

## EXPLANATION

James S. Farris<sup>1</sup>, Mari Källersjö<sup>1</sup>, Victor A. Albert<sup>2</sup>, Marc Allard<sup>3</sup>,  
 Arne Anderberg<sup>4</sup>, Brunella Bowditch<sup>5</sup>, Carol Bult<sup>6</sup>, James M. Carpenter<sup>7</sup>,  
 Timothy M. Crowe<sup>8</sup>, Jan De Laet<sup>9</sup>, Kirk Fitzhugh<sup>10</sup>, Darryl Frost<sup>11</sup>,  
 Pablo Goloboff<sup>12</sup>, Christopher J. Humphries<sup>13</sup>, Ulf Jondelius<sup>14</sup>, Darlene Judd<sup>15</sup>,  
 Per Ola Karis<sup>4</sup>, Diana Lipscomb<sup>3</sup>, Melissa Luckow<sup>16</sup>, David Mindell<sup>17</sup>,  
 Jyrki Muona<sup>18</sup>, Kevin Nixon<sup>16</sup>, William Presch<sup>19</sup>, Ole Seberg<sup>20</sup>, Mark E. Siddall<sup>21</sup>,  
 Lena Struwe<sup>2</sup>, Anders Tehler<sup>22</sup>, John Wenzel<sup>23</sup>, Quentin Wheeler<sup>24</sup>,  
 and Ward Wheeler<sup>25</sup>

<sup>1</sup>Molekylärsystematiska laboratoriet, Naturhistoriska riksmuseet, Box 50007, S-104 05 Stockholm, Sweden, <sup>2</sup>New York Botanical Garden, Bronx, New York 10458 U.S.A., <sup>3</sup>Department of Biological Sciences, George Washington University, Washington, D.C. 20052, U.S.A., <sup>4</sup>Sektionen för fanerogambotani, Naturhistoriska riksmuseet, Box 50007, S-104 05 Stockholm, Sweden, <sup>5</sup>Department of Biology, Trinity College, 125 Michigan Ave NE, Washington, D.C. 20009 U.S.A., <sup>6</sup>Institute for Genomic Research, 932 Clopper Rd, Gaithersburg, Maryland 20878 U.S.A., <sup>7</sup>Department of Entomology, American Museum of Natural History, Central Park West at 79th St, New York, New York 10024 U.S.A., <sup>8</sup>FitzPatrick Institute of Ornithology, University of Cape Town, Rondebosch 7700, Cape Town, South Africa, <sup>9</sup>Laboratory of Plant Systematics, Botanical Institute K. U. Leuven, Kard. Mercierlaan 92, B-3001 Heverlee, Belgium, <sup>10</sup>Department of Invertebrates, Los Angeles County Museum, Los Angeles, California U.S.A., <sup>11</sup>Department of Herpetology, American Museum of Natural History, Central Park West at 79th St, New York, New York 10024 U.S.A., <sup>12</sup>Instituto Superior de Entomologica, Instituto Miguel Lillo, Miguel Lillo 205, 4000 S.M. de Tucumán, Argentina, <sup>13</sup>Department of Botany, The Natural History Museum, London SW7 5BD, U.K., <sup>14</sup>Sektionen för evertebratzoologi, Naturhistoriska riksmuseet, Box 50007, S-104 05 Stockholm, Sweden, <sup>15</sup>Department of Entomology, Texas A&M University, College Station, Texas 77843, U.S.A., <sup>16</sup>L. H. Bailey Hortorium, 462 Mann Library, Cornell University, Ithaca, New York 14853, U.S.A., <sup>17</sup>Department of Ornithology, Museum of Zoology, The University of Michigan, Ann Arbor, Michigan 48109 U.S.A., <sup>18</sup>Eläinmuseo, P. Rautatiekatu 13, 00014 Helsingin Yliopisto, Helsinki, Finland, <sup>19</sup>Department of Biology, California State University, Fullerton, California 92634 U.S.A., <sup>20</sup>Botanical Institute, University of Copenhagen, Gothersgade 140, DK-1123 Copenhagen K, Denmark, <sup>21</sup>Virginia Institute of Marine Science, School of Marine Science, The College of William and Mary, Gloucester Point, Virginia 23062, U.S.A., <sup>22</sup>Botaniska Institutionen, Stockholms universitet, S-106 91 Stockholm, Sweden, <sup>23</sup>Department of Entomology, Ohio State University, Columbus, Ohio 43210 U.S.A., <sup>24</sup>Department of Entomology, Cornell University, Ithaca, New York 14853 U.S.A., <sup>25</sup>Department of Invertebrates, American Museum of Natural History, Central Park West at 79th St, New York, New York 10024 U.S.A.

Received for publication 9 June 1994; accepted 31 August 1994

### Introduction

Kluge (1993) criticized the three-taxon approach (of Nelson and Platnick, 1991; hereinafter 3ta), while Platnick (1993) viewed those criticisms as largely misplaced, but each author considered his position consistent with Farris' (1983) discussion of the relationship between parsimony and explanatory power. In that latter regard the disagreement would appear to result from overlooking some aspects of Farris' treatment, and our purpose here is to show how this misunderstanding can be resolved.

### Parsimony

Platnick (1993: 271) put the matter succinctly.

"Farris [(1983)] showed that those cladograms which imply fewer steps are to be preferred on grounds of their ability to summarize, describe, and explain the data. Do similar arguments justify the three-taxon approach?"

Fewer steps, that is, according to the standard approach, as Platnick called it. On the same page he concluded,

"The two approaches both use parsimony, and rely on the same type of justification for that use."

By parsimony he meant that each method minimizes extra steps of some kind, but that analogy need not mean that both are covered by the same justification. To show where differences may arise, we begin by summarizing Farris' (1983) discussion.

The connection between extra steps and explanatory power has several parts. There is first the matter of what is explained (Farris, 1983: 18).

"Genealogies provide only a single kind of explanation. A genealogy does not explain by itself why one group acquires a new feature while its sister group retains the ancestral trait. . . A genealogy is able to explain observed points of similarity among organisms just when it can account for them as identical by virtue of inheritance from a common ancestor."

Similarities not so explained might be produced by mistakes in observation or coding, possibilities that can be evaluated by checking the data. Similarities still unexplained after such checking must be attributed *ad hoc* to homoplasy. But a theory loses explanatory power when an *ad hoc* hypothesis must be introduced (cf. Farris, 1983: 17ff).

Requirements for *ad hoc* hypotheses are then to be minimized, but the counting must be done carefully (Farris, 1983: 19).

"It is thus important to ensure that the homoplasies combined in such totals are logically independent, since otherwise their number need not reflect required *ad hoc* hypotheses."

Explanatory power is reduced when a new *ad hoc* hypothesis is required, but not when all or part of an old one is repeated.

This means, for example, that (as is well known) logically interdependent characters should not be used as if they comprised separate evidence. Another source of dependence will be particularly pertinent here (Farris, 1983: 20).

"Suppose . . . that a putative genealogy distributes [the 20 terminals showing feature X] into two distantly related groups A and B of 10 terminals each. There are 100 distinct two-taxon comparisons of members of A with members of B, and each of those similarities in X considered in isolation comprises a homoplasy. Those homoplasies do not constitute independent required

hypotheses, however. The genealogy . . . is consistent with identity by descent of X within each group. If X is identical by descent in any two members of A, and also in any two members of B, then the A-B similarities are all homoplasies if any one of them is. The genealogy thus requires but a single ad hoc hypothesis of homoplasy. Of course the numbers in the groups do not matter; the same conclusion would follow if they were 15 and 5, or 19 and 1."

Despite the apparent complexity of such cases, the problem of counting hypotheses turns out to have a simple solution (Farris, 1983: 20).

"A hypothesis of homoplasy logically independent of others is required precisely when a genealogy requires an additional origin of a feature."

In standard parsimony each extra step corresponds to an additional origin of feature, and this leads at last to the main result. The number of independent *ad hoc* hypotheses of homoplasy required by a genealogy increases with the number of extra steps. Ability to explain observed similarities thus decreases as extra steps increase.

### Dependence

That finding, however, pertains to extra steps as calculated in the standard method. The extra steps in 3ta do not generally correspond in number to origins of features; rather they count discarded three-taxon statements, those not conforming to the considered tree. To arrive at his conclusion, that 3ta is justified in the same way as standard parsimony, Platnick equated discarded statements with *ad hoc* hypotheses. But this disregards possible dependence among those statements.

The effect seen in Farris' 1983 example can become more pronounced in 3ta. Three-taxon statements include two-taxon similarities, and their number is further compounded by choices of the third taxon. This is seen in Fig. 1, where 20 terminals have apomorphic state 1 of a character, 12 others showing plesiomorphic state 0. In Case 1, groups A and B each comprise 10 of the 20 1s. As before, only a

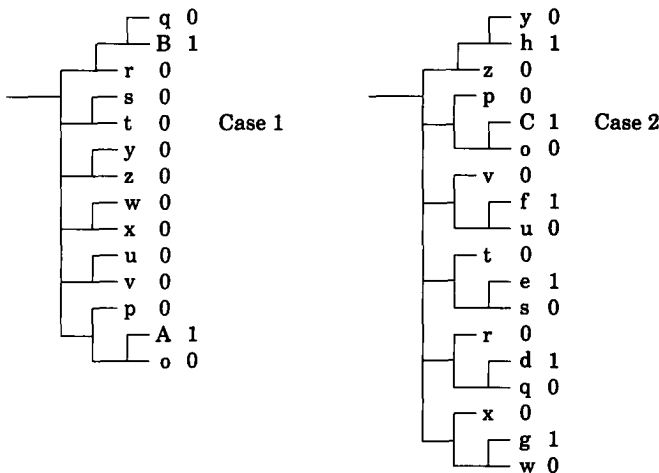


Fig. 1. Two cases illustrating dependence among three-taxon statements. For explanation see text.

single hypothesis of homoplasy is required. But 1200 three-taxon statements are discarded.

Counting discarded three-taxon statements usually amounts to counting each independent hypothesis of homoplasy several times, but simple inflation is not the greatest difficulty. Worse is that the degree of redundancy is not uniform, so that the two counts need not be well correlated.

The algebraic basis for Case 1 is that the number of discarded three-taxon statements is proportional to the product of the sizes (numbers of included terminals) of the groups into which the apomorphic state is divided. Because that product becomes relatively smaller when the groups are of unequal size, dependence among statements can produce results like that seen in Case 2. There group C consists of 15 of the 1s, each of the other five being placed on a separate branch of the tree.

Compared to Case 1, explanatory power is now decreased: the number of standard extra steps—of independent hypotheses of homoplasy required—increases to five. But the number of discarded three-taxon statements decreases to 1020. Counting each discarded statement as an *ad hoc* hypothesis can lead to the wrong conclusion when explanatory ability is to be compared between different trees.

Platnick (1993) did not discuss dependence, but others have addressed related subjects. Under "Character Dependence and Redundancy" Nelson (1993: 264) suggested,

"Each of A(DE), B(DE), and C(DE) is a statement of homology (relationship) relative to a different taxon (A, B, or C). The statements are separate evidence because each one, or pair, or all three might fit a particular tree, and the remainder, if any, might not."

That kind of separateness is not sufficient for independence of *ad hoc* hypotheses. Say that terminals *a* and *b* belong respectively to groups A and B of Fig. 1. Statements *s(ab)*, *x(ab)* and *y(ab)* (as well as others) are separate in Nelson's sense. But all are discarded by the tree of Case 1, although only a single hypothesis of homoplasy is required.

Nelson's reasoning, in fact, would even make pairwise similarities "separate evidence". Various trees could fit one, two, or all three of the two-taxon similarities XD, YD and ZD. But a tree might discard all three, yet require just one hypothesis of homoplasy. This could occur if X, Y and Z were grouped together while separated from D.

Dependence among *ad hoc* hypotheses is not identical to dependence among observations, since separate observations may be covered by a single *ad hoc* hypothesis. The two kinds of dependence are nonetheless connected. If "separate" characters are interdependent, "separate" *ad hoc* hypotheses covering them will generally be interdependent as well.

Nelson and Ladiges (1992) knew that three-taxon statements can be interdependent, and they proposed to take this into account by assigning a fractional weight to each statement. The weight is the ratio  $p/q$  between ( $q$ ), the number of statements derivable from some particular standard character, and ( $p$ ), the number of such statements taken to be independent.

That technique may affect the relative weights of statements from different standard characters, but any two statements from the same character always have the same weight. The weighting would thus have no effect on the outcome of the preceding example, since only one standard character is involved. Although

employing information on dependence, this method cannot undo the deleterious effects of dependence.

### Stems

Another difficulty is touched on by Platnick's (1993: 271) "only".

"A cladogram can only explain the homology statement expressed in a tree-taxon statement by common cause (ancestry) if that tree-taxon statement can be accommodated at one of its nodes."

In trying to assess explanation, 3ta counts every accommodated (undiscarded) statement, in effect proceeding as if "only" meant "always", and this can lead to contradictions.

To see this, first observe that, while Platnick mentioned just one homology, a three-taxon statement actually makes a further claim. According to statement H(IJ), I and J share an apomorphy, while H shows the plesiomorphy. To explain that statement, a tree must account for both parts as the result of inheritance. The apomorphic similarity between I and J cannot be attributed to parallelism, but neither can the plesiomorphy of H be attributed to reversal.

A convenient illustration is provided by Platnick's (1993: his Fig. 1) examples, reproduced in Fig. 2. There six terminals *a*, *f* show apomorphic state 1 of the character, nine others *r*, *z* having plesiomorphic state 0, for a total of 135 three-taxon statements. State 0 is presumed present at the root of all three trees.

That both apomorphies and plesiomorphies must be accounted for at once is seen in Tree 4, which (as Platnick pointed out) discards all 135 statements. For  $r(ab)$ , for example, Tree 4 is consistent with a single origin of either state, but not with single origins for both. If the 1s shared by *a* and *b* are not parallel, the 0 of *r* must be a reversal. If that 0 is not a reversal, the 1s must be parallel.

When a tree does explain a three-taxon statement, the same requirements determine the condition of stem species. This is seen in Tree 3, which accommodates  $r(ab)$ . If *a* and *b* share 1s because of inheritance, their most recent common ancestor—which is the stem of (*a*, *c*)—must have state 1. If the 0 of *r* is not reversal, each ancestor of *r* must have state 0. This establishes the state in the stems of (*r*, *z*), of (*d*, *z*) and of (*a*, *z*).

The difficulty arises when requirements from different statements conflict. Tree 1 accommodates  $r(af)$  and  $x(ef)$ , but it cannot explain both at once. Explaining the shared 1s of *a* and *f* by inheritance demands that all 11 stems between *a* and *f* have state 1. But to account for the 0 of *x*, every ancestor of *x*—including eight of those same stems—must have state 0 instead.

Other pairs of statements show similar contradictions. It is easily found that while Tree 1 seems to accommodate 65 of the statements, it can account for at most 45 of them together. They are those obtained by combining any one of *r*, *s*, *t* with any two 1s.

Platnick meant his examples to illustrate differences between the two methods. Tree 1 requires five standard extra steps and discards 70 (accommodates 65) statements. Tree 3 requires only one standard extra step, but discards 81 (accommodates 54) statements. Tree 3 would be preferred by standard parsimony,

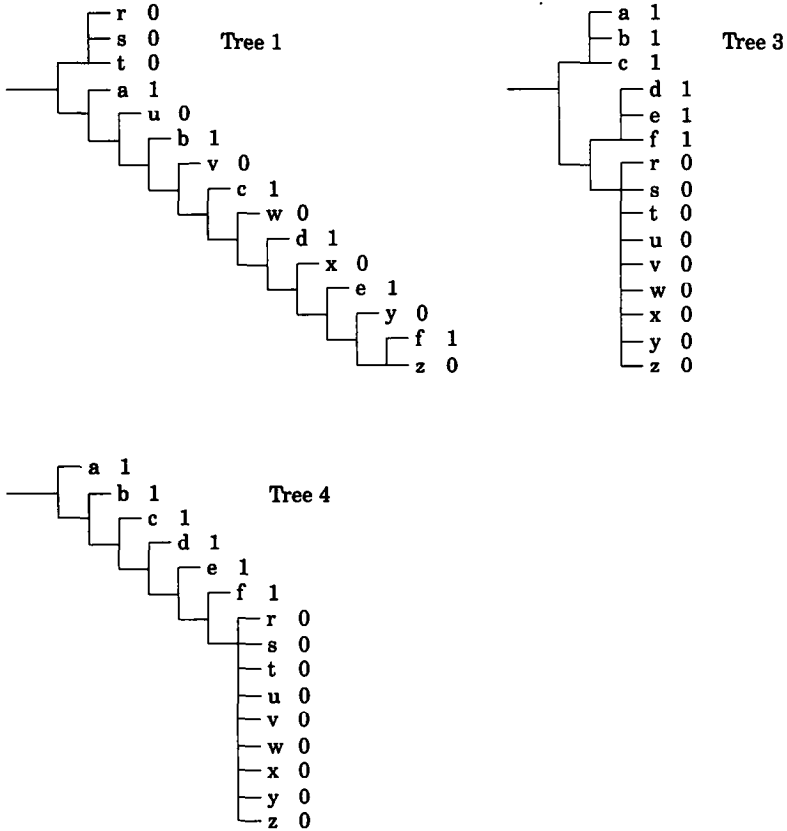


Fig. 2. Three examples after Platnick (1993: his Fig. 1). For explanation see text.

whereas—he supposed—Tree 1 would provide the better explanation of three-taxon statements.

Similar discrepancies can be produced when 3ta is misled by redundancy, as seen before. But redundancy aside for the moment, the statement counts employed in 3ta do not provide a reliable indication of relative explanatory power. While Tree 1 seems to fit 65 statements, it can explain only 45 of them. Tree 3 conforms to just 54 statements, but it can explain them all<sup>1</sup>. Tree 3 should then be preferred, and the improved assessment agrees in this case with the redundancy-free evaluation provided by standard parsimony.

Programs for parsimony analysis reconstruct the states of stem species by optimization procedures (cf. Farris, 1970), and applying optimization to three-taxon statements has proved controversial (Platnick, 1993: 267).

“Using a standard parsimony program to optimize them is simply a misuse of the program.”.

For this reason it is important to note that the present discussion does not rely

<sup>1</sup>Assign state 1 to the stems of (*a* . *c*) and (*d* . *f*), state 0 to all other stems.

on standard optimization. The conditions attributed here to stem species are just those that follow logically from the premise that a three-taxon statement is explained by inheritance and common ancestry. The point of making those deductions is simply to show that explanations of accommodated statements may be mutually exclusive, which is to say that what 3ta counts does not measure explanation.

This difficulty arises because the three-taxon statements involve the same character. If  $r(af)$  and  $x(ef)$  instead described independent binary variables, the contradiction would vanish. While this drawback of 3ta seems superficially different from the redundancy problem seen earlier, the two thus have a common cause. Three-taxon statements concerning the same features are interrelated, but Nelson (quoted above) took them as separate evidence, and 3ta treats them accordingly.

### Data

Platnick (1993: 271) saw no difficulty in the 3ta evaluation, but felt instead that 3ta might properly give results different from those of standard parsimony. This is because:

"[The two methods] differ with regard to what it is they are trying to explain: minimally [in standard parsimony], a raw observation that one taxon has one feature, or [in 3ta] a hypothesis that only two out of three taxa share a homologous feature."

And (p. 271),

"The notion that systematic data constitute a normal character  $\times$  taxon matrix is not an intrinsically cladistic notion; indeed, that type of matrix seems to have originated with pheneticists. Consider an alternative view, that the three-taxon matrix instead constitutes systematic data."

The origin of that conception of what the standard method explains is a complete mystery<sup>2</sup>, but the general sense of the argument seems clear. The relative merits of trees may reasonably change when the evidence is written in a different notation. Nelson (1993: 261) expressed a similar view.

"[Kluge (1993)] assumes . . . that a matrix of binary characters (his matrix 1) is original data and that the corresponding three-item matrix (his matrix 2) is transformation. Why not assume the opposite and justify other conclusions?"

Justify 3ta, he meant<sup>3</sup>. As to why to assume the opposite (p. 261):

"Empirical data originate in observation. Do data originate there as binary characters or as three-item statements? . . . When and how do [mammary glands] originate in human understanding as a binary character, or as one or more three-item statement, rather than a source of milk to a young and hungry mouth?"

All this was intended to suggest that three-taxon statements may somehow be the units to count in assessing explanatory power. But even if that premise were accepted, the three-taxon approach would still give the wrong counts. In the last example, the relative explanatory power of trees is misjudged precisely because the statements that 3ta counts as accommodated cannot all be explained.

<sup>2</sup>As is the reason for Platnick's theory on the history of data matrices. Pheneticists had simply adopted the sort of matrix long employed in multivariate statistics.

<sup>3</sup>Kluge's (1993) use of "original" and "transformed" followed Nelson and Platnick's (1991).

Not that these comments actually provide any support for that premise. Fascinating thought it is, Nelson's theory of human perception is completely irrelevant. The issue is not how we notice that mammals share mammary glands, or how we formulate that discovery, but rather just how shared features bear on phylogenetic inferences.

While he avoided psychology, Platnick followed Nelson in regarding evaluation as a matter of notation and, in so doing, he missed one of the main points of Farris' (1983) treatment. The counts used in standard parsimony are not chosen because the data are written as a standard matrix, but because they reflect the number of independent *ad hoc* hypotheses of homoplasy. It is that independence that is crucial to assessing explanatory power. The fault of the three-taxon approach is simply that it treats as independent, quantities that are in fact interrelated.

### Acknowledgements

This work was supported by NFR grants 10204-300 to J. S. Farris and 09858-303 to M. Källersjö. We thank Prof. K. Bremer for his excellent advice and strong encouragement.

### REFERENCES

- FARRIS, J. S. 1970. Methods for computing Wagner trees. *Syst. Zool.* 19: 83-92.
- FARRIS, J. S. 1983. The logical basis of phylogenetic analysis. *In*: N. I. Platnick and V. A. Funk (eds), *Advances in Cladistics*, Vol. 2: Proceedings of the Second Meeting of the Willi Hennig Society. Columbia Univ. Press, New York, pp. 7-36.
- KLUGE, A. G. 1993. Three-taxon transformation in phylogenetic inference: Ambiguity and distortion as regards explanatory power. *Cladistics* 9: 246-259.
- NELSON, G. 1993. Reply. *Cladistics* 9: 261-265.
- NELSON G. AND P. Y. LADIGES. 1992. Information content and fractional weight of three-item statements. *Syst. Biol.* 41: 490-494.
- NELSON, G. AND N. I. PLATNICK. 1991. Three-taxon statements: A more precise use of parsimony? *Cladistics* 7: 351-366.
- PLATNICK, N. I. 1993. Character optimization and weighting: Differences between the standard and three-taxon approaches to phylogenetic inference. *Cladistics* 9: 267-272.