

A comparison of information-theoretically optimal classifications with estimates of phylogenetic relationships in the squid family Cranchiidae (Cephalopoda: Oegopsida)

By R. S. VOSS, G. F. ESTABROOK and NANCY A. VOSS

Received on 23. August 1982

Introduction

Criteria by which some biological classifications may be judged better or worse than others have been debated for centuries, but since 1859 most systematists have striven to create taxonomies that reflect genealogical relationships, and the familiarity of such terms as "monophyletic", "paraphyletic" and "polyphyletic" reflects a widespread concern about the nature of correspondences between classifications and phylogenies. Some systematists, however, have expressed pessimism about the ability of modern researchers to discover true phylogenies, and question the utility of phylogenetic classifications even were it possible to ascertain their historical accuracy. J. S. L. GILMOUR, an early and articulate critic of evolutionary taxonomy, proposed that biological classifications be comprised exclusively of "... natural groups [that] class together individuals which have a large number of attributes in common" (GILMOUR 1940: 466). The utility of natural groups, so defined, is usually equated with the information about characters conveyed by knowledge of group memberships (SOKAL and SNEATH 1963; FARRIS 1979).

That classifications should (1) correspond to known or inferred phylogenies and (2) be maximally informative about characters are both intuitively appealing ideal properties, and it is therefore of interest to inquire whether or not they need be viewed irreconcilable as optimality criteria. In this paper we define ways in which classifications can correspond to tree diagrams of evolutionary history, propose quantitative measures of taxonomic information content and examine the relationships between phylogenetic reconstructions and informative classifications of pelagic squids in the family Cranchiidae.

Classifications and phylogenies

The (finite) collection of biological entities to be classified is called the Study, S . A level classification of S is any partition, P_k , of S into subsets such that the union of the subsets exhausts the membership of S and the intersection of any two subsets is the empty set. Thus, every member of S belongs to one and only one of the subsets comprising P_k , and these subsets are called the classes of P_k .

By a phylogeny of S is meant a conventional tree diagram on which the members of S occupy the terminal nodes; internal nodes represent shared ancestors, and the lines connecting nodes represent phyletic lineages. A phylogeny and three level classifications for a study with five member taxa are illustrated in Figure 1.

A level classification of S corresponds to a phylogeny of S by monophyly (is said to be monophyletic) if and only if each of its classes contains all and only the descendants of a single common ancestor. Classification P_1 of Figure 1 corresponds to the accompanying

U. S. Copyright Clearance Center Code Statement: 0044-3808/83/2102-0081/\$ 2.50/0

Z. zool. Syst. Evolut.-forsch. 21 (1983) 81-95

© 1983 Verlag Paul Parey, Hamburg und Berlin

ISSN 0044-3808/InterCode: ZZSEAA

phylogeny by monophyly, but classification P_2 does not because the class $P_{21} = (A, B, C)$ does not contain all and only the descendants of any common ancestor.

A level classification of S corresponds to a phylogeny of S by convexity (is said to be convex) if and only if by inclusion of appropriate ancestors and phyletic lines all of its classes

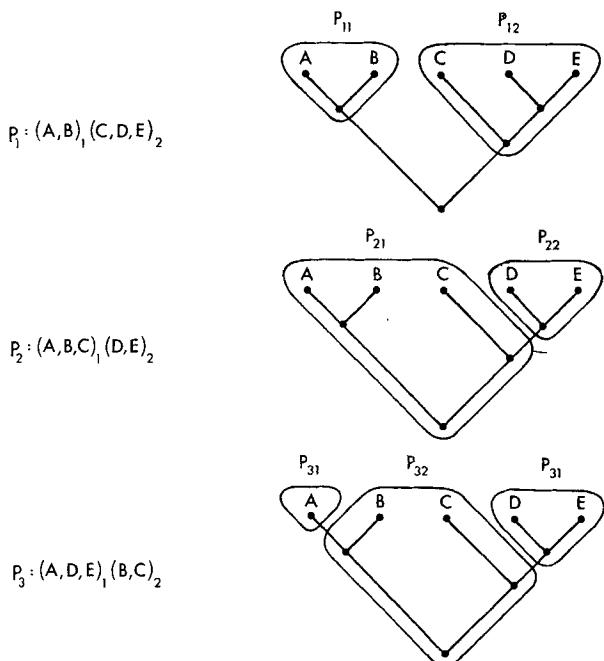


Fig. 1. Three alternative level classifications for a study collection with five member taxa

can be simultaneously construed as the result of uninterrupted descent from a single common ancestor each. Thus, classification P_2 (Figure 1), while not monophyletic, is convex, for it is possible to enlarge its classes such that both can be construed, simultaneously, as descended each from a single ancestor along uninterrupted phyletic lines. By contrast, classification P_3 does not correspond to the accompanying phylogeny (is neither monophyletic nor convex) because it is impossible to construe both of its classes at once as continuous entities on the adjacent tree diagram – at least one class must have arisen twice in the evolutionary history of the study collection.

Obviously, correspondence by monophyly is only a special case of correspondence by convexity (all monophyletic partitions are convex, but not all convex partitions are monophyletic). Systematists have sometimes used “paraphyletic” and “polyphyletic” to mean “convex but not monophyletic” and “not convex”, respectively, but these older terms are usually employed as descriptors of individual groups (classes) of taxa. We emphasize that convex correspondence, or the lack thereof, is here considered a property of level classifications, not of groups (classes) considered individually. The distinction is important: either P_{31} or P_{32} could be drawn as convex on the phylogeny in Figure 1 by inclusion of just those phyletic line segments necessary to connect all member taxa; however, *both could not be so drawn simultaneously*. When, in traditional taxonomic dialogue, an individual group (e.g., Reptilia) is said to be paraphyletic or polyphyletic, such a group is implicitly understood to be embedded in a level classification (e.g., the set of all amniote classes) the simultaneous extension of whose disjoint subsets on the phylogeny in question gives meaning to the statement.

Measures of information

Effective application of a taxonomic optimality criterion is facilitated when numbers can be calculated to quantify the property that the classification is desired to maximize. Below, we briefly describe mathematical measures of classificatory information content developed by ESTABROOK (1967, 1971), to which papers the interested reader is referred for more extended discussions.

A qualitative character, K_i , defined for S , is also a partition of the study into a collection of exhaustive and mutually exclusive subsets with the same properties as those attributed earlier to the classes of a level classification. The subsets of K_i are called character states.

The information associated with any partition of S , be that a level classification or a qualitative character, can be intuited as the average number of yes-no questions in a maximally efficient questioning scheme required to determine the subset membership of any element, x , equiprobably drawn from S . For a qualitative character, K_i , of S , the expression

$$H(K_i) = - \sum_{j=1}^{n_i} \frac{|K_{ij}|}{|S|} \log_2 \frac{|K_{ij}|}{|S|}$$

where $|K_{ij}|$ is the number of elements in state K_{ij} , and $|S|$ is the number of elements in S , measures the information associated with character K_i . This, SHANNON's familiar entropy statistic, assumes high values when the states of K_i are many with the members of S evenly distributed among them, and low values when the states of K_i are few or the members of S unevenly distributed. In effect, if one knew only what fraction of the membership of S belonged to each state of K_i , then $H(K_i)$ is a measure of the uncertainty one would experience in attempting to determine, by shrewd guessing, the state of K_i to which x is properly assignable.

A level classification (P_k) of S "contains" information about a qualitative character (K_i) of S when knowing the class of P_k to which x belongs reduces uncertainty about the state membership of x in K_i . We define

$$H(K_i/P_{kj}) = - \sum_{l=1}^{n_i} \frac{|K_{il} \cap P_{kj}|}{|P_{kj}|} \log_2 \frac{|K_{il} \cap P_{kj}|}{|P_{kj}|}$$

which is simply that information associated with K_i restricted to class P_{kj} of classification P_k , or, equivalently, the amount of uncertainty remaining about state membership of x in K_i when x is known to be a member of class P_{kj} of P_k . The expected value of this statistic is just the sum of the values it assumes over the classes of P_k , weighted by the frequencies of those classes in S :

$$H(K_i/P_k) = \sum_{j=1}^{n_k} \frac{|P_{kj}|}{|S|} H(K_i/P_{kj})$$

Intuitively, this defines the amount of information associated with character K_i that is not "in" classification P_k – the uncertainty remaining, on average, about the state membership of x in K_i given knowledge of the class membership of x in P_k . Since knowing the class of P_k to which x belongs cannot, on average, convey less information than the (prior) knowledge that x is a member of S , the value of $H(K_i/P_k)$ varies between $H(K_i)$ when knowledge about P_k conveys no information about K_i , and 0 when the classes of P_k coincide in membership with the states of K_i .

The amount of information shared by a level classification and a qualitative character

$$R(K_i, P_k) = H(K_i) - H(K_i/P_k)$$

is simply the amount of information associated with the character, K_i , less the information that remains in K_i when classification P_k is known. This function equals 0 when P_k is uninformative about K_i , and assumes its maximum value, $H(K_i)$, when character state and class memberships coincide. Of interest as a taxonomic optimality criterion, however, is the amount of information conveyed by a level classification about all of the characters that describe the elements of S . If there are m such characters, then the total information effectively contained by P_k about them can be written

$$C_1(P_k) = \sum_{i=1}^m R(K_i, P_k) / H(P_k)$$

The information associated with the classification itself, $H(P_k)$, in addition to being an index of uncertainty in the sense previously defined, is also a measure of the capacity of P_k to distinguish differences among the elements of S and therefore establishes an upper bound for the total character information that can be contained by the classification. In effect, $C_0(P_k)$ is a function both (1) of the number and degree of correspondences between classes and character states and (2) of the raw discriminatory ability of P_k in S . Thus, for two alternative dichotomous classifications that both correspond with a like number of character state distributions in S , $C_0(P_k)$ will usually be greater for that classification whose classes are more nearly equal in membership, that is, for which $H(P_k)$ is highest.

An alternative optimality measure, therefore, is the quotient

$$C_0(P_k) = \sum_{i=1}^m R(K_i, P_k)$$

which can be construed as that fraction of the information associated with P_k that is useful for preserving information about the i^{th} character, summed over all m characters of the study. Unlike $C_0(P_k)$, $C_1(P_k)$ does not increase with enhancement merely of the discriminatory capacity of P_k in S , and is therefore a more faithful index of correspondences between classes and character states than the former statistic.

A relationship between the information theoretic measures derived above and the taxonomist's ability to retrieve character information from knowledge of a classification remains to be established. As HULL (1970) correctly remarks, biological classifications provide only names for specified subsets of a study and "contain" information about the attributes of organisms only indirectly, through some extraneous vehicle for character summarization (e.g., taxon diagnoses) that is indexed by the names of the classes. Nevertheless, while a classification per se does not directly communicate character information, the efficacy with which the indexed diagnoses are able to do so is a direct consequence of the subset memberships in which the substance of the classification consists. Efficacy of diagnosis is obviously maximized when class and character state memberships coincide, and "... a classification is [therefore] most informative concerning a particular character when the groups of the classification describe the distribution of states of that character as simply as possible: the distribution of each state is described by (the membership of) one of the groups [classes] of the classification" (FARRIS 1979: 499). It is, of course, precisely when the states of character K_i correspond in membership with the classes of P_k that $R(K_i, P_k)$, the information common to both, is greatest; the functions $C_0(P_k)$ and $C_1(P_k)$ that sum this shared quantity (or functions thereof) over all characters would thus seem to be reasonable measures of classificatory information content.

Cranchiid phylogeny

Squids of the family Cranchiidae are among the most numerous pelagic cephalopods in the world oceans and exhibit a broad array of adult and larval morphologies. The family was recently revised at the generic level by Voss (1980) who recognized the two subfamilies and 13 genera listed in Table 1 (Tables 1-7 see Appendix p. 91). Voss and Voss (1983) subsequently analyzed phylogenetic relationships among the cranchiid genera; their methods, data and results are summarized briefly below.

Voss and Voss (1983) based their phylogenetic reconstruction on analysis of 14 qualitative morphological characters that reflect variation in anatomical features associated with reproduction, locomotion, feeding, digestion, excretion, structural support and concealment from predators. Characters were selected based on their intrageneric constancy and because the variants of the morphological expressions they represent could be coded as discrete states with minimal ambiguity. Identifications of primitive and derived states of

cranchiid characters were based on (outgroup) comparisons with seven other oegopsid squid families, or on analyses of ontogenetic transformations within the Cranchiidae. Table 2 presents the distribution of states of the original 14 characters among the 13 genera of the cranchiid study, Figure 2 provides character state trees, and Table 3 presents the matrix that results from additive binary recoding of the raw characters of Table 2 given the state trees of Figure 2. VOSS and VOSS (1983) employed the WAGNER Tree method of FARRIS (1970; see also KLUGE and FARRIS 1969 and FARRIS et al. 1970) and the method of Character Compatibility Analysis (ESTABROOK et al. 1977; see also MEACHAM 1980, 1981) to derive maximally-corroborated hypotheses of cranchiid evolutionary history. Both methods are phylogenetic sensu HENNIG (1966) in that recency of common ancestry is inferred on the basis of shared, derived resemblances (synapomorphies) only.

The WAGNER Tree for the cranchiid study is illustrated in Figure 3. All relationships are completely resolved in this phylogenetic reconstruction, and the topology of the diagram requires a minimum of

45 character state transitions in order to derive the observed phenotypes of extant cranchiid genera from the hypothesized morphology of their most recent common ancestor. VOSS and VOSS (1983) tested several other trees in the near neighborhood of that illustrated in Figure 3, but found none that were more parsimonious. The consistency index for the WAGNER estimate of cranchiid relationships is .69, indicating an unusually good fit of tree to data.

Character Compatibility Analysis revealed two maximal cliques of 22 pairwise-compatible binary factors, and fully 71% (22/31) of the derived cranchiid attributes represented in Table 2 could therefore have resulted from unique and unreversed character transformations. The tree illustrated in Figure 4 is supported by the 21 binary factors comprising the intersection of the two largest cliques.

The results of both WAGNER and Character Compatibility analyses support hypotheses of monophyletic origin for the traditional subfamilies Cranchiinae and Taoniinae, and for at least two of the three generic groups that VOSS (1980) discerned within the Taoniinae (Table 1). The two estimates differ only with respect to their ability to resolve the relationships of *Liguriella* and *Sandalops*; these taoniin genera exhibit no synapomorphies, apart from those supporting their subfamilial membership, that are compatible with all of the other binary factors in the two maximal cliques. That the pair forms a monophyletic group with *Bathothauma* and *Helicocranchia*, as shown on the WAGNER Tree, is in fact supported

Character Number(s)	Tree
1	$\begin{array}{c} c \\ \uparrow \\ b < a > d \end{array}$
2, 3	$d < a > b > c$
4, 5, 7, 8, 12, 13	$a > b$
6	$b < a > c > d > e > f$
9, 11	$a > b > c$
10	$b < a > c$
14	$\begin{array}{c} c \quad d \\ \uparrow \quad \uparrow \\ b < a > e > f \end{array}$

Fig. 2. Tree diagrams illustrating the estimates of polarities of states of 14 cranchiid characters. Hypotheses of polarity are presented in the right-hand column, and the characters whose states are believed to have evolved in the sequences illustrated are listed to the left

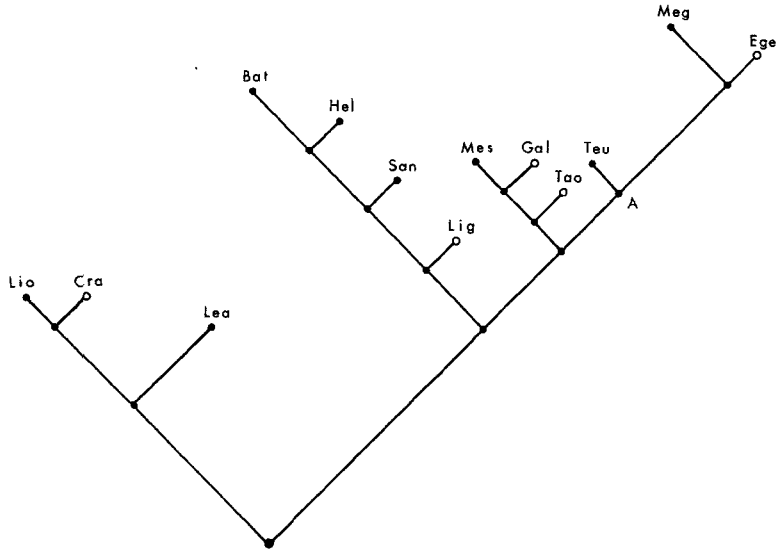


Fig. 3. WAGNER reconstruction of cranchiid phylogeny; names of cranchiid genera are abbreviated as in Table 2. Phyletic lines are drawn proportional to the estimated amounts of morphological evolution (number of character state transitions) that separate extant cranchiids from their hypothetical ancestors or hypothetical ancestors from one another. External nodes drawn as open circles represent extant cranchiids that appear indistinguishable from reconstructed ancestral phenotypes; such nodes have been removed an arbitrary one branch length unit from the appropriate ancestors

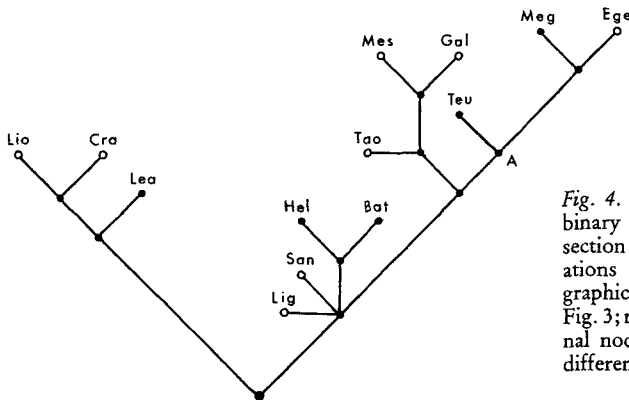


Fig. 4. Phylogeny determined by the 21 binary factors that comprise the intersection of the two largest cliques. Abbreviations of cranchiid generic names and graphical conventions are the same as for Fig. 3; rotations of branches on their internal nodes do not constitute substantive differences between this tree and the preceding figure

by no character state transition that is unique and unreversed in that reconstruction, and their relationships within the subfamily appear to be genuinely equivocal.

Voss and Voss (1983) considered both adaptive correlations and biomechanical linkages among cranchiid characters, but neither phenomenon seems to provide an adequate explanation for the robust cladistic structure evident in the analyses summarized above. Whether the tree diagrams of Figures 2 and 3 are true representations of cranchiid evolutionary history cannot be known with certainty, but as contemporary inferences of genealogy they are surely well supported by the available morphological data.

Informative hierarchies

Fully resolved hierarchic classifications consist of nested, dichotomous partitions of successively smaller subsets of the study. Hierarchies comprised of dichotomous partitions that are each maximally informative about character expressions at successive levels can be derived by a divisive procedure with the aid of the information-theoretic measures introduced earlier. The desired optimality statistic, $C_0(P_k)$ or $C_1(P_k)$, is first calculated for all possible dichotomous partitions of the entire study, and that partition selected for which the chosen measure assumes its highest value. All dichotomous partitions of each of the newly-formed subsets are subsequently tested in the same way, and so forth until a completely resolved hierarchy for the study is completed.

The result of applying such an algorithm to the cranchiid data set for whole characters (Table 2), and maximizing $C_0(P_k)$ at each level is shown in Figure 5. The number associated with each stem is the value of the information-preserving statistic for the partition effected below it. We note that the values assumed by $C_0(P_k)$ decrease monotonically from higher to lower levels in the classification as class memberships become increasingly homogeneous with respect to character state expressions.

Partitions of the cranchiid study diagrammed in Figure 5 are all convex on the estimates of evolutionary history illustrated in Figures 3 and 4. The first partition in the information tree corresponds to the monophyletic subfamilial classification (Table 1) and effects correspondence between class memberships and the distribution of eight character states (1 d, 4 a, 4 b, 6 b, 7 a, 7 b, 12 a, 12 b). The most informative partition of the subfamily Cranchiinae reflects the distribution of states of characters 5, 11, 13 and 14 and is also monophyletic. The most informative dichotomy within the Taoniinae distinguishes members of the *Sandalops* group (Hel, Bat, San, Lig) from the six higher taoniin genera; the partition is monophyletic on the WAGNER estimate of cranchiid relationships (Figure 3) and is at least convex on the Character Compatibility Analysis results (Figure 4). Subsequent partitioning of the *Sandalops* group is convex but not demonstrably monophyletic. Separation of the six higher taoniins into the monophyletic *Taonius* (Tao, Gal, Mes) and *Megalocranchia* (Ege, Meg, Teu) groups is also effected in this information-maximizing hierarchy, and corresponds to the distribution of five character states (3 b, 3 c, 8 a, 8 b, 9 a) among members of this taoniin subset. Within the *Megalocranchia* group, all of the information in characters 1, 5, 6 and 14 is contained in the monophyletic partition (Teu) (Ege, Meg), but no uniquely best partition of the *Taonius* group is possible because *Galiteuthis* is with equal information-preserving efficacy classed either with *Taonius* or *Mesonychoteuthis*, each arrangement corresponding to the state distribution of one character (9 or 11).

Use of $C_1(P_k)$ as the classificatory optimality measure results (Figure 6) in a hierarchy of squids that differs from the preceding only with respect to partitioning the six higher taoniin genera: *Egea* and *Megalocranchia* are separated as a group apart from *Teuthowenia*, *Taonius*, *Galiteuthis* and *Mesonychoteuthis*. The partition is convex (but not monophyletic) on the tree diagrams of Figures 3 and 4, and results in classificatory correspondence with six character states (1 a, 1 b, 5 a, 5 b, 14 c, 14 e). A subsequent partition separates *Teuthowenia* from members of the (monophyletic) *Taonius* group.

Table 4 presents values of $C_0(P_k)$ for the first, second and third most informative dichotomous partitions of successively smaller subsets of the cranchiid study where subset memberships at each level is determined by the preceding most informative choice. As can be seen, the most informative partition at each level is not always so by a large margin. Particularly striking is the fact that partitioning the entire family into monophyletic subfamilies is hardly more character-informative than is a non-convex alternative that places *Egea* and *Megalocranchia*, two higher taoniins, together with the three cranchiin genera. For the first dichotomous partition of the Taoniinae, a non-convex alternative is likewise only a

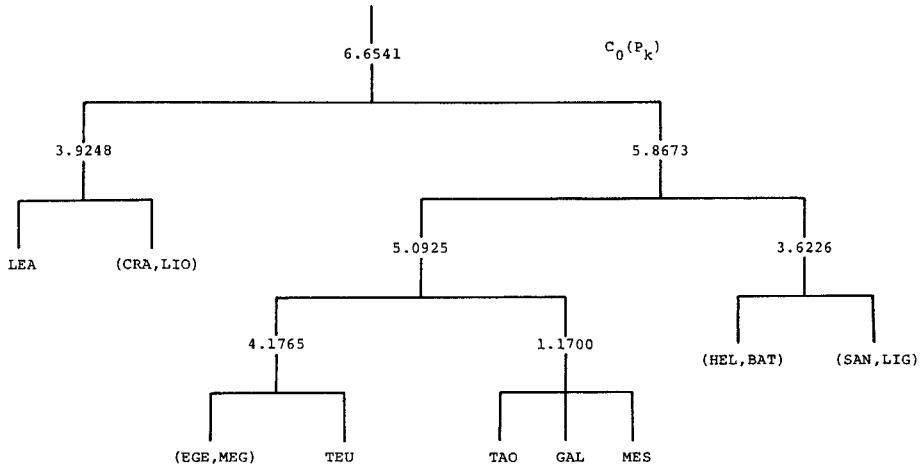


Fig. 5. The hierarchy resulting from application of a divisive algorithm that maximizes $C_0(P_k)$ at each level to the whole (unfactored) character data of Table 2. The number associated with each stem is the value of $C_0(P_k)$ for the partition effected below it

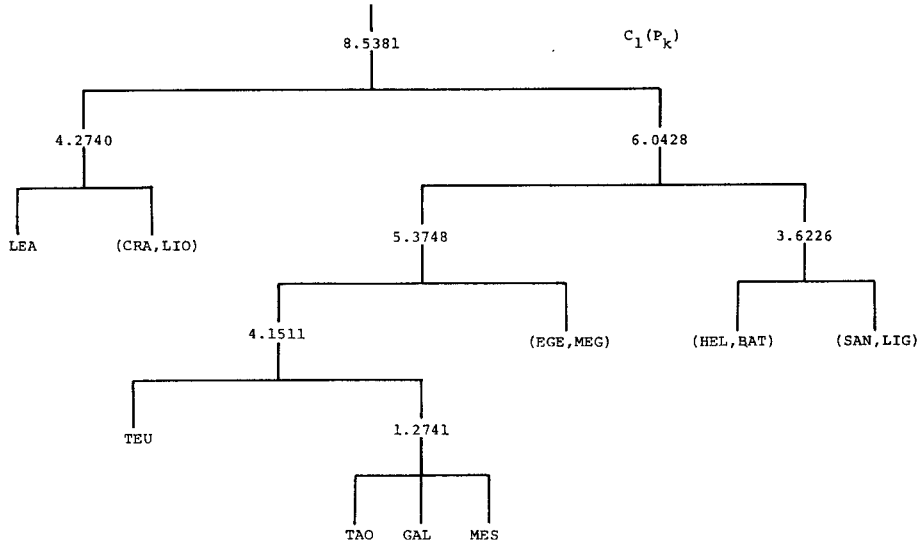


Fig. 6. The hierarchy resulting from application of a divisive algorithm that maximizes $C_1(P_k)$ at each level to the whole (unfactored) character data of Table 2. The number associated with each stem is the value of $C_1(P_k)$ for the partition effected below it

little less informative than the monophyletic choice. Partitions that result in highest and next-highest values of $C_1(P_k)$ at each level are all at least convex (Table 5).

The non-convex level classifications of the whole family and of the subfamily Taoniinae mentioned above reflect memberships in states 8b and 11c, respectively, that Voss and Voss (1983) believed to represent instances of homoplastic evolution, but why alternative, convex classifications are not more informative to a greater degree might also be attributable to the fact that most of the characters of the study are multistate (Figure 2), and no provision has been made, thus far, for taking the ordering of states into account. Additive binary recoding is one way of introducing this information, and the results of applying our divisive

algorithms to the factored data of Table 3 are presented in Tables 6 and 7. As can be seen in Table 6 where $C_0(P_k)$ is the optimized measure, one of the effects of binary recoding is to eliminate the non-convex alternative as second choice to the monophyletic subfamilial partition; a non-convex alternative persists as the next-most-informative taoniin dichotomy, but binary recoding has rendered the monophyletic partition relatively more informative than before. Another change in classification has been effected by recoding the characters, however, and that is to make the (convex) separation of *Egea* and *Megalocranchia* as a group apart from the other four higher taoniins more informative than the monophyletic partition; the same result is obtained when $C_1(P_k)$ is the maximized statistic (Table 7).

Discussion

We have shown that nested, dichotomous level classifications that are each maximally informative about qualitative character expressions within successively smaller subsets of the squid family Cranchiidae are all convex on phylogenetic reconstructions based on the same data. To the best of our knowledge, this relationship between independently derived phylogenetic reconstructions and information-preserving hierarchies has not hitherto been investigated, and it is therefore of interest to inquire whether our results may be expected to obtain generally, for other study collections.

A qualitative character supports an estimate of evolutionary relationships for a study collection if and only if the character (as a partition of the study) is convex on the tree diagram in question. If most of the characters of a study form a pairwise compatible collection (can simultaneously support a common phylogenetic reconstruction) then there may, accordingly, be a strong tendency for maximally informative level classifications of the study to be convex on the best-supported estimate of relationships insofar as correspondences between classes and character states are thereby likely to be maximized. States of characters that are not convex on otherwise well supported phylogenies are usually hypothesized to describe non-homologous morphological expressions, but since examples of apparently convergent or reversed evolution are usually identified as such because they are uncommon and incompatible with many other characters, putative homoplasy (to the extent that it can be recognized) is unlikely to render maximally informative a level classification that is not convex on a credible phylogenetic reconstruction. In the cranchiid example, the distribution of character state 11 c within the Taoniinae is at least partly responsible for the high information content of the non-convex partition (Hel, Bat, San, Lig, Mes) (Tao, Gal, Ege, Meg, Teu), but it is the very scarcity of character states with similar distributions that both supports an hypothesis of convergent derivation for 11 c (Voss and Voss 1983) and prevents the classificatory dichotomy that reflects its distribution from being maximally informative about other characters of the study. In short, while it cannot be known with certainty whether most-informative level classifications will necessarily be convex on tree diagrams of true evolutionary history, it does seem plausible that they will often be so on estimates of relationships that are well supported by characters.

Maximally informative dichotomous classifications may conceivably fail to be convex on phylogenetic reconstructions, even in the absence of evident homoplasy, if many unifactored multistate characters are present in the data. This results from the fact that no level classification can contain all of the information about a multistate character unless there are as many classes in the classification as states in the character. Thus, if a character (K_i) has n states, and a classification (P_k) $n-1$ classes, then at least one of the classes of P_k must contain members of two or more states of K_i and P_k will only be convex on phylogenetic estimates supported by K_i if P_k is convex on the state tree of K_i . If K_i , for example, has the state tree $b \leftarrow a \rightarrow c$, and more than half of the members of S belong to K_{ia} , then for dichotomous classification P_k , $R(K_i, P_k)$ will be greatest when $P_k = (\text{members of } K_{ia}) (\text{members of } K_{ib} \cup K_{ic})$ and P_k will fail to be convex on any tree that K_i supports. If there are

several characters distributed like K_i in S , then even in the absence of (real or inferred) homoplasy $C_0(P_k)$ and (perhaps) $C_1(P_k)$ may be maximal for some partition that is not convex on the best-supported phylogeny for S . Additive binary recoding eliminates this source of difficulty by reducing multistate characters to two-state factors in which each state is the complement of exactly one other state; in the cranchiid study, two non-convex partitions that were nearly as character-informative as convex alternatives before (Table 4) but not after (Table 6) binary recoding illustrate this phenomenon.

That alternative level classifications of the cranchiid study are ranked differently when evaluated by $C_0(P_k)$ and $C_1(P_k)$ (compare Tables 4 and 6 with 5 and 7) relates to properties of these statistics discussed earlier: both measures assume high values when class and character state correspondences are maximized, but $C_0(P_k)$ is also a function of the raw discriminatory ability of the partition ($H(P_k)$). Thus, the relatively high values assumed by $C_0(P_k)$ for the non-convex partitions in Table 4, though partially ameliorated by character recoding for reasons discussed above, also reflect the near-evenness (and associated high discriminatory capacity) of those dichotomies. $C_1(P_k)$, by contrast, does not assume high values for non-convex cranchiid classifications under either method of character coding treated here.

The close correspondence between the hierarchies of Figures 3 and 4 and those of Figures 5 and 6 suggests that classificatory monophyly may often serve effectively to communicate character information as well as to convey explicit hypotheses of relationships, but maximally character-informative partitions of cranchiids are not always demonstrably monophyletic. This may be due to a real absence of well-corroborated cladistic structure, as was detailed above for the genera of the *Sandalops* group, or no uniquely most informative partition may exist for a given collection (the *Taonius* group example). Some informative departures from monophyly, however, are attributable to inequalities in reconstructed evolutionary rates. As can be seen in both the Wagner and Character Compatibility hypotheses of cranchiid phylogeny (Figures 3 and 4), from the shared internal node A the ancestral lineage of *Egea* and *Megalocranchia* is estimated to have diverged morphologically to a greater extent than did the ancestral lineage of *Teuthowenia*, and the effect of this disproportionately rapid divergence was to increase the number of character states that distinguish *Teuthowenia* and members of the *Taonius* group on the one hand from *Egea* and *Megalocranchia* on the other. The convex (but not monophyletic) partition of these six genera in Figure 6 therefore correspond to the distribution of states of four binary factors (1b, 5b, 14c, 14e) while the monophyletic partition is coincident with only three (3c, 8b, 9b).

That evolutionary rate inequalities can render monophyletic level classifications less than maximally character-informative is not an original observation (see BOTTJER 1980, for a recent exposition). FARRIS (1979), however, has recently advanced the thesis that character information is always best communicated by classifications that are monophyletic with respect to most-parsimonious phylogenetic reconstructions. In thus arguing, FARRIS quantifies information as a function of the entire hierarchy, and would, therefore, permit an included level classification to be less informative about characters than an alternative partition if that choice facilitated information retrieval elsewhere in the hierarchy. We prefer to ground measures of classificatory information content on the relationships between level classifications and characters not only because the problem of discovering hierarchies that maximize measures of such relationships is solveable, but also because the choice situation confronting working taxonomists usually concerns alternative level classifications (different subgeneric arrangements, for example) not completely-resolved hierarchies that would, for many studies, be unwieldy as formal classifications.

Acknowledgements

The data analyzed in this paper derive from research supported by National Science Foundation grants DEB-7713945 and DEB-8105193 to N. A. Voss. Thanks are also due to the Division of Biological Sciences, University of Michigan, for support of computing services.

Appendix

Table 1
Classification of Voss, 1980

Family Cranchiidae
Subfamily Cranchiinae
<i>Cranchia</i>
<i>Liocranchia</i>
<i>Leachia</i>
Subfamily Taoniinae
<i>Sandalops</i> group
<i>Helicocranchia</i>
<i>Bathothauma</i>
<i>Sandalops</i>
<i>Liguriella</i>
<i>Taonius</i> group
<i>Taonius</i>
<i>Galiteuthis</i>
<i>Mesonychoteuthis</i>
<i>Megalocranchia</i> group
<i>Egea</i>
<i>Megalocranchia</i>
<i>Teuthowenia</i>

Table 2
Primary data matrix

Columns represent cranchiid genera; rows represent characters. The entry for a given row × column is the state label appropriate to the corresponding genus and character. States labelled 'a' are hypothesized plesiomorphs, states b–f are presumed apomorphs; character state trees are presented in Fig. 2

Character Number	Taxa													
	<i>Cra</i>	<i>Lio</i>	<i>Lea</i>	<i>Hel</i>	<i>Bat</i>	<i>San</i>	<i>Lig</i>	<i>Tao</i>	<i>Gal</i>	<i>Mes</i>	<i>Ege</i>	<i>Meg</i>	<i>Teu</i>	
1	d	d	d	c	c	c	a	a	a	a	b	b	a	
2	a	b	a	a	d	a	b	c	c	c	c	c	c	
3	a	a	a	d	d	a	a	b	b	b	c	c	c	
4	b	b	b	a	a	a	a	a	a	a	a	a	a	
5	a	a	b	b	b	b	b	b	b	b	a	a	b	
6	b	b	b	c	c	d	d	e	e	e	e	e	f	
7	a	a	a	b	b	b	b	b	b	b	b	b	b	
8	b	b	b	a	a	a	a	a	a	a	b	b	b	
9	a	a	a	a	a	a	a	b	c	c	a	a	a	
10	a	a	a	a	b	b	b	a	a	a	a	c	a	
11	c	c	a	c	c	c	c	b	b ¹	c	b	a	b	
12	b	b	b	a	a	a	a	a	a	a	a	a	a	
13	a	a	b	b	b	b	b	b	b	b	b	b	b	
14	b	b	d	e	f	a	e	e	e	e	c	c	e	

Abbrev. of taxa for this and subsequent tables and figures: *Cra*: *Cranchia*; *Lio*: *Liocranchia*; *Lea*: *Leachia*; *Hel*: *Helicocranchia*; *Bat*: *Bathothauma*; *San*: *Sandalops*; *Lig*: *Liguriella*; *Tao*: *Taonius*; *Gal*: *Galiteuthis*; *Mes*: *Mesonychoteuthis*; *Ege*: *Egea*; *Meg*: *Megalocranchia*; *Teu*: *Teuthowenia*

¹ Expressions of character 11 vary among species of *Galiteuthis* (see Voss and Voss 1982); state 'b' was arbitrarily chosen to represent the genus in the present study.

Table 3

Factored data matrix

Columns are labelled as in Table 1. Binary factors are labelled with the character state that is the apomorph for the transformation represented by each factor

Factor Label	Taxa													
	<i>Cra</i>	<i>Lio</i>	<i>Lea</i>	<i>Hel</i>	<i>Bat</i>	<i>San</i>	<i>Lig</i>	<i>Tao</i>	<i>Gal</i>	<i>Mes</i>	<i>Ege</i>	<i>Meg</i>	<i>Teu</i>	
1b	0	0	0	0	0	0	0	0	0	0	1	1	0	
1c	0	0	0	1	1	1	0	0	0	0	0	0	0	
1d	1	1	1	0	0	0	0	0	0	0	0	0	0	
2b	0	1	0	0	0	0	1	1	1	1	1	1	1	
2c	0	0	0	0	0	0	0	1	1	1	1	1	1	
2d	0	0	0	0	1	0	0	0	0	0	0	0	0	
3b	0	0	0	0	0	0	0	1	1	1	1	1	1	
3c	0	0	0	0	0	0	0	0	0	0	1	1	1	
3d	0	0	0	1	1	0	0	0	0	0	0	0	0	
4b	1	1	1	0	0	0	0	0	0	0	0	0	0	
5b	0	0	1	1	1	1	1	1	1	1	0	0	1	
6b	1	1	1	0	0	0	0	0	0	0	0	0	0	
6c	0	0	0	1	1	1	1	1	1	1	1	1	1	
6d	0	0	0	0	0	1	1	1	1	1	1	1	1	
6e	0	0	0	0	0	0	0	1	1	1	1	1	1	
6f	0	0	0	0	0	0	0	0	0	0	0	0	1	
7b	0	0	0	1	1	1	1	1	1	1	1	1	1	
8b	1	1	1	0	0	0	0	0	0	0	1	1	1	
9b	0	0	0	0	0	0	0	1	1	1	0	0	0	
9c	0	0	0	0	0	0	0	0	1	1	0	0	0	
10b	0	0	0	0	1	1	1	0	0	0	0	0	0	
10c	0	0	0	0	0	0	0	0	0	0	0	1	0	
11b	1	1	0	1	1	1	1	1	1	1	1	0	1	
11c	1	1	0	1	1	1	1	0	0	1	0	0	0	
12b	1	1	1	0	0	0	0	0	0	0	0	0	0	
13b	0	0	1	1	1	1	1	1	1	1	1	1	1	
14b	1	1	0	0	0	0	0	0	0	0	0	0	0	
14c	0	0	0	0	0	0	0	0	0	0	1	1	0	
14d	0	0	1	0	0	0	0	0	0	0	0	0	0	
14e	0	0	0	1	1	0	1	1	1	1	0	0	1	
14f	0	0	0	0	1	0	0	0	0	0	0	0	0	

Table 4

First, second and third highest values of $C_0(P_k)$ for alternative partitions of successively smaller subsets of the cranchiid study using whole (unfactored) characters; only the partitions effecting the highest and next-highest values of $C_0(P_k)$ at each level are shown

Whole Characters	
$C_0(P_k)$	Partition
6.6541	(<i>Cra, Lio, Lea</i>) (<i>Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu</i>)
6.4137	(<i>Cra, Lio, Lea, Ege, Meg</i>) (<i>Hel, Bat, San, Lig, Tao, Gal, Mes, Teu</i>)
6.3346	
5.8673	(<i>Hel, Bat, San, Lig</i>) (<i>Tao, Gal, Mes, Ege, Meg, Teu</i>)
5.0797	(<i>Hel, Bat, San, Lig, Mes</i>) (<i>Tao, Gal, Ege, Meg, Teu</i>)
4.9010	
5.0925	(<i>Tao, Gal, Mes</i>) (<i>Ege, Meg, Teu</i>)
4.9356	(<i>Tao, Gal, Mes, Teu</i>) (<i>Ege, Meg</i>)
3.1873	
1.1700	(<i>Tao, Gal</i>) (<i>Mes</i>)
1.1700	(<i>Tao</i>) (<i>Gal, Mes</i>)
0.5033	
4.1765	(<i>Ege, Meg</i>) (<i>Teu</i>)
2.8431	(<i>Ege, Teu</i>) (<i>Meg</i>)
1.5098	
3.6226	(<i>Hel, Bat</i>) (<i>San, Lig</i>)
2.6790	(<i>Hel, Bat, San</i>) (<i>Lig</i>)
2.4902	
3.9248	(<i>Cra, Lio</i>) (<i>Lea</i>)
1.9248	(<i>Cra, Lea</i>) (<i>Lio</i>)
1.2582	

Table 5

First, second and third highest values of $C_1(P_k)$ for alternative partitions of successively smaller subsets of the cranchiid study using whole (unfactored) characters

Other conventions as in Table 4

Whole characters	
$C_1(P_k)$	Partition
8.5381	<i>(Cra, Lio, Lea) (Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu)</i>
7.3528	<i>(Cra, Lio) (Lea, Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu)</i>
6.7427	
6.0428	<i>(Hel, Bat, San, Lig) (Tao, Gal, Mes, Ege, Meg, Teu)</i>
5.9745	<i>(Hel, Bat, San, Lig, Tao, Gal, Mes, Teu) (Ege, Meg)</i>
5.5611	
5.3748	<i>(Tao, Gal, Mes, Teu) (Ege, Meg)</i>
5.0925	<i>(Tao, Gal, Mes) (Ege, Meg, Teu)</i>
4.4170	
4.1511	<i>(Tao, Gal, Mes) (Teu)</i>
2.2451	<i>(Tao, Teu) (Gal, Mes)</i>
1.8369	
1.2741	<i>(Tao, Gal) (Mes)</i>
1.2741	<i>(Tao) (Gal, Mes)</i>
0.5481	
3.6226	<i>(Hel, Bat) (San, Lig)</i>
3.3022	<i>(Hel, Bat, San) (Lig)</i>
3.0695	
4.2740	<i>(Cra, Lio) (Lea)</i>
2.0961	<i>(Cra, Lea) (Lio)</i>
1.3701	

Table 6

First, second and third highest values of $C_0(P_k)$ for alternative partitions of successively smaller subsets of the cranchiid study using factored characters

Other conventions as in Table 4

Factored Characters	
$C_0(P_k)$	Partition
8.4982	<i>(Cra, Lio, Lea) (Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu)</i>
8.4824	<i>(Cra, Lio, Lea, Hel, Bat, San, Lig) (Tao, Gal, Mes, Ege, Meg, Teu)</i>
7.9272	
7.8964	<i>(Hel, Bat, San, Lig) (Tao, Gal, Mes, Ege, Meg, Teu)</i>
6.6012	<i>(Hel, Bat, San, Lig, Mes) (Tao, Gal, Ege, Meg, Teu)</i>
6.5122	
6.1540	<i>(Tao, Gal, Mes, Teu) (Ege, Meg)</i>
6.0592	<i>(Tao, Gal, Mes,) (Ege, Meg, Teu)</i>
3.9465	
3.6790	<i>(Tao, Gal, Mes) (Teu)</i>
2.5565	<i>(Tao, Teu) (Gal, Mes)</i>
1.6129	
1.1700	<i>(Tao, Gal) (Mes)</i>
1.1700	<i>(Tao) (Gal, Mes)</i>
0.5034	
3.8678	<i>(Hel, Bat) (San, Lig)</i>
2.7354	<i>(Hel, Bat, San) (Lig)</i>
2.7354	
5.7615	<i>(Cra, Lio) (Lea)</i>
2.4282	<i>(Cra, Lea) (Lio)</i>
1.7615	

Table 7

First, second and third highest values of $C_1(P_k)$ for alternative partitions of successively smaller subsets of the cranchiid study using factored characters

Other conventions as in Table 4

Factored characters	
$C_1(P_k)$	Partition
10.9042	(<i>Cra, Lio, Lea</i>) (<i>Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu</i>)
9.3867	(<i>Cra, Lio</i>) (<i>Lea, Hel, Bat, San, Lig, Tao, Gal, Mes, Ege, Meg, Teu</i>)
8.5188	
8.1327	(<i>Hel, Bat, San, Lig</i>) (<i>Tao, Gal, Mes, Ege, Meg, Teu</i>)
7.7356	(<i>Hel, Bat, San, Lig, Tao, Gal, Mes, Teu</i>) (<i>Ege, Meg</i>)
7.3894	
6.7015	(<i>Tao, Gal, Mes, Teu</i>) (<i>Ege, Meg</i>)
6.0592	(<i>Tao, Gal, Mes</i>) (<i>Ege, Meg, Teu</i>)
5.1467	
4.5348	(<i>Tao, Gal, Mes</i>) (<i>Teu</i>)
2.5565	(<i>Tao, Teu</i>) (<i>Gal, Mes</i>)
1.9881	
1.2741	(<i>Tao, Gal</i>) (<i>Mes</i>)
1.2741	(<i>Tao</i>) (<i>Gal, Mes</i>)
0.5482	
3.8678	(<i>Hel, Bat</i>) (<i>San, Lig</i>)
3.3717	(<i>Hel, Bat, San</i>) (<i>Lig</i>)
3.3717	
6.2741	(<i>Cra, Lio</i>) (<i>Lea</i>)
2.6442	(<i>Cra, Lea</i>) (<i>Lio</i>)
1.9183	

Summary

The distribution of states of 14 qualitative morphological characters among 13 genera of cranchiid squids is analyzed in order to compare information-theoretically optimal classifications with phylogenetic reconstructions based on the same data. The information shared between a level classification and a qualitative character is first defined mathematically and then expanded to derive measures for the information contained by a level classification about all of the characters describing a study collection. Hierarchies of nested, dichotomous partitions that are each maximally informative about character state distributions within successively smaller subsets of the cranchiid study are obtained by application of a simple divisive algorithm. Level classifications of squids contained in such hierarchies were all found to be convex on the results of WAGNER Tree and Character Compatibility analyses. Potentially confounding effects of putative homoplasy and of character coding are discussed, and the possibility that character-informative level classifications may generally be expected to be convex on well-supported phylogenetic reconstructions is suggested.

Zusammenfassung

Vergleich einer informationstheoretisch optimalen Klassifikation mit einer phylogenetischen Verwandtschaftsanalyse am Beispiel der Kalmar-Familie Cranchiidae (Cephalopoda: Oegopsidae)

Basierend auf der Verteilung von 14 qualitativ-morphologischen Merkmalen von 13 Gattungen der Cranchiidae wird untersucht, inwieweit daraus abgeleitete „optimale Klassifikationen“ im Sinne der Informationstheorie mit auf phylogenetischen Rekonstruktionen basierenden „Stammbaumdarstellungen“ übereinstimmen.

Der Informationsgehalt verschiedener Klassifikationen wird in bezug auf einzelne Merkmale mathematisch definiert und zu einem Maß für die in einer Klassifikation enthaltene Information über alle dazu verwendeten Merkmale erweitert. Hierarchische Klassifikationen, deren dichotom aufgebaute Teilgruppen jeweils eine maximale Information über die Verteilung der betrachteten Merkmale aufweisen, sind „konvex“ mit phylogenetischen Stammbaumdarstellungen.

Literature

- BOTTJER, P. D., 1980: FARRIS' "information content" and phylogenetic versus evolutionary classification: the philosophical differences remain. *Syst. Zoology* 29, 382-385.
- DUNCAN, T.; ESTABROOK, G. F., 1976: An operational method for evaluating classifications. *Syst. Botany* 1, 373-382.
- ESTABROOK, G. F., 1967: An information theory model for character analysis. *Taxon* 16, 86-97.
- 1971: Some information theoretic optimality criteria for general classification. *Math. Geol.* 3, 203-207.
- ESTABROOK, G. F.; STRAUCH, J. G.; FIALA, K. L., 1977: An application of compatibility analysis to the Blackliths' data on orthopteroid insects. *Syst. Zoology* 26, 269-276.
- FARRIS, J. S., 1970: Methods for computing Wagner Trees. *Syst. Zoology* 19, 83-92.
- 1979: The information content of the phylogenetic system. *Syst. Zoology* 28, 483-519.
- FARRIS, J. S.; KLUGE, A. G.; ECKHARDT, M. J., 1970: A numerical approach to phylogenetic systematics. *Syst. Zoology* 19, 172-189.
- GILMOUR, J. S. L., 1940: Taxonomy and philosophy. In: *The New Systematics*. Ed. by J. HUXLEY. Oxford: Clarendon. 461-474.
- HENNIG, W., 1966: *Phylogenetic Systematics*. (Transl. by D. D. DAVIS and R. ZANGERL). Chicago: Univ. Illinois Press.
- HULL, D. L., 1970: Contemporary systematic philosophies. *Ann. Rev. Ecol. Syst.* 1, 19-54.
- KLUGE, A. G.; FARRIS, J. S., 1969: Quantitative phyletics and the evolution of anurans. *Syst. Zoology* 18, 1-32.
- MEACHAM, C. A., 1980: Phylogeny of the Berberidaceae with an evaluation of classifications. *Syst. Botany* 5, 149-172.
- 1981: A manual method for character compatibility analysis. *Taxon* 30, 591-600.
- SOKAL, R. R.; SNEATH, P. H. A., 1963: *Principles of Numerical Taxonomy*. San Francisco: W. H. Freeman.
- VOSS, N. A., 1980: A generic revision of the Cranchiidae (Cephalopoda: Oegopsida). *Bull. Mar. Sci.* 30, 365-412.
- VOSS, NANCY A.; VOSS, R. S., 1983: Phylogenetic relationships in the cephalopod family Cranchiidae (Oegopsida). *Malacologia* 23, 397-426.
- Authors' addresses:* ROBERT S. VOSS and GEORGE F. ESTABROOK, Museum of Zoology and Division of Biological Sciences, University of Michigan, Ann Arbor, MI 48109; NANCY A. VOSS, Rosenstiel School of Marine and Atmospheric Sciences, University of Miami, Miami, FL 33149

Drosophila paulistorum* egg length**By IRA B. PERELLE¹, ST. B. DANIELS and L. EHRLMAN*Received on 10. November 1982Introduction**

It is generally recognized that natural selection can be categorized into three major types: directional selection, in which an extreme phenotype is favored; disruptive selection, in which extreme phenotypes are favored; and stabilizing selection, in which an intermediate phenotype is favored. Recently, CURTSINGER (1976 a, b) determined that egg length in *Drosophila melanogaster* is influenced by stabilizing selection. He found that intermediate sized eggs show an adaptive advantage vis-à-vis hatchability, the criterion he employed to determine fitness.

¹ Prof. Dr. PERELLE is affiliated with Mercy College, Dobbs Ferry, New York 10522 and with the Division of Natural Sciences, State University, Purchase, New York 10577 where this work was done in collaboration with the other two authors.