Character mapping and cladogram comparison versus the requirement of total evidence: does it matter for polychaete systematics?

KIRK FITZHUGH*

Natural History Museum of Los Angeles County, 900 Exposition Boulevard, Los Angeles, California 90007, USA (kfitzhug@nhm.org)
* To whom correspondence and reprint requests should be addressed. E-mail: kfitzhug@nhm.org

Abstract

Fitzhugh, K. 2014. Character mapping and cladogram comparison versus the requirement of total evidence: does it matter for polychaete systematics? *Memoirs of Museum Victoria* 71: 67–78.

The practice of partitioning data for the inferences of phylogenetic hypotheses has become a routine practice in biological systematics. Two popular approaches: (i) mapping 'morphological' characters onto 'molecular' phylogenies, and (ii) comparing 'morphological' and 'molecular' phylogenies, are examined in light of what is known as the requirement of total evidence. Inferences of phylogenetic hypotheses, indeed all taxa, occur by a type of non-deductive reasoning known as abduction. The intent of abduction is to offer at least tentative causal accounts that explain character data. The rational acceptance of abductively derived hypotheses is subject to conditions of the requirement of total evidence as a matter of the evidential support for those hypotheses. It is shown that both character mapping and comparisons of cladograms using partitioned datasets are procedures that severely reduce the credibility of phylogenetic hypotheses. This problem is alleviated by acknowledging the formal structure of the why-questions we ask in relation to character data, for which phylogenetic hypotheses serve as answers.

Keywords

abductive inference, biological systematics, cladograms, phylogenetic hypotheses, requirement of total evidence

"The requirement of total evidence is not itself controversial." (Kelly, 2008: 64)

Introduction

Biological systematics has entered a state of complacency, where research agendas tend to follow prescribed methodological rules that satisfy requirements for using particular software packages or programs that lead to phylogenetic (or otherwise) hypotheses, or claim to provide empirical assessments of those hypotheses. This state of affairs might be expected if we adhere to Kuhn's (1970) notion of 'normal science' (but see Popper, 1970). Regardless of the consensus that might obtain in a field of science, this does not afford the accepted protocols and methods immunity from critique. There is, for instance, the expectation that scientific inquiry operates within the constraints set by the basic principles of rational reasoning (Williamson, 2000; Thagard, 2004), where the acceptance of propositions is governed by evidential support. That support comes either in the form of evidence leading to inferences of hypotheses and theories or the subsequent evidence supplied during empirical testing. If approaches to inquiry agreed upon among a group of scientists are identified as leading to less than rational conclusions due to the exclusion of evidence, either during the formulation or testing of hypotheses/theories, then the intended goal(s) of such inquiries and associated methods must

be judged relative to the criteria that determine the credibility of those hypotheses/theories. Whitehead's (1925: 18) admonition remains relevant: "The progress of biology and psychology has probably been checked by the uncritical assumption of half-truths. If science is not to degenerate into a medley of *ad hoc* hypotheses, it must become philosophical and must enter into a thorough criticism of its own foundations."

Among presentations at the 11th International Polychaete Conference that addressed phylogenetic relationships, the most common approach was that of 'character mapping'. Phylogenetic hypotheses, implied by cladograms, are inferred for sets of sequence data, and via those diagrams various conclusions are drawn regarding the evolution of morphological traits (cf. Halanych et al., 2001; Bleidorn et al., 2003; Hall et al., 2004; Halanych, 2005; Rousset et al., 2006; Schulze, 2006; Struck et al., 2007; Colgan et al., 2008; Kupriyanova and Rouse, 2008; Wiklund et al., 2008; Vrijenhoek et al., 2009; Zanol et al., 2010; Struck et al., 2011; Magesh et al., 2012; Goto et al., 2013). Interestingly, the inverse—obtaining transformation series via the mapping of nucleotides on cladograms inferred from 'morphological' characters—is never considered. An equally widespread approach involves comparisons of cladogram topologies inferred from different datasets for the same group of organisms (cf. Rousset et al., 2003, 2004; Eeckhaut and

Lanterbecq, 2005; Halanych, 2005; McHugh, 2005; Kupriyanova et al., 2006; Sperling et al., 2009; Zrzavý et al., 2009; Parry et al., 2014). The popularity of character mapping and cladogram comparisons is by no means limited to polychaetes, as perusals of such journals as *Molecular Phylogenetics and Evolution*, *Nature*, and *Systematic Biology* will attest. Regardless of their popularity, the problems surrounding these techniques are so significant as to preclude their use. This paper will identify the epistemic difficulties in light of the necessary principle of rationality known as 'the requirement of total evidence'.

Why systematics?

Determining that protocols such as cladogram comparisons and character mapping are problematic requires that we first acknowledge the intent of reasoning in biological systematics. The overarching goal of scientific inquiry is to acquire causal understanding of the phenomena we observe/describe, which also affords opportunities for predictions into the future (Hempel, 1965; Hanson, 1958; Rescher, 1970; Popper, 1983, 1992; Salmon, 1984a; Van Fraassen, 1990; Strahler, 1992; Mahner and Bunge, 1997; Hausman, 1998; Thagard, 2004; Nola and Sankey, 2007; de Regt et al., 2009; Hoyningen-Huene, 2013). As a field of science, we should expect the objective of systematics to be consistent with that of other fields. The consequence is that the aim of systematics is to causally account for the differentially shared characters we observe among organisms, whether extant or represented as fossils (Fitzhugh, 2012, 2013, and references therein). Consider the actions of compiling observation statements into a data matrix and 'inferring cladograms.' The implied intent would have to be that of explaining, by way of past evolutionary events, differentially shared characters. The primacy of explanation in systematics is, however, rarely cogently articulated and has led to a tendency to only focus on the diagrammatic qualities of cladograms, 'phylogenies' or 'trees', with inordinate attention on 'groups' and topologies, rather than recognising that cladograms are composite hypotheses representing at least three classes of causal events: (i) character origin/fixation among individuals of reproductively isolated ancestral populations and (ii) subsequent population-splitting events (Fitzhugh, 2012: Figs 1, 4; 2013: Fig. 1), as well as (iii) species hypotheses, which are inferred prior to cladograms-as-hypotheses, denoting more proximate accounts of character origin/fixation among individuals of reproductively isolated populations observed in the present. Causal events (i)–(iii) are implied by the 'interior branches', 'nodes' and 'terminal branches', respectively, that make up cladograms. Needless to say, cladograms typically convey nothing in the way of specifics regarding the causal events they are intended to imply. There are additional classes of hypotheses utilised in systematics (cf. Hennig, 1966: Fig. 6; Fitzhugh, 2012: Table 1; 2013: Table 1), but the emphasis in this paper will be on those that are phylogenetic. Presenting a diagram as a 'phylogeny' minimally assumes that it causally accounts for specifiable characters that were the basis for the inference, e.g. a data matrix. To assert that cladograms do not have to meet such an obligation would reduce them to nothing more than rhetorical devices with little or no scientific utility.

Phylogenetic reasoning

Acknowledging cladograms, trees, phylogenies, etc., as sets of explanatory accounts providing at least initial causal understanding of select characters of organisms necessitates that we identify the particular type of reasoning employed to move from observation statements, as data matrices *partim*, to cladograms. Inferring tentative causes from observed effects is known as abductive reasoning, or abduction (Peirce, 1878, 1931, 1932, 1933a, 1933b, 1934, 1935, 1958a, 1958b; Hanson, 1958; Achinstein, 1970; Fann, 1970; Reilly, 1970; Curd, 1980; Nickles, 1980; Thagard, 1988; Josephson and Josephson, 1994; Baker, 1996; Hacking, 2001; Magnani, 2001; Psillos, 2002, 2007, 2011; Godfrey-Smith, 2003; Norton, 2003; Walton, 2004; Aliseda, 2006; Fitzhugh, 2005a, 2005b, 2006a, 2006b, 2008a; 2008b; 2008c, 2009, 2010a; Schurz, 2008). Abduction has the form:

[1]   • auxiliary theory(ies)/hypotheses, *b*
       • theory(ies) relevant to observed effects, *t* (e.g. 'common ancestry')
       • observed effects, e_1 (e.g. shared characters)
       ─────────────────────────────
       • explanatory hypothesis(es), *h* (e.g. cladograms).

Abduction is non-deductive, as indicated by the double line separating premises (upper) from conclusion(s) (lower); deductive arguments are denoted by a single line separating premises and conclusion. Operationally, while abduction supplies hypotheses that at least initially account for observed effects, potential test evidence required to empirically evaluate the causal claims in hypotheses are predicted deductively:

[2]   • auxiliary theory(ies)/hypotheses, *b*
       • theory(ies) relevant to the observed effects, *t*
       • specific causal conditions presented in explanatory hypothesis via **[1]**
       • proposed conditions needed to perform test
       ─────────────────────────────
       • observed effects, e_1, originally prompting *h* (cf. **[1]**)
       • 'predicted test evidence', i.e. effects related as closely as possible to the specific causal conditions of the hypothesis.

Induction *sensu stricto* is the subsequent act of testing hypotheses:

[3]   • auxiliary theory(ies)/hypotheses, *b*
       • theory(ies) relevant to observed effects, *t*
       • test conditions performed
       • confirming/disconfirming evidence, e_2 (observations of 'predicted test evidence' in **[2]**, or alternative observations)
       ─────────────────────────────
       • *h* is confirmed/disconfirmed.

Note that the premises in **[3]** comprise the 'test evidence'. But of this evidence, it is the observations that ensue from the act of testing (third and fourth premises), either in the form of 'predicted test evidence' inferred in **[2]** or alternative results, that stand as 'test evidence' that confirms or disconfirms, respectively, the hypothesis.

While there is the assumption that the premises used in inferences of any kind are true, only deduction can provide a

conclusion that is guaranteed true if the premises are true. In other words, the rules for valid deduction limit the conclusion to being a restatement of what is in the premises (Salmon, 1984b; Copi and Cohen, 1998). The conclusions derived from abductions and inductions are probabilistic rather than certain since the content of conclusions can extend beyond that of the premises.

From a Bayesian perspective, abduction provides the basis for the prior probability of a hypothesis, $P(h \mid e_1, b)$, and induction the posterior probability, $P(h \mid e_2, b)$. Note that 'evidence' in both **[1]** and **[3]** consists of the respective premises (Longino, 1979; Salmon, 1984b; Achinstein, 2001; Fitzhugh, 2012).[1] It is worth mentioning that while we speak of evidence as the premises in any form of inference that leads to conclusions, 'evidence' in the form of character data allowing for abductive inferences of cladograms is in sharp contrast to the 'test evidence' required to empirically evaluate those hypotheses (cf. Fitzhugh, 2006a, 2010a, 2012).

Regarding systematics, the inferences of phylogenetic hypotheses, indeed all taxa, are abductive (Fitzhugh, 2006a, 2012, 2013). Following the form in **[1]**, inferences of phylogenetic hypotheses should exhibit the following schematic structure:

[4] • *Phylogenetic theory*: If character $x(0)$ exists among individuals of a reproductively isolated, gonochoristic or cross-fertilising hermaphroditic population, and character $x(1)$ originates by mechanisms $a, b, c \ldots n$, and becomes fixed within the population by mechanisms $d, e, f \ldots n$ (= ancestral species hypothesis), followed by event(s) $g, h, i \ldots n$, wherein the population is divided into two or more reproductively isolated populations, then individuals to which descendant species hypotheses refer would exhibit $x(1)$.
   • *Observations (effects)*: Individuals to which specific hypotheses *x-us* and *y-us* refer have ventrolateral margins with appendages in contrast to smooth as seen among individuals to which other species hypotheses (*a-us*, *b-us*, etc.) refer.

   • *Causal conditions (phylogenetic hypothesis* **X-us***)*: Ventrolateral margin appendages originated by some unspecified mechanism(s) within a reproductively isolated population with smooth ventrolateral margins, and the appendage condition became fixed in the population by some unspecified mechanism(s) (= ancestral species hypothesis), followed by an unspecified population-splitting event(s) that resulted in two or more reproductively isolated populations.

Note that while the formal name *X-us* would be graphically represented as a cladogram, i.e. ((*a-us*, *b-us* (*x-us*, *y-us*))), what is significant is that such a diagram implies the 'causal conditions' of character origin/fixation and population-splitting events.

The form of the 'phylogenetic theory' in **[4]** is determined by a necessary conceptual link that must exist between 'observed effects' in the form of differentially shared characters and the 'phylogenetic theory' (Fitzhugh, 2012); that link being the why-questions we implicitly or explicitly ask (Fitzhugh, 2006c, 2012) regarding those effects:

[5] 'Why do individuals to which specific hypotheses *x-us* and *y-us* refer have ventrolateral margins with appendages in contrast to smooth, as seen among individuals to which other species hypotheses (*a-us*, *b-us*, etc.) refer?'

As we are confronted with surprising or unexpected phenomena requiring explanation, in the form of differentially shared characters among organisms, what follows are the why-questions that prompt abductive inferences to phylogenetic hypotheses. The analyses by Fitzhugh (2006c, 2008b, 2012, 2013) have shown that those why-questions are located within the data matrix, where the designations of outgroups contribute to what is known as the contrastive nature of why-questions (Salmon, 1984a, 1989; Sober, 1986, 1994; Van Fraassen, 1990; Lipton, 2004; Fitzhugh, 2006a; 2006b; 2006c). This contrastive form distinguishes what is in need of explanation ('Why do individuals to which specific hypotheses *x-us* and *y-us* refer have ventrolateral margins with appendages …'), from what has been previously explained ('… in contrast to smooth, as seen among individuals to which other species hypotheses (*a-us*, *b-us*, etc.) refer?'). Why-questions seek common cause answers by way of the fact that observation statements of shared similarities carry the presupposition that those statements are true (Bromberger, 1966; Sober, 1986, 1988; Marwick, 1999; Sintonen, 2004; Schurz, 2005). Given this presupposition, explaining those similarities should involve causes that maintain as much as possible the truth of the observation statements, and that is achieved by way of a theory that ensures common causes as much as possible. Hence, the 'phylogenetic theory' in **[4]** is consistent with the presuppositions of why-questions implied in data matrices, and thus necessary. The impact of this issue on phylogenetic inference, especially regarding so-called 'likelihood' and 'Bayesian' methods, will be mentioned later (cf. 'Defeasible arguments against the requirement of total evidence').

The requirement of total evidence

It was noted in the previous section that abduction, like induction *sensu stricto*, is non-deductive, such that regardless of the truth of the premises, conclusions are only probable, as opposed to certain *qua* deduction. The consequence is that 'initial' credibility of abductive conclusions, tentative though they are, must be judged against the content of the premises. Excluding evidence that has the potential, either positively or negatively, to alter belief in, or support for a conclusion directly impinges on acceptance of that conclusion. While there are no general rules of non-deductive logic dictating the content of premises, there is the principle known as 'the requirement of

---

1 The prior probability, $P(h \mid e_1, b)$, is typically shown as $P(h)$. Since the evidence in abduction is known, i.e. $P(e_1) = 1$, then $P(h \mid e_1, b)$ = $P(h)$. As noted by Williamson (2000: 187), "… e itself should not be built into the background information, for that would give $P(e)$ the value 1, in which case $P(h \mid e)$ and $P(h)$ would be equal and e would not be evidence for anything". The negative implications for how systematists routinely refer to character data as 'supporting evidence' for cladogram topologies are significant (cf. Fitzhugh, 2012).

total evidence' that determines the degree to which rational credibility should be assigned to hypotheses (Carnap, 1950; Barker, 1957; Hempel, 1962, 1965, 1966, 2001; Salmon, 1967; 1984a, 1984b, 1989, 1998; McLaughlin, 1970; Sober, 1975; Fetzer, 1993; Fetzer and Almeder, 1993; Fitzhugh, 2006b; Kelly, 2008; Neta, 2008). Carnap (1950: 211, emphasis original) provided the first explicit description of the requirement:

"'*Requirement of total evidence*': in the application of inductive logic to a given knowledge situation, the total evidence available must be taken as basis for determining the degree of confirmation."

While the context of Carnap's characterisation is inductive, the requirement applies to all non-deductive reasoning. Failure to consider this more inclusive application has led some systematists (e.g. Wheeler, 2012: 73) to incorrectly justify the requirement via the conflation of phylogenetic inference with testing.

If the goal of scientific inquiry is the continued pursuit of causal understanding of phenomena we encounter, and evidence is that which justifies belief in the hypotheses that afford us that understanding, then deciding what evidence to consider in the derivations of beliefs will be of paramount importance. The requirement of total evidence provides the basis for choosing. Hempel (1962: 138) cogently describes the situation: "The general consideration underlying the requirement of total evidence is obviously this: If an investigator wishes to decide what credence to give to an empirical hypothesis or to what extent to rely on it in planning his actions, then rationality demands that he take into account all the relevant evidence available to him; if he were to consider only part of that evidence, he might arrive at a much more favorable, or a much less favorable, appraisal, but it would surely not be rational for him to base his decision on evidence he knew to be selectively biased."

In speaking of systematics, with the popular approaches of comparing phylogenetic hypotheses inferred from different datasets, or mapping characters on to a pre-existing set of hypotheses, i.e. cladograms, Hempel's (1966: 177, emphasis original) remarks are particularly illuminating: "When two sound inductive arguments thus conflict, which conclusion, if any, is it reasonable to accept, and perhaps act on? If the available evidence includes the premises of [two different] arguments, it is irrational to base our expectations concerning the conclusions exclusively on the premises of one or the other of the arguments; the credence given to any contemplated hypothesis should always be determined by the support it receives from the *total* evidence available at the time ... What the requirement of total evidence demands, then, is that the credence given to a hypothesis *h* in a given knowledge situation should be determined by the inductive support, or confirmation, which *h* receives from the total evidence *e* available in that situation."

In the event one is determining the plausibility of a hypothesis, whether as the product of abduction or induction, the requirement of total evidence provides a basis for assuring that plausibility is considered by way of all relevant evidence available to an investigator.[2] This is a matter of judging what premises are being used to support a particular conclusion, cf. **[1]** and **[3]**. Note that Hempel (1966) speaks of rationality when it comes to deciding theory or hypothesis acceptance. Scientific inquiry is rational to the extent we accept that theories and hypotheses are true, and that they lead to true beliefs, given available evidence. The requirement of total evidence is one of the basic tools to ensure rationality.

Since our present interest is with abduction specifically, it would be useful to look at an example of the implications of the requirement of total evidence on that type of reasoning. Consider the following abductive argument, where I attempt to explain why my lawn is wet:

[6]   • When it rains, the grass gets wet
  • My lawn was wet this morning

  • It must have rained last night.

The basis for the abduction would follow from the (contrastive) why-question (cf. **[5]**), 'Why is my lawn wet in contrast to being dry?' Questioning the initial plausibility of the hypothesis would entail determining if there are available premises that have been excluded or not considered. For instance, if we consider other premises (in italics), the plausibility of the conclusion in **[6]** drops substantially:

[7]   • *My lawn sprinklers turn on automatically at 4 am every day*
  • My lawn was wet this morning
  • *The grass is dry in adjacent yards*

  • The lawn sprinklers watered the lawn last night.

Notice that the contradictory conclusions in **[6]** and **[7]** are permissible given their respective premises. The requirement of total evidence imposes no rules on how non-deductive reasoning itself should take place, but rather provides a necessary criterion of rationality for accepting the conclusions from inferences based on available evidence, i.e. the premises. If we are aware of the additional premises in **[7]**, it would be

2 It is routine in systematics that inclusion of 'all relevant/available evidence' in abduction might not be immediately practical. For instance, it is often the case that various classes of 'morphological' characters are known across a group of organisms, but other classes of characters , e.g. cilia patterns, internal anatomy, ultrastructure, nucleotide sequences, etc., are sporadically available. Inclusion of these latter data can necessitate an abundance of 'unknown' (i.e. '?') codings, resulting in explanations (transformation series) that are largely uninformative within the scope of organisms considered. It might be more effective to delay explaining these latter observations until more inclusive coverage is attained. This is *not* to suggest that some classes of characters should be explained separately from others. The requirement of total evidence stipulates an ideal for inclusion of evidence. The goal with regard to abduction is to get as close as possible to that ideal within the limits of epistemic feasibility. Alternatively, if one wishes, for instance, to explain sequence data for a limited group of organisms for which other data are readily available, e.g. 'morphological' characters, the requirement of total evidence decisively mandates that these latter data be explained within the same abductive inference as those sequence data.

less rational to accept the conclusion in [6]. We recognise that considering these latter premises makes the initial conclusion less credible relative to the causal account that relies on the more inclusive available evidence that can affect plausibility:

[8] • There are no records of rainfall last night
• My lawn sprinklers turn on automatically at 4 am every day
• My lawn was wet this morning
• The grass is dry in adjacent yards
====================================
• The lawn sprinklers watered the lawn last night.

An analogous situation will be examined in the next section for phylogenetic inferences.

The requirement of total evidence and systematics: epistemic issues

With the basics of abductive reasoning and the requirement of total evidence presented in the previous sections, we can identify implications for two common approaches in systematics: comparing cladograms inferred from different datasets, and mapping characters on cladograms inferred from other data.

Comparing cladograms

The practice of inferring phylogenetic hypotheses from separate sets of why-questions *qua* partitioned datasets, with subsequent comparisons of topologies, also known as 'taxonomic congruence,' has a lengthy history (e.g. Mickevich, 1978). The most popular approach at present is to compare cladogram topologies inferred from 'morphological' and sequence data, respectively, or between 'morphological' and different sets of sequence data.

Using the schematic example in fig. 1A, the most basic problem with cladogram comparison can be identified. Separate abductive inferences (cf. [1], [4]) accounting for observations in datasets α and β are implied by the respective topologies, (*a-us* (*b-us* (*c-us*, *d-us*))) and ((*a-us*, *b-us*) (*c-us*, *d-us*)). The letters on each cladogram 'node' indicate hypotheses of population-splitting events necessary to explain the data in conjunction with hypotheses of character origin and fixation ['transformation series,' i.e. $n(0 \rightarrow 1)$]. Whether or not the theories used (cf. ***Phylogenetic theory*** in [4]) in the two inferences are the same will not matter at the moment. Note that the respective conclusions are contradictory in that they hypothesise the past existence of different sets of causal conditions. Strictly speaking, however, the causal events of character origin/fixation are assumed to be independent of one another. This assumption is required for the fact that we ask separate why-questions (cf. [5]) regarding different characters, and operate under the view that those observations need to be explained by separate or independent causal events of character origin and fixation among members of reproductively isolated ancestral populations (Fitzhugh, 2006a, 2006c, 2008c, 2012). But when we take population-splitting events into account, problems with cladogram comparison become apparent.

Consider population-splitting event B in fig. 1A. In conjunction with the hypotheses of character origin/fixation of characters 2(1), 3(1), and 4(1) among members of an ancestral population, splitting event B also explains the presence of those characters among individuals to which specific hypotheses *b-us*, *c-us* and *d-us* also refer. Next consider population-splitting events E and F in the other cladogram. Hypothesis E partially explains character 7(1) among individuals to which *a-us* and *b-us* refer, while hypothesis F accounts in part for character 8(1) among individuals to which *c-us* and *d-us* refer. What is immediately apparent is that hypothesis B contradicts hypotheses D, E and F, and vice versa. The plausibilities of the individual hypotheses are compromised because they account for respective observations with conflicting causal events of the same class. Hypothesis B could not be rationally accepted relative to hypotheses D–F. Contradictory sets of population-splitting events are decisive for acknowledging that the composite hypotheses represented by cladograms impinge on our ability to rationally explain all available, relevant observations. It is also the case that the separate hypotheses of character origin/fixation implied by the two cladograms call into question the credibility of those classes of hypotheses. For instance, explaining characters 2(1) through 5(1) influence rational acceptance of hypotheses for characters 7(1) and 8(1), and vice versa. The solution is to infer causal accounts for both sets of characters within the same inference (fig. 1B). Indeed, this is a constraint immediately apparent from the perspective mentioned earlier, that why-questions (cf. [5]) determine the conceptual link between observation statements and the theory that must be uniformly applied to those statements (cf. [4]).

Related to the issue of contradictory population-splitting events just described (fig. 1A), there is an additional problem that has received insufficient attention. It is not uncommon, especially with the separate inferences of phylogenetic hypotheses for 'morphological' and sequence data, that different theories are employed. As the only solution to rationally decide between contradictory hypotheses of population-splitting events is to apply the requirement of total evidence (fig. 1B), this also entails that the same theory(ies) be used for all available observations being explained. The matter of what theory(ies) to use in the inferences of phylogenetic hypotheses lies beyond the scope of this paper.[3] Regardless, there are substantial epistemic difficulties associated with most phylogenetics-related theories due to the fact that relations between observations, why-questions, and abductive inferences required to answer those questions have been largely overlooked (Fitzhugh, 2006a, 2006b, 2006c, 2008c, 2012, 2013). In lieu of combining data, the only alternative is to segregate out those why-questions that would not require phylogenetic hypotheses as answers, but rather one of the other classes of hypotheses, e.g. intraspecific or specific. Such attention to detail is, however, rarely considered.

An obvious consequence of the analysis presented thus far is that phrases of the form 'Morphological and molecular

---

3 Albeit the ***Phylogenetic Theory*** in [4] is sufficient for the why-questions asked in systematics (cf. [5]) (Fitzhugh, 2012). In terms of presenting causal events accounting for shared characters, cladograms are remarkably vague in their details.
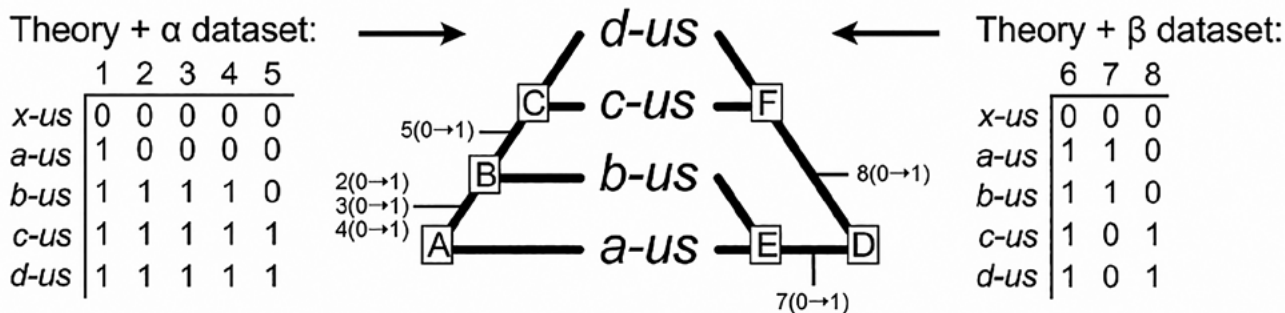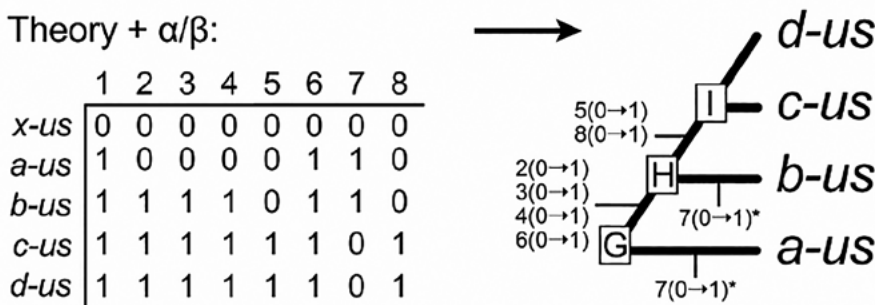
Figure 1. Example of the error of cladogram comparisons. A, phylogenetic hypotheses inferred from separate sets of premises. Letters on cladogram 'nodes' indicate population-splitting events relevant to the various hypotheses of character origin/fixation within ancestral populations. The requirement of total evidence precludes such a comparison of cladogram topologies because explanations of characters 1(1)–5(1) by population-splitting events A–C (left cladogram) contradict explanations of 6(1)–8(1) by population-splitting events D–F. See text for further discussion. B, explaining observations in accordance with the requirement of total evidence, correcting the problem in 'A'.

phylogenies for group *X* disagree (or agree)' are epistemically meaningless. There can be no disagreement/agreement due to the fact that the objective of phylogenetic inference is not to obtain 'trees'. Cladograms, as branching structures, are only as scientifically informative as the hypotheses of past causal events that can be discerned from such diagrams, as answers to why-questions. To speak of 'disagreement' among 'phylogenies' or cladograms as branching structures is to commit the fallacy of reification; treating cladograms as the tangible objects of interest rather than the actual hypotheses implied by those diagrams. The only disagreements that can be referred to among cladograms inferred from different sets of data are hypotheses of character origin/fixation within ancestral populations and subsequent population-splitting events (cf. fig. 1A); both being the result of failing to follow the requirement of total evidence (*pace* fig. 1B).

Character mapping

The popular alternative to separate inferences of phylogenetic hypotheses for partitioned data is the use of cladogram topologies based on one set of data as the 'framework' for

determining phylogenetic hypotheses for other data not involved in the inference of a cladogram (i.e. not present in the premises; cf. [1], e_1). As with cladogram comparisons discussed earlier, the issue here will be to show that decisions regarding the plausibility of phylogenetic hypotheses are compromised because mapping involves inferential processes separate from inferences of the cladograms-as-phylogenetic hypotheses upon which characters are mapped.

Fig. 2A presents an abductive inference for a set of observed effects—dataset α—where the cladogram implies at a minimum the two classes of causal events of character origin/fixation and subsequent population splitting. Also represented are the separately inferred species hypotheses, *a-us* through *d-us*. Using this cladogram topology, additional observations—dataset β—are then 'mapped' on to 'branches' of the cladogram (fig. 2B), generally in a presumptive effort to 'optimise' placements of characters to minimise *ad hoc* hypotheses of homoplasy.

Character mapping fails as a scientifically viable approach because it is in essence a variant of cladogram comparison. As discussed in the previous section, the phylogenetic hypotheses
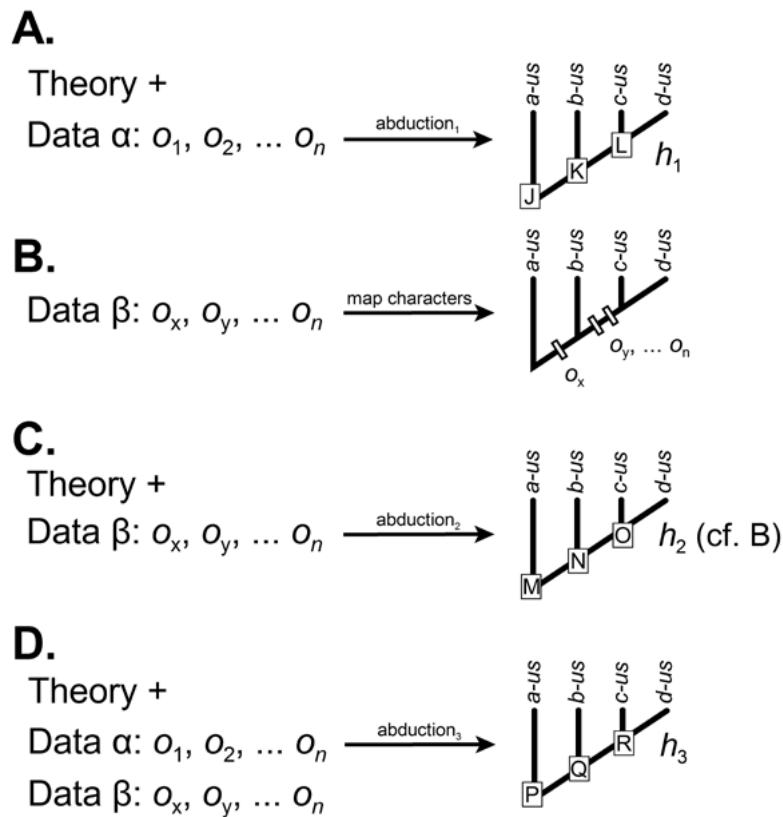
Figure 2. Example of the error of character mapping. A, phylogenetic hypotheses are inferred for a set of characters. Numbers on cladogram 'nodes' indicate population-splitting events relevant to the various hypotheses of character origin/fixation within ancestral populations (not shown; cf. fig. 1). B, a different set of characters are 'mapped' onto the branches of the cladogram in 'A'. C, the 'mapped' characters in 'B' actually refer to phylogenetic hypotheses inferred separately from the hypotheses implied by the cladogram in 'A' and 'B'. D, explaining observations in accordance with the requirement of total evidence, correcting the problem in 'B' and 'C'. See text for further discussion.

(fig. 2A) inferred using dataset α are only relevant to those characters, as explanatory accounts. While mapping (fig. 2B) gives the appearance of conjoining additional observations to these hypotheses to produce a more inclusive set of explanations, this is not the case. Regardless of what characters are mapped on to a previously inferred cladogram, the transformation series for the mapped characters do in fact represent consequences of inferential acts, albeit quite vague, that are wholly separate from the initial inference (fig. 2C). As composite hypotheses, cladograms h_1 and h_2 in fig. 2A and 2B/C, respectively, refer to different sets of explanatory accounts. The fact that the cladograms have the same topologies has no epistemic standing. Topologies of branching diagrams are immaterial. What matters are the causal events conveyed by those diagrams as answers to why-questions. The population-splitting events in h_1 (fig. 2A) only pertain to explanations of α-type characters, while events in h_2 (fig. 2B/C) only relate to β-type characters, yet both sets of hypotheses refer to classes of events that directly impinge on the credibility of those hypotheses. Per the requirement of total evidence, the only solution is that both sets of characters must be explained via the same abductive inference (fig. 2D).

Defeasible arguments against the requirement of total evidence

Cladogram comparisons and character mapping have become accepted practices in biological systematics on the basis of two common arguments endorsing the partitioning of character data: (i) sets of characters are so different in quality, or subject to radically dissimilar causal processes, as to require separate treatment, and (ii) classes of data with inordinately disparate representation will result in the 'signal' or 'noise' from the larger class 'overwhelming' what can be offered by the smaller class. Most often the perceived need for partitioning falls along the arbitrary lines of 'morphology' and nucleotide or amino acid sequences. Partitioning has never been defended on the basis of presenting a valid alternative to the requirement of total evidence that indicates the requirement is defective and at the same time establishes that partitioning promotes a more rational evaluation of hypothesis credibility in relation to abductive reasoning (cf. Fitzhugh, 2006b, 2008c). In this section, arguments (1) and (2) are shown to be invalid.

'Characters cannot be combined'

Claiming that a particular class of data, e.g. nucleotide sequences, is fundamentally different from another class, e.g. 'morphology,' such that phylogenetic hypotheses explaining the former must be inferred separately from phylogenetic hypotheses explaining the latter suffers from several basic oversights. Recall that aligning systematics with all fields of science requires acknowledging that the objective is to acquire causal understanding of differentially shared characters among organisms. This goal, via why-questions (cf. [5]) leading to abductive inferences (cf. [4]), provides the conceptual link between our observation statements of the properties of organisms and the explanatory hypotheses referred to as taxa (Fitzhugh, 2005b, 2008b, 2009, 2010b, 2012, 2013; Nogueira et al., 2010, 2013). There are two aspects of this conceptual link that have been almost uniformly overlooked in systematics, especially with regard to developments of algorithms for phylogenetic inference: the why-questions related to our observations (cf. [5]) and the nature of abductive reasoning required to provide at least initial answers to those questions (cf. [1], [4]). Indeed, while principles of phylogenetic inference have developed around notions like parsimony, 'likelihood,' and 'Bayesianism,'[4] the latter two have no relevance to abduction, and parsimony is only worthy of consideration in the context of the why-questions to which abduction is directed (Sober, 1975; Fitzhugh, 2006a, 2006b, 2012). All in all, what stands as the basis for phylogenetic inference is correctly applying abduction to why-questions, not deciding whether to use [sic] parsimony, 'likelihood,' or 'Bayesianism.'

What precludes data partitioning on the basis that classes of data are either qualitatively different or the products of substantively different causal processes is that the why-questions invariably have the form shown in [5]. The very nature of observation statements of shared similarities determines that why-questions seek common cause answers (cf. 'Reasoning and the requirement of total evidence', above)—a perspective that is at odds with 'likelihood' and 'Bayesian' methods in systematics (Fitzhugh, 2006a, 2012). The standard argument for 'likelihood' and 'Bayesian' phylogenetic inferences is that they take into consideration rates of sequence evolution (Felsenstein, 2004;

Schmidt and von Haeseler, 2009; Ronquist et al., 2009). But once one invokes rates, this must place *a priori* constraints on our observation statements, rather than introducing rates within the abductive framework for explaining those observations relative to other observations by way of phylogenetic hypotheses. This is a direct consequence of basic logic and rationality: the assumption that premises are true propositions (Williamson, 2000). For observation statements of shared similarities to have the status of evidence/premises in abduction (e.g. [4]: ***Observations (effects)***), those statements must be regarded as true. The conjunction of a theory of substitution rates and shared similarities is a contradiction. Rates of sequence evolution must be considered at the point one proceeds from perceptions to observation statements. For instance, rather than accepting that individuals to which species hypotheses *x-us*, *y-us* and *z-us* refer have nucleotide A at position 234, in contrast to T, as observed among individuals to which species hypotheses *a-us*, *b-us* and *c-us* refer, a theory of substitution rates must first be used to determine which nucleotides are in fact A while others are A'. In other words, accepting a theory of substitution rates requires that one's perceptions of A first be subjected to an initial abductive inference distinguishing some A's as shared similarities that are distinct from A''s (other shared similarities). Upon making this distinction, the subsequent why-question would have the form, "Why do individuals to which species hypothesis *x-us* refers have an A at position 234, whereas individuals to which species hypotheses *y-us* and *z-us* refer have A' (in contrast to T, as observed among individuals to which species hypotheses *a-us*, *b-us* and *c-us* refer)?" The form of the why-question is a necessary consequence of applying the theory of substitution rates at the proper epistemic juncture, i.e. prior to the abductive inference of phylogenetic hypotheses, [4].[5] The subsequent abductive inference directed at all relevant shared similarities would again seek common cause answers in the form of phylogenetic hypotheses.

With the correct utilisation of why-questions that require phylogenetic hypotheses as answers, there are no differences between characters that could warrant the partitioning of data that leads to cladogram comparison or character mapping. Similarly, attempts to develop methodological criteria to determine the extent to which data should be combined, such as the incongruence length difference test (Farris et al., 1995; Barker and Lutzoni, 2002), are nullified due to the fact that they operate under the incorrect assumption that cladograms can be empirically compared for the purpose of deciding whether or not the respective explanations of partitioned data should be discarded in lieu of being explained *en masse*.

---

4 These terms are placed in quotes because their application to abductive reasoning is erroneous (Fitzhugh, 2012). The likelihood principle refers to the probability of observing *test evidence* for a particular hypothesis, $P(e \mid h)$ (Hacking, 1965; Howson and Urbach, 1993; Lipton, 2008), while Bayesianism addresses changes in belief in hypotheses, as posterior probabilities $P(h \mid e)$, subsequent to the 'introduction of test evidence' (Salmon, 1967; Howson and Urbach, 1993; Hacking, 2001). The methods known as 'maximum likelihood' and 'Bayesianism' in systematics incorrectly conflate the abductive inferences of hypotheses with the testing of those hypotheses—a long-standing view created by equating abductive evidence, i.e. the premises in [1] and [4], with test evidence (cf. [2], [3]). This mistake has been extended to include the concept of statistical consistency (Felsenstein, 1981, 2004), where preferred methods should 'converge' on true [sic] hypotheses with the addition of more and more 'test' evidence (*not* abductive evidence). As noted by Fitzhugh (2012, see also references therein), consistency is a perspective that is meaningless in the context of abduction.

---

5 I doubt any systematist would find this manoeuvre practical, much less readily operational. But the only alternative is to maintain the integrity of observation statements of shared similarities in both why-questions and abductive inferences (cf. [5], [4], respectively). As with any field of science, calling into question whether or not shared similarities should be explained by way of some hypothesis of common cause is something considered during the process of empirical hypothesis testing, not the inferences of those hypotheses. This is yet one more reason why 'likelihood' and 'Bayesian' approaches to abductive reasoning are misguided.

'One set of data will overwhelm other data'

The intuitive appeal of the idea that the large number of nucleotides or amino acids comprising sequence data can have negative effects on the 'signal' offered by 'morphological' characters derives from two misconceptions. First, it is senseless to regard characters as either 'signal' or 'noise.' To invoke this distinction introduces the incorrect presumption that one has already explained observations prior to the abductive inferences of phylogenetic hypotheses, or is relying on specious 'support' measures like the bootstrap or Bremer index (Fitzhugh, 2006a, 2012) subsequent to inferring explanations. As the intent of phylogenetic inference is to provide answers to specifiable why-questions regarding our observation statements, there are no concepts of 'signal' and 'noise' that are applicable. Second, presuming that explaining one set of characters negatively impinges on explanations of other sets of characters requires introducing some sort of extra-evidential justification for partitioning, of which there is none. Characters considered in abductive inferences to phylogenetic hypotheses are equivalent from the perspective that they require the same explanatory structure. That equivalence is determined by the fact that the why-questions being asked (cf. **[5]**), and which are implied by a data matrix (Fitzhugh, 2006c), invoke a theory of common ancestry (cf. **[4]**, *Phylogenetic theory*) applicable to all the observations. Rather than introducing *ad hoc* maneuvers to ensure obtaining unwarranted, preordained results, answers to why-questions need to be evaluated through the standard approach of seeking test evidence that either confirms hypotheses or points to alternatives.

Conclusions

Rationality is a fundamental feature of scientific inquiry, for it enables making empirical choices between competing hypotheses or theories. In the context of abductive reasoning, being the source of hypotheses throughout biological systematics, objectively determining initial degrees of belief between hypotheses is a matter of considering the content of premises (cf. **[1]**, **[4]**, **[6]**–**[8]**). The requirement of total evidence ensures that the basis for initially accepting one hypothesis over another, i.e. $P(h_1 \mid e_1, e_2, \ldots e_n) > P(h_2 \mid e_1)$, is a rational decision. That initial acceptance is not the same as acceptance subsequent to subjecting hypotheses to empirical testing (cf. **[2]**, **[3]**), in which case the requirement of total evidence would also apply when taking into account test evidence. Regardless of properly adhering to the requirement of total evidence, the hypotheses implied by cladograms are profoundly meager causal constructs, lacking in the details needed to even consider them worthy of testing (Fitzhugh, 2012). But, this inherent limitation does not justify the tradition of uncritical thinking that has developed within, and has become a mainstay of biological systematics.

The lack of proper consideration of the requirement of total evidence within systematics has probably been mainly due to outright disagreement with the principle and/or not fully understanding it, coupled with the historical failure to embrace abductive reasoning, and perhaps no awareness regarding the importance of rationality in science. Overlooking these factors figures prominently in, for instance, Felsenstein's (2004: 536) mistaken view that a 'total evidence debate' exists in systematics. What might be perceived as a debate is actually the conjunction of multiple misunderstandings of reasoning. No valid dispute exists on the subject within the scope of logic (Hempel, 1965; Kelly, 2008; Neta, 2008) that could warrant the perception that the requirement can be bypassed in systematics. Unless systematics is successful at devising its own unique protocols for ensuring rational reasoning—which has not been the case—there is no denying the import of the requirement of total evidence. It is an ironic twist that scientists are quick to criticise such pursuits as creationism/intelligent design because they fail at leading to scientifically acceptable conclusions. Given the choice between the well-tested theory of natural selection and an untested theory of a non-natural designer, reliance on the latter is acknowledged as offering less rational understanding than the former. Yet, we see cladogram comparisons and character mapping deemed acceptable, even though they too violate the same basic tenet of rationality. The success of scientific inquiry stands on consistently recognising the essential necessary elements for rational reasoning. Systematics cannot afford to depart from those standards by ignoring the requirement of total evidence.

References

Achinstein, P. 1970. Inference to scientific laws. Pp. 87–111 in: Stuewer, R.H. (ed.), *Vol. V: Historical and philosophical perspectives of science. Minnesota studies in the philosophy of science.* University of Minnesota Press: Minneapolis.

Achinstein, P. 2001. *The book of evidence.* Oxford University Press: New York. 290 pp.

Aliseda, A. 2006. *Abductive reasoning: logical investigations into discovery and explanation.* Springer: Dordrecht. 225 pp.

Baker, V.R. 1996. Hypotheses and geomorphological reasoning. Pp. 57–85 in: Rhoads, B.L. and Thorn, C.E. (eds), *The scientific nature of geomorphology. Proceedings of the 27th Binghamton Symposium in Geomorphology held 27–29 September 1996.* John Wiley & Sons: New York.

Barker, F.K., and Lutzoni, F.M. 2002. The utility of the incongruence length difference test. *Systematic Biology* 51: 625–637.

Barker, S.F. 1957. *Induction and hypothesis.* Cornell University Press: New York. 203 pp.

Bleidorn, C., Vogt, L., and Bartolomaeus, T. 2003. A contribution to sedentary polychaete phylogeny using 18S rRNA sequence data. *Journal of Zoological Systematics and Evolutionary Research* 41: 186–195.

Bromberger, S. 1966. Why-questions. Pp. 86–111 in: Colodny, R.G. (ed.), *Mind and cosmos: essays in contemporary science and philosophy.* University of Pittsburgh Series in the Philosophy of Science, 3. University of Pittsburgh Press: Pittsburgh.

Carnap, R. 1950. *Logical foundations of probability.* University of Chicago Press: Chicago. 607 pp.

Colgan, D.J., Hutchings, P.A., and Beacham, E. 2008. Multi-gene analyses of the phylogenetic relationships among the Mollusca, Annelida, and Arthropoda. *Zoological Studies* 47: 338–351.

Copi, I.M., and Cohen, C. 1998. *Introduction to logic.* Prentice-Hall: New Jersey. 714 pp.

Curd, M.V. 1980. The logic of discovery: an analysis of three approaches. Pp. 201–219 in: Nickles, T. (ed.), *Scientific discovery, logic and rationality.* D. Reidel Publishing Company: Dordrecht.

de Regt, H.W., Leonelli, S., and Eigner, K. 2009. Focusing on scientific understanding. Pp. 1–17 in: de Regt, H., Leonelli, S. and Eigner, K. (eds), *Scientific understanding: philosophical perspectives.* University of Pittsburgh Press: Pittsburgh.

Eeckhaut, I., and Lanterbecq, D. 2005. Myzostomida: a review of the phylogeny and ultrastructure. *Hydrobiologia* 535–536: 253–275.

Fann, K.T. 1970. *Peirce's theory of abduction.* Martinus Nijhoff: The Hague. 62 pp.

Farris, J.S., Kallersjo, M., Kluge, A.G., and Bult, C. 1995. Testing significance of incongruence. *Cladistics* 10: 315–319.

Felsenstein, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution* 17: 368–376.

Felsenstein, J. 2004. *Inferring phylogenies.* Sinauer Associates Inc.: Sunderland. 664 pp.

Fetzer, J.H. 1993. *Philosophy of science.* Paragon House: New York. 197 pp.

Fetzer, J.H., and Almeder R.F. 1993. *Glossary of epistemology/ philosophy of science.* Paragon House: New York. 149 pp.

Fitzhugh, K. 2005a. Les bases philosophiques de l'inférence phylogénétique: une vue d'ensemble. *Biosystema* 24: 83–105.

Fitzhugh, K. 2005b. The inferential basis of species hypotheses: the solution to defining the term 'species'. *Marine Ecology* 26: 155–165.

Fitzhugh, K. 2006a. The abduction of phylogenetic hypotheses. *Zootaxa* 1145: 1–110.

Fitzhugh, K. 2006b. The 'requirement of total evidence' and its role in phylogenetic systematics. *Biology & Philosophy* 21: 309–351.

Fitzhugh, K. 2006c. The philosophical basis of character coding for the inference of phylogenetic hypotheses. *Zoologica Scripta* 35: 261–286.

Fitzhugh, K. 2008a. Fact, theory, test and evolution. *Zoologica Scripta* 37: 109–113.

Fitzhugh, K. 2008b. Abductive inference: implications for 'Linnean' and 'phylogenetic' approaches for representing biological systematization. *Evolutionary Biology* 35: 52–82.

Fitzhugh, K. 2008c. Clarifying the role of character loss in phylogenetic inference. *Zoologica Scripta* 37: 561–569.

Fitzhugh, K. 2009. Species as explanatory hypotheses: refinements and implications. *Acta Biotheoretica* 57: 201–248.

Fitzhugh, K. 2010a. Evidence for evolution versus evidence for intelligent design: parallel confusions. *Evolutionary Biology* 37: 68–92.

Fitzhugh, K. 2010b. Revised systematics of *Fabricia oregonica* Banse, 1956 (Polychaeta: Sabellidae: Fabriciinae): an example of the need for a uninomial nomenclatural system. *Zootaxa* 2647: 35–50.

Fitzhugh, K. 2012. The limits of understanding in biological systematics. *Zootaxa* 3435: 40–67.

Fitzhugh, K. 2013. Defining 'species', 'biodiversity', and 'conservation' by their transitive relations. Pp. 93–130 in: Pavlinov, I.Y. (ed.), *The species problem – ongoing problems.* InTech: New York.

Godfrey-Smith, P. 2003. *Theory and reality: an introduction to the philosophy of science.* University of Chicago Press: Chicago. 272 pp.

Goto, R., Okamoto, T., Ishikawa, H., Hamamura, Y., and Kato, M. 2013. Molecular phylogeny of echiuran worms (Phylum: Annelida) reveals evolutionary pattern of feeding mode and sexual dimorphism. *PLoS ONE* 8: e56809.

Hacking, I. 1965. *Logic of statistical inference.* Cambridge University Press: New York. 232 pp.

Hacking, I. 2001. *An introduction to probability and inductive logic.* Cambridge University Press: New York. 302 pp.

Halanych, K.M. 2005. Molecular phylogeny of siboglinid annelids (a.k.a. pogonophorans): a review. *Hydrobiologia* 535–536: 297–307.

Halanych, K.M., Feldman, R.A., and Vrijenhoek, R.C. 2001. Molecular evidence that *Sclerolinum brattstromi* is closely related to Vestimentiferans, not to frenulate pogonophorans (Siboglinidae, Annelida). *The Biological Bulletin* 201: 65–75.

Hall, K.A., Hutchings, P.A., and Colgan, D.J. 2004. Further phylogenetic studies of the Polychaeta using 18S rDNA sequence data. *Journal of the Marine Biological Association of the United Kingdom* 84: 949–960.

Hanson, N.R. 1958. *Patterns of discovery: an inquiry into the conceptual foundations of science.* Cambridge University Press: New York. 241 pp.

Hausman, D.M. 1998. *Causal asymmetries.* Cambridge University Press: New York. 300 pp.

Hempel, C.G. 1962. Deductive-nomological vs. statistical explanation. Pp. 98–169 in: Feigl, H. and Maxwell, G. (eds), *Minnesota studies in the philosophy of science, Vol. 3.* University of Minnesota Press: Minneapolis.

Hempel, C.G. 1965. *Aspects of scientific explanation and other essays in the philosophy of science.* The Free Press: New York. 505 pp.

Hempel, C.G. 1966. Recent problems of induction. Pp. 112–134 in: Colodny, R.G. (ed.), *Mind and cosmos.* University of Pittsburgh Press: Pittsburgh.

Hempel, C.G. 2001. *The philosophy of Carl G. Hempel: studies in science, explanation, and rationality.* Fetzer, J.H. (ed.) Oxford University Press: New York.

Hennig, W. 1966. *Phylogenetic systematics.* University of Illinois Press: Urbana. 263 pp.

Hoyningen-Huene, P. 2013. *Systematicity: the nature of science.* Oxford University Press: New York. 287 pp.

Howson, C., and Urbach, P. 1993. *Scientific reasoning: the Bayesian approach.* Open Court: Chicago. 470 pp.

Josephson, J.R., and Josephson, S.G. (eds) 1994. *Abductive inference: computation, philosophy, technology.* Cambridge University Press: New York. 306 pp.

Kelly, T. 2008. Common sense as evidence: against revisionary ontology and skepticism. *Midwest Studies in Philosophy* 32: 53–78.

Kuhn, T.S. 1970. *The structure of scientific revolutions.* University of Chicago Press: Chicago. 210 pp.

Kupriyanova, E.K., Macdonald, T.A., and Rouse, G.W. 2006. Phylogenetic relationships within Serpulidae (Sabellida, Annelida) inferred from molecular and morphological data. *Zoologica Scripta* 35: 421–439.

Kupriyanova, E.K., and Rouse, G.W. 2008. Yet another example of paraphyly in Annelida: molecular evidence that Sabellidae contains Serpulidae. *Molecular Phylogenetics and Evolution* 46: 1174–1181.

Lipton, P. 2004. *Inference to the best explanation.* Routledge: New York. 232 pp.

Lipton, P. 2008. Inference to the best explanation. Pp. 193–202 in: Stathis, P. and Curd, M. (eds), *The Routledge companion to philosophy of science.* Routledge: New York.

Longino, H.E. 1979. Evidence and hypothesis: an analysis of evidential relations. *Philosophy of Science* 46: 35–56.

Magesh, M., Kvist, S., and Glasby, C.J. 2012. Description and phylogeny of *Namalycastis jaya* sp. n. (Polychaeta, Nereididae, Namanereidinae) from the southwest coast of India. *ZooKeys* 238: 31–43.

Magnani, L. 2001. *Abduction, reason, and science: processes of discovery and explanation.* Kluwer Academic: New York. 205 pp.

Mahner, M., and Bunge, M. 1997. *Foundations of biophilosophy.* Springer: New York. 423 pp.

Marwick, P. 1999. Interrogatives and contrasts in explanation theory. *Philosophical Studies* 96: 183–204.

McHugh, D. 2005. Molecular systematics of polychaetes (Annelida). *Hydrobiologia* 535–536: 309–318.

McLaughlin, A. 1970. Rationality and total evidence. *Philosophy of Science* 37: 271–278.

Mickevich, M.F. 1978. Taxonomic congruence: Rohlf and Sokal's misunderstanding. *Systematic Zoology* 29: 162–176.

Neta, R. 2008. What evidence do you have? *The British Journal for the Philosophy of Science* 59: 89–119.

Nickles, T. 1980. Introductory essay: scientific discovery and the future of philosophy of science. Pp. 1–59 in: Nickles, T. (ed.), *Scientific discovery, logic and rationality*. D. Reidel Publishing Company: Dordrecht.

Nogueira, J.M.D.M., Fitzhugh, K., and Hutchings, P. 2013. The continuing challenge of phylogenetic relationships in Terebelliformia (Annelida : Polychaeta). *Invertebrate Systematics* 27: 186–238.

Nogueira, J.M.D.M., Fitzhugh, K., and Rossi, M.C.S. 2010. A new genus and new species of fan worms (Polychaeta: Sabellidae) from Atlantic and Pacific Oceans—the formal treatment of taxon names as explanatory hypotheses. *Zootaxa* 2603: 1–52.

Nola, R., and Sankey, H. 2007. *Theories of scientific method: an introduction*. McGill–Queen's University Press: Ithaca. 381 pp.

Norton, J.D. 2003. A material theory of induction. *Philosophy of Science* 70: 647–670.

Parry, L., Tanner, A., and Vinther, J. 2014. The origin of annelids. *Palaeontology* DOI: 10.1111/pala.12129.

Peirce, C.S. 1878. Illustrations of the logic of science. Sixth paper—deduction, induction, and hypothesis. *Popular Science Monthly* 13: 470–482.

Peirce, C.S. 1931. *Collected papers of Charles Sanders Peirce, Vol. 1. Principles of philosophy*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 393 pp.

Peirce, C.S. 1932. *Collected papers of Charles Sanders Peirce, Vol. 2. Elements of logic*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 535 pp.

Peirce, C.S. 1933a. *Collected papers of Charles Sanders Peirce, Vol. 3. Exact logic*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 433 pp.

Peirce, C.S. 1933b. *Collected papers of Charles Sanders Peirce, Vol. 4. The simplest mathematics*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 601 pp.

Peirce, C.S. 1934. *Collected papers of Charles Sanders Peirce, Vol. 5. Pragmatism and pragmaticism*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 455 pp.

Peirce, C.S. 1935. *Collected papers of Charles Sanders Peirce, Vol. 6. Scientific metaphysics*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 462 pp.

Peirce, C.S. 1958a. *Collected papers of Charles Sanders Peirce, Vol. 7. Science and philosophy*. Hartshorne, C., Weiss, P. and Burks, A.W. (eds). Harvard University Press: Cambridge. 415 pp.

Peirce, C.S. 1958b. *Collected papers of Charles Sanders Peirce, Vol. 8. Reviews, correspondence and bibliography*. Burks, A.W. (ed.). Harvard University Press: Cambridge. 352 pp.

Popper, K.R. 1970. Normal science and its dangers. Pp. 51–58 in: Lakatos, I. and Musgrave, A. (eds), *Criticism and the growth of knowledge*. Cambridge University Press: New York.

Popper, K.R. 1983. *Objective knowledge: an evolutionary approach*. Oxford University Press: New York. 395 pp.

Popper, K.R. 1992. *Realism and the aim of science*. Routledge: New York. 420 pp.

Psillos, S. 2002. Simply the best: a case for abduction. Pp. 605–625 in: Kakas, A.C. and Sadri, F. (eds), *Computational logic: logic programming and beyond*. Springer: New York.

Psillos, S. 2007. *Philosophy of science A–Z*. University Press: Edinburgh. 280 pp.

Psillos, S. 2011. An explorer upon untrodden ground: Peirce on abduction. Pp. 117–151 in: Gabbay, D., Hartmann, S. and Woods, J. (eds), *The handbook of the history of logic. Vol. 10: inductive logic*. Elsevier B.V.: Oxford.

Reilly, F.E. 1970. *Charles Peirce's theory of scientific method*. Fordham University Press: New York. 200 pp.

Rescher, N. 1970. *Scientific explanation*. The Free Press: New York. 242 pp.

Ronquist, F., van der Mark, P., and Huelsenbeck, J.P. 2009. Bayesian phylogenetic analysis using MRBAYES. Pp. 210–266 in: Lemey, P., Salemi, M. and Vandamme, A.-M. (eds), *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press: New York.

Rousset, V., Pleijel, F., Rouse, G.W., Erséus, C., and Siddall, M.E. 2006. A molecular phylogeny of annelids. *Cladistics* 22: 1–23.

Rousset, V., Rouse, G.W., Féral, J.-P., Desbruyères, D., and Pleijel, F. 2003. Molecular and morphological evidence of Alvinellidae relationships (Terebelliformia, Polychaeta, Annelida). *Zoologica Scripta* 32: 185–197.

Rousset, V., Rouse, G.W., Siddall, M.E., Tillier, A., and Pleijel, F. 2004. The phylogenetic position of Siboglinidae (Annelida) inferred from 18S rRNA, 28S rRNA and morphological data. *Cladistics* 20: 518–533.

Salmon, W.C. 1967. *The foundations of scientific inference*. University of Pittsburgh Press: Pittsburgh. 157 pp.

Salmon, W.C. 1984a. *Scientific explanation and the causal structure of the world*. Princeton University Press: Princeton. 305 pp.

Salmon, W.C. 1984b. *Logic*. Prentice-Hall Inc.: Englewood Cliffs. 180 pp.

Salmon, W.C. 1989. Four decades of scientific explanation. Pp. 3–219 in: Kitcher, P. and Salmon, W.C. (eds), *Scientific explanation. Minnesota studies in the philosophy of science, Vol. XIII*. University of Minnesota Press: Minneapolis.

Salmon, W.C. 1998. *Causality and explanation*. Oxford University Press: New York. 434 pp.

Schmidt, H.A., and von Haeseler, A. 2009. Phylogenetic inference using maximum likelihood methods. Pp. 181–198 in: Lemey, P., Salemi, M. and Vandamme, A.-M. (eds), *The phylogenetic handbook: a practical approach to phylogenetic analysis and hypothesis testing*. Cambridge University Press: New York.

Schulze, A. 2006. Phylogeny and genetic diversity of palolo worms (*Palola*, Eunicidae) from the tropical North Pacific and the Caribbean. *The Biological Bulletin* 210: 25–37.

Schurz, G. 2005. Explanations in science and the logic of why-questions: discussion of the Halonen–Hintikka-approach and alternative proposal. *Synthese* 143: 149–178.

Schurz, G. 2008. Patterns of abduction. *Synthese* 164: 201–234.

Sintonen, M. 2004. Reasoning to hypotheses: Where do questions come? *Foundation of Science* 9: 249–266.

Sober E. 1975. *Simplicity*. Oxford University Press: New York. 189 pp.

Sober, E. 1986. Explanatory presupposition. *Australasian Journal of Philosophy* 64: 143–149.

Sober, E. 1988. *Reconstructing the past: parsimony, evolution, and inference*. MIT Press: Cambridge. 265 pp.

Sober, E. 1994. *From a biological point of view: essays in evolutionary biology*. Cambridge University Press: New York. 268 pp.

Sperling, E.A., Vinther, J., Moy, V.N., Wheeler, B.M., Sémon, M., Briggs, D.E., and Peterson, K.J. 2009. MicroRNAs resolve an apparent conflict between annelid systematics and their fossil record. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 276: 4315–4322.

Strahler, A.N. 1992. *Understanding science: an introduction to concepts and issues*. Prometheus Books: Buffalo. 409 pp.

Struck, T.H., Paul, C., Hill, N., Hartmann, S., Hösel, C., Kube, M., Lieb, B., Meyer, A., Tiedemann, R., Purschke, G., and Bleidorn, C. 2011. Phylogenomic analyses unravel annelid evolution. *Nature* 471: 95–98.

Struck, T.H., Schult, N., Kusen, T., Hickman, E., Bleidorn, C., McHugh, D., and Halanych, K.M. 2007. Annelid phylogeny and the status of Sipuncula and Echiura. *BMC Evolutionary Biology* 7: 57.

Thagard, P. 1988. *Computational philosophy of science*. The MIT Press: Cambridge. 240 pp.

Thagard, P. 2004. Rationality and science. Pp. 363–379 in: Mele, A. and Rawlings, P. (eds), *Handbook of rationality*. Oxford University Press: Oxford.

Van Fraassen, B.C. 1990. *The scientific image*. Clarendon Press: New York. 235 pp.

Vrijenhoek, R.C., Johnson, S.B., and Rouse, G.W. 2009. A remarkable diversity of bone-eating worms (*Osedax*; Siboglinidae; Annelida). *BMC Biology* 7: 74–86.

Walton, D. 2004. *Abductive reasoning*. The University of Alabama Press: Tuscaloosa. 303 pp.

Wheeler, W.C. 2012. *Systematics: a course of lectures*. Wiley–Blackwell: Oxford. 426 pp.

Whitehead, A.N. 1925. *Science and the modern world. Lowell lectures, 1925*. The Macmillan Company: New York. 212 pp.

Wiklund, H., Nygren, A., Pleijel, F., and Sundberg, P. 2008. The phylogenetic relationships between Amphinomidae, Archinomidae and Euphrosinidae (Amphinomida: Aciculata: Polychaeta), inferred from molecular data. *Journal of the Marine Biological Association of the United Kingdom* 88: 509–513.

Williamson, T. 2000. *Knowledge and its limits*. Oxford University Press: New York. 340 pp.

Zanol, J., Halanych, K.M., and Fauchald, K. 2010. Phylogeny of the bristle worm family Eunicidae (Eunicida, Annelida) and the phylogenetic utility of noncongruent *16S*, *COI* and *18S* in combined analyses. *Molecular Phylogenetics and Evolution* 55: 660–676.

Zrzavý, J., Říha, P., Piálek, L., and Janouškovec, J. 2009. Phylogeny of Annelida (Lophotrochozoa): total-evidence analysis of morphology and six genes. *BMC Evolutionary Biology* 9: 189.