



TaxonSearch: a relational database for suprageneric taxa and phylogenetic definitions

Paul C. Sereno¹, Steve McAllister², Stephen L. Brusatte³

¹ *Department of Organismal Biology and Anatomy, University of Chicago, 1027 E. 57th St., Chicago, IL 60637 USA*

Email: dinosaur@uchicago.edu

² *Email: steve@mcallister.tv*

³ *Department of Geophysical Sciences, University of Chicago 5734 S. Ellis Ave., Chicago, IL 60637 USA*

Received: 15 August 2005 - Accepted: 4 November 2005

Abstract

Over the last 25 years, phylogenetic analysis and phylogenetic taxonomy have narrowed the meaning of suprageneric taxa. In phylogenetic analysis, suprageneric taxa identify clades rather than an unspecified mixture of monophyletic groups (clades), paraphyletic or polyphyletic grades, and redundant taxa erected solely to occupy rank. In phylogenetic taxonomy, clades are circumscribed by definition rather than a variable assortment of diagnostic characters. This phylogenetic approach has reinvigorated taxonomy and substantially increased the number of suprageneric taxa in current use. Basic information about suprageneric taxa, nevertheless, remains scattered across a vast taxonomic literature. We introduce a web-based application called *TaxonSearch*, which aims to provide practicing taxonomists with an efficient tool for logging, locating, and sharing information regarding suprageneric taxa (author, citation, definition, composition, history, temporal duration). An example compilation, "Stem Archosauria," is composed of 789 taxonomic records that include all suprageneric taxa ever applied to these animals. These records may be searched, sorted or summarized in many ways. Given the pace of phylogenetic work and the increasing number of phylogenetic definitions, efficiently locating information about suprageneric taxa is a growing and critical need.

Introduction

Over the last 25 years, there has been a dramatic increase in resolution of branch points in the tree of life driven principally by three factors: an increasing volume of

phylogenetic research, widespread use of computer-assisted analysis (Swofford et al., 1996; Felsenstein, 2004), and the influx of molecular data (Hillis et al., 1996). Increased phylogenetic resolution has generated a concomitant increase in the

number of suprageneric taxa, which serve as labels for branch points (clades) of interest.

Darwin (1859) famously predicted that taxonomy would come to reflect knowledge of genealogy, and that knowledge is currently undergoing rapid advance. Modern taxonomists, nonetheless, are in the unenviable position of riding a tidal wave of interest in biodiversity and its evolutionary history while lacking the tools to efficiently log, locate, and share basic information regarding suprageneric taxa. Such information includes original authorship, bibliographic citation, phylogenetic definition, taxonomic composition, taxonomic history, and estimated temporal duration.

Taxonomy is now subject to an international effort to create web-accessible databases that log living diversity and attempt to integrate specimen inventories, morphology, and biogeography (Bisby, 2000). Under the rubric "biodiversity informatics," the undertaking has at its core a compilation of Linnaean binomials (genus plus species). Typically only a few ranked suprageneric taxa (e.g. Phyla, Class, Order) are included for organizational purposes. Such higher-level taxa have long been understood as a mixed lot: some represent monophyletic clades; some constitute paraphyletic or polyphyletic assemblages; others are byproducts of a redundant ranking convention.

The aim of phylogenetic taxonomy is to change that circumstance by basing taxonomic names solely on genealogy. In phylogenetic taxonomy, the content of a taxon is governed (specified) by a phylogenetic definition that points to a particular branch point, or clade (de Queiroz and Gauthier, 1990, 1992; Sereno, 2005). Phylogenetic definitions, nevertheless, have added yet another dimension to the dispersed literature governing taxonomy (Minelli, 2003).

The Void

You are a taxonomist in the middle of a manuscript describing a new organism. You

decide it forms a new clade with several other organisms, and you begin to draft an indented taxonomy. But which higher-level taxonomy to use, and who are the original authors? You intend to name a new clade—but has it already been coined? Or is there an older taxon that has fallen from use that might fit the bill? What phylogenetic definition might best accommodate others that have been proposed? And who proposed those definitions and in what publications? And how old are the oldest fossils of closely related clades?

Few on-line taxonomic compilations for extant or extinct clades can efficiently answer many of these questions. Taxonomists typically reach for the nearest relevant taxonomic treatise, summary, or paper and search for answers. Using dinosaurs as an example, a second edition of *The Dinosauria* (Weishampel et al., 2004) would provide helpful clues and might be the first time a specialist in this field would consult. This work, however, does not provide taxonomic background to the higher taxa included in the many disparate indented taxonomies; there is no appendix of comparable taxa that were coined but not included; there is no compilation of recommended taxonomic definitions or log of previous phylogenetic definitions; and citations to the taxonomic literature are not available on-line and are dispersed among hundreds of references to other topics. The taxonomic summaries that are included, furthermore, often perpetuate mistaken spellings, attributions or citations, which are cumbersome to correct in printed literature. And, finally, none of these standard resources facilitate on-line communication with the taxonomists that use them.

Current Taxonomic Databases

Biodiversity Informatics

The rapidly expanding field of web-based bioinformatics functions as a repository for global biodiversity surveys at the species level (Bisby, 2000; Gewin, 2002). Dozens of relational database applications have arisen that integrate species, specimen numbers, geographic ranges, literature citations,

images and other information, while limiting reference to higher taxa to a few mandatory Linnean ranks (Table 1). In such surveys, variation in supraspecific taxa between taxonomists is viewed as an unnecessary evil (Bisby, 2000; Godfray, 2002; Mallet and Willmott, 2002). As a consequence, input of suprageneric taxa is usually limited to the most commonly used, mandatory Linnean ranks (Phylum, Class, Family). Indented taxonomies that incorporate higher taxa are often presented as a given, without literature citation or any other related information. Species, genera, and suprageneric taxa whose members are extinct are ignored, as is the temporal duration of taxa. These efforts are strongly focused on living species in the Recent.

Similarly, databases have arisen with a focus on extinct organisms. *The Paleobiology Database*, the most elaborate, aims to provide collection-level information on the spatial, temporal, and environmental distribution of fossils and includes taxonomic lists keyed to a taxonomic hierarchy, references and range data (Table 1). The temporal and spatial range of extinct species is the principal focus of this database. Taxa are nested within a single ranked hierarchy, the basis for which is not given; there is little associated phylogenetic information and no link to phylogenetic definitions. Other taxonomic databases, such as *Mammals of the World*, *Dinobase*, or *DinoData* (Table 1), are compilations of species organized within a singular taxonomic hierarchy. They also include distributional data and author citations, although none are designed to allow efficient searching/sorting or effectively manage phylogenetic definitions.

Taxonomy Databases

Several recent applications have appeared that incorporate "multiple taxon concepts" (Berendsohn, 1995, 1997) (*MoReTax*, *Nomenclurator*, *Prometheus*, *Taxonomer*, Table 1). They present information models to establish patterns in character-based, ranked taxa as conceived in traditional Linnean taxonomy and allow more than one concept to be associated with a particular taxon name. "The circumscription of a taxon

may vary from one author to another, the system the authors adhere to may be different . . . The data model must allow the preservation of these different concepts, and document the errors" (Berendsohn, 1997:288). The term 'classification' is here understood as "the placing of a plant . . . or group of plants." *Nomenclurator*, for example, defined a taxon as "an abstraction of one or more individuals for species or lower classification levels, or taxa of lower rank for classification levels higher than species" (Ytow et al., 2001:84). "Meaningful circumscription of the taxon," the authors explained, "relies on a knowledge of closely related taxa which may not be available at the time of original definition," and the taxon concept is "at best a summary of it as perceived by the author at the time of publication and in the context of contemporary knowledge" (Ytow et al., 2001:84). *Taxonomer* described a taxon as "an objective entity . . . linked to the biological world via a properly designated type specimen" and a taxon concept as "a purely abstract, subjective construct that ultimately exists only in the mind of the taxonomist" (Pyle, 2004:18). Taxa and their respective taxon concepts are linked by "protonyms" and "assertions," which constitute the treatment or descriptions of taxa by various authors.

Nomenclurator, *Taxonomer* and several other taxonomic applications are ill-suited for phylogenetic taxonomists, because taxa are not regarded as phylogenetic entities circumscribed by phylogenetic definitions. In Linnaean taxonomy, specific and supraspecific taxa are often delineated by "differential diagnoses" (Mayr et al., 1953), which require only a differential, or phenetic, difference between taxa. Taxa created under this paradigm may be monophyletic, paraphyletic or even polyphyletic. Attempting to track or synthesize such unwieldy traditional taxon concepts in a computerized environment is a daunting task that has encouraged the formulation of data standards (Taxonomic Databases Working Group; <http://www.tdwg.org>). As can be seen from the conceptual framework summarized in the paragraph above, nevertheless, there is little here for the phylogenetic taxonomist. Indeed, this

Table 1. Sampling of database projects available or in progress that involve compilations of biological taxa.

Name	Home	Website/Documentation	Content
BibMaster	Global Biodiversity Information Facility, Madrid, Spain	http://www.gbif.es/index_in.php	Biodiversity database using standard Linnean ranks
BONAP	North Carolina Botanical Garden, Chapel Hill, NC	http://www.bonap.org/summary.html	Biodiversity database using standard Linnean ranks
Dinobase	University of Bristol, Bristol, UK	http://palaeo.gly.bris.ac.uk/dinobase/dinopage.html	Database for dinosaur genera, species, and general information
DinoData	Lisbon, Portugal	http://www.dinodata.net/	Database for dinosaur genera, species, and general information
EMBL (Reptile Database)	European Bioinformatics Institute, Cambridgeshire, UK	http://www.embl-heidelberg.de/~uetz/Reptiles.html	Genus/species by family of nonavian reptiles using standard Linnean ranks
GBIF (Global Biodiversity Information Facility)	University of Copenhagen, Denmark	http://www.gbif.org/	Facilitating digitisation and global dissemination of primary biodiversity data
GTI (Global Taxonomy Initiative)	Australian Biological Resources Study, Canberra, Australia	http://www.biodiv.org/programmes/cross-cutting/taxonomy/	Biodiversity database using standard Linnean ranks
ITIS (Integrated Taxonomic Information System)	multi-institutional	http://www.itis.usda.gov/index.html	Biodiversity database coordination program
Linnaeus II	Expert Center for Taxonomic Identification, University of Amsterdam, Netherlands	http://www.eti.uva.nl/Products/Linnaeus.html	Biodiversity database for species, hierarchies, identification keys, distributions
MacTaxon	Katholieke Universiteit Leuven, Belgium	http://www.kuleuven.be/bio/sys/MacTaxonlite/home.html	Receives genus/species and locality data for iMap
Mammal Species of the World (MSW)	Smithsonian Institution, Washington, DC	http://nmgoph.si.edu	Family, genera, and species information about mammals
MoReTax	Botanic Garden and Museum, Berlin, Germany	http://www.bgbm.org/BioDiversity/Projects/MoreTax/default.html	Database model that manages to account for varying taxon concepts
Nomenclator	Bishop Museum, Honolulu, HI	http://www.nomenclator.org/MultipleViews.html	Biodiversity database model managing competing taxon concepts in traditional taxonomy
The Paleobiology Database	UCSB, Santa Barbara, CA	http://www.paleodb.org/	Extensive range, geographic, and taxonomic data for invertebrates and some vertebrates
PhyloCode	Ohio University, Athens, OH	http://www.ohiou.edu/phylocode	Database planned for established phylogenetic definitions and taxa
PLATYPUS	Australian Biological Resources Study, Canberra, Australia	http://www.deh.gov.au/biodiversity/abrs/online-resources/software/platypus/index.html	Biodiversity database using standard Linnean ranks
Prometheus	Napier University, Edinburgh, UK	http://www.dcs.napier.ac.uk/~prometheus/prometheus_1/taxa2.htm	Biodiversity database model trying to account for varying taxon concepts in traditional taxonomy

Species 2000	University of Reading, UK	www.sp2000.org/secretariat.html	Database of botanic collections and taxonomic literature using standard Linnean ranks
SysTax	University of Ulm, Germany	http://www.biologie.uni-ulm.de/systax	Biodiversity database for species, input genus/species by family
TAXIS (Taxonomic Information System)	University of Helsinki, Finland	http://www.bio-tools.net/index.htm	Interactive database for recording taxonomic data
TaxLink	University of Göttingen, Germany	Gradstein, <i>Taxon</i> 2001:1075-1084	Program that manages competing taxon concepts
Taxonomer	Bishop Museum, Honolulu, HI	Pyle (2004), <i>PhyloInformatics</i> 1:1-54	Diverse input including taxonomic data that manages competing taxon concepts
The Fossil Record 2	University of Bristol, UK	Benton (1993) <i>The Fossil Record 2</i> ; http://palaeo.gly.bris.ac.uk/frwhole/FR2intro.html	Family (class/phylum) data through time
ToL (The Tree of Life)	University of Arizona, Tuscon, AZ	http://tolweb.org/tree/phylogeny.html	Interlinked web pages containing information about biodiversity
W³TROPIC OS	Missouri Botanical Garden, St. Louis, MO	http://mobot.mobot.org/W3T/Search/vast.html	Botanic database using standard Linnean ranks

elaborate search for a rationale to define higher taxa under a traditional paradigm underscores the futility of such an exercise in the mind of a phylogenetic taxonomist.

The Tree of Life

The web-based *The Tree of Life* project links together phylogenies for the major groups of organisms (Table 1). The primary aim is to present a single tree of life for professional taxonomists, teachers, and the general public that allows rapid location of characters, images, general references or other information pertaining to a particular clade or species. In these regards, it is a superb resource. It does not aim to be a repository for available suprageneric taxa, encyclopedic literature citations, taxonomic intricacy, phylogenetic definitions, or estimated temporal duration.

PhyloCode

The *PhyloCode* (Cantino and de Queiroz, 2004) outlines a nomenclature or set of rules for naming clades and species (Table 1). The intention is to compile an officially-sanctioned database for “established” phylogenetic definitions and their corresponding new and “converted” taxon names. Although this database has yet to be constructed and the rules governing

phylogenetic nomenclature are still under active debate, the intention is clear enough: to create a single protocol and registry for surpaspecific taxa and, possibly, species as well. This goal—a unitary, universal taxonomy—is lauded by some as modern and necessary (Mallet and Willmott, 2002) and decried by others as profoundly impractical and authoritarian, including international organizations governing botanical and zoological nomenclature (Ereshevsky, 2001; Fortey, 2002; Greuter et al., 2000; ICZN, 1999; Kluge 2005; Sereno, 2005a).

Whatever happens with this initiative, the *PhyloCode* was never intended as a repository for either all higher-level taxa ever coined or all previous phylogenetic definitions. This information, nevertheless, is important information for taxonomists to bear in mind.

TaxonSearch: Conceptual Overview

TaxonSearch is a web-based relational database designed to assist in the compilation and rapid recall of basic information about suprageneric taxa (author, date of publication, bibliographic reference), their associated phylogenetic definitions

(author, date of publication, definitions, type of definition, is best specifiers, etc.), and their approximate age range in millions of years. *TaxonSearch* is devoted to suprageneric taxa. Although it is true that a genus with more than a single species constitutes a clade, *TaxonSearch* is best restricted to suprageneric taxa to clearly distinguish its higher-level purview from the principal focus of most biodiversity databases. *TaxonSearch* will locate records for suprageneric taxa based on any aspect, or combination of aspects, of the features that are logged for each taxon in the clade under consideration. Given the pace of phylogenetic work and the increasing

number of phylogenetic definitions, efficiently locating information about suprageneric taxa is a growing and critical need.

Users can access *TaxonSearch* at <http://www.taxonsearch.org>. The home page provides basic information about *TaxonSearch*, including the structure of the user interface, the range of search functions, and the protocol for creating a new *TaxonSearch* compilation for eventual posting (Figure 1). Users may send general comments or questions about the site by email. To access taxonomic compilations, the user must enter *TaxonSearch*, which

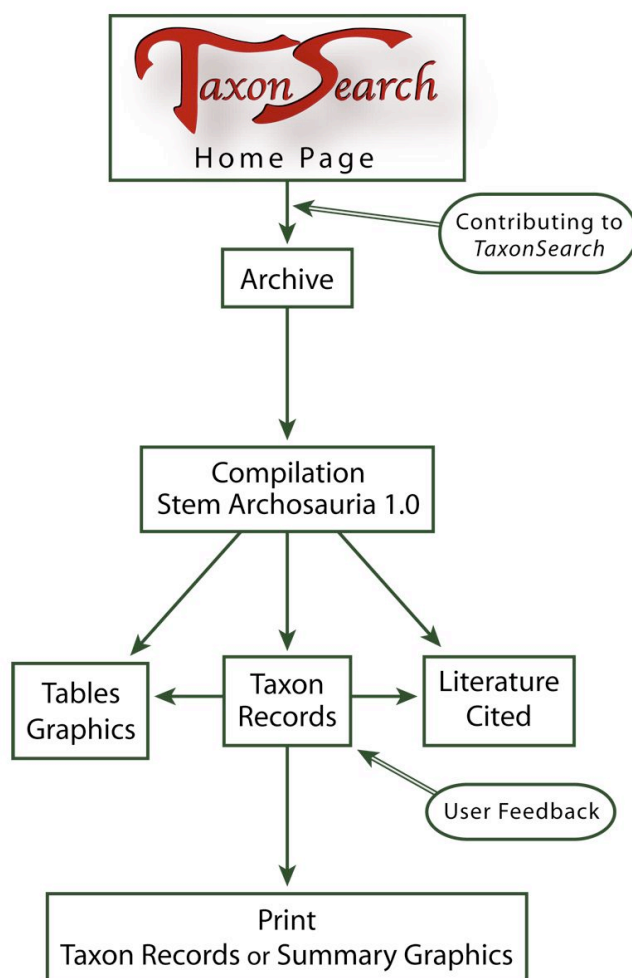



Figure 1. Conceptual overview of *TaxonSearch*, accessible at <http://www.taxonsearch.org>. The home page provides background information on using *TaxonSearch* and access to the archive, which lists posted compilations. Selecting the compilation Stem Archosauria 1.0 (Serenó, 2005b) takes you to this compilation's home page and access to its 789 taxon records, associated tables, graphics, and cited literature. After initial α - and β -development, users may create and submit new *TaxonSearch* compilations on-line or post feedback on individual taxon records in posted compilations.



Stem Archosauria 1.0

Compilation Stem Archosauria
Version 1.0
Original Compiler(s) Paul C. Sereno
2° Compiler(s) none
Date Posted November 07, 2005
Citation Sereno, P. C. 2005. Stem Archosauria—*TaxonSearch*
 URL <http://www.taxonsearch.org/Archive/stem-archosauria-1.0.php>
 [version 1.0, 2005 November 7]

[Home](#)
[Archive](#)

RECORDS

[Browse](#)
[Search](#)
[Summarize](#)

[Tables](#)
[Graphics](#)
[Literature Cited](#)
[Glossary](#)

No. Taxon Records 789
Keywords Archosauria, Crurotarsi, Crocodylomorpha, Dinosauria, Ornithischia, Saurischia, Sauropoda, Theropoda, Aves
Notes on Records Includes all suprageneric taxa within Archosauria that lie outside crown clades Crocodylia (extant crocodylians) and Neornithes (extant birds), excluding taxa within Pterosauroomorpha.
Associated Publication(s) P. C. Sereno. 2005. The logical basis of phylogenetic taxonomy. *Systematic Biology* 54:595-619.

 Sereno, P. C. in review. A phylogenetic taxonomy for Archosauria (Tetrapoda: Reptilia). *Journal of Vertebrate Paleontology*.

 Sereno, P. C., McAllister, S., and Brusatte, S. L. in press. *TaxonSearch*: a relational database for suprageneric taxa and phylogenetic definitions. *PhyloInformatics*.

Acknowledgements The author would like to thank S. McAllister for assistance in bringing *TaxonSearch* online and S. Brusatte for assistance in the compilation of taxon records.

Figure 2. The home page for the compilation Stem Archosauria includes important information about the compilation as well as access to its taxon records, tables, graphics, literature cited and glossary of terms. The functions available for viewing or sorting taxon records include *Browse* (which lists all records available), *Search* (which sorts records according to user input), and *Summarize* (which graphically summarizes user-selected taxon records).

opens an archive that lists all posted compilations. As these become more numerous, a search function using taxon keywords will allow the user to quickly locate a compilation of interest. Opening a compilation brings the user to that compilation's home page, which lists pertinent information including the version, compiler(s), date posted, proper citation, number of taxon records, and notes by the compiler(s) (Figure 2). Taxon records contain all pertinent information regarding the taxon. Users may browse, search, or graphically summarize sets of taxon records. Users may email feedback on any taxon record for posting with that record. Users cannot edit or change information in a *TaxonSearch* compilation. Compilers, likewise, cannot edit or change information in the taxon records of a posted *TaxonSearch* compilation. Compilers, however, are free to post a new version of the compilation. *TaxonSearch* compilations, thus, are archival, with a version number and posting date, and must be linked initially

to a publication in a peer-reviewed journal that introduces the compilation.

TaxonSearch would be much less effective if there were multiple or overlapping compilations for the same taxa by multiple compilers. One would need to open and compare more than a single compilation, and duplicate records would abound. Maintenance of the site, therefore, prohibits more than a single record for any taxon. How then does *TaxonSearch* differ from unitary proposals in taxonomy? Taxon records in *TaxonSearch* compilations, first, are open to posted commentary from other taxonomists. Each taxon record provides the option to post user feedback, which is then available to anyone accessing that record. Second, original compilers have a time limit (tentatively set at five years) for significantly updating their compilations to incorporate new taxonomic developments and user commentary as seen fit. If there are no substantive updates during that time

from the original compiler, or if that compiler voluntarily opens the compilation for revision, another taxonomist or group of taxonomists will have access to the original compilation for revision. The original compiler and all secondary compilers, in this case, are acknowledged on the home page of the compilation. With this protocol, taxonomic records in *TaxonSearch* compilations are open to differing points of view and are periodically updated with a minimum of site management.

Implementation

TaxonSearch was initially developed as a proof of concept in FileMaker Pro 7.0 (by PSC). All 789 taxon records in the initial compilation Stem Archosauria were entered (by PCS and SLB). The website was created (by SM in collaboration with PCS) using PHP4 with a MySQL database running on a Linux server in a shared hosting environment. Summary graphs of taxon records are generated using JGraph, a PHP graphing package available without fee for educational and non-commercial use.

User Interface

In the following paragraphs, *italics* is used for headings in the (1) navigational sidebar, (2) principal sections of the taxon record, and (3) names for individual fields. "Quotes" are used for user input. The field names in each taxon record are defined in Table 2. Background for the terms, structure of phylogenetic definitions, and history of phylogenetic taxonomy is given in Sereno (2005a). The initial compilation, Stem Archosauria (Figure 2), covers all supragenetic taxa excluding those within the archosaurian crown clades, Crocodylia and Neornithes, and those within Pterosauromorpha.

TaxonSearch contains a collection of taxon records (one for each suprageneric taxon) that are contained within an encompassing clade (i.e. named monophyletic group). As is the case with Stem Archosauria, the clade (Archosauria) must be monophyletic. All subclades within this clade, however, need

not be included. In this particular case, the compilation includes 789 taxon records, but excludes taxa within the crown clades, Crocodylia and Neornithes, and one diverse subclade, Pterosauromorpha. Taxa within these archosaurian subclades are available for compilation by other taxonomists.

Each taxon record has six color-differentiated sections (*Clade*, *Taxon*, *Active Phylogenetic Definition*, *Inactive Taxon Status*, *Current Age Range*, *Definitional History*), which organize the various fields. These sections and their respective fields are described below (Figures 3-7; see Table 2 for field definitions). The navigational sidebar on the left provides accessory information to the compilation of taxon records (*Tables*, *Graphics*, *Literature Cited*, *Glossary*). *Tables* contain tabular information related to the compilation; in the case of Stem Archosauria, an indented taxonomy linked to definitions summarizes all taxa with active phylogenetic definitions in the compilation. *Graphics* contain diagrammatic summaries related to the compilation; in the case of Stem Archosauria, labeled cladograms summarize the relative positioning of node and stem-based taxa in the compilation. *Literature Cited* contains full bibliographic references to citations in the compilation. The completeness of this bibliography is at the discretion on the compiler(s). A search function allows efficient location of a bibliographic reference by author, date or key words. Refinement or modification of sidebar supporting data (*Tables*, *Graphics*, *Literature Cited*) does not constitute revision of the compilation, which is reserved for changes in taxon records. The *Glossary* contains definitions for terms used in phylogenetic taxonomy and in taxon records. Shortened definitions appear for terms in taxon records when overlain by the cursor.

Clade Section

The encompassing clade is listed in the *Clade* field at the top of each record (Figure 3). The *Clade* section allows subdivision of taxon records into subclades via a pop-down menu. In the example compilation, Stem Archosauria, the encompassing taxon

View Details: [Prev Record](#) [Next Record](#)

Clade	
Subclade	Archosauria: Avemetatarsalia to Neornithes

Taxon Feedback	
Taxon	Abelisauridae
Nominal Author	Bonaparte and Novas 1985
2° Nominal Author	
Taxon Status	<input checked="" type="radio"/> active <input type="radio"/> inactive
Comments	Abelisauridae was first defined by Novas (1997) and Sereno (1998) as a node-based taxon. Wilson et al. (2003) presented a second-order revision of Abelisauridae as a stem-based taxon in opposition to the stem-based taxon Noasauridae, which came to be recognized as its sister taxon. The advantage of a stem-based definition of Abelisauridae is that it forms a stable node-stem triplet, identifying a major dichotomy in the history of the Abelisauoidea. As a node-based taxon, in contrast, Abelisauridae would have excluded the recently described basal member <i>Rugops</i> (Sereno et al. 2004). Taxonomists appear to have followed this arrangement, regarding <i>Rugops</i> as an abelisaurid. The active definition, thus, is a first-order revision of that stem-based definition, which uses species. The revised definition also cites for exclusion two additional species (<i>Coelophysis bauri</i> , <i>Passer domesticus</i>) that ensure stability of the taxonomic content of Abelisauridae, should abelisaurids be placed elsewhere among basal theropods. Padian (1999) and Tykoski and Rowe (2004) have followed the initial definitions of Novas (1997) and Sereno (1998). Tykoski and Rowe (2004), nonetheless, were unaware of the earlier second-order revision by Wilson et al. (2003).
Potential Synonymy	Abelisauria

Figure 3. The *Taxon* section of a taxon record (Abelisauridae) in the *TaxonSearch* compilation Stem Archosauria showing the nominal author, the status of the taxon, commentary on its historical usage and phylogenetic definition, and potential synonymy.

("Archosauria") is subdivided by menu into two subclades ("Archosauria: Crurotarsi to Crocodylia"; "Archosauria: Avemetatarsalia to Neornithes"). The *Clade* section, thus, allows the compiler to label taxon records by subclade. The user then is able to sort taxon records for a particular subclade. In Stem Archosauria, for example, is it possible to use the search function and the pop-down menu in the *Clade* section to determine that 202 and 582 taxon records are located within the clades Crurotarsi and Avemetatarsalia, respectively.

Taxon Section

The second section, *Taxon*, includes author and status information for a given taxon (Figure 3). A *Nominal Author* is the author

that coined a taxon. It lists the original author for the taxon or the first author of a taxon in the family series (superfamily, family, subfamily). A *2° Nominal Author* is an author that recognized a new rank within a family group that was originally named by a nominal author. Thus, although Linnean ranks are not used or recognized in *TaxonSearch*, nominal authorship for taxa in the family group (with suffixes -oidea, -dae, and -nae) follows the International Code of Zoological Nomenclature (ICZN, 1999:Art. 43), recognizing as nominal author the first author of a taxon at any rank within the family group.

TaxonSearch is designed to include all taxa, even those without phylogenetic definitions or those that have fallen from use. An *active*

taxon is one that has a complete phylogenetic definition and is in current use, or favored for such use, by the compiler. An *inactive* taxon, by contrast, is one that is rejected by the compiler. The terms *active* and *inactive*, thus, describe the status of the taxon, as interpreted by the compiler. If the taxon is regarded as *active*, any accessory comments about the taxon or a short history of its use may be input in the *Comments* field. The *Potential Synonymy* field contains taxa that may be synonymous with the taxon

under consideration. The purpose of this field is to alert the user to comparable taxa in the literature and in the compilation, not to evaluate their specific interrelationship (i.e. junior versus senior synonyms; subjective versus objective synonyms).

Further information regarding active taxa is input in the *Active Phylogenetic Definition* section, whereas further information regarding inactive taxa is input in the

Active Phylogenetic Definition		Feedback
Active Definition	The most inclusive clade containing <i>Carnotaurus sastrei</i> Bonaparte 1985 but not <i>Noasaurus leali</i> Bonaparte and Powell 1980, <i>Coelophysis bauri</i> (Cope 1889), <i>Passer domesticus</i> (Linnaeus 1758).	
Shorthand	> <i>Carnotaurus sastrei</i> but not <i>Noasaurus leali</i> , <i>Coelophysis bauri</i> , <i>Passer domesticus</i> .	
Definitional Author	Sereno 2005	
Definition Status	<input type="radio"/> original <input type="radio"/> textual substitution <input checked="" type="radio"/> first-order revision <input type="radio"/> second-order revision	
Definition Type	<input type="radio"/> node <input checked="" type="radio"/> stem	
Node-Stem Triplet	<input type="radio"/> yes <input type="radio"/> no	
Other Triplet Taxa		
Specifiers		
Specifier A	<i>Carnotaurus sastrei</i>	
Specifier(s) B		
Specifier(s) C	<i>Noasaurus leali</i> , <i>Coelophysis bauri</i> , <i>Passer domesticus</i>	
Qualifiers		
+Taxon		
-Taxon		
Datum		
Taxonomic Content	<input checked="" type="radio"/> stem (entirely extinct) <input type="radio"/> crown (extant-bounded) <input type="radio"/> mixed (extant/extinct-bounded)	
Publication Year	2003	

Figure 4. The *Active Phylogenetic Definition* section of a taxon record (Abelisauridae) in the *TaxonSearch* compilation Stem Archosauria showing the active definition (longhand and shorthand versions), the definitional author, the status of the active definition, the definitional type and possible role in a node-stem triplet, information on its specifiers and qualifiers (if any), the nature of its taxonomic content and the year the definitional type and intended clade were first published.

Inactive Taxon Status section as detailed below. User feedback regarding information displayed in the *Taxon* section may be reviewed or submitted for posting by email.

Active Phylogenetic Definition Section

The third section, *Active Phylogenetic Definition*, includes the *Active Definition*, its *Shorthand* version, the *Definitional Author*, and structural information about the definition—*Definition Status*, *Definition Type*, *Specifiers*, *Qualifiers*, the nature of its *Taxonomic Content*, etc.) (Figure 4). The *Active Definition* field shows the longhand phylogenetic definition that includes the full text of the relational phrase about inclusiveness and the authors of the species used as specifiers. The author and date listed after each specifier always refers to the original author of the species; if an author is shown in parentheses, the species named by that author was initially referred to a different genus than the one in the active definition. The *Shorthand* field shows the shorthand version of a phylogenetic definition, which includes symbols for clade inclusiveness, (<, >), “and” or “but not” as connecting words, and the species that act as specifiers. More information on symbolic, shorthand and longhand versions of phylogenetic definitions is available elsewhere (*Glossary* in the navigational sidebar; Sereno, 2005a).

Commas are omitted between author and date in phylogenetic definitions and elsewhere in the database. This style is much less confusing when composing phylogenetic definitions that often list several species and authors seriatim, each of which in turn must be separated by a comma. Omitting the comma between author and date differs from a recommendation in the International Code of Zoological Nomenclature (ICZN, 1999:Rec. 22A.2.1).

The *Definition Status* field records the historical origin of the *Active Definition*. It is *original* if it constitutes the first phylogenetic definition for a taxon. If there are prior definitions in the literature (tabulated

chronologically in the *Definitional History* section), then the active definition is either a textual variant (*textual substitution*) of an earlier definition or a *first-* or *second-order* revision (Table 2). Textual substitution involves equivalent re-wording of the relational statement of a phylogenetic definition, “Least inclusive clade” and “the common ancestor and all descendants,” for example, are equivalent relational phrases. Substituting one for the other does not constitute definitional revision, and so the definitional author listed would be the one associated with the first published definition. The specifiers in definitions that have undergone textual substitution, however, must be exactly the same as the original definition. In a *first-order revision*, the revisor intends to identify the same clade. The most common first-order revisions involve substituting alternative, but functionally comparable, specifiers, that identify the same clade. *Second-order revision* involves change of definitional type (node versus stem or “apomorphy-based” definition) or the selection of alternative specifiers. In second-order revision, the intention is to identify a different clade (see Sereno, 2005a).

Definition Type is either *node* or *stem*, because all complete phylogenetic definitions are either node- or stem-based (Sereno, 2005a). “Apomorphy-based” definitions, for example, are shown elsewhere to constitute stem-based definitions that use one internal specifier and a form qualifier (Sereno, 2005a). The database supports form-qualified definitions. Form-qualified definitions, nevertheless, are not recommended because they are demonstrably less stable and often difficult to interpret. Taxonomic qualifiers (species) are entered into + *Taxon* and - *Taxon* fields, which designate mandatory inclusion or exclusion, respectively, for the taxon name to apply. The *Datum* field is for temporal qualifiers in phylogenetic definitions, such as “and any extant taxa.” A datum qualifier (“and any extant species”), for example, may be used in a stem-based phylogenetic definition for a crown clade.

Finally, the fields *Taxonomic Content* and *Publication Year* record basic information

Table 2. Glossary of terms used in phylogenetic taxonomy (Sereno, 2005a, b) and *TaxonSearch* records.

Term	Definition
<i>active taxon</i>	Status of a taxon that has a complete phylogenetic definition and is actively employed by the compiler
<i>active definition</i>	Current compiler-accepted phylogenetic definition for a taxon
<i>clade</i>	monophyletic group
<i>compiler</i>	Person creating or authorizing taxon records for a <i>TaxonSearch</i> compilation; an <i>original compiler</i> is the first person to create a compilation; a <i>secondary compiler</i> is a person who gains access to, and modifies, a compilation originally created by someone else
<i>crown clade</i>	Clade defined by extant specifiers and whose immediate outgroup is extinct
<i>datum</i>	Time horizon used as a qualifier in a phylogenetic definition
<i>definitional author</i>	Author of a phylogenetic definition
<i>definitional component</i>	A functional unit of a phylogenetic definition, which includes paradigms, specifiers, and qualifiers
<i>definitional revision</i>	Alteration of one or more components of a previous definition that results in either <i>first-</i> or <i>second-order revision</i>
<i>definitional type</i>	The fundamental kind of definition (node-based, stem-based), which are distinguished by the polarity of inclusion (least, most inclusive) and distinctive symbolic notation [T = (< A and BQ); T = (> AQ but not C)] Symbols: T = taxon; () = phylogenetic definition; <, > = polarity of inclusion; A-C = specifiers; Q = qualifier
<i>earliest record</i>	Earliest age (Ma = million years ago) assigned to any included species
<i>first-order revision</i>	Definitional revision that does not alter the definitional type or intended clade
<i>inactive taxon</i>	Status of a taxon that has no phylogenetic definition or is not currently in use by the compiler
<i>latest record</i>	Latest age (Ma) assigned to any includes species, or "Recent" (= 0) if there are extant members
<i>mixed clade</i>	Clade with living and extinct members that can only be defined by using at least one extinct specifier (because of basal extinct members)
<i>monophyletic</i>	A group composed if a real or hypothetical common ancestor and all of its descendants
<i>node-based</i>	Statement specifying the membership of a taxon as the least inclusive clade that contains at least two internal specifiers
<i>node-stem triplet</i>	Trio of taxa with complementary definitions consisting of a node-based taxon and two subordinate stem-based taxa
<i>nominal author</i>	Author who first coined a taxon name
<i>2° nominal author</i>	Author who coined a new taxon name within a family group that was already named by an earlier author
<i>paradigm</i>	Fundamental manner in which a phylogenetic definition circumscribes a taxon; expressed in terms of the polarity of inclusion
<i>phylogenetic definition</i>	Statement specifying the membership of a taxon
<i>potential synonymy</i>	Comparable taxa that may be regarded as either junior or senior synonyms by some taxonomists

<i>publication year (field in Active Phylogenetic Definition)</i>	Year that a phylogenetic definition for an active taxon was first published which has the same definitional type and which identifies the same clade; year that the fundamental definitional concept was first published
<i>qualifier</i>	Species or an attribute of a species or specimen cited in a phylogenetic definition as a dependent clause that qualifies clade membership
<i>range</i>	Duration of a taxon in millions of years (My) between the <i>Earliest</i> and <i>Latest Records</i> . "Recent" is the <i>Latest Record</i> for a taxon with extant members
<i>second-order revision</i>	Definitional revision that intentionally identifies a different clade by using a different definitional type (node- vs stem-based) and/or by using alternative specifiers or qualifiers to intentionally identify a different clade
<i>shorthand</i>	Shortened version of a phylogenetic definition that retains specifiers but eliminates authors, shortens qualifiers, and substitutes symbols for functions (e.g. < = "the least inclusive clade containing")
<i>specification</i>	Delimitation of clade membership by a phylogenetic definition
<i>specifier</i>	Species cited in a phylogenetic definition as an independent phylogenetic reference point
<i>specifier A</i>	Mandatory internal specifier
<i>specifier(s) B</i>	Optional internal specifier(s)
<i>specifier(s) C</i>	Optional external specifier(s)
<i>stem-based</i>	Statement specifying the membership of a taxon as the most inclusive clade that contains at least one internal specifier
<i>stem clade</i>	Clade composed solely of extinct members whose immediate outgroup contains at least one living member
<i>symbolic notation</i>	Symbolic representation of a phylogenetic definition (e.g. T= (< A and BQ); general symbolic notation for a stem-based definition) Symbols: T = taxon; () = phylogenetic definition; < = "The least inclusive clade containing"; A = mandatory internal specifier; C = optional external specifier(s); Q = qualifier
<i>taxon</i>	Named clade (= monophyletic group)
<i>taxon qualifier</i>	Qualifier in a phylogenetic definition composed of species cited for inclusion (positive or "+ taxon") or exclusion (negative or "- taxon"); can be used to limit the potential taxonomic content of a taxon
<i>taxonomic content</i>	Existing or potentially existing taxa or specimens that by definition are included within a taxon
<i>textual substitution</i>	Comparable phraseology between phylogenetic definitions that does not involve changes in a definitional component (paradigm, specifier, qualifier) and, as a result, does not constitute definitional revision

about the active phylogenetic definition. If all members are extinct, it is considered a *stem* taxon; if the taxon is defined using extant specifiers and if the immediate outgroup is extinct, it is considered a *crown* taxon, following definitions in Sereno (1999, 2005a); if the taxon has at least one extant, and one extinct, specifier, it is considered a *mixed* taxon. The *Publication Year* records

the earliest published phylogenetic definition for a particular taxon that uses the same definitional type and identifies the same clade as the active definition. This date, thus, marks the first use of a definition that actually, or intentionally, identified the same clade as the active phylogenetic definition. It may be the same or earlier than the date of the active phylogenetic definition (date in

Inactive Taxon Status Feedback	
Rejection Criteria	<input checked="" type="checkbox"/> junior synonym <input type="checkbox"/> historically polyphyletic <input type="checkbox"/> incomplete definition <input checked="" type="checkbox"/> no definition <input type="checkbox"/> apomorphy-based <input type="checkbox"/> redundant <input type="checkbox"/> other
Critique	This taxon has not been in wide use during the twentieth century despite an attempt by Bakker (1986) to revive it as a replacement for <i>Ornithischia</i> Seeley 1888. No phylogenetic definition has been proposed.

Figure 5. The *Inactive Taxon Status* section of a taxon record (Predentata) in the *TaxonSearch* compilation Stem Archosauria showing the rejection criteria that led to its inactive status and an explanatory critique.

the *Definitional Author* field), which may have undergone subsequent first-order revision.

***Inactive Taxon Status* Section**

The fourth section, *Inactive Taxon Status*, provides the opportunity for the compiler to explain why a particular taxon ought to be regarded as *inactive* (Figure 5). There are only a handful of reasons to reject use of a suprageneric taxon, and these are listed under *Rejection Criteria*. The taxon may constitute a *junior synonym* or may be *redundant* with another taxon; the taxon may lack a phylogenetic definition (*no definition*); proposed definitions for the taxon may be incomplete (*incomplete definition*); or there may be other reasons (*other*). Explanatory notes are entered in the adjacent field *Critique* field. under *Rejection Criteria*. The taxon may constitute a *junior synonym* or may be *redundant* with another taxon; the taxon may lack a phylogenetic definition (*no definition*); proposed definitions for the taxon may be incomplete (*incomplete definition*); or there may be other reasons (*other*). Explanatory notes are entered in the adjacent field *Critique* field.

***Current Age Range* Section**

The fifth section, *Current Age Range*, lists the *Earliest* and *Latest Records* for an active taxon and its *Range* in millions of years (Figure 6). The earliest record for a taxon (either node- or stem-based) is the oldest age for any included species. Stem-based taxa are treated the same, even though their sister taxa might suggest an older divergence date. Thus, a stem-based taxon like Coelurosauria would have as its earliest record the oldest coelurosaurian (tentatively, Middle Jurassic *Proceratosaurus*); it would not be based on the age of its sister taxon (Early Jurassic *Cryolophosaurus*). The *Range* duration, thus, is not affected by variation in the most likely sister taxon.

The numerical ages and stage assignments used in the *Current Age Range* section are from Gradstein et al. (2004), rounded to the nearest million. The *Earliest Record* field, for example, lists the stage assigned to the earliest fossil, using the lower time boundary for that stage. The latest record may be "Recent" if there are extant members, in which case the age assignment is zero. The *Range* in millions of years equals the age assigned to the *Earliest Record* minus that given for the *Latest Record*. If the *Earliest* or *Latest Record* is uncertain and more than a single stage is given, the midpoint of the stage range can be used. Literature citations, evidence used in establishing age assignments, or the use of mid points for

Current Age Range		Feedback
Earliest Record	Aptian/Albian (112 Ma)	
Latest Record	Maastrichtian (66 Ma)	
Range (My)	46	
Basis for Range	The earliest record comes from the mid Cretaceous (Aptian/Albian) of Africa, as yet unnamed abelisaurid and noasaurid from Niger (Sereno et al. 2004). The latest record comes from the Maastrichtian of Madagascar and India (Sampson et al. 1998; Wilson et al. 2003). Dates are rounded to the nearest million, the earliest from the midpoint of the two stages indicated and the latest from the end of the stage (Gradstein et al. 2004).	

Figure 6. The *Current Age Range* section of a taxon record (Abelisauridae) in the *TaxonSearch* compilation Stem Archosauria showing the earliest and latest records, the estimated temporal range, and the evidence supporting the temporal data.

multiple stage ranges are given in the *Basis for Range* field. In Stem Archosauria, the temporal range is based on the timescale in Gradstein et al. (2004) and, like other aspects of the taxon record, is not open for revision except when posting a new version of the compilation. Thus, the dates are tied to a particular cited timescale and cannot be globally adjusted by the user, as is the case with some other databases (e.g. *The Paleobiology Database*).

More elaborate temporal and biogeographic data were not included for two reasons. First, doing so would have introduced substantial ambiguity and error, given different interpretations of the content of a taxon. It must be kept in mind that a phylogenetic definition may circumscribe somewhat different taxonomic content when applied to an alternative phylogenetic hypothesis. Second, there already exists a database devoted to temporal and biogeographic compilation (*The Paleobiology Database*; Table 1). *TaxonSearch* may well link taxonomic records to such data in the future, which will be determined by user feedback and need.

Definitional History Section

The sixth and final section, *Definitional History*, lists the *First Definition*, its author (*Definition 1 Author*), its definitional type (*Definition 1 Type*), its specifiers (*Definition*

1 Specifiers), and qualifiers (*Definition 1 Qualifiers*) should there be any (Figure 7). This section records the historical sequence of phylogenetic definitions that have been proposed for a taxon. Definitions that differ only by textual substitution are not considered revisions (Sereno, 2005a) and so are not listed as distinct definitions in the definitional history section. Most active taxa have one or two phylogenetic definitions. In the Stem Archosauria compilation, Maniraptora and Aves have the greatest number of distinct phylogenetic definitions (seven and eight, respectively). These fall into two (second-order) groups that differ in the preferred taxonomic content of the taxon (one a mixed more inclusive clade anchored on *Archaeopteryx lithographica*; the other a crown clade bounded by extant species).

Using TaxonSearch

Example Compilation: Stem Archosauria

Stem Archosauria 1.0 (Sereno 2005b) is the initial version of a compilation containing 789 taxonomic records that document all suprageneric taxa and phylogenetic definitions that have ever been used in the taxonomy of stem archosaurs—i. e., all archosaurs excluding those within crown crocodylians (Crocodylia), crown birds (Neornithes), and Pterosauroomorpha.

Definitional History #1	
Definition 1	Previously listed taxa [<i>Abelisaurus comahuensis</i> , <i>Carnotaurus sastrei</i> , <i>Xenotarsosaurus bonapartei</i> , <i>Indosaurus matleyi</i> , <i>Indosuchus raptorius</i> , <i>Majungasaurus crenatissimus</i>] and all descendants of their common ancestor.
Author	Novas 1997:1
Type	<input checked="" type="radio"/> node <input type="radio"/> stem
Specifiers	<i>Abelisaurus comahuensis</i> , <i>Carnotaurus sastrei</i> , <i>Xenotarsosaurus bonapartei</i> , <i>Indosaurus matleyi</i> , <i>Indosuchus raptorius</i> , <i>Majungasaurus crenatissimus</i>
Qualifiers	

Definitional History #2	
Definition 2	<i>Abelisaurus</i> , <i>Carnotaurus</i> , their most recent common ancestor and all descendants.
Author	Sereno 1998:64
Type	<input checked="" type="radio"/> node <input type="radio"/> stem
Specifiers	<i>Abelisaurus</i> , <i>Carnotaurus</i>
Qualifiers	

Definitional History #3	
Definition 3	<i>Carnotaurus</i> , <i>Abelisaurus</i> , and all descendants of their most recent common ancestor.
Author	Padian et al. 1999:70
Type	<input checked="" type="radio"/> node <input type="radio"/> stem
Specifiers	<i>Carnotaurus</i> , <i>Abelisaurus</i>
Qualifiers	

Definitional History #4	
Definition 4	The most inclusive clade containing <i>Carnotaurus sastrei</i> but not <i>Noasaurus leali</i> .
Author	Wilson et al. 2003:30
Type	<input type="radio"/> node <input checked="" type="radio"/> stem
Specifiers	<i>Carnotaurus sastrei</i> , <i>Noasaurus leali</i>
Qualifiers	

Definitional History #5	
Definition 5	<i>Abelisaurus comahuensis</i> and <i>Carnotaurus sastrei</i> and all descendants of their most recent common ancestor.
Author	Tykoski and Rowe 2004:64
Type	<input checked="" type="radio"/> node <input type="radio"/> stem
Specifiers	<i>Abelisaurus comahuensis</i> , <i>Carnotaurus sastrei</i>
Qualifiers	

Figure 7. The *Definitional History* section of a taxon record (Abelisauridae) in the *TaxonSearch* compilation Stem Archosauria showing in chronological order the five available phylogenetic definitions that have been proposed for the taxon Abelisauridae.

These three subclades were excluded because they are currently under consideration by other phylogenetic taxonomists and are beyond the expertise of the compiler.

Searching and Sorting

Users log-on to *TaxonSearch* (www.taxonsearch.org), enter the database, and open a particular posted compilation. Users can search for a particular taxon record, sort for a particular set of records, open related information (indented taxonomy; labeled cladograms), or locate bibliographic references. Practically an infinite number of queries can be answered instantaneously about authors, taxa, dates, definitions, and literature citations using *Search* on the navigational sidebar. Besides the description below, further examples are given on the home page of *TaxonSearch* (in *Using TaxonSearch*).

Searching

Users can search for a particular taxon record by entering part or all of the taxon name in the *Taxon* field. Entering "Pred", for example, locates the taxon record for *Predentata*, a taxon coined by Marsh (1894) that has fallen from use and is generally regarded as equivalent to *Ornithischia*. The user learns from the *Comments* field in the *Taxon* section that "This taxon has not been widely used during the twentieth century despite an attempt by Bakker (1986) to revive it as a replacement for *Ornithischia* Seeley 1888" (Figure 5). As the record is scanned, the user learns that Hennig (1915) introduced a spelling variant (*Praedentata*), that no phylogenetic definition has ever been proposed, and that the compiler (P. Sereno) regards the taxon as a potential synonym of *Ornithischia*. The taxon is considered *inactive*.

Sorting

If the user remembered the nominal author (O. C. Marsh), and was wondering how many taxa were coined by this author among stem archosaurs, entering "Marsh" in the *Nominal Author* field sorts 67 taxon

records. If the user wanted to determine how Marsh's taxa have fared in the current literature, selecting in addition the *active* radio button in the *Taxon Status* field reduces sorted taxon records to 22—meaning that 22 suprageneric taxa coined by O. C. Marsh are regarded as active by the compiler, whereas 45 have fallen from use for various reasons. It is easy to check these taxon records using *previous record* or *next record* controls at the top of the page. If the user wanted to determine if any of Marsh's taxa apply to crurotarsal archosaurs rather than dinosaurs and closest of kin, "Archosauria: Crurotarsi to Crocodylia" can be selected from the *Subclade* menu and "Marsh" entered into the *Nominal Author* field. O. C. Marsh proposed three suprageneric taxa among crurotarsal archosaurs (*Belodontia*, *Hallopoda*, *Hallopodidae*), none of which are in current use or regarded as active by the compiler.

Output

In addition to displaying individual taxon records, *TaxonSearch* provides two graphical summaries of taxon records in a compilation when *Summarize* is selected in the navigational sidebar (Figure 8). The first, a pie chart, summarizes the percentages of active versus inactive taxa and basic information about the structure of active phylogenetic definitions. The second, a graph, summarizes the temporal accumulation of active taxa and their phylogenetic definitions. This graphical output can be calculated for all, or a subset, of taxon records in the compilation.

Taxon/Definitional Status

The *Taxon* section includes the status of a taxon, i.e. whether it is in current use (*active*) or rejected (*inactive*) by the compiler. The quantity of each of condition, as a percentage of total records under consideration, is given at the core of the pie chart (Figure 8A.). The outer ring of the pie chart provides additional information about each basic subdivision of taxon records.

Taxa with *active* status are subdivided based on the *Definition Type* field in the

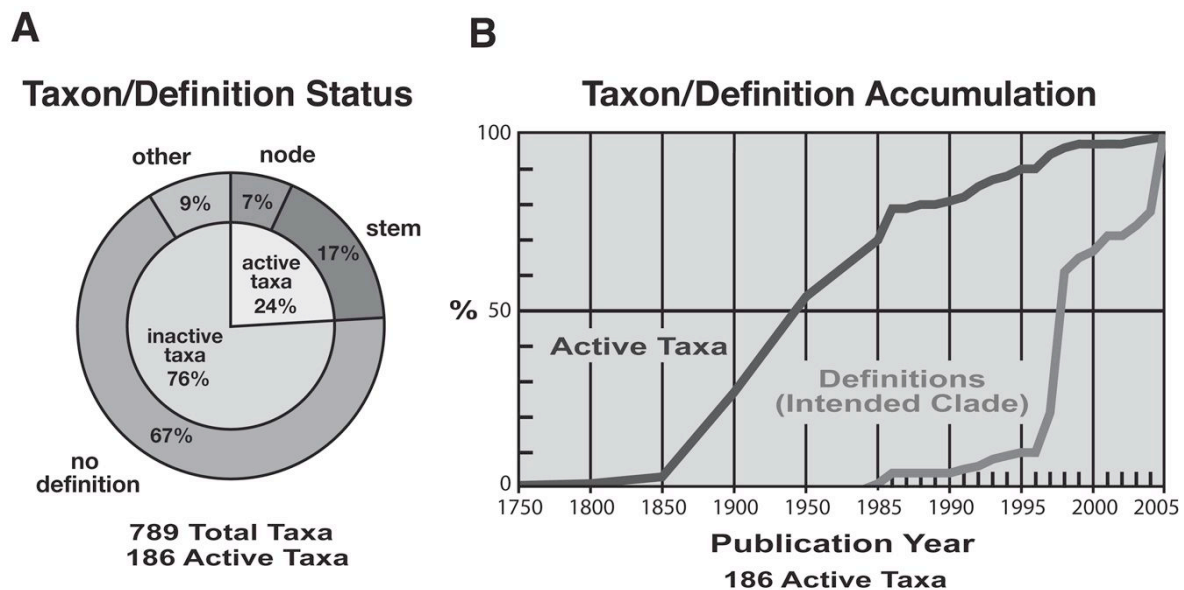


Figure 8. *TaxonSearch* output. **A**, Pie chart showing the status of taxa and their respective definitions (“active taxa” denotes those in current use by the compiler; “inactive taxa” includes those regarded as potential synonyms or rejected for other reasons by the compiler). **B**, Accumulation profile for the publication date for active taxa and the publication date for its intended phylogenetic definition (after Sereno, 2005a).

Active Phylogenetic Definition section. The percentage of active taxa with each kind of phylogenetic definition (node, stem) is shown in the outer ring. *Inactive* status is subdivided into two categories (no definition, other). The reasons for inactive status for any particular taxon are given in the *Rejection Criteria* field in the *Inactive Taxon Status* section. The percentage of inactive taxa with *no definition* is calculated, with the remainder binned as “other.” All percentages are calculated against the total number of taxon records considered (the number below the pie chart).

Taxon/Definitional Accumulation

This graph shows the cumulative percentage by year of publication of active taxa and their intended phylogenetic definitions (Figure 8B.). *TaxonSearch* creates the graph by binning the dates from the *Nominal Author* field in the *Taxon* section into the time units described below, and compiles a cumulative authorship curve for active taxa. *TaxonSearch* compiles a similar cumulative authorship curve for the *Publication Year* field in the *Taxon* section, which is the date that the definitional type

and intended clade of the active definition was first published.

The earliest taxon names, such as Crocodylia or Aves, were published in the mid 1700’s, and so the first 200 years of the graph has dates grouped in 50-year intervals, starting in 1750 up to 1950. Then there are smaller intervals between 1950-1985. Because the first phylogenetic definitions appeared in 1986, the final interval (1985-2005) is plotted per year. This gives the user a general idea when taxon names were coined and when their respective phylogenetic definitions (i.e. the first version that identified the intended clade) were first published.

The graph for Stem Archosauria suggests that the number of new suprageneric taxa is beginning to tail off in the last decade after a steady increase over the last 150 years. Perhaps the main outlines of archosaur phylogeny are now known and named. The number of taxa with phylogenetic definitions has risen sharply following publication of the first general protocol for phylogenetic definitions (de Queiroz and Gauthier, 1990).

Compiling

Users can become compilers and work on-line to create new *TaxonSearch* compilations, adding, saving and editing taxon records for clades of interest. *TaxonSearch* compilations that are posted (and thus linked to a publication) are archival and cannot be altered by users, compilers or site managers. New versions of the compilation, nevertheless, may be posted at any time. The latest version of the compilation is by default the one that is accessed by users, although earlier versions may be opened as well.

Management

Users can access but cannot edit taxon records in any version of any compilation that is posted in the archive of *TaxonSearch*. Users may submit by email general comments ("Contact Us") or feedback for posting to any taxon record in any compilation; comments will receive minimal screening for suitability before posting, and the posting will include the author, date, and subject.

As described above (*TaxonSearch*: Conceptual Overview), compilers have no more than five-years after posting to substantively update their compilations, or they will become available for updating and revision by secondary compilers. Prior to posting, newly created compilations will be checked for completeness and non-overlap with compilations already in the database. Much of this will be automated, but minimal site management will be maintained. Should *TaxonSearch* expand and achieve broad use, management and site support will be turned over to an elected managerial board.

Conclusions

Taxonomists in general and phylogenetic taxonomists in particular are in need of a web-accessible tool for organizing and searching data on suprageneric taxa. The utility and success of *TaxonSearch* will be measured in the coming years by the number and frequency of users, number of

posted compilations, the amount of user feedback, and the degree to which it helps stimulate consensus among phylogenetic taxonomists.

Acknowledgements

The authors thank C. Abraczinskas for preparation of the final figures and J. Wilson for comments on an earlier draft. This work was supported by the National Science Foundation (to PCS).

References

- Bakker, R. T.** 1986. *The Dinosaur Heresies*. Bath Press, Avon. 481 pp.
- Berendsohn, W. G.** 1995. The concept of "potential taxa" in databases. *Taxon* **44**:207-212.
- Berendsohn, W. G.** 1997. A taxonomic information model for botanical databases: the IOPI Model *Taxon* **46**:283-309.
- Bisby, F.** 2000. The quiet revolution: biodiversity informatics and the internet. *Science* **289**:2309-2312.
- Cantino, P. D., and K. de Queiroz** 2004. PhyloCode: A phylogenetic code of biological nomenclature, Version 2b (June 17, 2004) <http://www.ohio.edu/phylocode/>.
- Darwin, C.** 1859. *The Origins of Species by Natural Selection*. John Murray, London. 1000 pp.
- de Queiroz, K., and J. Gauthier** 1990. Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. *Systemic Zoology* **39**:307-322.
- de Queiroz, K., and J. Gauthier** 1992. Phylogenetic taxonomy. *Annual Review of Ecology and Systematics* **23**:449-480.
- Ereshefsky, M.** 2001. *The Poverty of the Linnean Hierarchy*. Cambridge University Press, Cambridge. 316 pp.
- Felsenstein, J.** 2004. *Inferring Phylogenies*. Sinauer Associates, Inc., Sunderland. 664 pp.
- Fortey, P. L.** 2002. PhyloCode—pain, no gain. *Taxon* **51**:43-54.
- Gewin, V.** 2002. Taxonomy: All living things, online. *Nature* **418**:362-363.
- Godfray, H. C. J.** 2002. Challenges for taxonomy. *Nature* **417**:17-19.
- Gradstein, F. M., Ogg, J. G., and Smith A. G.** (eds.) 2004. *A Geologic Time Scale 2004*. Cambridge University Press, Cambridge. 589 pp.
- Greuter, W., J. Msneill, R. Barrie, H.-M. Burdet, V. Demoulin, T. S. Filguerias, D. H. Nicolson, P. C. Silva, J. E. Skog, P. Threhane, N. J. Turland and D. L. Hawksworth** (eds. and compilers). 2000. *International Code of Botanical Nomenclature (Saint Louis Code) adopted by the 16th International Botanical Congress, St. Louis, Missouri, July-August 1999*. Regnum Vegetabile 138. Koelz Scientific Books, Königstein. 474 pp.
- Hennig, E.** 1915. *Kentrurosaurus aethiopicus*, der Stegosauridae des Tendaguru. *Sitzungsberichte der Dorpater Naturforscher-Gesellschaft* **1915**:219-147.
- Hillis, D. M., C. Moritz, and B. K. Mable** (eds.) 1996. *Molecular Systematics*. 2nd ed. Sinauer Associates, Inc., Sunderland. 655 pp.
- ICZN [International Commission on Zoological Nomenclature]** (ed.) 1999. *International Code of Zoological Nomenclature*. The International Trust for Zoological Nomenclature, London. 306 pp.
- Kluge, A.** 2005. Taxonomy in theory and practice, with arguments for a new phylogenetic system of taxonomy. Pages 7-52 in *Ecology and Evolution in the Tropics: A Herpetological Perspective* (M. A. Donnelly, B. I. Crother, C. Guyer, M. H. Wake, and M. E. White, eds.), University of Chicago Press, Chicago.
- Mallet, J., and K. Willmott** 2002. Taxonomy: renaissance or Tower of Babel? *Trends in Ecology and Evolution* **18**:57-59.
- Marsh, O. C.** 1894. The typical Ornithopoda of the American Jurassic. *American Journal of Science* **48**:85-90.
- Mayr, E., E. G. Linsley, and R. L. Usinger** 1953. *Methods and Principles of Systematic Zoology*. McGraw-Hill Book Co., Inc., New York. 328 pp.
- Minelli, A.** 2003. The status of taxonomic literature. *Trends in Ecology and Evolution* **18**:75-76.
- Pyle, R. L.** 2004. *Taxonmer*: a relational data model for managing information relevant to taxonomic research. *PhyloInformatics* **1**:1-54.
- Sereno, P. C.** 1999. Definitions in phylogenetic taxonomy: Critique and rationale. *Systematic Biology* **48**:329-351.
- Sereno, P. C.** 2005a. The logical basis of phylogenetic taxonomy. *Systematic Biology* **51**:1-25.
- Sereno, P. C.** 2005b. Stem Archosauria—*TaxonSearch*. URL <http://www.taxonsearch.org/Archive/stem-archosauria-1.0.php> [version 1.0, 2005 November 7]

- Swofford, D. L., G. J. Olsen, P. J. Waddell, and D. M. Hillis** 1996. Phylogenetic inference. Pages 407-514 in *Molecular Systematics* (D. M. Hillis, C. Moritz, and B. K. Mable, eds.). Sinauer Associates, Inc., Sunderland.
- Weishampel, D. B., P. Dodson, and H. Osmólska** (eds) 2004. *The Dinosauria*. University of California Press, Berkeley. 861 pp.
- Ytow, N., D. R. Morse, and D. M. Roberts** 2001. Nomenclator: a nomenclatural history model to handle multiple taxonomic views. *Biological Journal of the Linnean Society* **73**:81-98.