

## METHODS AND TECHNIQUES

## Characters as groups: a new approach to morphological characters in phylogenetic analysis

Bruce K. Kirchoff<sup>1</sup>, Scott J. Richter<sup>2</sup> & David L. Remington<sup>1</sup>

<sup>1</sup> *Department of Biology, University of North Carolina at Greensboro, Greensboro, North Carolina 27402-6170, U.S.A. kirchoff@uncg.edu (author for correspondence)*

<sup>2</sup> *Department of Mathematical Sciences, University of North Carolina at Greensboro, Greensboro, North Carolina 27402-6170, U.S.A.*

A new method for working with morphological characters is described and explored in experiments using human participants. The method uses direct comparison and sorting of images to produce hierarchical character-cladograms. A character-cladogram is a graphical representation of a single character that serves as a hypothesis of phylogeny based on that character. Each dichotomy in the character-cladogram represents a character state. Character states are unnamed, thus avoiding problems that arise through the application of verbal labels. Experiments with human participants are used to explore the conditions under which direct comparison produces reliable (consistent from investigator to investigator) and valid (in agreement with an independent estimate of phylogeny) characters. Participants were drawn from students taking a course in plant diversity at UNC Greensboro, and professional plant morphologists attending the Botany 2004 meetings. The students were randomly assigned to trained and untrained groups. Training was carried out using a method that has been shown to change a participant's mode of visual processing from analytic (the mode used by visual novices) to holistic (an additional mode only employed by visual experts). Morphologists (no specialists of the taxonomic group) were included in the study to investigate the effects of disciplinary expertise on the ability to describe character-cladograms. They received no additional training beyond that available to them as disciplinary experts. The results suggest an improvement in both reliability and validity after the training regime. We found no support for the idea that the morphologists differed from untrained students in their ability to produce reliable or valid character-cladograms. Disciplinary expertise may not translate into the ability to make reliable and valid assessments of similarity in an unfamiliar visual domain. Based on these results, we suggest a method for creating morphological characters and character states.

**KEYWORDS:** character, character state, complex characters, configural processing, Costaceae, holistic processing, morphology, phylogeny, Zingiberaceae

### INTRODUCTION

Morphological characters remain an essential tool for phylogenetic studies, even in the current age of molecular systematics (Scotland & al., 2003; Jenner, 2004; Wiens 2004; Smith & Turner, 2005; Lee, 2006). Consequently, the reliability and validity of the methods used to identify characters and define character states is highly relevant to systematics research. Although identifying appropriate characters has been a longstanding issue in systematics, advances in phylogenetic methods have not been fully matched by improvements in methods for selecting and evaluating characters (Neff, 1986; Kemp, 1999: 63; Judd & al., 2002: 46–51; Richards, 2003; Mishler, 2005).

This paper explores a solution to the problem of delimiting morphological characters by applying insights from cognitive psychology. We do not deal with the

many theoretical aspects of the character concept, which have been presented by others (see Wagner, 2001), but suggest new ways of thinking about the terms character and character state, and propose operational definitions of these terms. We view characters as arising through an interaction between the objective object of study (the organism) and the investigator. Characters depend not only on the structure of the organism, but on the ability of the investigator to reliably dismember the organism into meaningful units, and on his or her ability to assess the similarity of these units across study taxa (Richards, 2003). The first process produces characters, the second character states.

At a minimum, characters and character states should be consistent from investigator to investigator (they should be reliable), and they should be good estimates of the process being investigated (they should be valid). Reliability can be investigated by statistical

comparisons of different subjects' performance in character and character state recognition tasks. Validity is more difficult to measure. Because we do not know the course of evolution with certainty, we can only estimate a character's validity by comparing it to the best current estimates of evolutionary relationships based on independent data. Although this method is imperfect, we can use it to estimate the quality of a character in experimental tests of character creation methods. Better characters will have higher agreements with independent estimates of phylogeny.

Our work involves a reconceptualization of the terms "character" and "character state." We approach characters as groups of complex features that are divided into hierarchical clusters, called character states (Fig. 1). A complex feature is a part of an organism that is identified in the process of homology assessment, and is not decomposed into smaller constituent units. A part is an internally integrated system that is isolated from its surround (McShea, 2001; McShea & Venit, 2001). In the method we present here, parts are represented photographically

for the creation of character states. Photographic representation is used to show variation in the parts across the study taxa. For example, much of the variation in plant ovary structure can be represented with cross sections at the middle of the locules (Fig. 1). Prior to subdivision into hierarchical groups, the collection of these photographs serves as a pre-operational character: a character that has not been divided into character states. To complete the process of character-state description, the photographs are laid on a table and sorted into hierarchal, dichotomous groups by direct comparison of the images. We call the hierarchal diagram that results from the sorting process a character-cladogram (Fig. 1).

A character-cladogram is a dichotomously branching diagram that serves as a hypothesis of relationships among taxa based on a single character. Character-cladograms differ from character-state trees (Swofford & Begle, 1993: 13) in not encoding transformation series. They differ from networks in that a network can have internal reticulations (e.g., Holland & al., 2004), and because the internal nodes of networks are sometimes

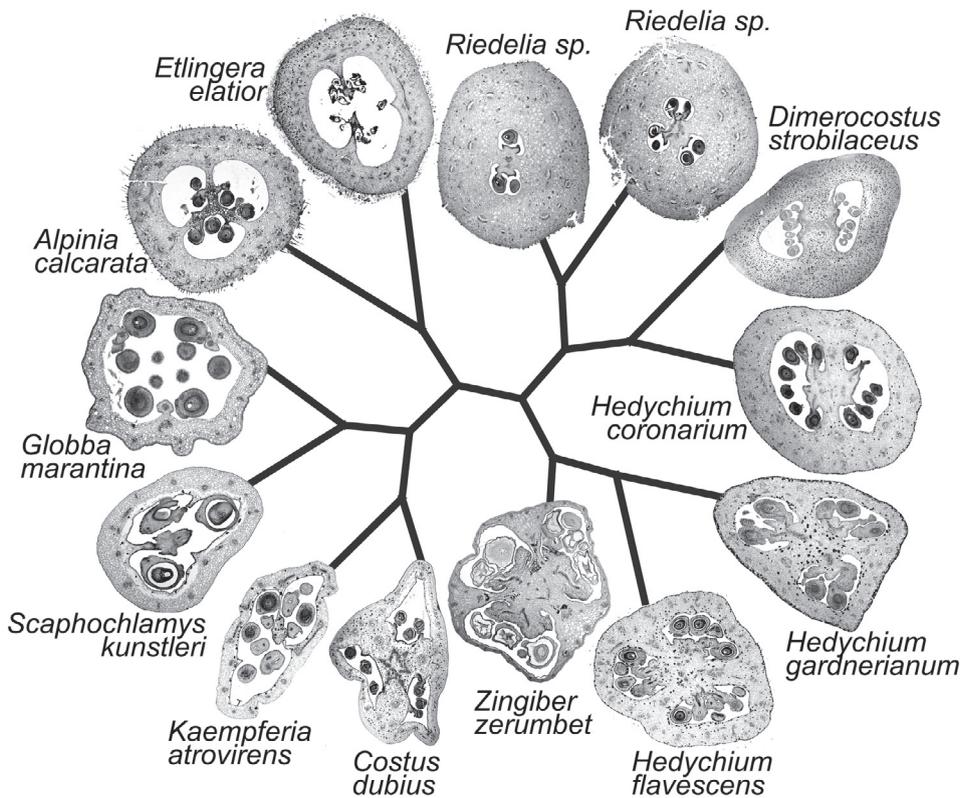


Fig. 1. Character-cladogram of the character "ovary structure at mid-locular level," the position from which the photographs were taken. Each terminal represents one taxon, or operational taxonomic unit. Variation within a taxon is represented by photographs of each variant (e.g., *Riedelia* sp.). Character states are the nested, hierarchical groups (clans) and are unnamed (Wilkinson & al., 2007). For instance, the group represented by the photographs of *Zingiber zerumbet*, *Hedychium flavescens*, and *Hedychium gardnerianum* form one character state, as does the nested group consisting of *Hedychium flavescens*, and *Hedychium gardnerianum*. This character-cladogram was created by the participant who was trained for nine hours (see Materials and Methods).

labeled. Although all of the character-cladograms discussed in this paper are unrooted, character-cladograms may also be rooted based on the position of an outgroup. We defer the discussion of rooted character-cladograms to a later time. Character-cladograms can be displayed graphically to convey the content of the character (Fig. 1), represented as NEXUS tree statements, or coded for use in phylogenetic analysis with Matrix Representation with Parsimony or as step-matrices (Farris, 1973; Swofford & Begle, 1993:13).

Because we advocate the use of complex features, our definitions of the terms “character” and “character state” differ from those in common use in phylogenetic practice. What we call a character is more akin to a suite of correlated conventional (atomistic) characters than to a single character. For convenience, and to distinguish them from conventional characters, we will refer to these non-atomistic characters as “complex characters.” We recognize that there is a continuum between complex and atomistic characters, and that no character can be fully complex or atomistic.

Our concept of a character state is even more divergent from that in current use. For us, a character state is a section (“clan”) of a hierarchical character-cladogram whose terminals are the complex features of the taxa (Fig. 1) (Wilkinson & al., 2007). If we restrict ourselves to a single binary partition of the character-cladogram, our concept of a character state approaches that in conventional use. For instance, if we segregate *Etingera*, *Alpinia*, *Globba*, *Scaphochlamys*, *Kaempferia*, and *Costus* from *Riedelia*, *Dimerocostus*, *Hedychium*, and *Zingiber* based on the central dichotomy in Fig. 1, we can code the result as a conventional two-state character. However, the complex nature of the features still distinguishes these characters from those in conventional use.

The use of complex features and direct comparison allows us to apply insights from cognitive psychology to help improve performance in the character state description (sorting) task. Cognitive psychology is relevant because, like color, the perception of form depends on the presence of an observer (Landa & Fairchild, 2005). Character states arise when observed variation is partitioned into discrete units by a human observer. Even when the ordination is done quantitatively, decisions must be made as to how to delimit the states (Gift & Stevens, 1997; Guerrero & al., 2003). These decisions have the potential to be influenced by human cognitive abilities. We cannot eliminate this source of variation, but we can learn its parameters and how to limit it through appropriate training. Recent results in cognitive psychology provide the conceptual and methodological tools to do this. These results stem from research on human visual processing modes, and have been experimentally verified over several decades (Scapinello & Yarmey,

1970; Carey & Diamond, 1977; Jacoby & Brooks, 1984; Gauthier & Tarr, 1997; Gauthier & al., 1998). We discuss these results more fully below.

In addition to proposing an operational definition of character and character-state, this paper presents preliminary evidence to suggest that subjects who are trained to be holistic visual processors are able to produce character-cladograms that are both more reliable and more valid than are those produced by subjects who process features analytically. Although these results are not conclusive, they suggest that work with complex characters holds promise. Characters described through direct comparison of complex features will be particularly useful in areas such as flower development where complex shapes must be compared. They may also find use in palaeontology, where molecular characters are not available.

Direct comparison of complex features was previously explored by Kirchoff & al. (2004) in the context of determining if complex data allow the production of more reliable and valid character states than simple data. Eight groups of plant systematists were given one of two classes of drawings of plant parts and asked to divide them into character states. The first class of drawings consisted of simple outlines of cotyledons extracted from a published phylogenetic analysis (Thiele & Ladiges, 1996). Four groups of systematists received these simple drawings. The other four groups received triplets of drawings from the same paper (Thiele & Ladiges, 1996): a cotyledon, seedling leaf, and inflorescence bract. The triplets were used to simulate complex data such as might be garnered by looking at part of a plant. Each trial resulted in four characters, one for each group of systematists. Analysis revealed little agreement among systematists when they classified the simple data. However, using the complex data the systematists produced smaller groups, i.e., more precisely defined character states. These states had greater inter-investigator agreement, and had significantly greater agreement with an independent assessment of phylogeny (Mast & Givnish, 2002). Kirchoff & al. (2004) related these results to research on visual processing mechanisms.

**Visual processing mode and complex characters.** — Human beings possess two distinct modes for processing visual information (stimuli). Holistic (or configural) processing is the propensity to view an object as a whole rather than as composed of parts, whereas analytic (or featural) processing is the tendency to focus on the parts more than the whole (Diamond & Carey, 1986). Holistic processing involves simultaneously attending to all aspects of an object so that it is difficult to selectively attend to a single part and ignore the others, even when specifically instructed to do so (Farah & al., 1995, 1998; Goldstone, 2000). Analytic processing involves selectively attending to distinct, separate features of a stimulus

(e.g., the number of locules in an ovary, thickness of ovary wall, etc.). Analytic processing is the mode employed by novices, whereas visual experts in some perceptual domain process objects holistically (Jacoby & Brooks, 1984; Gauthier & Tarr, 1997; Gauthier & al., 1998).

The holistic/analytic distinction is relevant to character description because visual processing mode affects which aspects of an object are used in categorizations (Jacoby & Brooks, 1984; Diamond & Carey, 1986). Analytic processors will tend to focus on isolated features of the organism, whereas holistic processors also have the ability to look at relationships between features, and at the stimulus as a whole. The difference between analytic and holistic processing is reflected in conventional systematic practice in the difference between character and character state description. Characters are recognized through homology assessment, a process that depends largely on configural (holistic) information. Two of Remane's (1952) criteria for determining homology (similarity in position, linkage with intermediate forms) are configural in this sense. Similarity in position can only be assessed in relationship to other parts, whereas linkage with intermediate forms is established based, in part, on the internal configuration of the parts. Likewise, two of Patterson's (1982) three similarity criteria are holistic (topographic identity), or depend on configural information (compositional correspondence). On the other hand, conventional morphological character states are more commonly described analytically. For instance, both the presence of fructose oligosaccharides (Rudall & al., 2000, character 35) and the presence of a hairy surface on a butterfly egg (Freitas & Brown, 2004, character 2) are analytic characters. Neither depends on configural information. Both can be assessed based on the observation of one isolated aspect of the organism. The vast majority of conventional characters are described analytically.

Theoretical considerations suggest a relationship between a scientist's ability to process visual information holistically and his or her ability to describe better complex characters through direct comparison. Organisms possess a hierarchical organization in which complex features are highly integrated and conserved over large taxonomic groups (Riedl, 1978). A recent population genetic model demonstrates that when genes affect more than two characters that are under simultaneous stabilizing selection, morphological variation is restricted (Wagner, 1998; Waxman & Peck, 1998). The restricted states show many interdependencies, similar to the interdependencies seen in complex features (Riedl, 1978). Interdependencies are also uncovered by correlation studies, which suggest that suites of characters are evolving together (Hedenäs, 2002; Kangas & al., 2004). These interdependencies could be either adaptive or genetic (e.g., pleiotropic). The fact that mutant phenotypes are sometimes named based on

complex phenotypes suggests that at least some interdependencies have a genetic basis (Niku & Taipale, 2003; Zebrafish Nomenclature Committee, 2006). These results suggest that organisms possess features that can best be perceived with a holistic method.

In our previous work we found we could explain the higher quality of the complex characters by applying the analytic/holistic processing distinction to the character state description task (Kirchoff & al., 2004). Inspection of the character states suggested that those based on complex data were consistent with holistic processing. The character states described using simple data were consistent with analytic processing. The preliminary experiments reported below support this idea: subjects who have been trained to view photographs holistically seem to produce better characters than do analytic processors.

Several factors appear to be necessary for holistic processing to take place. First, holistic processing develops as a result of experience (Diamond & Carey, 1986; Gauthier & Tarr, 1997, 2002; Gauthier & al., 1998; Goldstone 2000; Gauthier & al., 2003). Although the most work has been done with faces, experienced dog show judges have been shown to process dog photographs holistically (Diamond & Carey, 1986), car enthusiasts process photographs of cars holistically (Gauthier & al., 2003), and even novel created "organisms" (Greebles) can be processed holistically by expert subjects (Gauthier & Tarr, 1997; Gauthier & al., 1998). Second, the objects often share a common configuration: similar parts are in similar positions (Wells & Hryciw, 1984; Diamond & Carey, 1986; Tanaka & Sengco, 1997). Objects with a common configuration cannot be distinguished based on the presence or absence of parts, so subjects must turn to information about the relative sizes and positions of the parts (i.e., their configuration). Finally, the complexity of the stimulus may also be important, although evidence for this is indirect (Kirchoff & al., 2004; Nisbett & Miyamoto, 2005; Miyamoto & al., 2006).

It was initially assumed that many years of experience were needed to become a holistic processor (Diamond & Carey, 1986). More recent work has shown that only a few hours of intense training can substitute for this experience (Gauthier & Tarr, 1997; Gauthier & al., 1998). We used this finding in designing our training regime, described below.

The experimental investigations presented here were done as part of a preliminary study to test the validity of these ideas. The sample sizes are small, and some conditions are not parallel across the experiments. The experiments were exploratory in nature, and were not designed to test rigorously formed hypotheses. We recognize these limitations, but report the results because they suggest that a holistic approach to character description shows promise.

## MATERIALS AND METHODS

### Feature identification and photography. —

Thirteen photographs of ovary cross sections at the mid-locular level were taken from ten species representing eight genera of the families Zingiberaceae and Costaceae (Zingiberales; Table 1). Plant ovaries in the study taxa satisfy all of Remane's (1952) and Patterson's (1982) criteria for homology. Only mature flowers or flower buds close to maturity were sectioned. The species were chosen to represent the range of diversity in these two families. Sections from two variants of one collection (*Riedelia* sp., trilocular and bilocular variants) were included to present features that had similar overall forms (Gestalts) but differed in an analytical feature, the number of locules (Fig. 1). Photographs of three species of *Hedychium* were used for similar reasons. One of these photographs was of a bilocular variant of a normally trilocular species (*Hedychium coronarium*). The other two were from typical trilocular ovaries. Of Costaceae included in the study, *Costus dubius* is trilocular and *Dimerocostus strobilaceus* is bilocular (Newman & Kirchoff, 1992).

Photographs were taken of whole cross sections, a single locule, of the septa alone, and of the epidermis and immediately adjacent tissues (Fig. 2). The photographs of the septa were edited in Adobe Photoshop CS (Adobe Systems Inc., San Jose, California, U.S.A.) to remove the ovules. Without editing, some photographs would have contained ovules and others would not. The presence of ovules in some would have provided analytical clues to their identity.

**Participants.** — Participants in the experiments were either professional plant morphologists attending the Botany 2004 meetings, or undergraduate students studying biology at the University of North Carolina

at Greensboro. One first-semester masters student also participated, but his experience with plant anatomy was no greater than that of the undergraduates. Three morphologists and 10 students participated in the experiments. Although both men and women participated in all of the groups, we will refer to all participants as "he" in order to disguise their identities. The morphologists were all actively engaged in research and were actively publishing. All had, or have since received, federal funding. Although the morphologists all had extensive prior experience with plant anatomy, none were experts in ovary anatomy, and none had any significant experience with Zingiberaceae or Costaceae. All of them had published at least one paper dealing with floral anatomy, between 6 and 60 papers on plant anatomy, and had between 5 and 35 years of experience in plant anatomy research. The students each had two classes in which they had seen sections of higher plants (Principles of Biology II, Plant Diversity) for a total exposure time of approximately five to ten hours. The students were randomly assigned to trained (six students) and untrained (four students) groups, and were paid an hourly wage for their participation.

**Training.** — The training consisted only of familiarizing one group of participants with the photographs described above. No conceptual knowledge of plant or ovary structure was conveyed, nor were descriptions of any for the photographs provided to the participants. The participants were told neither the species names, nor that two of the photographs came from the same species. The training merely provided an opportunity for the students to become very familiar with the photographs by performing a task that engaged their attention. The training regimes were based on the methods used in cognitive psychology, but with a simpler protocol. Protocols similar to those applied here have been shown to change a participant's visual processing mode from

Table 1. Collection information for study species.

Species	Family	Number locules	Garden Accession No.	Voucher	Deposited
<i>Dimerocostus strobilaceus</i> O. Kuntze subsp. <i>strobilaceus</i>	Costaceae	2	Lyon 78.1074	<i>Kirchoff 88-148</i>	BISH
<i>Costus dubius</i> (Afzel.) K. Schum.	Costaceae	3	Waimea 76p941	<i>Kirchoff 88-163</i>	BISH
<i>Alpinia calcarata</i> Roscoe	Zingiberaceae	3	Lyon 80.712	<i>Nagata 2831</i>	HLA
<i>Etilingera elatior</i> (Jack) R.M. Smith	Zingiberaceae	3	Wiamea 74p345	<i>Lau 2616</i>	BISH
<i>Riedelia</i> sp.	Zingiberaceae	3	Lyon L72.0552	<i>Nagata 1444</i>	HLA
<i>Riedelia</i> sp.	Zingiberaceae	2 <sup>a</sup>	Lyon L72.0552	<i>Nagata 1444</i>	HLA
<i>Globba marantina</i> L.	Zingiberaceae	3	FTG 71-348	<i>Kirchoff 84-20</i>	FTG
<i>Kaempferia atrovirens</i> N.E. Br.	Zingiberaceae	3	Waimea 80p258	<i>Lau 2621</i>	BISH
<i>Scaphochlamys kunstleri</i> (Baker) Holttum	Zingiberaceae	3	FTG 83-266	<i>Kirchoff 84-23</i>	FTG
<i>Zingiber zerumbet</i> (L.) Smith	Zingiberaceae	3	Lyon unaccessioned	<i>Kirchoff 88-149</i>	BISH
<i>Hedychium gardnerianum</i> Sheppard ex. Kev. Gawl.	Zingiberaceae	3	Lyon 67-1053	—	—
<i>Hedychium coronarium</i> Koenig	Zingiberaceae	2 <sup>a</sup>	Lyon 67.1051	—	—
<i>Hedychium flavescens</i> Carey ex Roscoe	Zingiberaceae	3	Waimea 74p382	—	—

<sup>a</sup>Bilocular variant.

analytic to holistic (Gauthier & Tarr, 1997; Gauthier & al., 1998). Each trained student was first given prints of the 13 photographs of whole cross sections with common English names affixed to their backs. The participants studied the photographs until they felt comfortable with their names, and were then given identification quizzes until they could correctly name all of them twice in sequence without error. This procedure was repeated three more times with the other types of photographs: a single locule, the septa alone, and the epidermis (Fig. 2). The average amount of time it took the students to complete the training was 2 hrs. and 26 min. (range = 1 hr. 50 min. to 3 hrs. and 10 min.; SD = 34 min.) scattered over three non-contiguous days.

When designing the training we did not anticipate how quickly the students would complete the tasks. The psychological literature suggests that approximately nine hours of training are required to shift a participant from analytic to holistic processing (Gauthier & Tarr, 1997; Gauthier & al., 1998). The duration and simplicity of the task make it unlikely that the students received sufficient training to become full holistic processors (Isabel Gauthier, pers. comm.). To investigate the effect of increasing the training time one untrained student was invited back after completing the exercise, given the same type of training as described above, and asked to perform the sorting task a second time. In this case training was not

stopped when the student could correctly name the images twice in succession, but was continued for nine hours regardless of his proficiency.

**Character state description.** — All of the participants in each group were given prints of the same 13 photographs of full cross sections as were studied by the trained group, and were asked to sort them into hierarchical, dichotomous clusters. The morphologists received color photographs that were printed to scale, while the students received black and white prints that were all printed to the same size (approximately 5 by 7 inches). The difference in stimuli was due to modifications we made based on feedback from the morphologists, the first group to be tested. Conversations with the morphologists following the trials suggested that they did not make use of the color or size information. Because scaled color photographs are difficult to produce, and because we needed to produce several sets of photographs for the training step, we decided to eliminate the color and size information from the trials with the students. We recognize that this difference in stimuli reduces the validity of statistical comparisons between morphologists and students. Compounding this problem, one of the morphologists worked with the photographs as part of a related task immediately before participating in the experiment. This task lasted about one-half hour, and may have influenced his ability to sort the photographs. We will refer to the morphologist

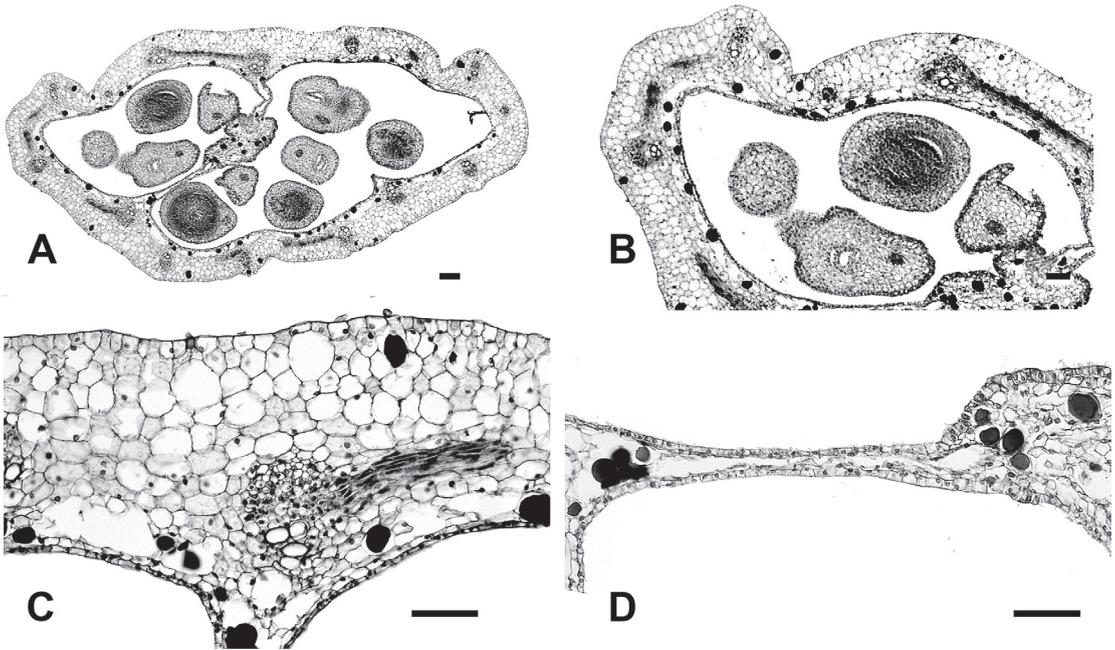


Fig. 2. A–D, *Kaempferia atrovirens*. Representative photographs of the type used in training participants to be holistic visual processors. A, full cross section at the mid-locular level; B, a single locule from the same level; C, ovary wall/epidermis from the mid-locular level; photographs of the comparable region from other species show only the outermost few layers of cells; D, septum from the middle of the locules. Scale bars = 100  $\mu$ m.

who had prior experience with the photographs as “morphologist three.”

To create the character-cladograms the photographs were laid on a table in haphazard order and sorted hierarchically into dichotomous groups by the following procedure, which we specified. Each participant subdivided the photographs into two groups whose members they judged most similar to each other. Each subgroup was then divided into two groups, and so on until the smallest groups contained only one or two photographs. A record was kept of group membership at each stage of this process. The results were 13 unrooted character-cladograms representing the opinions of the 13 participants on the relationships among the taxa based on these photographs (Fig. 1).

We used three methods to investigate inter-investigator agreement (reliability). Estabrook & al.’s (1985) and Day’s (1986) Quartet Symmetric Difference (QSD, a tree distance dissimilarity measure) as implemented in Component (Page, 1993) was used to perform pair-wise comparisons between subjects in a test group. We also converted the character-cladograms from each test group into a single data matrix using Matrix Representation with Parsimony (MRP; Farris, 1973), and performed a bootstrap analysis of the MRP matrix with a random addition sequence and 200 bootstrap replications (PAUP\* 4.0 beta; Swofford, 2002). The bootstrap values, and Rohlf’s (1982)  $CI_2$  index as implemented in PAUP\* (Swofford, 2002), were used as measures of the agreement of the 50% majority rule consensus trees with the character-cladograms. Rohlf’s (1982)  $CI_2$  is a dissimilarity measure based on the number of fully resolved trees that are consistent with a consensus tree.

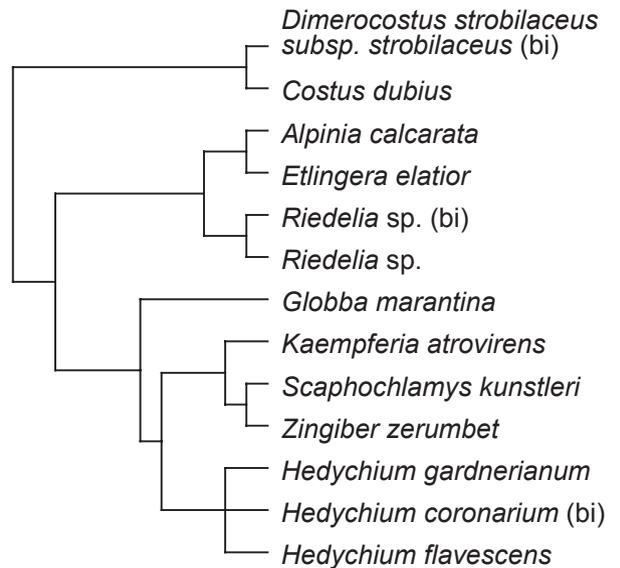
In order to investigate agreement with an independent assessment of phylogeny (validity) we used a phylogenetic tree based on Kress & al.’s (2002) independent assessment of the phylogenetic relationships in Zingiberaceae, with the two members of Costaceae placed as the sister group to the Zingiberaceae (Fig. 3). This placement of the Costaceae is based on published phylogenetic analyses of the order (Dahlgren & Rasmussen, 1983; Kirchoff, 1988; Kress, 1990, 1995; Kress & al., 2001). Similarity between the character-cladograms and this independent estimate of the phylogeny was measured both with QSD, and by converting the character-cladograms into MRP matrices and optimizing each matrix on the model phylogeny (Fig. 3) using the Consistency (CI) and Retention (RI) Indices of Kluge & Farris (1969) and Farris (1989).

Means and medians of the comparison measures are used to summarize the data (Tables 2, 3). All statistical analyses used non-parametric Wilcoxon rank-sum tests, as we had no reason to suspect that the data would be normally distributed. One-sided tests were used to test specific hypotheses about participant responses. For all

measures we expected trained to perform better than untrained participants, and morphologists to perform better than untrained students. For some analyses we combined the morphologists with the untrained students in order to compare the participants who were trained by us, with all of those who were not. Statistical tests on  $CI_2$  are not possible because  $CI_2$  is computed based on a single consensus tree for each class of subjects.

**Visual processing mode of the participants.**

— Direct determination of the visual processing mode used by the participants is possible (Gauthier & Tarr, 1997; Gauthier & al., 1998), but was beyond the scope of this study. An indirect assessment of these modes is



**Fig. 3. Phylogenetic relationship of the study taxa from Kress & al. (2002), with the two Costaceae (*Dimerocostus strobilaceus subsp. strobilaceus*, *Costus dubius*) placed as the sister-group to the Zingiberaceae; (bi), bilocular flower, all others trilocular.**

**Table 2. Reliability of the character-cladograms.**

	QSD <sup>a</sup>		Rohlf's $CI_2$ <sup>a</sup>
	Mean/median	$p^b$ $p$	
Morphologists	0.54/0.60		$3.27 \times 10^{-9}$
Untrained Students	0.59/0.61	0.45	$2.06 \times 10^{-7}$
Trained Students	0.50/0.58	0.07	$6.55 \times 10^{-10}$
Untrained Participants <sup>c</sup>	0.57/0.60		— <sup>d</sup>

<sup>a</sup>QSD and  $CI_2$  are dissimilarity metrics. QSD is a tree distance metric.

<sup>b</sup>All p-values are one sided.

<sup>c</sup>Untrained Participants = Morphologists + Untrained Students

<sup>d</sup> $CI_2$  was not calculated for all Untrained Participants because there is no bootstrap tree for this group.

possible by looking at the placement of congeneric species in the 50% majority rule Consensus Trees (Fig. 4), and by listening to, and sometimes soliciting, comments from the participants after they had completed their sorts. We did not conduct formal participant interviews, partly because this was a preliminary study intended only to demonstrate the concepts, and partly because participants typically cannot easily or accurately report how they performed a task when stimuli are processed holistically (e.g., Biederman & Shiffrar, 1987).

## RESULTS

For the measures of reliability (QSD and Rohlf's  $CI_2$ ), untrained students were the most dissimilar, followed by morphologists and finally trained students (Table 2). This suggests that training designed to promote holistic processing leads to more consistent performance. For QSD used as a measure of validity (i.e., in comparisons with the model tree), morphologists scored highest (most dissimilar), followed by untrained students and then trained students (Table 3). For the two measures of agreement between the character-cladograms and the model tree (CI and RI) trained students had the highest agreement, followed by morphologists and finally untrained students (Table 3). These results also suggest an improvement due to familiarization with the stimuli.

Statistical analyses provide moderate support for the idea that training improves performance (Tables 2, 3). For QSD used as a measure of reliability, there is moderate support for the supposition that the trained students had greater inter-investigator agreement than both the untrained students, and all of the untrained participants (untrained students + morphologists) (Table 2). There is no support for the contention that morphologists performed better than untrained students (Table 2).

A more detailed analysis of this result is warranted because two of the morphologists identified analytic criteria as the basis for their sorts (see below), whereas

the responses of the third suggest that he used more configural criteria. Pairwise comparisons between the three morphologists show QSD = 0.31 for the analytic pair, whereas comparisons between these two and the third morphologist produced QSD's of 0.60 and 0.70. These results support the suggestion that the third morphologist was using different sorting criteria.

Turning to the measures of validity, there is moderate statistical support from QSD and CI that trained students produced character-cladograms that agreed better with the model tree than did those of the untrained participants (untrained students + morphologists), but less convincing support for this supposition for RI (Table 3). Comparisons between trained and untrained students show moderate support for better agreement as measured by CI, but less convincing support from QSD and RI (Table 3). There is no statistical evidence that morphologists produced more valid character-cladograms than untrained students.

A slightly different picture emerges if we consider the responses of the three morphologists individually. For RI the two morphologists who identified analytic criteria for their sorts both produced trees with RI = 0. Given this, their CI scores (0.513, 0.480) must be attributed to autapomorphy, which is not informative of relationships. The morphologist whose criteria were more holistic produced CI = 0.67 and RI = 0.37, indicating that he identified some groups that agreed with the model tree. His RI was higher than all of those produced by untrained students (0.25, 0.08, 0.00, 0.00), and higher than all but those of two of the trained students (0.41, 0.39, 0.29, 0.23, 0.00, 0.00).

QSD presents a similar pattern. Morphologist three produced a character cladogram that had a lower QSD (0.57) than his colleagues (0.76, 0.74), and of all but one of the untrained students (0.72, 0.67, 0.67, 0.54). It compares favorably to the mean of the trained students (0.58; Table 2), and is lower than or equal to three of these measures (0.73, 0.64, 0.57, 0.54, 0.52, 0.49).

To further explore the idea that training improves agreement with an independent assessment of phylogeny, one untrained student was given nine hours of training

**Table 3. Validity of the character-cladograms.**

	QSD <sup>a</sup>			CI			RI		
	Mean/median	<i>p</i> <sup>b</sup>	<i>p</i>	Mean/median	<i>p</i>	<i>p</i>	Mean/median	<i>p</i>	<i>p</i>
Morphologists	0.69/0.74	0.20	0.18	0.55/0.51	0.31	0.04	0.12/0.00	0.57	0.16
Untrained Students	0.65/0.67			0.51/0.50			0.08/0.04		
Trained Students	0.58/0.51	0.06	0.58/0.59	0.07	0.22/0.26	0.13			
Untrained Participants <sup>c</sup>	0.67/0.67		0.53/0.50		0.10/0.00				

<sup>a</sup>QSD is a dissimilarity, tree distance metric.

<sup>b</sup>All *p*-values are one sided.

<sup>c</sup>Untrained Participants = Morphologists + Untrained Students

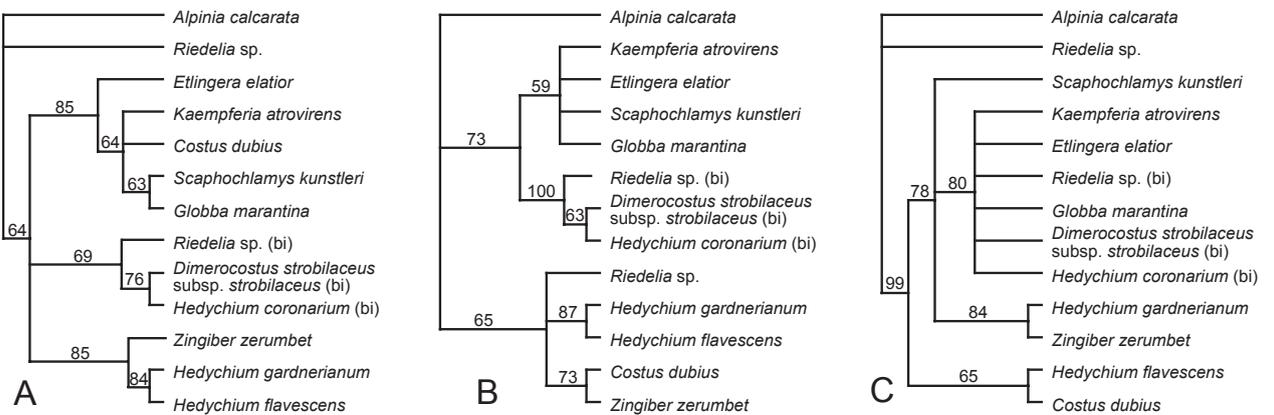
with the photographs, and then was asked to sort them again. His second sort was quite dissimilar to his first (QSD=0.73), and was much more similar to the independent assessment of phylogeny than any of the other sorts (QSD=0.26 compared to a mean of 0.70 for the morphologists, 0.65 for untrained students, and 0.58 for trained students).

**Visual processing mode of the participants; results of participant interviews.** — Two of the morphologists were eager to explain their criteria for sorting the photographs. These criteria were primarily analytic (septa/no septa; three/two carpels; incomplete/complete septa; thin/thick septum; two/more than two ovules). These subjects used one criterion for each dichotomous division of the photographs. One of the morphologists actually wrote out all of his criteria on his character-cladogram without being prompted to do so. Morphologist three (the one with previous experience with the stimuli) offered his criteria in the form of a short narrative, only after gentle prompting. “Initially looked at all; several individuals XS’s [cross sections] didn’t relate to others; these were separated out initially. Then groups of ‘similar’ XS’s were put together based on ovary wall structure, ovule size and shape, integument thickness/etc., I considered locule number some.” These criteria are mainly configural. None of the morphologists mentioned color or size of the photograph as important characteristics, though ovule size and integument thickness do clearly depend on the overall size of the photographs.

In general, neither the trained nor the untrained students offered any spontaneous comments on why they had made their decisions. At the end of the experiments we asked the student who had been trained for nine hours to explain his reasons for sorting the stimuli (Fig. 1). Several interesting facts emerged from this discussion.

In general, he did not restrict his comments to any one level of the sort, but made general comments on the types of features that he looked at throughout the process. After several attempts to get him to be more specific, he pointed to several features that he had used. The first of these was the distribution of the tanniferous idioblasts, structures that he had never been exposed to and did not know. What he said was (approximately) “See those black dots? They have different arrangements in the different pictures.” The observation that the distribution of tanniferous idioblasts differs in the cross sections is very sophisticated. It is not a character regularly used in phylogenetic analyses, and is not one mentioned by any of the morphologists. When asked if he had then based his decisions on the distribution of the tanniferous idioblasts, he replied (again approximately) “No, not exactly.” He went on to explain that he had used a number of features, and relationships between features (ovary wall thickness relative to septa thickness, for instance) to make his decisions. This identification of configural features is suggestive of a subject who was sorting based holistic visual processing.

An analysis of the placement of congeneric photographs in the 50% majority rule consensus tree supports the supposition that the trained students were using holistic processing, whereas the morphologists and untrained students were analytical processors (Fig. 4). The two ovaries of *Riedelia* sp. come from the same collection (Table 1), and differ primarily in the number of locules (Fig. 1). Participants using analytic processing should be more likely to place the bilocular *Riedelia* with other bilocular species, whereas those using holistic processing should be more likely to place the two photographs of *Riedelia* together. The consensus trees for both the trained students and the morphologists possess bilocular clades (*Riedelia*



**Fig. 4. A–C, 50% majority rule (unrooted) bootstrap trees from the three participant groups. A, consensus of 366 bootstrap trees for the trained students; B, consensus of 173 bootstrap trees for the morphologists; C, consensus of 698 trees for the untrained students. (bi), bilocular flower, all others trilocular.**

[bi], *Dimerocostus*, *Hedychium coronarium*), whereas the untrained students were less consistent in their placement of taxa. This is reflected in the large polytomy in the untrained students' consensus tree (Fig. 4). The support for the bilocular clades is weak in the trained students' tree (69%) and perfect (100%) in the morphologists tree, indicating that the trained students were less likely to place the bilocular species together. An analysis of the placement of the *Riedelia* photographs in the individual character-cladograms reveals that 25% of the untrained students placed the *Riedelia* sections together in a group of two, whereas 33% of the morphologists and 50% of the trained students did the same. These results support the contention that the trained students were more likely to view the photographs holistically than either of the untrained groups (untrained students or morphologists). This is consistent with the hypothesis that the trained students were using holistic processing.

A similar analysis can be done for the three species of *Hedychium* included in the study. The photographs of both *H. flavescens* and *H. gardnerianum* are of trilocular ovaries (Fig. 1). That of *H. coronarium* is of a bilocular variant of an otherwise trilocular species. The consensus trees (Fig. 4) show that the trained students frequently, but not exclusively, placed *H. coronarium* with the other bilocular species. On the other hand the morphologists always placed the three bilocular species together, but they differed in whether *H. coronarium* was grouped with *Riedelia* (bi) or *Dimerocostus*. The consensus tree for the untrained students indicates that they had greater variation in their placement of the species of *Hedychium*. The three species occur together only as part of a six-species polytomy (Fig. 4C).

## DISCUSSION

**Training, visual processing mode, and good characters.** — The experiments described here provide preliminary evidence that subjects are able to describe character states with higher reliability and validity when they are trained to view the data holistically. Untrained students, and morphologists who have no prior experience with the stimuli, appear to sort them based on analytical characteristics. Their sorts tend to have lower inter-investigator agreement, and agree less well with an independent assessment of phylogeny.

The lack of sizable differences among the comparison measures (Tables 2, 3) for the trained and untrained groups is most likely due to the short training time and the ease of the training task (Isabel Gauthier, pers. comm.). This suggests that repeating the experiments with longer training times will lead to more substantial differences between the trained and untrained groups.

Comparing the morphologists with the untrained students, we found no statistical support for the supposition that the morphologists produced character-cladograms with higher validity than untrained students. There was mixed evidence for the supposition that their character-cladograms were more reliable.

The lack of a large difference between these groups is more striking given the fact that the morphologists had access to information on the relative size of the ovaries whereas the untrained students did not. Based on their comments at the end of the experiments, it seems likely that the morphologists ignored both the size and color of the photographs. Their training as disciplinary experts may have led them to disregard these aspects. The color of stained anatomical sections is notoriously variable, and the structure and arrangement of tissues rather than their size are generally the foci of anatomical studies.

It seems likely that the morphologists would have performed better if they had greater experience with floral anatomy. The holistic processing literature suggests that only training in the specific stimulus domain leads to holistic processing (Gauthier & Tarr, 2002). Disciplinary experts are expected to be analytic processors in stimulus domains with which they have little experience. None of the morphologists were experts in floral anatomy. Whether greater experience in floral anatomy would have allowed them to process the photographs holistically remains uncertain, though the results of morphologist three are suggestive. If his improved performance can be attributed to his brief (ca. ½ hour) exposure to the photographs prior to his sort, it suggests that the time it takes to train a disciplinary expert to be a holistic processor may be significantly under the nine hours needed to train a novice. Additional studies with more morphologists are needed to test this hypothesis.

**Definitions: characters, character states and parts.** — The results presented here suggest the possibility of modifying the definitions of the character and character state to bring them more into line with the operational procedures we propose. The most commonly used definition of character (any feature that is shared among organisms: Hennig, 1966; Davis & Heywood, 1973; Colless, 1985; Stuessy, 1990; Fristrup, 1992, 2001) and character state (one or more alternative manifestations of a character: Michener & Sokal, 1957: 137; Colless, 1985; Kitching & al., 1998: 201) give little attention to the fact that different processes are involved in describing characters and character states. Characters are, or at least should be, described based on homology assessment. Character states are based on similarity assessments within the set of homologous features, i.e., within a character. With these factors in mind, we propose the following provisional definitions.

**Morphological character:** A group of parts (one per taxon, or operational taxonomic unit) that have been

judged homologous to each other through the application of one or more homology criteria.

**Morphological character state:** A subgroup of parts belonging to a single character. Character states are described through the explicit comparison of homologous parts, and result in clusters of similar parts (character-cladograms), which may be hierarchical.

Essential to these definitions is a clear understanding of the structure of a part. A part is an internally integrated system that is isolated from its surround (McShea 2001; McShea & Venit 2001), where the surround can be thought of as the next higher level in the hierarchy of homology. A part must have both lower level constituents, and a higher level context (Salthe, 1985). Parts are often hierarchically nested. Thus, flowers are parts that can be decomposed into the lower-level parts called sepals, petals, etc. Petals are parts but, except in rare cases, they do not have parts. Any division of a petal will produce a feature that is either not isolated from its surround, or does not have lower-level constituents.

A few additional examples will help clarify the nature of parts. A portion of an ovary wall (Fig. 2C) is not a part because its higher-level boundary conditions are not well defined. The lateral margins of the portion are arbitrary. Similar arguments apply to the septa (Fig. 2D). Both of these aspects can inform character state description at the next higher level, the level of the whole cross section (Fig. 1). The whole plant is also not a part, in this case because it has no surround.

While these definitions do not stress the need to use complex parts, prior results suggest that character states described based on complex parts will be more reliable and valid than those based on simple parts (Kirchoff & al., 2004). The use of complex stimuli is also a prerequisite for holistic processing.

We recognize that these definitions of character and character state will not be applicable to all cases. We are certainly not suggesting that all use of conventional characters be abandoned. We are suggesting that further experiments be undertaken with complex characters, including experiments with their use in phylogeny reconstruction. The definitions given above are intended to be a step in a reconceptualization of these important terms.

**Operational definitions of character and character state.** — We suggest the following method for creating characters and character states. (1) *Taxa selection:* The more precisely the taxa can be specified, the more accurate and meaningful the characters will be. For instance, if a character is intended to be valid across a family, sufficient species must be chosen to sample the variability expressed in that family. (2) *Homology assessment:* Identify the homologies of the parts using well established criteria for homology assessment (Remane, 1952; Platnick, 1979; Patterson, 1982; Pimentel &

Riggins, 1987; de Pinna, 1991; Brower & Schawaroch, 1996). If homologies cannot be established across the study taxa, either enlarge your concept of the part, or reduce the size of the study group until unambiguous homologies can be established. The attempt to establish homologies when comparable features are missing from some taxa is an indication that the features may be too narrowly (or too broadly) defined. For instance, if we were interested in the relationships of the families of the Fagales (Soltis & al., 2000; Hufford, 1992), then evaluating them on the characteristics of their petals would exclude meaningful comparison with the Casuarinaceae and Myricaceae, which lack petals. Broadening part definition to the level of the whole flower and representing the parts pictorially, allows the inclusion of all families. The absence of petals is one way in which the flower can be structured. The whole configuration of the flower may suggest relationships not implied by describing a character based on absence of petals. (3) *Character description:* The accurate assessment of homology leads to the description of characters: collections of the homologous parts across the study taxa. At this point in the process a character has no character states. (4) *Pictorial representation of the variability in the character:* Once characters have been established, use photographic or other pictorial methods to represent the variability in each character. This variability has two components: parts vary across taxa; and parts often have internal structure. Create explicit representations of the variability, comparable to the photographs used in the experiments described above (Fig. 2). For simplicity, we will refer to these representations as “photographs.” (5) *Character state description:* Lay the photographs out on a table and sort them dichotomously into character states to create character-cladograms. To assure reliable and valid character states the person who does the sorts should be unfamiliar with the taxa in question, but should be trained on the stimuli for a sufficient amount of time to allow him to process the photographs holistically. Increased confidence in the character states can be obtained by having three to five independent investigators sort the photographs, then summarizing their results as a bootstrap or jackknife tree. Nodes with weak bootstrap support (< 63%; Farris & al., 1996) can be collapsed. (6) *Character coding:* The resulting character-cladogram can be coded using Matrix Representation with Parsimony or as stepmatrices, and used in phylogenetic analyses. PHYLIP (Felsenstein, 2004) provides options for treating the whole MRP representation of a character-cladogram as a single character. In PAUP\* (Swofford, 2002) each column of the MRP representation must be treated as an independent character.

## CONCLUSIONS

Our results provide preliminary evidence that investigators can be trained to describe complex morphological characters, and that these characters may be more reliable and valid than those produced without training. These results, coupled with knowledge gained from cognitive psychology, suggest a method for character description, the reliability and validity of which can now be formally tested relative to that of traditional approaches. If results of more formal tests are successful, these methods will provide a valuable tool for morphological systematics, and will help resolve long-standing problems with identifying and representing characters and character states.

## ACKNOWLEDGMENTS

The authors thank Lucinda McDade, Chelsea Specht, Norm MacLeod, Rod Page, Peter Forey, and Rolf Rutishauser for their comments on the manuscript. Ed Wisniewski, Isabel Gauthier, and Jim Tanaka provided consultation on various aspects of cognitive psychology and experimental design. All responsibility for the opinions expressed in the paper remains the full responsibility of the authors.

## LITERATURE CITED

- Biederman, I. & Shiffrar, M.M.** 1987. Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual-learning task. *J. Exp. Psychol. Learn. Mem. Cogn.* 13: 640–645.
- Brower, A.V.Z. & Schawaroch, V.** 1996. Three steps of homology assessment. *Cladistics* 12: 265–272.
- Carey, S. & Diamond, R.** 1977. From piecemeal to configurational representation of faces. *Science* 195: 312–314.
- Colless, D.** 1985. On “character” and related terms. *Syst. Zool.* 34: 229–233.
- Dahlgren, R.M.T. & Rasmussen, F.** 1983. Monocotyledon evolution: characters and phylogenetic estimation. *Evol. Biol.* 16: 255–395.
- Davis, P.H. & Heywood, V.** 1973. *Principles of Angiosperm Taxonomy*. Van Nostrand, Princeton.
- Day, W.H.E.** 1986. Analysis of quartet dissimilarity measures between undirected phylogenetic trees. *Syst. Zool.* 35: 325–333.
- De Pinna, M.C.C.** 1991. Concepts and tests of homology in the cladistic paradigm. *Cladistics* 7: 367–394.
- Diamond, R. & Carey, S.** 1986. Why faces are and are not special: an effect of expertise. *J. Exp. Psychol. Gener.* 115: 107–117.
- Estabrook, G.F., McMorris, F.R. & Meacham, C.A.** 1985. Comparison of undirected phylogenetic trees based on subtrees of four evolutionary units. *Syst. Zool.* 34: 193–200.
- Farah, M.J., Tanaka, J.N. & Drain, M.** 1995. What causes the face inversion effect? *J. Exp. Psychol. Human Percept. Perform.* 21: 628–634.
- Farah, M.J., Wilson, K.D., Drain, M. & Tanaka, J.N.** 1998. What is special about face perception? *Psychol. Rev.* 105: 482–498.
- Farris, J.S.** 1973. On comparing the shapes of taxonomic trees. *Syst. Zool.* 22: 50–54.
- Farris, J.S.** 1989. The retention index and rescaled consistency index. *Cladistics* 5: 417–419.
- Farris, J.S., Albert, V.A., Källersjö, M., Lipscomb, D. & Kluge, A.G.** 1996. Parsimony jackknifing outperforms neighbor-joining. *Cladistics* 12: 99–124.
- Felsenstein, J.** 2004. *PHYLIP (Phylogeny Inference Package)*, vers. 3.6. Department of Genome Sciences, University of Washington, Seattle. (<http://evolution.genetics.washington.edu/phylip/getme.html>)
- Freitas, A.V.L. & Brown, K.S., Jr.** 2004. Phylogeny of the Nymphalidae (Lepidoptera). *Syst. Biol.* 53: 363–383.
- Fristrup, K.** 1992. Character: current usages. Pp. 45–51 in: Keller, E.F. & Lloyd, E.A. (eds.), *Keywords in Evolutionary Biology*. Harvard University Press, Cambridge.
- Fristrup, K.** 2001. A history of character concept in evolutionary biology. Pp. 13–35 in: Wagner, G. (ed.), *The Character Concept in Evolutionary Biology*. Academic Press, San Diego.
- Gauthier, I., Curran, T., Curby, K.M. & Collins, D.** 2003. Perceptual interference supports a non-modular account of face processing. *Nature Neurosci.* 6: 428–432.
- Gauthier, I. & Tarr, M.J.** 1997. Becoming a “greeble” expert: Exploring mechanisms of face recognition. *Vision Res.* 37: 1673–1682.
- Gauthier, I. & Tarr, M.J.** 2002. Unraveling mechanisms for expert object recognition: bridging brain activity and behavior. *J. Exp. Psychol. Human Percept. Perform.* 28: 431–436.
- Gauthier, I., Williams, P., Tarr, M.J. & Tanaka, J.** 1998. Training ‘greeble’ experts: A framework for studying expert object recognition processes. *Vision Res.* 38: 2401–2428.
- Gift, N. & Stevens, P.F.** 1997. Vagaries in the delimitation of character states in quantitative variation—an experimental study. *Syst. Biol.* 46: 112–125.
- Goldstone, R.L.** 2000. Unitization during category learning. *J. Exp. Psychol. Human Percept. Perform.* 26: 86–112.
- Guerrero, J.A., De Luna, E. & Sánchez-Hernández, C.** 2003. Morphometrics in the quantification of character state identity for the assessment of primary homology: an analysis of character variation of the genus *Artibeus* (Chiroptera: Phyllostomidae). *Biol. J. Linn. Soc.* 80: 45–55.
- Hedenäs, L.** 2002. Important complexes of intercorrelated character states in pleurocarpous mosses. *Lindbergia* 27: 104–121.
- Hennig, W.** 1966. *Phylogenetic Systematics*. University of Illinois Press, Urbana.
- Holland, B.R., Huber, K.T., Moulton, V. & Lockhart, P.J.** 2004. Using consensus networks to visualize contradictory evidence for species phylogeny. *Molec. Biol. Evol.* 21: 1459–1461.
- Hufford, L.** 1992. Rosidae and their relationships to other nonmagnoliid dicotyledons: a phylogenetic analysis using morphological and chemical data. *Ann. Missouri Bot. Gard.* 79: 218–248.
- Jacoby, L.L. & Brooks, L.R.** 1984. Nonanalytic cognition: memory, perception, and concept learning. *Psychol. Learn. Motiv.* 18: 1–47.

- Jenner, R.A.** 2004. Accepting partnership by submission? Morphological phylogenetics in a molecular millennium. *Syst. Biol.* 53: 333–342.
- Judd, W.S., Campbell C.S., Kellogg, E.A., Stevens, P.F. & Donoghue, M.J.** 2002. *Plant Systematics a Phylogenetic Approach*, 2nd ed. Sinauer Associates, Sunderland.
- Kangas, A.T., Evans, A.R., Thesleff, I. & Jernvall, J.** 2004. Nonindependence of mammalian dental characters. *Nature* 432: 211–214.
- Kemp, T.S.** 1999. *Fossils and Evolution*. Oxford University Press, Oxford.
- Kirchoff, B.K.** 1988. Floral ontogeny and evolution in the ginger group of the Zingiberales. Pp. 45–56 in: Leins, P., Tucker, S.C. & Endress, P.K. (eds.), *Aspects of Floral Development*. J. Cramer, Berlin.
- Kirchoff, B.K., Richter, S.J., Remington, D.L. & Wisniewski, E.** 2004. Complex data produce better characters. *Syst. Biol.* 53: 1–17.
- Kitching, I.J., Forey, P.L., Humphries, C.J. & Williams, D.M.** 1998. *Cladistics*, 2nd ed. Oxford University Press, Oxford.
- Kluge, A.G. & Farris, J.S.** 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* 18: 1–32.
- Kress, W.J.** 1990. The phylogeny and classification of the Zingiberales. *Ann. Missouri Bot. Gard.* 77: 698–721.
- Kress, W.J.** 1995. Phylogeny of the Zingiberales: Morphology and molecules. Pp. 443–460 in: Rudall, P.J., Cribb, P.K., Cutler, D.F. & Humphries, C.J. (eds.), *Monocotyledons: Systematics and Evolution*. Royal Botanic Gardens, Kew.
- Kress, W.J., Price, L.M., Hann, W.J. & Zimmer, E.Z.** 2001. Unraveling the evolutionary radiation of the families of the Zingiberales using morphological and molecular evidence. *Syst. Biol.* 50: 926–944.
- Kress, W.J., Price, L.M. & Williams, K.J.** 2002. The phylogeny and a new classification of the gingers (Zingiberaceae): evidence from molecular data. *Amer. J. Bot.* 89: 1682–1696.
- Landa, E.R. & Fairchild, M.D.** 2005. Charting color from the eye of the beholder. *Amer. Sci.* 93: 436–443.
- Lee, M.S.Y.** 2006. Morphological phylogenies and the universe of useful characters. *Taxon* 55: 5–7.
- Mast, A. & Givnish, T.J.** 2002. Historical biogeography and the origin of stomatal distributions in *Banksia* and *Dryandra* (Proteaceae) based on their cpDNA phylogeny. *Amer. J. Bot.* 89: 1311–1323.
- McShea, D.W.** 2001. Parts and integration: Consequences of hierarchy. Pp. 27–60 in: Jackson, J.B.C., Lidgard, S. & McKinney, F.K. (eds.), *Evolutionary Patterns: Growth, Form and Tempo in the Fossil Record*. University of Chicago Press, Chicago.
- McShea, D.W. & Venit, E.P.** 2001. What is a part? Pp. 259–284 in: Wagner, G. (ed.), *The Character Concept in Evolutionary Biology*. Academic Press, San Diego.
- Michener, C.D. & Sokal, R.R.** 1957. A quantitative approach to a problem in classification. *Evolution* 11: 130–162.
- Mishler, B.D.** 2005. The logic of the data matrix in phylogenetic analysis. 2005. Pp. 57–70, 201–217 in: Albert, V.A. (ed.), *Parsimony, Phylogeny, and Genomics*. Oxford University Press, Oxford.
- Miyamoto, Y., Nisbett, R.E. & Masuda, T.** 2006. Culture and the physical environment: holistic versus analytic perceptual affordances. *Psychol. Sci.* 17: 113–119.
- Neff, N.A.** 1986. A rational basis for a priori character weighting. *Syst. Zool.* 35: 110–123.
- Newman, S.W. & Kirchoff, B.K.** 1992. Ovary structure in the Costaceae (Zingiberaceae). *Int. J. Pl. Sci.* 153: 471–487.
- Niku, M. & Taipale, M.** 2003. *Clever Gene Names*. <http://tinman.vetmed.helsinki.fi/index.html>
- Nisbett, R.E. & Miyamoto, Y.** 2005. The influence of culture: holistic versus analytic perception. *Trends Cogn. Sci.* 9: 468–473.
- Page, R.D.M.** 1993. *COMPONENT: Tree Comparison Software for Microsoft Windows*, vers. 2.0. Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow. (<http://taxonomy.zoology.gla.ac.uk/rod/cpw.html>)
- Patterson, C.** 1982. Morphological characters and homology. Pp. 20–74 in: Joysey, K.A. & Friday, A.E. (eds.), *Problems of Phylogenetic Reconstruction*. Academic Press, London.
- Pimentel, R.A. & Riggins, R.** 1987. The nature of cladistic data. *Cladistics* 3: 201–209.
- Platnick, N.I.** 1979. Philosophy and the transformation of cladistics. *Syst. Zool.* 28: 537–546.
- Remane, A.** 1952. *Die Grundlagen des natürlichen Systems der vergleichenden Anatomie und der Phylogenetik*. Gesst und Portig, Leipzig.
- Richards, R.** 2003. Character individuation in phylogenetic inference. *Philos. Sci.* 70: 264–279.
- Riedl, R.** 1978. *Order in Living Organisms*. Wiley, Chichester, New York.
- Rohlf, F.J.** 1982. Consensus indices for comparing classifications. *Math. Biosci.* 59: 131–144.
- Rudall, P.J., Stobart, K.L., Hong, W.-P., Conran, J.G., Furness, C.A., Kite, C.G. & Chase, M.W.** 2000. Consider the lilies: systematics of the Liliales. Pp. 347–359 in: Wilson, K.L. & Morrison, D.A. (eds.), *Monocots: Systematics and Evolution*. CSIRO Publishing, Collingwood.
- Salthe, S.N.** 1985. *Evolving Hierarchical Systems*. Columbia University Press, New York.
- Scapinello, K.F. & Yarmey, A.D.** 1970. Role of familiarity and orientation in immediate and delayed recognition of pictorial stimuli. *Psychonomic Sci.* 21: 329–331.
- Scotland, R., Olmstead, R.G. & Bennett, J.R.** 2003. Phylogeny reconstruction: the role of morphology. *Syst. Biol.* 52: 539–548.
- Smith, N.D. & Turner, A.H.** 2005. Morphology's role in phylogeny reconstruction: perspectives from paleontology. *Syst. Biol.* 54: 166–173.
- Soltis, D.E., Soltis, P.S., Chase, M.W., Mort, M.E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Prince, L.M., Kress, J.W., Nixon, K.C. & Farris, J.S.** 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcl*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.
- Stuessy, T.F.** 1990. *Plant Taxonomy: The Systematic Evaluation of Comparative Data*. Columbia University Press, New York, New York.
- Swofford, D.L.** 2002. *PAUP\*: Phylogenetic Analysis Using Parsimony (and Other Methods)*, vers. 4.0b10 Sinauer Associates, Sunderland.
- Swofford, D.L. & Begle, D.P.** 1993. *User Manual for PAUP: Phylogenetic Analysis Using Parsimony, Version 3.1*. Smithsonian Institution, Washington, D.C.

- Tanaka, J.W. & Senge, J.A.** 1997. Features and their configuration in face recognition. *Memory Cogn.* 25: 583–592.
- Thiele, K. & Ladiges, P.Y.** 1996. A cladistic analysis of *Banksia* (Proteaceae). *Austral. Syst. Bot.* 9: 661–733.
- Wagner, G.** 1998. Complexity matters. *Science* 279: 1158–1159.
- Wagner, G.** 2001. *The Character Concept in Evolutionary Biology*. Academic Press, San Diego.
- Waxman, D. & Peck, J.R.** 1998. Pleiotropy and the preservation of perfection. *Science* 279: 1210–1213.
- Wells, G.L. & Hryciw, B.** 1984. Memory for faces: encoding and retrieval operations. *Memory Cogn.* 12: 338–244.
- Wiens, J.J.** 2004. The role of morphological data in phylogeny reconstruction. *Syst. Biol.* 53: 653–661.
- Wilkinson, M., McInerney, J.O., Hirt, R.P., Foster, P.G. & Embley, T.M.** 2007. Of clades and clans: terms for phylogenetic relationships in unrooted trees. *Trends Ecol. Evol.* 22: 114–115.
- Zebrafish Nomenclature Committee.** 2006. *Zebrafish Nomenclature Guidelines*. University of Oregon, Eugene. ([http://zfin.org/zf\\_info/nomen.html](http://zfin.org/zf_info/nomen.html))