

- GAREY, M. R., AND D. S. JOHNSON. 1979. Computers and intractability: A guide to the theory of NP-completeness. W. H. Freeman, San Francisco.
- HILLIS, D. M. 1996. Inferring complex phylogenies. *Nature* 383:130–131.
- HILLIS, D. M., J. P. HUELSENBECK, AND C. W. CUNNINGHAM. 1994. Application and accuracy of molecular phylogenies. *Science* 264:671–677.
- KELLOGG, E. A., AND C. S. CAMPBELL. 1987. Phylogenetic analyses of the Gramineae. Pages 310–322 in *Grass systematics and evolution* (T. Soderstrom, K. Hillu, C. Campbell, and M. Barkworth, eds.). Smithsonian Institution Press, Washington, D.C.
- KIM, J. 1996. General inconsistency conditions for maximum parsimony: Effects of branch lengths and increasing numbers of taxa. *Syst. Biol.* 45:363–374.
- LES, D., D. K. GARVIN, AND C. F. WIMPEE. 1991. Molecular evolutionary history of ancient aquatic angiosperms. *Proc. Natl. Acad. Sci. USA* 88:10119–10123.
- MADDISON, D. R. 1991. The discovery and importance of multiple islands of most-parsimonious trees. *Syst. Zool.* 40:315–328.
- MADDISON, D. R., M. RUVOLO, AND D. L. SWOFFORD. 1992. Geographic origins of human mitochondrial DNA: Phylogenetic evidence from control region sequences. *Syst. Biol.* 41:111–124.
- MADDISON, W. P., M. J. DONOGHUE, AND D. R. MADDISON. 1984. Outgroup analysis and parsimony. *Syst. Zool.* 33:83–103.
- MADDISON, W. P., AND D. R. MADDISON. 1992. *MacClade: Interactive analysis of phylogeny and character evolution*, version 3.04. Sinauer, Sunderland, Massachusetts.
- MISHLER, B. D. 1994. Cladistic analysis of molecular and morphological data. *Am. J. Phys. Anthropol.* 94:143–156.
- OLMSTEAD, R. G., B. BREMER, K. M. SCOTT, AND J. D. PALMER. 1993. A parsimony analysis of the Asteridae sensu lato based on *rbcL* sequences. *Ann. Mo. Bot. Gard.* 80:700–722.
- OLMSTEAD, R. G., AND J. D. PALMER. 1994. Chloroplast DNA systematics: A review of methods and data analysis. *Am. J. Bot.* 8:1205–1224.
- PAGE, R. D. M. 1993. On islands of trees and the efficiency of different methods of branch swapping in finding most-parsimonious trees. *Syst. Biol.* 42:200–210.
- PENNY, D., M. A. STEEL, P. J. WADDELL, AND M. D. HENDY. 1995. Improved analyses of human mtDNA sequences support a recent African origin for *Homo sapiens*. *Mol. Biol. Evol.* 12:863–882.
- SAITOU, N., AND M. NEI. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4:406–425.
- SOBER, E. 1988. *Reconstructing the past: Parsimony, evolution, and inference*. MIT Press, Cambridge, Massachusetts.
- SWOFFORD, D. L. 1990. PAUP: Phylogenetic analysis using parsimony, version 3.0. Illinois Natural History Survey, Champaign.
- SWOFFORD, D. L. 1993. PAUP: Phylogenetic analysis using parsimony, version 3.1.1. Illinois Natural History Survey, Champaign.
- SWOFFORD, D. L., G. J. OLSEN, P. J. WADDELL, AND D. M. HILLIS. 1996. Phylogenetic inference. Pages 407–514 in *Molecular systematics*, 2nd edition (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer, Sunderland, Massachusetts.
- TEMPLETON, A. 1992. Human origins and analysis of mitochondrial DNA sequences. *Science* 255:737–739.
- YEATES, D. K. 1995. Groundplans and exemplars: Paths to the tree of life. *Cladistics* 11:343–357.
- ZHARKIKH, A., AND W.-H. LI. 1993. Inconsistency of the maximum-parsimony method: The case of five taxa with a molecular clock. *Syst. Biol.* 42:113–125.

Received 5 August 1996; accepted 18 October 1996
Associate Editor: David Cannatella

Syst. Biol. 46(3):563–569, 1997

Stratigraphic Indices and Tree Balance

REBECCA HITCHIN AND MICHAEL J. BENTON

Department of Geology, University of Bristol, Wills Memorial Building, Queens Road, Bristol BS8 1RJ, England;
E-mail: r.hitchin@bristol.ac.uk (R.H.), mike.benton@bristol.ac.uk (M.J.B.)

Siddall (1996) claimed that (1) the number of internal nodes in a cladogram (n) is correlated with tree balance (Heard's [1992] index of imbalance, I_m), (2) Huel- senbeck's (1994) stratigraphic consistency index (SCI) is correlated with the recip-

cal of n (Siddall, 1996: fig. 1), and (3) SCI is therefore correlated with I_m (Siddall, 1996: fig. 6).

Siddall tested these assertions with a database of 14 cladograms taken from Huel- senbeck's (1994) original SCI analysis and

found a highly significant relationship between SCI and I_m . He concluded that the SCI "can be used neither as an unbiased summary statistic of stratigraphic fit nor as an alternative optimality criterion" (Siddall, 1996:115).

The SCI is the only stratigraphic index that has been explicitly mathematically analyzed for biases caused by tree balance. There are two other commonly used indices, the relative completeness index (RCI; Benton and Storrs, 1994, 1996; Benton, 1995; Benton and Simms, 1995; Benton and Hitchin, 1996) and the Spearman rank correlation (SRC; Norell, 1992, 1993; Norell and Novacek, 1992a, 1992b), and we tested these along with the SCI for correlation with tree balance. We used a larger database (357 published cladograms) than Siddall's to test the validity of his claims 2 and 3, i.e., the relationships between SCI and n and between SCI and I_m .

METHODS AND MATERIALS

Our data set contained 357 trees, of which 120 were for fishes, 175 for land vertebrates, and 62 for echinoderms. The data produced are heterogeneous and could be analyzed as a whole. However, they are separated into taxonomic groups that reflect underlying differences in life environment (terrestrial/aquatic) and classification (vertebrate/invertebrate). These groupings follow previous work in this field (Benton and Simms, 1995; Benton and Hitchin, 1996; Hitchin and Benton, 1997). The cladograms were extracted from published cladistic analyses and were used in a pectinate form for the SRC test but in a full form for the SCI and RCI tests.

Methods of cladogram manipulation for polytomies and the measurement of I_m followed Heard (1992) and Siddall (1996). All cladograms with polytomies were removed from the analysis (thus reducing the data set from the original 376 cladograms; Benton and Hitchin, 1996) because polytomy reduction can affect the resultant SCI and I_m results.

Tree balance was measured by Heard's (1992) index of imbalance (I_m):

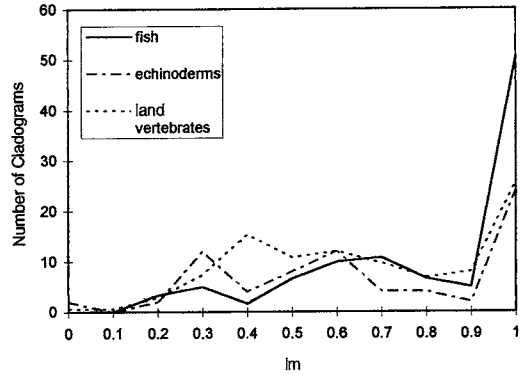


FIGURE 1. Distribution of I_m values in the data set. Fishes: $\bar{x} = 0.78$, $SD = 0.02$. Land vertebrates: $\bar{x} = 0.63$, $SD = 0.02$. Echinoderms: $\bar{x} = 0.62$, $SD = 0.05$.

$$I_m = \sum \frac{|T_r - T_l|}{(n-1)(n-2)/2'}$$

where T_r = number of terminal taxa subtended to the right of the node, T_l = number of terminal taxa subtended to the left of the node, and n = number of terminal taxa.

The index results range from 0 in a symmetrical or balanced cladogram to 1 in a completely pectinate (imbalanced) cladogram. The outgroup node was not included in the calculation of I_m to match the method of SCI calculation.

The SCI, SRC, and RCI are used in their logarithmically transformed form to avoid the problem of raw data variances being correlated with the means.

CONSISTENCY AND BALANCE

Siddall's (1996) first assertion is that I_m decreases with increasing number of terminal taxa (n). For any given number of taxa, there is a finite number of possible tree topologies, which increases dramatically with increasing n (Mooers, 1994). For example, for $n = 3$, there can only be one fully imbalanced topology. For $n = 5$, three distinct unbalanced topologies can be formed, but only one of these is fully imbalanced (pectinate) and the other two show some degree of symmetry. For $n = 6$, six topologies are possible; again only one is completely imbalanced, and the other five show varying degrees of symmetry.

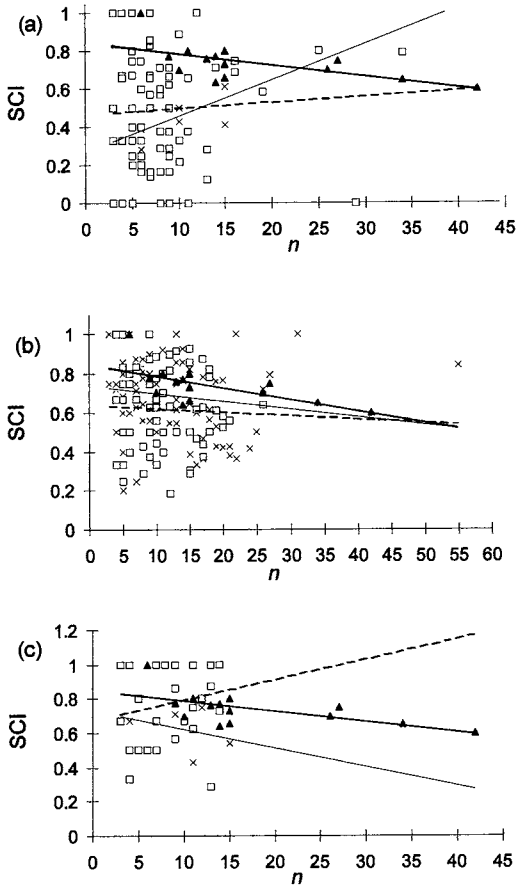


FIGURE 2. Relationship between n and SCI. \square = unmodified data; \times = modified data; \blacktriangle = Siddall's data; --- = unmodified data best fit line; light line = modified data best fit line; heavy line = Siddall's data best fit line. Siddall's data were plotted for comparison and were not used in the calculation of the relationship. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

Thus, the amount of balance (i.e., a decrease of Im) should increase with a greater number of nodes, an expectation that is fulfilled by the logarithmic decrease of the expected value of Im with an increase in n , assuming a Markovian null model (Siddall, 1996: fig. 5). The validity of this first process in Siddall's correlation of SCI and tree balance is not in doubt.

Siddall's second and third premises depend upon each other: if SCI is correlated with $1/n$, then it is inevitable that SCI is correlated with Im . However, the validity of the relationship with n is in doubt, even

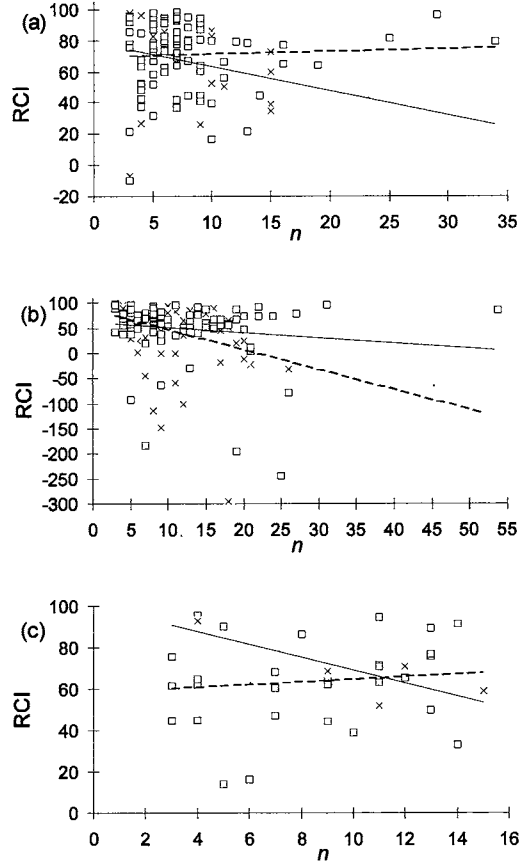


FIGURE 3. Relationship between n and RCI. \square = unmodified data; \times = modified data; --- = unmodified data best fit line; — = modified data best fit line. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

though Huelsenbeck's data set showed a relationship between SCI and n of $r = -0.597$ and $P = 0.022$. Because of this doubt, the relationship between SCI and Im is also in question.

RESULTS

Distribution of Im Values through the Data Set

Of the fish cladograms tested, 48% had an Im value of 1, showing that they had a perfectly pectinate structure (Fig. 1). No cladograms showed a perfectly balanced structure; the lowest value was 0.11.

The land vertebrate and echinoderm data sets both had one cladogram that was fully balanced. However, 24% of the land

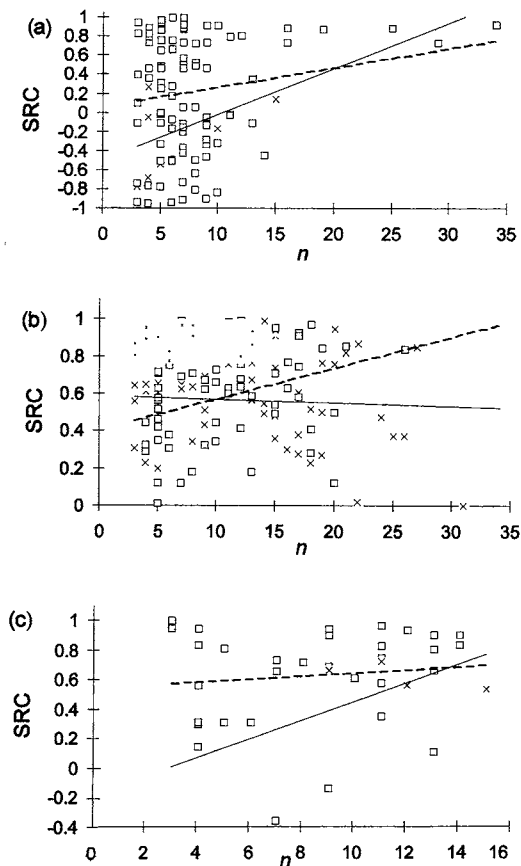


FIGURE 4. Relationship between n and SRC. \square = unmodified data; \times = modified data; --- = unmodified data best fit line; — = modified data best fit line. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

vertebrate cladograms were fully pectinate, whereas only 18% of the echinoderm cladograms showed this imbalanced state (Fig. 1).

Relationship between Stratigraphic Indices and n

In the plots of our 357 cladograms, no significant relationships were found, in contrast to Siddall's strong negative relationship (Figs. 2–5; Table 1). The $\ln(\text{SCI})$, $\ln(\text{RCI})$, and $\ln(\text{SRC})$ showed no significant relationship at a 95% significance level with any of the subsets tested.

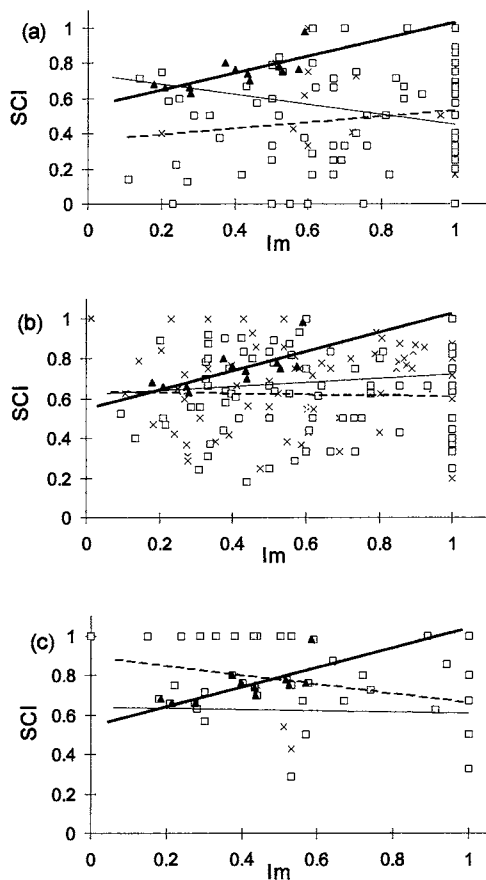


FIGURE 5. Relationship between Im and SCI. \square = unmodified data; \times = modified data; \blacktriangle = Siddall's data; --- = unmodified data best fit line; light line = modified data best fit line; heavy line = Siddall's data best fit line. Siddall's data were plotted for comparison and were not used in the calculation of the relationship. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

Relationship between Stratigraphic Indices and Im

Siddall (1996) claimed a positive relationship between SCI and Im , with a significance of $P = 0.002$. In our more extensive plots, we found a significant relationship between $\ln(\text{SCI})$ and Im for echinoderms (Fig. 5; Table 2), although the correlation was negative ($r = -3.78$, $P = 0.019$). All other plots of $\ln(\text{SCI})$, $\ln(\text{RCI})$, and $\ln(\text{SRC})$ with Im gave nonsignificant positive correlations (Figs. 5–7; Table 2).

TABLE 1. Summary statistics for the relationships between stratigraphic indices and n .

Group	ln(SCI)		ln(RCI)		ln(SRC)	
	r	P	r	P	r	P
Fishes	-1.8	0.062	-0.45	0.684	-0.45	0.707
Land vertebrates	5.0	0.660	3.0	0.051	0.0	0.424
Echinoderms	13.0	0.019	10.3	0.064	3.8	0.069

DISCUSSION

Distribution of Im Values through the Data Sets

The set of published cladograms of all three groups of organisms used here were biased towards imbalance, with the fishes showing the largest percentage (48%) of totally pectinate trees. Both land vertebrates and echinoderms had between 20% and 30% totally pectinate trees, with a relatively large number of trees showing an Im value of 0.3–0.6. Fishes had relatively few cladograms with an Im of <1.

This preponderance towards pectinate trees, especially for large cladograms, has been noted several times (Guyer and Slowinski, 1991; Mooers, 1994) because tree shape can give insights into the forces that shape the phylogenies. However, there are several ways in which the production method of the cladogram can bias the resulting topology, one of the most important being completeness of the tree (i.e., whether a full monophyletic group is presented) (Mooers, 1994). Also, the preponderance noted may have been caused by error. When there is an error in the phylogenetic method, more trees are produced that are asymmetric relative to the expectation under a random branching (Markovian) model of cladogenesis (Huelsenbeck and Kirkpatrick, 1996).

In further study of this data set, we will attempt to elucidate whether this pattern of

imbalance truly is an evolutionary phenomenon or simply is a by-product of the selection of taxa.

Tree Balance and Stratigraphic Indices

We cannot confirm Siddall's finding of a strong correlation of SCI or RCI with either n or Im. Siddall may have uncovered spurious relationships between SCI and n and Im because of the small size of his tested sample of cladograms. There is also little evidence for a correlation of SRC with n or Im, although the SRC value for a cladogram can be influenced by the number of nodes contained within the cladogram (Norell and Novacek, 1992a, 1992b; Hitchin and Benton, 1997).

One use of the SCI is as an optimality criterion for resolving series of equally parsimonious trees, but this scenario is not analyzed here. Tree balance, discussed by Siddall (1996), is probably a problem when the SCI is used in this way because the preferred tree is often imbalanced (Mooers, 1994). Thus, it is important to notice tree balance when using the SCI.

However, Siddall (1996) also commented that his postulated links between SCI and n and Im weaken the suitability and utility of the index as a descriptive statistic. We suggest, however, that this relationship does not exist, at least within our larger data set of cladograms of land and aquatic vertebrates and invertebrates. The SCI

TABLE 2. Summary statistics for the relationships between the stratigraphic indices and Im.

Group	ln(SCI)		ln(RCI)		ln(SRC)	
	r	P	r	P	r	P
Fishes	0.0	0.842	-0.6	0.524	-1.6	0.184
Land vertebrates	-0.3	0.461	-0.8	0.323	0.32	0.760
Echinoderms	-0.9	0.576	-0.9	0.589	1.1	0.556

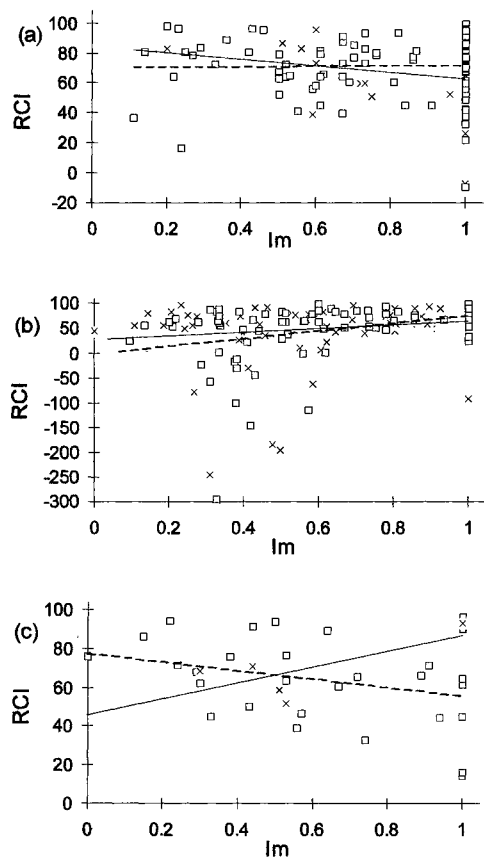


FIGURE 6. Relationship between Im and RCI. \square = unmodified data; \times = modified data; --- = unmodified data best fit line; — = modified data best fit line. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

seems the best available statistic for measuring the fit of observed stratigraphy and estimated phylogenies, and the use of SCI should be promoted, although caution is advised because of other biases of the test (Hitchin and Benton, 1997). In the same way, the use of RCI should not be compromised by any suggestion of a possible bias of tree balance but should be promoted for use in describing the completeness of the fossil record for estimated phylogenies.

ACKNOWLEDGMENTS

Funding for this work was provided by the Leverhulme Trust (individual grants 1988 and grant F182/AK). John Huelsenbeck and Mark Siddall provided informative and interesting comments on this manuscript.

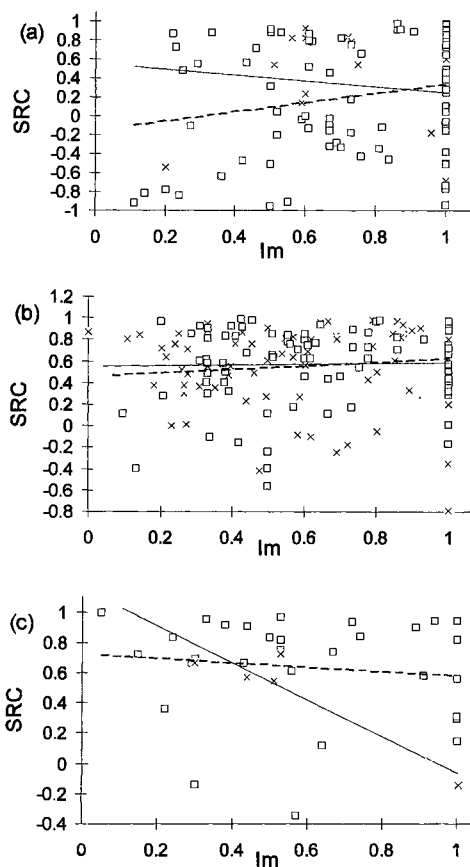


FIGURE 7. Relationship between Im and SRC. \square = unmodified data; \times = modified data; --- = unmodified data best fit line; — = modified data best fit line. (a) Fishes. (b) Land vertebrates. (c) Echinoderms.

REFERENCES

- BENTON, M. J. 1995. Testing the time axis of phylogenies. *Philos. Trans. R. Soc. Lond. B* 349:5-10.
- BENTON, M. J., AND R. HITCHIN. 1996. Testing the quality of the fossil record by groups and by major habitats. *Hist. Biol.* 12:111-157.
- BENTON, M. J., AND M. J. SIMMS. 1995. Testing the marine and continental fossil records. *Geology* 23: 601-604.
- BENTON, M. J., AND G. W. STORRS. 1994. Testing the quality of the fossil record: Paleontological knowledge is improving. *Geology* 22:111-114.
- BENTON, M. J., AND G. W. STORRS. 1996. Diversity in the past: Comparing cladistic phylogenies and stratigraphy. Pages 19-40 in *Aspects of the genesis and maintenance of biological diversity* (M. E. Hochberg, J. Clobert, and R. Barbault, eds.). Oxford Univ. Press, Oxford, England.
- GUYER, C., AND J. B. SLOWINSKI. 1991. Comparison of observed phylogenetic topologies with null expect-

- tations among three monophyletic lineages. *Evolution* 45:340–350.
- HEARD, S. P. 1992. Patterns in tree balance among cladistic, phenetic, and randomly generated phylogenetic trees. *Evolution* 46:1818–1826.
- HITCHIN, R., AND M. J. BENTON. 1997. Congruence between parsimony and stratigraphy: Comparisons of three indices. *Paleobiology* 23:20–32.
- HUELSENBECK, J. P. 1994. Comparing the stratigraphic record to estimates of phylogeny. *Paleobiology* 20: 470–483.
- HUELSENBECK, J. P., AND M. KIRKPATRICK. 1996. Do phylogenetic methods produce trees with biased shapes? *Evolution* 50:1418–1424.
- MOOERS, A. Ø. 1994. Tree balance and tree completeness. *Evolution* 49:379–384.
- NORELL, M. E. 1992. Taxic origin and temporal diversity: The effect of phylogeny. Pages 89–118 in *Extinction and phylogeny* (M. J. Novacek and Q. D. Wheeler, eds.). Columbia Univ. Press, New York.
- NORELL, M. E. 1993. Tree-based approaches to understanding history: Comments on ranks, rules and the quality of the fossil record. *Am. J. Sci.* 293A:407–417.
- NORELL, M. E., AND M. J. NOVACEK. 1992a. Congruence between superpositional and phylogenetic patterns: Comparing cladistic patterns with fossil records. *Cladistics* 8:319–337.
- NORELL, M. E., AND M. J. NOVACEK. 1992b. The fossil record and evolution: Comparing cladistic and paleontologic evidence for vertebrate history. *Science* 255:1690–1693.
- SIDDALL, M. E. 1996. Stratigraphic consistency and the shape of things. *Syst. Biol.* 45:111–115.

Received 26 June 1996; accepted 21 August 1996
Associate Editor: David Cannatella

Syst. Biol. 46(3):569–573, 1997

Stratigraphic Indices in the Balance: A Reply to Hitchin and Benton

MARK E. SIDDALL

Museum of Zoology, University of Michigan, Ann Arbor, Michigan 48109, USA; E-mail: msiddall@umich.edu

Upon reading Huelsenbeck's (1994) paper describing the stratigraphic consistency index (SCI) and inspection of the SCI values he had obtained for 14 data sets (Huelsenbeck, 1994: table 1), I noted that there seemed to be a negative relationship between SCI and the number of taxa or number of nodes (n). Insofar as the global consistency index (Kluge and Farris, 1969) was already known to exhibit a size bias, I thought it reasonable to investigate this same possibility in relation to the SCI. In so doing, I reported (Siddall, 1996) that Huelsenbeck's 14 trees showed a strongly negative relationship between SCI and number of taxa, and I postulated that this relationship likely had something to do with expectations of imbalance for data sets of increasing numbers of taxa. This suggestion seemed to be reasonably well corroborated in that there was a stronger relationship between SCI and Heard's (1992) index of imbalance (I_m) than there was between SCI and number of taxa (Sid-

dall, 1996: fig. 5). In light of my having outlined the logic as to why there might be a balance bias for the SCI (Siddall, 1996: figs. 2–4), I suggested that its use as an alternative optimality criterion (Huelsenbeck, 1994) was specious.

Benton and Hitchin (1996) revisited the issue of goodness of fit between the fossil record and cladograms with their relative completeness index (RCI). Moreover, they have examined the behavior of the RCI, the SCI, and the Spearman rank correlation coefficient (SRC) approach (Norell and Novacek, 1992) in relation to hundreds of data sets (Hitchin and Benton, 1997). These authors concluded that they "cannot confirm Siddall's finding of a strong correlation of SCI . . . with either n or I_m " and further argued that the "SCI seems the best available statistic for measuring the fit of observed stratigraphy and estimated phylogenies, and the use of SCI should be promoted" (Hitchin and Benton, 1997:567–568, emphasis added). It is difficult to anticipate how