Molecular Phylogenetics and Evolution
Vol. 16, No. 1, July, pp. 113–126, 2000
doi:10.1006/mpev.2000.0772, available online at http://www.idealibrary.com on IDEAL®

Factors Influencing Phylogenetic Inference: A Case Study Using the Mammalian Carnivores

Olaf R. P. Bininda-Emonds[1]

Department of Zoology, University of Oxford, Oxford OX1 3PS, United Kingdom

Received August 9, 1999; revised December 21, 1999

Phylogenetic reconstruction has undergone numerous developments in tree selection criteria (e.g., phenetics, cladistics, maximum-likelihood), available data sources (morphology versus molecules, and subsets of the latter), and practical limits on study size. Together with study age, I examined the effects of these variables on inferences of phylogeny for the mammalian order Carnivora. The raw data comprised 274 source trees spread among 13 carnivore taxa (generally families), which I divided into categories for each variable and combined using the supertree technique matrix representation with parsimony analysis. Incongruence between the resultant tree topologies or the underlying data was assessed using four comparison measures, each with slightly different properties: the triplet measures "do not conflict" and "explicitly agree," the partition metric, and the incongruence length difference metric. Except for a few cases reflecting historical problem areas in carnivore systematics, no significant differences in incongruence levels were found among the different categories within each variable, between the variables themselves, or between the taxa. Thus, most estimates of carnivore phylogeny cannot be distinguished from one another (and may even point toward the same solution) regardless of the methodology or data source employed. This conclusion held regardless of the comparison measure used.

INTRODUCTION

Phylogenetic inference has undergone numerous developments. Early systematists clustered taxa based on their overall morphological similarity. Since then, and particularly over the past 40 years, phylogenetics has seen three important advances: (1) the development of rigorous clustering techniques and tree selection criteria, (2) the discovery of molecular data sources, and (3) changes in the number of taxa that can be analyzed.

Of these three advances, the most attention has been focused on the accuracy of different clustering techniques and of different data sources. Much has been written about the clash between proponents of two of the first rigorous tree selection criteria: phenetics and cladistics (see Hull, 1980 for an overview). Despite its eventual victory, cladistics (really parsimony) is now challenged by more complex model-based algorithms (e.g., maximum-likelihood, neighbor-joining). Simulation studies examining the relative efficacy of these techniques generally conclude that, although certain methods such as UPGMA or Lake's invariants perform consistently worse, most methods possess shortcomings under certain specific sets of conditions and are generally similar (Nei, 1991; Hillis and Huelsenbeck, 1993; Charleston *et al.*, 1994; Hillis *et al.*, 1994; Tateno *et al.*, 1994; Huelsenbeck, 1995; Siddall, 1998).

Development of the newer tree selection criteria is linked to the rise of molecular information as a data source. Although in use for only about 30 years, molecular data have arguably surpassed morphological data in popularity. Molecular systematics has also advanced rapidly from initial karyological and immunological studies to karyotypic banding studies to amino acid sequences and finally to DNA sequence data of various forms. Despite the perceived conflict between morphological and molecular data (Goodman, 1989; Graur, 1993; de Jong, 1998), numerous studies and reviews indicate that the two data sources do not produce substantively different answers on the whole (e.g., Hillis, 1987; Sanderson and Donoghue, 1989; Novacek, 1992; Patterson *et al.*, 1993). Proponents of the character congruence or "total evidence" approach (*sensu* Kluge, 1989) argue that the best answers are obtained when all the available data, morphological and molecular, are combined and analyzed simultaneously.

Study size can also affect accuracy. Size can refer to either the number of characters or the number of taxa in an analysis; however, I restrict myself to discussing the latter. Recently, it has been asserted that larger analyses using more complete taxon sampling yield

[1] Current address: Section of Evolution and Ecology, University of California, Davis, CA 95616. Fax: (530) 752-1449. E-mail: orbininda@ucdavis.edu.

1055-7903/00 $35.00

more reliable results (see Hillis, 1996, 1998 and references in the latter) because the judicious inclusion of as many taxa as possible avoids errors in phylogenetic inference introduced through long-branch attraction (Swofford *et al.*, 1996) or assumptions of monophyly and selection of exemplar taxa (Arnold, 1981; Donoghue *et al.*, 1989; Lecointre *et al.*, 1993; Bininda-Emonds *et al.*, 1998). Study size shows an interesting history. In early systematic studies, the simple, roughly intuitive clustering used allowed for large numbers of taxa (e.g., Gregory and Hellman, 1939). Ironically, the advent of more rigorous clustering techniques decreased study sizes initially because the more complicated calculations limited the number of taxa that could be worked with, either by hand or by early computers. It is only relatively recently that advances in computer technology combined with algorithmic shortcuts (e.g., the branch and bound algorithm: Hendy and Penny, 1982; heuristic tree search strategies: see Swofford *et al.*, 1996; parsimony jackknifing: Farris *et al.*, 1996; matrix representation with parsimony analysis: Baum, 1992; Ragan, 1992; compartmentalization: Mishler, 1994) have allowed systematic studies to become larger and more inclusive than ever before (e.g., Chase *et al.*, 1993; Purvis, 1995a; Van de Peer and de Wachter, 1997; Källersjö *et al.*, 1998; Bush *et al.*, 1999).

The question remains whether the advances listed above (which I refer to as "variables" hereafter) have changed our phylogenetic estimates to any substantial degree or in a consistent manner. Certainly, differences have been advocated by proponents of a specific methodology, but their impact has rarely been examined in a large-scale comparative or statistical framework. A recent supertree of all extant species of Carnivora (Bininda-Emonds *et al.*, 1999) is ideal in this regard. The 274 source trees used in this study span from 1970 to 1995 and reflect changes in methodology over this period. I therefore used these source trees to test the effect of the three variables above plus study age on estimates of carnivore phylogeny.

MATERIALS AND METHODS

Variables under Examination

I subdivided the 274 source trees from Bininda-Emonds *et al.* (1999) into the following categories for each of five variables:

(1) Tree selection criteria—discrete character, distance data, and "intuitive parsimony."

I followed Nei (1991) in dividing the many formal clustering techniques into two broad classes: those for discrete character data (e.g., parsimony and maximum-likelihood) and those for distance data (e.g., phenetics, UPGMA, neighbor-joining, minimum-evolution, and morphometrics). I added a third category (intuitive parsimony) for studies that derived a phylogeny from a set of data without apparent recourse to any formal clustering algorithm.

(2) Data source—morphological, molecular, and total evidence.

I classified studies according to whether they were purely morphological, were purely molecular, or used both data types simultaneously (total evidence).

(3) Molecular source—sequence, karyology, and "other."

This variable is a subclass of the previous variable. Rapid advances in molecular systematics have led to many distinct types of molecular data. Sequence includes DNA and amino acid sequence studies, restriction site analyses (e.g., restriction-fragment length polymorphisms), and microsatellite studies. Any form of chromosomal analysis falls under karyology. Other is a catchall category for techniques that were not common enough among the source studies to merit their own category (e.g., allozymes, hybridization, serology, and immunology).

(4) Size of study—small and large.

The number of terminal taxa is unsuitable as a measure of study size because studies that (implicitly) include all species can still provide limited phylogenetic information (e.g., taxonomies). Instead, a fairer representation of size is the amount of hierarchical information (i.e., number of nodes) that each study contains, although this measure is biased against studies with numerous polytomies. "Small" studies had $<50\%$ of the maximum number of informative nodes for the taxonomic group. The only exception was Hyaenidae, which with four extant species, and therefore two informative nodes at most, could not possess $<50\%$. Therefore, small studies for hyaenids were $\leq 50\%$.

(5) Date of study—1970s (and before), 1980s, and 1990s.

This variable essentially examines the effects of the other variables acting in concert. It is reminiscent of Benton and Storrs (1994), who asked whether our increased knowledge and surveying of the fossil record with time have changed the interpretations derived from it. In the present case, some correlations with the other variables are apparent. Studies from the 1970s tend to be morphological, relatively small, and analyzed using parsimony, phenetics, or intuitive parsimony. By the 1990s, studies are predominantly molecular, larger, and analyzed using parsimony, maximum-likelihood, or some algorithm for distance data. Although the 1990s encompasses only 6 years in the sample (to the end of 1995), the recent explosion in systematic effort means that sample sizes for the 1990s were roughly equal to those for each of the 1970s and 1980s (see Table 1).

TABLE 1

Distribution of Research Effort among Different Variables and Their Categories

		Date			Study size		Clustering technique			Data source			Molecular technique		
Taxon	Total[a]	1970s	1980s	1990s	Small	Large	Discrete	Distance	Intuitive	Morphology	Molecular	Both	Sequence	Karyology	Other
Higher groups (12)	62/202	15/51	21/80	26/81	44/82	18/130	37/138	19/51	12/37	20/85	36/109	6/18	19/58	9/27	17/46
Canidae (34)	36/180	15/76	7/34	16/80	34/146	3/40	16/86	7/34	12/58	15/86	19/85	4/19	16/64	4/27	4/17
Felidae (36)	40/282	5/41	22/179	15/85	36/182	5/114	19/187	20/131	6/34	13/94	24/152	5/59	11/44	9/112	10/64
Herpestidae (25)[b]	9/53	4/35	4/14	3/20	9/51	—	3/20	2/13	7/40	5/29	5/31	—	2/16	3/20	2/13
Hyaenidae (4)	6/8	3/3	3/4	2/3	4/4	3/5	5/7	—	2/2	6/8	—	—	—	—	—
Lutrinae (13)	6/37	2/13	4/25	2/6	3/7	4/34	5/28	—	3/23	5/27	—	2/14	—	—	—
Mephitinae (9)	5/18	—	2/9	4/12	3/8	3/13	4/15	2/6	—	2/5	3/10	2/9	2/7	2/9	3/12
Mustelidae (45)	30/155	5/26	18/105	7/43	27/121	2/45	10/84	11/47	10/47	12/75	15/51	3/48	5/21	7/62	8/61
Otariidae (14)	15/46	7/19	4/15	6/14	13/33	3/14	6/17	5/8	7/25	10/36	5/8	2/4	3/5	2/4	4/7
Phocidae (19)	21/120	8/39	5/32	10/55	17/71	5/52	11/78	8/22	7/34	9/68	12/44	2/14	7/31	4/18	5/23
Procyonidae (18)	7/27	—	3/15	5/18	6/19	2/14	6/25	2/8	2/8	3/21	5/12	—	4/10	2/8	—
Ursidae (8)	28/50	6/9	8/20	16/26	24/34	5/19	15/30	11/19	7/11	4/6	21/42	5/7	13/24	5/8	10/21
Viverridae (34)	9/90	3/35	6/61	2/11	7/43	3/56	2/40	3/16	6/51	7/86	3/13	—	—	2/10	2/11

Note. Presented as number of source trees/number of matrix elements. Number of terminal taxa within each group is given in parentheses.

[a] Numbers within a variable may not equal the "Total" because of the inclusion of Wozencraft (1993) to seed each analysis and because the same source tree may appear in more than one category for some variables.

[b] The actual number of herpestid species is 37; however, 12 extremely poorly known species (see Bininda-Emonds *et al.*, 1999) were always excluded from every analysis.

Constructing Supertrees

Following Bininda-Emonds *et al.* (1999), I derived supertrees for the various carnivore taxa (generally families and the subfamilies Lutrinae and Mephitinae within Mustelidae; see Fig. 1, Table 1) for each category using the supertree technique matrix representation with parsimony analysis (MRP). Briefly, MRP codes each node of a source tree in turn using additive binary coding (Farris *et al.*, 1970). If a taxon is descended from a given node, it is scored as 1; otherwise, it is scored as 0 unless it is missing from that source tree, in which case it receives ? (see Baum, 1992; Ragan, 1992). The "matrix elements" thus created are really statements of membership—is the taxon a member of this cluster?—and not characters in the usual sense. However, MRP can also be viewed as stripping the homoplasy from a source data set to leave only a single synapomorphy for each internal node on the associated source tree. To create the supertree, the matrix representations of all source trees are combined into a single matrix and analyzed using parsimony, which is the most efficient means of recovering the supertree (Baum and Ragan, 1993). Because the matrix elements are not true characters, homoplasy should merely be interpreted as incongruence between source tree representations and not as convergence or similar concepts from character evolution.

Parsimony analysis used PAUP* 4.0b2 (Swofford, 1999). I did not account for differential signal strength among source trees. This information was often not

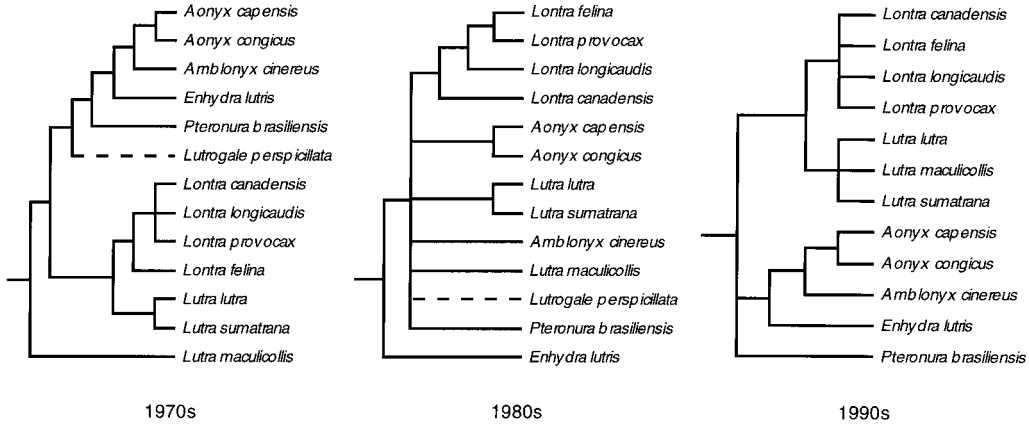

FIG. 1. Supertrees for the three categories of the variable "study date" for Lutrinae (otters). Note that taxon sets are not identical due to the use of safe taxonomic reduction to derive the supertree for the 1990s (no information existed for *Lutrogale perspcillata*). Tree topologies were therefore compared by pruning off *Lutrogale* from the other two supertrees (as indicated by dashed branches) to create a set of agreement subtrees.

provided or was given by different metrics, which may not yield equivalent information (see Bininda-Emonds and Bryant, 1998). Strict consensus was used to summarize equally most parsimonious solutions.

Although MRP is remarkable in maintaining resolution even when source trees conflict strongly (Purvis, 1995b), the subdivision that I employed often rendered the category sample sizes too small for good resolution. Frequently, a species lacked any information for a given category, making all possible positions equally parsimonious and reducing the supertree to a bush under strict consensus. I counteracted this by using Wozencraft's (1993) taxonomy to provide a backbone of limited clustering information for most species. Only analyses of interfamilial relationships ("higher groups") could not be seeded in this manner because the taxonomy did not extend beyond the family level.

In more extreme cases, I used "safe taxonomic reduction" (Wilkinson, 1995) to identify and remove taxa that limited resolution, but could not otherwise influence the tree topology because their character states were completely redundant with a more completely known taxon and so did not possess any novel clustering information (see also Bininda-Emonds *et al.*, 1999). Since tree comparison metrics (see below) can only compare trees with identical sets of terminal taxa, I created a set of agreement subtrees (see Page, 1993) within each variable by pruning the same taxon (or taxa) from the supertrees of the remaining categories (see Fig. 1). The contrasting approaches of safe taxonomic reduction versus forming agreement subtrees reflect that the strategy for obtaining the best inference depends on the amount of information in the matrix. For poorly resolved trees, the best inference is obtained when problematic taxa that contain no unique clustering information are removed and the analysis is rerun. Such is not the case with well-resolved trees within the same variable. Here, safe taxonomic reduction potentially removes nontrivial information and may alter the inferred relationships. Instead, the best inference is obtained when all taxa are analyzed and the required taxa are subsequently pruned from the supertree. Analyses involving Herpestidae always excluded at least the 12 poorly known species identified by Bininda-Emonds *et al.* (1999).

One limitation of this study was that not every source tree from the period 1970 to 1995 inclusive was sampled; those appearing in smaller or less publicized journals were more likely to be overlooked or to be unobtainable. However, other than a deliberate tendency by Bininda-Emonds *et al.* (1999) to underrepresent taxonomies (as they are summaries of primary sources rather than primary sources themselves), there is no systematic bias among taxa or categories. Therefore, the results and trends herein are likely to still be valid.

Comparison of Supertrees within Each Variable

The supertree from Bininda-Emonds *et al.* (1999) is unsuitable as a reference tree because of nonindependence: the source trees in each category all contributed to this tree. Categories with more source trees had greater input into the topology of the supertree and so will resemble it more closely than smaller categories. To avoid this problem, I instead compared all combinations of supertrees within a variable to one another (e.g., all pairwise combinations of 1970s, 1980s, and 1990s for the variable "date of study").

With this approach, it is not possible to state whether one category yields a "better" answer than another, but only whether it produces an answer significantly different from those of the remaining categories. The corresponding hypotheses for this question are:

H_O: The supertrees for the different categories within a variable are all equally different from one another.

H_A: The supertrees for one or two categories is/are significantly different from those for the remaining categories.

Note that the construction of these hypotheses is such that categories may yield trees that are very different from one another, but still not produce a significant result so long as all trees are equally different from one another.

Two subsidiary questions examine how any differences are distributed among the higher carnivore taxa and among the variables themselves. Are differences among the categories larger or smaller in some taxonomic groups (within each variable) or some variables? For instance, supertrees constructed for the different tree selection criteria may be more incongruent than those from different data sources. In other words, the choice of tree selection criteria may have a larger impact on the inferred phylogeny than does the data source used. Similarly, some carnivore taxa may be more difficult to reconstruct for a given variable (e.g., felid supertrees may show more dependence on the data source used) than others are. Again, the corresponding hypotheses are:

H_O: The differences between the supertrees for the categories are equal among all taxonomic groups (within each variable) or among all variables.

H_A: The supertrees for one or more taxonomic groups/variables have values significantly different from those for the remaining taxonomic groups/variables.

Comparison Metrics

Differences between the two or three supertrees within a variable were quantified using four metrics with slightly different properties. The first three met-

rics compared tree topologies only: the partition metric (d_S; Robinson and Foulds, 1981) and two metrics from triplet analysis (Estabrook *et al.*, 1985; Day, 1986). The partition metric describes the number of clades found in one tree or the other, but not both. As such, it treats polytomies as being real events ("hard"; see Maddison, 1989) and can indicate a large difference between two trees even if they differ in the placement of only a single taxon (Page, 1993).

Triplet analysis determines whether each possible pair of triplets between two trees has the same topology, different topologies, or whether either one or both are unresolved. Of the many metrics available to summarize triplet analysis (see Day, 1986), I used only the dissimilarity metrics "do not conflict" (DC), which gives the proportion of resolved triplets that are different, and "explicitly agree" (EA), which adds triplets that are unresolved in either or both trees to this value (Estabrook *et al.*, 1985).

A potential shortcoming of tree comparison metrics is their independence from the data underlying the topologies. Two categories yielding similar supertree topologies might be highly incongruent at the level of the "raw data" (i.e., membership statements in the source trees in this study). I tested for data incongruence using the incongruence length difference metric (D_{XY}; *sensu* Farris *et al.*, 1994) of Mickevich and Farris (1981). D_{XY} quantifies the amount of incongruence in a combined analysis of two test matrices that is over and above ("extra to") the sum of the incongruences within each test matrix. In all cases, incongruence is defined as the length of the most parsimonious solution(s) minus the ideal, minimal length were there no homoplasy present (i.e., the number of additional steps over the ideal, minimal length). D_{XY} is then the "extra" incongruence present in the combined matrix divided by the total incongruence of the combined matrix (see Kluge, 1989 for a worked example). A permutation test based on D_{XY} exists to determine whether two test matrices display significant data incongruence (Farris *et al.*, 1994); it is implemented in PAUP* as the partition homogeneity test.

D_{XY} can also compare solutions where safe taxonomic reduction has yielded different sets of terminal taxa. This is because the removed taxa are "redundant" with other taxa in the matrix (see above) and so do not affect tree length, number of "synapomorphies," or consistency index, the latter being a measure of incongruence (Wilkinson, 1995). Taxa were removed from the combined analysis if they were removed from both individual analyses.

I used COMPONENT (Page, 1993) to obtain values of d_S, DC, and EA and PAUP* to obtain most parsimonious lengths and number of "synapomorphies" for calculating D_{XY} and to perform the partition homogeneity test. For all metrics, higher values indicate increasingly different solutions. Both triplet measures and D_{XY} are bounded by 0 and 1. I standardized d_S by dividing it by its maximum value of $2n - 6$, where n = number of taxa (Steel and Penny, 1993). My use of Wozencraft's (1993) taxonomy to seed each analysis will cause all measures of incongruence to be slightly lower than they should be. However, because this is true for all analyses and because I am making comparative rather than absolute statements, this should not be a problem.

I tested each of the three sets of hypotheses using nonparametric Kruskal–Wallis tests because the data often were not normally distributed, nor possessed equal variances. The critical value to reject each null hypothesis was 0.05, corrected for multiple comparisons (i.e., the different metrics) within each variable using a sequential Bonferroni technique (Rice, 1989).

Sliding Window Analysis

In addition to the preceding statistical framework, I examined how results regarding two long-standing carnivore systematic questions have changed over time. These questions are the affinity of pinnipeds (including whether they are mono- or diphyletic) and that of the red panda, *Ailurus fulgens* (see Bininda-Emonds and Russell, 1996 and Pecon Slattery and O'Brien, 1995, respectively, for reviews). Although the giant panda, *Ailuropoda melanoleuca*, was a subject of similar uncertainty historically, its status as a primitive ursid (see O'Brien *et al.*, 1985) is now largely unquestioned. Also, because Bininda-Emonds *et al.* (1999) constrained the giant panda to be an ursid because of this, alternative placements for it did not exist among the source trees.

To examine these changes in opinion over time, I used a "sliding window" form of time-series analysis. I ordered the 62 source trees that provided any clustering information about the carnivore families (higher groups) in ascending chronological order and secondarily by author name in ascending alphabetical order. Supertrees were determined for contiguous, overlapping sets of 15 trees (e.g., source trees 1–15, 2–16, 3–17, and so on). This number of source trees provided sufficient clustering information to avoid the use of safe taxonomic reduction.

For each supertree from a window of 15 source trees, the sister groups for pinnipeds (ursid, mustelid, both, neither, or not applicable) and red panda (ursid, procyonid, both, or neither) were determined, as well as whether the pinnipeds were monophyletic or not. I also compared the supertree from each window to the supertree of Bininda-Emonds *et al.* (1999) using the above metrics. Although the issue of nonindependence remains, it is less problematic here because all windows are about equal in size and therefore nonindependent to roughly the same degree. D_{XY} could not be applied because the matrix for the supertree contains all the source trees (and their incongruences) for a

TABLE 2

Examination of Whether Categories within a Variable Are Producing Significantly Different Inferences of Phylogeny

Comparison	Comparison metric			
	DC	EA	d_s	D_{XY}
Tree selection criterion				
Discrete–distance	0.056	0.470	0.294	0.136
Discrete–intuitive	0.072	0.304	0.248	0.110
Distance–intuitive	0.022	0.475	0.219	0.103
SE	0.012–0.036	0.060–0.090	0.031–0.046	0.029–0.044
H (*df* = 2)	2.196	2.957	2.994	0.210
P value	0.3335	0.2280	0.2238	0.9002
Data source				
Morphology–molecular	0.057	0.423	0.311	0.130
Morphology–both	0.071	0.386	0.291	0.158
Molecular–both	0.058	0.389	0.275	0.143
SE	0.029–0.031	0.059–0.078	0.038–0.045	0.030–0.058
H (*df* = 2)	0.320	0.168	0.210	0.039
P value	0.8628	0.9194	0.9005	0.9805
Molecular source				
Sequence–karyology	0.035	0.494	0.273	0.225
Sequence–other	0.007	0.429	0.228	0.223
Karyology–other	0.013	0.384	0.165	0.082
SE	0.007–0.024	0.058–0.094	0.042–0.058	0.043–0.059
H (*df* = 2)	1.662	1.569	2.119	4.819
P value	0.4356	0.4563	0.3466	0.0899
Study size				
Small–large	0.063	0.328	0.308	0.119
SE	0.032	0.032	0.042	0.031
H (*df* = 1)	n/a: insufficient comparisons			
Date				
1970s–1980s	0.061	0.441	0.372	0.190
1970s–1990s	0.069	0.374	0.258	0.178
1980s–1990s	0.104	0.418	0.386	0.164
SE	0.026–0.042	0.049–0.084	0.047–0.072	0.031–0.041
H (*df* = 2)	0.286	0.600	1.859	0.287
P value	0.8668	0.7407	0.3947	0.8663

Note. Presented as mean values, range of standard errors (SE), and results of a Kruskal–Wallis test (*H* corrected for ties; *df* = degrees of freedom). None of the differences were significant at the 0.05 level (corrected for multiple comparisons).

given window and no extra incongruence is possible (i.e., D_{XY} would always be 0). Instead, I quantified data incongruence using PAUP* to constrain the solution for each window to the topology of the supertree. I then divided the number of extra steps that this required by the length of the most parsimonious solution for that window to derive a simple "incongruence metric." Note that this metric does not have an upper bound of 1 as do the other metrics.

RESULTS

Distribution of Previous Research Effort

Numbers of source trees and matrix elements ("characters"; a proxy for average tree size) for each taxon are presented in Table 1. Most taxa had at least one source tree in every category. Two notable exceptions were hyaenids and lutrines, which lacked any purely molecular source trees. The amount of phylogenetic effort directed among carnivore taxa was discussed by Bininda-Emonds *et al.* (1999): interfamilial relationships of carnivores (higher groups), canids, felids, mustelids, phocids, and ursids are comparatively well studied (i.e., high numbers of source trees), while herpestids, hyaenids, lutrines, mephitines, procyonids, and viverrids are poorly researched.

Trends are also apparent within each variable. For tree selection criteria, studies using discrete characters were normally the most common for each taxon. This largely reflects both the long history of use of parsimony and its simplicity. Parsimony is the only criterion to span all 25 years examined in this study and was largely unchallenged for a time after the demise of phenetics. Despite the large number of techniques currently available, parsimony remains popular, partly due to its ease of calculation compared to the newer, more computationally intensive techniques. The distribution of research effort among the remaining catego-

TABLE 3

Examination of Whether Differences between Supertrees for Various Categories Are the Same across All Variables

Variable	Comparison metric			
	DC	EA	d_s	D_{XY}
Tree selection criterion	0.051	0.416	0.256	0.117
Data source	0.062	0.402	0.294	0.143
Molecular source	0.019	0.436	0.222	0.175
Study size	0.063	0.328	0.308	0.119
Date	0.079	0.411	0.341	0.176
SE	0.009–0.032	0.035–0.044	0.023–0.042	0.021–0.031
H ($df = 4$)	9.966	1.821	9.054	4.696
P value	0.0410	0.7686	0.0598	0.3200

Note. Presented as mean values, range of standard errors (SE), and results of a Kruskal–Wallis test (H corrected for ties; df = degrees of freedom). There was no significant difference among variables at the 0.05 level when corrected for multiple comparisons.

ries depended on how well studied a taxon is. Well-known taxa tended to be analyzed by the more rigorous distance data criteria, while intuitive parsimony tended to predominate in studies of poorly known taxa.

Molecular data generally yielded more source trees for each taxon than morphological data, but produced fewer matrix elements. Thus, although molecular data are more popular, they are available for fewer species. This appears to have two root causes. First, the greater research effort is replicated among the different molecular data types for the same small group of well-known species. Second, molecular data have until recently been relatively expensive and difficult to obtain, again limiting their accumulation for most species. The various molecular data sources are all about equally well represented, although sequence data may be slightly more numerous. Total evidence studies were generally few in number and small in size.

There were fewer trees from large studies, but their larger size compared to trees from small studies meant that the number of matrix elements was about equal in each. Finally, study date showed increased numbers of trees and matrix elements with time. In fact, the 1990s was often the largest category, despite not including any source trees from after 1996.

Statistical Comparisons

An insufficient number of comparisons ($df = 1$) prevented study size from being analyzed; otherwise, no significant differences were found for the remaining within-variable comparisons (Table 2). This indicates that the "raw data" or resultant supertrees for each category within a variable are all equally different. This was true for all variables and comparison metrics. In only one case (molecular technique measured by D_{XY}) did the P value drop below 0.10. In general, the metrics showed good agreement with one another, as indicated by their similar P values. (As the metrics measure incongruence slightly differently [see "Discussion"], their average values need not be comparable or mean the same thing.)

Also, the variables did not differ significantly among themselves with respect to the level of discordance exhibited by their respective categories (Table 3). However, the metrics were less harmonious on this point. Both EA and the D_{XY} clearly failed to reject the null hypothesis of no differences between the variables, while DC and d_S indicated results approaching significance (uncorrected $P < 0.10$).

Similarly, essentially no significant differences existed among the carnivore taxa (Table 4). Although all variables except study size had one or more metrics indicating results bordering on significance (P values were generally below 0.15 and often below 0.10), the only significant difference obtained when correcting for multiple comparisons was for study date as measured by EA. For this one case, the mean rank from the Kruskal–Wallis test for felids was noticeably higher than those of the remaining taxa, hinting that inferences of felid phylogeny have differed more over time than those of other carnivore taxa. In contrast, higher groups and hyaenids had noticeably lower values, which indicate more stable phylogenetic inferences over time.

The partition homogeneity test revealed significant data incongruence between categories for only a few of the larger taxa (Table 4). Canids displayed differences for all five variables, felids, higher groups, and mustelids for three variables, and viverrids for study size only. In most of these 15 cases, all categories within a variable were significantly incongruent from one another. Although it is easier to detect incongruence for larger matrices, many of the smaller taxa displayed low P values for the partition homogeneity test as well (e.g., higher groups and lutrines).

For Tables 2–4, EA generally had the highest values, followed by D_{XY}, d_S, and DC.

TABLE 4

Examination of Whether Differences between Supertrees for Various Categories Are the Same across All Carnivoran Taxa for a Given Variable

Taxon	Comparison metric			
	DC	EA	d_s	D_{XY}
Tree selection criterion				
Canidae	0.008	0.566	0.290	0.320 (0.001*)[a]
Felidae	0.144	0.708	0.323	0.087 (0.197)
Herpestidae	0.000	0.508	0.049	0.074 (0.939)
Higher groups	0.048	0.097	0.333	0.049 (0.166)
Hyaenidae	0.000	0.250	0.000	0.000 (1.000)
Lutrinae	0.381	0.399	0.550	0.167 (0.522)
Mephitinae	0.000	0.357	0.250	0.000 (1.000)
Mustelidae	0.179	0.447	0.246	0.190 (0.081)
Otariidae	0.009	0.742	0.303	0.067 (0.768)
Phocidae	0.001	0.390	0.250	0.052 (0.897)
Procyonidae	0.000	0.340	0.067	0.037 (0.855)
Ursidae	0.024	0.178	0.333	0.093 (0.540)
Viverridae	0.022	0.265	0.355	0.263 (0.043)
SE	0.000–0.092	0.000–0.258	0.019–0.057	0.012–0.131
H (df = 12)	19.994	22.729	21.260	18.336
P value	0.0672	0.0301	0.0467	0.1059
Data source				
Canidae	0.082	0.481	0.258	0.236 (0.001*)[a]
Felidae	0.082	0.690	0.394	0.148 (0.276)
Herpestidae	0.082	0.257	0.191	0.278 (0.080)
Higher groups	0.077	0.093	0.333	0.069 (0.012*)[b]
Hyaenidae	n/a	n/a	n/a	n/a
Lutrinae	0.000	0.350	0.150	0.517 (0.176)
Mephitinae	0.079	0.357	0.222	0.111 (0.438)
Mustelidae	0.177	0.568	0.381	0.251 (0.001*)[c]
Otariidae	0.037	0.434	0.242	0.147 (0.470)
Phocidae	0.000	0.142	0.167	0.033 (0.918)
Procyonidae	0.000	0.518	0.133	0.125 (0.409)
Ursidae	0.012	0.488	0.467	0.033 (0.550)
Viverridae	0.000	0.361	0.371	0.000 (0.993)
SE	0.000–0.101	0.000–0.121	0.015–0.085	0.002–0.111
H (df = 11)	14.135	22.387	18.498	18.535
P value	0.2160	0.0215	0.0707	0.0700
Molecular source				
Canidae	0.001	0.298	0.226	0.308 (0.001*)[d]
Felidae	0.046	0.767	0.374	0.240 (0.001*)[a]
Herpestidae	0.000	0.520	0.059	0.229 (0.126)
Higher groups	0.052	0.512	0.519	0.210 (0.001*)[a]
Hyaenidae	n/a	n/a	n/a	n/a
Lutrinae	n/a	n/a	n/a	n/a
Mephitinae	0.079	0.321	0.222	0.111 (0.839)
Mustelidae	0.004	0.706	0.270	0.208 (0.200)
Otariidae	0.000	0.419	0.061	0.333 (0.065)
Phocidae	0.000	0.229	0.125	0.056 (0.889)
Procyonidae	0.000	0.266	0.067	0.000 (1.000)
Ursidae	0.000	0.232	0.267	0.000 (1.000)
Viverridae	0.000	0.361	0.000	0.000 (0.995)
SE	0.000–0.079	0.006–0.197	0.029–0.088	0.000–0.167
H (df = 10)	19.308	15.893	21.683	12.380
P value	0.0365	0.1027	0.0168	0.2604
Study size				
Canidae	0.095	0.229	0.468	0.155 (0.012*)[a]
Felidae	0.393	0.413	0.470	0.118 (0.001*)[a]
Herpestidae	n/a	n/a	n/a	n/a
Higher groups	0.023	0.118	0.333	0.104 (0.002*)[a]
Hyaenidae	0.000	0.250	0.000	0.000 (1.000)
Lutrinae	0.064	0.432	0.350	0.053 (1.000)
Mephitinae	0.000	0.357	0.250	0.000 (1.000)
Mustelidae	0.098	0.495	0.405	0.232 (0.003*)[a]

TABLE 4—*Continued*

Taxon	Comparison metric			
	DC	EA	d_s	D_{XY}
Study size				
Otariidae	0.003	0.266	0.182	0.250 (0.226)
Phocidae	0.011	0.315	0.281	0.083 (0.648)
Procyonidae	0.000	0.518	0.133	0.000 (1.000)
Ursidae	0.036	0.214	0.400	0.100 (0.398)
Viverridae	0.029	0.333	0.419	0.333 (0.002*)[a]
SE	n/a	n/a	n/a	n/a
H (*df* = 11)	11.000	11.000	11.000	11.000
P value	0.4433	0.4433	0.4433	0.4433
Date				
Canidae	0.019	0.574	0.355	0.218 (0.001*)[e]
Felidae	0.041	0.784	0.475	0.191 (0.001*)[a]
Herpestidae	0.000	0.559	0.069	0.247 (0.190)
Higher groups	0.087	0.116	0.352	0.080 (0.078)
Hyaenidae	0.167	0.167	0.667	0.222 (0.264)
Lutrinae	0.155	0.530	0.400	0.269 (0.033)
Mephitinae	0.238	0.357	0.333	0.333 (0.251)
Mustelidae	0.291	0.583	0.341	0.305 (0.001*)[f]
Otariidae	0.037	0.275	0.182	0.120 (0.653)
Phocidae	0.023	0.261	0.354	0.146 (0.371)
Procyonidae	0.000	0.518	0.100	0.000 (1.000)
Ursidae	0.012	0.309	0.333	0.097 (0.597)
Viverridae	0.015	0.347	0.312	0.051 (0.843)
SE	0.000–0.124	0.000–0.153	0.000–0.333	0.016–0.111
H (*df* = 12)	21.520	26.097	15.664	17.239
P value	0.0433	0.0104*	0.2071	0.1408

Note. Presented as mean values (*n* = the 3 categories, 2 for "size of study") with ranges of standard errors (SE), and results of a Kruskal–Wallis test (*H* corrected for ties; *df* = degrees of freedom). *P* values for a partition homogeneity test follow D_{XY} in parentheses. An asterisk indicates a significant difference among taxa (or categories for the partition homogeneity test) at the 0.05 level (corrected for multiple comparisons).

[a] All categories differ significantly.
[b] Morphology and molecules differ significantly.
[c] Morphology differs significantly from remaining categories.
[d] Sequence and karyology differ significantly.
[e] 1970s and 1990s differ significantly.
[f] 1980s differs significantly from remaining categories.

Sliding Window Analysis

Both portions of the sliding window analysis revealed changes over time in inferences of relationships among the higher groups. The different metrics (Fig. 2) show that inferences from the 1980s resembled the supertree of Bininda-Emonds *et al.* (1999) to the greatest degree. This could be because the 1980s had the largest number of source trees or because these source trees were highly congruent (for any of a number of possible reasons) and therefore provided a more concerted signal to influence the overall supertree. Adding studies from either the 1970s or the 1990s increased disagreement with the supertree to about equal levels. It is not possible to discern whether the corresponding changes in topology are similar, nor to qualify the importance of the differences in Fig. 2. However, the previous section showed that higher groups possessed a slightly lower level of conflict for study date compared to other carnivore taxa and one that was not significant as measured by the partition homogeneity test (Table 4).

As in the previous section, the metrics show good agreement and track each other relatively closely. This is true despite d_S being on a scale different from those of the remaining metrics. It consistently possessed the highest values (unlike in the previous section), rarely dropping below the maximum of 0.170 obtained by DC, EA, and the simple incongruence metric (despite the latter being unbounded above). Values for DC were always less than or equal to those of EA, while the incongruence metric could be higher or lower than either.

The inferred sister groups of pinnipeds and red panda (Fig. 3) also display clear time effects. Both taxa are inferred to have ursid affinities in the 1990s, but mustelid affinities in the 1980s. No pattern emerges from the 1970s, but, due to the distribution of studies, there is no extended period when studies from this decade dominate.

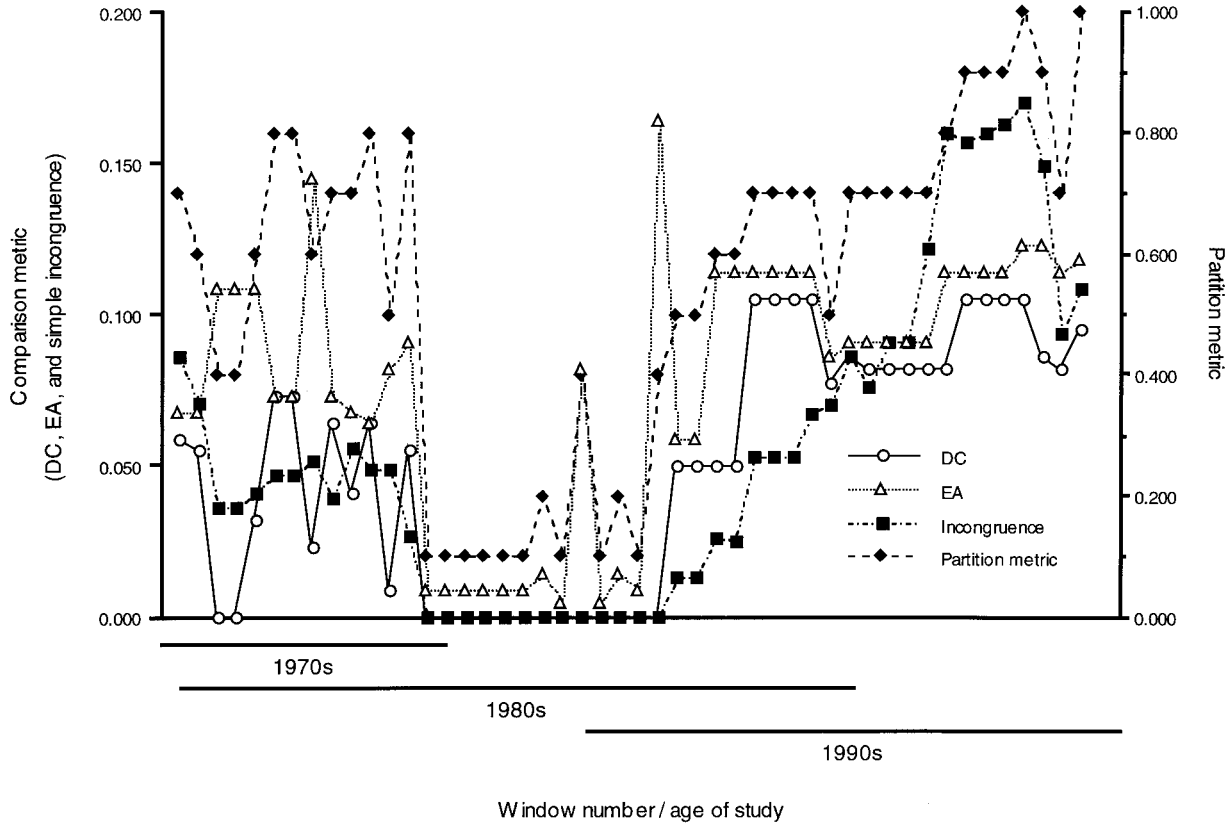

FIG. 2. Sliding window analysis examining the changes in higher level carnivore phylogeny over time using four comparison metrics. For each window of 15 source trees, supertrees were determined and compared to the supertree for higher groups in Bininda-Emonds *et al.* (1999; their Fig. 1).

Interestingly, a majority of windows (24 versus 14 for the next largest group) inferred that the red panda has an ursid sister group, despite the overall supertree placing it as the sister group to musteloids (mustelids and procyonids). Bininda-Emonds *et al.* (1999) noted a similar discrepancy: although more studies indicated a red panda–ursid pairing, such a topology was not globally most parsimonious. However, windows promoting procyonid affinity (1980s) spanned the same period when resemblance to the supertree is the greatest (Fig. 2). The inferred sister group of pinnipeds was split about equally between mustelids and ursids. Most windows indicate a monophyletic Pinnipedia; only 6 windows at the interchange of the 1970s and 1980s indicated polyphyly. This accords well with the time of greatest popularity of the diphyly hypothesis.

DISCUSSION

A consistent picture emerged whereby no large-scale differences between the supertrees for the different categories within a variable nor in the level of discordance between categories for the variables themselves or between different taxa within each variable were found. The comparatively few significant differences detected within certain taxa by the partition homogeneity test correspond to specific long-standing problem areas in carnivore systematics. Overall, the lack of significant differences suggests that we cannot distinguish between most estimates of carnivore phylogeny. The absolute differences that do exist could therefore be interpreted as random variation around some measure of central tendency (which is hopefully the true tree). Similarly, no single advancement in phylogenetic methodology (i.e., any one variable) stood out as being more influential than the rest on our phylogenetic inferences. Although these results are strictly applicable to carnivores only, they are encouraging, particularly for "total evidence" and supertree construction (*sensu* Sanderson *et al.*, 1998) approaches to combining phylogenetic information. Both approaches appear able to give larger and/or more robust phylogenies, but implicitly depend on homogeneity among the different data sets or source trees (see Bull *et al.*, 1993; Huelsenbeck *et al.*, 1996).

Much has been written concerning the relative merits or deficiencies of different tree selection criteria (in particular, cladistics versus any other criterion) or data sources. For the former, evidence has been marshaled based on philosophical discussions of method-

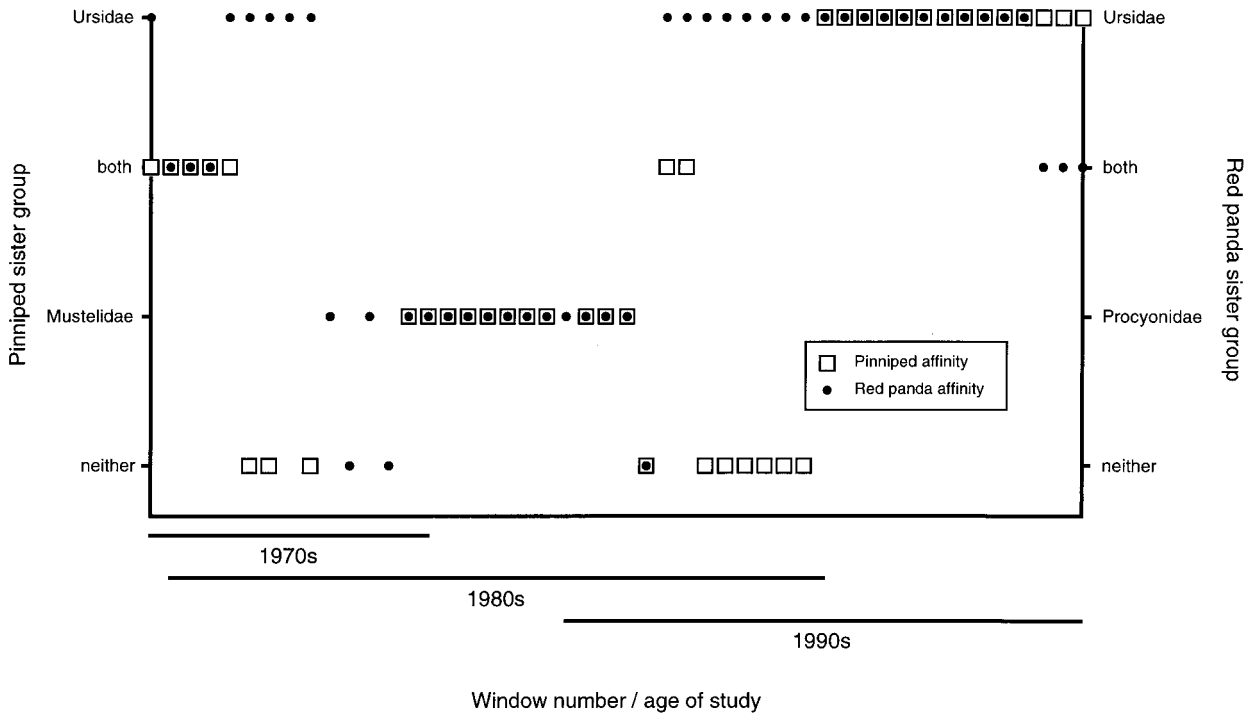

FIG. 3. Sliding window analysis examining changes in the inferred sister groups of pinnipeds and the red panda (*Ailurus fulgens*) over time. Each window represents a supertree of 15 source trees. Pinniped polyphyly is indicated by the lack of an open square for that window.

ological validity (e.g., Cracraft and Helm-Bychowski, 1991; Kluge and Wolf, 1993) or on outcomes of experimental studies involving a known phylogeny or simulations (e.g., Nei, 1991; Hillis *et al.*, 1992, 1994; Hillis and Huelsenbeck, 1993; Tateno *et al.*, 1994; Huelsenbeck, 1995; Siddall, 1998). Although the conclusions so reached may be valid under the specific conditions in which they were obtained, my results indicate that their practical manifestations are not easily discernable (see also Charleston *et al.*, 1994). In particular, tree selection criterion displayed the fewest within-taxon significant differences as measured by the partition homogeneity test (Table 4).

For the latter "conflict" between molecules and morphology, my results reinforce the underpublicized statement that morphological and molecular data give largely concordant results, or at least not any more incongruent than studies within each data type (Hillis, 1987; Patterson *et al.*, 1993; Smith and Littlewood, 1994). The similar metric values for data source and molecular source in Tables 3 and 4 provide strong evidence for this latter assertion, as do the similar number of significant differences from the partition homogeneity test (Table 4). Instead, differences between morphological and molecular data, when they occur, often result from different sets of assumptions or methods of analysis (Hillis, 1987). Even the celebrated example of this conflict within carnivores, that of pinniped monophyly versus diphyly, has recently come to a similar conclusion. Historically, most morphological and paleontological studies of pinniped evolution favor a diphyletic origin of the group (e.g., Flower, 1869; McLaren, 1960; Tedford, 1976; de Muizon, 1982), whereas most molecular studies indicate a monophyletic origin (e.g., Sarich, 1969a,b; Arnason, 1974, 1977; de Jong, 1982; Wayne *et al.*, 1989). However, much of the conflict derived not from the different data types *per se*, but from questions about character applicability and inclusion for the morphological data (see Wyss, 1988, 1989; and the exchange between Repenning [Repenning, 1990] and Berta and Wyss [Berta and Wyss, 1990] in particular). Many characters potentially uniting all pinnipeds were often excluded *a priori* because they were held to be under similar functional constraints due to being aquatic adaptations. The resolution of this issue has seen numerous morphological studies supporting a monophyletic Pinnipedia (e.g., Wyss, 1987; Flynn *et al.*, 1988; Wyss and Flynn, 1993; Berta and Wyss, 1994; Bininda-Emonds and Russell, 1996).

The lack of an absolute reference tree tempers any statement as to whether some categories are "better" than others are; however, the analyses provide indirect evidence that weakly reject such statements. For variables with three categories, arguments can be constructed as to why one category might be better or worse at estimating the one true tree. For instance, the more rigorous tree selection criteria for discrete and

distance data should yield better estimates of the true tree than intuitive parsimony. The same might be said of source trees from the 1990s, given the advancements in phylogenetic methodology and theory. Therefore, supertrees for discrete and distance data criteria should resemble each other more closely than either does with that from intuitive parsimony, and those from the 1970s and 1980s should be more dissimilar than either would be with that of the 1990s (as each contain at least some signal and thus should resemble the more accurate 1990s tree to a greater extent). However, these hypotheses are not borne out. Table 2 reveals that the discrete–distance comparison was never the lowest of the three for any metric and the 1970s–1980s comparison showed no consistent pattern. The more rigorous partition homogeneity test also failed to support these hypotheses (see Table 4).

The comparison metrics used herein showed remarkable concordance despite their different properties. Methodologically, the metrics could be differentiated by the level at which they measure incongruence (i.e., "raw data" or tree topologies). One might expect data comparison metrics (D_{XY} and the simple incongruence metric) to produce answers different from those of tree comparison metrics (DC, EA, and d_S), given that no linear correspondence has been established between patterns of data set incongruence and patterns of differences in the resultant topologies. Highly congruent data sets should produce similar trees, but the reverse does not automatically follow. Depending on the amount, direction, and distribution of conflict, incongruent data sets could still yield similar trees (e.g., if the incongruence was relatively rare, randomly distributed, or confined to homoplastic characters). Even large amounts of incongruence could be tolerated so long as the signal-producing characters within the data sets were congruent.

Although this study does not rigorously examine the point, it appears that data set incongruence does translate monotonically to differences in tree topologies. Both D_{XY} and the simple incongruence metric track the tree comparison metrics relatively closely through a wide range of apparent incongruence (Tables 2–4, Fig. 2). Thus, tree comparison metrics may provide a quick estimate of the incongruence between partitions of raw data, a finding with important practical implications given that they are generally more accessible than data comparison metrics. However, this inference requires confirmation and is tempered by the directionless aspect of all four metrics: they merely indicate the magnitude of the difference, not the direction in which it lies. Thus, two solutions can be equally different compared to some reference point, but very different from one another (within the limits of the triangle inequality). Similarly, it does not follow that two data sets that are incongruent in a certain manner will produce trees that differ from one another along an analogous direction. Unfortunately, the directionless nature of both tree space and the incongruence between data sets makes it unlikely that a comparison metric will ever be developed that incorporates directional information.

ACKNOWLEDGMENTS

I thank Andy Purvis for many stimulating early discussions. Paul Harvey, Colleen Kelly, Rod Page, Robert Scotland, two anonymous reviewers, and especially Mike Charleston and Mike Sanderson provided many helpful suggestions on how to improve this paper. Financial support was provided by Alberta Heritage, the United Kingdom's Overseas Research Scholarship plan, and an NSERC postgraduate scholarship.

REFERENCES

Arnason, U. (1974). Comparative chromosomal studies in Pinnipedia. *Hereditas* **76:** 179–226.

Arnason, U. (1977). The relationship between the four principal pinniped karyotypes. *Hereditas* **87:** 227–242.

Arnold, E. N. (1981). Estimating phylogenies at low taxonomic levels. *Z. Zool. Syst. Evolutionsforsch.* **19:** 1–35.

Baum, B. R. (1992). Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41:** 3–10.

Baum, B. R., and Ragan, M. A. (1993). Reply to A. G. Rodrigo's "A comment on Baum's method for combining phylogenetic trees." *Taxon* **42:** 637–640.

Benton, M. J., and Storrs, G. W. (1994). Testing the quality of the fossil record: Paleontological knowledge is improving. *Geology* **22:** 111–114.

Berta, A., and Wyss, A. R. (1990). Response to "Oldest pinniped." *Science* **248:** 499–500.

Berta, A., and Wyss, A. R. (1994). Pinniped phylogeny. *In* "Contributions in Marine Mammal Paleontology Honoring Frank C. Whitmore, Jr.," (A. Berta and T. A. Deméré, Eds.), Vol. 29, pp. 33–56. Proceedings of the San Diego Society of Natural History, San Diego.

Bininda-Emonds, O. R. P., Bryant, H. B., and Russell, A. P. (1998). Supraspecific taxa as terminals in cladistic analysis: Implicit assumptions of monophyly and a comparison of methods. *Biol. J. Linn. Soc.* **64:** 101–133.

Bininda-Emonds, O. R. P., and Bryant, H. N. (1998). Properties of matrix representation with parsimony analyses. *Syst. Biol.* **47:** 497–508.

Bininda-Emonds, O. R. P., Gittleman, J. L., and Purvis, A. (1999). Building large trees by combining phylogenetic information: A complete phylogeny of the extant Carnivora (Mammalia). *Biol. Rev.* **74:** 143–175.

Bininda-Emonds, O. R. P., and Russell, A. P. (1996). A morphological perspective on the phylogenetic relationships of the extant phocid seals (Mammalia: Carnivora: Phocidae). *Bonn. Zool. Monogr.* **41:** 1–256.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., and Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42:** 384–397.

Bush, R. M., Fitch, W. M., Bender, C. A., and Cox, N. J. (1999). Positive selection on the H3 hemagglutinin gene of human influenza virus A. *Mol. Biol. Evol.* **16:** 1457–1465.

Charleston, M. A., Hendy, M. D., and Penny, D. (1994). Effects of sequence length, tree topology and number of taxa on the performance of phylogenetic methods. *Comput. Biol.* **1:** 134–151.

Chase, M. W., Soltis, D. E., Olmstead, R. G., Morgan, D., Les, D. H., Mishler, B. D., Duvall, M. R., Price, R. A., Hills, H. G., Qiu, Y. L., Kron, K. A., Rettig, J. H., Conti, E., Palmer, J. D., Manhart, J. R., Sytsma, K. J., Michaels, H. J., Kress, W. J., Karol, K. G., Clark, W. D., Hedren, M., Gaut, B. S., Jansen, R. K., Kim, K. J., Wimpee, C. F., Smith, J. F., Furnier, G. R., Strauss, S. H., Xiang, Q. Y., Plunkett, G. M., Soltis, P. S., Swensen, S. M., Williams, S. E., Gadek, P. A., Quinn, C. J., Eguiarte, L. E., Golenberg, E., Learn, G. H., Graham, S. W., Barrett, S. C. H., Dayanandan, S., and Albert, V. A. (1993). Phylogenetics of seed plants: An analysis of nucleotide sequences from the plastid gene *rbc*L. *Ann. Mo. Bot. Gard.* **80:** 528–580.

Cracraft, J., and Helm-Bychowski, K. (1991). Parsimony and phylogenetic inference using DNA sequences: Some methodological strategies. *In* "Phylogenetic Analysis of DNA Sequences" (M. M. Miyamoto and J. Cracraft, Eds.), pp. 184–220. Oxford Univ. Press, Oxford.

Day, W. H. E. (1986). Analysis of quartet dissimilarity measures between undirected phylogenetic trees. *Syst. Zool.* **35:** 325–333.

de Jong, W. W. (1982). Eye lens proteins and vertebrate phylogeny. *In* "Macromolecular Sequences in Systematics and Evolutionary Biology" (M. Goodman, Ed.), pp. 75–114. Plenum, New York.

de Jong, W. W. (1998). Molecules remodel the mammalian tree. *Trends Ecol. Evol.* **13:** 270–275.

de Muizon, C. (1982). Phocid phylogeny and dispersal. *Ann. S. Afr. Mus.* **89:** 175–213.

Donoghue, M. J., Doyle, J. A., Gauthier, J., Kluge, A. G., and Rowe, T. (1989). The importance of fossils in phylogeny reconstruction. *Annu. Rev. Ecol. Syst.* **20:** 431–460.

Estabrook, G. F., McMorris, F. R., and Meacham, C. A. (1985). Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. *Syst. Zool.* **34:** 193–200.

Farris, J. S., Albert, V. A., Källersjö, M., Lipscomb, D., and Kluge, A. G. (1996). Parsimony jackknifing outperforms neighbor-joining. *Cladistics* **12:** 99–124.

Farris, J. S., Källersjö, M., Kluge, A. G., and Bult, C. (1994). Testing significance of incongruence. *Cladistics* **10:** 315–319.

Farris, J. S., Kluge, A. G., and Eckhardt, M. J. (1970). A numerical approach to phylogenetic systematics. *Syst. Zool.* **19:** 172–191.

Flower, W. H. (1869). On the value of characters of the base of the cranium in the classification of the order Carnivora, and on the systematic position of *Bassaris* and other disputed forms. *Proc. Zool. Soc. Lond.* **1869:** 4–37.

Flynn, J. J., Neff, N. A., and Tedford, R. H. (1988). Phylogeny of the Carnivora. *In* "The Phylogeny and Classification of the Tetrapods." Vol. 2—"Mammals" (M. J. Benton, Ed.), pp. 73–116. Clarendon, Oxford.

Goodman, M. (1989). Emerging alliance of phylogenetic systematics and molecular biology: A new age of exploration. *In* "The Hierarchy of Life" (B. Fernholm, K. Bremer, and H. Jörnvall, Eds.), pp. 43–61. Excerpta Medica, Amsterdam.

Graur, D. (1993). Molecular phylogeny and the higher classification of eutherian mammals. *Trends Ecol. Evol.* **8:** 141–147.

Gregory, W. K., and Hellman, M. (1939). On the evolution and major classification of the civets (Viverridae) and allied fossil and recent Carnivora: A phylogenetic study of the skull and dentition. *Proc. Am. Philos. Soc.* **81:** 309–392.

Hendy, M. D., and Penny, D. (1982). Branch and bound algorithms to determine minimal evolutionary trees. *Math. Biosci.* **59:** 277–290.

Hillis, D. M. (1987). Molecular versus morphological approaches to systematics. *Annu. Rev. Ecol. Syst.* **18:** 23–42.

Hillis, D. M. (1996). Inferring complex phylogenies. *Nature* **383:** 130–131.

Hillis, D. M. (1998). Taxonomic sampling, phylogenetic accuracy, and investigator bias. *Syst. Biol.* **47:** 3–8.

Hillis, D. M., Bull, J. J., White, E. M., Badgett, M. R., and Molineux, I. J. (1992). Experimental phylogenetics: Generation of a known phylogeny. *Science* **255:** 589–592.

Hillis, D. M., and Huelsenbeck, J. P. (1993). Success of phylogenetic methods in the four-taxon case. *Syst. Biol.* **42:** 247–264.

Hillis, D. M., Huelsenbeck, J. P., and Cunningham, C. W. (1994). Application and accuracy of molecular phylogenies. *Science* **264:** 671–677.

Huelsenbeck, J. P. (1995). Performance of phylogenetic methods in simulation. *Syst. Biol.* **44:** 17–48.

Huelsenbeck, J. P., Bull, J. J., and Cunningham, C. W. (1996). Combining data in phylogenetic analysis. *Trends Ecol. Evol.* **11:** 152–158.

Hull, D. L. (1980). "Science as a Process," Univ. of Chicago Press, Chicago.

Källersjö, M., Farris, J. S., Chase, M. W., Bremer, B., Fay, M. F., Humphries, C. J., Petersen, G., Seberg, O., and Bremer, K. (1998). Simultaneous parsimony jackknife analysis of 2538 *rbc*L DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Pl. Syst. Evol.* **213:** 259–287.

Kluge, A. G. (1989). A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.* **38:** 7–25.

Kluge, A. G., and Wolf, A. J. (1993). Cladistics: What's in a word? *Cladistics* **9:** 183–199.

Lecointre, G., Philippe, H., Vân Lê, H. L., and Le Guyader, H. (1993). Species sampling has a major impact on phylogenetic inference. *Mol. Phylogenet. Evol.* **2:** 205–224.

Maddison, W. P. (1989). Reconstructing character evolution on polytomous cladograms. *Cladistics* **5:** 365–377.

McLaren, I. A. (1960). Are the Pinnipedia biphyletic? *Syst. Zool.* **9:** 18–28.

Mickevich, M. F., and Farris, J. S. (1981). The implications of congruence in *Menidia*. *Syst. Zool.* **30:** 351–370.

Mishler, B. D. (1994). Cladistic analysis of molecular and morphological data. *Am. J. Phys. Anthropol.* **94:** 143–156.

Nei, M. (1991). Relative efficiencies of different tree-making methods for molecular data. *In* "Phylogenetic Analysis of DNA Sequences" (M. M. Miyamoto and J. Cracraft, Eds.), pp. 90–128. Oxford Univ. Press, Oxford.

Novacek, M. J. (1992). Mammalian phylogeny: Shaking the tree. *Nature* **356:** 121–125.

O'Brien, S. J., Nash, W. G., Wildt, D. E., Bush, M. E., and Benveniste, R. E. (1985). A molecular solution to the riddle of the giant panda's phylogeny. *Nature* **317:** 140–144.

Page, R. D. M. (1993). "COMPONENT user's manual (version 2.0)," Trustees of The Natural History Museum, London.

Patterson, C., Williams, D. M., and Humphries, C. J. (1993). Congruence between molecular and morphological phylogenies. *Annu. Rev. Ecol. Syst.* **24:** 153–188.

Pecon Slattery, J., and O'Brien, S. J. (1995). Molecular phylogeny of the red panda (*Ailurus fulgens*). *J. Hered.* **86:** 413–422.

Purvis, A. (1995a). A composite estimate of primate phylogeny. *Philos. Trans. R. Soc. Lond. B* **348:** 405–421.

Purvis, A. (1995b). A modification to Baum and Ragan's method for combining phylogenetic trees. *Syst. Biol.* **44:** 251–255.

Ragan, M. A. (1992). Phylogenetic inference based on matrix representation of trees. *Mol. Phylogenet. Evol.* **1:** 53–58.

Repenning, C. A. (1990). Oldest pinniped. *Science* **248:** 499.

Rice, W. R. (1989). Analyzing tables of statistical tests. *Evolution* **43:** 223–225.

Robinson, D. F., and Foulds, L. R. (1981). Comparison of phylogenetic trees. *Math. Biosci.* **53:** 131–147.

Sanderson, M. J., and Donoghue, M. J. (1989). Patterns of variation in levels of homoplasy. *Evolution* **43:** 1781–1795.

Sanderson, M. J., Purvis, A., and Henze, C. (1998). Phylogenetic supertrees: Assembling the trees of life. *Trends Ecol. Evol.* **13:** 105–109.

Sarich, V. M. (1969a). Pinniped origins and the rate of evolution of carnivore albumins. *Syst. Zool.* **18:** 286–295.

Sarich, V. M. (1969b). Pinniped phylogeny. *Syst. Zool.* **18:** 416–422.

Siddall, M. E. (1998). Success of parsimony in the four-taxon case: Long-branch repulsion by likelihood in the Farris Zone. *Cladistics* **14:** 209–220.

Smith, A. B., and Littlewood, D. T. J. (1994). Paleontological data and molecular phylogenetic analysis. *Paleobiology* **20:** 259–273.

Steel, M. A., and Penny, D. (1993). Distributions of tree comparison metrics—Some new results. *Syst. Biol.* **42:** 126–141.

Swofford, D. L. (1999). "PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4," Sinauer, Sunderland, MA.

Swofford, D. L., Olsen, G. J., Waddell, P. J., and Hillis, D. M. (1996). Phylogenetic inference. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz, and B. K. Mable, Eds.), pp. 407–514. Sinauer, Sunderland, MA.

Tateno, Y., Takezaki, N., and Nei, M. (1994). Relative efficiencies of the maximum-likelihood, neighbor-joining, and maximum-parsimony methods when substitution rate varies with site. *Mol. Biol. Evol.* **11:** 261–277.

Tedford, R. H. (1976). Relationships of pinnipeds to other carnivores (Mammalia). *Syst. Zool.* **25:** 363–374.

Van de Peer, Y., and de Wachter, R. (1997). Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site rate variation in 18S rRNA. *J. Mol. Evol.* **45:** 619–630.

Wayne, R. K., Benveniste, R. E., Janczewski, D. N., and O'Brien, S. J. (1989). Molecular and biochemical evolution of the Carnivora. *In* "Carnivore Behavior, Ecology, and Evolution" (J. L. Gittleman, Ed.), Vol. 1, pp. 465–494. Cornell Univ. Press, Ithaca, NY.

Wilkinson, M. (1995). Coping with abundant missing entries in phylogenetic inference using parsimony. *Syst. Biol.* **44:** 501–514.

Wozencraft, W. C. (1993). Order Carnivora. *In* "Mammal Species of the World: A Taxonomic and Geographic Reference" (D. E. Wilson and D. A. Reeder, Eds.), pp. 279–348. Smithsonian Institution Press, Washington, DC.

Wyss, A. R. (1987). The walrus auditory region and the monophyly of pinnipeds. *Am. Mus. Novit.* **2871:** 1–31.

Wyss, A. R. (1988). Evidence from flipper structure for a single origin of pinnipeds. *Nature* **334:** 427–428.

Wyss, A. R. (1989). Flippers and pinniped phylogeny: Has the problem of convergence been overrated? *Mar. Mamm. Sci.* **5:** 343–360.

Wyss, A. R., and Flynn, J. J. (1993). A phylogenetic analysis and definition of the Carnivora. *In* "Mammalian Phylogeny: Placentals" (F. S. Szalay, M. J. Novacek, and M. C. McKenna, Eds.), pp. 32–52. Springer–Verlag, New York.