Nomenclature and Classification, Principles of

F. Christian Thompson
Systematic Entomology Laboratory, U.S. Department of Agriculture

Classification has two meanings in English: the process by which things are grouped into classes by shared characters and the arrangement of those classes. Identification is the process of observing characters and thereby classifying things. Biological classifications are arrangements of organisms. The ability to classify is common to all animals, for to survive animals must group other organisms into at least three classes: Those to be eaten, those to be avoided, and those to associate with, especially members of their own class. For scientists, classification is formalized into a nested or hierarchical set of hypotheses: hypotheses of characters, groups (taxa), and relationships among the groups. Individual specimens of organisms are observed and characteristics noted. So, for example, we may observe that some people are black, others yellow or white, and conclude as Linnaeus did that there are different groups of humans *(Homo sapiens)*. This is a hypothesis that skin color is a useful character. Further testing of this character hypothesis has shown that skin color among humans does not delimit natural groups; hence we reject skin color as a character for humans as well as those groups this character defines. Color, however, is a very useful character for classifying many other groups. Then there is the hypothesis of a group. Groups of biological organisms are called taxa (taxon, singular) and these taxa are hierarchically arranged in our classifications. Taxa in a classification have rank, with the basic (basal, bottom) rank being designated as species. Some of the higher ranks are genus, tribe, family, order, class, phylum, and kingdom.

Nomenclature is a system of names along with the procedures for creating and maintaining that system. Classification, in its second definition, is the structure for nomenclature, being the model on which names are arranged. Names form the essential language of biology and are the means that we use to communicate about our science. To avoid a Tower of Babel, a common system of nomenclature is required, especially an effective, efficient, system that has a minimal cost.

NAMES AND CLASSIFICATION

Names are tags. Tags are words, short sequences of symbols (letters) used in place of something complex, which would require many more words to describe. Hence, tags save time and space. Instead of a long description, we use a short tag. A scientific name differs from a common name in that the scientific name is a unique tag. In nonscientific languages, such as English, there may be multiple tags (common names) for the same organism. For example, imagine the various words in English that are used to describe *H. sapiens.* In computer (database) jargon, data elements that are used to index information are termed keys, and keys that are unique are called primary keys. Scientific names are primary keys. The word "key" has another meaning in English, which is "something that unlocks something." Scientific names are those critical keys that unlock biosystematic information, which is all that we know about living organisms. To repeat: scientific names are tags that replace descriptions of objects or, more precisely, concepts based on objects (specimens). Scientific names are unique within a classification, there being only one valid scientific name for a particular concept, and each concept has only one valid scientific name.

Scientific names are more than just primary keys to information. They represent hypotheses. To most systematists, this is a trivial characteristic that is usually forgotten and thereby becomes a source of confusion later. To most users, this is an unknown characteristic that prevents them from obtaining the full value from scientific names. If a scientific name were only a unique key used for storing and retrieving information, it would be just like a social security number. *H. sapiens* is another unique key used to store and retrieve information about humans, but that key also places that information into a hierarchical classification. Hierarchical classifications allow for the storage at each node of the hierarchy of the information common to the subordinate nodes. Hence, redundant data, which would be spread throughout a nonhierarchical system, are eliminated. Biological classifications, however, do more than just hierarchically store information. If one accepts a single common (unique) history for life (phylogeny) and agrees that our biological classifications reflect this common history in their hierarchical arrangement, then biological classifications allow for prediction, namely that some information stored at a lower hierarchical node may belong to a higher node; that is, may be common to all members of the more inclusive group. Such predictions take the following form: if some members of a group share a characteristic that is unknown for other members of the same group, then that characteristic is likely to be common to all members of the group. So scientific names are tags, unique keys, hierarchical nodes, and phylogenetic hypotheses. Thus systematists pack a lot of information into their names and users can get a lot from them.

Scientific names are hypotheses, not proven facts. Systematists may and frequently do disagree about hypotheses. Hypotheses, which in systematics range from what is a character to what is the classification that best reflects the history of life, are always prone to falsification, hence to change. Disagreements about classification can arise from differences in paradigm and/or information. Systematists use different approaches to construct classifications, such as cladistic versus phylogenetic versus phenetic methods. Given the same set of data that underlies a given hierarchy, cladists will derive classifications different from those derived by pheneticists (Fig. 1). Even among cladists, there can be differences about the rank (genus, family, order) and thereby the hierarchical groups used. These are disagreements based on paradigm. There can be disagreement about the hypotheses that underlie the information used to construct the classifications, such as what are the characters. And disagreement can arise among systematists because individuals use different information. While disagreements will affect the ability to predict, they need not affect the ability to retrieve information.

The desirable attribute that must be preserved to ensure complete access to information across multiple classifications is uniqueness. Our scientific nomenclature must guarantee that any scientific name that is used in any classification is unique among all classifications. This can be assured by having two primary keys. Unfortunately, having two keys increases the overhead of our information systems. So most systematists and *all* users want to avoid this problem by mandating that there be only *one* classification. Although in theory there is only one correct classification, as there was only one history of life, in reality there have been multiple classifications in the past, there may be multiple classifications in use today, and there will be multiple classifications in the future. That is the price of scientific progress, of the increase in our knowledge of the world. If information is to be retrieved across time—that is, if we want to extract information stored under obsolete classifications, and if we want to avoid dictating "the correct" classification—then we need a nomenclatural system that supports two unique keys.

The two keys for our language of biodiversity are the valid name and the original name. The valid name is the correct name for a concept (taxon) within a classification; the original name is the valid name in the classification in which it was

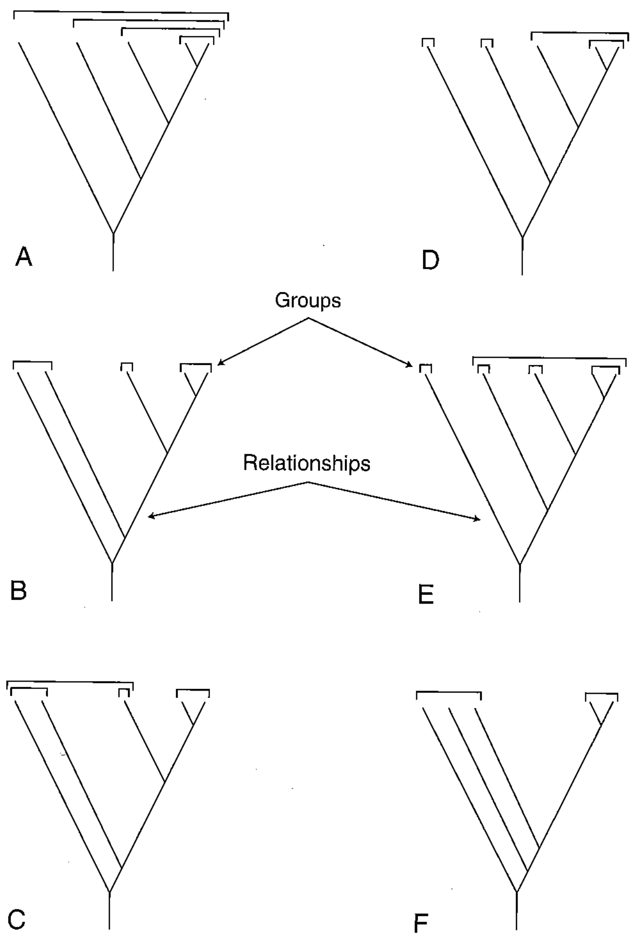

FIGURE 1 Multiple classifications for identical cladistic hypotheses. The brackets along the top of the trees indicate the groups recognized. Cladistic classifications are shown for trees A, D, and E and phenetic classifications for B, C, and D.

proposed. Valid names may be different among classifications, but the original name is invariant across all classifications (Table I). Valid names are the best names to use because they provide the full value of scientific names. These are the names that provide a basis for prediction. The original name is useful only for information retrieval across multiple classifications. Although valid and original names may be and frequently are the same, users must know the differences between them. Specifically, they need to know that a valid name is a powerful inference tool, that a valid name provides for prediction about unknown attributes of the organism that bears the name. But they must understand that there may be multiple valid names in the literature and/or in use and that valid names represent hypotheses that may change as our knowledge is tested and improved. So most importantly, if there are multiple valid names in use, then there are conflicting scientific hypotheses being advocated, and users must select the name that best serves their purpose. If users do not want to decide, do not want to use classifications to organize and synthesize their information, then they may use the original name to index their information, being assured that it will always be a unique key.

TABLE I Multiple Classifications (rows) and Primary Keys (columns) to Information

Year	Valid name	Original name	Authority
1776	*Musca balteata*	*Musca balteata*	De Geer
1822	*Syrphus balteatus*	*Musca balteata*	Meigen
1843	*Scaeva balteata*	*Musca balteata*	Zetterstedt
1917	*Episyrphus balteatus*	*Musca balteata*	Matsumura
1930	*Epistrophe balteata*	*Musca balteata*	Sack
1950	*Stenosyrphus balteatus*	*Musca balteata*	Fluke
Today	*Episyrphus balteatus*	*Musca balteata*	Vockeroth

There are other problems today with our classifications: synonymy, having two names for the same concept, and homonymy, having the same name for different concepts. These problems are, however, largely the result of ignorance. If we knew all names and their types and could agree on what are species, then by applying the rules of nomenclature we could immediately eliminate all synonymy and homonymy problems. Homonymy is eliminated by the rule of uniqueness. Synonymy is addressed by the rules of typification, which tie a physical instance of a concept to a name, and is resolved by logic of circumscription and the convention of priority (or usage). The name of a concept is the name affixed to one and only one of the types that falls within its circumscription (Fig. 1). The name used is determined by which name is the oldest (priority) or most widely used (usage). The specific rules for resolving homonymy and synonymy, as well as for the proper formation and documentation of names are our codes of nomenclature. These rules, however, do not address the problem of multiple classifications, nor can they establish order under conditions of ignorance of the universe of applicable names and their typification.

There is one final problem: the species problem. This is the problem of what is the basic unit of information and/or data. There is also the question of what species are and whether species are real or hypotheses. Species may be a category (rank) in classifications or a unit of information. The best current review on these questions is by Wheeler and Meier, but for nomenclature the species (or more precisely the species group, which includes the subspecies category) is considered to be a basic unit of information. The problem is that the species is not a data element. The species is not an indivisible unit, but consists of information, that is, data derived from specimens that have been identified as belonging to that species. Mistakes can be made during this identification, which is after all another hypothesis. Information is ultimately not derived from species, but from specimens. Biological information management really begins with specimen data management. The problems of specimen-based data management are not intractable but are readily addressed by the use of unique identifiers, such as bar codes, another form of unique keys.

The species problem is also one of circumscription, the definition of the limits of a taxon. A group with the same name and type may be more or less inclusive depending on the characters used to define its limits. Zoologists differ from botanists in not considering circumscription to be a problem, since minimally all identically named taxa have at least some characteristics in common. The problem of how much is held in common, therefore, is best resolved by enumeration of the included taxa or specimens. The history of circumscription can be tracked by use of an additional key that uniquely identifies the person who defined the limits and the date of that action. Sufficient for our purposes is to know that specimen-based data will always be summarized into species-based information units and that all species-based information should be specimen based.

PARADIGMS AND CLASSIFICATIONS

The information that is embedded in nomenclature comes from the classification used. As noted, classifications consist of hierarchically nested groups of taxa, with the basic unit being the taxa ranked as species. Paradigms are theories about scientific knowledge and its organization. The first classifications developed by Linnaeus and Fabricius were largely based on Aristotelian essentialism/typology. Things were grouped together because they shared the essences of the group, which is the type. Later, when evolution was articulated as a paradigm, classifications were based on phylogeny, which is the genealogical hypothesis of relationship. More recently, when computers began to appear, classifications were proposed on the basis of statistical measures of overall or phenetic similarity. Finally, different ways of deciphering phylogeny were developed, and so, different ways of translating phylogenetic information into a hierarchical classification were proposed (phylogenetic vs cladistic methods). Over the past half-century, much has been written about the relative merits of phenetics, evolutionary systematics, and cladistics, but the inescapable conclusion for predictive and, therefore, maximally informative classifications, is that the cladistic paradigm is mandatory. Schuh provides a good summary of the arguments for cladistic classifications.

Regardless of the paradigm followed, all approaches leave unsolved the problem of how to translate the result of taxonomic analysis, be it a tree or a branching diagram of overall similarities, into a hierarchical classification. There are only two approaches to the translation of an analysis into a classification: subordination or sequencing. For subordination, each clade/branch becomes a recognized (named) taxon and a rank indicator provides a key to the relative level of subordination. Subordination works best when the phylogeny/branching diagram is balanced, that is, when each branching point divides the remaining terminal taxa into equal sized groups. For example (Fig. 2, Table II), 8 species could be clustered into 4 genera and 2 subfamilies, whereas a fully pectinate analysis would yield 7 genera clustered into 5 named ranks (subfamily, infrafamily, supertribe, tribe, subtribe). For sequencing, only the terminal clades/branches are recognized, but their order is preserved and suitably indicated to encode their sequential level of subordination. This method is highly efficient for analyses that result in pectinate trees. The pectinate example could be reduced to 7 sequenced genera. Sequencing does not work when the analysis is balanced. Given that most analyses are neither fully balanced nor fully pectinate, a mixture of subordination and sequencing should be used as long as the classification properly identifies which methodology was used for each portion. Wiley provides a full set of conventions to deal with these issues as well as others that involve the placement of fossil groups *(plesion)* or groups of uncertain or changeable position *(sedis mutabilis)* or unknown relationships *(incertae sedis).*

Beyond the translation of a taxonomical analysis into a hierarchical classification, another challenge remains, that is, what groups to formally name and what ranks to assign to

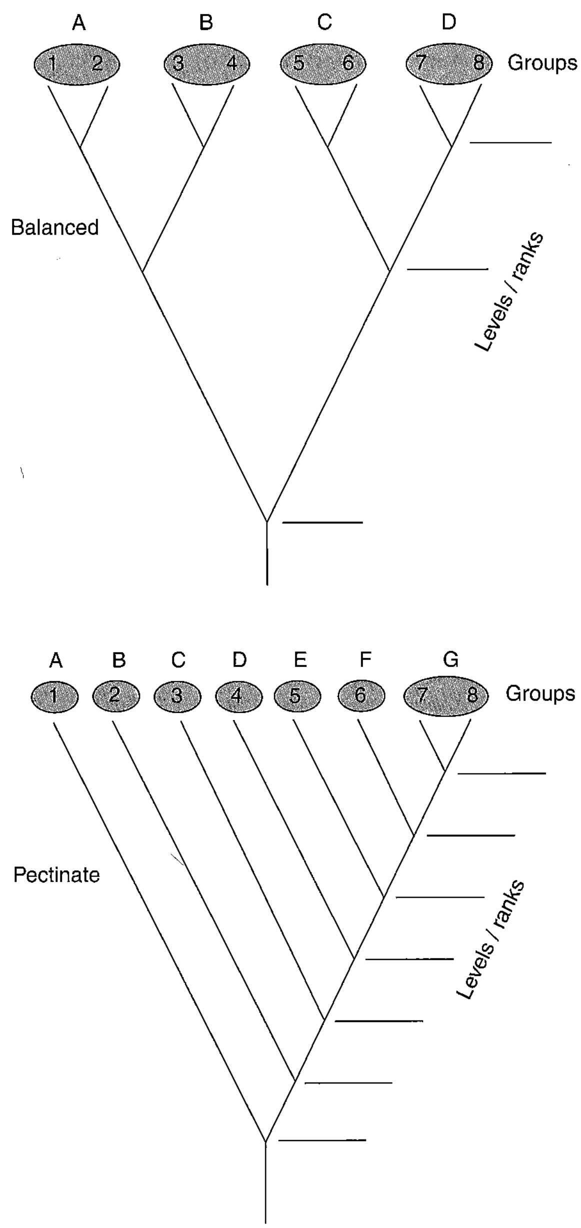

FIGURE 2 (Top) Balanced and (bottom) pectinate cladistic hypotheses. See Table II for the different classifications that result from these hypotheses.

TABLE II Classification: Sequencing and Subordination

Subordinated classification for Fig. 2A balanced analysis	Subordinated classification for Fig. 2B pectinate analysis	Sequenced classification for Fig. 2B pectinate analysis
Family A-idae	Family A-idae	Family A-idae
Subfamily A-inae	Subfamily A-inae	Genus A
Genus A	Genus A	Genus B
Genus B	Subfamily B-inae	Genus C
Subfamily C-inae	Infrafamily B-ites	Genus D
Genus C	Genus B	Genus E
Genus D	Infrafamily C-ites	Genus F
	Supertribe C-idi	Genus G
	Genus C	
	Supertribe D-idi	
	Tribe D-ini	
	Genus D	
	Tribe E-ini	
	Subtribe E-ina	
	Genus E	
	Subtribe F-ina	
	Genus F	
	Genus G	

those named groups. Obviously, when a group is fully resolved taxonomically, there could be as many named groups as there are terminal taxa. No school of taxonomy insists on naming all of them, but other than that there is no method nor any consensus among taxonomists on what taxa to name. This general problem is usually referred to by the names of the extreme views on either side, the "splitters" and the "lumpers," or those who would recognize many groups versus those who would recognize only a few. The merit of splitting is that the more taxa named, the more hierarchical information is embedded into those names and classification itself. Unfortunately that also leads to a loss of utility inasmuch as less information is summarized in each taxon.

Consider birds, the best-known group of organisms. Some 9700 species are clustered into 204 families and 2004 genera. Their scientific nomenclature is largely meaningless to many users, such as bird-watchers. For bird-watchers, common names, which more closely follow the original Linnaean classification, such as ducks *(Anas)* or hawks *(Falco)* or hummingbirds *(Trochilus)*, are more meaningful groups than the oversplit genera. On the other hand, mosquitoes, some 3500 species, are clustered into only 34 genera. The important disease vectors, such as *Anopheles* for malaria and *Aedes* for yellow fever and dengue, remain large groups where the scientific name and common name are the same and are useful to doctors, public health workers, and other entomologists. The problem of the appropriate rank for groups recognized is similar. Naturally, splitters must have a greater series of rank indicators to express their fully named hierarchies. So, although there are relatively few species of birds, they are clustered into a large number of families (204), whereas flies (order Diptera) comprise 16 times as many species clustered into fewer families (142)!

The ranking issue also brings with it the question of equivalency. Obviously a family of birds is not an equivalent unit of biodiversity or of anything else in comparison to a family of flies. Rank equivalence is an important issue because many biologists want to make comparisons across different groups of organisms. Biological comparison should never be made on the basis of taxonomical categories above the rank of species. For example, studies that base conclusions on the circumstance that one treatment or niche has more families, than another are totally meaningless because the units being compared are not equivalent. Biological comparison should be made only on the basis of cladistically defined sister-group relationships, since sister groups are of equal age.

The entomologist Willi Hennig proposed in 1966 an objective method for assigning ranks that also allowed for biological comparisons: rank should reflect the hypothesized age of origin of the taxon. His suggestion has been rejected by all on the ground that the approach would cause a major upheaval in the traditional ranks of groups. For example, humans, placed in a separate kingdom by some (Psychozoa by Huxley in 1957), would be clustered among the apes and lemurs as nothing more than a species group. For entomology, some of the larger groups, like Coleoptera, Diptera, and Lepidoptera, which go back to Aristotle, would change in rank if not content. So after more than 2000 years of using those concepts, no one wants to split up the groups or change their rank.

These issues of classifications are largely ignored by working taxonomists, most of whom focus on their specialty and do not concern themselves with global classifications. Entomologists generally do not care how birds are classified, nor do beetle workers even worry about how flies are classified. Entomologists also tend to take a pragmatic, utilitarian approach, such that conservative ranks and grouping are used among mosquitoes and other economically important insects. In summary, a few general guidelines should be followed:

1. Only monophyletic taxa should be recognized and named.
2. Subordination or sequencing should be used as is most appropriate given the analysis and always should be annotated.
3. "Empty" taxa should never be named (i.e., if a family contains only a single genus, there is no a need to name a subfamily or tribe simply because these ranks are used elsewhere in the classification).
4. The fewer taxa named, the more useful the classification generally will be to nonspecialists.
5. Traditional groups and ranks should be preserved where possible.

CODES OF NOMENCLATURE

Because names are critical for communication and information retrieval, nomenclature needs to be universal, precise, and accurate. Universality requires that the same methodology be used by all and that methodology ensure stable nomenclature over time. Precision requires that only one result be derived from an individual application of the rules of nomenclature. Accuracy requires that names be consistently and precisely tied to the hypotheses they denote. The International Code of Zoological Nomenclature (ICZN) ensures the implementation of these basic functions in our scientific names and classification. This is achieved through a series of rules organized into chapters and articles.

Stability of nomenclature should not be confused with stability of taxonomic hypotheses (taxa) and classifications. As knowledge improves and more characters are discovered and analyzed, resulting in improved understanding of relationships among organisms, taxa and classifications will change. So, as more is known about the history of life, old Aristotelian groups like reptiles will be replaced by better defined ones, and the name Reptilia will drop from our classifications. But in other well-characterized groups, such as spiders (order Araneae) or flies (order Diptera), which have proven to be natural, the names shall remain unchanged in our classification.

The current ICZN is the product of a long evolution that began with the system of binominal nomenclature introduced by Carolus Linnaeus, a Swedish professor of natural history. This system was the direct result of an earlier government biodiversity project. The Swedish crown had some far-flung possessions and wanted to know what use could be made of them. Linnaeus was sent to investigate, to survey what today is called biodiversity, and to write a report characterizing his findings with recommendations on how to use them. At the time, there was only a binary system of nomenclature: one word for the genus, with the species being described by a series of adjectives. Given the diversity Linnaeus found, he did not want to waste time repeating long strings of adjectives that were required to characterize the biodiversity. So, because the base characterizations were in his flora of Sweden, he used a combination of the genus name and single word (an epithet) for each species to form a unique key to those descriptions (Stearn gives more details).

The system was an immediate success. Linnaeus codified the system, built and maintained a universal information database for all names (his *Systema Naturae,* 10th edition in 1758), and trained a cadre of students to carry on his work. The students dispersed and converted others. But since there could be only one master, Linnaeus, they divided nature up. There was to be no more *Systema Naturae.* For entomology, the student in charge was Johann Christian Fabricius. Fabricius defined his principles in his *Philosophia entomologica* and produced a series of *Systemae* for insects, the last comprehensive one being published in 1792 to 1794.

For the next 50 or so years, there was a significant increase in the number of animals discovered, described, and named, but little concern for nomenclature, which became muddled. This led a group of British zoologists, in 1843, to propose a formal set of rules, now known as the Strickland code, from which the modern ICZN evolved. After their effort, there was another half-century of new codes being proposed for various groups of animals (birds, fossils, insects) or nationalities (English, French, German). This proliferation led to zoologists joining forces and working toward an international code for all animals, resulting in the establishment of the International Commission on Zoological Nomenclature in 1895 and *Règles Internationales de la Nomenclature Zoologique* in 1905. Although a few entomologists (e.g., Banks and Caudell in 1912) continued to work on a specialized code for insect names, these development efforts were quickly abandoned.

For the next half-century, the *Règles* and the commission operated well, but clearly improvements were needed. So after the World War II, the task of revision began. After a series of international meetings, the American entomologist J. C. Bradley, then president of the commission, wrote out a draft that in 1962 became the second edition. The next edition, in 1985, and the current one, in 1999, were largely the work of Curtis W. Sabrosky, an American entomologist, David Ride, an Australian mammalogist, and Richard Melville, a British paleontologist.

The challenge in writing codes of nomenclature is making a set of rules that demand the best practices of taxonomists today, but also preserve the names created by past workers. Hence, to accommodate the work of the past, a code makes general provisions followed by a series of exceptions qualified by dates. Also, in zoology, there are two options for preserving history. A provision can be made in the code to solve a problem, or an appeal can be made to the commission to set aside the code to preserve an old name. Changing the code affects all occurrences of a problem, but a ruling of the commission applies only to a particular occurrence. In the past the commission frequently took many years to rule on cases. Hence, for Sabrosky and others, changing provisions of the code became the preferred method of addressing problems of old names. Hence, the current ICZN has many clauses that exist only to make old names available and to preserve their customary usage. Unfortunately, in adding to the ICZN provisions of these kinds, Sabrosky and others frequently created unforeseen problems. So the ICZN must be used carefully, since for almost every positive statement there is usually an exception. Other linguistic constructions may also be confusing, such as the frequent use of the phrase "as such." For example, this phrase, used in Article 1.3.2, requires subsequent workers to decipher the intent of the original author. If the original author knowingly was describing an aberrant specimen, then the scientific name does not enter into nomenclature, but if the author thought the specimen was typical of the taxon being described, then the name must be considered to be available for use in nomenclature. Finally, the ICZN uses a number of specialized terms or special definitions for words; these are all covered in its glossary.

INTERNATIONAL CODE OF ZOOLOGICAL NOMENCLATURE

The current 4th edition of the ICZN consists of a preamble, a series of 18 chapters comprising 90 articles, recommendations and examples, and a glossary. The book also contains a preface, an introduction, and three appendices (the first two are general recommendations and the last is the Constitution of the International Commission on Zoological Nomenclature). The text is in both English and French, and there is a single combined index. This book, bounded in green, is the "official" edition published by the commission, but other official versions have been approved and published in different languages.

The preamble sets the objectives and basic principles of the code: "to promote stability and universality" of names and to ensure that the name of each taxon is unique. The preamble also declares that taxonomy is independent of nomenclature. The articles are the definitive rules, with examples of how they are applied in specific cases as well as recommendations of appropriate practices. The glossary defines each term used so that the rules can be interpreted consistently. The following is a summary of the code by major topics (in parentheses after the topic are the articles covered).

Scope (Arts. 1–3) The scope of zoological nomenclature is restricted to names for animals published starting in 1758, the date of the 10th edition of Linnaeus' *Systema Naturae.* Here, the ICZN uses the verb "deemed" to declare for nomenclature that *Systema Naturae* was published on 1 January and before Clerck's *Aranei Svecici* (spiders of Sweden), neither of which is true, since Clerck's work was actually published in 1757! Exclusions are also listed, such as hypothetical concepts. That simply means that if the Loch Ness monster is not real, then its name *Nessiteras rhombopteryx* Scott, is not a name covered by the ICZN.

Publication (Arts. 7–9), Dates (Arts. 21–22), and Authorship (Arts. 50–51) Although zoological nomenclature is a language for communication, the names that are regulated are those "published." Since taxonomy is based on some 250 years of work, the definition of publication used by the ICZN is based on printed works. For names and nomenclatural acts to be within the coverage of the ICZN, they must have been first published in a printed work in numerous copies available to the public and for the permanent scientific record. This definition excludes some printed works, such as daily newspapers, which are not published for the permanent and scientific record. The ICZN rules exclude the evolving digital world, such as the Internet. This assures that all users read the same material in determining what are the appropriate names. The ICZN does accept new digital media such as CD-ROM or DVD disks that are "stamped out," not printed. The ICZN provides rules to determine who is the author of these works and their dates of publication.

Names (Arts. 4–6) and Their Formation [Arts. 11 (11.2–11.3, 11.7–11.9), 25–49] Beyond falling within the scope of zoological nomenclature and having been published, names must be properly formed. They must be written with the Latin alphabet, and they must agree with various other requirements, many of which reflect the origin of the system at a time when all scholarly works were written in Latin. The ICZN groups scientific names into three kinds: family-group names, the names of taxa above the genus and species; genus-group names, the names for groups of species that form the first part of the binomen; and species-group names, the specific (epithet) names. The ICZN prescribes that names of higher taxa, such as superfamilies (-oidea), families (-idae), subfamilies (-inae), tribes (-ini), and subtribes (-ina), have specific suffixes; that generic and subgeneric names have gender and be nouns or be treated as such; and that specific names (epithets) be either nouns (and invariant) or adjectives (and whose ending agrees with the gender of the generic name with which it is combined).

Availability (Arts. 10, 12–20) Given that all the foregoing conditions are fulfilled, names and nomenclatural acts must meet additional requirements if they are to be made available for consideration under the ICZN and if they are to be held to be valid. The distinction between available and valid is critical. A valid name is the correct name to be used for a taxon, that is, a hypothesis of a group; an available name is any name that meets the requirements of the ICZN.

The additional requirements that names must follow to be available are as follows: (1) they must be formed as part of the system of binominal nomenclature, and (2) they must represent taxa that are considered to be valid. How taxa are defined is further regulated in a series of articles that are applied according to time. Before 1930 the standards were simple, such as attaching a previously unpublished name to an illustration, but current requirements are much more rigorous. For example, one must explicitly declare that a new name is being proposed, fully document the hypothesis (a description that "purports" to differentiate the taxon from all others), and designate the type for the name (which for a species is usually a dead specimen); if extant, the specimen must be deposited in a named, bona fide collection, or plans for such deposition must be furnished. (see Arts. 13.1.1, 16.1, and 16.4).

Validity (Arts. 23–24) A key problem for nomenclature is the existence of two or more names for the same taxon, for only one name can be considered to be correct or valid. This reflects the more general problem in science of who is to be given the credit for new ideas when multiple people claim to have arrived at them first. The general principle involved is that of priority—credit should go to the first person who published the idea. This principle was set forth by Henry Oldenburg, the first secretary of the Royal Society and editor of its *Philosophical Transactions* (1664–1677). Unfortunately, sometimes priority fails because the first to publish may have been forgotten, and someone else has become recognized as the first. In scientific nomenclature, this means that one name may become so familiar to many people that to change it to a name that is less widely known would cause instability. Hence, the ICZN provides a saving clause to allow for a widely used and familiar name to be preserved as the valid name when an older, obscure name is rediscovered. So, the ICZN provides a statement of the principle of priority (Art. 23.1), how it is to be applied in various situations, and, finally, when the principle should not be used (reversal of precedence, Art. 23.9).

Homonymy (Arts. 52–60) Another key problem of nomenclature occurs when two or more names that are the same apply to different taxa. This is known as homonymy. Because some names consisting of different Latin letters may mean the same thing, the ICZN defines "same": for example, the epithet pairs *microdon* and *mikrodon,* and *litoralis* and *littoralis* are considered to be the same (Art. 58). Then the ICZN dictates how homonymy is to be resolved: the senior name is to be retained, and the junior name is to be replaced; but there are exceptions as explained in Articles 52 to 60.

Typification (Arts. 61–76) Since names are only tags from scientific hypotheses, the question of whether two or more names are synonyms involves both nomenclature and taxonomy. Taxonomy in this regard can be defined as the circumscription of character space (Fig. 3), that is, the definition of taxa. Nomenclature is then the name, the designation of types for the name, and the rules for selecting among multiple types (see earlier remarks under Validity). Thus, a nominal taxon is only a name and a type; a taxon, the hypothesis, includes at least one nominal taxon. For species-group names, the type is a specimen (holotype, neotype, or lectotype) or a group of specimens (syntypes), which is or are the ultimate source of character information. For genus-group names, the type is a species name (nominal taxon), and for family-group names the type is a genus-group name. The determination of the type/genus of a family-group name is self-evident because that genus is basis for the family-group name itself (the type of the family Muscidae is the genus *Musca*). For both genus-group and species-group names, the types are designated either by the original author of the name or by subsequent workers.

An author may declare that a particular species/specimen is the type (original designation/holotype), may include only one species/specimen in the taxon (monotypy/holotypy) or may for genus-group names include a species-group name that is the same as the genus-group name (absolute or Linnaean tautonymy). If the type is not clearly fixed in the original publication, the ICZN provides rules for determining what species/specimens may be designated the type by subsequent workers.

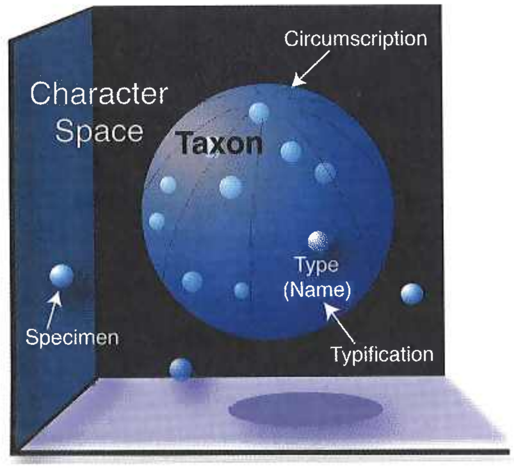

FIGURE 3 Circumscription and typification. Axes represent characters; small spheres represent specimens plotted against those axes; the large sphere represents the circumscription of a taxon; one specimen sphere is a type and its location determines the name for the taxon.

For genus-group names, all species included in a newly defined genus are eligible to be designated as the type species. But this applies only to genus-group names proposed before 1931, for as noted earlier, after 1930 type designation became a requirement of availability. If no species were originally included, then those first subsequently included are to be considered. To subsequently designate a type species, a worker merely, but unambiguously, declares one of these originally included species as the type.

For species-group names, when no type is designated in the original publication, all specimens upon which the author based the name (including specimens not seen by the author but referred to by bibliographic citation) are eligible to be designated lectotype and are called syntypes until such a lectotype is designated.

Collectively all specimens that are either holo-, lecto-, neo-, or syntypes are termed primary types. Other specimens studied by the author may be termed paratypes (or one may be an allotype if of a sex different from the holotype), but these secondary types have no nomenclatural significance. When all primary types are no longer extant (lost or destroyed), a subsequent worker may designate any specimen as the neotype to objectively define the nominal species-group taxon. Naturally, there are recommendations and restrictions about which specimen may make a more appropriate neotype. Provisions are also made for types that have been misidentified; that is, the characters used by an author to define a taxon do not agree with those of the nominal type. When this happens, workers are free to select as type either one that agrees with the characters or the named type.

Exceptions (Arts. 78, 80–83) and Registration (Art. 79) The stated objectives of the ICZN are to promote stability and universality in the names of animals, and thus all its provisions must be subservient to these goals. Hence, the ICZN provides means by which any provision of the ICZN (except those that deal with its authority and exception handling) can be set aside in a particular situation. These articles outline how one appeals to the commission and how the commission arrives at its opinion, which is then published in its *Bulletin of Zoological Nomenclature.* These plenary powers and specific powers are restricted to specific cases, usually involving only a few names or works. The ICZN has provided a procedure to rule on a whole set of names at once and has created a *List of Available Names in Zoology.* This specific power allows international groups of zoologists to propose a set of names that may be approved by the commission, thereby fixing all the relevant details about those names (their spelling, authorship, place and date of publication, and typification) and giving those names precedence over other names. Names not on the list are thereafter excluded from zoological nomenclature.

Authority (Arts. 77, 84–90) The last section of the ICZN includes a series of rules explaining the derivation and perpetuation of its authority, as well as the various regulations governing the particular edition of the ICZN. These state that the ICZN is prepared by the International Commission on Zoological Nomenclature with the participation of the zoological community under the authority of a single organization (originally the International Congresses of Zoology, now the International Union of Biological Sciences). For the future, authority can be delegated to other international organizations as specified. The effective date of the fourth edition is given as January 1, 2000, and all previous editions no longer have any force.

OTHER CODES

Currently there are five different codes of nomenclature in use for organisms-one each for plants, cultivated plants, bacteria, viruses, and animals. All these codes address the same problem, the need for universal, stable, and precise sets of names for organisms, and all are similar in their methodology. However, differences are significant and can cause difficulties for those developing databases that cover all life. Hence, in the early 1990s, an effort was undertaken to develop a single code of nomenclature for all life. Meetings were held and a draft BioCode was published, but nothing further has happened.

All the codes in use today are based on and have their origin in the binominal system established by Linnaeus. Although the Linnaean system has evolved from its topological roots into one adapted to the Darwinian evolutionary model, some believe that the system cannot fully express human knowledge about the cladistic relationships among organisms. A PhyloCode has been proposed to address these perceived failures. Unfortunately, the PhyloCode adds more uncertainty, since names for clades can be based on three different methods for defining groups, and clade names have no rank, which means that virtually all information content is lost (as discussed by Benton, Forey, and Platnick).

INTERNATIONAL COMMISSION ON ZOOLOGICAL NOMENCLATURE

Zoologists realized early on that no code of nomenclature could be perfect, always able to resolve all situations in a manner that promotes stability and universality in scientific names. Hence, zoologists established an international group of specialists with the authority not only to develop and maintain the ICZN, but to make rulings on specific names and situations. What the commission is empowered to do has been outlined in this article and is covered in Articles 78 to 81. How the commission operates is set in Articles 77 and 83 and in its constitution, which is published as Appendix 3 of the ICZN. To appeal to the commission, a worker prepares a proposal and submits it to the commission. The proposal is then published in the *Bulletin of Zoological Nomenclature* for public comment. After 6 months, the commission may vote on the proposal, and the ruling will later be published in the *Bulletin.* Each proposal is treated separately on its own merits. The commission does not set precedents or follow case law.

CONCLUSIONS

No modern science places as much emphasis on priority as taxonomy and nomenclature. This emphasis requires specialists to be familiar with at least a century of published work and sometimes 300 years' worth. Some may question the value of such a long view, in as much as most sciences look back only a decade or so. Beyond the moral and ethical considerations, however, much can be gained by understanding the past. Recognition of taxa is an innate ability of humans. Ernst Mayr once noted that the primitive natives of New Guinea knew and had names for 137 of the 138 local species of birds that took western scientists years to formally describe. So, previous workers who failed to generate cladistic classifications and were not aware of the proper names for their taxa may well have recognized and characterized natural groups. So, one tries to understand how one's predecessors, who looked at the same organisms, decided how to organize their observations into taxa. Thus all who want to truly master nomenclature and classification are well advised to examine carefully what their precedessors did, appreciating what Newton once wrote: "If I have seen further, it is by standing upon the shoulders of Giants."

See Also the Following Articles

Biodiversity • History of Entomology • Museums and Display Collections • Phylogeny

Further Reading

Banks, N., and Caudall, A. N. (1912). The entomological code—a code of nomenclature for use in entomology. Washington, DC.

Benton, M. J. (2000). Stems, nodes, crown clades, and rank-free lists: is Linnaeus dead? *Biol. Rev.* **75,** 633–648.

Forey, P. L. (2001). The PhyloCode: description and commentary. *Bull. Zool. Nomencl.* **58,** 81–96.

Hennig. W. (1966). "Phylogenetic Systematics." University of Illinois Press, Urbana.

Huxley, J. S. (1957). The three types of evolutionary process. *Nature* **180,** 454–455.

International Commission on Zoological Nomenclature. (1999). "International Code of Zoological Nomenclature," 4th ed. International Trust for Zoological Nomenclature, London.

Mayr, E., Linsley, E. G., and Usinger, R. L. (1953). "Methods and Principles of Systematic Zoology." McGraw Hill, New York.

Melville, R. V. (1995). "Towards Stability in the Names of Animals. A History of the International Commission on Zoological Nomenclature 1895–1995. International Trust for Zoological Nomenclature, London.

Schuh, R. T. (2000). "Biological Systematics. Principles and Applications." Cornell University Press, Ithaca, NY.

Stearn, W. T. (1957). An introduction to the *Species Plantarum* and cognate botanical works of Carl Linnaeus. *In* "C. Linnaeus, *Species Plantarum.*" [A facsimile of the first edition of 1753]. The Ray Society, London.

Thompson, F. C. (1996). Names: the keys to biodiversity. *In* "BioDiversity II" (Reaka-Kudla, M. L., Wilson, D. E. and Wilson, E. O. eds.), pp. 199–212. Joseph Henry Press, Washington, DC.

Wheeler, Q. D., and Meier, R. (eds.). (2000). "Species Concepts, Phylogenetic Theory: A Debate." Coumbia University Press, New York.

Wiley, E. O. (1981). "Phylogenetics. The Theory and Practice of Phylogenetic Systematics." Wiley, New York.