors focus most of their paper on their finding that two spinosaurid taxa show clear, significant adaptations for sub-aqueous foraging.

The broadest impact of the manuscript is the comparative dataset and subsequent analysis. The authors are indeed correct when they point out that while bone compactness has been used on a taxon-by-taxon basis as a means of assessing foraging habits, that a comparative framework for understanding how osteosclerotic changes are related to activity pattern changes is lacking. Therefore, I think that their research in this particular area is pioneering and will be widely cited by researchers using the comparative method in palaeontology, by those seeking to derive more information from osteohistology, and by those studying aquatic adaptations in a range of extinct organisms (in fact, dinosaurs may be the least important of these!).

The more discipline-specific impact of this paper is its contribution to a rampant debate about the habits of spinosaurids. More than a dozen papers over the last decade have investigated some aspect of spinosaurid palaeobiology, with some studies reaching diametrically opposed conclusions about their aquatic habits, based on different means of assessment (e.g., buoyancy vs isotopes vs cranial geometry vs axial biomechanics). Fabbri et al.'s finding presents a statistically defensible means of answering this question from a completely novel perspective – that is, while the debate about spinosaur habits is not in itself novel, the finding presented here involves a novel source of data, setting this manuscript apart from others on the topic. My impression is that this issue has a broad public audience as well as a more focused scientific audience.

Based on these two merits, I think that the manuscript is worth publication in your journal. However, I have a number of concerns about the manuscript which I lay out below, in no particular order.

Thank you for your insightful comments. We modified the manuscript accordingly.

1. Allometry. One of my immediate concerns about bone compactness was its relationship with organismal size. I think that many of the scientific readers of this paper will have similar concerns in the back of their minds. However, my concerns were alleviated when I found abundant evidence (in the form of the statistical model tests, and the more qualitative assessment of supplemental figure 4) that allometric changes are a poor explainer of bone compactness differences among the sampled taxa. However, I found little treatment of this in the main text, and I would urge the authors to make more explicit, reasoned arguments against allometry.

We added additional statements in the results and discussion regarding the influence of allometry on our results and ecological inference of spinosaurids.

a. Also on this topic: while the main figures look nice, I found the black and white femoral slices compared between foraging groups to be a more convincing explanation of the differences between subaqueous foragers and other categories. The first in-text figure does not provide that comparative sample set, but instead focuses only on spinosaurids. Researchers interested in citing the paper for non-dinosaurian foraging strategy research will undoubtedly get more out of a main-text figure with a broader array of comparisons, in the style of several of the convincing figures in the supplement.

We added a comparison of femoral cross sections per different ecology in Figure 1. Moreover, we added 10 Extended Supplementary Figures in the main text to image all novel femoral and costal cross sections collected for this study. These include non-avian dinosaurs, living birds (including waders, flightless, and subaqueous foraging taxa), marine reptiles, aquatic and terrestrial mammals, and lepidosaurs.

b. More on allometry: Another observation I think clearly shows that allometry isn't a good explanatory variable is the stark cross-sectional differences when looking in comparison between Spinosaurus/Baryonyx and the comparable-massed Tyrannosaurus rex. That's just an observation and the authors can make of it what they want.

We labelled giant predatory dinosaurs, such as *Tyrannosaurus*, in Figure 2a in our previous version of the manuscript exactly for this reason. This was probably not enough to point out how large body size predatory dinosaurs. Therefore, we added a sentence in the main text explicitly stating such comparison between *Baryonyx* and *Spinosaurus* on the one hand, and *Tyrannosaurus, Tyrannotitan, Torvosaurus Suchomimus,* (femur dataset) and a large carcharodontosaurid (dorsal rib dataset) on the other. These taxa remain labelled in figures 2 and 3.

2. Deep diving versus shallow diving. A casual reader or a person poorly versed in diving categories might be confused or remain unconvinced by the massive differences between the cross-sectional bone geometries of deep diving animals (e.g., many cetaceans, ichthyosaurs that are full of spongiosa) and shallow, subaqueous foragers (e.g., manatees and crocodilians, where the bones have thick cortices plus spongiosa). I found these striking, and only discovered an explanation of these differences in the supplement. A pass through the dataset shows that the global compactness values of these types of taxa are more in line with those of non-aquatic animals (although stem members like Basilosaurus show heightened bone compaction), further confusing me. It might be worth explaining these differences in the main text – for example, you can certainly state that the trabecular morphology of spinosaurids resembles shallow divers such as crocodilians rather than obligate aquatic deep divers
such as ichthyosaurs and whales. In fact, perhaps deep diving could be its own category in some of your statistical tests?

As stated above, the figures aimed at visualizing bone tissue arrangement across femur diaphysis between these ecological categories, which were previously placed in the supplementary information, are now moved to the main text as extended data figures. In order to further show which taxa included in the analyses are graviportal or deep divers we 1) placed asterisks in figures 2a and 3a for visualization, and 2) added two tables (Supplementary Tables 5 and 6) to list such taxa. We did not add a new category for

3. Graviportalism as a locomotor category. I was struck by how a graviportal taxon like Antetonitrus fell neatly in between the cross sectional geometry of T rex and Spinosaurus. And graviportal postures certainly must represent a unique loading regime that clearly results in stereotyped osteosclerosis, as you indicate in your citation of relevant mammalian literature. Did you ever consider graviportality as a categorical variable in your various analyses? Its signal seems quite strong from a qualitative viewpoint, and it might be important to show that you can differentiate it from aquatic taxa.

4. Utility of quadratic DFA. I am not familiar with the use of quadratic DFA, but I'm wondering as to the utility of this approach. It doesn't seem to incorporate phylogenetic information explicitly (especially when compared to pFDA) and given the high phylogenetic signal in the dataset, I'd be worried interpreting its results without necessary caveats about shared ancestry. It's also difficult to fully understand the output image you show in Supplemental Figure 8 (more explanation would help). Moreover, the pFDA analysis gives similar results, incorporates phylogeny, and involves a straightforward output table that is easily interpretable (and which has been used in several recent papers where the comparative method is applied to palaeontological questions). But see my comment next about the sample set of the pFDA.

5. Inclusion of samples in pFDA: I found your statement that "We excluded non-archosaurian taxa from the training set for this analysis as many of them are subaqueous-foragers and as such they may unbalance the phylogenetic signal" (main text, pg. 16, line 358 and on) to be curious. Do you mean that, e.g., you had different sample sizes of sub-aqueous vs non-diving animals in archosaurs versus mammals? Different sample sizes of categories have

long been realized to cause issues for DFA and similar analyses, however these can either be addressed by subsampling (e.g., method employed by Chapelle et al., 2020 for bipedal vs quadrupedal locomotor categories) or by the insertion of prior category membership proportions. If you didn't mean that, could you explain why you made this statement? It seems to act to undermine the general applicability of your results as it stands…

Our previous dataset was taxonomically biased, because only aquatic ecologies were characterizing the mammalian branch. This was problematic for our analyses. In order to circumvent such problem, we added terrestrial mammals and lepidosaurs in the novel version of the femur dataset. All our new analyses based on femora were ran on the entire amniote sample. The same applies to the dataset of dorsal ribs.

6. Please correct typos in taxon names in supplemental worksheet

a. E.g., "Ichtyosaurus"

We corrected the spelling

7. Lack of informative clade labels on supplemental trees. While the phylogenetic work of this paper is secondary to the diving hypothesis, some researchers will undoubtedly cite it for its systematic findings. Can you please provide trees with explicit labels showing the positions of major clades that you recover? And, if possible, typical support values?

We included labels of the major clades and related that to node supports

8. Unsupported, vague, or inefficient statements in body text: The main text contains several passages that seem out of place with the comparative biological approach of this paper. These make unsubstantiated claims or read as hyperbolic. They can be amended or deleted without making any change in the significance of your findings. For example:

a. "This finding reveals spinosaurids to be aquatic specialists with surprising ecological disparity, including subaqueous foraging behavior in Spinosaurus and Baryonyx and potentially wading habits in Suchomimus" (main text, page 2, line starting at 37).

i. You examine three spinosaurids, and find that two of them do the same thing and one does not. Is that "surprising ecological disparity" in the group? I find that hard to assess.

We believe so. First of all, it is the first time that subaqueous foraging is directly linked to non-avian dinosaurs with such statistical support. What is even more surprising is that, although *Baryonyx* and *Suchomimus* share morphological similarity (based at least on the available remains), they possess completely different bone densities. This is a striking example of how a clade (Baryonichinae) might show low morphological disparity while being ecologically diverse. Additionally, *Baryonyx* and *Spinosaurus* show the same possibility of being subaqueous foragers, although being postcranially very different. Again, this shows how ecological diversity is underestimated in the fossil record, because our inferences mainly rely on anatomical proxies.

ii. "Potentially wading habits" is really not supported by your analysis at all. Your novel data consists of global bone compactness indices – and on those indices Suchomimus looks identical to many non-waders, for example Rhamphorhynchus. So the inference of wading is based on prior data that has been well-published. I would just indicate that you firmly predict Baryonyx and Spinosaurus as shallow divers, and surprisingly that Suchomimus lacks such osteosclerotic adapations, suggesting it has secondarily reverted to non-diving life habits (while acknowledging, as you do later on, that other parts of the skeleton suggest at least some affinity to water).

We agree with the reviewer and changed the main text accordingly. Because of this, we also modified the title, which is now only referring to "subaqueous foraging"

b. "The divergence of spinosaurids and their acquisition of adaptation to aquatic environments followed the faunal turnover and diversification of tetanuran theropods initiated by the emplacement of the Karoo-Ferrar large igneous province during the late Early Jurassic16,17, which led to a remarkable increase in ecomorphological disparity among predatory dinosaurs." (page 2, lines starting at 39).

i. The K-F LIP and the divergence of the spinosaurids you mention and include on your tree are separated by at least 40mya. I find it hard to make a compelling argument that these form a causal linkage. This argument is perpetuated elsewhere in the paper, with relatively little rhetorical support. I'd suggest deleting it altogether, or instead providing definitive statements upon which it can be evaluated as a hypothesis.

As mentioned before in our response to reviewer 1, we deleted any reference to the Karoo-Ferrar large igneous province and limited our statements to the origin of the clade (Jurassic) and the acquisition of aquatic habits (probably Cretaceous)

c. "Our findings greatly expand the ecological disparity of non-avian dinosaurs. Furthermore, the appearance of aquatic adaptations among spinosaurids is linked to a major burst in evolutionary diversification of the here-recovered clade Carnosauria that followed the emplacement of the Karoo-Ferrar large igneous province during the Jurassic16,17. (page 5, lines starting 106).

i. Your findings expand the ecological niche space of non-avian dinosaurs by exactly one category, which is already in the literature as a fairly well-developed hypothesis. So I would avoid making this statement, it seems hyperbolic. The importance of your finding is rather that a novel method, that has statistical power, has resolutely supported the notion of diving in some spinosaurids, resolving a major controversy in favor of an expanded niche.

We deleted any reference to the Karoo-Ferrar and our suggestion of increased ecological disparity among non-avian dinosaurs.

ii. What do you mean by a "major burst in evolutionary diversification"? I don't understand if this is in terms of lineages/per time or in terms of ecospace of some sort? Perhaps I missed this in the text and the supplement, but if not then I think this should be explicitly stated and if necessary quantified. Or you could just leave this sort of statement out – it would not detract from the study in any way.

We mean lineages per time: tetanurans are shown to rapidly diverge in the Jurassic in our analyses. This was previously quantified by Rauhut and Pol (2019).

d. "In conclusion, our data reveal that Spinosauridae, a geographically widespread clade of predatory dinosaurs, was ecomorphologically adapted to life in water. Our findings greatly increase the ecological disparity of non-avian dinosaurs, further challenging the longstanding hypothesis that this group was restricted to terrestrial environments. Spinosaurids were part of the rapid radiation of Tetanurae following the emplacement of the Karoo-Ferrar large igneous province during the late stages of the Early Jurassic16,17. Increased diversification appeared in concert with morphological innovation and high rates of homoplasy across vertebrate clades, including tetanurans17. Exploration of ecological regimes may have facilitated ecomorphological radiation linked to the invasion of previously uninhabited environments. Among these major ecomorphological radiations, multiple dinosaur clades independently evolved aerial capabilities32. Our study demonstrates that ecomorphological

radiations among non-avian dinosaurs also included the invasion of aquatic systems." (page 10, beginning line 226)
i. Please see previous comments about ecological disparity. You might even consider that since you show that ornithomimosaurs and haszkaraptorines are not divers, your analysis has reduced the disparity!

Please, see our response to your comments and the new version of the manuscript
ii. You present no data supporting the rapid tetanuran radiation post K-F LIP emplacement. So why are you concluding that you did?

Please, see our response to your comments and the new version of the manuscript
iii. Where is the data supporting high rates of homoplasy across vertebrate clades? What does this mean in the context of this paper and why is it important?

This was quantified by Rauhut & Pol (2019). We used the same phylogenetic dataset for our analyses. It is important because the higher degree if homoplasy at the divergence of tetanurans found by Rauht & Pol (2019) is found to correlate with terrestrial ecologies among the earliest diverging taxa. Therefore, our results show that ecological diversification happened secondarily to the early divergence of the major clades of tetanurans.
iv. What is an ecomorphological radiation in the context of exploration of ecological regimes? You're describing three taxa and a single theropod aquatic "invasion". While I do understand that sampled fossils really represent a much greater amount of evolutionary innovation and cladogenesis, these statements are so vague as to be nonsensical. I'd just leave them out entirely. Again, this will not diminish your work.

We changed the conclusion paragraph according to this and previous comments.
Sincerely,

Jonah Choiniere

Referee #3 (Remarks to the Author):

The study uses a broad survey of bone compactness metrics among archosaurs to predict aquatic behavior in non-avian dinosaurs. The key results of the work are 1) that subaqueous foraging behavior in extant archosaurs can be predicted by bone compactness with a high degree of accuracy, and 2) strong support is found for aquatic ecology in spinosaurids, a subject of much recent interest and controversy. The analysis is rigorously performed and well documented, and the writing and figures are excellent. The conclusions are original and will be of great interest to people who study dinosaur ecology. More broadly, the refinement of methods to predict behavior from bone compactness will be useful for people who study paleoecology and evolutionary biomechanics. My only major suggestion is that the limitations of the approach and how the data should be interpreted could be made clearer. This would not only avoid misinterpretation but also increase the manuscript's impact by making
the implications obvious to a broader audience.

Many thanks for these comments. We included more explicit statements regarding the limitations of our approach in distinguishing ecological categories in the main text. We hope that the new paragraph clarifies this and how certain methodological limitations can be circumvented with the addition of complementary proxies.

Introduction and literature: The literature cited seems appropriate, but previous studies have used different definitions or terminology, and some clarification would be helpful for those outside the field of ecology. For example, Quemeneur et al. 2013 used aquatic, semi-aquatic, and terrestrial as categories. How do these categories relate to probability of subaqueous foraging? Why is probability of subaqueous foraging used instead in this study – because it is more specific and/or more closely linked with bone compactness, or because it is a behavior rather than an ecological category? It seems that in this study anatomical variables other than bone histology are used to infer aquatic habits. Please clarify the relationship between aquatic/semi-aquatic, diving, and subaqueous foraging, ideally with examples from extant taxa.

Thanks for this comment. The reason we preferred to focus on subaqueous foraging and sustained flight instead of using previous categorizations (e.g. aquatic, semiaquatic, and terrestrial) are the following: subaqueous foraging and sustained flight are extreme behaviors related to the ecology of the studied taxa (see below our explanation of Liem's paradox). We find that semiaquatic and aquatic do not fully capture such behaviors among extant taxa, therefore limiting the understanding of ecological adaptations. For example, Canoville et al. (2016) and Quemeneur et al. (2013) categorize penguins and crocodiles as aquatic and semiaquatic, respectively. This was confusing for us: both taxa lay eggs on land and spend a significant amount of time in terrestrial environments, although they are dependent on fully submerged active pursuit behavior for hunting. The same applies to hippos and *Ornithorhynchus*: these share similar ecologies to penguins and crocodilians, but were scored as amphibious, contrary to penguins. Additionally, cetaceans, which are completely adapted to water, were scored as penguins (aquatic), even though their ecologies are significantly different. We thought that, perhaps, the best way of characterizing these different degrees of adaptations to water was to categorize what all these animals had in common, which is subaqueous foraging.

Regarding the relationship between our categories and the ones used in previous studies, we could say that our interpretation of subaqueous foraging does partially match the categories amphibious/semiaquatic and aquatic. However, we additionally distinguished between frequent and infrequent subaqueous foraging. For example, *Ursus maritimus* was previously scored as semiaquatic: this is correct in our opinion. However, such behavior is generally rare, especially when compared to other amniotes, such as crocodilians and *Ornithorhynchus*. In order to distinguish a commonly diving forager (such as *Ornithorhynchus*) from others, we added "frequent" and "infrequent" to better capture these differences among taxa.

We agree that this should be clarified in the materials and methods. Therefore, we added a paragraph in the method section explicitly explaining why we changed the general *modus operandi* and how we applied such categorizations.

Data & methodology: The approach has been validated in this and previous studies. Although almost all the extant taxa in this study are birds, previous studies have shown similar results for lepidosaurs and mammals. The analysis has been carefully chosen to separate the phylogenetic signal in the data from a functional signal. The data are

comprehensive and very well presented and the methods are sufficient for replication. One suggestion: I don't quite get the connection between Liem's paradox and the inclusion of flying as a variable - I thought Liem's paradox was that highly specialized taxa are also "jacks of all trades." Clarification would be appreciated.

We invoke Liem's paradox to justify that our ecological coding relies on the capability/frequency of two different abilities (flying and subaqueous foraging) to define the niche of each species used in our analyses. That's precisely because although very specialised taxa can do many things besides the hability they are specialised for, they can do things non-specialised taxa cannot accomplish.

The results support the conclusions. I'm a bit concerned that the limitations of interpretation are not clear enough and that readers may misinterpret or over-interpret the results. As I understand it, this method cannot be used to distinguish between animals that feed aquatically but do not submerge (e.g., storks) and completely terrestrial animals, nor between deep-diving fully aquatic animals (e.g., cetaceans) and semiaquatic animals that forage subaqueously (e.g., ducks).

Our methods can distinguish between animals that fully submerged in shallow waters (such as penguins) and those that do not fully submerged or show only infrequent subaqueous foraging (such as storks, ducks, pelicans, and wading birds in general), terrestrials, and volant. Having said that, the reviewer is correct: our method can not distinguish between waders, deep divers, gravioportal and terrestrial animals. Please note that while the femur has a higher predictability power for terrestrial and flying animals, dorsal ribs have a much lower signal for the distinction between terrestrial, flying, and deep diving animals.

We included an explicit paragraph in the main text stating the limitations of our approach.

Presumably this is because the relationship between bone compactness and aquatic lifestyle/behavior beyond two or three broad categories is not well understood (as pointed out by previous studies). However, in the supplementary information a distinction is made between the histology of "deep divers" such as ichthyosaurs and cetaceans and swimmers in shallow water, which include Spinosaurus (supp. Fig. 2), and other studies have discussed qualitative differences in bone histology between taxonomic groups and types or degrees of aquatic adaptation (Alexandra Houssaye, P. Martin Sander, Nicole Klein, Adaptive Patterns in Aquatic Amniote Bone Microanatomy—More Complex than Previously Thought, Integrative and Comparative Biology, Volume 56, Issue 6, December 2016, Pages 1349–1369, https://doi.org/10.1093/icb/icw120 and references therein). Could the writers comment on this, and perhaps how future studies might go about distinguishing between different degrees of aquatic adaptation and/or different types of aquatic locomotor or feeding behaviors (e.g., looking at differences in compactness index of vertebrae, ribs, and humerus where available)? This is set up by paragraph 1 of the introduction which discusses various types and degrees of aquatic adaptation found in extant animals, but it is not revisited in the light of the results (other than Suchomimus).

We included a paragraph discussing such features in the main text and moved Supplementary Figure 2 to the main text as part of Figure 1. The addition of a new dataset focused on dorsal ribs allowed us to evaluate how the investigation of multiple skeletal elements is beneficial for ecological inference: gravioportal animals show increased bone

density in the limbs, but lack this in non-supporting skeletal elements. This trend is observed only in graviportal animals, such as sauropods.

Also, how should one interpret a result like the 34% probability of subaqueous foraging in Megalosaurus: as an inconclusive
result or an indication of semi-aquatic adaptation? I understand the limitations of interpreting fossils with unique anatomical characteristics (223-224), but I think it could be valuable to hear the authors' educated opinions on possible ecological interpretations of spinosaurids. We believe that 34% does not represent a great statistical support. Although this threshold might be sound as subjective, we wanted to be conservative in our interpretations. The same applies to the unnamed Kem Kem averostran: this taxon is predicted as subaqueous forager for slightly more than 50% of the times, which is clearly not a strong result. Moreover, none of these non-avian dinosaurs actually show any sign of osteosclerosis, such as the presence of trabeculae and spongiosa in the medullary cavity, supporting our conclusion of non-subaqueous foraging among other clades. For all these reasons, we reject the possibility of subaqueous foraging among these taxa. On the other hand, Baryonyx and Spinosaurus show osteosclerosis in all the investigated skeletal elements, which is in line with the high statistical support for their ecological inference.

Referee #4 (Remarks to the Author):

It's not entirely clear to me what the conclusions of this paper are, because the fundamental language used is unclear – what exactly _is_ subaqueous foraging? Is it completely submerged, swimming pursuit? Being an ambush predator lying in wait under the water? Does it extend as far as wading in deep water, and the subaqueous applies only to that being foraged? Are crocodiles and sharks both "subaqueous foragers"? The reason this matters is because it directly affects the novelty of the study – it is well established that Spinosaurus is a piscivore (e.g. 1,2,3,4,5), and therefore probably spent time in the water, and certainly foraged from under the water's surface. What is contentious (2,3,4,6), is whether it actively swam in pursuit of prey, or was an ambush predator, either wading or partially/fully submerged. This paper does not answer that, even though its primary conclusion is that "Spinosauridae… [were] ecomorphologically adapted to life in water". I don't know if the authors want to hedge their bets, or if their data simply cannot discriminate between types of aquatic lifestyles, but in this regard, there is little novel about the conclusions.

Thank you for your comments. Our definition of subaqueous foraging is fully submerged active behavior (not all subaqueous foragers swim, e.g. *Hippopotamus* and *Tapirus* show underwater walk). Our results completely refute the "wading hypothesis" and previous suggestions that animals like *Spinosaurus* and *Baryonyx* were merely submerging their snouts: all wading animals have hollow bones and/or low bone density. This is clearly not the case in *Spinosaurus* and *Baryonyx* . On the other hand, *Suchomimus* is inferred as a non-diving animal. We modified the introduction to explain what we mean by subaqueous foraging.
The second major part of the study is in showing that bone compactness is increased in aquatic (sub-aqueously-foraging) animals. Again, this is not particularly novel (7, 8, 9, 10), though admittedly most previous work has been on mammals.

Previous work mainly focused on specific clades or elements. Our study includes the largest dataset of bone density ever built – by a wide margin - and covers the entire clade Amniota. That being said, the methods and data are an excellent resource, and clearly contribute to the literature. My biggest concern is that only a single bone, the femur, is used in the study. Is this always going to be the most representative of lifestyle?

Our original choice of the femur for this study was based on multiple factors: 1) this element has been heavily studied among dinosaurs because of function/gait; 2) it has been heavily sampled for osteohistological studies, therefore it allows a broad coverage of non-avian dinosaurs (which is the main group of interest in this study); 3) it has been used as a proxy for size estimation among extant and extinct taxa. In order to test our conclusions, we added a new dataset focused on dorsal ribs (175 amniote taxa including spinosaurids). We found the same patterns recovered for the femur: bone density correlates with subaqueous foraging. We found very high possibility that both *Spinosaurus* and *Baryonyx* were subaqueous foragers, while *Suchomimus* is found not to be a subaqueous foraging animal.

A number of waders are listed (ln 134-137), but all of these are relatively small, and also fly. Yes, subsequent models incorporate flight into the relationship with bone compactness, but if Spinosaurus were a wader, would we perhaps expect dense hind-limb bones, but typical theropod pneumatisation elsewhere?

Osteosclerosis is also found among small sized birds and mammals. Allometry is not strongly correlated with bone density (please see Table 1, Supplementary Tables 3-4, 7-10). All waders are found to have hollow bones across the skeleton. This is not the case for subaqueous foragers. *Baryonyx* and *Spinosaurus* are shown to be osteosclerotic across the whole postcranium in our paper (in addition to our analyses based on the femur and dorsal ribs, please see Extended Data Figure 1-10). Finally, pneumatization is absent across the postcranium of spinosaurids, with the exception of the cervicals (Benson et al. 2012)

Benson, R. B., Butler, R. J., Carrano, M. T., & O'Connor, P. M. (2012). Air-filled postcranial bones in theropod dinosaurs: physiological implications and the 'reptile'–bird transition. *Biological Reviews*, *87*(1), 168-193.

This could presumably be tested through histology/CT scanning of other bones in the skeleton, and seeing if they are as compact. The authors could also look at other bones (e.g. vertebrae) from their sample of 141 taxa and explore if other parts of the skeleton are more or less indicative of a subaqueous lifestyle.

We added a new dataset based on dorsal ribs (175 taxa): these elements were previously shown to be informative skeletal elements for inferring ecological adaptations. We applied the same methods to this dataset (testing the relevance of the proxy among extant taxa first to then infer ecological adaptations among non-avian dinosaurs). We found the exact same results: *Spinosaurus* and *Baryonyx* are subaqueous foragers, while *Suchomimus* is recovered as a non-diving animal. This further demonstrate that :1) presence or absence of osteosclerosis is not limited to few skeletal elements, but is widespread in the postcranial skeleton; 2) bone density is a valid proxy for ecological inference; 3) bone density is linked to aquatic adaptations in shallow waters.

This of course is difficult to apply to Spinosaurus, because no matter how large the sample size of extant

taxa (or extinct taxa of conclusive aquatic affinity), there is nothing closely analogous to Spinosaurus (very large, potentially wader/shoreline predator, potentially bipedal on land) – which is presumably why the study was carried out with that animal as the focus in the first place.

Our datasets include very large and bipedal predatory dinosaurs: *Tyrannosaurus, Tyrannotitan, Suchomimus*, and *Torvosaurus* in the femur dataset, and a carcharodontosaurid in the dorsal rib dataset. All these taxa are found to be terrestrial, with no or very weak probability of being subaqueous foragers. On the other hand, we find that the similarly sized *Baryonyx* and *Spinosaurus* are strongly predicted as subaqueous foragers. Our results reject previous hypothesis depicting *Baryonyx* and *Spinosaurus* as shoreline waders.

Minor comments:

Line 119: use of (1) and (2) for two variables that have 0, 1 and 2 might be confusing, perhaps use letters and numbers: "taxa were scored using two categorical explanatory variables that encode the presence of (a) subaqueous foraging (0, unable, 1, able…"

Changed

Line 121: Would "two independent variables" not be simpler than "two independently varying variables"?

Line 193: do these results point to all Spinosaurids as being aquatic? If so, how can only some of them be subaqueous foragers?

Not all spinosaurids are subaqueous foragers: this is the novelty of the study in addition to the dataset and methods. Our study clearly demonstrates that anatomy alone is not a very good proxy for ecological inference and that the ecological diversity in the fossil record is underestimated.

Data availability: Please provide direct links to all CT data used, in accordance with standard open data practice (11)

We will be sharing all CT scans (freely available in morphosource) and black-and-white sections collected for this study (extended data figures 1-9) upon publication of the manuscript.

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

In this revision, the authors successfully address my suggest changes and criticisms. The paper is more clear, succinct, and integrated (with a larger dataset) and this paper will lead to similar future studies on bone density. Quickly going through the other reviewers, it appears that the authors satisfy most of the previous concerns and criticisms. I recommend publication as is from my view.

Referee #2 (Remarks to the Author):

Dear Authors,

I have read your revised manuscript and its associated files. Thank you for your responses to my comments and for the various clarifications and improvements you made to the manuscript and its accompanying analysis.

I have no further substantive comments and sincerely hope that the work will soon be published.

Referee #3 (Remarks to the Author):

I appreciate the information that has been added about the limitations and the definitions of aquatic behaviors. It is now very clear what the study shows and what the broader implications are. I have a few remaining questions about interpretation of data, but overall I think the manuscript is very good.

1. Regarding the classification rate of the discriminant analysis: the overall rate of 83-85% does not really indicate how well it would work in extinct taxa because many of the taxa in the analysis are obviously aquatic or terrestrial based on their general anatomy (e.g., elephants, sea turtles). Can you give any indication of how it performs in more difficult cases, such as semi-aquatic animals? This would give a more realistic estimate of how useful the method will be for inferring ecology of extinct animals.

2. I still don't understand fully how the authors interpreted the probabilities generated by the discriminant functional analysis. For the taxa with 34 and 50% probabilities and without signs of osteosclerosis, the possibility of subaqueous foraging was rejected. Does this mean the null hypothesis is that a certain taxon did not engage in subaqueous foraging? What would be the threshold to reject this hypothesis?

Minor comments
Lines 40-41: unclear what "secondarily" refers to; secondary to what?

Lines 68-69: what do "these" and "this" refer to – suggestions and controversy?
Line 178: rib, not "ribs"

Referee #4 (Remarks to the Author):

The authors have clearly added a lot since the last version of the manuscript. In particular, the clarification of language, and additional datasets have addressed my prior concerns as much as could be hoped.

Providing all the data are made openly available, as the authors have committed to, I am sure this paper will be a major contribution to the field.

Author Rebuttals to First Revision:

We are grateful for the insightful and thorough comments, which improved and strengthened the results and impact of our manuscript. We followed all reviewers' comments and included several additional paragraphs in the methods section to provide further details on our methodologies. We also improved our formatting of the manuscript: figures 1-3 are formatted to be published at 2/3 column size (1 and 2) and at single column size (figure 3). Table 1 in the main text is shortened, representing the three best models for each dataset, rather than the best four. Although previous reviews required further discussion of the results, and consequently led to a considerable increase in word count, we managed to shorten the manuscript of 348 words. The current length of the main text (including the abstract) is 2544 words. References were corrected according to the formatting guide. Finally, we uploaded all the CT scan datasets used in this study in morphosource (see Supplementary dataset for a list of taxa uploaded and their respective links).

We hope that the editorial board and the reviewers will be pleased with our improved version of the manuscript.

Referee #1 (Remarks to the Author):

In this revision, the authors successfully address my suggest changes and criticisms. The paper is more clear, succinct, and integrated (with a larger dataset) and this paper will lead to similar future studies on bone density. Quickly going through the other reviewers, it appears that the authors satisfy most of the previous concerns and criticisms. I recommend publication as is from my view.

Referee #2 (Remarks to the Author):

Dear Authors,

I have read your revised manuscript and its associated files. Thank you for your responses to my comments and for the various clarifications and improvements you made to the manuscript and its accompanying analysis.

I have no further substantive comments and sincerely hope that the work will soon be published.

We want to thank the reviewers for the insightful comments that improved the manuscript

Referee #3 (Remarks to the Author):

I appreciate the information that has been added about the limitations and the definitions of aquatic behaviors. It is now very clear what the study shows and what the broader

implications are. I have a few remaining questions about interpretation of data, but overall I think the manuscript is very good.

Thank you for these additional comments. We included further details regarding our methods and interpretations in the main text. We hope that the reviewer will find these satisfactory.

1. Regarding the classification rate of the discriminant analysis: the overall rate of 83-85% does not really indicate how well it would work in extinct taxa because many of the taxa in the analysis are obviously aquatic or terrestrial based on their general anatomy (e.g., elephants, sea turtles). Can you give any indication of how it performs in more difficult cases, such as semi-aquatic animals? This would give a more realistic estimate of how useful the method will be for inferring ecology of extinct animals.

Our analyses are based on bone density and maximum shaft diameter only. Anatomy is completely excluded from our quantitative inference, although it is taken into consideration in our discussion. This is explicitly stated in our new version of the manuscript. Correct classification rates vary between categories, e.g. subaqueous forager or non-, because dependent on the ecological classification in extant animals or extinct taxa with undisputed ecological adaptations (e.g. subaqueous foraging in ichthyosaurs and plesiosaurs) and they show a range of bone density values. In other words, classification rates are not symmetrical. Specific classification rates for subaqueous foraging and non-subaqueous foraging are provided in both the main text and Supplementary Information. Overall, the correct classification rate is high, especially when considering the range of taxa included in the dataset: it includes both taxa deeply adapted to a life in water, such as cetaceans, and taxa that show an amphibious lifestyle, but still retain anatomical traits shared with terrestrial animals (e.g. *Hippopotamus* and *Tapirus*). Therefore, we think our dataset appropriately captures the complexity of ecologies between fully terrestrial and fully aquatic lifestyles, including some extreme cases like graviportality and pelagic specialists.

2. I still don't understand fully how the authors interpreted the probabilities generated by the discriminant functional analysis. For the taxa with 34 and 50% probabilities and without signs of osteosclerosis, the possibility of subaqueous foraging was rejected. Does this mean the null hypothesis is that a certain taxon did not engage in subaqueous foraging? What would be the threshold to reject this hypothesis?

Our phylogenetic discriminant function analysis was repeated using 100 phylogenetic trees (with differing branch lengths to account for stratigraphic uncertainty). Each time, the variables (bone compactness and diameter) from the training set of taxa with known ecologies, together with the phylogenetic structure of data, are used to generate the discriminant functions. Those functions will be subsequently used to predict the ecologies in a handful of taxa with unknown ecologies (including spinosaurids). A given species is predicted as a subaqueous forager if the posterior probability is 50% or more. This is because our inference has only two possible outcomes: subaqueous forager or not. We summarised

our results by providing the median value of those 100 posterior probabilities and the number of times a particular taxon is predicted as subaqueous forager (median probability of 50% or more). This gives us two proxies of the likelihood of each taxon to be an actual subaqueous forager. For instance, a taxon could be predicted 100 times as subaqueous forager with a median probability of 51% which means the evidence for this extinct species to be an actual subaqueous forager is very weak and this inference is consider very unlikely. Median probabilities need to be within the range of 80-100% to be considered strong evidence for subaqueous foraging. A clear example of this is the unnamed averostran from the Kem Kem group, which is predicted as subaqueous forager in the analysis, although the median probability is only 58% (Supplementary Table, 6). Therefore, although in each round of the analysis this taxon passed the threshold of 50%, the evidence for this taxon to be an actual subaqueous forager is very weak. Additionally, the presence of an open medullary cavity rejects osteosclerosis in this taxon, although its bone cortex is relatively thicker than its sister taxa. If those values are compared with the ones recovered for spinosaurids, interstation of the results should be easily accomplished.

Minor comments
Lines 40-41: unclear what "secondarily" refers to; secondary to what?

We eliminated "secondarily"
Lines 68-69: what do "these" and "this" refer to – suggestions and controversy?

"these" is referred to previously suggested hypotheses. This refers to the confusion affecting our understanding of ecological diversity in the fossil record. We rephrased the sentence.
Line 178: rib, not "ribs"

Changed

Referee #4 (Remarks to the Author):

The authors have clearly added a lot since the last version of the manuscript. In particular, the clarification of language, and additional datasets have addressed my prior concerns as much as could be hoped.

Providing all the data are made openly available, as the authors have committed to, I am sure this paper will be a major contribution to the field.

Thanks for your comments. We uploaded all datasets in morphosource and they will be fully accessible upon publication of the manuscript.

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #3 (Remarks to the Author):

The responses have cleared up my remaining questions. A very strong manuscript, and I'm looking forward to seeing it published.

Referee #4 (Remarks to the Author):

No further comments - the Authors have addressed all my concerns and the manuscript is in a publishable state.