

Evidence, Content and Corroboration and the Tree of Life

E. Kurt Lienau · Rob DeSalle

Received: 28 August 2008 / Accepted: 18 September 2008 / Published online: 18 November 2008
© The Author(s) 2008. This article is published with open access at Springerlink.com

Abstract We examine three critical aspects of Popper’s formulation of the ‘*Logic of Scientific Discovery*’—evidence, content and degree of corroboration—and place these concepts in the context of the Tree of Life (ToL) problem with particular reference to molecular systematics. Content, in the sense discussed by Popper, refers to the breadth and scope of existence that a hypothesis purports to explain. Content, in conjunction with the amount of available and relevant evidence, determines the testability, or potential degree of corroboration, of a statement; content distinguishes scientific hypotheses from metaphysical assertions. Degree of corroboration refers to the relative and tentative confidence assigned to one hypothesis over another, based upon the performance of each under critical tests. Here we suggest that systematists attempt to maximize content and evidence to increase the potential degree of corroboration in all phylogenetic endeavors. Discussion of this “total evidence” approach leads to several interesting conclusions about generating ToL hypotheses.

Keywords Corroboration · Tree of life · ToL · Karl Popper · Genomics · Systematics · HGT · Evolution · Vantage points · Darwin · Philosophy of science

E. Kurt Lienau (✉)
Department of Biology, New York University, New York, NY 10003, USA
e-mail: ekl@amnh.org

E. Kurt Lienau · R. DeSalle
Sackler Institute for Comparative Genomics, American Museum of Natural History,
New York, NY 10024, USA

1 Introduction

Because of genome sequencing initiatives, the breadth of organisms sampled and the completeness with which they are described offers evolutionary biologists new opportunity for great advances in understanding life on Earth. For the first time, systematists are beginning to make and test potentially credible ToL hypotheses. The biggest challenge to achieving the goal of a well corroborated theory of the evolutionary history of all of life on our planet (which is by and large prokaryotic) is eliminating the potentially confounding effects homoplasy, due in large part to horizontal gene transfer (HGT), may have on recovering vertical, or organismal, history in phylogenetic analysis. A telling example of the importance of the philosophical ramifications of this endeavor concerns deciding whether or not a bifurcating tree of life (ToL) exists for *all* life on Earth. Some scientists feel that the problem of eliminating the effects of homoplasy is insurmountable because one can never know if one finds the “true” or “real” tree or instead recovers a tree-shaped network of connections that more closely reflects horizontal signal than the actual history of life (Bapteste et al. 2008; Doolittle and Bapteste 2007). In other words, these authors claim that a bifurcating tree of life is impossible to attain, at least for the bacterial and archaeal branches. They also suggest that the current state of microbial evolutionary study rejects the very hypothesis that there is a bifurcating tree of life for prokaryotic organisms.

We disagree with both of these conclusions.

First, we suggest that the “ToL analysis is impossible” conclusion is problematic because it is a specific expression of a general question that applies to all of science, that being, “Do our hypotheses faithfully describe reality?”. The answer is, of course, unknowable. The correspondence of human knowledge to a universal reality is ultimately uncertain and, as pointed out by Popper, there are two distinct ways of approaching the problem. One approach is to take a conventionalist position, which is to say that, because science is inseparable from (presumably fallacious) human thought, there is no reasonable expectation of any correspondence of scientific theory to reality because we cannot control for the potential confounding affects of our perceptions. The other position is that, through application of critical tests to our hypotheses and by continuously changing those hypotheses in response to falsification (as in Popper’s *The Logic of Scientific Discovery*), we come closer to describing reality by determining what is universally true through consistency with critical test. There is no objective way to choose between these two positions, but as scientists, it is logical to think that we approach some approximation of how things actually are.

With this in mind, we suggest, contra Doolittle and Bapteste (2007), that there is ample background evidence to support the existence of a distinct bifurcating history of life on Earth, even for prokaryotes, because the majority of genetic material is passed from one cell to another vertically (e.g., Beiko and Hamilton 2006; Beiko and Ragan 2008). We suggest, in turn, that any claim that the ToL hypothesis has been refuted must offer a hypothesis that is at least equally or better corroborated. In the absence of such an alternate, critically tested hypothesis that explains the graded

similarity of all life, including prokaryotic organisms, we choose to retain Darwin's "tree of naturally selected forms".¹

Our question is: "How do we best interpret genomic character data in a way that informs us of the evolutionary history of *both* the organisms that make up life (the ToL) *and* the potentially horizontally transferred genes that make up the genomes of those organisms?"

2 Reinterpreting the LSD for the Genomic ToL

Faith (2006) criticizes Helfenbein and DeSalle (2005) for writing: "It might be timely to reverse roles and allow science to play the part of the horse pulling the cart of philosophy to finish the journey begun decades ago by the dedicated workers interested in making logical sense of their science" (p. 279). We agree with this criticism; the philosophy of science, of course, logically precedes the gathering and testing of data. A better analogy to how we should approach the ToL problem, perhaps, is that we should fit our scientific cart with a new hitch to the horse of philosophy. In other words, Popperian philosophy should be revisited in order that it is better applied to the novel systematic problems produced by genomic technology with special reference to the Tree of Life problem.

Faith and Trueman (2001), Faith (2004), Helfenbein and DeSalle (2005), Faith (2006) and Lienau (2007) attempted to make inroads into this challenge. The goal of the current paper is to re-examine the relationship of content, corroboration and evidence to systematic science in light of the existence of massive amounts of genomic data for large numbers of taxa. We conclude that evidence (e) in ToL systematics is the observation of similarity among organisms that can be formulated in to potentially falsifying hypotheses of relationship (homology characters, or the ideographic transformation series of Grant and Kluge (2004) and synapomorphies) to test hypotheses of relationship (h). We point out that these observations are dynamic, and that data can be arranged in myriad different ways, each arrangement corresponding to an observational vantage point that tests the content of hypotheses of relationship independently in a total evidence, super-matrix framework.

¹ We disagree with the Doolittle and Baptiste (2007) assertion that the Darwin's tree is the explanandum, or what needs to be explained; it is a hypothesis, or explanans, of why organisms look similar by degrees. To wit, Darwin writes "In considering the origin of species, it is quite conceivable that a naturalist reflecting on the mutual affinities of organic beings...might come to the conclusion that each species had not been independently created, but had descended, like varieties, from other species. Nevertheless, such a conclusion would be unsatisfactory, until it could be shown how the innumerable species of this world have been modified..." (Darwin 2004). This passage shows that Darwin (the naturalist) had come to the conclusion that the affinities, or similarities, of organisms are best explained by hypothesizing that they are the result of descent from common species (a tree). That he added natural selection to the tree of descent hypothesis and that one can also use a tree to depict the overall similarity of organisms without a hypothesis of evolutionary history (i.e., phenetics) is beside the point.

3 Content and Corroboration: Some Definitions

“I believe that these two ideas – content and degree of corroboration – are the most important logical tools in my book.” (Popper 2002, p. 411)

Content is a quality of a particular theory or hypothesis. Popper described content as follows: “...the content of a theory, which is the same as its improbability, determines its testability and its corroboration” (Popper 2002, p. 411). A higher level of content, then, corresponds to a higher level of testability or greater relative difference in statement from background knowledge (*b*). Popper was very clear about the relationship of content to testability when he stated that, “To a higher degree of universality or precision corresponds a greater (logical or) empirical content, and thus a higher degree of testability” (Popper 2002, p. 106).

Precision is the narrowness of space in the universe of possible statements that statements consistent with a hypothesis occupy. Precision is the level of description that is reached by the predictions of the hypothesis. Higher levels of precision correspond to finer and finer descriptions of reality; this increases testability by limiting the number possible events that are consistent with the theory. *Precision is exclusion.* (Lienau 2007)

Universality is the number of classes that are derivable from a hypothetical statement (i.e., must be true if the hypothesis is true). This can be perceived as the phenomena that are provided for by, and thus can be used to test, the hypothesis; universality is the scope and range of demands the hypothesis makes on reality which in turn can vary in levels of specificity or precision. *Universality is inclusion.* (Lienau 2007)

4 Degree of Corroboration (DOC)

Popper proposed the concept of *degree of corroboration* (herein referred as DOC, when convenient) as relative tentative confirmation or confidence in one hypothesis over another based upon the performance of each hypothesis under critical tests. Degree of corroboration is not equivalent to statistical or formal probability, but is based on a distinct “absolute” probability calculus Popper wrote to eliminate contradictions and logical absurdities that arise when one attempts to use a formal probability calculus (logical or statistical) to assess the support of a hypothesis (or any relatively universal statement) by evidence (Appendix ix in Popper 2002).

The idea of corroboration is central to empirical science. “Corroborate”, unlike “confirm” carries with it the acknowledgement that science involves testable conjectures that must be able to be falsified by experience and further, that this knowledge is tentative and meant to be replaced by new thoughts, or hypotheses that can be more easily tested and may stand up better to test. Thus, scientific knowledge can never be verified, only tentatively accepted. Popper saw the corroboration of a theory (or hypothesis, *h*) by evidence (*e*) as a continuum that ranged in degrees from refutation to independence, or irrelevance, to corroboration.

Helfenbein and DeSalle (2005, p. 272) interpreted Popper's definition of degree of corroboration to be "the degree to which a hypothesis stands up to severe tests, the appraisal of the worth of the hypothesis". 'Truth' and 'confirmation' have no role in the absolute calculus of corroboration. Truth, according to Popper (2002), is atemporal (Helfenbein and DeSalle 2005), because "what is true in the present was true in the past and will be true in the future" (Popper 2002). Corroboration is temporal because the outcome of tests is unknown before they are implemented: a particular test may have corroborated a hypothesis yesterday, and another more severe test corroborated it today; most importantly, though, is that another, more severe test may refute the hypothesis tomorrow. Popper rejected the word "confirmation" precisely because confirmation indicates static truth, which, according to him, does not exist in the realm of science. From the history of science, an intuitively satisfying example of this situation is the rejection during the twentieth century of Newton's hypotheses of gravity in favor of Einstein's propositions, framed within his relativity theory.

Popper developed several equations for this calculus based on the desiderata of his new 'absolute' probability system. The most novel aspect of the calculus was the ability to define zero probability for a hypothesis (h) (which occurs in the denominator). Popper's formula for DOC^2 is shown below:

$$C(e, h) = \frac{P(e, h) - P(e)}{P(e, h) - P(eh) + P(e)}$$

This equation can be read as "the corroboration of evidence (e) given hypothesis (h) is equal to the absolute probability of e given h minus the absolute probability of e alone divided by the absolute probability of e given h minus the absolute probability of h unified with e (where the two probabilities intersect) plus the absolute probability of e alone.

Popper then adds the variable b , to account for evidence present, but outside of the test in question (background knowledge):

$$C(e, h, b) = \frac{P(e, hb) - P(e, b)}{P(e, hb) - P(eh, b) + P(e, b)}$$

This complex equation has several components to it. For instance, $p(e, hb)$ can be read "the probability of the evidence given the hypothesis and the background knowledge". The numerator of the equation is then this term minus the "probability of the evidence given the background knowledge alone", giving a measure of the explanatory advantage of the hypothesis for the evidence at hand. The denominator in the equation, while complex looking and having little intuitive significance, is deemed to be "the simplest normalization factor" (Popper 1983, p. 240).

² The terms in the following equations require some basic definitions: *evidence* (e) comes in the form of basic statements, which are the statements we agree upon as observation of occurrences. These statements are, as all statements must be, conjectures based upon perception. Hypotheses (h) are falsifiable conjectures made in the form of universal statements, which range from natural laws to the lower level potentially falsifying hypotheses that we use to test the predictions of more universal hypotheses. Background knowledge (b) is the set of assumptions on which our analyses are based; we take these statements as the given truth (and so must take care in deciding just what they should be).

The three components of corroboration, then, are theory (i.e., hypothesis, h), evidence (e) and background knowledge (b). Corroboration can only be accomplished if h is falsifiable (i.e., it must contain testable content that is different from b) and gives predictions consistent with e (i.e., demands few *ad hoc* hypotheses to fit the evidence). A test must “prove its mettle” (Popper 2002) by maximizing the value in the numerator, “ $p(e, hb) - p(e, b)$ ” with respect to competing hypotheses. The numerator will be positive as the probability of e becomes greater due to h relative to b (this is corroboration). Negative values for the numerator are obtained if the probability of e in the light of h is less than that given b alone (this is falsification). No corroboration results if both probabilities are the same and the value of the numerator is zero; the result is a tautology. When a statement is corroborated, $C(h, e, b)$ can be any value from just larger than 0 to 1.0. Where, then, does corroboration become significant? Popper suggested the arbitrary value of $1/2$ as a good lower value of significance in the following quote: “the support given by $e-h$ becomes significant only when $p(e, hb) - p(e, b) > 1/2$ ” (Popper 1983, p. 240).

A multitude of papers have been written about the applicability of particular optimality criteria, in light of Popper’s logic (Carpenter et al. 1998; Farris 1983, 1995, 1999, 2000; Felsenstein 1978, 1979, 1981, 1985; Kluge 2001; Siddall 2001) and indeed Popper’s logic itself, (de Queiroz and Poe 2003; Rieppel 2003, 2005) to phylogenetic systematics. We, of course, do have our own ideas about these topics, but will not address them here. Instead, we will enter into the discussion that Faith and Trueman (2001) and Faith (2004) started, regarding the issue of corroboration in modern molecular systematics with regard to the ToL and the ‘species problem’. They conclude, “for building the ToL, we must combine various pieces of evidence from separate studies in a meaningful way to evaluate ToL hypotheses” (Faith 2004, p. 1). This statement is hard to argue with, but the critical wording in this quote is ‘meaningful way’. What is the most meaningful way to combine the evidence, in order to increase the universality (i.e., the breadth and number of taxa included in the analysis) and precision of inferences in the ToL problem?

There are currently two ways to accomplish this task, both with their own arguments as to what ‘meaningful’ means and both with major ramifications as to how the ToL is constructed. The first approach utilizes the topological information from various data sources separately. Trees that are generated from different studies are evaluated in a ‘meaningful way’ and combined into a supertree. Perhaps the least ‘meaningful’ way to produce a supertree is through consensus of all of the trees involved: doing this does not consider DOC for each of the individual trees, and completely disregards the potential for the existence of composite signal or history that is more than simply the sum of the histories in each separate analysis (hidden support; Gatesy et al. 1999). Perhaps a more meaningful way to produce a supertree (in as much as this is possible at all) is by matrix representation, using parsimony (MRP; Bininda-Emonds and Sanderson 2001) because the result of such an analysis is an approximation of a ‘total evidence’ approach. However, the fundamental problem with supertree approaches is that “it is impossible to translate accurately the relative character support from different source trees into a supertree data set” (Gatesy et al. 2002). In other

words, there is really no ‘meaningful’ way to evaluate the corroboration of ToL hypotheses generated from supertree approaches, because there is no way to ascertain what the relationships of the taxa included in one phylogeny are to taxa in another phylogeny without actually taking the trouble to analyze those relationships under a common optimality criterion. To counter this suggestion, Faith (2004) and Faith and Trueman (2001) developed a ‘combined corroboration approach’ that utilizes the corroboration equations of Popper discussed above in conjunction with Bayesian and X^2 statistics. We point out that application, like this, of Popper’s corroboration formulae to non-absolute statistical systems are problematic because of the inability of non-absolute statistical systems to incorporate zero in the denominator.

The total evidence approach (Kluge 1989), suggests all available character information needs to be incorporated into a combined matrix. This method can also be called the ‘supermatrix approach’ Its advantage is that as long as independence of characters is justified, the hypotheses generated by a supermatrix will be meaningfully evaluated, that is to say tested as critically as possible given the data at hand. Faith (2004, p. 12) has stated that attempts to construct a “supermatrix may capture character support in some form, but may not allow for integration of separate forms of corroboration assessment, related to different kinds of data, from the separate studies”.

We disagree with this statement on several grounds.

First, researchers *can* ascertain the contribution of different partitions to support in a total-evidence framework (Gatesy et al. 1999). Second, Faith’s position is tantamount to suggesting that different sources of characters should be weighted differentially, a statement with broad philosophical implications (e.g., Kluge 1997) and that cannot not be defended as empirical. Third, and most importantly, the supermatrix or Matrix of Life (MoL) can accommodate any type of data from any source and any time, provided computer power and researcher will are strong enough. The combination of all of the available evidence in totality to levy tests upon competing hypothesis is, indeed, the goal of the total evidence approach.

Further, we suggest that once the information is combined into a single supermatrix, one steps outside the realm of corroboration of partitions of data, to enter into the realm of overall corroboration, making assessment of partition corroboration relevant only to questions specific to the partition tested or to inquiries into the relationship of corroboration among partitions. Overall character corroboration is the only ‘meaningful’ way to assess the ToL; as we elaborate below, we think that combined corroboration from various vantage points can be applied to produce a meaningful way of evaluating hypotheses. We do, however, agree with Faith’s (2004, p. 1) assertion that it is important to resist “a priori justification of one form of evidence and focus ‘skeptically’ on many different kinds of evidence”. Again, the wording of the statement is critical; what does “many different kinds of evidence” mean in the context of corroboration? We therefore discuss next the role of ‘evidentiary vantage points in analysis space’ as a means to assess what is meant by “many different kinds of evidence”.

5 Evidence, Analysis Space and Vantage Points

Systematists have had different ideas about what the evidence (e) is in systematic analysis. Many have identified the synapomorphy as the evidence that tests hypotheses of relationship (e.g., Kluge 2003); others have written that the fit of the data to the tree (Faith 1992) or the character distribution (Novacek 2001) is evidence (e). More recently, Grant and Kluge (2004) have argued that evidence (e) is the ideographic character (the hypothesis of transformation of one biological attribute to another).

However, we see systematic evidence (e) as agreed upon conjectures of the perceived similarity of biological traits that are tentatively accepted as fact for the duration of an analysis. Because synapomorphies are hypotheses of homology that are tested within the context of the analysis through the process of reciprocal illumination (Hennig 1966), and indeed change throughout the examination of different tree topologies during the search, they are not accepted as the tentative ground on which our analyses stand. Rather, observations of similarity (be it of nucleotide sequence or something we can point at, like spines) are the evidence (e) that corroborate potentially falsifying hypotheses (which can be both the ideographic characters of Grant and Kluge (2004) and synapomorphies). The fit of these characters to the particular hypothesis of relationship at hand is not evidence for that hypothesis, as (Faith 1992) would argue, but is instead the extent to which the evidence at hand *corroborates* the hypothesis, a concept that is logically distinct from evidence (e). Rather, evidence (e) is observation formulated into basic statements that are meant to test, and therefore may or may not be consistent with, particular hypotheses.

Popper wrote, “[B]asic statements have the form of singular existential statements... [they] must therefore satisfy the following conditions: (1) From a universal statement, without initial conditions, no basic statements can be deduced. On the other hand, (2) a universal statement and a basic statement may contradict each other.” (Popper 2002, p. 82) and “Nor do we need say of basic statements that they are ‘true’ or ‘false’ for we may interpret their acceptance as a conventional decision, and the accepted states as a result of this decision.” (Popper 2002, p. 274). Popper used the word “occurrence” to describe a basic statement in “realistic language”; basic statements are what we agree upon as the stopping point for what we accept as evidence. They are the bricks of the building of science both in the edification of the falsifiable theories we endlessly test in search of truth, and as what we decide upon as the tentative ground of facts we use to accept or reject those theories. Popper wrote, “It should be noticed that the conjunction of two basic statements, p and r which do not contradict each other, is in turn a basic statement” (Popper 2002, p. 84).

The basic statements (evidence) of the systematist are character observations, or primary homology statements (dePinna 1991). These observations of topographic similarity provide the evidence (e) for systematic analysis. A common basic statement used in molecular systematics is the conjunction of three statements such as “Gene x is in Taxon y ” and “Gene x^1 is in taxon z ” and “Gene x in taxon y is most similar to gene x^1 in taxon Z out of all genes that we have sequences for.” The falsifying hypothesis that this statement corroborates might be “Gene x from taxon

y is most closely related by common descent to gene x^1 from taxon z”, (the ideographic character of Grant and Kluge (2004)) or “gene x/x^1 is a synapomorphy that is consistent with the hypothesis that taxon y is most closely related by common descent to taxon x out of all of the taxa we have on hand.” These are singular hypotheses of the existence of homology that can be consistent with, or contradict other hypotheses of homology made by different characters about the evolutionary history of the same taxa and, therefore also can contradict assertions of evolutionary history made by a phylogeny in the process of reciprocal illumination (Hennig 1966). Each statement of similarity, or likeness, is independent and satisfies the conditions of a basic statement in that (using Popper’s terms), (1) A primary homology statement of similarity is not derivable from the universal and related higher subclass, the phylogenetic tree or the natural law of descent with modification (2) On the other hand, a statement of similarity can corroborate a homology hypothesis that itself contradicts the more universal statement that is a phylogenetic tree. This is so because a tree hypothesis predicts the relationships that should exist in any particular subset of the relationships implied in the tree. Therefore, a tree can predict the negation of a homology hypothesis. On suboptimal trees, we accept the falsification provided by the homology hypotheses. On optimal trees, we call this contradiction homoplasy, and those basic homology statements that “cause” this contradiction offer alternate hypotheses about the evolutionary history of the gene and or taxa-range in question (Farris 1983). A finite class of the entire possible class of basic homology hypotheses will be permitted, or consistent with any particular representation of the relationships of the taxa (tree). Those data that can be arranged into a totally self-consistent arrangement may be said to be part of one complex compound basic statement the relative weight of which may be determined by the number of observations, or events it represents. In this sense, constructing the weightiest basic homology statement is the goal of the total evidence strategy of Kluge (1998) and Kluge and Wolf (1993), (Lienau 2007).

Evidence (*e*) in a Popperian framework is observation; transformation series of characters in phylogenetic systematics are hypotheses of homology. Hypotheses (*h*) are not evidence. From the grandest, most character and taxon rich tree of life hypothesis to the least inclusive potentially falsifying hypothesis of character state transformation, hypotheses are conjectures that are not agreed upon by convention, but must be corroborated, at some level, by an observation that is agreed upon by convention. In other words, hypotheses are not evidence of themselves. In phylogenetic analysis, hypotheses can only be challenged by other hypotheses; we take as given that these hypotheses are corroborated by evidence that is observed and agreed upon as a basic statement of fact *sensu* (Popper 2002). This evidentiary observation is the primary homology statement *sensu* (dePinna 1991). As Grant and Kluge (2004) put it, “[w]hat is actually observed are the parts of an individual organism, yet those parts can be individuated according to a variety of criteria, each of which may be relevant to a given problem or in a particular inferential context.” It is the individuation of the parts in a manner specific to the evolutionary context that is the basic, “primary” statement of homology. This statement is always based, at its root, on an observation of similarity. The criteria we use to accept two or more physical entities (such as genes) as the same can and *should* vary with the demands

of the particular question being posed (Lienau 2007). For instance, one would choose different similarity criteria to group genes together as gene families for a tree of life study then for an analysis of a genus of bats.

Further, evidence from the same data source can be examined from different vantage points, thus increasing the precision of the description of taxa. This approach can be thought of as analogous to using a bispocopic camera to produce a three-dimensional image of a subject, instead of using only one aperture to produce a two dimensional image. Several systematic techniques have been designed that use this strategy; one approach was discussed in the context of nucleic acid—amino acid sequences by Agosti et al. (1996). In that study, the authors concluded that combination of information from the two vantage points of the same source of data actually increased information content and hence had the ability to increase DOC. Specifically, they suggested that the use of both nucleic acid and amino acid sequences helped to establish a character weighting approach without resorting to *a priori* notions of character quality, but read (Simmons et al. 2002) for an alternative view. In a similar vein, Fitch and Smith (1983), Gatesy et al. (1993), and Wheeler et al. (1995) discussed alignment problems, pointing out to solutions when dealing with the ‘alignment space’. In the context of alignment space, different cost values for gaps and changes in alignments can produce different vantage points of test. Comparison of alignments from different cost combinations can produce columns in the alignments that are identical, while other columns may vary from alignment to alignment. The aforementioned authors designated the columns that are identical in all alignments as ‘alignment unambiguous’ and those that are different or have no match across alignments as ‘alignment ambiguous’.

Gatesy et al. (1993) and Wheeler et al. (1995) indicated two ways of dealing with the alignment ambiguous columns. The first, named ‘culling’, simply removes any alignment ambiguous column. While always yielding more consistent results, the culling procedure also removes potential tests from the analysis. They also suggested an approach called ‘elision’, a procedure that concatenates alignments from different cost regimes. They pointed out that those columns that are alignment unambiguous get higher weight than columns deemed alignment ambiguous, due to the more frequent occurrence of the former type of columns from alignment to alignment. Wheeler (1995) also introduced a ‘sensitivity analysis approach’ to the alignment space problem. This approach generates values estimated for percentage of shared and congruent groups (PSG and PCG, respectively) for a wide range of alignment cost regimes and chooses a solution based on optimized values for these metrics with a well corroborated phylogeny. Wheeler (1995) also utilized the ‘incongruence length difference’ metric (ILD, Mickevich and Farris 1981) in these sensitivity analyses, as a measure of congruence. All of these approaches to the alignment space, though, have in common the problem of arbitrariness in the gap cost regime.

6 Content, Consistency and Degree of Corroboration in ToL Analyses

In accordance with Lienau et al. (2006) and Lienau (2007), we suggest that when choosing among different vantage points of observation in phylogenetic analysis,

one should prefer those that are composed of the most severely tested, well-corroborated homology hypotheses. For example, in whole genome ToL studies, statements of gene presence or absence corroborate potentially falsifying hypotheses, or ideographic characters of transformation sensu (Grant and Kluge 2004), of gene gain (through duplication or HGT) and loss during genomic evolution that can be tested by their agreement with the overall phylogenetic patterns suggested by all gene gain and loss statements in the analysis, using measurements such as the consistency index (CI; Kluge and Farris 1969). However, it is important to note that corroboration of a hypothesis under a Popperian framework is determined not only by consistency (i.e., the extent to which the data do not disagree) but also by information content (i.e., the strength and specificity of the predictions made by the hypothesis). The extent to which phylogenetic hypothesis is refutable is determined by the boldness of the hypothesis, or resolution of the tree, and the amount of relevant information available to test it. New metrics that incorporate these two concepts, like the one developed by Lienau et al. (2006), may aid us in our ability to determine the DOC (as opposed to degree of consistency) of a particular ToL hypothesis.

If, during the process of inference of the ToL from the MoL, the criterion to which primary importance is ascribed to is DOC; its increase would be best achieved by adding characters and taxa; in general terms, those studies with more characters (e.g., whole genome studies) will have a higher DOC than those with fewer characters (e.g., single gene phylogenies) provided both contain the same number of taxa. On the other hand, only content is increased by adding taxa, because using more taxa in an analysis increases the universality, or the breadth of 'reality' that is explained by the hypothesis, but not the number of tests levied on that hypothesis. In this context, only content and therefore potential testability (as opposed to DOC) would be increased. Because potential DOC depends both on the content and on the amount of evidence used to test that content, any total evidence, whole genome (many characters, or vantage points) ToL (many taxa) study would have high potential DOC both because of the large number of characters, which would bring a multitude of critical tests to bear on competing phylogenetic hypotheses, and the large breadth of taxa, which would increase content and therefore testability.

Supertree ToL studies (i.e., those that use many sources of topological information, but lack information about character transformation series that span the entire range of taxa in question) have high content because they almost always use a wide breadth of taxonomic sampling. However, supertree approaches, despite Faith and Trueman (2001) arguments to the contrary, do not increase potential DOC, but rather a single component of potential DOC—namely, content. In order to truly increase the DOC of ToL analyses, phylogenetic systematists should use the maximum amount of character information possible as evidence to test the high content implicit in the hypotheses of species relationship that include many, many taxa.

On the basis of these arguments, we conclude that before we worry about the ToL, we must worry even more about the MoL. That is, we should take great care in just how we collect and assemble our evidence before we use it to test hypotheses of

evolutionary history. Focusing on the construction of the Tree of Life before we determine the philosophical desiderata of constructing the Matrix of Life, as in supertree analyses, is truly putting the cart (that is, the ToL) before the horse (i.e., the MoL).

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- Agosti D, Jacobs D, Desalle R (1996) On combining protein sequences and nucleic acid sequences in phylogenetic analysis: the homeobox protein case. *Cladistics* 12(1):65–82
- Bapteste E, Susko E, Leigh J, Ruiz-Trillo I, Bucknam J, Doolittle WF (2008) Alternative methods for concatenation of core genes indicate a lack of resolution in deep nodes of the prokaryotic phylogeny. *Mol Biol Evol* 25:83–91
- Beiko RG, Hamilton N (2006) Phylogenetic identification of lateral genetic transfer events. *BMC Evol Biol* 6:15
- Beiko RG, Ragan MA (2008) Detecting lateral genetic transfer: a phylogenetic approach. *Methods Mol Biol* 452:457–469
- Bininda-Emonds ORP, Sanderson MJ (2001) Assessment of the accuracy of matrix representation with parsimony analysis supertree construction. *Syst Biol* 50:565–579
- Carpenter JM, Goloboff PA, Farris JS (1998) PTP is meaningless, T-PTP is contradictory: a reply to trueman. *Cladistics* 14:105–116
- Darwin C (2004) *The origin of species by means of natural selection*. Barnes and Noble, New York (1st publication, 1859)
- depinna MC (1991) Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7:361–394
- de Queiroz K, Poe S (2003) Failed refutations: further comments on parsimony and likelihood methods and their relationship to Popper's degree of corroboration. *Syst Biol* 52:352–367
- Doolittle WF, Bapteste E (2007) Pattern pluralism and the tree of life hypothesis. *Proc Natl Acad Sci USA* 104:2043–2049
- Faith DP (1992) Systematics and conservation: on predicting the feature diversity of subsets of taxa. *Cladistics* 8:361–373
- Faith DP (2004) From species to supertrees: Popperian corroboration and some current controversies in systematics. *Aust Syst Bot* 17:1–16
- Faith DP (2006) Science and philosophy for molecular systematics: which is the cart and which is the horse? *Mol Phylogenet Evol* 38:553–557
- Faith DP, Trueman JWH (2001) Towards an inclusive philosophy for phylogenetic inference. *Syst Biol* 50:331–350
- Farris JS (1983) The logical basis of systematic analysis. *Advances in cladistics: proceedings of the second meeting of the Willi Hennig Society*, pp 7–36
- Farris JS (1995) Conjectures and refutations. *Cladistics* 11:105–118
- Farris JS (1999) Likelihood and inconsistency. *Cladistics* 15:199–204
- Farris JS (2000) Corroboration versus “strongest evidence”. *Cladistics* 16:385–393
- Felsenstein J (1978) Cases in which parsimony or compatibility methods will be positively misleading. *Syst Zool* 27:401–410
- Felsenstein J (1979) Alternative methods of phylogenetic inference and their interrelationship. *Syst Zool* 28:49–62
- Felsenstein J (1981) A likelihood approach to character weighting and what it tells us about parsimony and compatibility. *Biol J Linn Soc* 16:183–196
- Felsenstein J (1985) Phylogenies from gene frequencies a statistical problem. *Syst Zool* 34:300–311
- Fitch W, Smith T (1983) Optimal sequence alignments. *Proc Natl Acad Sci USA* 80:1382–1386
- Gatesy J, Wheeler W, DeSalle R (1993) Alignment ambiguous nucleotide sites and the exclusion of data. *Mol Phylogenet Evol* 2:152–157

- Gatesy J, O'Grady P, Baker RH (1999) Corroboration among data sets in simultaneous analysis: hidden support for phylogenetic relationships among higher level artiodactyl taxa. *Cladistics* 15:271–313
- Gatesy J, Matthee C, DeSalle R, Hayashi C (2002) Resolution of a supertree/supermatrix paradox. *Syst Biol* 51:652–664
- Grant T, Kluge AG (2004) Transformation series as an ideographic character concept. *Cladistics* 20: 23–31
- Helfenbein KG, DeSalle R (2005) Falsifications and corroborations: Karl Popper's influence on systematics. *Mol Phylogenet Evol* 35(1):271–280
- Hennig W (1966) *Phylogenetic systematics*. Translated by D. Dwight Davis and Rainer Zangerl. University of Illinois Press, Urbana
- Kluge AG (1989) A concern for evidence and a phylogenetic hypothesis for relationships among Epicrates (Boidae, Serpentes). *Syst Zool* 38:1–25
- Kluge AG (1997) Sophisticated falsification and research cycles: consequences for differential character weighting in phylogenetic systematics. *Zool Scr* 26(4):349
- Kluge AG (1998) Total evidence or taxonomic congruence: cladistics or consensus classification. *Cladistics* 14:151–158
- Kluge AG (2001) Philosophical conjectures and their refutation. *Syst Biol* 50:322–330
- Kluge AG (2003) On the deduction of species relationships: a precis. *Cladistics* 19:233–239
- Kluge AG, Farris JS (1969) Quantitative phyletics and the evolution of anurans. *Syst Zool* 18:1–32
- Kluge AG, Wolf AJ (1993) Cladistic: what's in a word? *Cladistics* 9:183–199
- Lienau EK (2007) *Horizontal gene transfer and the tree of life: an empirical approach to systematics in the genome age: biology*. New York University, New York
- Lienau EK, DeSalle R, Rosenfeld JA, Planet PJ (2006) Reciprocal illumination in the gene content tree of life. *Syst Biol* 55(3):441–453
- Mickevich MF, Farris JS (1981) The implications of congruence in Menidia. *Syst Zool* 30:351–370
- Novacek MJ (2001) Character distributions are phylogenetic evidence. *Paleobiology* 27:172–173
- Popper KR (1983) *Realism and the aim of science*. Rowman and Littlefield, Totowa
- Popper KR (2002) *The logic of scientific discovery*. Routledge Classics, London, p 513 (ISBN 0-415-27844-9. 1st publication, 1935, last modified, 1980)
- Rieppel O (2003) Popper and systematics. *Syst Biol* 52:259–271
- Rieppel O (2005) Proper names in twin worlds: monophyly, paraphyly, and the world around us. *Org Divers Evol* 5:89–100
- Siddall ME (2001) Philosophy and phylogenetic inference: a comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. *Cladistics* 17:395–399
- Simmons MP, Ochoterena H, Freudenstein JV (2002) Conflict between amino acid and nucleotide characters. *Cladistics* 18:200–206
- Wheeler WC (1995) Sequence alignment, parameter sensitivity, and the phylogenetic analysis of molecular data. *Syst Biol* 44:321–331
- Wheeler W, Gatesy J, DeSalle R (1995) Elision: a method for accommodating multiple molecular sequence alignment with alignment ambiguous sites. *Mol Phylogenet Evol* 4:1–9