

## THE EMERGING STATISTICAL PERSPECTIVE IN SYSTEMATICS: A COMMENT ON MARES AND BRAUN

SCOTT J. STEPPAN\* AND JACK SULLIVAN

*Smithsonian Institution, Laboratory of Molecular Systematics, 4210 Silver Hill Road,  
Suitland, MD 20746 (SJS)*

*Department of Biological Sciences, University of Idaho, Moscow, ID 83844 (JS)*  
*Present address of SJS: Department of Biological Science, Florida State University,  
Tallahassee, FL 32306-1100*

Recent advances in systematics have resulted in the development of statistical approaches to phylogeny estimation and hypothesis testing. We present this emerging statistical perspective, distinguish it from the nonstatistical approach as commonly practiced, and demonstrate its application to 2 systematic proposals from the recent literature on rodents: Mares and Braun's (1996) retention of *Andalgalomys* contra Steppan (1995) and the conclusion of Engel et al. (1998) that *Calomys* is a multitribal composite taxon. Systematic arguments and recommendations by Mares and Braun are critiqued in detail. We argue that the statistical approach provides a more objective and powerful basis for evaluating alternative phylogenetic hypotheses and a more stable foundation for nomenclatural decisions.

**Key words:** *Calomys*, consistency index, hypothesis testing, nomenclature, phyllotines, statistics, systematics

---

Systematics has been one of the most vital and dynamic fields in biology over the past several decades. Rapid advances in theory, data collection, software, and computers have resulted in important changes in how systematic studies are conducted, how evidence is interpreted, and the nature of how hypotheses are presented and tested. Coupled with these advances has been a reawakening among biologists to the central role that phylogeny plays in evolutionary studies. This welcome maturation in the scientific character of the field was made possible by the long history of careful morphological studies at all levels. A major force in this progression is the emergence of a statistical perspective in phylogenetic analysis (e.g., Huelsenbeck and Rannala 1997). The statistical approach includes subtle but fundamental changes in the nature of systematic questions and emphasizes

rigorous hypothesis testing rather than data exploration and explanation. Given these recent developments, we believe it is timely to discuss some consequences of the statistical approach and to apply it to recent examples from the mammalian literature. Here, we have applied a statistical approach to the systematics of the phyllotine rodents *Graomys*, *Andalgalomys*, and *Calomys*; these analyses illustrate the importance of statistical testing of competing hypotheses.

This paper was prompted by a recent systematic study of a new species of *Graomys* (*Andalgalomys*) by Mares and Braun (1996). In their discussion, they argued for retaining *Andalgalomys* as a genus distinct from *Graomys*, contrary to the synonymy proposed by Steppan (1995), who provided the most recent revision addressing *Andalgalomys*. They presented a series of criticisms of the phylogenetic analysis for the tribe Phyllotini on which Steppan (1995) based his decision to synonymize *Andal-*

\* Correspondent: [steppan@bio.fsu.edu](mailto:steppan@bio.fsu.edu)

*galomys* with *Graomys*. These criticisms questioned the reliability of Steppan's analysis as a whole and were followed by recommendations to systematists on how to interpret aspects of phylogenetic analyses and how classificatory decisions should be made. We argue that most of the criticisms are invalid and that the recommendations are either erroneous or outdated. Because some of Mares and Braun's (1996) arguments commonly are found in the mammalian systematic literature, we believe that a response is of general significance.

Herein, we address the following issues. First, we describe the emerging statistical perspective in phylogenetics and how it differs from the manner in which systematic inquiry has been conducted traditionally. Second, we summarize briefly the taxonomic history of *Andalgalomys* and the origin of this specific nomenclatural debate. Third, we discuss general issues behind each of Mares and Braun's (1996) criticisms. Fourth, we demonstrate the statistical approach by analyzing 2 data sets. The 1st is Steppan's (1995) original data set modified according to Mares and Braun's (1996) recommendations and further enlarged by the inclusion of 4 more taxa (2 of which were described subsequent to Steppan's 1995 publication) that may be members of the *Andalgalomys*-*Graomys* clade. The 2nd is a data set from a recent molecular study (Engel et al. 1998) involving another phylotine, *Calomys*. Finally, we summarize the relevant evidence for and against synonymizing *Andalgalomys* and present our recommendations. We emphasize here that the important issue is the practice of systematics and the validity of recommendations made to systematists, not the specific issue of the nomenclatural ranking of *Andalgalomys*.

#### STATISTICAL APPROACH TO PHYLOGENETIC INFERENCE

Recent history of phylogenetic inference has focused primarily on optimal trees; that is, phylogenetic analyses typically have in-

volved searching for the best tree under a certain optimality criterion. Parsimony is the classic optimality criterion in cladistics, and the tree(s) that requires the fewest number of evolutionary changes is the preferred estimate of phylogeny (or the tree that maximally explains the data). In reality, however, the shortest tree(s) represents a point estimate and, from a statistical perspective, is an inadequate representation of the evolution of a group in the absence of an assessment of reliability of that estimate. Inadequacy of optimal trees (as point estimates) has led to the statistical approach to phylogeny reconstruction and also has motivated development of such commonly used measures of phylogenetic reliability as the nonparametric bootstrap (Felsenstein 1985) and the nonstatistical decay (or Bremer support) index (Bremer 1994). Use of these measures of nodal reliability is a great advance over treatment of optimal topologies as a single hypothesis because a phylogenetic tree actually represents a series of nested hypotheses; these measures of nodal support represent estimates of how strongly data support each of these hypotheses.

However, from a hypothesis-testing perspective, these approaches have the weakness that the series of nested hypotheses are nonindependent, and it may not be appropriate to test them simultaneously. Assigning statistical probabilities to bootstrap or decay values ranges from problematic to impossible. Nonetheless, they minimally provide a qualitative estimate of nodal robustness, which is a decided improvement over no estimate at all.

Another, more powerful, approach involves estimating how strongly a particular data set refutes a particular hypothesis. Preferably, these hypotheses are proposed before analysis of the data. A number of methods have been developed that test significance of the difference between scores of 2 trees in either a parsimony or likelihood framework; these tests include Templeton's tests (winning-sites test and Wilcoxon signed-ranks test—Larson 1991;

Templeton 1987), the Kishino–Hasegawa test (Kishino and Hasegawa 1989), and the parametric bootstrap (Huelsenbeck et al. 1996).

Such tools raise phylogenetic analysis out of the realm of simply explaining data and into the more rigorous realm of statistical hypothesis testing. A particular taxonomic or other evolutionary hypothesis predicts a certain phylogenetic relationship or set of relationships. An investigator then can collect data to test explicitly the hypothesis of interest by finding the best (least suboptimal) tree that is consistent with that hypothesis; that is, the best tree constrained to fit relationships predicted by the hypothesis under examination. The difference in optimality score between the tree constrained to fit the hypothesis and the unconstrained (optimal) tree (i.e., difference in number of steps or difference in log-likelihood score) becomes the test statistic. The significance of the magnitude of the test statistic can be assessed according to several methods. The relevant issue is whether or not the constrained tree (consistent with the hypothesis under examination) is significantly worse than the unconstrained tree (optimal tree). If so, the data at hand refute that hypothesis. Thus, statistical phylogenetics is concerned with the ability of a particular data set to refute a particular hypothesis at a given significance level rather than with the exact topology of the optimum tree under the chosen criterion. Because this is still a young field and application of statistical inference to historical events is a difficult undertaking, the tests presented here are likely to improve as the field develops.

#### TAXONOMIC HISTORY OF *ANDALGALOMYS*

*Graomys pearsoni* was described by Myers (1977). Soon afterward, Williams and Mares (1978) described the new species *olorogi*, created the genus *Andalgalomys* for it, and reassigned *pearsoni* from *Graomys* to *Andalgalomys*. They concluded that the new genus was most closely related to *Cal-*

*omys* and *Eligmodontia*. In describing a new subspecies of *A. pearsoni*, Olds et al. (1987) revised the diagnosis of *Andalgalomys* and compared it with other phyllotine genera. They concluded that it was phenetically more similar to *Graomys* than to *Calomys* or *Eligmodontia*. Braun (1993) conducted cladistic analyses and found that *Graomys* was paraphyletic with respect to *Andalgalomys*; *G. griseoflavus*, *G. domorum*, and *Andalgalomys* were in an unresolved polytomy, with the following topology: (other phyllotines (*G. edithae* (*G. domorum*, *G. griseoflavus*, (*A. pearsoni*, *A. olorogi*))). Thus, Braun's (1993:17, 19, figures 8 and 9) cladistic results supported conclusions of Olds et al. (1987) and further suggested that *Andalgalomys* should be synonymized within *Graomys* (although Braun did not discuss these nomenclatural consequences for the 2 taxa). Steppan (1993, 1995) conducted phylogenetic analyses with a somewhat larger data set and effectively confirmed Braun's cladogram regarding *Andalgalomys*. In Steppan's (1995) analyses, *G. griseoflavus* was more closely related to *A. pearsoni* than to *G. domorum*; support for that relationship was considered strong enough (83% bootstrap score, decay index of 4) to justify subsuming *Andalgalomys* within *Graomys*. Subsequently, Mares and Braun (1996) described a new species of *Andalgalomys* and argued for retention of that genus, contrary to Steppan's (1995) conclusion. Most recently, Anderson (1997) followed Steppan (1995) and synonymized *Andalgalomys* within *Graomys*.

#### GENERAL ISSUES

*Consistency index.*—Statements by Mares and Braun (1996) regarding low consistency indices (CIs) and retention indices (RIs) for the trees of Steppan (1995) reflect common misinterpretations of these measures. Because these indices are used widely and misinterpreted frequently, we discuss this issue in more detail here. First, Mares and Braun (1996:939) suggested that the low CI and RI of the phyllotine analyses in Steppan

(1995) indicated a lack of “agreement among trees.” This conclusion is erroneous because the CI is a measure of congruence among characters. For their statement to be correct, they should have referred to character state trees for individual characters; they made no such stipulation. Mares and Braun (1996:939) went on to say that “this would suggest the relationships among taxa may not be resolved.” Resolution could be determined easily by an examination of the trees, and they were fully resolved in the relevant region. We presume that Mares and Braun (1996) meant that phylogenies were not accurately resolved. It is this interpretation of CIs that reflects the widespread misconception that these indices are indicators of phylogenetic accuracy or reliability, under the assumption that there is an inverse relationship between phylogenetic accuracy and homoplasy. Concepts of resolution, accuracy, reliability, and robustness often are confused. Resolution is the degree to which nodes are estimated to be bifurcating. Accuracy is the degree to which the estimated tree reflects true phylogeny. Resampling techniques such as bootstrapping give measures of robustness that are used as surrogates for reliability. Steppan (1993, 1995) addressed this issue by pointing out that CIs for the sigmodontine and phyllo-tine trees were near both the mean for published studies with similar numbers of taxa (Archie 1989) and the expected values (Sanderson and Donoghue 1989) for matrices of those sizes. Moreover, Steppan (1993) discussed problems associated with using CI as an indicator of robustness or accuracy. Relative CIs also can be influenced strongly by character selection, thus making comparisons between studies questionable. Nonmolecular systematists typically choose to exclude phenotypic characters that appear to be excessively homoplastic on the basis of a priori expectations of relationship. Those expectations usually are unstated. Steppan (1995) chose the conservative approach; he included some highly homoplastic characters that might other-

wise be excluded, thus reducing the possibility of biasing results and lowering CIs further.

Interpretation of CI and RI as indicators of reliability has been refuted convincingly; Sanderson and Donoghue (1989) and Archie (1989) demonstrated that CI is correlated highly with number of taxa. So tight is this correlation that these indices provide virtually no information except in the context of number of taxa. Furthermore, Sanderson and Donoghue (1996) showed that there was no relationship between CI and several measures of support (decay index or Bremer support and bootstrap percentages). Trees with low CI can be robust and well resolved, and the inverse also is possible. Additionally, Hillis (1996) showed that precisely the condition that yields low CIs, large number of taxa, can facilitate accurate estimation of phylogeny. Consequently, comprehensive studies that follow the recommendation to sample a phylogeny thoroughly will tend to produce trees with a low CI. Neither CI nor RI is an estimate of phylogenetic resolution, accuracy, reliability, or robustness, and if reported as tree statistics, these indices should be discussed only in the context of number of taxa.

*Heuristic analyses.*—Mares and Braun (1996) criticized Steppan’s (1995) use of heuristic searches in his phylogenetic analysis by misciting Swofford and Olsen (1990). Specifically, Swofford and Olsen stated that heuristic searches are not guaranteed to produce optimal trees; the statement by Mares and Braun (1996:939) that heuristic approaches “do not produce optimal trees” is a subtle, but important, misinterpretation. In fact, under many conditions, heuristic algorithms do an excellent job of finding optimal trees. Criticizing a study for employing heuristic strategies in tree searches is analogous to criticizing use of sample statistics to estimate parametric values. The rigor of the heuristic search is the relevant issue in phylogenetic inference, just as adequacy of sampling scheme (rather than the fact that, for example, sample

means are used to estimate population means) is the important issue in classical statistics. More importantly, for data sets of 47 taxa analyzed by Steppan (1995) or 43 taxa analyzed by Braun (1993), exact searches are computationally impossible. To disregard all analyses that use heuristic searches would limit studies to no more than ca. 18 taxa.

*Taxon subsets.*—Mares and Braun (1996) reanalyzed part (<10%) of the data set of Steppan (1995) by analyzing 3 relevant taxa, *G. griseoflavus*, *G. domorum*, and *G. (A.) pearsoni*, using *Eligmodontia morgani* as an outgroup. They correctly reported that this restricted dataset yielded most-parsimonious trees with *G. griseoflavus* and *G. domorum* as sister species, leaving *Graomys* sensu stricto monophyletic with respect to *G. (A.) pearsoni*. However, inclusion of even 1 more species in the analysis (e.g., any *Phyllotis*) can result in a paraphyletic *Graomys*, as when Steppan's (1995) characters were analyzed.

For a variety of reasons, including bisecting long branches, uncovering homoplasy, and more accurately estimating ancestral character states, inclusion of more taxa usually is better than analyzing fewer taxa (e.g., Hillis 1996). Exceptions to this general guideline include testing the influence of problematic taxa, sensitivity analyses, or estimating robustness to taxon sampling (e.g., taxon jackknifing). Discarding most (>90%) of the data should not be expected to give more reliable results; alternative hypotheses should be tested with more information rather than less. Thus, in the following analysis, we re-examine the question with additional data.

*Consensus trees.*—Steppan (1995) used consensus trees to summarize relationships that were observed in several most-parsimonious trees. Mares and Braun (1996) reiterated Carpenter's (1988) criticism of this use of consensus trees because there can be no single character optimization possible for the consensus tree that is not longer than that on the most-parsimonious trees. How-

ever, even critics of statistical approaches to phylogeny view Steppan's use of consensus trees to summarize multiple equally parsimonious trees as valid (DeSalle and Brower 1997). Furthermore, the criticism of Mares and Braun (1996) loses its significance under the statistical framework discussed above because the unresolved consensus tree may actually be as good an estimate of the true tree as are any of the most-parsimonious trees. Even if this criticism were valid in general, relationships among *Graomys* and *Andalgalomys* are fully resolved in the consensus trees published by Steppan (1995:64, 65, figures 22 and 23). Carpenter's (1988) criticism refers only to choosing among equally parsimonious resolutions. Thus, the fact that figures 22 and 23 of Steppan (1995) are consensus trees is irrelevant to the specific hypothesis of paraphyly of *Graomys*. Even under a traditional point-estimate framework, all most-parsimonious trees agree that *Graomys* is paraphyletic, with a subclade (*Andalgalomys* plus *G. griseoflavus*) supported by a decay index of 4 and a bootstrap value of 83%.

*Trenchant characters.*—Mares and Braun (1996) criticized Steppan's (1995) omission of trenchant characters that distinguish *Andalgalomys* from *Graomys*. The concept of trenchant characters reflects neither a traditional cladistic perspective nor a statistical perspective but is instead a remnant of precladistic typological thinking. Trenchant characters are those that serve to distinguish 1 group from another. In Steppan's study, those trenchant characters cited by Mares and Braun (1996) were autapomorphies present only in *G. (A.) pearsoni*; because they contained no information as to order of branching, they were excluded. With inclusion of additional taxa in the phylogenetic analysis (*olrogi* and *roigi*), those characters become synapomorphies uniting the new species with *pearsoni*; they are nonetheless irrelevant to the issue at hand, specifically whether *Graomys* (sensu stricto) is paraphyletic with respect to these 3 species. Trenchant characters are relevant only in

the context of establishing taxonomic ranks, an entirely arbitrary endeavor. Given Steppan's (1995) results, were he to have retained *Andalgalomys* as a genus, he would have been required also to confer generic rank on *domorum* to avoid paraphyletic taxa.

Some trenchant characters that Mares and Braun (1996:938) listed for distinguishing *Andalgalomys* from *Graomys* include a whitish subauricular spot, a small notch on the postcingulum of M1 (variably present in *Andalgalomys* and absent in paratypes of *olrogi*), and a small anteriorly directed anteroloph of M2 (often lost with wear). These dental traits in particular do not seem to qualify as trenchant characters, given the highly variable nature of these structures in sigmodontine molars.

*Missing data.*—In an ideal situation, all characters could be evaluated for all taxa in an analysis, but often this is not the case for a variety of reasons. Mares and Braun (1996) raised the possibility that missing data for *domorum* and *pearsoni* (9.2% and 8.2% of characters coded, respectively) could bias Steppan's (1995) analysis in some unpredictable way and cited Wilkinson (1995a, 1995b) in support. Although this issue is of interest from both statistical and traditional viewpoints, the point raised by Wilkinson (1995a:111) is not relevant in the present case because, as he noted, "arbitrary resolutions [resulting from zero-length branches caused by missing data or other causes] cannot affect the topology of strict consensus trees." The phyllotine branches in question in Steppan's (1995) analyses have nonzero lengths and thus are not subject to the arbitrary resolution discussed by Wilkinson (1995a). The 2nd study they chose to cite (Wilkinson 1995b) focused on a situation where large blocks of data were missing (>50%) as is often the case with partial preservation of fossil material. Again, the problem caused by missing data in that case is polytomies produced by many most-parsimonious trees, a situation that is not a problem in Steppan's (1995)

analysis because *Graomys*–*Andalgalomys* is fully resolved.

#### PHYLOGENETIC ANALYSES

For reasons given above and because the relevant portion of the topology may be unstable, we have expanded the data matrix of Steppan (1995) and reanalyzed it. Four species have been added (Appendix I): *Salinomys delicatus* (Braun and Mares 1995) and *G. (A.) roigi* (Mares and Braun 1996), both described since Steppan's (1995) publication; *A. olrogi*, the type species; and the new genus and species *Tapecomys primus* from Tapecua, Bolivia (Anderson and Yates, 2000). *Tapecomys primus*, designated species nova by Steppan (1993), previously was found to be closely related to the *Graomys* complex (Steppan 1993) but was removed from the 1995 analysis pending formal description. Two characters have been added from Mares and Braun's (1996) list of trenchant characters that define *Andalgalomys*: whitish subauricular spot and large anterior palatal pits. The other listed characters were variably present in *Andalgalomys*, insufficiently distinct to define new character states, or already addressed as parts of other characters (e.g., small anteroloph in M2 addressed in character 21—Steppan 1995). Mares and Braun (1996) were correct in pointing out the discrepancy between Steppan's (1995) assessment of the alisphenoid strut in *A. pearsoni* and that reported by other workers (Mares and Braun 1996; Olds et al. 1987). Steppan (1995) had reported a strut for *pearsoni* (the others reported it absent), and the phylogenetic analysis resulted in the hypothesis that presence of the strut was a synapomorphy shared by *Graomys* sensu stricto and *Andalgalomys*. A re-examination of the *pearsoni* material revealed alisphenoid struts in 3 specimens, partial struts in 1, and absence of struts in 2, whereas examination of additional material from the American Museum of Natural History found struts present in 1 of 4 specimens examined. Thus, contrary to other reports, alisphenoid struts are present in

*pearsoni* but not consistently, and we have accordingly modified the coding for that species to account for this polymorphism. The struts are absent from the 1 specimen of *roigi* and the 3 specimens of *olrogi* available for examination.

The new matrix contains 51 taxa and 100 characters. Characters were treated as by Steppan (1995). Heuristic searches were conducted with 100 random-addition sequences, with tree bisection-reconnection branch swapping and multiple parsimonious trees saved at each step. We found 22 equally parsimonious trees (841 steps, CI = 0.206, RI = 0.545; Fig. 1). In all of these, *A. roigi* and *A. pearsoni* are sister taxa, *Andalgalomys* is monophyletic, *Tapecomys* is sister to *Andalgalomys*, *Graomys* sensu stricto is paraphyletic, and *G. griseoflavus* is sister to the *Andalgalomys*-*Tapecomys* clade. Within the statistical phylogenetics framework, we can evaluate strength of support for a paraphyletic *Graomys* and thereby the taxonomic recommendations of Steppan (1995). The best tree constrained for monophyly of *Graomys* sensu stricto is 842 steps; the test statistic for the hypothesis of interest (reciprocal monophyly) is 1. The significance of the test statistic can be assessed by means of parametric and non-parametric tests described above. The most powerful of these tests that is applicable to morphological data is the Kishino-Hasegawa test (Kishino and Hasegawa 1989), which examines the differences between 2 trees on a character-by-character basis. Under the null hypothesis of no significant difference between a pair of topologies, the distribution of single-character length differences between the trees being compared can be assumed to follow a bivariate normal distribution (Kishino and Hasegawa 1989); this distribution can be used to estimate the probability of obtaining a test statistic as extreme as that observed under the null hypothesis of no significant difference between the trees being compared. In this case, the test statistic is not significant ( $P = 0.829$ ). Therefore, application of the sta-

tistical approach has demonstrated that these data can refute neither the taxonomic hypothesis of Mares and Braun (1996) nor Steppan's (1995) hypothesis of paraphyly. Disagreements over the appropriate rank for *Andalgalomys* cannot be resolved by data collected thus far, and additional data will be required, perhaps DNA sequence data.

Although current morphological data cannot distinguish between paraphyly and monophyly of *Graomys* with respect to *Andalgalomys*, the statistical approach has demonstrated the degree of confidence that can be placed in competing hypotheses and has shown that decisions regarding paraphyly of *Graomys* drawn from these data are arbitrary. Given the predominantly low bootstrap values in parsimony analysis and the small magnitude of difference in tree length, the difficulty encountered in distinguishing statistically between these alternatives using these data is not altogether surprising. However, a 2nd example of the power of this approach relates to a recent hypothesis regarding placement of another phyllotine, *Calomys callosus*. Engel et al. (1998) hypothesized, on the basis of analyses of mitochondrial DNA sequence data, that *Calomys* is polyphyletic and that *C. callosus* is not a phyllotine but instead is a member of an oryzomyine clade. Analyzing the same set of morphological data discussed above produces 162 best trees consistent with that hypothesis (if *C. callosus* is constrained to be included in a oryzomyine clade); these trees are 20 steps longer than the best trees, so in this case the test statistic is 20. The 162 best topologies constrained to fit the predictions of Engel et al. (1998) were tested against the best unconstrained topology. The corresponding  $P$ -values range from 0.015 to 0.060. The probability of observing a difference of 20 steps by chance alone under the null hypothesis is low, and these morphological data refute the position of *C. callosus* as a member of the oryzomyine clade. Thus, despite low resolution afforded by these data, we are still able to reject the hypothesis of

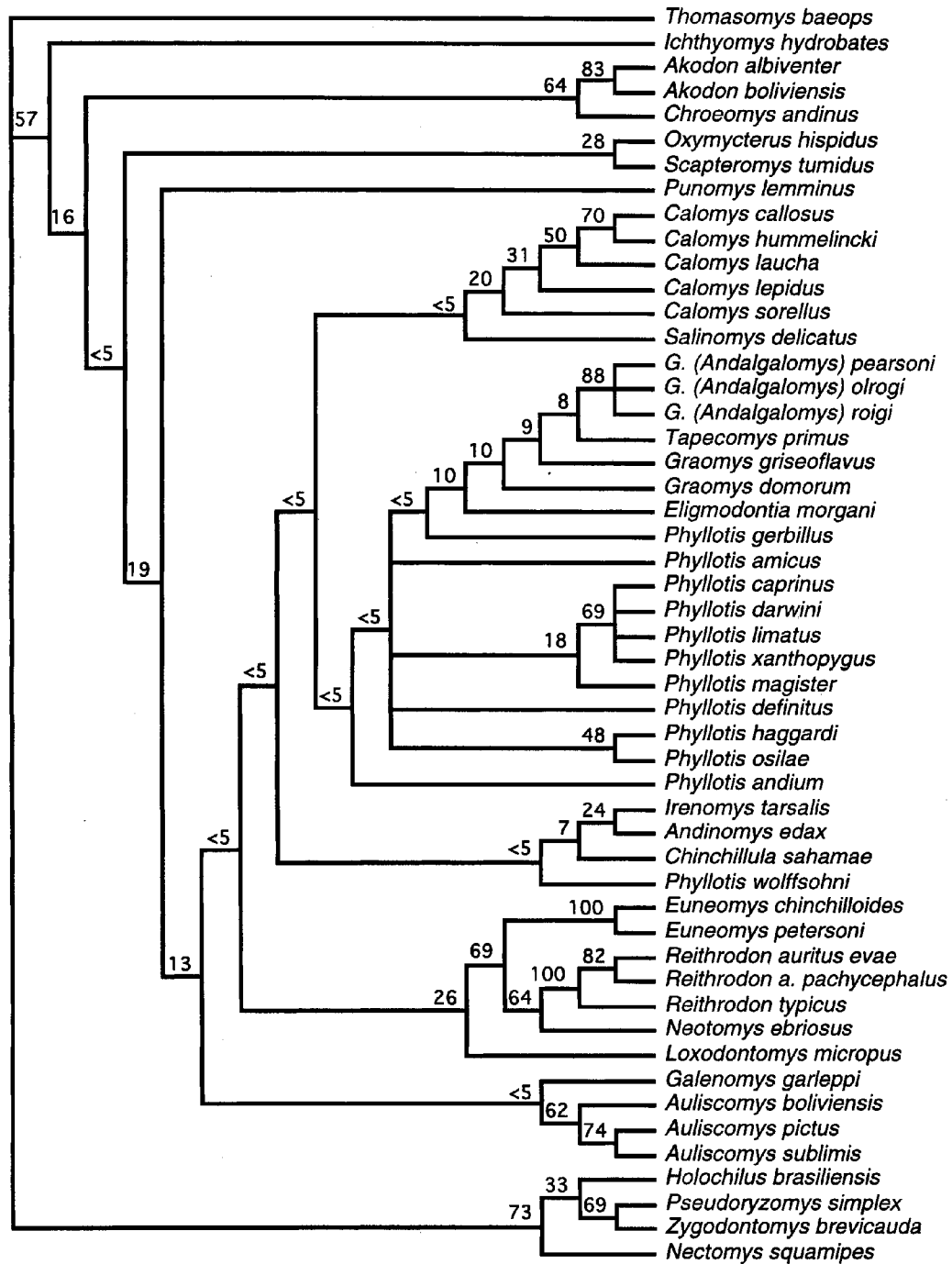


FIG. 1.—Strict consensus of 22 equally parsimonious trees found by heuristic search using PAUP\* (version 4.0d59-d62—D. L. Swofford, pers. comm.). Starting tree was obtained by stepwise addition with 100 random input orders, and branch swapping was conducted using tree bisection–reconnection with multiple parsimonious trees saved at each step. This strategy also was used in constrained searches (trees not shown). Numbers above branches represent bootstrap values; 1,000 bootstrap replicates were conducted.



Engel et al. (1998). Interestingly, Engel et al. (1998) conducted the reciprocal test; that is, they constrained their data to fit the taxonomic hypothesis (a monophyletic *Calomys*) and were able to reject the monophyly of their 3 sequences for *Calomys*. Thus, the possibilities arise that the sequence for *C. callosus* examined by Engel et al. (1998) either is misidentified or represents a contaminant, that there is a gene-tree/species-tree conflict with regard to this taxon and its mitochondrial DNA, or our analyses and/or those of Engel et al. (1998) fail to account for some unknown systematic error. This area of conflict merits further exploration.

#### SYSTEMATIC RECOMMENDATIONS

Modern systematic conventions specify that a taxon name is applicable only if, in addition to rules of the nomenclatural code dealing with validity, 2 phylogenetic conditions are met: the taxon must be monophyletic and all taxa of equal rank must be monophyletic with respect to it. If the 1st condition is satisfied but a taxon of equal rank is paraphyletic with respect to the taxon in question, one must either subsume the taxon of interest within the more inclusive taxon (retaining the name if desired by lowering its rank) or elevate other lineages of the more inclusive taxon to equal rank until no paraphyletic taxa remain. Stepan (1995) followed the former procedure on grounds that the evidence was strong enough to accept the paraphyly of *Graomys* with respect to *Andalgalomys*. Thus, the fundamental question is the paraphyly of *Graomys*.

Here, we summarize the evidence for and against paraphyly of *Graomys*. The 2 major observations favoring paraphyly of *Graomys* sensu stricto and the synonymy of *Andalgalomys* are the available phylogenetic analyses. The most-parsimonious trees from this reanalysis indicate that *domorum* lies outside an *Andalgalomys griseoflavus* clade. Similarly, in the most-parsimonious trees of Braun (1993), *edithae* and possibly

*domorum* are outside that clade. The point estimates of phylogeny suggest a paraphyletic *Graomys*.

As evidence against paraphyly of *Graomys*, although not addressing its monophyly, Mares and Braun (1996) raised a series of criticisms that we have shown to be overstated or invalid. A reasonable criticism not raised by Mares and Braun (1996) is that 3 of the 4 putative synapomorphies reported by Stepan (1995) involve conformational changes in M3, characters that could be correlated structurally with each other. If so, then bootstrap values and Bremer support for this clade as reported by Stepan (1995) would be inflated. Even if these characters are independent, bootstrap support in our reanalysis is low and the decay index is small; alternative hypotheses are statistically indistinguishable on the basis of these data, and there is no longer strong evidence for paraphyly. To the extent that *Andalgalomys* is morphologically distinctive, one may choose to recognize that formally. The issue of morphological distinctiveness touched on by Mares and Braun (1996) is relevant in regard to ranking so long as all ranking decisions are consistent with the topology. A series of synapomorphies support *Andalgalomys*, although they are fewer and less trenchant than Mares and Braun (1996) contended.

In conclusion, phylogenetic evidence for the position of *Andalgalomys* is equivocal, and taxonomic resolution is therefore arbitrary. Until the question of paraphyly of *Graomys* can be tested in a comprehensive study with additional data, the decision regarding the status of *Andalgalomys* should be based on consistency and clear communication among systematists. The 1st described species of *Andalgalomys* originally was placed in *Graomys* (Myers 1977), but for the 20 years following Myers' (1977) work, systematists recognized *Andalgalomys*, so that taxonomy is well understood and most people working with the South American murid rodents would interpret the name *Graomys* to exclude *Andalgalomys*.

The most recent and comprehensive revision dealing with this issue synonymized *Andalgalomys* (Steppan 1995). Given the current state of knowledge and the need for a consistent nomenclature, we recommend that the common usage of *Andalgalomys* as a genus be retained provisionally until additional data can resolve the phylogeny of *Graomys*.

#### ACKNOWLEDGMENTS

We are grateful to P. Myers for confirming some character assessments on *A. pearsoni*. We thank M. Mares, J. Braun, and S. McLaren for making available specimens of *Salinomys*, *A. roigi*, and *A. olrogi*. S. Anderson and T. Yates graciously allowed us to include the undescribed species from Tapecua, Bolivia, in the analysis. A. Thistle, R. Hoffmann, J. Huelsenbeck, D. Swofford, J. Wilgenbusch, D. Wilson, and 2 anonymous reviewers provided many thoughtful comments that improved the accuracy and clarity of the manuscript. We thank D. Swofford for making the test version of PAUP\* available to us.

#### LITERATURE CITED

- ANDERSON, S. 1997. Mammals of Bolivia, taxonomy and distribution. *Bulletin of the American Museum of Natural History* 231:1–652.
- ANDERSON, S., AND T. L. YATES. 2000. A new genus and species of phyllotine rodent from Bolivia. *Journal of Mammalogy* 81:18–36.
- ARCHIE, J. W. 1989. A randomization test for phylogenetic information in systematic data. *Systematic Zoology* 38:239–252.
- BRAUN, J. K. 1993. Systematic relationships of the tribe Phyllotini (Muridae: Sigmodontinae) of South America. Special Publication, Oklahoma Museum of Natural History, Norman.
- BRAUN, J. K., AND M. A. MARES. 1995. A new genus and species of phyllotine rodent (Rodentia: Muridae: Sigmodontini: Phyllotini) from South America. *Journal of Mammalogy* 76:504–521.
- BREMER, K. 1994. Branch support and tree stability. *Cladistics* 10:295–304.
- CARPENTER, J. M. 1988. Choosing among multiple equally parsimonious cladograms. *Cladistics* 4:291–296.
- DESALLE, R., AND A. V. Z. BROWER. 1997. Process partitions, congruence, and the independence of characters: inferring relationships among closely related Hawaiian *Drosophila* from multiple gene regions. *Systematic Biology* 46:751–764.
- ENGEL, S. R., K. M. HOGAN, J. F. TAYLOR, AND S. K. DAVIS. 1998. Molecular systematics and paleobiogeography of the South American sigmodontine rodents. *Molecular Biology and Evolution* 15:35–49.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791.
- HILLIS, D. M. 1996. Inferring complex phylogenies. *Nature* 383:130–131.
- HUELSENBECK, J. P., D. M. HILLIS, AND R. JONES. 1996. Parametric bootstrapping in molecular phylogenetics: applications and performance. Pp. 19–45 in *Molecular zoology: advances, strategies, and protocols* (J. D. Ferraris and S. R. Palumbi, eds.). Wiley-Liss, New York.
- HUELSENBECK, J. P., AND B. RANNALA. 1997. Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* 276:227–232.
- KISHINO, H., AND M. HASEGAWA. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA data, and the branching order in Hominoidea. *Journal of Molecular Evolution* 29:170–179.
- LARSON, A. 1991. Evolutionary analysis of length variable sequences: divergent domains of ribosomal RNA. Pp. 221–248 in *Phylogenetic analysis of DNA sequence data* (M. M. Miyamoto and J. Cracraft, eds.). Oxford University Press, New York.
- MARES, M. A., AND J. K. BRAUN. 1996. A new species of phyllotine rodent, genus *Andalgalomys* (Muridae: Sigmodontinae), from Argentina. *Journal of Mammalogy* 77:928–941.
- MYERS, P. 1977. A new phyllotine rodent (genus *Graomys*) from Paraguay. *Occasional Papers of the Museum of Zoology, University of Michigan* 676:1–7.
- OLDS, N., S. ANDERSON, AND T. L. YATES. 1987. Notes on Bolivian mammals 3: a revised diagnosis of *Andalgalomys* (Rodentia, Muridae) and the description of a new subspecies. *American Museum Novitates* 2890:1–17.
- SANDERSON, M. J., AND M. J. DONOGHUE. 1989. Patterns of variation in levels of homoplasy. *Evolution* 43:1781–1795.
- SANDERSON, M. J., AND M. J. DONOGHUE. 1996. The relationship between homoplasy and confidence in a phylogenetic tree. Pp. 67–89 in *Homoplasy: the recurrence of similarity in evolution* (M. J. Sanderson and L. Hufford, eds.). Academic Press, San Diego, California.
- STEPPAN, S. J. 1993. Phylogenetic relationships among the Phyllotini (Rodentia: Sigmodontinae) using morphological characters. *Journal of Mammalian Evolution* 1:187–213.
- STEPPAN, S. J. 1995. Revision of the leaf-eared mice Phyllotini (Rodentia: Sigmodontinae) with a phylogenetic hypothesis for the Sigmodontinae. *Fieldiana: Zoology* 80:1–112.
- SWOFFORD, D. L., AND G. J. OLSEN. 1990. Phylogeny reconstruction. Pp. 411–501 in *Molecular systematics* (D. M. Hillis and C. Moritz, eds.). 1st ed. Sinauer Associates, Inc., Publishers, Sunderland, Massachusetts.
- TEMPLETON, A. 1987. Nonparametric inference from restriction cleavage sites. *Molecular Biology and Evolution* 4:315–319.
- WILKINSON, M. 1995a. Arbitrary resolutions, missing entries, and the problem of zero-length branches in parsimony analysis. *Systematic Biology* 44:108–111.

- WILKINSON, M 1995b. Coping with abundant missing entries in phylogenetic inference using parsimony. *Systematic Biology* 44:501–514.
- WILLIAMS, D. F., AND M. A. MARES. 1978. A new genus and species of phyllotine rodent (Mammalia: Muridae) from northwestern Argentina. *Annals of the Carnegie Museum* 47:193–221.

Submitted 28 April 1998. Accepted 29 January 1999.

Associate Editor was Troy L. Best.

#### APPENDIX I

*Specimens examined.*—Museum abbreviations are as follows: American Museum of Natural History, AMNH; Carnegie Museum of Natural History, CMNH; Oklahoma Museum of Natural

History, OMNH; University of Michigan Museum of Zoology, UMMZ.

*Graomys (Andalgalomys) olrogi.* Argentina. Catamarca, 15 km W Andalgalá (CMNH 44029: holotype, 44022, 44023). Total 3.

*Graomys (Andalgalomys) pearsoni.* Bolivia. (AMNH 262344–262347). Paraguay. Nueva Asunción: km 620, Trans-Chaco Road (UMMZ, uncataloged, T. W. Nelson field numbers 184, 193, 201, 202, 233). Total 9.

*Graomys (Andalgalomys) roigi.* Argentina. San Luis, 6 km W Haultaran, Parque Provincial Sierra de los Quijades (OMNH 23795). Total 1.

*Salinomys delicatus.* Argentina. San Luis, Departamento. Capital, 15 km E Salinas de Besedero (OMNH 23602). Total 1.