The case for sensitivity: a response to Grant and Kluge

Gonzalo Giribet[a]* and Ward C. Wheeler[b]

[a]*Museum of Comparative Zoology, Department of Organismic & Evolutionary Biology, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA;* [b]*Division of Invertebrate Zoology, American Museum of Natural History, Central Park West at 79th Street, New York, NY 10024-5192, USA*

Recently, Grant and Kluge (2005) have criticized sensitivity analysis *sensu* Wheeler (1995) and its justification by Giribet (2003). Grant and Kluge state that sensitivity analysis "is neither scientific nor heuristic" (p. 603), and therefore "remains a method in search of scientific justification." (p. 603).

Much of Grant and Kluge's criticism is based on their personal views ("our philosophy" p. 598) and their definitions of science and scientific utility. This leads them to a series of assertions by which other methods, in this case sensitivity analysis, are evaluated. Certainly, these methods may come up short in their eyes, but that is not of universal concern (e.g., Miller and Hormiga, 2004). Yet clearly our arguments have not been sufficiently precise. Here we will try to remedy this shortcoming.

What is parsimony?

Parsimony can be described generally as follows. A cladogram or tree, *T* is defined by a set of vertices *V* and edges or branches *E*.

$$T = (V, E)$$

Observed data (*D*) are a subset of all possible observations (*D'*).

$$D \subseteq D'$$

The transformation cost σ specifies an edit cost between all potential observations (states).

$$\sigma : D' \times D' \to \mathbb{R}$$

χ is the Erdös and Székel (1994) mincut-maxflow operator that parsimoniously assigns states (*D'*) given a topology *T* and its vertices [*V(T)*].

$$\chi : T \times V(T) \to D'$$

The most parsimonious cost of a tree is then $T(\sigma, \chi)$. As χ depends only on the observed data (*D*) given *T*, $T(\sigma, \chi)$ can be replaced by $T(\sigma, D)$. The "best" or most parsimonious cost *C* is derived from the topology that minimizes the tree cost.

$$C = \min_T T(\sigma, D)$$

So far, we should all agree. Our differences arise from the transformation weight set σ. Grant and Kluge (2005, citations therein) state that these values must all be unity. We disagree. Clearly, this is not a mathematical restriction and there are mathematical arguments that require this not be so, but more on that later.

What is a phylogenetic hypothesis?

A phylogenetic hypothesis is a topology *T* given some data *D* and a model or weight scheme σ.

$$T(\sigma, D)$$

Again, σ is the crucial difference. *D* is given, and we usually care most about *T*, but we cannot separate it from σ. In a probabilistic world, we could simply perform a minimization jointly over tree and weight-scheme space, keeping the minimum. With parsimony, however, we cannot perform this global minimization. The σ-values are not absolute natural statements, but interpretive values that allow us to optimize variation on a topology. The optimality values based on different σ scenarios are not numerically comparable. Each σ will yield its own *T*.

Any fixed σ (not only σ = 1), causes this problem to disappear. This may be simplifying, but is not a requirement of the formalism.

Corresponding author:
E-mail address: ggiribet@oeb.harvard.edu

What are we testing?

At its core, a phylogenetic test is a comparison between alternate hypotheses, adjudicated by optimality score (cost, i.e., $T(\sigma, D)$ above). This test, however, as mentioned above and in opposition to Grant and Kluge (2005) is a comparison of both topology (T) and model (σ) given the same set of observations (D). σ is an ineradicable component of the optimality function. Simply because we cannot explicitly perform the joint minimization does not remove the requirement. Clearly, if σ is fixed to any set of values (e.g., 1), this component is removed and the topology comparisons can be made directly and optimality scores are entirely comparable.

Sensitivity

The notion of sensitivity analysis is derived from decision theory, incarnated here, the decision to accept a hypothesis. Given that σ parameters are likely to be unknown, any topological choice from among the universe of parsimonious topologies may or may not be tightly linked with the specific choice of σ-values. Hence, the attraction of *robust* decisions.

> By *robustness*, we mean the sensitivity of the decision to assumptions (in the analysis) that are uncertain (Berger, 1985, p. 69)

Sensitivity analysis is the study of robustness. This has nothing to do with *stability*, defined here and generally as variation in outcome due to changes in data. Robustness is not a test *sensu* Grant and Kluge (2005), as all the σ-specific topologies are the results of observation-based tests. Nor is it "heuristic" by their definition, as it does not lead to bold, novel notions. Yet, it is a useful even necessary means of understanding the relationship between assumption and result.

Stability

Grant and Kluge's (2005, pp. 600–601) criticism of stability as a goal of systematics is largely irrelevant. We have not argued for the preference of taxonomic stability over scientific methodology, hence, the arguments of Giribet (2003) stand. First, we would rather not introduce a taxonomic rank for a clade that is unstable to parameter choice—contrary to their argument that "all groups supported by the available evidence deserve to be reflected in formal taxonomy, regardless of their relative degrees of support" (Grant and Kluge, 2005, p. 601). We find this an irresponsible practice given that in taxonomy, unlike other disciplines, poor science cannot be ignored. This has resulted in

profligate name generation, cluttering up the literature as these taxon names are immortal (Wenzel, 1997; Prendini and Wheeler, 2005). Revising taxonomy based on an optimal, but unstable result is not "incorrect" in any sense, but does not promote the unambiguous conveyance of information via naming that taxonomy embodies. This caution does not bow to convention but exemplifies the different roles systematic analysis and taxonomy play. It escapes our understanding as to why the alternative of naming only stable and/or well-supported clades would serve "to promote paraphyly and social convention over evidence of monophyly" (Grant and Kluge, 2005, p. 601). They clearly confound consideration with authoritarianism.

The second argument of Grant and Kluge (2005) against stability (fiat or polytomies) misses the point. The issue is not whether stability should be the goal of systematics (Grant and Kluge, 2005), but rather whether it is a desirable property for the analysis of a particular data set (Giribet, 2003).

Congruence as optimality

Given that all T_σ are the results of tests *sensu* Grant and Kluge (2003) in that observations have been brought to bear to distinguish among hypotheses, that case where $\sigma = 1$ is no more nor less a test than any other. Based on this equality of merit, Wheeler (1995) stated that the presentation of such a sensitivity space might well be the end-point of analysis. Wheeler (1995, 1999; Wheeler and Hayashi, 1998; Wheeler et al., 2006) and Giribet (2003) among others, have argued for congruence-based measures to choose a "best" T from among all the T_σ.

Although there is no consensual measure of congruence, the notion of character congruence as an optimality criterion is certainly inherent in parsimony. That cladogram that minimizes character congruence (measured in a partition-free character-based manner, e.g., Wheeler et al., 2006) will also be the most parsimonious. Such a criterion can be applied *ab initio* via a search over both topological and model space even though we tend to separate these steps for computational reasons.

Whether or not investigators employ some form of congruence-based optimality (with or without data partitions), the edit cost matrix is an ineradicable component of cladogram optimality.

Metricity and transformation

Metricity places limits on the relative costs of transformations (Fig. 1a). In this sense, metricity reduces the

$$d\,(A,C) \leq d\,(A,\text{-}) + d\,(\text{-},C)$$
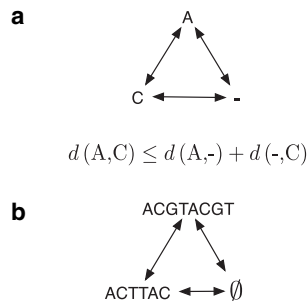
$$d\,(\text{ACTTAC},\text{ACGTACGT}) \nleq d\,(\text{ACTTAC},\emptyset) + d\,(\emptyset,\text{ACGTACGT})$$

Fig. 1. Metricity among three states. (a) Triangle inequality limits on indel transformation cost [d(A,) and d(–,C)]. (b) non-metric transformation case where nucleotide and indel costs are set to unity (1) as well as locus origin and loss [d(ACTTAC,$\emptyset$) and d($\emptyset$,ACGTACGT)].

universe of possible σ-values. The simple case of sequence insertion and deletion is an example (Wheeler, 1993) where the cost of an indel must be no less than one-half the transformation cost between any two nucleotides. Obviously, σ = 1 fits this stricture.

When we enlarge the world of possible transformations, however, homogeneous weighting may no longer satisfy this requirement. Consider the case of three sorts of transformation: nucleotide substitution, nucleotide insertion–deletion, and locus or gene insertion–deletion. Given the locus complement variation in organellular and genomic data sets, this is not a contrived situation. Further consider a case where there are homologous loci in two taxa (A = ACTTAC and A′ = ACGTACGT) that vary sequence and length (Fig. 1b). If we apply the σ = 1 weighting scheme, the transformation cost between A and A′ will be simply the number of indels and substitutions to interconvert the two. If locus insertion–deletion costs are also 1, the distance between A and A′ must be no greater than 2. This is due to the lower cost path of A → $\emptyset$ → A′ (always 2) versus A → A′. Given that observed loci often differ by more than two bases substitutions, σ = 1 will inevitably lead to non-metric, illogical results (e.g., all sequence transformation mediated by locus indels).

Summary

Our points are few and simple. Grant and Kluge (2005) have asserted a philosophy of science that excludes differential transformation weighting. This view removes sensitivity analysis from science. We claim that differential weighting is a necessary component of parsimony-based systematics and sensitivity analysis its logical investigation. Additionally, we show that very simple situations can lead homogeneous weighting into non-metric transformation costs and meaningless results.

References

Berger, J.O., 1985. Statistical Decision Theory and Bayesian Analysis, 2nd edn. Springer-Verlag, New York.

Erdös, P., Székel, L.A., 1994. On weighted multiway cuts in trees. Math Programming, 65, 93–105.

Giribet, G., 2003. Stability in phylogenetic formulations, and its relationship to nodal support. Syst. Biol. 52, 554–564.

Grant, T., Kluge, A.G., 2003. Data exploration in phylogenetic inference: scientific, heuristic, or neither. Cladistics 19, 379–418.

Grant, T., Kluge, A.G., 2005. Stability, sensitivity, science, and heurism. Cladistics, 21, 597–604.

Miller, J.A., Hormiga, G., 2004. Clade stability and the addition of data: a case study from erigonine spiders (Araneae: Linyphiidae, Erigoninae. Cladistics, 20, 385–442.

Prendini, L., Wheeler, W.C., 2005. Scorpion higher phylogeny and classification, taxonomic anarchy, and standards for peer review in online publishing. Cladistics, 21, 446–494.

Wenzel, J.W., 1997. When Is a Phylogenetic Test Good Enough?, Vol. 173. Origin of Biodiversity in Insects: Phylogenetic Tests of Evolutionary Scenarios. Mémoires du Muséum National d'Histoire Naturelle, Paris.

Wheeler, W.C., 1993. The triangle inequality and character analysis. Mol. Biol. Evol. 10, 707–712.

Wheeler, W.C., 1995. Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. Syst. Biol. 44, 321–331.

Wheeler, W.C., 1999. Measuring topological congruence by extending character techniques. Cladistics, 15, 131–135.

Wheeler, W.C., Hayashi, C.Y., 1998. The phylogeny of the extant chelicerate orders. Cladistics, 14, 173–192.

Wheeler, W.C., Ramírez, M.J., Aagesen, L., Schulmeister, S., 2006. Partition-free congruence analysis: implications for sensitivity analysis. Cladistics, 22, 256–263.