ACTA ZOOLOGICA BULGARICA

Acta Zool. Bulg., December 2024, 76 (4): 587-590

Research Critique

Published online 15 November 2024
DOI: https://doi.org/10.71424/azb76.4.002767

Who Would Care about the Science? Critique of the Paper by Gündemir et al. (2023), with Comments on Peer Review Responsibility

Bartosz Borczyk

Department of Evolutionary Biology and Conservation of Vertebrates, Faculty of Biological Sciences, University of Wrocław, Sienkiewicza 21, 50-335 Wrocław, Poland; E-mail: bartosz.borczyk@uwr.edu.pl; ORCID: 0000-0001-9630-1809

Abstract: This article presents critique of the most important errors and flaws of the recently published article by Gündemir et al. entitled "Shape differences of the *Carina sterni* in birds of various locomotion types". Numerous problems arise, including study design, statistical approach, interpretation of the results and others. It also raises the question on the responsibility of editors, reviewers and erosion of ethos of science.

Key Words: *carina sterni*; morphology; critique; scientific methodology

Introduction

Scientific journals, especially the leading ones, are overloaded by new submissions. New publishing consortia are arising, often with high publication fees and still find hundreds of authors willing to pay just to publish their papers. However, pressure for publication (famous publish-or-perish attitude) often translates into quantity (and not quality) of papers (Sarewitz 2016). It may also lead to unethical or controversial practices. These may include self-reviewing (by creating false accounts for reviewers and subsequently suggesting them as reviewers) or reviewers "consortia" (one author reviews the work of another who reciprocates with a favourable review), data falsification or fabrication and others that add to erosion of the ethos of science (e.g. Ferguson et al. 2014, Walker & da Silva 2015).

Recently, the respected journal *Anatomia, Histologia, Embryologia* published a paper by Gündemir et al. (2023) dedicated to shape differences of sternal crests in birds that use different modes of locomotion. Unfortunately, I find the paper overloaded by trivial, but cardinal, errors. The data and sampling are inadequate and poorly analysed. The discussion is inadequate, too. To give some overview of the problem, I point out some of the most striking errors and mistakes (they are listed and commented below); however, the list is not exhaustive. My intention is not to provide peer-review-like comments but point out the most obvious errors and flaws of the Gündemir et al. (2023) paper.

Major objections to the Gündemir et al. (2023) paper

Gündemir et al. (2023) claim the presence of sternums in chondrichthyans should attract serious attention from any comparative anatomist, since this structure is crucial for tetrapod locomotion. Authors wrote: "In fish, the sternum was described by Howes (1891) at the shark *Notidanus indicus* [= *Notorynchus cepedianus* (Péron, 1807), broadnose sevengill shark – BB] being represented by two cartilages." Howes (1891) recalled an earlier paper by

Parker (1980); however, the cartilages that Parker described as sternums are not homologous to tetrapod sternum. Moreover, it is clearly indicated in the Howes' note: "My friend Prof. T. J. Parker has in these pages (Dec. 11, 1890, p. 142) lately recorded the existence of a sternum in the shark *Notidanus indicus*. The anterior of the two cartilages, which he figures, has been already described by Haswell (Proc. Linn. Soc. N.S.W., vol. ix., part 1) [Haswell 1884 – BB]. In view of Parker's conclusions, it is interesting to note that he speaks of it (p. 23) as 'temptingly like the pre-sternal,' but that 'the presence of such an element in the skeleton of any group nearer than Amphibia seems to preclude this explanation."

Surprisingly, Gündemir et al. (2023) referred to the paper by Feneck et al. (2021), who clearly stated that the first trace of the sternum is found in *Ichthyostega* (Pierce et al. 2013), a famous transitional fossil between 'fishes' and tetrapods – far away from sharks on the phylogeny. The chondrichthyan-osteichthyan (Osteichthyes: ray-finned fishes, coelacanths, lungfishes and tetrapods) divergence took place in the Silurian, approximately 460 mya; *Ichthyostega* is dated approximately 360 mya (Benton & Donoughe 2007, Inoue et al. 2010, Kumar et al. 2017).

In the next paragraph, one can find statement that caecilians (Order Gymnophiona Müller, 1832) are reptiles ("Reptiles, excluding snakes, turtles, caecilians, and legless lizards (…)" (Gündemir et al. 2023). However, the caecilians are not reptiles (Class Reptilia Laurenti, 1768) and this is a common name for legless amphibians (Class Amphibia Gray, 1825). This obvious error has its probable roots in the sentence: "In snakes, caecilians, and legless lizards that have undertaken complete forelimb reduction, the sternum is absent" from Feneck et al. (2021) to who Gündemir et al. (2023) referred. Despite the unfortunate listing of caecilians between two reptile groups, there is no indication that Feneck et al. (2021) recognised these legless amphibians as reptiles.

Further, Gündemir et al. (2023) try to discuss adaptations to different locomotion styles represented by "walking" (turkey and chicken), "flying" (pigeon), and "swimming" (duck and goose). Such a study could be of great value for understanding mechanisms of bird evolution, morphological integration and trade-offs that result from historical and functional constraints (e.g., Lowi-Merri et al. 2021). First, the bird sternum plays a great role in flight ability but not all flight styles are the same: flipper, burst, soaring, etc. However, its morphology is also constrained and shaped by the habitat (aquatic, terrestrial, arboreal, etc.), phylogeny (historical constraints), allometry (hummingbirds, the smallest flying birds, can weigh 2 g with a 5 cm wingspan, whereas wandering albatross *Diomedea exulans* weighs 12,700 g with a 370 cm wingspan), and other factors (Lowi-Merri et al. 2021). Thus, the simple classification proposed by Gündemir et al. (2023) does not reflect the actual diversity of locomotion types within their studied species, which includes different ways of terrestrial locomotion (Galliformes are the most terrestrial of the studied species but all others frequently move on the ground), different flight patterns, wing-shapes and relative size, animal size, etc. Disentangling effects of such different factors on sternum form (or any other structure) need at least two basic assumptions to be fulfilled. First, one needs a biomechanically justified classification of locomotion styles. This is lacking in the paper of Gündemir et al. (2023) and this itself makes their study useless. Second, one needs a more diverse locomotory pattern within closely related species and similar locomotory patterns among more distantly related forms to sort the effect of phylogeny (historical constraints) from biomechanical needs (Harvey & Pagel 1991). This is not met here, since the "walking" birds are represented by two galliforms, and "swimming" birds are represented by two anseriforms. The "flying" birds are represented by a single species from a separate lineage. Aside from small taxonomical sampling, there is neither within-lineage variation nor any shared characteristics between more distantly related forms that would allow any evolutionary reasoning.

Further, all of the studied species are highly domesticated. Thus, the locomotion abilities of these species might be seriously restrained by the artificial selection on meat and egg production and rapid growth (with exception of some pigeon breeds selected for flight ability). Beside this, authors ignore the size differences between these species. Pigeons weigh up to 300 grams, whereas male turkeys may reach 12,000 grams, which is a 40-times size difference. Thus, potential differences might simply result from allometry (e.g., Schmidt-Nielsen 1984, Harvey & Pagel 1991, Lowi-Merri et al. 2021), which is ignored by Gündemir et al. (2023).

Even more ridiculous is the interpretation of statistics in the analysed paper. Among numerous methods for quantifying and comparing the shape of a given structure, one can use shape differences measures, such as Procrustes distances (PD) (Dryden & Mardia 1998), which was used by Gündemir et al. (2023). The Procrustes analysis is a method of shape analysis based on landmark configuration (e.g., Rohlf & Slice 1990, Dryden & Mardia 1998,

Zelditch et al. 2004). PD is calculated as the square root of the sum of squared differences in the positions of the landmarks calculated by Procrustes analysis in two shapes. As such, it is used to describe the difference between landmark configurations (Rohlf & Slice 1990, Dryden & Mardia 1998). In very rough simplification, these distances are the measure of how averaged shapes of several objects in the given category differ between these groups, and larger values of PD indicate bigger shape differences. Now, let us see what did Gündemir et al. (2023) compared and what passed the rigorous reviewers and editorial peers! Well, they used the p-values as the measure of shape distance (sic!) in the following way: the lower the p-value is, the bigger the distance between the groups is. At the same time, they present (in their Table 1) the PD (and Mahalanobis distances), which shows different patterns than those 'deduced' from the inspection of p-values:

"The most different distribution was between the duck and turkey (p: 0.0003) and between the turkey and pigeon (p: 0.0008). The third-largest difference was between the pigeon and duck (p: 0.0023)" (Gündemir et al. 2023).

Inspection of the Table 1 of Gündemir et al. (2023) shows the largest PDs were, in fact, between duck and turkey (PD: 0.2252), goose and turkey (PD: 0.2236), and pigeon and chicken (PD: 0.2038). The Procrustes distances for turkey and pigeon (PD: 0.1937) and pigeon and duck (PD: 0.1718) are smaller. P-value is the probability of the data observed to have occurred under the null hypothesis. This is covered in the first chapters of any introductory course for statistics (e.g., Dytham 2011). It is commonly accepted that the threshold for statistically significant results is $p < 0.05$ and this often leads to thinking like "low p-value good, high p-value bad" (e.g., Dushoff et al. 2019). However, the p-value has nothing to do with the distance between two means/medians because, depending of the variation of studied traits, it may be a very sharp separation with small distance or a large overlap with large distances.

Finally, authors conclude that: "Taxonomy can be made on the carina sterni with geometric analysis. Less variation in shape was observed among birds belonging to the same habitat in taxonomic studies." However, they did not test the use of geometric morphometrics in the context of taxonomy. Thus, such conclusions, although they may be correct, are unjustified. Considering the above-mentioned problems, any discussion of their results is simply not supported by their study.

These are only a few problems with methodology and a few serious errors; however, this is not an exhaustive list. The described circumstances, however, give fair insight into the serious fail of the peer and editorial services. In my opinion, these flaws should be more than enough justification for hard rejection of the paper. However, that was not the case, and one can wonder how external reviewers and editors overlooked these cardinal errors and flaws.

Problems behind the case-paper

Several problems arise with such publications. First, it brings a mess into science and society, since misleading, erroneous results will be taken and referenced to as valid findings. The risk is especially high because published papers are usually read and accepted with a lesser degree of scepticism and critical thinking than unpublished findings, prior to the peer-review process. Although morphology of the avian sternum is not in the centre of public interest, in areas related to public health, it is enough to mention the infamous Andrew Wakefield paper linking vaccinations with autism and the impact is has for anti-science and anti-vaccinator movements (e.g., Bosques-Padilla & Gómez-Almaguer 2016). Although retraction of such papers may be seen as a partial cure for science, retracted papers are still cited and referenced to by other authors (Texteira da Silva & Bornemann-Cimenti 2017). Second, most research is financed via government grants and other public resources. Thus, it is important that the diminishing funding is not wasted on poorly designed and methodologically flawed research-like activities. Aside from eroding funds for such research, it may bring yet another consequence of extrapolation of poor-science over the whole science society. It is not impossible for some demagogues to use such poor papers in public debate as an example to justify further cuts of science funding.

Finally, it is devastating for younger generations of future scientists who see that scientific scrutiny may be just an empty-slogan, and even papers containing serious factual and methodological errors may be published in respected journals. Thus, why bother following scientific ideology if any paper can find its way into public scientific space? Going further, there is another crucial question: are students expected to respect an academic teacher publishing such flawed papers?

There is a discussion on the pros and cons of the peer-review system. Personally, I find it crucial. My paper is not to criticise the system as it is but to stress the need for taking peer duty responsibly, both at the level of peer-reviewers and editors as well as at the level of the university evaluation process. Peer service is time consuming and requires high train-

ing in critical thinking, wide knowledge and an open mind. As cases of failures in peer-review services show, the quality of peer services may be critical for societies far behind sophisticated issues, attracting scientists' attention in narrow, highly specialised areas. This puts even more responsibility on the reviewers and editors. However, when evaluation is coming, it is usually the number (not even their quality) of published papers or grants one won that is counted, not the other kinds of science-society service and activity. If peer-reviews are taken into account, it is again the number of such reviews that are included rather than the quality of the reviews. This leads to unethical predatory-review practices, recently discussed by Al-Khatib & Teixeira da Silva (2019a, 2019b) and Texteira da Silva (2020).

Post scriptum

All critical remarks regarding this paper were known to at least the corresponding authors of the Gündemir et al. (2023) paper prior to submission to *Anatomia, Histologia, Embryologia*. Although I feel it is fair to disagree with reviewers' comments when one is convinced to be right in his/her view, some mistakes are undisputable (e.g., caecilians as reptiles, sternums in sharks, p-values as distance measures, etc.). None of these have been considered or improved by authors. After the paper by Gündemir et al. (2023) appeared online as an early-view article, I attempted (twice) to contact the editor-in-chief of *Anatomia, Histologia, Embryologia* (via editorial office e-mail and directly to personal e-mail). However, I was denied the ability to publish a critique of the paper. Instead, the editor of *Anatomia, Histologia, Embryologia*, Dr. Fred Sinowatz (personal e-mail to BB, 06.02.2023), suggested to contact the corresponding authors of the paper by Gündemir et al. (2023) to ask them to send corrections for the paper. Given the seriousness of the allegations, I find such an approach unprofessional.

References

Al-Khatib A. & Teixeira da Silva J.A. 2019a. Rewarding the quantity of peer review could harm biomedical research. Biochemia Medica 29: 020201.

Al-Khatib A. & Teixeira da Silva J.A. 2019b. Is biomedical research protected from predatory reviewers? Science and Engineering Ethics 25: 293–321.

Benton M. J. & Donoughe P. C. J. 2007. Paleontological evidence to date the tree of life. Molecular Biology and Evolution 24: 26–53.

Bosques-Padilla F. J. & Gómez-Almaguer D. 2016. A new form of fraud in scientific publishing: supplanting or hacking the scientific review process. Medicina Universitaria 18: 1–2.

Dryden I. L. & Mardia K. V. 1998. Statistical Shape Analysis (Vol. 4). Wiley, Chichester.

Dushoff J., Kain M. P. & Bolker B. M. 2019. I can see clearly now: reinterpreting statistical significance. Methods in Ecology and Evolution 10: 756–759.

Dytham C. 2011. Choosing and Using Statistics. A Biologist's Guide. 3rd edition. Wiley, Blackwell.

Feneck E. M., Bickley S. R. B. & Logan M. P. O. 2021. Embryonic development of the avian sternum and its morphological adaptations for optimizing locomotion. Diversity 13: 481.

Ferguson C. Marcus A. & Oransky I. 2014. The peer-review scam. Nature 515: 480–482.

Gündemir M. G., Szara T., Spataru C., Demircioglu I., Turek B., Petrovas G. & Spataru M. C. 2023. Shape differences of the *Carina sterni* in birds of various locomotion types. Anatomia, Histologia, Embryologia 52: 190–196.

Harvey P. H. & Pagel M. D. 1991. The Comparative Method in Evolutionary Biology. Oxford Series in Ecology and Evolution. Oxford University Press, 239 p.

Haswell W. A. 1884. Studies on elasmobranch skeleton (Plates 1 and 2). Proceedings of the Linnean Society of New South Wales 9: 71–119.

Howes G. 1891. The morphology of the sternum. Nature 43: 269.

Inoue J. G., Miya M., Lam K., Tay B.-H., Danks J. A., Bell J., Walker T. I. & Venkatesh B. 2010. Evolutionary origin and phylogeny of the modern holocephalans (Chondrichthyes: Chimaeriformes): a mitogenomic perspective. Molecular Biology and Evolution 27: 2576–2586.

Kumar S., Stecher G., Suleski M. & Hedges S. B. 2017. TimeTree: a resource for timelines, timetrees, and divergence times. Molecular Biology and Evolution 34: 1812–1819.

Lowi-Merri T. M., Benson R. B. J., Claramunt S. & Evans D. C. 2021. The relationship between sternum variation and mode of locomotion in birds. BMC Biology 19: 165.

Parker T. J. 1890. Notes from the Otago University Museum. Nature 43: 141–142.

Pierce S., Ahlberg P., Hutchinson J., Molnar J. L., Sanchez S., Tafforeau P. & Clack J. A. 2013. Vertebral architecture in the earliest stem tetrapods. Nature 494: 226–229.

Rohlf F. J. & Slice D. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Systematics Biology 39: 40–59.

Sarewitz D. 2016. The pressure to publish pushes down quality. Nature 533: 147.

Schmidt-Nielsen K. 1984. Scaling: Why is animal size so important? Cambridge University Press, Cambridge.

Teixeira da Silva J.A. 2020. Are negative reviews, predatory reviewers or failed peer review rewarded at Publons? International Orthopaedics 44: 2193–2194.

Teixeira da Silva J. A. & Bornemann-Cimenti H. 2017. Why do some retracted papers continue to be cited? Scientometrics 110: 365–370.

Walker R. & da Silva P.R. 2015. Emerging trends in peer review–survey. Frontiers in Neurosciences 9: 169, 2015.

Zelditch M. L., Swiderski D. L., Sheets H. D. & Fink W. L. 2004. Geometric Morphometrics for Biologists. A Primer. Amsterdam et al.: Elsevier Academic Press.

Received: 15.11.2023
Accepted: 01.11.2024