

Toward a phylogenetic system of biological nomenclature

Kevin de Queiroz and
Jacques Gauthier

Despite the widely held belief that modern biological taxonomy is evolutionary, some of the most fundamental concepts and principles in the current system of biological nomenclature are based on a non-evolutionary convention that pre-dates widespread acceptance of an evolutionary world view by more than a century. The development of a phylogenetic system of nomenclature requires reformulating these concepts and principles so that they are no longer based on the Linnean categories but on the tenet of common descent.

Kevin de Queiroz is at the Dept of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC 20560, USA; Jacques Gauthier is at the Dept of Herpetology, California Academy of Sciences, Golden Gate Park, San Francisco, CA 94118, USA.

Biological nomenclature is the basic language by which scientists communicate about the diversity of living things. This language consists of the names given to countless extinct and extant species and the nested series of higher taxa to which they belong. In biological nomenclature, as in any language, communication is most effective if the meanings of its terms, in this case taxon names, are explicit, universal and stable^{1,2}. That is, the association between a name and a taxon should be clear, all biologists should use the same name for the same taxon, and a name should not designate different taxa, nor a taxon be designated by different names, at different times. In an attempt to achieve these goals, biologists have developed elaborate sets of concepts and principles (rules) governing the use of taxon names. These concepts and principles form a nomenclatural system, which is spelled out in books called the codes of nomenclature³⁻⁵ and which makes up an important part of the taxonomic system, the entire unified body of principles and rules governing taxonomic practice.

Ever since the revolution in biology precipitated by Darwin's seminal work⁶, taxonomists have increasingly come to

accept the proposition that taxonomy is to be based on the tenet of evolution. Species have come to be equated with population lineages⁷, and higher taxa have come to be equated with clades⁸. Given that population lineages and clades are the entities biologists name as taxa, then taxon names have implicit, if not explicit, phylogenetic meanings. Consequently, biological nomenclature requires rules that promote explicit, universal and stable names of particular species and clades.

Unfortunately, the current nomenclatural system fails to accomplish these goals. Under an evolutionary interpretation of higher taxa and their names, the non-evolutionary basis of the current system causes nomenclatural ambiguity, parochialism and instability. Here, we identify the most fundamental principle of nomenclature, and we illustrate how this principle can be reformulated with an evolutionary basis in order to promote nomenclatural explicitness, universality and stability. Our use of various terms relevant to this discussion is summarized in Box 1.

The current system

The current nomenclatural system is clearly non-evolutionary. After all, this system is based on conventions developed by Linnaeus⁹ more than 100 years before the widespread acceptance of the tenet of common descent. Although the first international codes of nomenclature came into existence after the publication of *On the Origin of Species*^{2,10-12} and currently include elements that were not present in Linnaeus' work, the codes are Linnean rather than evolutionary in that their most fundamental concepts and principles are based on the Linnean taxonomic categories. Moreover, these concepts and principles make reference neither to common descent nor to any other evolutionary phenomenon. But most importantly, under evolutionary interpretations of higher taxa and their names, the current system fails to accomplish its own stated purpose.

The failure of the current system is evident in cases involving the taxonomic practices of unification and division, commonly referred to as 'lumping' and 'splitting', which are intimately tied to the Linnean categories (Box 1). In situations involving either lumping or splitting,

taxon names change their evolutionary meanings, while evolutionary taxa either change or lose their names (Box 2). In the case of lumping, names change their designations from less-inclusive to more-inclusive taxa. In the case of splitting, the situation is reversed; names change their designations from more-inclusive to less-inclusive taxa. These changes in the meanings of taxon names are manifested in the synonymizing of names that previously referred to different taxa (in the case of lumping) and in assigning previously synonymous names to different taxa (in the case of splitting).

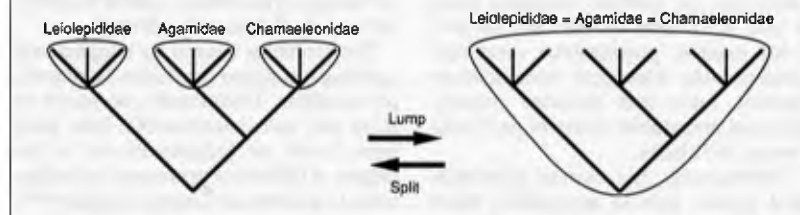
The problems caused by lumping and splitting transcend particular taxonomic philosophies. Traditionally, decisions to lump and split supraspecific taxa have been based on judgements as to the degree of difference necessary for assignment to a particular Linnean category^{13,14}. Lately, however, such decisions have often been prompted by the desire to bring taxonomy into line with cladistic views^{8,15-17} on what constitutes a valid taxon – that is, to eliminate paraphyletic taxa (Box 3). But regardless of the reason for lumping or splitting, the result is the same: names change their associations

Box 1. Terminology

A **nomenclatural system** governs the names of taxa. A **taxon** is a named group of organisms, whether an abstract class or a concrete **species** (population lineage) or **clade** (group of species sharing an exclusive common ancestry). The names of taxa are **taxon names**. To **unite** or **lump** is to combine two or more taxa assigned to the same Linnean category into a single taxon at that categorical level (e.g. if several families are lumped, their included species are now considered to make up a single family). To **divide** or **split** is to partition a single taxon assigned to a particular Linnean category into multiple taxa at that categorical level (e.g. if a single family is split, its included species are now considered to constitute several families). A **paraphyletic taxon** consists of a common ancestor and some, but not all, of its descendants; a **monophyletic taxon** (**named clade**) consists of a common ancestor and all of its descendants. **Definitions** are statements specifying the meanings of taxon names, whereas **diagnoses** specify how taxa or their component parts are recognized. The **content** of a taxon is the set of its component organisms or species. **Synonyms** are names that are spelled differently but designate the same taxon. **Homonyms** are names that are spelled the same but designate different taxa. The principle of **priority** states that the oldest of several synonyms is the valid name of a taxon. A **type** is a specimen or subordinate taxon to which the name of a taxon is permanently attached.

Box 2. Nomenclatural changes caused by lumping and splitting

Lumping the three families Leiolepididae, Agamidae and Chamaeleonidae results in the names of three different taxa (left) becoming names of a single inclusive taxon (right). Although only one of these names would be considered the valid name of the inclusive taxon, all three names nonetheless become synonyms. Meanwhile, the three taxa originally designated by these names must either go without names or be assigned new ones (e.g. 'Leiolepidinae', 'Agaminae', 'Chamaeleoninae'). Conversely, splitting the single family results in the synonymous names of a single inclusive taxon (right) becoming the non-synonymous names of three different taxa (left). Meanwhile, the inclusive taxon previously designated by one of these names must either go nameless or be assigned a new name (e.g. 'Chamaeleonidea'). Furthermore, if the three less-inclusive taxa had been named previously (e.g. as subfamilies), they would have to change their names (e.g. from 'Leiolepidinae' to 'Leiolepididae', etc.).



from one clade or ancestor to another, and clades either change or lose their names.

In short, the current nomenclatural system's basis in the Linnean taxonomic categories promotes neither explicitness, universality nor stability with regard to the phylogenetic meanings of taxon names. Whenever opinions differ regard-

ing the categorical assignments of taxa, authors use the same names for different taxa and different names for the same taxa¹⁸. Variation among authors in the associations between taxa and taxon names implies that the meanings of those names are not universal, and hence the names are equivocal. Needless to say, this situation generates considerable

ambiguity. Furthermore, even when a proposed change in categorical assignment is generally accepted, the meanings of taxon names are unstable. That is, although the associations between taxa and names may be universal at a given time, these designations change from one time to another. This not only hinders communication, thus undermining the very purpose of nomenclature, it also contradicts the avowed goals¹⁻⁵ of the codes.

The phylogenetic system

Definitions of taxon names

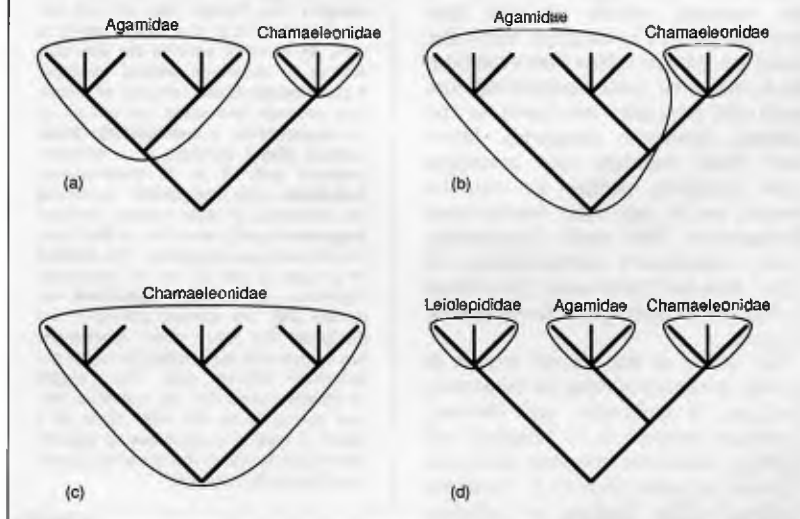
How is biology to achieve a system of nomenclature that promotes explicit, universal and stable meanings of taxon names with regard to what they signify about common ancestry? The answer is related to what is perhaps the most fundamental principle of nomenclature, namely, the manner in which taxon names are defined. Although the codes say very little about definitions explicitly, emphasizing instead the concepts of synonymy and homonymy and the principle of priority, all of these concepts and principles rest ultimately on definitions. Priority is a criterion used to establish the valid name from among several synonyms (names that are spelled differently but designate the same taxon) or homonyms (names that are spelled the same but designate different taxa)^{1,3-5}. But the designations of taxon names are established by definitions. Thus, definitions are the foundation of biological nomenclature.

Traditionally, biologists and philosophers have thought that taxon names are defined by specifying the properties that are necessary and sufficient for taxon membership^{19,20}, and they have generally assumed that those properties take the form of organismal traits^{21,22}. But when we consider just what it is about a name that remains constant in the face of taxonomic changes or differences in opinion under the current nomenclatural system, we see that the implicit definitions of taxon names are based not on organismal traits but rather on nomenclatural types and Linnean categories. 'Agamidae', for example, is effectively defined as the taxon containing the (type) genus *Agama* that is assigned to the family category. There is nothing intrinsically evolutionary about such a definition, for it makes no reference to any evolutionary phenomenon.

In contrast with traditional definitions, phylogenetic definitions of taxon names²¹⁻²³ (Box 4) are stated in terms neither of organismal traits nor of types and Linnean categories. Instead, they are stated in terms of common descent and the phylogenetic entities deriving their

Box 3. Elimination of paraphyletic taxa

Under the current system, paraphyletic taxa are eliminated either by lumping or by splitting, both of which result in nomenclatural changes. (a) Shows relationships implied by an earlier taxonomy, while (b) revised hypothesis of relationships based on new data or methods of analysis, and according to which Agamidae is paraphyletic. Under the current system, the paraphyletic family Agamidae can be eliminated either by lumping it with its derivative family into a single monophyletic family (c), or by splitting it into several monophyletic families (d). In the former case, 'Chamaeleonidae' changes its designation from a less-inclusive clade to a more-inclusive one, and the clade previously designated by that name must either be assigned a new name (e.g. 'Chamaeleoninae') or go nameless. In the latter case, 'Agamidae' changes its designation from a more-inclusive paraphyletic taxon to a less-inclusive monophyletic taxon, and the two newly named clades may have to change their names (e.g. if they were previously named as the subfamilies 'Leiolepidinae' and 'Agaminae').



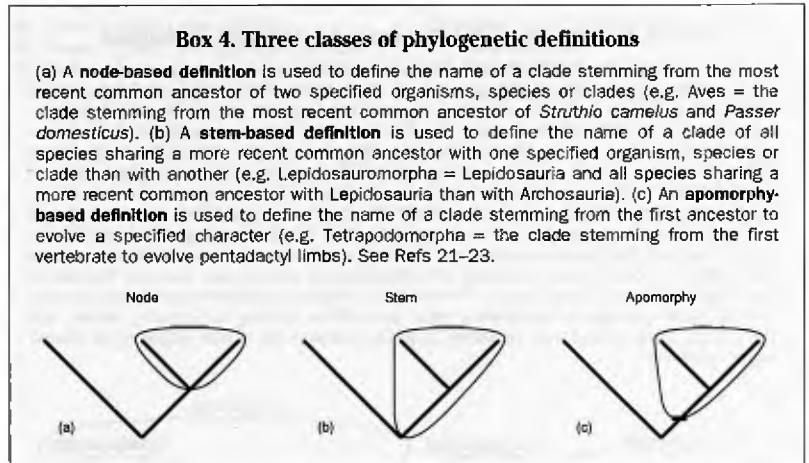
existence from that process. For example, the name 'Agamidae' might be defined as the clade stemming from the most-recent common ancestor of *Agama* and *Leiolepis*. Such a definition is thoroughly evolutionary in that the concept of common ancestry is fundamental to the meaning of the name.

The reason that the current nomenclatural system fails under an evolutionary view of taxa relates directly to the non-evolutionary manner in which it defines taxon names. By defining taxon names on the basis of Linnean categories, the current system effectively treats Linnean categorical assignments as if they are more fundamental to the meanings of taxon names than are the associations of those same names with monophyletic taxa²³. Consequently, that which remains universal and stable is the association of each taxon name with a particular Linnean category rather than with a particular clade.

Phylogenetic definitions solve this problem by making the associations between taxon names and monophyletic taxa the most fundamental aspect of the meanings of those names. In so doing, phylogenetic definitions provide the basis for developing a nomenclatural system with explicit, universal and stable evolutionary meanings of taxon names (Box 5). To the extent that definitions are clearly formulated, the associations of taxon names with particular clades will be unambiguous. Similarly, to the extent that all biologists adopt the same definition for each name, both now and in the future, the evolutionary meanings of taxon names will be universal and stable. Of course, simply adopting a phylogenetic approach to definitions does not guarantee clear formulation or widespread acceptance of particular definitions. Nevertheless, explicitness, universality and stability can be promoted by adopting secondary rules and recommendations designed to promote clear formulation and widespread acceptance. A phylogenetic perspective on other nomenclatural concepts and principles serves to illustrate this point.

Synonymy

Because definitions are central to biological nomenclature, basing definitions on common ancestry has implications for other nomenclatural concepts and principles that, together with phylogenetic definitions, form the core of the phylogenetic system of nomenclature. Although the fundamental concepts and principles of phylogenetic nomenclature have counterparts or analogs in the current taxonomic system, they differ significantly from their non-evolutionary



counterparts as a result of the fundamental difference concerning definitions.

One of the most important nomenclatural concepts is that of synonymy. Because definitions of taxon names in the current system are based on types and Linnean categories, synonymy depends on the categorical assignments of the taxa containing the types upon which the names in question are based (Box 2). For example, if one of the clades containing both *Agama* and *Leiolepis* is assigned to the family category, then the names 'Agamidae' and 'Leiolepididae' are considered synonyms. But if the clades assigned to the family category are less inclusive, such that the type genera *Agama* and *Leiolepis* belong to different family level taxa, then the names 'Agamidae' and 'Leiolepididae' are not considered synonyms. The Linnean basis of definitions is the ultimate cause of changes in, or disagreement about, synonymy resulting from taxonomic lumping and splitting.

Ambiguity and instability of this kind are solved by phylogenetic definitions, which carry with them an evolutionary criterion of synonymy. In the phylogenetic system, synonymy has nothing to do with the Linnean categories. Names are synonymous if they designate the same clade^{21,23}. Thus, in contrast with the situation under the current system, 'Chamaeleonidae' would not be considered a synonym of 'Agamidae' in the context of any of the relationships illustrated in Boxes 2, 3 and 5, because the definitions of the two names (Box 5) indicate that they designate different clades. Moreover, disagreements about phylogenetic synonymy reflect disagreements about objective phylogenetic relationships²¹, while those about Linnean synonymy reflect disagreements regarding subjective Linnean categorical assignments.

Priority

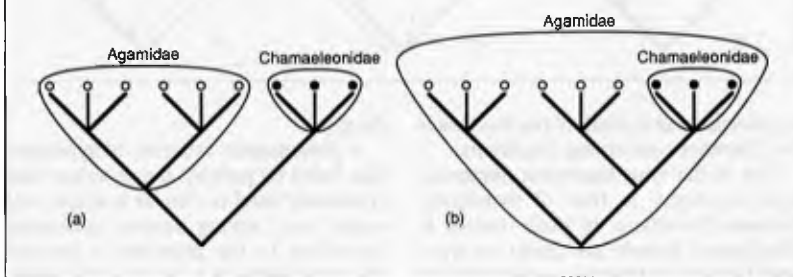
A phylogenetic criterion of synonymy also bears on priority, the principle most commonly used to choose a single valid name from among several synonyms. According to the principle of priority, the valid name of a taxon is the oldest name applied to it. Recent proposals for nomenclatural reform within the Linnean system have sought to restrict the application of the principle of priority^{24,25}. In contrast, the relevance of phylogenetic definitions to priority concerns not the use of priority in establishing the valid names of taxa but rather the concept of priority itself. Because the phylogenetic and Linnean systems derive their definitions from different first principles, the two systems are again in sharp contrast. In the Linnean system, priority is established by first use of a name in association with a particular Linnean category; in the phylogenetic system, priority is established by first use of a name to designate a particular clade^{21,23}. Thus, in contrast with the situation under the current system, 'Chamaeleonidae' would not have priority over 'Agamidae' in the context of any of the relationships illustrated in Boxes 2, 3 and 5, because the two names are not synonymous, that is, they designate different clades.

Knowledge of relationships and freedom of taxonomic opinion

Despite striking differences between the phylogenetic and Linnean systems of nomenclature, the phylogenetic system accords with a basic principle of the current codes in preserving freedom of taxonomic thought and action³⁻⁵. In other words, taxonomists still must determine the contents and diagnostic characters of taxa. Consequently, adopting a phylogenetic system of nomenclature requires neither agreement about, nor a detailed knowledge of, phylogenetic relationships,

Box 5. Nomenclatural stability under phylogenetic definitions

Under a phylogenetic system of nomenclature, names retain their associations with particular clades or ancestors despite changes in ideas about relationships. (a) Relationships implied by an earlier taxonomy according to which the following definitions are formulated: Agamidae = the clade stemming from the most recent common ancestor of the species represented by open circles; Chamaeleonidae = the clade stemming from the most recent common ancestor of the species represented by filled circles. (b) Revised hypothesis of relationships based on new data or methods of analysis. Under phylogenetic definitions, 'Chamaeleonidae' retains its association with the clade stemming from the most recent common ancestor of the species represented by filled circles, although the chamaeleonid species are now also thought to have descended from this ancestor. The manner in which the definitions are stated ensures that no names designate paraphyletic taxa, and neither splitting nor lumping occurs, but hierarchical relationships may be altered (e.g. Chamaeleonidae is now judged to be nested within Agamidae).



which admittedly are lacking for many higher taxa. For example, under the definition of 'Mammalia' as the clade stemming from the most recent common ancestor of monotremes and therians²⁶, taxonomists are free to accept²⁷ or reject^{26,28} the conclusion that monotremes and therians share a more recent common ancestor with one another than either does with various fossil taxa (e.g. *Kuehneotherium*). And even if the relationships of these fossils remain unresolved, this does not affect the definition of 'Mammalia,' which designates the clade stemming from a particular common ancestor regardless of our knowledge concerning its precise contents. Thus, taxonomists are free to differ in their opinions about relationships, included species and diagnostic characters in accordance with their interpretations of the available evidence, and they are free to change their views about relationships, contents and diagnoses in light of new data and improved analytical methods²¹.

Conclusion

A phylogenetic perspective on taxonomic definitions reveals fundamental incompatibilities between the current nomenclatural system and an evolutionary view of higher taxa. Specifically, the acts of lumping and splitting – which generate the problems discussed above and are basic to taxonomic revision in the current system – are nonsensical in a phylogenetic context. As singular phylogenetic entities (clades), taxa are not things that taxonomists can unite or

divide, but only things that they can discover and name. Furthermore, because a phylogenetic approach to nomenclature would preserve the spellings of taxon names in the face of changing ideas about relationships, the Linnean rank-associated endings would be rendered meaningless. For example, Chamaeleonidae might turn out to be nested within Agamidae (Box 5). Thus, phylogenetic definitions of taxon names not only form the basis of a nomenclatural system that is compatible with an evolutionary view of higher taxa, they also highlight the impossibility of developing such a system based on the Linnean categories.

Taxonomy is fundamental to biology. It provides a reference system that permits communication and access to the literature, as well as a context for comparative biology. In order to carry out these functions effectively, taxonomy must be governed by a unified body of concepts and principles designed to accomplish its practical goals within an appropriate theoretical context. The Linnaean categories have served as the foundation of the taxonomic system for more than 200 years. But if evolution is truly the unifying theory of biology, then the Linnean categories no longer provide an appropriate theoretical context. Modern biology requires a taxonomic system based on evolutionary concepts and principles. The New Systematics^{14,29,30} and Phylogenetic Systematics^{8,15-17} initiated the development of such a system by granting the tenet of evolution a central role in concepts of species and higher

taxa²³, respectively. The next logical step in this process is to reformulate the concepts and principles of biological nomenclature so that they too are based on the tenet of evolution.

Acknowledgements

We thank A.M. Bauer, D.C. Cannatella, J.M. Clark, M.J. Donoghue, R. Etheridge, M.T. Ghiselin, D.R. Frost, R.J. O'Hara, C. Patterson, T. Rowe, E.O. Wiley and the late R. Estes for helpful discussions; M.J. Donoghue, M.T. Ghiselin, and R.J. O'Hara also provided comments on an earlier version of the paper.

References

- Jelley, C. (1989) *Biological Nomenclature*, Edward Arnold
- Ride, W.D.L. (1988) in *Prospects in Systematics* (Hawksworth, D.L., ed.), pp. 332–353, Clarendon Press
- International Association of Microbiological Societies (1992) *International Code of Nomenclature of Bacteria*, American Society for Microbiology
- International Commission on Zoological Nomenclature (1985) *International Code of Zoological Nomenclature*, International Trust for Zoological Nomenclature
- International Botanical Congress (1988) *International Code of Botanical Nomenclature*, Koeltz Scientific Books
- Darwin, C. (1859) *On the Origin of Species by Means of Natural Selection*, John Murray
- O'Hara, R.J. (1993) *Syst. Biol.* 42, 231–246
- Hennig, W. (1966) *Phylogenetic Systematics*, University of Illinois Press
- Linnaeus, C. (1758) *Systema Naturae Per Regna Tria Naturae, Secundum Classes, Ordines, Genera, Species, Cum Characteribus, Differentiis, Synonymis, Locis, Laurentii Salvii*
- McNeill, J. and Greuter, W. (1986) in *Biological Nomenclature Today: A Review of the Present State and Current Issues of Biological Nomenclature of Animals, Plants, Bacteria, and Viruses* (Ride, W.D.L. and Younes, T., eds), pp. 3–25, IRL Press
- Ride, W.D.L. (1986) in *Biological Nomenclature Today: A Review of the Present State and Current Issues of Biological Nomenclature of Animals, Plants, Bacteria and Viruses* (Ride, W.D.L. and Younes, T., eds), pp. 26–35, IRL Press
- Nicolson, D.H. (1991) *Ann. Mo. Bot. Gard.* 78, 33–56
- Simpson, G.G. (1960) *Principles of Animal Taxonomy*, Columbia University Press
- Mayr, E. (1969) *Principles of Systematic Zoology*, McGraw-Hill
- Wiley, E.O. (1981) *Phylogenetics: The Theory and Practice of Phylogenetic Systematics*, John Wiley & Sons
- Ax, P. (1987) *The Phylogenetic System: The Systematization of Organisms on the Basis of Their Phylogenesis*, John Wiley & Sons
- Eldredge, N. and Cracraft, J. (1980) *Phylogenetic Patterns and the Evolutionary Process*, Columbia University Press
- Griffiths, G.C.D. (1976) *Syst. Zool.* 25, 168–173
- Buck, R.C. and Hull, D.L. (1966) *Syst. Zool.* 15, 97–111
- Ruse, M.E. (1973) *The Philosophy of Biology*, Hutchinson University Library

- 21 de Queiroz, K. and Gauthier, J. (1990) *Syst. Zool.* 39, 307-322
- 22 de Queiroz, K. (1992) *Biol. Philos.* 7, 295-313
- 23 de Queiroz, K. and Gauthier, J. (1992) *Annu. Rev. Ecol. Syst.* 23, 449-480
- 24 Savage, J.M. (1990) *Syst. Zool.* 39, 424-425
- 25 Stace, C.A. (1991) *Nature* 350, 466
- 26 Rowe, T. (1988) *J. Vert. Paleontol.* 8, 241-264
- 27 Hopson, J.A. and Crompton, A.W. (1969) *Evol. Biol.* 3, 15-72
- 28 Kemp, T.S. (1983) *Zool. J. Linn. Soc.* 77, 353-384
- 29 Huxley, J., ed. (1940) *The New Systematics*, Clarendon Press
- 30 Mayr, E. (1942) *Systematics and the Origin of Species*, Columbia University Press