The manuscript of Amson and Nyakatura is an interesting contribution to the existing literature on the relationship between trabecular and cortical bone and behaviour of extant xenarthrans, and tests whether the behaviour of fossil taxa can be reconstructed. In their previous publications, the authors have been the first researchers to apply modern methods to investigate trabecular bone structure in non-primate taxa, and I am pleased to see the potential utility of this aspect of bone structure explored.

The humerus and radius are analysed amongst extant taxa which differ in both their locomotor mode, and in the frequency and mode of fossorial behaviour. Several fossil xenarthrans were analysed in order to determine their locomotor and fossorial behaviour. Although, locomotor and fossorial behaviour could not be not confidently assigned to the fossil taxa included in the analysis, *Hapalops* was found to have sloth-like morphology. I am pleased to see the authors discuss the complexities of this methodology, including the problems arising when interpreting bone structure in such large-bodied, extinct species.

I have divided my comments on the manuscript into major and minor suggestions, and have also noted spelling and grammatical errors.

Major comments

1. The manuscript would greatly benefit from a more detailed and nuanced interpretation of the various bone parameters measured.

Cortical and trabecular bone are analysed collectively in the manuscript, however, the two regions of bone perform very different functions. Cortical bone cross sectional geometric properties reveal the potential for bones to resist compressive and bending forces, whereas trabecular bone structure is adapted to compressive forces at the articular surface. As such, the two types of bone would not necessarily be expected to co-vary directly, as assumed by the authors.

The manuscript should more clearly explain how these two different bone structures reflect mechanical loading.

2. A separate, but related comment is that the biomechanical hypotheses tested are not clearly explained. I understand that often the kinetics/kinematics of behaviours such as climbing and digging are unknown, but the authors should elaborate on the differences that would be expected for their behavioural categories. For example, what differences might there be in the variability of loading (as measured by DA) between fossorial and non-fossorial groups? Or, how might the overall the magnitude or orientation of loading differ between behavioural categories?

Although not feasible to incorporate in the present study, I wonder if elements of the hindlimb might be informative for future studies. The two behaviours of interest in these groups are arboreal/terrestrial locomotion and fossorial behaviour. I assume that digging in these groups only, or at least primarily, involves the forelimb, in which case you would expect the hindlimb to be influenced only by locomotion. As it stands, the elements in this study are influenced by both locomotion and digging, thus the morphology is likely due to a combination of two different loading regimes.

3. The cross-sectional geometry of the elements included in this study are highly complex, and at the sites analysed there are large muscle attachment sites. I would suggest that the authors discuss the possible impact of these muscle attachments sites on the cross-sectional geometry results, and for future analyses consider sampling locations on the diaphysis without prominent entheses.

4. Information should be included on sample size and extant species in the sample, and the details of CT scanning of the extant sample (i.e. where they were scanned and at what resolution).

5. I would like to see more information about the VOI placement protocol. The methods say that ROIs were selected from the centre of the epiphyses, however, looking at the images in Amson et al. 2017, this doesn't appear to be the case for either the humeral head or the radial trochlea, the MC3 was not included in this previous publication. A clearer description, and preferably a figure, should be included to explain further the VOI placement protocol.

6. p10, line 223: Was the total volume used here the size of the VOI? This is unlikely to be a good size proxy, because the VOIs were not scaled to the size of the epiphysis, rather as large a VOI as possible, avoiding cortical bone, was placed in the epiphysis. Although the TV is not used as a size proxy in the analysis, a measure of the size of the epiphysis would be more appropriate than the size of the VOI.

7. p12, line 265: Were the parameters normally distributed after log-transformation?

8. The paper should include a results table with the mean values for each taxonomic group or species, and the results for each fossil.

9. p13, Univariate Comparisons: The focus of this section is on the fossil taxa, but it would help the reader a brief description of how the extant groups differ from one another was included, for all parameters discussed.

10. Not enough information is given concerning the discriminant function analysis – I would expect the paper to include a table reporting data from this analysis, and additional information in the text. For example, a table with predicted group membership should be included for both the training data and the fossil taxa. It would also be informative to include the contribution of the variables to each function to better understand which variables are driving between group differences. How were the extant taxa grouped – it is unclear whether this is at the species level, generic level, or by a behavioural classification?

11. What is the potential influence of correlation between variables on the DFA, and on the PCA used for *Hapalops*? The included trabecular parameters are likely to be correlated with one another, for example BV/TV and Tb.Th.

<u>Minor comments</u>

Abstract: The authors overstate the sensitivity of trabecular architecture by using the phrase "extreme accuracy and sensitivity" in the abstract; in p1 line 44 "great accuracy and sensitivity"; and in p1 line 47 "great plasticity". Studies in primates have had very mixed results, in many cases the relationship between trabecular structure and behaviour is unclear. I would recommend these

phrases are adjusted to reflect that it is not known how accurately trabecular structure reflects loading.

P4, line 65: Tsegai et al. (2017) used the cortical thickness method developed by Treece et al. (2010; 2012). It is important to note that the focus was cortical bone thickness at the articular surface, rather than diaphyseal structure.

p4, line 88 and p7, line 153. Anteaters are described here as intermediate, it should be more specific, is this intermediate in their fossorialism or terrestrial/arborealism?

p9, line 209: Change “trabecular” to “cortical”

p17, line 401: Include other publications from the primate literature, as there any many studies which find DA, or primary trabecular orientation, to be informative (e.g. Ryan and Ketcham, 2002; Griffin et al., 2010; Barak et al., 2013; Su et al., 2013).

p19, line 460: This is in important point, which could be expanded upon. Is there any evidence for this in extant xenarthrans?

Spelling/grammatical comments:

p3, line 44: “excepted” should read “expected”

p3, line 49-54: Using “was” here sounds rather strange, I would recommend using “has been”. Also on p4 line 73, “parameters were” should read “parameters have never been”, also p18 line 430 change “was never” to “has never been”.

P4, line 68: Here and in a few other instances throughout the manuscript there is an error in the number of parentheses.

p5, line 91: “this is likely not true anymore” is confusing – please rephrase.

p5, line 95: Change teeth to tooth. It is either tooth morphology or morphology of teeth.

p5, line 97: In “This was found as challenging” change “as” to “to be”.

p5 line 101: Remove first “in”

p5 line 113: Add “a”: “Such a lifestyle”.

p10, line 224: Change “specific” to “species”.

p10, line 228: Add “taxa” after “extinct”.

p11, line 245: Change “Beside” to “Besides”

p12, line 270-271: There is an error with the ü in the PDF.

p12, line 274: Indent subheading

p16, line 383: Change “correlated to” to “correlated with”

p16, line 384: Change “studied taxa to influence the analysis” to “studied taxa from influencing the analysis”

p18, line 432: Change “as consistent with a fossorial” to “to be consistent with fossorial”

p18, line 434: Change “non-fossorial taxa to” to “non-fossorial taxa in”

p18, line 439: Move “is”, this should read “in neither case is the classification clear”

p 18, line 440: Change “more of their” to “additional”

p21, line 488: Change “investigations” to “investigation”

References

Barak MM, Lieberman DE, Raichlen D, Pontzer H, Warrener AG, Hublin J-J (2013) Trabecular evidence for a human-like gait in *Australopithecus africanus*. *PLoS ONE,* **8,** e77687.

Griffin NL, D’Août K, Ryan TM, Richmond BG, Ketcham RA, Postnov A (2010) Comparative forefoot trabecular bone architecture in extant hominids. *J Hum Evol,* **59,** 202-213.

Ryan TM, Ketcham RA (2002) The three-dimensional structure of trabecular bone in the femoral head of strepsirrhine primates. *J Hum Evol,* **43,** 1-26.

Su A, Wallace IJ, Nakatsukasa M (2013) Trabecular bone anisotropy and orientation in an Early Pleistocene hominin talus from East Turkana, Kenya. *J Hum Evol,* **64,** 667-677.

Treece GM, Gee AH, Mayhew PM, Poole KES (2010) High resolution cortical bone thickness measurement from clinical CT data. *Medical Image Analysis,* **14,** 276-290.

Treece GM, Poole KES, Gee AH (2012) Imaging the femoral cortex: Thickness, density and mass from clinical CT. *Medical Image Analysis,* **16,** 952-965.