

properties for phylogeny estimation. Recently Kluge and Wolf (1993) have questioned the existence of classes of characters as "mind-independent categories [p. 190]." They suggest that systematists must "question artificial subdivisions of evidence because there is no reason to believe those definitions have discoverable boundaries [p. 190]." Under this view, it would seem there is no reason even to consider comparing behavioral and morphological characters.

Without delving into what is meant by a "mind-independent category," we take the position that one can define classes of characters that have systematically different properties with respect to estimating phylogenetic relationships. For example, codons in pseudogenes generally evolve more rapidly than those in protein-coding genes (Li & Graur 1991). Similarly, within protein-coding genes, third base positions in codons tend to evolve more rapidly than first and second positions (Li et al. 1985). Different genes may even experience different branching histories (Avisé et al. 1983; Tajima 1983). Whether such categorizations are mind-independent or not, their recognition seems to be useful for phylogenetic studies — for example, in refining methods of phylogeny estimation (Bull et al. 1993; Miyamoto et al. 1994).

Although we believe that recognizing classes of characters is useful, it is not clear that traditional categorizations have meaning for phylogenetic analyses. Behavior, morphology, molecules, and physiology may be definable categories, but do characters in these different classes have different properties for phylogeny estimation?

In this chapter, we are concerned with the question of whether behavior and morphology have such different properties. As argued above, we consider it valid to examine different classes of characters. However, at least in this case, we do not consider it valid to assume that one class of characters is inferior to the other without empirical proof. Both behavior and morphology subsume an enormous diversity of traits that defy easy generalization. Furthermore, behavior and morphology are often strongly associated with each other. For example, the distinction between substrate spawning and mouthbrooding in cichlid fishes involves much more than the mouthbrooders simply picking their eggs up and orally brooding them. A large suite of morphological changes accompanies the appearance of mouthbrooding. Mouthbrooders typically have wider heads, more and longer gill rakers, and more yolk in their eggs. In addition, the mouthbrooding species have non-adhesive

## CHAPTER 7

# Comparing Behavioral and Morphological Characters as Indicators of Phylogeny

Peter H. Wimberger and Alan de Queiroz

(Writers share equal authorship)

"...we can understand on the principle of inheritance, how it is that the thrush of South America lines its nest with mud, in the same peculiar manner as does our British thrush...."

-- Darwin (1859, p. 243)

As the above quote indicates, Darwin recognized that behavioral traits, like morphology, could be inherited and thus reflect evolutionary affinities. Through the work of Whitman (1899) and Heinrich (1911) and, especially, early ethologists such as Lorenz (1941) and Tinbergen (1959), the idea that behavioral traits can and should be used as indicators of phylogeny has been maintained (McLennan et al. 1988; de Queiroz and Wimberger 1993). However, many biologists still view behavioral characters as inferior to morphology for estimating phylogenetic relationships. In the extreme, it has even been claimed that behavior is essentially worthless for reconstructing phylogenetic history (Atz 1970; Aronson 1981).

Initially we have to ask whether the dichotomy between behavioral and morphological characters is, in any sense, "real." The division of a systematic data set into behavioral and morphological subsets assumes at least the possibility that these classes of characters have different

eggs and a different developmental pattern than substrate spawners (Fishelson 1966; Chardon & Vandewalle 1971). The behavioral character, parental care mode, involves morphological change as well as behavioral change. Many behavioral traits show this dependence on morphology and, ultimately, since behavior is mediated by the nervous system, all behavior is dependent on morphology at some level. These considerations suggest there will be no absolute distinction between properties of behavioral and morphological characters.

Nonetheless, it could be that behavioral characters tend to be less reliable indicators of phylogeny. At least two reasons have been given for why this might be so. One is that behavioral characters are more evolutionarily labile than morphology (Atz 1970; Baroni Urbani 1989) and, therefore, the phylogenetic information in behavioral traits erodes more quickly through time than for morphology. Wilson (1975) states that, "behavior [is] the part of the phenotype most likely to change in response to long-term changes in the environment." In the extreme, this argument has been used to suggest that behavior is essentially uncorrelated with phylogeny. The second reason for dismissing behavioral characters is the claim that identifying homologous behavioral patterns is difficult, if not impossible (Atz 1970; Hodos 1976; Aronson 1981). Brooks and McLennan (1991) and de Queiroz and Wimberger (1993) discuss these ideas and their history more thoroughly.

Here we present several analyses bearing on the question of whether behavior and morphology should be considered different classes for phylogeny estimation. In an earlier study (de Queiroz & Wimberger 1993), we found no statistically significant difference in levels of homoplasy (evolutionary convergence and reversal) between behavioral and morphological characters from a wide variety of systematic studies. Systematic methodology provides a number of different measures of homoplasy. Here, we review our earlier results and present an additional analysis using a measure of homoplasy that may better reflect phylogenetic information content than the measure we originally used. Because differences in levels of homoplasy do not necessarily indicate differences between the estimated phylogenetic trees, we also make two kinds of direct comparisons of trees estimated from behavioral and morphological data. The first of these addresses the question of whether the behavioral and morphological trees are more similar than randomly

selected pairs of trees (Penny et al. 1982; Hendy et al. 1984), the implication being that such similarity indicates that both kinds of data reflect the true phylogeny. The second tree comparison addresses the question of whether behavioral and morphological data disagree more in their estimates of phylogeny than one would expect from random partitions of the total data set. This last analysis, which uses a test recently proposed by Swofford (1994), is preliminary in nature but represents perhaps the most appropriate method for assessing whether a data set should be partitioned into subsets prior to phylogenetic analysis, a controversial implication of the recognition of character classes (e.g., Bull et al. 1993; Kluge & Wolf 1993).

## *Levels of homoplasy in behavioral and morphological characters*

### *A. General procedure*

To assess the generality of statements made about the relative utility of behavioral and morphological characters for estimating phylogenetic relationships, we have examined a large number of systematic studies that used behavioral and/or morphological data. We compared levels of homoplasy in the two types of characters, the idea being that levels of homoplasy are generally inversely related to utility for phylogeny estimation (but see Goloboff 1991). We performed two kinds of comparisons. In the first, we compared homoplasy of behavioral and morphological characters within data sets that contained both kinds of characters. This analysis controls for potential differences among clades in rates of character evolution, tree topology, or other factors that might influence the level and/or measurement of homoplasy. In the second comparison, we examined overall levels of homoplasy for data sets containing either behavioral or morphological characters. This second analysis lacks the control by study group of the first analysis but provides a larger sample of taxa and a more powerful statistical test. The methods are fully described in de Queiroz and Wimberger (1993). Here we describe only the most salient points.

We collected morphological and behavioral data sets from the literature (Tables 7.1 and 7.2). We defined a behavioral character as any

character representing movement of the organism or its parts. Thus, behavioral characters ranged from simple, stereotyped movements to more complex characters such as mating system or nesting dispersion (solitary vs. colonial). Epiphenomena of behavior, such as nest architecture, were also classified as behavioral characters. Courtship and territorial behavior were, by far, the most common types of behavior used in systematic studies, followed by nest site and architecture which were primarily used in studies of social insects and birds (Fig. 1). Ideally, it would have been preferable to have a more uniform distribution of behavioral types represented in this study. Morphological characters included gross external characteristics, color pattern, osteological characters and features of the soft anatomy.

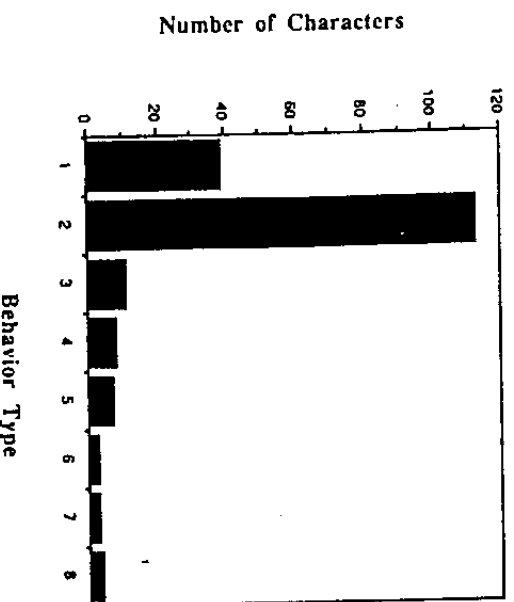


Figure 1. Behavioral characters used in the within-study comparison of CIs and RIs broken down into: (1) nest architecture and site; (2) courtship and territorial behavior; (3) parental care; (4) other reproductive behavior; (5) feeding and other maintenance (including locomotion); (6) general lifestyle; (7) social system; (8) other.

Table 7.1. Systematic data sets used for comparison of character CIs and RIs within data sets<sup>a</sup>

Taxon	Rank	# Taxa	MC	BC	MCI	BCI	MRJ	BRI
Polistine wasps	Genera	28	34	18	0.60	0.59	0.65	0.81
Eumachine wasps	Genera	25	25	1	0.86	1.00	0.85	1.00
Vespine wasps	Species/Genera	7	17	8	0.88	0.71	0.79	0.38
Vespid wasps	Subfamilies	6	20	8	0.78	0.96	0.63	0.96
<i>Apis</i> bees	Species	6	9	3	0.94	0.83	0.94	0.83
Meloid beetles	Genera	9	21	2	0.87	1.00	0.90	1.00
Arachnids	Orders	11	63	1	0.71	1.00	0.64	1.00
Shrimps	Orders	5	3	1	0.83	1.00	0.67	1.00
Gasterosteiform fishes	Genera	6	18	17	0.86	0.96	0.84	0.91
North American Hyliid frogs	Species	13	9	4	0.94	0.90	0.96	0.50
Pipid frogs	Species/Genera	7	52	2	0.88	0.75	0.79	0.50
<i>Triturus</i> newts	Species	9	2	11	0.52	0.81	0.62	0.78
Seaducks	Species/Genera	10	12	2	0.83	0.75	0.83	0.84
Alcid birds	Species	23	31	2	0.82	0.70	0.90	0.85
Pelecaniform birds	Families	7	26	5	0.92	1.00	0.87	1.00
<i>Ochthoeca</i> flycatchers	Genera	5	5	2	0.90	1.00	0.80	1.00
Today-tyrant flycatchers	Genera	9	8	2	1.00	0.75	1.00	0.50
<i>Myiobius</i> flycatchers	Genera	5	8	2	0.85	1.00	0.81	1.00
<i>Empidonax</i> flycatchers	Genera	7	6	3	1.00	0.83	1.00	0.83
Manakin birds	Species	19	36	29	0.90	0.84	0.94	0.83
Squamate reptiles	Families	19	61	3	0.68	0.67	0.70	0.77
Sand lizards	Species	10	36	2	0.87	0.67	0.85	0.67
Francolin birds	Species	13	17	5	0.82	0.67	0.85	0.87

<sup>a</sup>Rank is the taxonomic rank of the terminal taxa. MC and BC are the number of morphological and behavioral characters, respectively. MCI and BCI are the mean CIs for morphological and behavioral characters. MRJ and BRI are the mean RIs for morphological and behavioral characters. References are given in de Queiroz and Wimberger (1993), except for the francolin data set, which is from Crowe et al. 1992. See text and de Queiroz and Wimberger (1993) for further description.

When a single data set was used for a group we used the tree estimated by the author. For some groups we combined the available behavioral, morphological, and molecular data sets and, in these cases, we obtained trees using *PAUP 3.0* (Swofford 1989).

Table 7.2. Systematic data sets used for comparison of overall CIs for morphological and behavioral data sets\*

Taxon	Character Type	Rank	#Taxa	#Chars	CI
Polistine wasps	Behavior	Genera	28	39	0.65
Vespine wasps	Behavior	Species/Genera	7	13	0.79
Vespid wasps	Behavior	Subfamilies	6	12	1.00
Zygothrica flies	Behavior	Species	7	14	0.75
Gasterosteiform fishes	Behavior	Genera	6	23	0.85
<i>Trithrus</i> newts	Behavior	Species	9	13	0.73
Manakin birds	Behavior	Species	19	46	0.78
Cranes	Behavior	Species	13	39	0.64
<i>Spilanthes</i> composites	Morphology	Species	6	17	0.79
<i>Pegolattia</i> composites	Morphology	Species	9	19	0.79
Myrtaceae	Morphology	Species	14	31	0.72
Andropogoneae	Morphology	Species	20	28	0.44
<i>Montanoa</i> (Asteraceae)	Morphology	Species	25	51	0.63
<i>Eucalyptus</i> (Myrtaceae)	Morphology	Species	29	65	0.43
Pontederiaceae	Morphology	Species	37	65	0.39
Triticeae (Poaceae)	Morphology	Genera	29	126	0.31
Asteraceae	Morphology	Genera	29	81	0.56
Clusiaceae*	Morphology	Genera	68	94	0.32
Monocots	Morphology	Families	15	56	0.60
Centrospermae	Morphology	Families	20	71	0.35
Angiosperms*	Morphology	Families	47	61	0.26
Rhodophyta	Morphology	Orders	15	37	0.55
Green algae & bryophytes	Morphology	Classes	11	41	0.82
Seed plants	Morphology	Classes	20	31	0.62
Seed plants	Morphology	Classes	20	62	0.50
Steamer-ducks	Morphology	Species	4	23	1.00
<i>Anoxybus</i> geckos	Morphology	Species	6	34	0.84
Australian birds	Morphology	Species	9	33	0.81
Oligochaete annelids	Morphology	Species	11	83	0.64
<i>Stomias</i> teleosts	Morphology	Species	12	32	0.70
Mackerels	Morphology	Species	20	74	0.68
Anole lizards	Morphology	Species	24	31	0.60
<i>Bledius</i> beetles	Morphology	Species	35	72	0.43
Xantusiid lizards	Morphology	Genera	5	36	0.72
Leptopodomorph bugs	Morphology	Genera	10	49	0.80
Gomphaesurine odonates	Morphology	Genera	16	15	0.50
Stomiid teleosts*	Morphology	Genera	27	323	0.49
Mammals*	Mammals	Genera	57	707	0.45

Table 7.2. (cont.)

Taxon	Character Type	Rank	#Taxa	#Chars	CI
Plecoptera	Morphology	Families	22	113	0.63
Amniotes	Morphology	Classes	8	84	0.89
Cercerian flatworms	Morphology	Classes	9	39	0.93
<i>Collybia</i> basidiomycetes	Morphology	Species	11	25	0.53
Rusts	Morphology	Genera	17	24	0.53
Eukaryotes	Morphology	Classes	36	105	0.34

\*#Chars is the number of binary-character equivalents, i.e., the number of characters that would constitute the data set if all multistate characters were converted to binary (two-state) characters. References are given in de Queiroz and Wimberger (1993; for behavioral data sets) and Sanderson and Donoghue (1989; for morphological data sets). Asterisks indicate data sets not used in the analyses reported here. See the text and de Queiroz and Wimberger (1993) for further description.

We tested the assertion that behavioral characters are more prone to convergence and reversal (show more homoplasy) than morphological characters by examining two homoplasy measures, the consistency index (CI, Kluge & Farris 1969) and the retention index (RI, Farris 1989). The CI is the minimum number of steps (evolutionary character state changes) required by the character assuming no homoplasy divided by the number of steps required by the character on the specified tree. The minimum number of changes assuming no homoplasy is always just one less than the number of character states. Thus, if a two-state character changes once on the tree in question, it will have a CI of 1.0 indicating no homoplasy for that character on the given tree. If the character requires two changes on the tree (say from A→A' and then A→A) the CI will equal 0.5 (minimum number of changes = 1, actual number of changes = 2; CI = 1/2 = 0.5).

The character retention index (Farris 1989) is:

$$RI = \frac{(g - s)}{(g - m)}$$

where  $g$  is the greatest possible number of steps for the character on any tree,  $s$  is the actual reconstructed number of steps on the specified tree,

and  $m$  is the minimum possible number of steps on any tree. The  $RI$ , like the  $CI$ , compares the number of steps on the specified tree with the minimum number of steps. However, unlike the  $CI$ , the  $RI$  takes into account the maximum possible number of steps for the character given the number of taxa that have each state.

A character with a high  $CI$  is one that rarely changes and thus, it could be argued, should be given high weight (Farris 1969; Carpenter 1988b; Sundberg 1989); the same is not necessarily true of a character with a high  $RI$ . However, the  $RI$ , precisely because it does take into account the number of taxa that have each state, is probably a better measure of the phylogenetic information content of a character; as Farris (1989, p. 407) puts it, the  $RI$  "reflects the degree to which similarities apparent in the data can be retained as homologies on a tree." Because of the somewhat different implications of these two measures, we report results for both here.

We excluded phylogenetically uninformative characters (i.e., characters that were either invariant within the ingroup or in which derived character states were unique to single taxa) from our calculations because including these characters inflates the  $CI$ . (The  $RI$  cannot be calculated for such characters because the denominator is zero.) Ideally, because derived character states unique to single taxa (autapomorphies) represent evolutionary change, they should be included in the analysis (de Queiroz & Wimberger 1993). However, because some of the data sets we used excluded autapomorphies, we were unable to perform a separate analysis including them.

#### *B. Relative homoplasy levels within data sets*

To compare the average levels of homoplasy within data sets, we calculated the mean  $CI$  and  $RI$  for each character on the tree(s) given by the author(s) or that we estimated from data sets that included both behavioral and morphological characters (Table 7.1). The average morphological  $CI$  ( $MCI$ ), behavioral  $CI$  ( $BCI$ ), morphological retention index ( $MRI$ ), and behavioral  $RI$  ( $BCI$ ) were then calculated for each data set. Although the distribution of  $CI$ s is highly nonnormal, the distributions of the differences between  $MCI$ s and  $BCI$ s and  $MRI$ s and  $BCI$ s within data sets were not significantly different from normality (Lilliefors test for  $CI$ :  $N = 23$ , maximum distance = 0.114,  $p > 0.6$ ; for

$RI$ :  $N = 23$ , maximum distance = 0.134,  $p > 0.3$ ). As a result, we used paired  $t$ -tests to examine the difference between  $MCI$ s and  $BCI$ s, and  $MRI$ s and  $BCI$ s; each data set provided an independent paired comparison for the analysis.

Behavioral characters were no more prone to convergence and reversal than the morphological characters used in systematic studies. Paired  $t$ -tests indicated no significant differences between morphological and behavioral character  $CI$ s or  $RI$ s (for  $CI$ :  $X_{morph} - X_{behav} = -0.01 \pm .160$ ,  $t = -0.195$ ,  $p > 0.8$ ; for  $RI$ :  $X_{morph} - X_{behav} = 0.00 \pm 0.247$ ,  $t = 0.003$ ,  $p > 0.9$ ; for both tests  $N = 23$ ).

#### *C. Comparison of morphological and behavioral data sets*

The second approach we used to compare relative homoplasy levels was to examine the ensemble  $CI$  of data sets composed exclusively of morphological or behavioral characters. The ensemble  $CI$  is the sum of the numerators of the character  $CI$ s divided by the sum of their denominators. The morphological data sets were taken from Sanderson and Donoghue (1989) and the behavioral data sets were gleaned from the literature (Table 7.2).

The data were analyzed using analysis of covariance where number of characters and taxa were used as covariates because both variables affect the  $CI$  (Archie 1989; Sanderson & Donoghue 1989). The  $CI$ s were natural log transformed to make the variance independent of the means and because  $CI$  is curvilinearly related to number of taxa. We looked for differences in intercept of the log  $CI$  between morphological and behavioral characters at the mean number of taxa and characters for behavioral data sets because making the comparison at the mean for all data sets would have meant that the comparison was made outside the range of number of characters for the behavioral data sets.

The ANCOVAs indicated no significant difference between morphological and behavioral character  $CI$ s when either number of characters or number of taxa was used alone as a covariate, or when the two covariates were used together (Table 7.3, Fig. 2). Both covariates were significant when used alone; however, when used in the same analysis, only number of taxa had a significant effect. Interaction effects were not significant for any of the analyses.