Definitions in Phylogenetic Taxonomy: Critique and Rationale

PAUL C. SERENO

Department of Organismal Biology and Anatomy, University of Chicago,
1027 E. 57th Street, Chicago, Illinois 60637, USA; E-mail: dinosaur@uchicago.edu

Abstract.— A general rationale for the formulation and placement of taxonomic definitions in phylogenetic taxonomy is proposed, and commonly used terms such as "crown taxon" or "node-based definition" are more precisely defined. In the formulation of phylogenetic definitions, *nested* reference taxa stabilize taxonomic content. A definitional configuration termed a *node-stem triplet* also stabilizes the relationship between the trio of taxa at a branchpoint, in the face of local change in phylogenetic relationships or addition/deletion of taxa. Crown-total taxonomies use survivorship as a criterion for placement of node-stem triplets within a taxonomic hierarchy. Diversity, morphology, and tradition also constitute heuristic criteria for placement of node-stem triplets. [Content; crown; definition; node; phylogeny; stability; stem; taxonomy.]

Does one type of phylogenetic definition (apomorphy, node, stem) stabilize the taxonomic content of a taxon more than another in the face of local change of relationships? Is one type of phylogenetic definition more suitable for clades with unresolved basal relationships or uncertain outgroups? Which type of phylogenetic definition is preferable for clades whose members are entirely living or extinct?

Questions like these have not been rigorously addressed in phylogenetic taxonomy. Rather, attention has been focused on (1) the adaptation of traditional rules governing synonymy and redundancy for use within the phylogenetic system, and (2) the recommendation that "widely recognized" names are better restricted to crown taxa than to more-inclusive taxa with extinct basal members (de Queiroz and Gauthier, 1990, 1992; Rowe and Gauthier, 1992; Bryant, 1994, 1996; Lee, 1996). Despite increasing use of phylogenetic definitions in systematics, a general rationale has yet to be proposed for the formulation and placement of phylogenetic definitions, and many commonly used terms such as "crown taxon" and "node-based definition" have yet to be specifically defined.

LACK OF A GENERAL RATIONALE

De Queiroz and Gauthier (1990, 1992) first articulated the general structure of phylogenetic definitions, outlining apomorphy-, node-, and stem-based definitions (Fig. 1).

Most phylogenetic definitions have been constructed in the systematic literature since then without explanation or justification for the particular type of definition used. The justification given for preferential use of node- and stem-based definitions for crown and total taxa, respectively, is incomplete or inaccurate, as reviewed below.

Total taxa and crown taxa require stem-based and node-based definitions, respectively.— Given that a total taxon includes all currently known and potential extinct outgroups that are most closely related to a particular crown taxon (Table 1), the total taxon must have a stem-based definition (de Queiroz and Gauthier, 1992). The stem-based structure of the definition ensures the inclusion of all taxa up to, but excluding, the common ancestor shared with its sister total taxon. This justification is sufficient, because neither an apomorphy- nor a node-based definition would include all potential extinct outgroups.

A crown taxon, in contrast, does not require a node-based definition, although one is commonly assumed. De Queiroz and Gauthier (1992:469) simply remarked, "Names can be associated unambiguously with crown clades using node-based definitions." Likewise, Rowe and Gauthier (1992) and McKenna and Bell (1997) proposed node-based definitions for Mammalia without explaining why definitional type is preferable. Lee (1996:1103) remarked, "A crown-clade definition results when both taxa implicated in a node-based definition
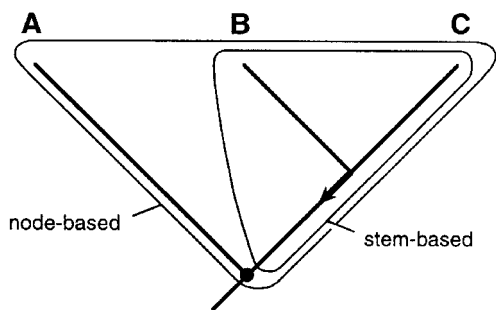
FIGURE 1. Node-based and stem-based phylogenetic definitions, which usually have been shown graphically by encircling portions of a cladogram (de Queiroz and Gauthier, 1992), are indicated here by a dot (node-based) or arrow (stem-based).

are represented by extant forms." And Wyss and Meng (1996:559) stated that the definitions of crown taxa "are thus node-based in their formulation," without presenting any reasoning for such a conclusion.

Any group of extant species may just as well be united by a stem-based definition. The crown taxon Amniota, for example, could have a stem-based definition, such as: "any Recent tetrapod more closely related to Mammalia than to Anura and all extinct descendants of their most recent common ancestor." A stem-based definition for a crown taxon, in fact, may be preferable if the basal relationships among extant taxa within a crown clade are poorly established, as seems to be the case with Amniota (see also Rodentia: Wyss and Meng, 1996) (Fig. 2b). All members of the crown taxon will be included, even if basal ingroup relationships are poorly resolved, because a stem-based definition is based on reference to an outgroup taxon (or taxa).

A single suboscine bird, for example, would suffice as an outgroup reference taxon for a stem-based definition of the crown taxon Oscines. A node-based definition of the same crown taxon would require listing many oscine subgroups to ensure inclusion of all living species currently regarded as oscines. On the other hand, if the basal dichotomy within a crown taxon is well established and outgroup relationships are uncertain (Fig. 2a), a node-based

definition will more effectively stabilize taxonomic content. Thus, there is no current justification for exclusive use of node-based definitions to delimit crown taxa. Node-based crown taxa, in conjunction with stem-based total taxa, however, yield a local definitional configuration that can provide justification for preferential use of node-based crown taxa (see *Node-Stem Triplet* below).

Node-based crown taxa are more stable, informative, and accurate than taxa defined by extinct species and have been (or should be) associated with "widely used" names.—Gauthier (1986:12) stated that restricting "Aves" to "living taxa" maximizes "stability and phylogenetic informativeness," and de Queiroz and Gauthier (1992:468) suggested that "biologists commonly use the widely known name Aves when making generalizations that apply to extant birds alone." With regard to the taxon Mammalia, Rowe and Gauthier (1992:372) observed that "no one in several centuries has mistakenly assigned a Recent mammal to some other taxon." McKenna and Bell (1997:32) remarked that "a crown group is likely to remain relatively stable."

Despite these and similar statements in the literature, node- or stem-based crown taxa are not demonstrably more stable in terms of *meaning* or *content*. Any node- or stem-based definition that specifically identifies a most recent common ancestor clearly specifies the boundaries of a clade, regardless of the living or extinct status of the reference taxa or the particular phylogeny to which the definition is applied (Lee, 1996).

In terms of taxonomic content, redefining Mammalia as a crown taxon does not immediately clarify which *extinct* taxa will be included or excluded. Mammalia as a crown taxon appears more stable only if extinct taxa are ignored and attention is focused on the great phylogenetic distance to their nearest *living* sister group. Even so, inclusion of some living mammals may be uncertain. Rowe and Gauthier's (1992:375) preferred node-based definition of Mammalia ("the last common ancestor of Monotremata and Theria"; see also McKenna and Bell, 1997:32,35.) may be synonymous with Theria, if monotremes

TABLE 1.   Definitions for terms used in this paper.

Term	Definition
Taxonomic definition	relational statement specifying the taxonomic content of a taxon
Taxonomic diagnosis	descriptive statement specifying the apomorphies (derived character states) that serve to identify members of a taxon
Taxonomic content	existing and potential taxa or individuals that by definition are included within a taxon
Reference taxon	taxon serving as a reference in a phylogenetic definition
Crown taxon	a living species, or a clade that can be defined by living species, whose immediate outgroup is extinct
Total taxon	clade composed of a crown taxon plus all extinct outgroups more closely related to it than to another crown clade
Stem taxon	an extinct species or clade for which immediate outgroup includes at least one living member
Node-based definition	statement specifying a clade composed of the most recent common ancestor of two or more reference taxa and all descendants
Stem-based definition	statement specifying a clade composed of all descendants more closely related to one reference taxon (or taxa) than another (or others)
Complementary definitions	phylogenetic definitions with the same reference taxa (as in a node-stem triplet)
Reciprocal definitions	phylogenetic definitions with the same reference taxa but in reverse order (as in opposing stem-based taxa of a node-stem triplet)
Node-stem triplet	trio of taxa with complementary definitions consisting of a node-based taxon and two subordinate stem-based taxa
Taxonomic equivalence	statement of equivalence (or equation) in which the taxonomic content of one taxon equals that of subordinate taxa plus their common ancestor (e.g., A = B + C, as in a node-stem triplet)

prove to be more closely related to marsupials, as indicated by some recent molecular evidence (Penny and Hasegawa, 1997). Stability of taxonomic content is not related to the living or extinct status of reference taxa in a phylogenetic definition, but rather to the particular choice of reference taxa. For many node-based definitions of crown taxa, stability is also related to the certainty with which a specified basal dichotomy has been correctly ascertained.

That "widely used" names, such as Aves, should be assigned to crown taxa has been justified by arguing that (1) neontologists should report their observations accurately, referring only to crown clades rather than crown clades plus some of their extinct outgroups, and that (2) doing so follows long-standing taxonomic practice (de Queiroz and Gauthier, 1992). The condition in extinct outgroups for characters that are not preserved, however, is ambiguous. To suppose that *Archaeopteryx* lacked an apomorphic molecular sequence found in crown-group birds (Neornithes) is no more justified

than to suppose that the ancient bird had this sequence (Lee, 1996). Accelerated character-state transformation, in other words, is not more correct, accurate, or informative than delayed transformation.

Historical and current usage, such as may be ascertained, also does not clearly associate "widely used" names with crown taxa (Bryant, 1994; Lee, 1996). Regarding Mammalia, Rowe and Gauthier (1992:375–376) stated that "its universal usage by comparative biologists, except for some paleontologists, has long been for the crown clade Mammals, viz., the last common ancestor of Monotremata and Theria, and all of its descendants," and that "current debate over the name Mammalia highlights the fact that many paleontologists continue to operate in a system influenced in fundamental ways by a pre-Darwinian world view." Darwin was fairly clear regarding his conception of the taxonomic content of Aves, Mammalia, Marsupialia, and other groups with living representatives. For Darwin, these taxa included fossil stem groups. In *The Origin of the Species* (1859:268), he wrote:

(a)

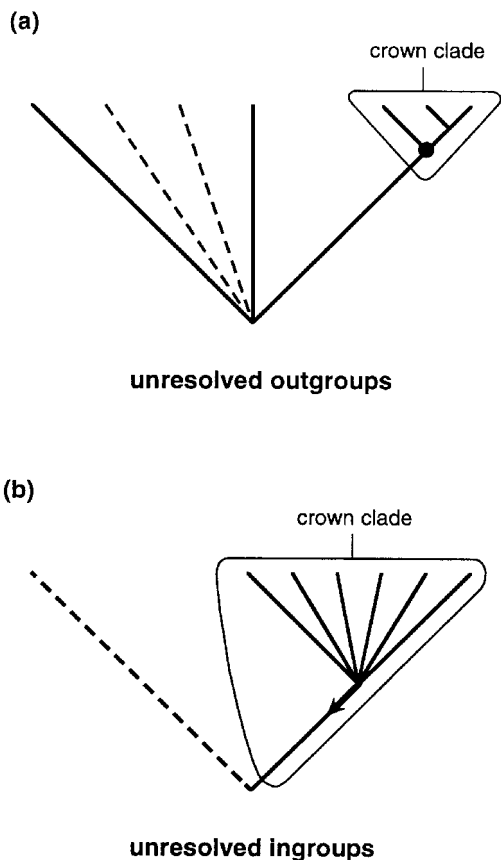**unresolved outgroups**

(b)

unresolved ingroups

FIGURE 2. Crown taxa, definitional types, and stability. (a) A crown clade with resolved basal relationships but with unresolved outgroup relationships may be defined with greatest stability as a node-based taxon. (b) A crown clade with unresolved basal relationships but with a stable extinct outgroup may be defined with greatest stability as a stem-based taxon. Crown groups are encircled; dashed lines indicate extinct taxa; a dot indicates a node-based definition; an arrow indicates a stem-based definition.

those groups which have, in known geological periods, undergone much modification, should in the older formations make some slight approach to each other; so that the older members should differ less in some of their characters than do the existing members of the same groups.

About mammals, Darwin (1859:522) wrote:

They [marsupials] appeared in an earlier geological period, and their range was formerly much more extensive than at present. Hence the Placentata are generally supposed to have been derived from the Im-

placentata, or Marsupials; not, however, from forms closely resembling the existing Marsupials, but from their early progenitors.

Most extant biologists also conceive of the taxonomic content of Mammalia in a similar fashion, whether they are evolutionary taxonomists, cladists, ecologists, or molecular biologists (e.g., Jollie, 1973:76–77; Futuyma, 1986:334; Walker, 1975). Higher taxa such as Mammalia and Aves have been associated with "key" characters and their associated functions. *Archaeopteryx*, with feathers, wings, and the capacity for flight, will always be considered a "bird" within Aves by the majority of biologists. Altering this equation for well-established monophyletic taxa is likely to engender confusion.

Stem-based definitions are less "consistent" than node-based definitions.—Schander and Thollesson (1995:264) suggested that stem-based definitions "refer to non-existing clades under some phylogenetic resolutions." They presented an example (Fig. 3a) in which relocation of one taxon (taxon A) results in a definition that refers to an "impossible" clade. Relocation of taxon A creates this internal inconsistency because the original ingroup reference taxon (taxon F) for stem-based taxon G now includes by definition the original outgroup taxon (taxon D). Schander and Thollesson (1995:264) concluded that "if a consistent system of taxonomy is wanted, stem-based definitions may better be avoided."

Similar internal "inconsistency," however, can be generated for node-based taxa as well (Fig. 3b). In this example, relocation of one taxon (taxon A) results in a definition that refers to an "impossible," or internally redundant, clade. Relocation of taxon A creates this internal inconsistency, because one of the original reference taxa (taxon E) for node-based taxon G now includes the other reference taxon (taxon F) and the common ancestor. Even if taxon E is regarded as a junior synonym of taxon G after relocation of taxon A, the definition of taxon G must be changed (because taxon E cannot be used as a reference taxon).

Rather than demonstrating any particular structural weakness of stem-based definitions, these examples (Fig. 3) better
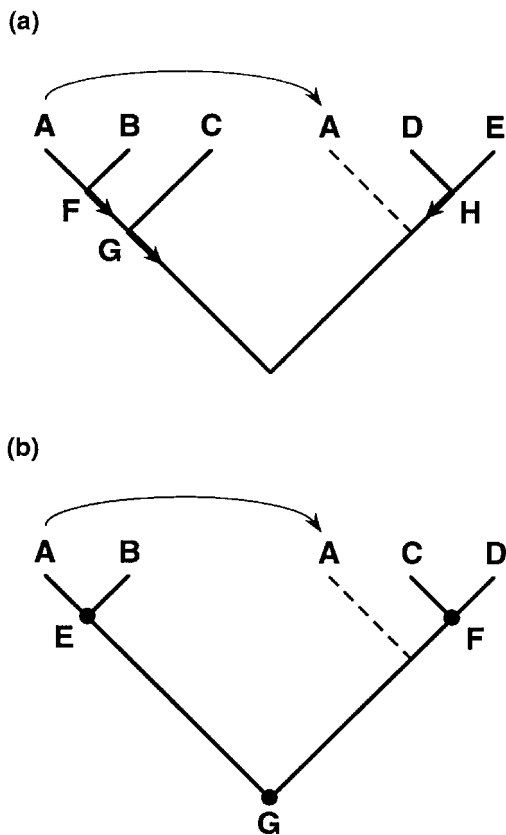
(a)

(b)

FIGURE 3. Problems associated with maximally inclusive reference taxa. (a) Relocation of taxon A results in an "inconsistent" stem-based definition for taxon G, because it defines a group that cannot exist under the alternative phylogenetic hypothesis (after Schander and Thollesson, 1995:Fig. 3). Schander and Thollesson defined stem-based taxon G as "F and all taxa sharing a more recent common ancestor with it than with D," and stem-based taxon F as "A and all taxa sharing a more recent common ancestor with it than with C." Taxon G is "inconsistent" after relocation of taxon A, because its ingroup reference taxon (taxon F) now includes by definition the outgroup reference taxon (taxon D). The "inconsistency," however, disappears if a less-inclusive ingroup reference taxon had been chosen for taxon G (i.e., taxon A or B rather than taxon F). In this case, taxon G exists after relocation of taxon A (either as a synonym of taxon A, or as a group including taxa B and C, depending on whether taxon A or B was used as the ingroup reference taxon). (b) Relocation of taxon A results in an "inconsistent" node-based definition for taxon G, because it defines a group that cannot exist under the alternative phylogenetic hypothesis. In this example, taxon G is node-based and defined as "taxon E, taxon F, their common ancestor, and all

descendants"; taxon E is defined as "taxon A, taxon B, their common ancestor, and all descendants"; and taxon F is node-based and defined as "taxon C, taxon D, their common ancestor, and all descendants." Taxon G is "inconsistent" after relocation of taxon A, because one of its reference taxa (taxon E) now includes by definition the other reference taxon (taxon F) and the common ancestor. The reference taxa for a node-based taxon are presumed to be exclusive of one another and their common ancestor; after relocation of taxon A, these conditions are violated for taxon G as originally defined. The "inconsistency," however, disappears if a less-inclusive reference taxon had originally been selected for taxon G (i.e., taxon A or B rather than taxon E). In this case, taxon G exists after relocation of taxon A (either unchanged in its taxonomic content, or as a less-inclusive taxon that excludes taxon B, depending on whether taxon B or A was used as an ingroup reference taxon, respectively). Dots indicate node-based definitions, and arrows indicate stem-based definitions.

illustrate problems that arise as a result of selecting maximally inclusive reference taxa—the most common kind of reference taxa currently chosen in phylogenetic definitions (see *Reference Taxa* below). If Schander and Thollesson had chosen a more nested ingroup reference taxon in their stem-based definition of taxon G (i.e., taxon A or B rather than taxon F; Fig. 3a), relocation of taxon A poses no special problems. In this case, taxon G either refers to the same clade minus taxon A (if taxon B is the ingroup reference taxon) or becomes a potential junior synonym of taxon A (if taxon A is the ingroup reference taxon). Similarly, for node-based taxon G (Fig. 3b), a more nested reference taxon (taxon A or B rather than taxon E) removes any inconsistency after relocation of taxon A. Taxon G refers either to the same clade (if taxon B is an ingroup reference taxon) or to a less inclusive clade that excludes taxon B (if taxon A is an ingroup reference taxon).

Node- or stem-based definitions better reflect original usage and content.—Gauthier (1986:12,13), for example, suggested that his stem-based definition for Haeckel's (1866) taxon Ornithurae ("Extant birds and all other taxa, such as *Ichthyornis* and Hesperornithes [sic], that are closer to extant birds than is *Archaeopteryx*") was "in keeping with its original intent" (Fig. 4). Chiappe (1991, 1996), on the other hand, argued that

Ornithurae is better defined as a less inclusive node-based taxon, with Hesperornithiformes and Neornithes (crown-group birds) as reference taxa (Fig. 4). Initially, Chiappe
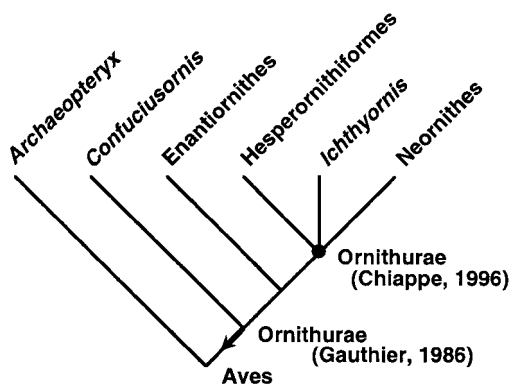

FIGURE 4. Cladogram of basal avians showing recent use of the higher taxon Ornithurae (Haeckel, 1866). The dot indicates a node-based definition, and the arrow indicates a stem-based definition.

(1991:337) reasoned that a node-based definition of Ornithurae was more appropriate because it excluded several newly discovered basal avians that lacked some of the synapomorphies in Gauthier's (1986) diagnosis of Ornithurae. Later Chiappe argued that a node-based definition of Ornithurae "better reflects the composition of the clade before to [sic] the application of phylogenetic definitions" (1996:205).

   Neither Gauthier's nor Chiappe's definitions have historical precedence. Haeckel (1866) coined Ornithurae ("bird tail") for the reduced number of vertebrae and unique coossification of the distal tail (as a single bone, the pygostyle) that characterizes all extant birds, in contrast to the condition in *Archaeopteryx*. Haeckel's taxon owes its existence to the then-recent discovery of *Archaeopteryx*, a primitive bird with a tail comprising over 20 vertebrae. Contrary to Chiappe (1996), Ornithurae predates the description of the extinct short-tailed genera *Hesperornis* (Marsh, 1872a) and *Ichthyornis* (Marsh, 1872b) or the taxon Hesperornithiformes (Fürbringer, 1888). Ornithu-

rae, as noted by Gauthier (1986), has been supplanted by Neornithes (Gadow, 1893), a taxon referring to crown-group avians. Based on original intent, an argument could be made that Haeckel's Ornithurae should supplant Neornithes on grounds of priority, or that it be defined explicitly as an apomorphy-based taxon based on the presence of a pygostyle, which is now known to be lacking in birds other than *Archaeopteryx* (Forster et al., 1998). In either case, the definition would circumscribe clades other than those identified by either Gauthier or Chiappe.

   For most taxon names, the original definition, if explicitly stated at all, is character-based or taxon-based (a list of included taxa) and lacks a relational phrase about ancestry that would specify *potential* membership. The *intent* of the original author of a taxon to include or exclude unknown, or repositioned, taxa usually cannot be interpreted unambiguously. This problem is not widely appreciated.

APOMORPHY-BASED DEFINITIONS

   An apomorphy-based definition specifies the boundaries of a clade by identifying the "first ancestor with a particular synapomorphy" (de Queiroz and Gauthier, 1990:310) and encompassing all of its descendants. An apomorphy-based taxon includes all descendants, whether or not the synapomorphy is maintained. Apomorphy-based definitions are subject to three problems that are not relevant to node-based and stem-based definitions: variation in character coding, optimization ambiguity, and homoplasy (Bryant, 1994; Schander and Thollesson, 1995; Sereno, 1998). These problems have long been associated with traditional use of "key" characters to define taxa. Thus far, few apomorphy-based definitions have been erected, and this definitional form should be avoided.

   Character ambiguity.—Consider an apomorphy-based definition for Aves (including *Archaeopteryx* and Ornithurae) based on the presence of "feathers," the "key character" usually associated with this taxon. First, the character coding—"Feathers: ab-

sent (0); present (1)"—is only one of several possibilities. A "feather" is composed of a rachis, barbs, barbules, and other components that may be better coded as separate characters. The character on which an apomorphy-based definition is based, therefore, may be recoded as two or more characters with independent character-state distributions. Indeed, this is an expectation for most qualitative or quantitative characters; that is, once the transformations become better known, intermediate states or multiple characters will emerge, as recent discoveries are beginning to reveal with regard to feathers (Chen et al., 1998; Ji et al., 1998), or with regard to the characters historically associated with Mammalia (Rowe, 1988; Rowe and Gauthier, 1992; Bryant, 1994; de Queiroz, 1994).

The apomorphy-based definition given as an example by de Queiroz and Gauthier (1990:310) is subject on the same grounds to interpretational ambiguity: "Tetrapoda" is defined as the "first vertebrate to possess digits (i.e., hands and feet rather than fins) and all of its descendants." What exactly constitutes a "digit" versus a "fin" or "paddle"? In fact, this is a significant question given recent fossil discoveries (Daeschler and Shubin, 1998). Can this feature, or list of features, be observed in extinct forms? And what happens if we learn that "digits" evolved first on the hands and later on the feet? These are common problems in the interpretation of character data, which should not be extended to taxonomic definitions.

Optimization ambiguity.—Apomorphy-based definitions also do not specify a single ancestor when the chosen apomorphy has an ambiguous optimization. Characters with ambiguous character-state optimization are commonplace in systematics and arise from missing data (lack of preservation, strong transformation) and homoplasy (Bryant, 1994; Schander and Thollesson, 1995; Sereno, 1998). There may be many equally parsimonious positions for a particular character state on the shortest cladogram.

Homoplasy.—Homoplasy may involve the stable placement of an apomorphy at more than one node (Bryant, 1994). This can create an impasse for an apomorphy-based definition, because the apomorphy might identify more than one clade. Invoking time as an arbiter (such as "the first taxon that has ...") invites ambiguities associated with temporal origin (such as missing ancestral lineages).

PHYLOGENETIC DEFINITIONS: TERMINOLOGY

The terminology used in the following discussion of phylogenetic taxonomic definitions is clarified below and tabulated (Tables 1, 2). Some of these terms are given more specific meaning than in the current literature; others are new or recently introduced (Sereno, 1998).

Taxonomic definition and *diagnosis* are defined here much as they have been described previously by phylogenetic systematists (de Queiroz and Gauthier, 1990, 1992) and their forebears (Buck and Hull, 1966; Farris, 1976; Ghiselin, 1984) (Table 1). De Queiroz and Gauthier (1992:461) defined phylogenetic definitions as "statements specifying the meanings of taxon names (words); they are stated in terms of ancestry." The "meaning of a taxon name" in phylogenetic taxonomy concerns its taxonomic content, or membership, as delimited by a relational statement. Taxonomic definitions, therefore, are regarded here as "extensional" (ostensive) statements of relationship that specify the taxonomic content of a taxon (de Queiroz, 1992).

Taxonomic diagnosis, on the other hand, involves the "intensional" characterization of members of a clade, as recognized by the shared apomorphies present in the common ancestor (Buck and Hull, 1966; Farris, 1976; Ghiselin, 1984; Rowe, 1987; de Queiroz and Gauthier, 1990, 1992; de Queiroz, 1992). According to de Queiroz and Gauthier (1992:461), diagnoses are "statements specifying how to determine whether a given species or organism is a representative of the taxon (clade) to which a particular name refers; they are most commonly stated in terms of characters." Strictly speaking, shared apomorphies (derived-character states), rather than characters, determine the membership of a particular clade or species.

Taxonomic content is used here to refer to all *existing* and *potential* taxa included by the phylogenetic definition of a taxon (Table 1). The smallest number of subordinate taxa that can fulfill this definition of taxonomic content is two; a redundant taxon (i.e., a single subordinate taxon) would have an identical phylogenetic definition and, thus, is uninformative in phylogenetic taxonomy (Farris, 1976). The taxonomic content of taxon C, therefore, could be "taxon A, taxon B, their most recent common ancestor, and all descendants," if taxa A and B are defined extensionally to encompass all existing and potential subordinate taxa currently within taxon C.

A list of included taxa is insufficient to unambiguously specify taxonomic content, because there will always be potential members of a taxon that lie outside any list of included taxa, whether or not the included taxa themselves have phylogenetic definitions. An exhaustive list of included taxa with stem-based definitions, for example, does not specify membership for their hypothetical immediate common ancestor (a node-based taxon) or more distant known or hypothetical outgroups (a stem-based taxon). Two reference taxa and a relational statement about ancestry are necessary and sufficient to specify taxonomic content, because they unambiguously specify membership for all existing and potential members of a taxon.

A *reference taxon* is a species or higher-level taxon that is used to delimit a phylogenetic definition (Table 1). For example, "taxon A, taxon B, their most recent common ancestor, and all descendants" uses A and B as reference taxa for a node-based taxonomic definition. Previously cited as "reference points"

(Schander and Thollesson, 1995), these important components of phylogenetic definitions have not been distinguished terminologically.

Several kinds of reference taxa can be identified (Fig. 5; Table 2). *Ingroup* and *outgroup* reference taxa are located within or outside the taxon they define, respectively (Fig. 5a, b). An *inclusive* reference taxon includes several ingroup or outgroup taxa; a maximally inclusive reference taxon includes the largest possible ingroup or outgroup reference taxon (Fig. 5c). Basal and nested reference taxa are less inclusive and, as their name suggests, are located at the base or nested within an ingroup or outgroup taxon (Fig. 5d, e).

A *crown taxon* denotes a living species, or a clade defined by using living reference taxa, whose immediate outgroup is extinct (Figs. 6, 7; Table 1). The extinct outgroup is a necessary component of the definition. Jefferies (1979:449) coined the term "crown group" to replace Hennig's (1969) *group, which Jefferies defined as "the latest common ancestor of all living members of group 1, plus all descendants of this ancestor, whether these are living or extinct." The existence of an immediate outgroup that is extinct is implied and was clearly present in the examples given by Hennig and Jefferies. Patterson (1981:207) also clearly used crown- and stem-groups in this manner, stating that the "stem-group contains fossil taxa that exhibit some, but not all of the characters of the crown-group." Crown- and stem-group terminology has been used exclusively for higher taxa bounded by living species that have extinct outgroups (e.g., de Queiroz and Gauthier, 1992; Forey, 1992; McKenna and Bell, 1997). Although it may

TABLE 2.   Definitions for the various kinds of reference taxa.

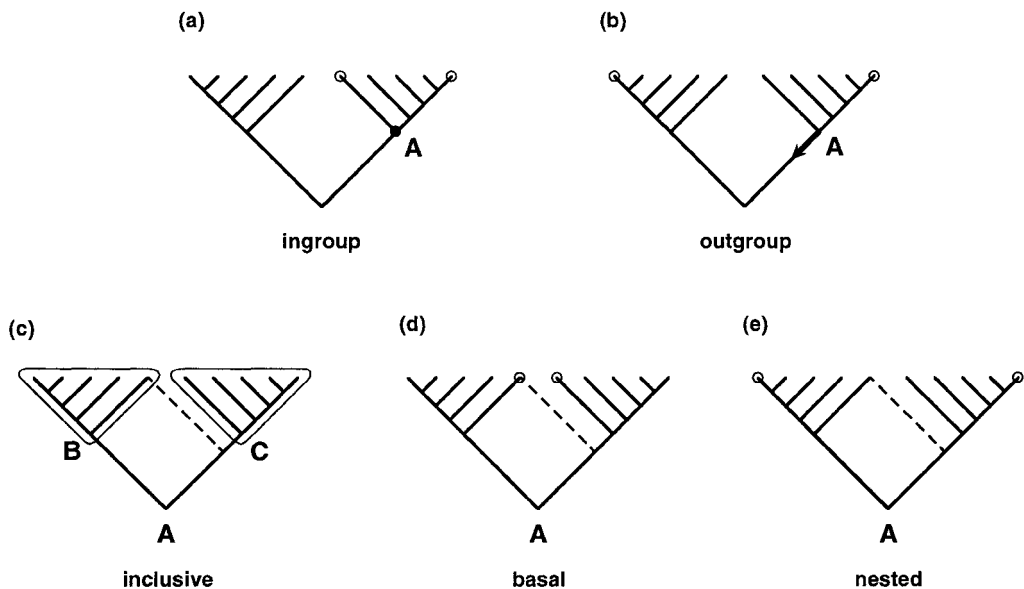| Reference taxon type	Definition
Ingroup	reference taxon included within a taxon by definition
Outgroup	reference taxon excluded from a taxon by definition
Inclusive	most inclusive reference taxon among ingroup or outgroup taxa
Basal	reference taxon near the basal furcation of a taxon
Nested	reference taxon distant from the basal furcation of a taxon
Vernacular	vernacular name used as a reference taxon (e.g., "birds" or "extant birds")

FIGURE 5. Kinds of reference taxa. (a) An ingroup reference taxon is included by definition within a taxon; a node-based definition is based on two ingroup reference taxa. (b) An outgroup reference taxon is excluded by definition from a taxon; a stem-based definition is based on one ingroup and one outgroup reference taxon. (c) An inclusive reference taxon is the most inclusive ingroup or outgroup reference taxon possible; inclusive reference taxa B and C for taxon A allow definitional ambiguity with change in the relationship of a basal taxon. (d) A basal reference taxon is located at, or near, the base of an ingroup or outgroup taxon; basal reference taxa allow definitional ambiguity when their relationships change. (e) A nested reference taxon is remote from the base of an ingroup or outgroup reference taxon; nested reference taxa are unaffected by changes in the relationships of basal taxa. Reference taxa are encircled; a dot indicates a node-based definition; an arrow indicates a stem-based definition; a dashed line indicates a change in position of a basal taxon.
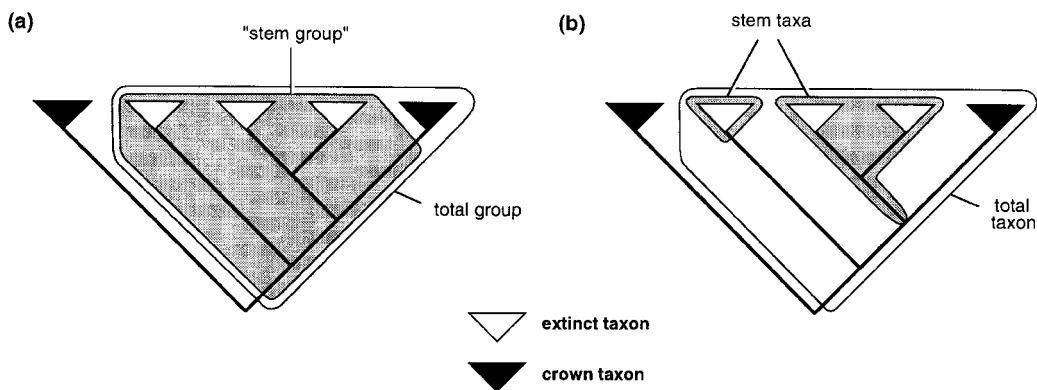

FIGURE 6. Crown, stem, and total taxa. (a) Total group composed of a crown group plus a paraphyletic "stem group," as conceived by Hennig (1969, 1983) and Jefferies (1979). (b) Total taxon composed of a crown taxon and monophyletic stem taxa, as used in this article. One stem taxon (left) is shown with a node-based definition and the other (right) with a stem-based definition.

(a)

crown taxa
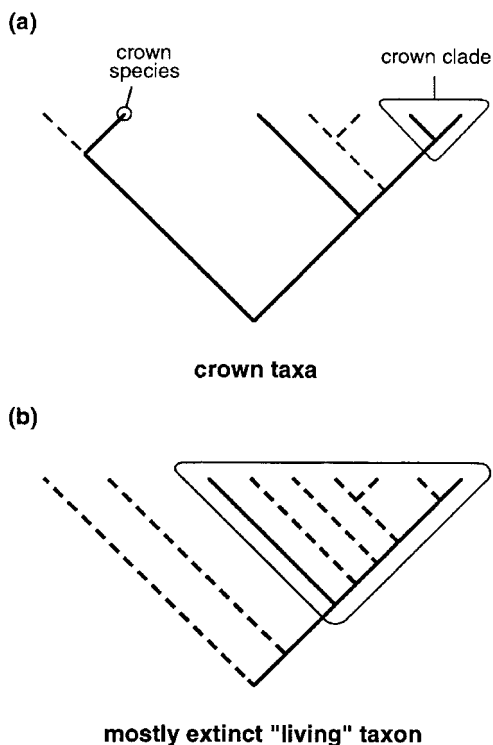
(b)

mostly extinct "living" taxon

FIGURE 7. Crown species or clades (enclosed) must have an extinct (stem) taxon (dashed) as an immediate outgroup. (a) Crown species and crown clades are delimited by the presence of an immediate outgroup that is extinct. (b) Often referred to as "extant" taxa, crown taxa may be composed primarily of extinct taxa. "Living" and "extinct" conditions are treated as if they constitute plesiomorphic and apomorphic character states, respectively, of an irreversible character "state of being." Crown clades, under this conceptualization, are plesiomorphy-based taxa bounded by two living taxa for which the immediate outgroup is extinct.

be hypothetically true that "every Recent group, whatever its position in the hierarchy, may be expected to have had a stem-group" (Patterson, 1981:207), clades bounded by living taxa with immediate outgroups that are also living have not been considered crown taxa.

Some recent definitions of crown taxa, however, have set aside Hennig's and Jefferies' formulations as well as general usage in the literature. A less restrictive meaning has been proposed based solely on reference taxa that include liv-

ing representatives. De Queiroz and Gauthier (1992:469) defined crown taxa as "clades within which both branches of the basal dichotomy are represented by extant descendants," without mention of the condition of the immediate outgroup. Yet, all of their examples of crown clades have immediate outgroups that are extinct. Regarding their craniate taxonomy, they stated that each stem-based taxon is composed of a "crown taxon plus all extinct taxa more closely related to it than to any other crown clade" (de Queiroz and Gauthier, 1992:475). Lee (1996:1103) proposed that "a crown clade definition results when both taxa implicated in a node-based definition ... are represented by extant forms." Although the condition of the immediate outgroup is never mentioned, Lee (1996:1103) stated that crown taxon definitions apply to clades that consist of "the most recent common ancestor of *all* extant forms, and all its descendants" (emphasis added). The presence of an immediate extinct outgroup is implied.

The definition of a crown taxon presented here, in contrast, does not specify ingroup structure (such as the presence or absence of a basal dichotomy) or definitional type (node- or stem-based) and explicitly states that the immediate outgroup(s) is extinct (Fig. 7a; Table 1). Crown taxa, so defined, can be mapped unerringly on a cladogram if (1) "living" and "extinct" conditions are treated as primitive and derived character-states, respectively, of the character "current state of being," and (2) if this character is regarded as irreversible (i.e., Camin–Sokal parsimony). The living state, in other words, is always regarded as plesiomorphic. The irreversible status of the character (state of being) is important, because it identifies crown groups on cladograms of living and extinct taxa that would have an ambiguous status if character-state reversal were allowed (Fig. 7b). This character-state conceptualization of the definition of a crown group captures the essence of crown taxa, which have always been used to refer to maximally inclusive taxa defined by extant representatives, not simply any "extant-

bounded" taxon. Discovery of extinct sister taxa provides the opportunity to recognize crown taxa.

Crown taxa are often regarded as clades that are bounded uniquely by living ingroups. The use of living reference taxa, however, is most often an arbitrary decision. Any extinct ingroup taxon that is most closely related to a living reference taxon could be used to delimit the same clade (Fig. 7b). A crown clade is defined uniquely by living reference taxa only when such taxa do not have extinct sister taxa (Fig. 7b), a fact that is not widely appreciated. Crown taxa, in other words, do not have any special biological significance beyond signaling that particular reference taxa have survived to an arbitrary datum (the Recent). Informal reference to crown taxa as "living" or "extant" taxa, therefore, can be misleading, because a crown taxon may be composed predominantly of extinct species, as is the case with the crown taxon Crocodylia (Fig. 7b; Brochu, 1997). "Extant crocodilians," if taken literally, refers to a small paraphyletic subset of Crocodylia. Explicit reference to "crown crocodilians" or simply "Crocodylia" is preferable to use of the terms "living" or "extant."

A living taxon is also essential to delimit a maximally inclusive clade that is entirely extinct—a *stem taxon*, as here defined. A stem taxon denotes an extinct species or clade, the immediate outgroup of which includes at least one living member (Fig. 6b). Many extinct taxa are not stem taxa, because their immediate outgroup is itself extinct. Such extinct taxa are members of more inclusive stem taxa. Stem taxa can have either node- or stem-based definitions (Fig. 6b). Stem taxa and their associated less inclusive crown taxon make up the taxonomic content of their respective total taxon.

Stem taxa, as here defined, should not be confused with "stem group" (Fig. 6a), which was first used in phylogenetic systematics by Hennig ("stammgruppe"; 1969, 1983). Hennig's (1983:15) vague formulation of the concept specified all extinct species "which can be shown probably to belong to a particular monophyletic group of the phylogenetic system but which are probably no

closer [sic] related to one subgroup among recent animals than to another" (translated from the German; Ax, 1987:224). Jefferies (1979) clarified the term to include all extinct outgroups of a crown taxon (Fig. 6a). Other authors have developed more convoluted definitions that approximate the same end. Wiley (1981:217), for example, defined a "stem group" as "all fossil species of a group more primitive than the most primitive Recent species of the same group." These "stem groups" are paraphyletic and often circumscribe the same "ancestral groups" and "evolutionary grades" that have muddied phylogenetic thinking since Darwin. Only a few cladists continue to endorse such concepts (Smith, 1994). Ax (1987:224) proposed a cumbersome solution to the problem of paraphyletic "stem groups" by dividing them into ancestral "stem lineages" and "representatives of the stem lineage." The latter are monophyletic taxa most simply referred to as stem taxa.

A *total taxon* is regarded here as a clade composed of a crown taxon plus all extinct outgroups more closely related to it than to another crown taxon (Fig. 3; Table 1). Hennig's (1969) and Jefferies' (1979) formulations of the concept are wordy but similar in effect, because they restricted inclusion of extinct outgroups to those most closely related to a particular crown taxon. De Queiroz and Gauthier (1992:470) defined a total clade as a "more inclusive clade consisting of the crown clade plus its extinct outgroups." This definition is incomplete, because it does not specify which extinct outgroups are included. A total taxon explicitly includes only those extinct outgroups most closely related to a particular crown taxon.

Node-based and *stem-based definitions* are constructed by (1) identifying reference taxa and (2) attaching to them a relational phrase that either includes their most recent common ancestor and all descendants (node-based) or limits inclusion to all descendants closer to one or more of the reference taxa (stem-based) (de Queiroz and Gauthier, 1992) (Fig. 1; Table 1). In node-based definitions, the phrase "least-inclusive clade" can replace "common ancestor and all descendants" for cladists sensitive to the allusion to

ancestors (Lee, 1998). Listing included taxa without a relational phrase about ancestry, or simply designating taxa as node- or stem-based without reference taxa, does not constitute a phylogenetic definition, because the boundaries of the taxon in question are not specified (de Queiroz and Gauthier, 1990; Bryant, 1996).

In the current literature, there exists some confusion over what is necessary and sufficient for node- and stem-based definitions. Bryant (1996:185) suggested that a more inclusive taxon be inserted within a stem-based definition, such as "all members of taxon A that are more closely related to taxon B than to taxon C." Including taxon A, however, is not necessary and may lead to internal conflicts within the definition (e.g., when taxon A is included within taxon B or C on an alternative phylogeny). The goal of a phylogenetic definition should be to unambiguously identify a specific common ancestor given *any* arrangement of taxa, rather than to identify such an ancestor only on a subset of phylogenetic hypotheses.

Meng et al. (1994) and Wyss and Meng (1996) introduced a "stem-modified node-based definition" to effectively unite crown-group rodents, whose basal relationships are uncertain. Their preferred definition of Rodentia ("the clade stemming from the most recent common ancestor of *Mus* and all Recent mammals more closely related to *Mus* than to Lagomorpha or members of any other eutherian 'order'"; Wyss and Meng, 1996:562) is better interpreted as a special stem-based rather than node-based definition, because it uses only one ingroup reference taxon (*Mus*) and one or more outgroup reference taxa (e.g., Lagomorpha). Their definition merely adds a phrase ("the clade stemming from the most recent common ancestor") to an otherwise typical stem-based definition based on living reference taxa. That phrase ensures inclusion of all extinct taxa within the crown clade. Bryant's (1996:185) reformulation of this definition, on the other hand, is clearly a node-based definition, because he specified only ingroup reference taxa ("the most recent common ancestor of the species referred to Rodentia by Wilson and Reeder

(1993) and all of its descendants"). These examples underscore the need for clarification of terms (Tables 1, 2).

A *node-stem triplet* (NST) is a definitional configuration involving three taxa, a node-based taxon composed of two stem-based subordinate taxa (Sereno, 1997, 1998; Fig. 8):

Taxon A: Taxa b and c, their most recent common ancestor, and all its descendants.

Taxon B: All taxa closer to taxon b than to taxon c.

Taxon C: All taxa closer to taxon c than to taxon b.
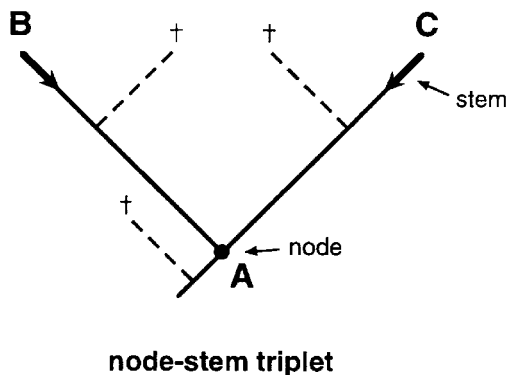

node-stem triplet

FIGURE 8. Definitional triumvirate termed a node-stem triplet, composed of a node-based taxon (A) and two subordinate stem-based taxa (B and C), which incorporates added or repositioned taxa (dashed lines with daggers) without changing the relative taxonomic content of taxon A, B, or C (as expressed by the equivalence statement A = B + C). A dot indicates a node-based definition; an arrow indicates a stem-based definition.

Taxa b and c (not shown in Fig. 8) constitute reference taxa within taxon B and C, respectively. Although not essential, the same reference taxa (b, c) may be used for the trio of taxa in a NST, which then specify *complementary* definitions. The pair of stem-based sister taxa, in addition, may have *reciprocal* definitions that utilize the same reference taxa in antipodal positions. The word "taxa" in the stem-based definitions listed above may optionally be replaced by reference to a more inclusive taxon (Bryant, 1996). A stem-based definition for Saurischia, for example, might read, "All dinosaurs more closely related

to Neornithes than to *Triceratops*" (Sereno, 1998).

Taxonomic equivalence is a statement (or equation) in which one taxon is shown to be equivalent in *current* and *potential* taxonomic content to subordinate taxa plus their common ancestor (Table 1). The NST described above is composed of a node-based taxon (A) and two subordinate stem-based taxa (B, C) and can be written as the equivalence statement "A = B + C," because taxon A, by definition, is equivalent to taxa B and C plus their most recent common ancestor.

PHYLOGENETIC DEFINITIONS: RATIONALE

The first step toward a phylogenetic taxonomy involved the distinction between definition and diagnosis (Buck and Hull, 1966; Griffiths, 1973, 1974; Farris, 1976; Ghiselin, 1984; Rowe, 1987). A second step involved the formulation of node- and stem-based definitions by use of reference taxa and relational statements about ancestry (de Queiroz and Gauthier, 1992). A third step, outlined here, involves the specification of guidelines for selecting effective reference taxa and criteria for positioning node- and stem-based definitions.

Phylogenetic rearrangements and *introduction of new taxa* are the principal factors that destabilize taxonomic content. Stability of taxonomic content is enhanced by choosing (1) effective reference taxa and (2) an effective configuration of phylogenetic definitions.

Reference Taxa

Position.—Given a particular phylogenetic hypothesis, the taxonomic content of a phylogenetic definition is altered only if the reference taxa specify an alternative common ancestor. Given a basal dichotomy with one reference taxon on each side, the identification of the immediate common ancestor is not affected by relocation of a reference taxon on its side of the basal dichotomy (Fig. 9, move 1). Relocation of a reference taxon to the opposing side or to a location outside the basal dichotomy identifies a different common ancestor

(Fig. 9, moves 2, 3). For a taxon with a basal dichotomy, an alternative common ancestor can be identified only if one reference taxon is relocated to the opposing side or outside the basal dichotomy. Are some reference taxa more likely to be relocated in this manner than others?
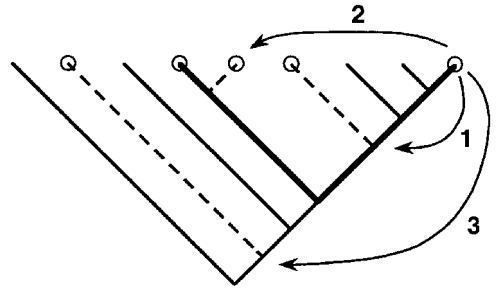

FIGURE 9. Three possible outcomes from relocation of a reference taxon. 1 = relocation on the same side of a basal dichotomy has no effect, 2 = relocation to the opposite side of the basal dichotomy identifies a less inclusive common ancestor, 3 = relocation outside the basal dichotomy identifies a more inclusive common ancestor. Reference taxa are encircled, and the original basal dichotomy is shown with heavy lines.

First, a taxon positioned near the basal dichotomy of a cladogram is always easier to relocate to the opposing side (i.e., requires fewer additional steps) than is a nested taxon, if all other factors are regarded as equal among ingroup taxa (e.g., completeness, missing data, and homoplasy). Unlike a basal taxon (Fig. 5d), a nested taxon shares synapomorphies that increase its phylogenetic (patristic) distance from the common ancestor (Fig. 5e). Thus, stability of taxonomic content is enhanced in phylogenetic definitions if nested reference taxa are chosen that are located at some distance (at least several nodes away) from the basal dichotomy. This distance is easy to assess quantitatively for various nested taxa on a given phylogeny.

Second, maximally inclusive reference taxa (Fig. 5c) have greater potential to create unnecessary taxonomic redundancy and internal inconsistencies after relocation of taxa (Fig. 3). This is true for both node-

and stem-based taxa. In the current litera-ture, maximally inclusive reference taxa are often used in phylogenetic definitions. In their tetrapod classification, for example, de Queiroz and Gauthier (1992:475) used suc-cessive maximally inclusive reference taxa for all definitions—a pattern here termed *re-cursive referencing* (Fig. 10a). Recursive refer-encing can have undesirable consequences because it shifts the burden of a more precise definition to a slightly less inclusive higher taxon. This higher taxon may not be de-fined, or it too may be affected by an alter-native configuration of basal relationships. Salgado et al. (1997), for example, chose Prosauropoda and Sauropoda as maximally inclusive reference taxa for Sauropodomor-pha. If a basal sauropodomorph is reposi-tioned, the definitions of Prosauropoda and Sauropoda must be determined to assess po-tential taxonomic consequences. However, Salgado et al. (1997) used Prosauropoda as a terminal taxon without definition. Sauropoda was defined as a node-based taxon on the basis of two maximally in-clusive reference taxa, *Vulcanodon* and Eu-sauropoda (all other sauropods). Thus, if the incompletely known basal sauro-pod *Vulcanodon* is reinterpreted as be-ing more closely related to prosauropods, Sauropoda and Sauropodomorpha would become synonyms.

If, on the other hand, Sauropodomorpha is defined on the basis of the nested ref-erence taxa—the prosauropod *Plateosaurus* and sauropod *Saltasaurus* (Sereno, 1998)— a change in the phylogenetic position of any other prosauropod or sauropod has no effect on the taxonomic con-tents of Sauropodomorpha or Sauropoda. Sauropodomorpha can be defined as "*Pla-teosaurus*, *Saltasaurus*, their most recent common ancestor, and all descendants"; and Sauropoda can be defined as "all sauropodomorphs more closely related to *Saltasaurus* than to *Plateosaurus*." The phylo-genetic position of *Vulcanodon* has no effect on the identification of respective common ancestors.

Completeness.—Poorly known taxa are more likely to be phylogenetically unsta-ble than well-known taxa (Wilkinson, 1995).
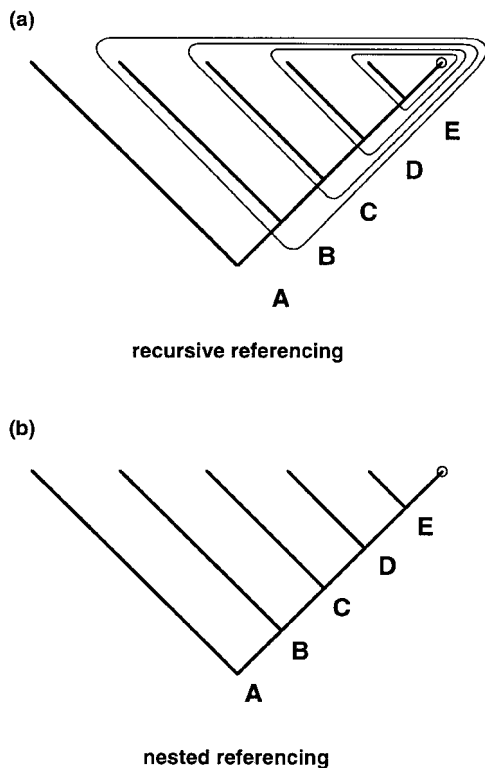

FIGURE 10. Referencing styles. (a) Recursive refer-encing selects a succession of the most inclusive ref-erence taxa. (b) Nested referencing selects that same remote reference taxon. Reference taxa are encircled.

Taxa with numerous missing entries in a phylogenetic analysis usually reduce phylo-genetic resolution by generating numerous equally parsimonious trees. Well-known ref-erence taxa, even if they are somewhat less nested, are preferable to poorly known taxa (Fig. 11). A single poorly known basal taxon is least stable and, therefore, least desirable as a reference taxon.

Use with Recent or extinct filters.—"Recent" ("extant" or "living") or "extinct" may be used in phylogenetic definitions of crown and stem taxa, respectively, to help to stabi-lize taxonomic content in the face of poorly resolved basal relationships. For a crown taxon with a stem-based definition, a Recent filter can restrict the boundaries of a clade to living taxa (Meng et al., 1994; Wyss and Meng, 1996). Rodentia, for example, could

be defined as "the least inclusive clade composed of Recent mammals more closely related to *Mus* than to Lepus." For stem taxa, in contrast, a stem-based definition using an "extinct" filter can restrict inclusion to extinct taxa when outgroup relationships are uncertain. Moas, for example, could be defined as "all extinct species more closely related to *Dinornis* than to *Struthio*," the ostrich.

Number.—The number of reference taxa that have been used in phylogenetic definitions has varied from two to infinity. Gauthier (1986:14), for example, defined Aves as a node-based taxon including "all descen-

nestedness versus completeness

● well known taxon

◐ poorly known taxon
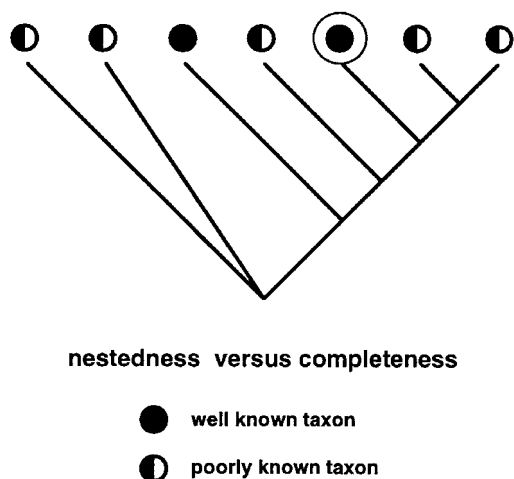
FIGURE 11. Nestedness and completeness in the selection of reference taxa. A nested, well-known taxon is more likely to constitute a stable reference taxon (encircled).

dants of the most recent common ancestor of Ratitae, Tinami, and Neognathae." Three reference taxa, rather than two, were chosen, presumably because of unstated uncertainty in the phylogenetic relationships among these three avian subgroups. De Queiroz and Gauthier (1990:310, 1992:461) suggested that phylogenetic definitions should utilize only two reference taxa—but then erected definitions for total taxa that used as reference taxa one crown clade plus "all extinct

taxa more closely related to it than to any other crown clade" (de Queiroz and Gauthier, 1992:475).

Listing more than a pair of reference taxa does not seem to compromise a node-based phylogenetic definition in any discernible way. Node-based definitions are constructed with what may be termed *internal* inclusion. A contradictory relationship between multiple ingroup reference taxa for a node-based definition is impossible to construct with internal inclusion, because a clade always exists that includes all ingroup reference taxa. Multiple ingroup reference taxa effectively stabilize the taxonomic content of a node-based taxon if basal relationships are uncertain (Fig. 12a).

For stem-based definitions, multiple outgroup reference taxa effectively stabilize the taxonomic content of a taxon if outgroup relationships are uncertain (Fig. 12b). Multiple ingroup reference taxa for stem-based definitions can create a contradictory situation, if one (or more) ingroup reference taxa are repositioned closer to the outgroup reference taxon (Fig. 12c) or to a position outside the clade as originally defined. In this case, the *external* inclusion fundamental to a stem-based definition creates a contradictory relationship among the original ingroup reference taxa, which are no longer more closely related to each other than to the outgroup reference taxon (Fig. 12c). In the example shown here, relocation of ingroup reference taxon D creates a contradictory relationship with a second ingroup reference taxon E, if taxon A has the stem-based definition of "all descendants more closely related to taxon D and E than to taxon B." An inclusive ingroup reference taxon, however, can manifest the same problems in this circumstance (Fig. 12c). If taxon A is defined as "all descendants more closely related to taxon C than to taxon B," and if taxon C is defined on the basis of taxon D, a contradictory definition results with relocation of taxon D as shown; taxon A now subsumes its original ingroup reference taxon C. Such contradictory definitions can be avoided if a nested ingroup reference taxon (or nested taxa) are selected, such as taxon E, its un-
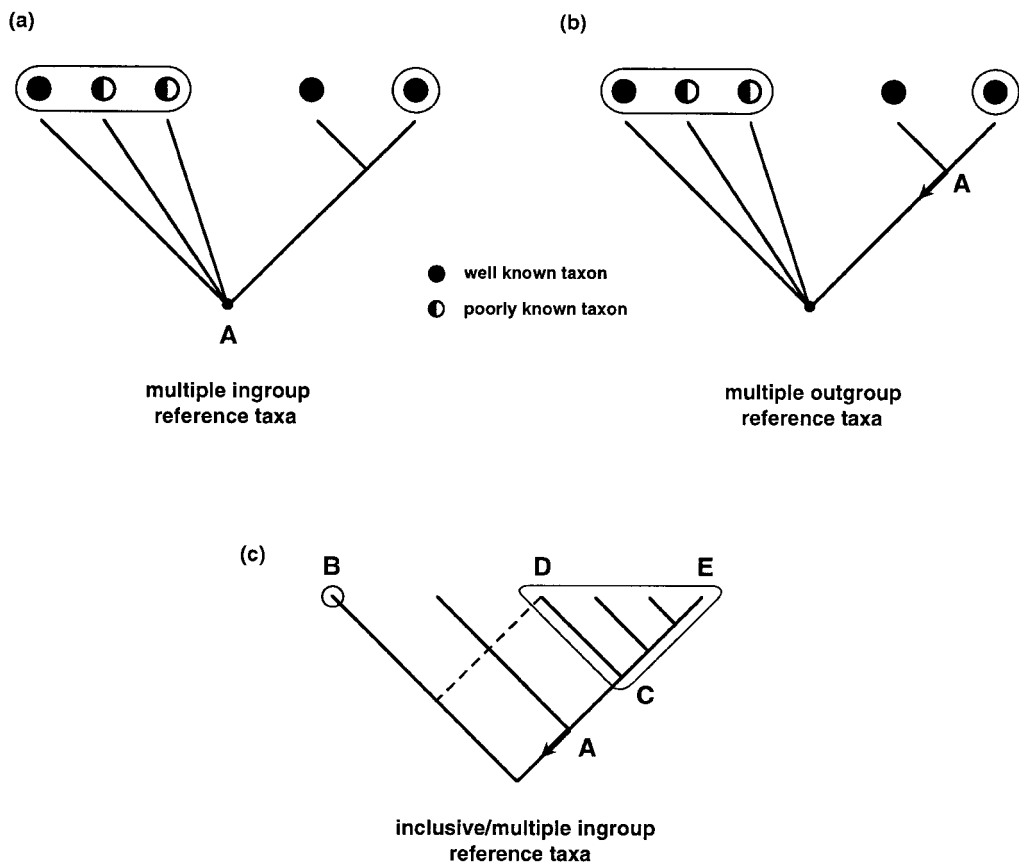
FIGURE 12. The effects of multiple reference taxa. (a) Multiple ingroup reference taxa effectively maintain the taxonomic content of node-based taxon A when basal relationships are unresolved or are unstable as a result of missing data. (b) Multiple outgroup reference taxa effectively maintain the taxonomic content of stem-based taxon A when outgroup relationships are unresolved or are unstable as a result of missing data. (c) Inclusive ingroup reference taxon C or multiple ingroup reference taxa (D and E) can result in definitional ambiguity for stem-based taxon A when basal taxon D changes its relationships; a single, nested ingroup reference taxon, such as taxon E, is preferable in stem-based definitions. Reference taxa are encircled; a dot indicates a node-based definition; an arrow indicates a stem-based definition; a dashed line indicates a change in position of a basal taxon.

labeled sister taxon, or both. Most contradictory definitions—which previously have been cited as a general weakness of stem-based definitions (Schander and Thollesson, 1995)—are the result of choosing maximally inclusive reference taxa rather than more-nested relatives. Still, multiple ingroup reference taxa may best be avoided in stem-based definitions.

Node-Stem Triplet

Given taxon A and subordinate taxa B and C (a simple dichotomy), there are only three resolved positions for relocation or addition of taxa (Fig. 8). If taxa A, B, and C have stem-based definitions, the following obtains with introduction of new taxa as shown. Taxa B and C will incorporate an additional taxon with no change in their relationship with the basal dichotomy or with taxon A. When taxon A incorporates a new taxon, however, it is removed from the original basal dichotomy; the taxonomic content of taxon A can no longer be expressed as taxon B plus taxon C, but rather must include an additional taxon. If taxa A, B, and C have node-

based definitions, the following obtains with introduction of new taxa (Fig. 8). The taxonomic content of taxon A is not affected by an additional outgroup taxon, which is excluded by definition. Taxa B and C, however, are removed from the basal dichotomy; the taxonomic content of taxon A, as a result, is no longer taxa B and C and their most recent ancestor, but rather must include two additional taxa.

Only one definitional configuration, a NST, preserves the taxonomic content of three taxa about a dichotomy, as expressed by the equivalence statement A = B + C. A NST is composed of a node-based taxon and two subordinate stem-based taxa. Addition of taxa to resolved locations around a NST cannot alter the simplest expression of relative taxonomic content, A = B + C.

When applying a NST to dichotomies on a cladogram, two observations are noteworthy. First, only a small subset of existing dichotomies in the history of life are labeled with a trio of taxon names suitable for a NST, such as Dinosauria = Ornithischia + Saurischia (Fig. 13). Naming all, or even most, clades is not a heuristic endeavor that would enhance communication. Second, NSTs cannot be constructed for adjacent dichotomies on a cladogram without coining new companion node- or stem-based taxa. An existing stem-based subordinate group of one NST, such as Ornithischia, cannot simultaneously function as a node-based group of a less-inclusive NST at the base of Ornithischia. A companion node-based taxon (composed of *Pisanosaurus* and other ornithischians) would need to be defined for this less-inclusive NST (Fig. 13). Coining numerous companion taxa that differ only in definitional type but not in current taxonomic content will not enhance communication. A complete NST network would require node- and stem-based taxa at
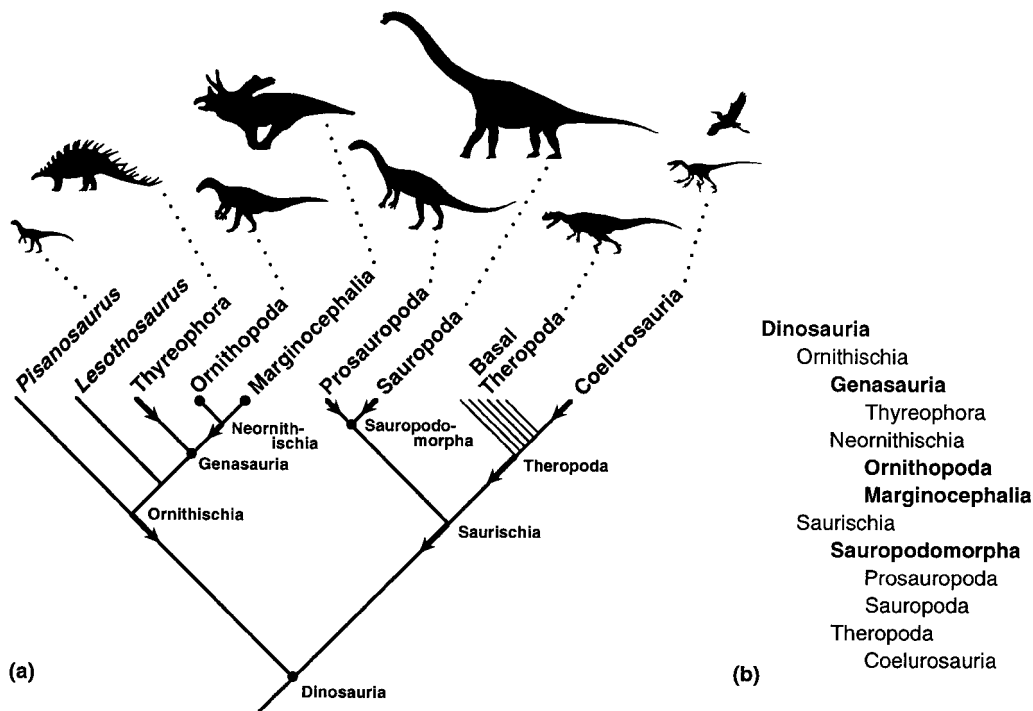

FIGURE 13. Recognition of NSTs in the higher-level taxonomy of Dinosauria (from Sereno, 1998). Stem-based taxa are represented by an arrow (a) or regular typeface (b). Node-based taxa are represented by a dot (a) or boldface (b).

each branchpoint. A more conservative rationale for placement of NSTs at named dichotomies is considered below.

Taxonomic Framework

Taxon names are applied to organismal diversity as tags for identification and information storage and retrieval (Farris, 1979). The following kinds of dichotomies have been "tagged" with a trio of names and used most often by systematists: (1) a higher-level crown clade with subordinate crown or total clades; (2) a clade composed of two diverse subclades; (3) a morphologically distinct clade composed of two distinctive subclades; and (4) a dichotomy with significant historical usage. Placement of NSTs is based on four criteria that mirror these traditional labeled dichotomies.

Survivorship.—Crown, stem, and total taxa are defined on the basis of survivorship (Table 1), which can provide the basis for NSTs (Fig. 14). The strength of this criterion lies in its unambiguous interpretation.
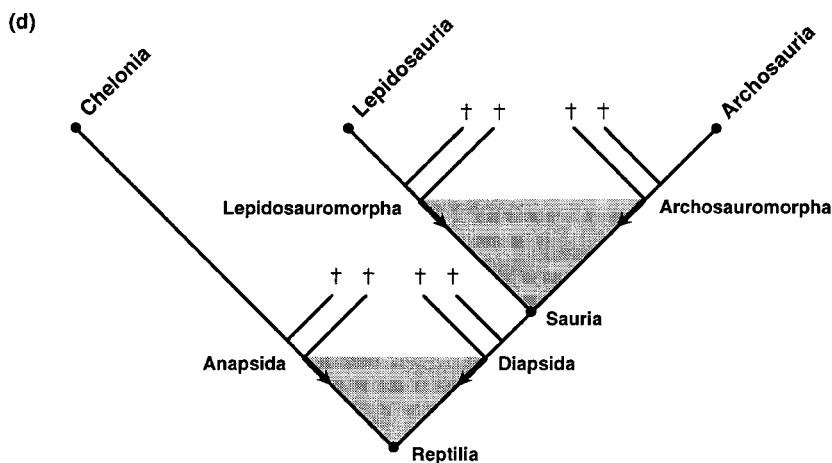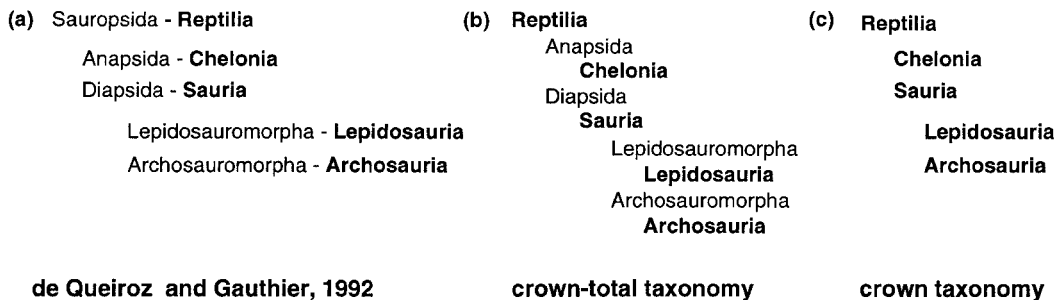
FIGURE 14. Survivorship as the basis for supraspecific taxonomy. (a) Indented crown-total taxonomy emphasizing the association between a total taxon and its less inclusive crown taxon (de Queiroz and Gauthier, 1992). (b) Indented crown-total taxonomy emphasizing NST structure. (c) Indented crown taxonomy omitting total taxa. (d) NST (shaded) form the framework of a crown-total taxonomy. Regular typeface (a–c) or an arrow (d) indicates a stem-based taxon; boldface (a–c) or a dot (d) indicates a node-based taxon; daggers indicate extinct taxa.

It is highly unlikely that a species believed to be extinct will be discovered alive and vice versa. The rapid extinction of the modern biota is a notable exception (Lucas, 1992); selection of a particular historical datum (such as the beginning of the Recent) may at least partially side-step this problem. NSTs based on higher-level crown taxa and their associated subordinate total taxa form the backbone of the tetrapod classification proposed by de Queiroz and Gauthier (1992). Because these authors emphasized the relationship between a crown clade and its more inclusive total clade (as shown by their indented classification; see Fig. 14a), the NST structure of their classification was not discussed and is not immediately apparent. Their classificatory scheme, nevertheless, consists of subordinate NSTs that can be tabulated as an indented crown-total taxonomy (Fig. 14b, d). One can list only crown groups (a crown classification; Fig. 14c), although the purpose of such a classification is not immediately apparent, given the inclusion of many extinct species within crown taxa.

Diversity.—Relative diversity is an important consideration in the erection and positioning of higher taxon names. Some groups are much more diverse than others, and this difference is often accorded biological significance. On a cladogram, the simplest diversity comparisons are between sister taxa on either side of a dichotomy (Fig. 15a). A diversity-based NST stabilizes the association of taxon names around a dichotomy that involves diverse sister groups and several low-diversity outgroups. A diversity-based NST ties taxa to diversity, despite relocation or addition of taxa near the basal dichotomy. There is only one way in which a diversity-based NST can be altered to dissociate the named trio of taxa and high diversity: relocation of most of the diversity of one of the sister groups but not the included reference taxon. Alternatively, one could view this dissociation as the relocation of a reference taxon but not of the majority of its previously associated diversity. Such phylogenetic rearrangements are unlikely (with nested reference taxa), and diversity remains one of the least ambiguous and most stable criteria for establishing NSTs.

A diversity-based NST recognizing Dinosauria and its two diverse subclades, Ornithischia and Saurischia, ties these three taxa to a major dichotomy in the evolution of archosaurs (Fig. 13a). Traditional systematists have maintained this dichotomy by interpreting newly discovered basal dinosaurs as either basal ornithischians or basal saurischians and by excluding outgroups from Dinosauria. The only way to dissociate diversity and a diversity-based NST positioned at Dinosauria is to relocate the deeply nested reference taxa *Triceratops* and Neornithes outside of Ornithischia and Saurischia, respectively (Sereno, 1998).

Morphology.—Some taxa are more distinctive morphologically than others. Phylogenetic analysis of morphological data rarely distributes apomorphies evenly across nodes, and those nodes characterized by profound transformations are often the subject of biological discourse. A morphology-based NST recognizes a dichotomy bounded by numerous apomorphies (Fig. 15b). The number of apomorphies (or the size of the morphologic "gap"), however, is less stable than diversity as a criterion for distinguishing some branchpoints from others. Discovery of new taxa inevitably entails the splitting of morphologic "gaps," and alternative optimization of character data with many missing entries may also dramatically shift the locus of character-state change.

Some morphologic "gaps," nonetheless, have remained remarkably stable, particularly those among extant species associated with a poor, or nonexistent, fossil record. Gap-bounded dichotomies also occur among extinct taxa. The distinctive clade Sauropodomorpha, for example, is composed of the distinctive subclades Prosauropoda and Sauropoda. Despite more than a century of discovery, morphologically intermediate species have yet to be uncovered that significantly reduce the distances at this dichotomy. A morphology-based NST at this dichotomy reflects historical experience—that newly discovered sauropodomorphs have been positioned among known prosauropod or sauropod taxa, whether this is the result of punctu-
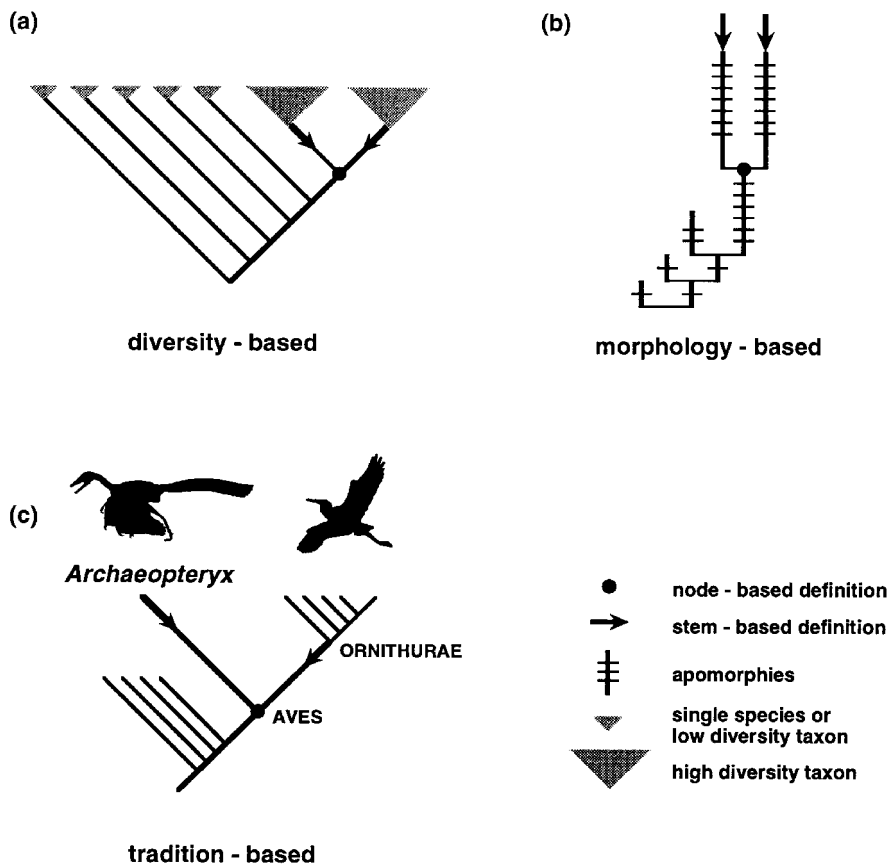
FIGURE 15.   Criteria for positioning NST. (a) Diversity-based NST, in which two high-diversity taxa have successive low-diversity outgroups. (b) Morphology-based NST, in which a distinctive (particularly apomorphic) taxon is composed of two distinctive subordinate taxa. (c) Tradition-based NST, in which a taxon and its subordinate taxa have a long historical association.

ated evolution, temporal gaps in the fossil record, or chance.

Tradition.—Some taxonomic names have a long-standing association with a particular dichotomy. Most often this association is based on a "key" character or set of characters and involves a taxon that is transitional in form (i.e., a "missing link" bounded by morphologic "gaps"). For the past century, for example, "Aves" has been associated with the dichotomy between *Archaeopteryx* and Ornithurae, a taxon comprising all other birds (Fig. 15c). *Archaeopteryx* preserves evidence of feathers, wings, a perching foot, and other striking adaptations reminiscent of those of living birds. Surely, however,

a "feather" or "wing" will be scored and interpreted in different ways by different systematists. Likewise, the long-standing association of several of these "key" synapomorphies with *Archaeopteryx* and Ornithurae clearly is oversimplified (Chen et al., 1998; Ji et al., 1998). Still, the name Aves may be better applied to its traditional dichotomy (Fig. 15c; Lee, 1996; Sereno, 1998) than relocated to a distant node, such as crowngroup avians (Gauthier, 1986), even if only to avoid confusion on the part of the majority of biologists. A tradition-based NST placed at the dichotomy between *Archaeopteryx* and Ornithurae stabilizes the long-standing historical interpretation of Aves and maintains

a close, albeit imprecise, association between Aves and synapomorphies involving modern feather, wing, and pedal design.

The four criteria discussed above are neither exact nor mutually exclusive. The number of taxa or apomorphies characterizing "high diversity" or "morphologically distinct" clades, respectively, is relative and imprecise. Of the four criteria, morphology-based dichotomies are least useful, because they are rare and relatively unstable over time and can exist solely as artifacts of missing data.

The best case for location of a NST is made when several of the criteria described above point to the same dichotomy, as is the case with the node Dinosauria. Within dinosaurs, the ceratopsid subgroups Ceratopsinae and Centrosaurinae provide another example, each more diverse and distinctive than their nearest outgroups (Fig. 16). When criteria for placement of a NST disagree, or when the proximity of another NST would necessitate coining new companion taxa, a NST need not be established (e.g., the dichotomy involving Saurischia, Sauropodomorpha, and Theropoda; Fig. 13a).

Subdivision of one group into two subordinate groups is the simplest form of nonredundant ordination in taxonomy and reflects the most prevalent phylogenetic pattern. The principal reasons that these dichotomies were named in the first place was to acknowledge survivorship to the Recent, to reference large-scale patterns in diversity, and to record major changes in morphology. Maintaining the integrity of such named dichotomies is a heuristic endeavor in concert with the aims of phylogenetic taxonomy and can be achieved by judicious choice of nested reference taxa and the NST configuration of phylogenetic definitions.

CONCLUSIONS

The foregoing critique of taxonomic definitions has resulted in the following major conclusions:

1. Currently there is no clear rationale for formulation and placement of phylogenetic definitions.
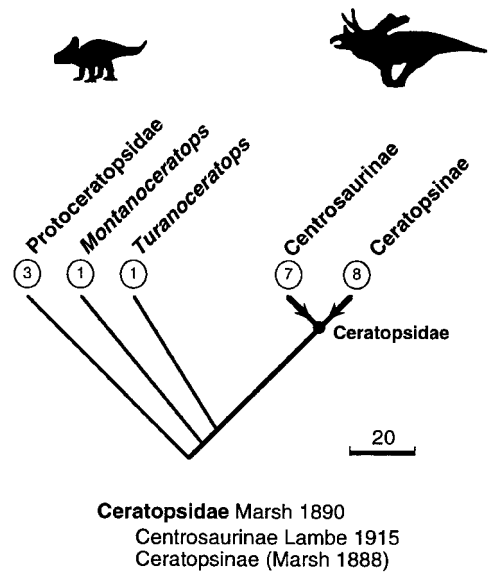2. Apomorphy-based definitions, unlike

FIGURE 16. NST among ceratopsian dinosaurs based on the criteria outlined in Figure 15. Number of genera for each terminal taxon are shown within circles, which highlights the diversity of ceratopsid subclades as compared with the outgroups. The thickened clade lines are scaled according to the number of synapomorphies that diagnose each clade, which underscores the morphologic disparity of Ceratopsidae and its two subgroups compared with adjacent, more inclusive nodes (scale bar equals 20 apomorphies; Sereno, 1998). The indented classification reveals a century-long tradition that recognizes Ceratopsidae and its two subgroups. Node-based taxa are represented by a dot (or boldface), and stem-based taxa are represented by an arrow (or regular typeface).

node-based and stem-based definitions, manifest many of the same ambiguities that have long been associated with traditional character-based taxa.

3. Crown taxa are defined with living reference taxa, but they require an immediate outgroup that is extinct. Crown taxa are not demonstrably more stable or informative than stem or total taxa, and "widely used" higher taxon names have not been historically associated with crown taxa.
4. Stem taxa are restricted to extinct clades (or species) whose immediate outgroup includes at least one living species. De-

fined in this way, a total taxon is composed of a crown taxon plus all stem taxa more closely related to it than to another crown taxon.

5. Well-known, nested reference taxa, rather than maximally inclusive reference taxa, stabilize taxonomic content. Use of multiple reference taxa, in addition, does not open phylogenetic definitions to internal conflict, except as ingroup reference taxa in a stem-based definition.

6. A NST configuration of complementary taxonomic definitions stabilizes the association of a trio of taxa (A, B, C) at a dichotomous branchpoint, such that the taxonomic content of taxon A will always equal that of taxa B and C plus their most recent common ancestor.

7. Survivorship, diversity, morphology, and tradition are heuristic criteria for placement of node-stem triplets.

REFERENCES

AX, P. 1987. The phylogenetic system. The systematization of organisms on the basis of their phylogenesis. Wiley and Sons, New York.

BROCHU, C. A. 1997. A review of "*Leidyosuchus*" (Crocodyliformes, Eusuchia) from the Cretaceous through Eocene of North America. J. Vertebr. Paleontol. 17:679–697.

BRYANT, H. N. 1994. Comments on the phylogenetic definition of taxon names and conventions regarding the naming of crown clades. Syst. Biol. 43:124–130.

BRYANT, H. N. 1996. Explicitness, stability, and universality in the phylogenetic definition and usage of taxon names: A case study of the phylogenetic taxonomy of the Carnivora (Mammalia). Syst. Biol. 45:174–189.

BUCK, R., AND D. L. HULL. 1966. The logical structure of the Linnaean hierarchy. Syst. Zool. 15:97–111.

CHEN, P., Z. DONG, AND S. ZHEN. 1998. An exceptionally well-preserved theropod dinosaur from the Yixian Formation of China. Nature 391:147–152.

CHIAPPE, L. M. 1991. Cretaceous avian remains from Patagonia shed new light on the early radiation of birds. Alcheringa 15:333–338.

CHIAPPE, L. M. 1996. Late Cretaceous birds of southern South America: Anatomy and systematics of Enantiornithes and *Patagopteryx deferrariisi*. Münch. Geowiss. Abh. 30:203–244.

DAESCHLER, E. B., AND N. SHUBIN. 1998. Fish with fingers? Nature 391:133.

DARWIN, C. 1859. On the origins of species by natural selection. John Murray, London.

DE QUEIROZ, K. 1992. Phylogenetic definitions and taxonomic philosophy. Biol. Philos. 7:295–313.

DE QUEIROZ, K. 1994. Replacement of an essentialistic perspective on taxonomic definitions as exemplified by the definition of "Mammalia." Syst. Biol. 43:497–510.

DE QUEIROZ, K., AND J. A. GAUTHIER. 1990. Phylogeny as a central principle in taxonomy: Phylogenetic definitions of taxon names. Syst. Zool. 39:307–322.

DE QUEIROZ, K., AND J. A. GAUTHIER. 1992. Phylogenetic taxonomy. Annu. Rev. Ecol. Syst. 23:449–480.

FARRIS, J. S. 1976. Phylogenetic classification of fossils with recent species. Syst. Zool. 25:271–282.

FARRIS, J. S. 1979. The information content of the phylogenetic system. System. Zool. 28:483–519.

FOREY, P. C. 1992. Fossils and cladistic analysis. Pages 124–136 *in* Cladistics: A practical course in systematics (P. L. Forey, C. J. Humphries, I. C. Kitching, R. W. Scotland, D. J. Siebert, and R. M. Williams, eds.) Oxford University Press, Oxford.

FORSTER, C. A., S. D. SAMPSON, L. M. CHIAPPE, AND D. W. KRAUSE. 1998. The theropod ancestry of birds: New evidence from the Late Cretaceous of Madagascar. Science 279:1915–1919.

FÜRBRINGER, M. 1888. Untersuchungen zur Morphologie und Systematik der Vögel. Holkema, Amsterdam.

FUTUYMA, D. J. 1986. Evolutionary biology. Sinauer, Sunderland, Massachusetts.

GADOW, H. 1893. Vögel. II Systematischer Theil. Dr. J. G. Bronn's Klassen und Ordnungen des Their-Reichs. C. F. Winter'sche Verlagshandlung, Leipzig.

GAUTHIER, J. A. 1986. Saurischian monophyly and the origin of birds. Mem. Calif. Acad. Sci. 8:1–55.

GHISELIN, M. T. 1984. "Definition," "character," and other equivocal terms. Syst. Zool. 33:104–110.

GRIFFITHS, G. C. D. 1973. Some fundamental problems in biological classification. Syst. Zool. 22:338–343.

GRIFFITHS, G. C. D. 1974. On the foundations of biological systematics. Acta Biotheor. 23:85–131.

HAECKEL, E. 1866. Generelle Morphologie der Organismen. Allgemeine Grundzuge der organischen Formen Wissenschaft, mechanisch begrundet durch die von Charles Darwin reformierte Deszendenz-Theorie. Georg Reimer, Berlin.

HENNIG, W. 1969. Die Stammesgeschichte der Insekten. Kramer, Frankfurt.

HENNIG, W. 1983. Stammesgeschichte der Chordaten. Fortschr. Zool. Syst. Evolutionsforsch. 2:1–208.

JEFFERIES, R. P. S. 1979. The origin of chordates—a methodological essay. Pages 443–477 *in* The origin of major invertebrate groups (M. R. House, ed.). Academic Press, London.

JI, Q., P. J. CURRIE, M. A. NORELL, AND S.-A. JI. Two feathered dinosaurs from northeastern China. Nature 393:753–761.

JOLLIE, M. T. 1973. Chordate morphology. Robert Krieger, Huntington, New York.

LEE, M. S. Y. 1996. Stability in meaning and content of taxon names: An evaluation of crown-clade definitions. Proc. R. Soc. London B 263:1103–1109.

LEE, M. S. Y. 1998. Ancestors and taxonomy. Trends Ecol. Evol. 13:26.

LUCAS, S. G. 1992. Extinction and definition of the Class Mammalia. Syst. Biol. 41:370–371.

MARSH, O. C. 1872a. Discovery of a remarkable fossil bird. Am. J. Sci. 3:56–57.

MARSH, O. C. 1872b. Notice of a new remarkable fossil bird. Am. J. Sci. 4:344.

MCKENNA, M. C., AND S. K. BELL. 1997. Classification of mammals above the species level. Columbia University Press, New York.

MENG, J., A. R. WYSS, M. R. DAWSON, AND R.-J. ZHAL. 1994. Primitive fossil rodent from Inner Mongolia and its implications for mammalian phylogeny. Nature 370:134–136.

PATTERSON, C. 1981. Significance of fossils in determining evolutionary relationships. Annu. Rev. Ecol. Syst. 12:195–223.

PENNY, D., AND M. HASEGAWA. 1997. The platypus put in its place. Nature 387:549–550.

ROWE, T. 1987. Definition and diagnosis in the phylogenetic system. Syst. Zool. 36:208–211.

ROWE, T. 1988. Definition, diagnosis and origin of Mammalia. J. Vertebr. Paleontol. 8:241–264.

ROWE, T., AND J. A. GAUTHIER. 1992. Ancestry, paleontology and definition of the name Mammalia. Syst. Biol. 41:372–378.

SALGADO, L., R. A. CORIA, AND J. O. CALVO. 1997. Evolution of titanosaurid sauropods. I: Phylogenetic analysis based on the postcranial evidence. Ameghiniana 34:3–32.

SCHANDER, C., AND M. THOLLESSON. 1995. Phylogenetic taxonomy—some comments. Zool. Scripta 24:263–268.

SERENO, P. C. 1997. The origin and evolution of dinosaurs. Annu. Rev. Earth Planet. Sci. 1997:435–489.

SERENO, P. C. 1998. A rationale for phylogenetic definitions, with application to the higher-level taxonomy of Dinosauria. Neues Jahrb. Geol. Paläont. Abh. 210:41–83.

SMITH, A. B. 1994. Systematics and the fossil record: Documenting evolutionary patterns. Blackwell Scientific, Oxford, England.

WALKER, E. P. 1975. Mammals of the world, 3rd edition. Johns Hopkins Univ. Press, Baltimore.

WILEY, E. O. 1981. Phylogenetics: The theory and practice of phylogenetic systematics. Wiley and Sons, New York.

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

WILSON, D. E., AND D. M. REEDER. 1993. Mammal species of the world, 2nd ed. Smithsonian Institute Press, Washington, D. C.

WYSS, A., AND J. MENG. 1996. Application of phylogenetic taxonomy to poorly resolved crown clades: A stem-modified node-based definition of Rodentia. Syst. Biol. 45:559–568.