Present study lead by Plateau & Foth and titled '*The impact of allometry on vomer shape and its implications for the taxonomy and cranial kinesis of crown-group birds*' deeply addresses a relevant aspect that strongly affect the morphological evolution of several anatomical structures: size. The approach you propose is essential to correctly understand how the influence of this central structure is in craniofacial and PPC configuration, and also in kinesis performance among other cranial traits. This is important specially when such structures are suggested as taxonomic distinguishing. Thus, it is important that you emphasize that, although there are important morphological differences between the main groups of modern birds, these are mostly due to (or related to) the evolution of body size, a fundamental biological trait, and not so much as shape changes themselves (i.e. vomer morphology is not independent, in this case from its size, and if you don't take raw size differences as a taxonomic trait, you shouldn't take as such a morphological trait associated with it).

The authors address this important question with a wide and dense combination of methods which look for the pure morphological differences that are not under the widespread influence of size, but I suggest below some commonplace aspects that I think they could contribute to the robustness of the study and make it more complete.

Methods

I would suggest to estimates the morphological disparity from the Procrustes shape variables for each group and compares it between groups with pairwise comparisons. Thus, you can compare statistically the morphological variation for each group and their distinct distribution across the morphospace, especially when you extract size and groups collapse, obtaining different significant results in size-affected shapes.

Also, another option to provide robustness to the hypothesis that vomer's morphology is not a good taxonomic character is estimating the phylogenetic signal both in shape variables, centroid size and centroid size proportional to skull size. You should obtain empirical K values lower to the predicted value of K = 1 for a Brownian Motion model of evolution indicating weaker phylogenetic signal and a higher degree of convergence. Additionally, it should be observed a similar distribution across the phylogeny between shape variables and size variables plotting them as continuous variables.

I would suggest to carry out the allometric regression with Phylogenetic Generalized Least Squares and group by group, running pairwise comparisons between their slopes in order to know how is the allometric influence. Allometry seems to affect specially in Paleognathae considering that when you extract size the difficulty to distinguish these birds from the rest increases. This aspect could be due to ratites reaches bigger sizes.

Finally, you discuss about kinesis and why is not be possible that its different types are related to vomer shape. In order to know this lack of correlation I would recommend grouping the sample by kinesis type,

estimating morphological disparities for each group and comparing statistically their distribution across the morphospace. It should collapse widely. Also, it would be interesting to explore by PGLS how different types of kinesis are related to different allometric scaling, because it could be that the shape-size relationships from distinct kinesis types are similar and not significantly different. Furthermore, to do that I would take into account as an important aspect for kinesis performance the vomer size relative to skull size for a better addressing of their significance on such trait.

Figures

For the figure 1 I suggest to add a close up of the region that anatomically distinguishes paleognathans from neognathans and a better schedule that could show the stiffness or mobility of the PPC (for example, marking or highlighting the structures that characterize each type of PPC).

For the PCAs (figure 2) I would suggest trying to plot vomer size and vomer size overall skull size as dot size. Thus, it could be visualized the polarization of size on the morphospace and the wide divergence between paleognathans and neognathans. Finally, the interpretation of the graphics would improve if you add silhouettes representing different groups and if you mark above from which analysis they come or if they are taking size into account or not.

There are no more comments that both authors and editors must take into account for the review of the manuscript.