

Reflections:

A biological method for diminishing the number of trees in phylogenetic studies

Luiz Ricardo L. Simone

Museu de Zoologia da Universidade de São Paulo
lrsimone@usp.br; lrsimone@gmail.com
OrcID: 0000-0002-1397-9823

Abstract

After several papers on phylogenetic analysis on Mollusca morpho-anatomy, along 3 decades, in different levels, the main criticism is not on the phylogenies themselves, but so in respect to the obtention of few or single cladograms. Here the secret is revealed: to seek for more characters. A brief discussion on this is performed, including tips on how to seek for more characters or states, and a discussion on how this noble professional activity has not been possible in present molecular approaches.

DOI:

Keywords: phylogeny, evolution, characters, matrix, database, morpho-anatomy.

Introduction

I have published dozens of papers on phylogeny in different branches of Mollusca, especially in gastropods and bivalves. These papers are dealing with phylogeny of Mollusca in general (e.g., Simone, 2009, 2011, 2021), up to papers dealing with phylogeny of close related species (e.g., Simone, 2006, 2022; Dornellas et al, 2020; Pastorino & Simone, 2021), as well as the phylogeny theory and philosophy themselves (e.g., Simone, 2020).

Sailing in a truly ocean of molecular phylogenies, I obstinately try to produce counterparts based on morpho-anatomy. Fighting against difficulties in research support, negatives of publication, depreciation of this kind of approach – pejoratively called as “traditional,” among other difficulties, it is expected that some feedbacks happen, mainly in the peer review processes. I was even

accused of being lazy of learning new techniques, as if molecular ones were terrible difficult to learn, and, on the contrary, morpho-anatomy was extremely easy.

However, along more than four decades of published papers, rare were the feedbacks on surveyed characters or even on resulted phylogenies, even though they have little congruences with those produced by molecular databases on the same taxa. On the other hand, the heaviest criticisms, mainly in peer-review processes, was that my phylogenetic results show few, or even single cladograms.

The referees did not save accusations. They usually credited the few or single cladogram, which supposedly should be a good factor, to artificial/fraudulent conduction of results, such as selection of characters (i.e., arbitrarily eliminating characters or states that cause noisy in the analysis), alignment of multistate characters, arbitrary weighting of characters, etc.

Of course, I never used such kind of artifices in my analyses, as much that all matrices are always available for reanalysis in all publications, and never a fraudulent contest was presented.

On the other hand, statistical juggling on the data matrix for decreasing the number of trees, which I also do not, or rarely use, is perfectly acceptable by the referees. Algorithms like strict consensus, majority consensus, successive weighting, etc., and even other kind of analyzes, such as the Bayesian and maximum likelihood, are amply used as mathematical, not biological, methods for decreasing the number of resulted cladograms to few, or even to one. These non-biological methods are accepted by biologists.

In this scenario, I think I can offer an additional method for decreasing the number of cladograms in a phylogenetic analysis, and reveal my secret:

- To seek for more characters -

Yes, is that simple. Of course, this is only, or mostly applicable to morphological approaches than to molecular ones, in which the data matrix is frozen.

In my experience, for a good resolution, a good number of characters is 4-times larger than the number of terminals (species usually). This is a good amount for diluting the influence of the homoplasies, as the majority of characters indicates the evolutive pathway. This 4-times more characters than taxa can be substituted by 4-times more apomorphic states than taxa, if there are much multistate characters in the analysis.

How to seek for more characters

In morphology, there is a saying that “between two characters, there is a third that you did not see.” Thus, seek for it. If you have a low-resolved region of a given cladogram, i.e., a polytomy, you can come back to your samples and try to extract 1 or 2 more characters or states. This mostly is sufficient to resolve the cladogram. Of course, despite you focus on the taxa of a given problematic region of your cladogram, you need to survey that character or state in all taxa of your analysis. Thus, you insert the character or state at the matrix and complete the states that all taxa have in it.

Sometimes you can extract information from some overlooked factor, like proportions, allocation of structures, interaction of organs and structures, topology, environmental or developmental features, etc., etc. As everything has potential of having a phylogenetic inference.

Thus, you must keep in mind the following topics:

1. It is always possible to add characters or states of characters, at least in morphological databases; mostly, the addition of 1-2 of them can resolve an unresolved portion of the cladogram.
2. Never select characters/states, i.e., eliminate those that are resulting in a topology that you do not agree with. The selection of characters is an induction of results, an ethical infringement.
3. If more characters or states are being included, they must be investigated in all taxa of your database.
4. It is indifferent if you insert 1 additional character or 1 additional state in an already existent character. At end, the result is the same.

Reflections

The extraction or search for characters in a morpho-anatomical study is the noble part of any project. It is the activity in which the professional uses his/her expertise, knowledge, coverage of taxa and readings, i.e., his/her biological wisdom. It is a brainstorm, exciting phase, in which the expert has an opportunity to meditate on each character, how has it evolved? Which is its ecological or physiological implications? Which were the environmental and genetic pressures that resulted in that? Etc., etc. Everything is fascinating issues that are stolen from the present generation of taxonomists. They obtain a mummified dendrogram from the computer, resulting from what the computer interprets as a database of 4 units (pyrimidines and purines), lacking synapomorphies capable of being analyzed.

Thus, the list of characters and the matrix of any phylogenetic analysis, at least in those based on morpho-anatomy, is not sacred, something immutable. On the contrary. The word “analysis” is just because everything must be analyzed, i.e., mentally, cerebrally processed, in order to offer a better, mostly resolved result. The result does not need to be impersonal, non-subjective, to approach Biology to the more technical sciences like physical and mathematical ones, something that only molecular approaches do. Do not be afraid to come back to your database and challenge it with the question: “what else am I not seeing?” Your brain, if well-prepared, will surprise you!

References

- Dornellas, AP; Couto, DR & Simone, LRL, 2020. Morphological phylogeny of the Tegulinae (Mollusca: Vetigastropoda) reinforces a Turbinidae position. *Cladistics* 36: 129-163. Doi 10.1111/cla.12400.
- Pastorino, P & Simone, LRL, 2021. Revision of the genus *Buccinanops* (Mollusca: Neogastropoda: Nassariidae), an endemic group of gastropods from the Southwestern Atlantic, including a new genus and accounts on the Buccinanopsinae classification *Journal of Zoological Systematics and Evolutionary Research* 56(6): 1209-1254. DOI: 10.1111/jzs.12479
- Simone, LRL, 2006. Morphological and phylogenetic study of the Western Atlantic *Crepidula plana* complex (Caenogastropoda, Calyptraeidae), with description of three new species from Brazil. *Zootaxa* 1112: 1-64. <http://www.moluscos.org/trabalhos/2006/Simone%202006%20Crepidula%20plana%20complex.pdf>

- Simone, LRL, 2009. Comparative morphology among representatives of main taxa of Scaphopoda and basal protobranch Bivalvia (Mollusca). *Papéis Avulsos de Zoologia* 49(32): 405-457. <http://www.moluscos.org/trabalhos/2009/Simone%202009%20Diasoma.pdf>
- Simone, LRL, 2011. Phylogeny of the Caenogastropoda (Mollusca), based on comparative morphology. *Arquivos de Zoologia* 42(4): 161-323. <http://www.moluscos.org/trabalhos/Caenogastro/Simone%202011a%20Caenogastropoda%20Phylogeny%20LIGHT.pdf>
- Simone, LRL, 2018. Phenotypic features of *Helicina variabilis* (Gastropoda: Neritimorpha) from Minas Gerais, Brazil. *Papéis Avulsos de Zoologia* 58: e20185832. <http://doi.org/10.11606/1807-0205/2018.58.32>
- Simone, LRL, 2020. Bosquejos de filogenia. Clube de Autores. Curitiba, 75 pp. <https://clubedeautores.com.br/books/search?where=books&what=bosquejos+de+filogenia>
- Simone, LRL, 2021. A proposal to improve the resolution of the high phylogeny and taxonomy of the Neogastropoda. *Malacopedia* 4(3): 31-37. <http://www.moluscos.org/trabalhos/Malacopedia/04-03Simone%202021%20Malacopedia-%20Neogastropoda.pdf>
- Simone, LRL, 2022. Additions to the genus *Anthinus* occurring in Minas Gerais and Goiás regions, Brazil, with description of five new species, one of them in the new related genus *Catracca* (Gastropoda, Eupulmonata, Strophocheilidae). *PLoS ONE* 17(8): e0273067. <https://doi.org/10.1371/journal.pone.0273067>