

## On the deduction of species relationships: a précis

Arnold G. Kluge\*

*Museum of Zoology, University of Michigan, Ann Arbor, MI 48109, USA*

Accepted 16 April 2003

If, contrary to my view, Popper's program cannot be executed in its purity, we may try something entirely different, with the hope that we shall not do injustice to the value of creativity in free imagination and rigorous criticism.

Agassi (1975, p. 35)

The deductive inference of phylogeny, first dealt with in some detail by Wiley (1975), continues to be called into question (for an early negation see Hull, 1983; but for a change of mind see Hull, 1999, p. 481). For example, Cleland (2002, p. 483) concludes "there is little in the practice of historical science that resembles what is prescribed by falsificationism." Olmstead (2001, p. 304) is even of the opinion that "Popper never addressed the question of how his philosophical framework related to the singular history of evolutionary diversification or to our ability to recover that pattern." Rieppel (2002, p. 30, 2003) is somewhat more specific in stating that "Popper's philosophy is not applicable to cladistic analysis of phylogenetic relationships, because ... there is no deductive link between a hypothesis of relationships and the character distribution on a tree ..." In his search for Kuhnian values (Kuhn, 1962), Richards (2002) claims to have discovered that cladistic parsimony cannot function as an unequivocal rule, because different cladograms depend on "the social factors that influence character individuation." Then there is de Queiroz and Poe (2001; see also Faith and Trueman, 2001), who maintain that Popper's degree of corroboration correctly interpreted is an inductive formalism, thereby tacitly denying an explication of falsificationism (Kluge, 2001). In the hope of achieving more critical discussions in the future, I describe phylogenetic systematics in hypothetico-deductive terms, sufficient I believe for it to be evaluated as a basis for

obtaining objective knowledge of species history. Philosophical formalisms and related arguments are placed in footnotes, or ignored altogether, with the bulk of the text consisting of the most familiar terminology and simplest examples of phylogenetic systematics that I can manage while still making my arguments.

Whether or not all scientific discoveries are united by the same underlying logic, there is a fundamental difference between the universal (predictive) and the historical (postdictive) sciences, the nomothetic and the idiographic, respectively. As Popper (1957, p. 143, italics in the original) underscored,

It is the distinction between the interest in universal laws and the interest in particular facts. I wish to defend the view, so often attacked as old-fashioned by historicists, that *history is characterized by its interest in actual, singular, or specific events, rather than laws or generalizations.*

With that distinction in mind, Popper (1957, p. 146) went on to state the condition for deducing historical events:

In so far as we are concerned with the historical explanation of typical events they must necessarily be treated as typical, as belonging to kinds or classes of events. For only then is the deductive method of causal explanation applicable.

That phylogenetic systematics exemplifies this method of inference may be judged from the following sets of arguments, 1–11, where *b* is background knowledge (descent, with modification), *c* is character, *e* is evidence (transformation series), *h* is hypothesis (cladogram, phylogenetic tree) and *n* is number of terminal taxa (species or monophyletic groups of species).

1. Darwin's (1859) first two principles of evolution, "descent, with modification," have long been considered sufficient, as background knowledge, *b*, in providing deductively testable hypotheses of species

\* Fax: +734-763-4080.

E-mail address: [akluge@umich.edu](mailto:akluge@umich.edu).

relationships.<sup>1</sup> In light of just these premises an  $h$  can be stated as ‘species A and B share a more recent common ancestral species than either does with species C,’ which reduces to the familiar expression (A,B)C (Hennig, 1966, p. 71; Lidén, 1990). Thus, the origin of species or monophyletic groups of species can be considered a “kind” of historical event.<sup>2</sup>

While the relative condition of species common ancestry,  $h$ , is specific enough to form a typical statement,<sup>3</sup> to assume Darwin’s first principle alone, “descent,” as  $b$ , is not a sufficient major premise (*contra* de Queiroz, 1992, p. 305; de Queiroz and Gauthier, 1992, p. 452), because descent is the property of the class of all monophyletic groups—the condition of any *particular* group or corresponding name being undefined. A polytomy is the most that one can hypothesize while assuming just “descent.” Further, there can be no empirical basis,  $e$ , for testing  $h$  without assuming something more as  $b$ , such as Darwin’s second principle, “with modification.”

Explicating phylogenetic systematics as a deductive science treats the kind of historical event in question as a numerical universal, not as a strictly universal (Stamos, 1996; see also Hull, 1983, p. 180; Kluge, 1999), where each of the typical statements, (A,B) or (A,C) or (B,C), stands for a finite, spatio-temporally restricted, and historically connected, group of things, and for which natural necessity, or lawfulness, cannot be claimed consequently (*contra* Cleland, 2002; *contra* Rieppel, 2002). While numerical universals may be verifiable in principle (Kitts, 1977, p. 189), it is clear that the typical kind of statement at issue in phylogenetic inference cannot be verified directly, by observation, because of the time passed, nor can it be estimated with subjective or objective (of the frequency kind) probability,<sup>4</sup> because each statement is necessarily unique (Kluge, 2002;

see also Hennig, 1950, p. 114). As Patterson (1978, p. 219) alludes to, “[t]he point at issue is ... not the distinction between universals and numerical universals, but between universals and [historical] individuals.”

Neither can this kind of typical statement be considered a natural kind. In the traditional sense of natural kind (Mill, 1843), there is no a priori definable intrinsic essential property of species. The cause of common ancestry can be described only with regard to indeterminate initial conditions, i.e., all of the unpredictable aspects of the situation of which the event is a part (Kluge, 2002). Boyd’s (1990, 1991, 1999) attempt to avoid the taint of traditional essentialism by specifying underlying causal homeostatic mechanisms of species, and groups of species, determined a posteriori as cluster kinds, is equally unsatisfactory. The a posteriori definition still leaves open the question “what makes a homeostatic mechanism a homeostatic mechanism for a particular kind” (Ereschfsky, 2001, p. 107). Moreover, whatever is “homeostatic” cannot, by definition, evolve!

2. A statement of  $h$  is but one part of a closed hypothesis set, as determined by the number of terminal taxa,  $n$  (Siddall and Kluge, 1997, p. 313).<sup>5</sup> Although there are not indefinitely many  $h_n$  to be falsified, the epistemic status of claims of relative recency of common ancestry depends on recognizing the singular nature of history—each species and transformation event being necessarily unique, and the alternative statements of  $h_n$  being *exclusive*, such as (A,B)C or (A,C)B or (B,C)A. An *inclusive* set of  $h_n$  statements, such as (A,B)C and (A,C)B and (B,C)A, which have been used to conditionally probabilify a hypothesis of common ancestry, would therefore represent an illogical thesis (Kluge, 2002).<sup>6</sup> There cannot be a relevant “situation” with regard to species where “[n]ature sometimes repeats herself, presenting researchers with multiple examples of a given type of historical phenomenon” (Cleland, 2002, p. 485), because each event of common ancestry is necessarily unique (Kluge, 2002).<sup>7</sup>

3. For any  $h_n$  there are  $n - 2$  basic, testable, statements, which represent the phylogenetic informativeness of the cladogram.<sup>8</sup> Not only is the testable phylogenetic

<sup>1</sup> Premises known to be false cannot serve as  $b$ , and therefore assuming just “descent, with modification,” is not considered *problematic* (Siddall and Kluge, 1997, p. 320). Further, the hierarchy of relative recency of common ancestry, given this interpretation of  $b$ , cannot be judged an a priori truth in phylogenetic systematics (*pace* Brady, 1994, p. 22). On the other hand, models are *problematic* and cannot be part of the auxiliary assumptions of phylogenetic systematics (see argument set 5, below), because they are counterfactual conditionals, of the general form ‘if  $p$  were to have happened  $q$  would have happened, where the supposition of  $p$  is contrary to the known fact not- $p$ .’

<sup>2</sup> The conditions defining a class or a kind, as opposed to those of a ‘natural’ kind, need not be lawfully related (Mahner and Bunge, 1997).

<sup>3</sup> A particular that instantiates a class concept is considered “typical” of that class. For example, a monophyletic Squamata is typical of the concept “clade,” just as a monophyletic Amniota is another instance of that class.

<sup>4</sup> This even includes maximum likelihood, where the inference involves determining whether the present evidence could have reasonably come from the model if the hypothesis is correct. Although a particular maximum likelihood estimate is a point probability, it is determined none-the-less in reference to the conditional (frequency) probabilities of the model, which are counterfactual conditionals (see Footnote 1).

<sup>5</sup> The set of  $h_n$  for any  $n$  is  $(2n - 3)!/2^{n-2}(n - 2)!$  (Edwards and Cavalli-Sforza, 1964, p. 73).

<sup>6</sup> Objective probability was not rejected out-of-hand, as a pseudo-problem (Grünbaum, 1989; Popper, 1983, p. 160), in the logical arguments that led me to judge  $h$  as necessarily unique (Kluge, 2002). Rather, I simply pointed out that deductive inference is not “burdened” by the problem of uncertainty (Watkins, 1984). What may be relevant to one epistemology, such as statistical consistency and model parameterization is to inductive inference, does not automatically carry over to another.

<sup>7</sup> Even the instrumentalist’s heuristic use of probability as the basis for prediction, and not explanation, cannot be based on an illogical thesis (Ariew, 1998).

<sup>8</sup> This is equivalent to the empirical content of a hypothesis in Popperian testability,  $1 - p(h, b)$ .

information of each  $h$  of a set of  $h_n$  precisely the same, all of the alternative statements for a given  $h_n$  are equally bold and simple (Popper, 1963, p. 61; 1972, p. 81, 143; 1982, p. 131). For example, for  $n = 3$  there is only one competing statement, (A,B) or (A,C) or (B,C). These preconditions of testability are not considered troublesome, because the phylogenetic system can ultimately yield causal explanations (see argument set 11, below). Given that increasing  $n$  increases the number of testable statements is reason enough, in the logic of scientific discovery (*sensu* Popper), to include more terminal taxa in a phylogenetic analysis of relationships (Kluge, 2001).

Since the number of competing  $h_n$  and their empirical content are matters of pure logic,<sup>8</sup> not the logic of scientific discovery, it is misleading to say that phylogenetic hypotheses are “reconstructed” or trees are “created.” To whatever extent creative genius may be involved in phylogenetic systematics (Popper, 1959, p. 32), it is limited to the invention of more efficient and effective methods for choosing among the logically possible alternatives for large numbers of  $n$ .

4. All typical statements of  $h_n$  are *prohibitive*, such as (A,B)C and not (A,C)B or (B,C)A, as required of a testable, falsifiable, hypothesis (Kluge, 2002). These kinds of propositions might be formulated as negative existential statements, where the falsifiability of the hypothesis is made clearer. For example, one could say that ‘a cladogram (A,B)C does *not* exist on which statements of transformation,  $e$ , can be causally explained as statements of homology’ (Kluge, 1999). However, such an exercise contradicts the logic that (A,B)C does exist, as a hypothesis.

5. Testability is a function of the improbability of  $h$  (e.g., Popper, 1983, pp. 236–240). Paraphrasing Kluge (1997, p. 88), assuming *only* “descent, with modification,”  $b$ , the evidence,  $e$ , for (A,B), (A,C), and (B,C) should be equally likely. However, if a large majority of one of those possible kinds of  $e$  were to be observed in an unbiased sample, say that which counts for (A,B), then this is improbable given only that  $b$ ,<sup>9</sup> but not under that  $b$  plus the postulated rooted cladogram (A,B)C. By extension, we can say that such a result is expected if (A,B)C were to be true, but not if it was false, and it is

under these circumstances of improbability that  $e$  can be considered to support  $h$  (see argument set 9, below).<sup>10</sup>

The improbability argument and the concept of support in Popperian testability are conditional on what is assumed. As Popper (1963, p. 61; see also Popper, 1959, pp. 272–273) pointed out, the simplest  $h$  that can be stated, in light of  $b$ ,

offer us the best chance to submit them to severe tests: the simpler theory has always a higher degree of testability than the more complicated one.

As set forth in argument sets 1–3 (above), it takes no more than Darwin’s first two principles of evolution to formulate a testable  $h_n$ , such as (A,B)C or (A,C)B or (B,C)A, and the hypothesis of  $e$  that is used to test any  $h_n$  (see argument set 6, below). What could be simpler than testing the set of such statements,  $h_n$ , where only one can be true, in light of  $b$ ?

Related to the effect that simplicity has on the degree of testability is the problem of *ad hoc* hypotheses (see argument set 10, below; Footnote 16). As Popper went on to argue:

One can show that the methodology of science (and the history of science also) becomes understandable in its details if we assume that the aim of science is to get explanatory theories which are as little *ad hoc* as possible: a ‘good’ theory is not *ad hoc*, while a ‘bad’ theory is. On the other hand one can show that the probability theories of induction imply, inadvertently but necessarily, the unacceptable rule: always use the theory which is the most *ad hoc*, i.e. which transcends the available evidence as little as possible.

For example, conformity to a particular phylogenetic hypothesis is frequently argued in terms of *ad hoc* additions to the system, such as other assumptions and evidence. Differential character weighting is *ad hoc* when those schemes presuppose what cannot be tested, independent of the phylogenetic analysis. Furthermore, weighting of any kind, a priori or a posteriori, correlates  $e$  by class, in reference to what is “good” or “bad,” but that violates the assumption that each  $e$  be an independent test of  $h_n$  (Kluge, 1998).

6. The transformation series concept of Hennig (1966, Fig. 21) constitutes evidence,  $e$ .  $e$  is temporal and historical, not because of ontogeny or tokogeny, but because the concept of species, of which  $e$  is a part, imposes it (Lidén, 1990). Thus, metaphysically speaking, there is a lineage system of change, i.e., one of evolution.

<sup>9</sup> This *desideratum* excludes all those model assumptions whereby  $e$  would be made more probable (see Footnote 1). Although (e.g., Popper, 1983, p. 238; italics in the original) used the calculus of probability to introduce his concept of improbability, if  $e$  is to support  $h$  it should not be probable on the background knowledge  $b$  alone by  $p(e, b) \ll 1/2$ , he went on to observe that this simply amounts to saying “the smaller  $p(e, b)$ , the stronger will be the support which  $e$  renders to  $h \dots$ ”, or with regard to a significant degree of corroboration,  $p(e, hb) - p(e, b) \gg 1/2$ . Likewise, my argument for the improbability of phylogenetic hypotheses can be stated in terms of conditional or logical probability (Kluge, 1997, p. 88; see also Kluge, 2001). See Footnotes 10, 14, and 15.

<sup>10</sup> This is severity of test in Popperian testability,  $S(e, h, b) = p(e, hb) - p(e, b)/p(e, hb) + p(e, b)$ . Popperianism does not have an absolute criterion for knowing the truth, but specifying the conditions for truth is a trivial exercise. For example, in the phylogenetic system, given three terminal taxa, the statement “(A,B)C is true” iff A and B share a more recent common ancestor than either does with C; “(A,C)B is true” iff A and C share a more recent common ancestor than either does with B; “(B,C)A is true” iff B and C share a more recent common ancestor than either does with A. See Footnote 9.

The basic statements that phylogenetic systematists actually use to test  $h_n$  are the observable outcomes of a hypothesized spatio-temporally restricted change, the independent heritable modifications that distinguish species from one another. These are phenotypic traits, the plesiomorphic/apomorphic characteristics of organisms that are the result of that kind of causal determinism. There is no one operational criterion for determining these states. It can be anything that leads to a testable hypothesis (Popper, 1959, p. 32), which in the present case are the characteristics of different species that are perceptually the *same thing*,  $c_s$ .<sup>11</sup> Ostension, but not intensional definition, such as that used to determine similarity relationships, may be employed in these efforts, because it is logically consistent with what is to be determined. While the a priori testability of these “thing hypotheses” remains little investigated, it is generally acknowledged as involving character analysis/reanalysis, where the tests of character compatibility and conjunction may be applied (Kluge, 1998).<sup>12</sup> Arguably, as the result of such a priori studies,  $c_s$  will become more severely tested for its objective reality, and consequently the more severely  $e$  can then be said to test  $h_n$  (see also argument sets 7 and 11, below).

Hennig (1966, p. 65) defined semaphoront, the character “bearer,” as the organism, which is in a certain theoretically infinitely small time span of its life, during which it is considered unchangeable. The importance of the concept of semaphoront in the present context lies in the fact that “in a system in which the genetic relationships between different things that succeed one another in time are to be represented, we cannot work with elements that change with time.” More generally, only by insisting on semaphoront comparisons can  $c_s$  be tested and causally explained historically. While topological similarity may be one of the criteria employed in phylogenetic systematics, in identifying character state relations among semaphoronts representative of different species, the pattern cladists’ principle of the topological invariance of taxa is not assumed (e.g., Brady, 1994, p. 27f). Obviously, social factors do not then necessarily influence the individuation of characters in phylogenetic systematics and thereby render the

application of parsimony “arbitrary or subjective” (*contra* Richards, 2002, p. 17).

It is common knowledge that only synapomorphies,  $c_s$ , not autapomorphies or symplesiomorphies, constitute tests of  $h_n$ , because only those empirical relations have the potential to falsify that kind of  $h$ . More specifically, an objective criterion of falsification resides in incongruent  $c_s$ , because not both can be true. Incongruent  $c_s$  constitute evidence for a different cladogram, only to the extent they have been sampled in an unbiased manner (see argument set 5, above).

The following explicates  $e$  as a test of  $h_n$ .<sup>13</sup> Assume a rooted three-taxon cladogram, (A,B)C, and three hypothesized binary-coded transformation series distributed as 110, 101 and 011 (state 1 is apomorphic). Thus, the congruent  $c_s$  has the taxonomic distribution A<sub>1</sub>, B<sub>1</sub>, C<sub>0</sub>, while the incongruent  $c_s$  have the distributions A<sub>1</sub>, B<sub>0</sub>, C<sub>1</sub> and A<sub>0</sub>, B<sub>1</sub>, C<sub>1</sub>. Next, consider the alternative conclusions which follow logically from *conjoining* the cladogram (A,B)C with the congruent and incongruent  $c_s$ : (conclusion 1) a cladogram *alone* does not declare the derived states of a congruent  $c_s$  are homologues (are of a common origin). As Sober (1988, pp. 132–135) correctly points out, ‘something more is needed’. (conclusion 2) However, a cladogram by itself does declare the derived states of an incongruent  $c_s$  are of independent origin (Farris, 1983, p. 13). Thus, it is only in this latter sense that  $c_s$  constitutes a decisive test, assuming only  $b$ —that an incongruent  $c_s$  can be said to count against a hypothesis of sister group relationships.

Some think that falsification in Popper’s logic of scientific discovery is an “asymmetrical, all-or-nothing affair” (e.g., Rieppel, 2003, pp. 262, 265, 269, 270). However, deductive arguments can never show that their conclusion of falsity is true. The decisiveness of the logical truth is irrelevant, for there is after all the conjunction of auxiliary assumptions, such as  $b$ . The most deductive arguments can offer scientists is rational, objective choice among empirical hypotheses. That choice is provided when falsification is applied simultaneously to all of the members of a set of  $h_n$ . This choice is not a function of having to assume “descent, with modification,” but is the logical consequence of any  $h$  being part of a closed hypothesis set, where falsification applies symmetrically. Those who define Popperian philosophy in terms of syntactical certainty fail to grasp its real goals, that of objectivity and rationality.

7. It also follows from the improbability argument above (see argument set 5), where only incongruent  $e$

<sup>11</sup> According to observational theory, pure observation reports are used to erect an inductivist epistemology. In the absence of such reports (as revealed, for example, by incongruent characters, regardless of a priori tests), the emphasis has shifted to the importance of perceptual theory (Ryle, 1949), which is readily adapted to Popper’s deductive epistemology (Popper, 1963, pp. 201–214); see Footnote 14. In the present case, synapomorphy,  $c_s$ , is epistemologically accessible.

<sup>12</sup> Character analysis/reanalysis in phylogenetic systematics does not include a test of similarity (*sensu* Patterson, 1982), because it involves induction, which can lead only to an abstract (class) concept of taxa. Denying the relevance of observational theory and similarity, phylogenetic systematics is left to focus on what may be perceptually the same thing (Kluge, 2003; see Footnote 12).

<sup>13</sup> When the consequent of an implication is denied in deductive logic it is *modus tollens*, e.g., if  $h$ , then not  $e/e$  is false/therefore,  $h$  is false (Rieppel, 2003). In the present case, if  $h_{(A,B)C}$  then not  $e_{1,0,1}$  or not  $e_{0,1,1}/e_{1,1,0}$  is false/therefore,  $h_{(A,B)C}$  is false; if  $h_{(A,C)B}$  then not  $e_{1,0,1}$  or not  $e_{0,1,1}/e_{1,1,0}$  is false/therefore,  $h_{(A,C)B}$  is false; if  $h_{(B,C)A}$  then not  $e_{1,0,1}$  or not  $e_{0,1,1}/e_{1,1,0}$  is false/therefore,  $h_{(B,C)A}$  is false.

counts as a falsifier, that severity of test increases with the number of those tests that have been carried out. Thus, the more tests the better (Popper, 1957, pp. 106–107; 1983, p. 238). In phylogenetic systematics, severity of test has been identified with the simultaneous analysis of  $c_s$ , and which is maximized in the results obtained from an unpartitioned data matrix (i.e., the total evidence of Kluge, 1997). Congruence corroborates shared character states as homologues (see argument sets 8–10, below).

8. The more severe the test  $e$ , as supporting evidence of  $h$ , in light of  $b$ , the greater the *power* of  $h$  to causally explain  $e$ , in light of  $b$ .<sup>14</sup> Delimiting  $e$  operationally, as an inferred transformation from plesiomorph to apomorph, the synapomorphous states,  $c_s$ , that can be optimized as spatio-temporally restricted on  $h$  are postdicted homologues (see argument set 10, below). The cause of common ancestry is not explained with regard to its effect, homology; quite the opposite, the phenomenon of homology is explained with regard to its cause, common ancestry. This is a typical example of the *when/where explanandum* of a deductive historical scheme of causal explanation, and it is without vicious circularity, i.e., epistemological dependence (*contra* Brady, 1994, pp. 11 and 13). That which is postdicted is not a necessary consequence of that which is assumed,  $b$ , because phylogeny is not defined in terms of homology (see Hennig, 1966, Fig. 21). Neither can circularity be claimed in the actual practice of phylogenetic systematics, because not all  $c_s$  fit the definition of homology, even though all homologues are  $c_s$  (Hull, 1967, p. 177; *contra* Brady, 1994, p. 25).

9. The corroboration of  $h$  by  $e$ , in light of  $b$ , is simply the measure of the degree of support given by  $e$  to  $h$ , in light of  $b$  (see argument set 5, above).<sup>15</sup> The alternative statements of  $h_n$  can be compared objectively for the degree to which they have been falsified by the critical evidence, and therefore the difference between optimal and suboptimal  $h_n$  is an objective measure of the magnitude of evidential support (Popper, 1959, 1983; Kluge, 2002, p. 591; Grant and Kluge, unpublished). A meaningful metric of this kind of support depends on the assumption that all of the  $c_s$  used in an analysis of  $h_n$  have achieved the same level of testability (Kluge, 1997,

p. 93). In any case, falsification in phylogenetics cannot be judged naive, because there is sophistication in the degree to which the competing  $h_n$  are evaluated in an objective manner.

10. Explanatory power<sup>14</sup> and degree of corroboration<sup>15</sup> are maximized by minimizing  $e$  on  $h$ , in light of  $b$ , according to the “rule of parsimony” (Popper, 1959, p. 145). According to the logic of scientific discovery, *enjoining parsimony protects the falsifiability of the phylogenetic system from going to zero* (see argument set 5, above).<sup>10,16</sup> In the practice of phylogenetic systematics, the hypothesis that has the greatest explanatory power and is the most corroborated, in light of  $b$ , is the  $h$  that requires the fewest transformations counted on the path between ancestor and descendant, i.e., the sum of the patristic differences (Farris, 1970).

The fact that character states must be optimized at each node to identify the most parsimonious  $h_n$  means, coincidentally, that nothing more than  $b$  need be assumed (such as the counterfactual conditionals of a stochastic model of evolution) to provide a causal explanation of  $c_s$  as homology (*contra* Sober, 1988, pp. 132–135). Also to be borne in mind, a  $c_s$  that counts *against* a particular  $h$ , such as (A,B), must count equally *for* one of the alternatives, such as (A,C) or (B,C). This is another product of pure logic.

The deductive argument for parsimony summarized above is not to be confused with justifications of goodness of fit (including maximizing explanatory power, by minimizing *ad hoc* assumptions) or simplicity (efficiency, improbability) of evolution. For example, Hennig’s (1966, p. 121) auxiliary principle, which has long served as the basis for goodness of fit in phylogenetic systematics (see also Kluge and Farris, 1969, p. 7; Wiley, 1975, p. 236; 1981, p. 20; Eldredge and Cracraft, 1980, p. 67; Farris, 1983, p. 8; Schuh, 2000, p. 116), states that “the presence of apomorphous characters in different species is always reason for suspecting kinship [i.e., that the species belong to a monophyletic group], and that their origin by convergence should not be assumed a priori. . . [P]hylogenetic systematics would lose all the ground on which it stands if the presence of apomorphous

<sup>14</sup> This is explanatory power in Popperian testability,  $E(h, e, b) = p(e, hb) - p(e, b)/p(e, hb) + p(e, b)$ . See Footnotes 9 and 10.

<sup>15</sup> This is degree of corroboration in Popperian testability,  $C(h, e, b) = p(e, hb) - p(e, b)/p(e, hb) - p(eh, b) + p(e, b)$ . If  $e$  neither supports nor undermines  $h$ , in light of  $b$ , then  $S(e, h, b) = E(h, e, b) = C(h, e, b) = 0$ .  $S(e, h, b), E(h, e, b)$  and  $C(h, e, b)$  are negative when  $e$  undermines  $h$ , in light of  $b$ , and  $S(e, h, b), E(h, e, b)$  and  $C(h, e, b)$  are positive when  $e$  supports  $h$ , in light of  $b$ .  $S(e, h, b) = E(h, e, b) = C(h, e, b) = -1$  only when  $e$  absolutely contradicts  $h$ , in light of  $b$ , and only when  $p(e, hb) = 1, p(e, b) = 0$ , and  $p(h, b) = 0$  can  $S(e, h, b) = E(h, e, b) = C(h, e, b) = +1$ . See Footnotes 9 and 10.

<sup>16</sup> When Farris (1983, p. 8) justified the use of parsimony with regard to explanatory power, by minimizing requirements for *ad hoc* hypotheses of homoplasy (see argument set 5, above; Popper, 1963, p. 61), he credited Wiley (1975) for the Popperian testability argument. However, Wiley’s (1975, pp. 237, 243) statement, “that a hypothesis of homology based on morphological comparisons which has been rejected the least number of times relative to other possible hypotheses is to be preferred over those other hypotheses,” lacks epistemological reason, and on its face is a justification for clique analysis, not parsimony. The preferred largest clique refers only to congruent  $e$ , whereas the most parsimonious cladogram refers to all  $e$ , congruent and incongruent (Kluge, 1976; see argument set 7, above). Also, Farris’s (1983) descriptive efficiency is an operational extension of his parsimony justification; it has no epistemological standing of its own.

characters in different species were considered first of all as convergences (or parallelisms), with proof to the contrary required in each case.” Secondly, Felsenstein (1983, p. 322; see also Humphries and Funk, 1984, p. 330; Richards, 2002, p. 13) underscores the connection between parsimony and simplicity of the evolutionary process when he states that “[p]arsimony methods try to find that tree which requires the fewest changes of character state. They originated as the ‘method of minimum evolution’ that Edwards and Cavalli-Sforza . . . applied to gene frequency data, and were applied to discrete morphological characters by Camin and Sokal . . .”

11. Increased knowledge can also result from a cycle of checking and rechecking the  $c_s$  that is incongruent with the most parsimonious, total evidence,  $h$ , in light of  $b$  (Kluge, 1998; see also Hennig, 1966). What are causally explained as historically the same things (the identity relation of homology) may be expected to increase as the systematic error of incongruent  $c_s$  is identified and removed with character state redefinition and/or character recoding.<sup>17</sup> Accordingly, each severely tested historical thing will come to be defined ostensibly with regard to a unique scientific term. There is one commensurable viewpoint for which there is a *direct* ostensibly defined scientific language. Naming need not be a social process in science (*contra* Kuhn, 1977, pp. 309–313), and therefore phylogenetic inference is not necessarily scientifically indeterminate (*sensu* Richards, 2002, p. 18). For all intents and purposes objective reality may be discovered in phylogenetic systematics.<sup>18</sup>

Given these 11 sets of arguments, what verdicts might be rendered by open-minded philosophers and scientists with regard to the Popperian nature of phylogenetic systematics? One possibility is that the case has not been made, because the purity of Popperianism cannot be executed (see epigraph). For example, Popperianism is concerned with making conditional scientific predictions, which requires scientific laws, and the most that any historical science can manage is the discovery of unique events and the retrodiction of trends. Another obvious verdict is that the combination of footnotes and text merely flesh out what Popper came to realize later in his life, that evolutionary theory could have scientific character. Relevant to the précis at hand, he said

<sup>17</sup> This heuristic use of incongruence is ampliative; it does not assume that homoplasy is something to be explained (Kluge, 1999, p. 431).

<sup>18</sup> The Duhem–Quine thesis of underdetermination, that we can never really know whether the incongruence is due to an error in  $h$ , or an error in  $e$ , or an error in  $b$ , applies equally to all possible  $h_n$  in phylogenetic systematics, because all are, at least in principle, tested simultaneously with the same  $e$ , assuming the same  $b$ . Thus, while scientific knowledge of phylogenetic systematics must be treated as tentative, our knowledge of species history is expected to grow with cycles of research, given the finite number of  $h$  being tested simultaneously (Kluge, 1998).

(Popper, 1980, p. 611), “It appears as if some people would think that the historical sciences are untestable because they describe unique events. However, the description of unique events can very often be tested by deriving from them testable predictions or retrodictions.” A detailed discussion of this expanded concept of what is empirical in evolutionary biology, and phylogenetic inference in particular, can be found in Stamos (1996, p. 165). Yet another verdict is that my arguments form a logical system for achieving *objective historical knowledge*, independent of the Popperian formalisms to which I appealed—an epistemology involving statements and contradiction, which is constrained by the sufficiency of “descent, with modification” as background knowledge. The salient features of this evolutionary system are the spatio-temporal restrictedness and historically contingent nature of life’s patterns and processes. This epistemology can have nothing to do with true belief, because its concern is for the testability of improbable hypotheses. Falsifiability is then a normative proposal about what we should and should not regard as scientific in the study of history. At least that much is Popperian (Notturmo, 2003).

### Acknowledgments

This paper was written at the Cladistics Institute, Harbor Springs, Michigan. Although we continue to disagree on some points, the content of several of my arguments benefited considerably from criticisms provided by Glenn Fox, Taran Grant, Maureen Kearney, and David Stamos. Taran Grant’s careful editing of the manuscript removed more than one “loose thread.” Dan Faith, acting as Associate Editor for Systematic Biology, provided the impetus for expanding the manuscript, so that it might better “stand on its own.”

### References

- Agassi, J., 1975. Science in flux. In: Cohen, R.S., Wartofsky, M.W. (Eds.), *Boston Studies in the Philosophy of Science*, vol. 28. Reidel, Boston, pp. 1–553.
- Ariew, A., 1998. Are probabilities necessary for evolutionary explanations? *Biol. Philos.* 13, 245–253.
- Boyd, R., 1990. Realism, approximate truth, and philosophical method. In: Savage, W. (Ed.), *Minnesota Studies in the Philosophy of Science XIV*. University Minnesota Press, Minneapolis, MN, pp. 355–391.
- Boyd, R., 1991. Realism, anti-foundationalism, and the enthusiasm for natural kinds. *Philos. Stud.* 61, 127–148.
- Boyd, R., 1999. Homeostasis, species, and higher taxa. In: Wilson, R.A. (Ed.), *Species: New Interdisciplinary Essays*. MIT Press, Cambridge, MA, pp. 141–185.
- Brady, R.H., 1994. Pattern description, process explanation, and the history of morphological sciences. In: Grande, L., Rieppel, O. (Eds.), *Interpreting the Hierarchy of Nature: From Systematic Patterns to Evolutionary Process Theories*. Academic Press, New York, pp. 7–31.

- Cleland, C.E., 2002. Methodological and epistemic differences between historical science and experimental science. *Philos. Sci.* 69, 474–496.
- Darwin, C., 1859. *The Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life*. John Murray, London [1964. Facsimile of the 1st edition. Harvard University Press, Cambridge, MA].
- de Queiroz, K., 1992. Phylogenetic definitions and taxonomic philosophy. *Biol. Philos.* 7, 295–313.
- de Queiroz, K., Gauthier, J.A., 1992. Phylogenetic taxonomy. *Annu. Rev. Ecol. Syst.* 23, 449–480.
- de Queiroz, K., Poe, S., 2001. Philosophy and phylogenetic inference: A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. *Syst. Biol.* 50, 305–321.
- Edwards, A.W.F., Cavalli-Sforza, L.L., 1964. Reconstruction of evolutionary trees. In: Heywood, V.H., McNeill, J. (Eds.), *Phenetic and Phylogenetic Classification*. Syst. Assoc. Publ., London, No. 6, pp. 67–76.
- Eldredge, N., Cracraft, J., 1980. *Phylogenetic Patterns and the Evolutionary Process: Method and Theory in Comparative Biology*. Columbia University Press, New York.
- Ereschetsky, M., 2001. *The Poverty of the Linnaean Hierarchy: A Philosophical Study of Biological Taxonomy*. Cambridge University Press, New York.
- Faith, D.P., Trueman, J.W.H., 2001. Towards an inclusive philosophy for phylogenetic inference. *Syst. Biol.* 50, 331–350.
- Farris, J.S., 1970. Methods for computing Wagner Trees. *Syst. Zool.* 19, 83–92.
- Farris, J.S., 1983. The logical basis of phylogenetic analysis. In: Platnick, N.I., Funk, V.A. (Eds.), *Advances in Cladistics II*. Columbia University Press, New York, pp. 7–36.
- Felsenstein, J., 1983. Methods for inferring phylogenies: A statistical view. In: Felsenstein, J. (Ed.), *Numerical Taxonomy*. Springer-Verlag, New York, pp. 315–334.
- Grant, T., Kluge, A.G. ms. Data Exploration: Scientific, heuristic, or neither. *Cladistics* [in review].
- Grünbaum, A., 1989. The pseudo-problem of creation in physical cosmology. *Philos. Sci.* 56, 373–394.
- Hennig, W., 1950. *Grundzüge Einer Theorie der Phylogenetischen Systematik*. Deutscher Zentralverlag, Berlin.
- Hennig, W., 1966. *Phylogenetic Systematics*. University Illinois Press, Urbana, IL.
- Hull, D.L., 1967. Certainty and circularity in evolutionary taxonomy. *Evolution* 21, 174–189.
- Hull, D.L., 1983. Karl Popper and Plato's metaphor. In: Platnick, N.I., Funk, V.A. (Eds.), *Advances in Cladistics II*. Columbia University Press, New York, pp. 177–189.
- Hull, D.L., 1999. The use and abuse of Karl Popper. *Biol. Philos.* 14, 481–504.
- Humphries, C.J., Funk, V.A., 1984. Cladistic methodology. In: Heywood, V.H., Moore, D.M. (Eds.), *Current concepts in plant taxonomy*. Systematics Association Special, vol. 25. Acad. Press, London.
- Kitts, D.B., 1977. Karl Popper, verifiability, systematic zoology. *Syst. Zool.* 26, 185–194.
- Kluge, A.G., 1976. Phylogenetic relationships in the lizard family Pygopodidae: An evaluation of theory, methods and data. *Misc. Publ. Mus. Zool., University Mich.* (152), 1–72.
- Kluge, A.G., 1997. Testability and the refutation and corroboration of cladistic hypotheses. *Cladistics* 13, 81–96.
- Kluge, A.G., 1998. Sophisticated falsification and research cycles: Consequences for differential character weighting in Phylogenetic Systematics. *Zool. Scripta* 26, 349–360.
- Kluge, A.G., 1999. The science of phylogenetic systematics: Explanation, prediction, and test. *Cladistics* 15, 429–436.
- Kluge, A.G., 2001. Philosophical conjectures and their refutation. *Syst. Biol.* 50, 322–330.
- Kluge, A.G., 2002. Distinguishing “or” from “and”, and the case for historical identification. *Cladistics* 18, 585–593.
- Kluge, A.G., 2003. The repugnant and the mature: Atemporal similarity and historical identity in phylogenetic inference. *Cladistics* [in revision].
- Kluge, A.G., Farris, J.S., 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18, 1–32.
- Kuhn, T.S., 1962. *The Structure of Scientific Revolutions*. University Chicago Press, IL [1970, 2nd edition].
- Kuhn, T.S., 1977. *The Essential Tension*. University Chicago Press, Chicago, IL.
- Lidén, M., 1990. Replicators, hierarchy, and the species problem. *Cladistics* 6, 183–186.
- Mahner, M., Bunge, M., 1997. *Foundations of Biophilosophy*. Springer, New York.
- Mill, J.S., 1843. *A System of Logic*. Longmans, Green, London.
- Notturmo, M.A., 2003. *On Popper*. Nelson Thomson Learning, Toronto.
- Olmstead, R., 2001. Phylogenetic inference and the writings of Karl Popper. *Syst. Biol.* 50, 304.
- Patterson, C., 1978. Verifiability in systematics. *Syst. Zool.* 27, 218–222.
- Patterson, C., 1982. Morphological characters and homology. In: Joysey, K.A., Friday, A.E. (Eds.), *Problems of Phylogenetic Reconstruction*. Acad. Press, New York, pp. 21–74.
- Popper, K., 1957. *The Poverty of Historicism*. Routledge and Kegan Paul, London.
- Popper, K., 1959. *The Logic of Scientific Discovery*. Harper & Row, New York [1968 edition].
- Popper, K., 1963. *Conjectures and Refutations: The Growth of Scientific Knowledge*. Harper and Row, New York [1968 edition].
- Popper, K., 1972. *Objective Knowledge: An Evolutionary Approach*. Clarendon Press, Oxford.
- Popper, K., 1980. *Evolution*. New Sci. 87, 611.
- Popper, K., 1982. *Unended Quest: An Intellectual Autobiography*. William Collins, Glasgow.
- Popper, K., 1983. *Realism and the Aim of Science*. Routledge, London [1992 reprint].
- Richards, R., 2002. Kuhnian values and cladistic parsimony. *Perspect. Sci.* 10, 1–27.
- Rieppel, O., 2002. Beyond the ‘logic of scientific discovery’: A theory of empirical classes. In: *Biodiversity in the information age.* [abstracts] Sixth International Congress of Systematic and Evolutionary Biology. Dept. Biol., University Patras, Patras, Greece p. 30.
- Rieppel, O., 2003. Popper and systematics. *Syst. Biol.* 52, 259–271.
- Ryle, G., 1949. *Concept of Mind*. Hutchinson's Library, New York.
- Schuh, R.T., 2000. *Biological Systematics*. Cornell University Press, Ithaca, New York.
- Siddall, M.E., Kluge, A.G., 1997. Probabilism and phylogenetic inference. *Cladistics* 13, 313–336.
- Sober, E., 1988. *Reconstructing the Past: Parsimony, Evolution, and Inference*. MIT Press, Cambridge, MA.
- Stamos, D.N., 1996. Popper, falsifiability, and evolutionary biology. *Biol. Philos.* 11, 161–191.
- Watkins, J., 1984. *Science and Scepticism*. Princeton University Press, Princeton, NJ.
- Wiley, E.O., 1975. Karl R. Popper, systematics, and classification: A reply to Walter Bock and other evolutionary taxonomists. *Syst. Zool.* 24, 233–242.
- Wiley, E.O., 1981. *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*. Wiley–Interscience, New York.