



The Ghost of Crises Past: A Reply to Mooi and Gill*

MATTHEW T. CRAIG

Department of Marine Sciences, University of Puerto Rico, Mayagüez, Puerto Rico, 00681, U.S.A. E-mail: matthew.craig@upr.edu

**In: Carvalho, M.R. de & Craig, M.T. (Eds) (2011) Morphological and Molecular Approaches to the Phylogeny of Fishes: Integration or Conflict?. Zootaxa, 2946, 1–142.*

The conclusion one might draw from all this is that the criterion used to define homology – i.e., the criterion of common ancestry – is a theoretical construct. Like truth, we must approximate it as best we can, and we have no touchstone to tell whether we have found it.---Colin Patterson, 1988: 621

Mooi & Gill (2010; hereafter M&G) presented a discussion of current methods in molecular systematics and contended that methodologies employed in the analysis of molecular data do not identify homology or synapomorphy, are phenetic in nature, and have resulted in a categorical movement “away from an intimate understanding of character distribution, homology, and the meaning of evidence” (M&G, p. 26). If they are correct in their assessment, systematic ichthyology (and systematic research in general) truly is in a state of crisis, and many of the advancements that have been made in the last two to three decades have, in reality, been steps backward. As with any discipline, papers pointing out the dangers of dogmatic thinking and the perils of uncritical analysis are often welcome “wake up calls” and serve a genuine and necessary purpose in scientific literature. Unfortunately, the paper by M&G appears nothing more than a rehashing of long-settled debates with not-so-subtle hints of deeply rooted animosity towards a few specific researchers. Their paper not only misses their intended mark (presumably to ensure the scientific integrity of systematic ichthyology), but also demonstrates a clear misunderstanding of how molecular data are accumulated, analyzed and interpreted.

The criticisms presented by M&G are numerous (so numerous, in fact, that it appears that the authors would be quite satisfied if genetic data were simply thrown out with the baby AND the bathwater) and many will be discussed in other papers in this issue. I will therefore focus on what I interpret as their chief concern, that molecular data and their resulting phylogenetic hypotheses do not provide evidence to support hypotheses of homology. Most of the comments herein are reviewed from literature written decades ago in which the theoretical basis for statements of homology in genetic data were discussed and associated concerns clarified, and in most cases, alleviated. Hypotheses of homology are at the heart of modern phylogenetic theory and systematic research; in fact, hypotheses of homology are central to all of comparative biology whether morphological or molecular. With the general acceptance of Darwin's thesis that all life is related through common ancestry, Sir Richard Owen's "homology" (Owen, 1843) has come to mean similarity *due* to common ancestry. This definition, however, does not highlight or aid in the discovery of homology, but rather is an explanation for it (Patterson, 1982). Homology may also be a relational term that defines groups; for example, the relation among winged forelimbs in tetrapods, genes in the genome, or base pairs in a gene (Forey et al., 1992). However one treats a statement of homology, there is a need to test the statement, and if a statement is testable, it is, by definition, a hypothesis. M&G repeatedly fail to recognize this important aspect of homology in their review, and seem to imply that they are more comfortable with *statements* of homology being made by those who have an “intimate understanding of character distribution” (M&G pg. 26). Does this imply that geneticists who have studied the structure, organization and evolution of the genome for decades do not have an intimate understanding of genes?

Since the dawning of the modern genetic era in biology, geneticists have studied genomes in ever increasing detail, and numerous orthologous genes have been identified. Orthologous genes and their sequences are homologues that reflect the descent of species (Fitch, 1970). They pass the congruence and similarity tests (Patterson,

1988), and are the equivalent of homologies in morphological data (Forey, et al., 1992). Beyond the orthologous gene, the more detailed sequence itself may be examined for a new set of homologous characters. Homology in sequence data may be detected in ways that are not possible with morphological data (Patterson, 1988). This is because of the one dimensional nature of gene sequences: an Adenine is an Adenine, a Cytosine a Cytosine, and so forth. If two sequences from an orthologous gene are aligned (the equivalent of creating a transformation series in morphological data), homologous base positions are inferred. If a set of bases is unique to a clade and involves transformation from an “ancestral” state polarized by outgroup comparison, it is also a synapomorphy.

I have tried to find flaw in this logic as the arguments presented by M&G are so vehement as to imply their truth. However there is no way to dispute the simple logic above. It may be that the issue lies with how sequences are aligned, the molecular equivalent of creating a transformation series, or how they are presented (or omitted) from publications. It is true that alignments of DNA data are usually not presented in many recent phylogenetic papers. This is an unfortunate demand made by editors in the interest of saving page space. In practice, however, it is usually quite simple to request an alignment from a researcher or to retrieve the raw sequence data from any number of publicly available databases (e.g., GenBank) and perform ones own alignment. If authors wish (and I hope that this becomes more commonplace), they can also submit their alignments as supplementary material made available through journal websites. Regardless of whether or not an alignment is readily available to an interested party, the ability for one to access the raw data used in molecular analyses far exceeds the accessibility of morphological data. As a representative example of a not uncommon practice in morphological studies, Johnson (1983:777) stated in his materials and methods “Cleared and stained specimens of representative genera from a wide variety of perciform families were examined for comparative purposes, including all nominal genera of the Percichthyidae and the serranid subfamilies Serraninae and Epinephelinae. Only the most pertinent material, that representing the Epinephelinae and the genera herein removed from the Serranidae, is listed.” Despite the enormous amount of attention to detail and the high quality of work by my colleague, no character matrix was presented that could be used to test his hypotheses of synapomorphy, and without a complete list of materials examined, it is nearly impossible to repeat his study or know the entire distribution of character states.

Another relevant point here is that, in most cases, a character is illustrated for ALL included taxa in a molecular matrix, but is never illustrated for all taxa in a morphological analysis. We assume that terms like “pointed” or “broad” are the same in every specimen examined in a morphological analysis, but we never get to judge that for ourselves unless we look at our own specimens. Add in this inherent subjectivity in discretizing morphological transformations (“intimate knowledge”?), and the pattern is clear: molecular data are less subjective and often far more transparent than are morphological data.

M&G are also concerned by the observation that when using optimization methods “topologies will change with the addition of taxa as the combination of character states changes and new optimal arrangements result – what were once homologies become homoplasies and vice versa” (M&G pg. 30). Of course they do. Hypotheses of homology are entirely dependent on the creation of a robust transformation series and a phylogenetic tree to test the hypotheses. If one adds taxa to an existing matrix, and new data derived from them help us to realize a more cogent and logical transformation series, why should we be tied to our previous hypothesis? The scientific method demands that we change our hypotheses when new data emerge that are contrary to the original hypothesis. Complaining that topologies change in the face of new data ignores this central concept in science.

Perhaps the most disheartening aspect of M&G’s position is evident in one of their recent publications. Gill and Mooi (2010) presented character evidence for the monophyly of the Microdesminae (a subfamily of goby fishes). Despite their concerns about the lack of homology and synapomorphy in molecular studies, the excellent drawings of osteological features and thorough review of past and recent systematic treatments of the Microdesminae, they provided no *test* of their hypotheses of homology. There is no phylogenetic tree, no explicit transformation series, and no identification of a sister group by which their characters were polarized. There are simply *statements* made by two experienced researchers with “intimate knowledge” of the character distributions within the Microdesminae.

This last point is particularly troubling; how do they know that their synapomorphies are valid without comparison to a sister group? The characters may very well be present in the sister group or have been reversed in a subclass that they excluded. Similarly, in the study mentioned above, Johnson (1983) listed synapomorphies for the family Serranidae (groupers and seabasses) in the same type of study, and noted that all were present in the genus *Trachinus* (then recognized in a separate family); but Johnson excluded *Trachinus* from his serranids with

little-to-no justification. Using molecular data, Smith and Craig (2007) not surprisingly found a monophyletic group comprised of *Trachinus* and all other “serranoid” genera examined in their study. This is not meant to dismiss or discredit the valuable work of Gill, Johnson or Mooi in any way; rather it is said to point out that investigators wishing to criticize others for not providing synapomorphies, etc., should hold themselves to the same standards.

The contentious debate surrounding the use of molecular versus morphological data harkens back to the movement away from evolutionary taxonomy to Hennigian argumentation and its derivatives used today. It is complete with the vociferous outbursts at society meetings, the scorn for “others”, and distrust of anyone using the “new” data. Casting aspersions, accusing investigators of presenting “manufactured morphological ‘evidence’” (M&G pgs. 34,35), and calling into question genetic analyses by making broadly generalized and uniformed statements, however, is nothing more than scare tactics similar to those used by Edison to convince the world that Tesla’s alternating current would surely cause death and destruction.

As a discipline, systematic ichthyology is losing talented students at an alarming rate. When asked, many students say that they do not want to enter such a cutthroat field; none wish to become the new targets (pers. obs.). Luckily, a majority of investigators in systematic biology have either accepted genetic data as a valid, alternative data type, or have embraced it wholeheartedly. However if this extreme animosity by prominent members of our field towards those who use genetic data in ichthyology persists, we will lose the war of attrition, and there will be no one to replace us in the future. This, I believe, is the true crisis in systematic ichthyology.

Acknowledgements

I thank the many researchers, morphologists and geneticists, who have engaged in fruitful, and at times heated, debate which has helped me to clarify my ideas presented herein. B. Erismann, P. Hastings, R. Mayden, J. McCosker, L. Rocha, L. Smith, C. Thacker, and E. Wiley provided useful reviews of this manuscript.

References

- Fitch, W.M. (1970) Distinguishing homologous from analogous proteins. *Systematic Zoology*, 19, 99–113.
- Forey, P.L., Humphries, C.J., Kitching, I.L., Scotland, R.W., Siebert, D.J. & Williams, D.M. (1992) Cladistics: A practical course in systematic. The Systematics Association Publication No. 10. Clarendon Press, Oxford.
- Gill, A. & Mooi, R. (2010) Character evidence for the monophyly of the Microdesminae, with comments on relationships to *Schindleria* (Teleostei: Gobioidae: Gobiidae). *Zootaxa*, 2442, 51–59.
- Johnson, G.D. (1983) *Nippon spinosus*: A primitive epinepheline serranid, with comments on the monophyly and intrarelationships of the Serranidae. *Copeia*, 1983, 777–787.
- Mooi, R. & Gill, A. (2010) Phylogenies without synapomorphies —a crisis in fish systematic: Time to show some character. *Zootaxa*, 2540, 26–40.
- Owen, R. (1843) Lectures on the comparative anatomy and physiology of the invertebrate animals. Longman, Brown, Green and Longmans, London.
- Patterson, C. (1982) Morphological characters and homology. In: Joysey, K.A. & Friday, A.E. (Eds.), *Problems in Phylogenetic Reconstruction*, Academic Press, London, pp. 21–74.
- Patterson, C. (1988) Homology in classical and molecular biology. *Molecular Biology and Evolution*, 5, 603–625.