The Auk 119(1):88–108, 2002

SEABIRD SUPERTREES: COMBINING PARTIAL ESTIMATES OF PROCELLARIIFORM PHYLOGENY

MARTYN KENNEDY[1] AND RODERIC D. M. PAGE
Division of Environmental and Evolutionary Biology, Institute of Biomedical and Life Sciences,
University of Glasgow, Glasgow G12 8QQ, United Kingdom

ABSTRACT.—The growing use of comparative methods to address evolutionary questions has generated an increased need for robust hypotheses of evolutionary relationships for a wide range of organisms. Where a phylogeny exists for a group, often more than one phylogeny will exist for that group, and it is uncommon that the same taxa are in each of the existing trees. The types of data used to generate evolutionary trees can also vary greatly, and thus combining data sets is often difficult or impossible. To address comparative questions for groups where multiple phylogenetic hypotheses already exist, we need to combine different hypotheses in a way that provides the best estimate of the phylogeny for that group. Here, we combine seven seabird phylogenies (based on behavioral, DNA–DNA hybridization, isozyme, life history, morphological, and sequence data) to generate a comprehensive supertree for the Procellariiformes using matrix representation with parsimony. This phylogeny contains 122 taxa and represents a conservative estimate of combined relationships presented in the original seven source trees. We compared the supertree with results of a combined sequence data supermatrix for 103 seabird taxa. Results of the two approaches are broadly concordant, but matrix representation with parsimony provides a more comprehensive and more conservative estimate of the phylogeny of the group because it is less influenced by the largest of the source studies (which uses a single, relatively quickly evolving gene). Genetic data sets that can be combined in a supermatrix approach are currently less likely to be available than phylogenies that can be combined using some form of supertree approach. Although there are limitations to both of those approaches, both would be simpler if all phylogenetic studies made both their data sets and trees they generate available through databases such as TREEBASE. *Received 8 December 2000, accepted 21 September 2001.*

RESUMEN.—El uso creciente de métodos comparativos para abordar preguntas evolutivas ha incrementado la necesidad de hipótesis robustas sobre relaciones evolutivas para una amplia variedad de organismos. Cuando existe una filogenia para un grupo, a menudo existirá más de una filogenia para ese grupo, y es poco frecuente que los mismos taxa estén presentes en cada uno de los árboles. Los tipos de datos usados para generar árboles evolutivos también pueden variar enormemente, por lo que generalmente es difícil o imposible combinar estos datos. Para responder a preguntas comparativas en grupos para los cuales existen múltiples hipótesis filogenéticas, necesitamos combinar diferentes hipótesis de manera que obtengamos la mejor estimación de la filogenia de estos grupos. Aquí combinamos siete filogenias (basadas en comportamiento, hibridización de ADN–ADN, isoenzimas, historia de vida, morfología y datos de secuenciamiento) para generar un super-árbol integrador para los Procellariiformes usando representación de matrices combinada con parsimonia. Esta filogenia contiene 122 taxa y representa una estimación conservativa de las relaciones presentadas en los siete árboles originales combinados. Comparamos el super-árbol con los resultados de una matriz de datos de secuencias combinada para 103 taxa de aves marinas. A grandes rasgos, los resultados de las dos aproximaciones concuerdan, pero la representación de matrices combinada con parsimonia brinda una estimación más integral y más conservativa de la filogenia del grupo, debido a que está menos influenciada por uno de los estudios utilizados (la fuente con más datos, que a su vez usa un único gen de evolución relativamente rápida). Los juegos de datos genéticos que pueden ser combinados en una super-matriz están por lo general menos disponibles que las filogenias que pueden ser combinadas usando alguna aproximación de tipo super-árbol. Aunque existen limitaciones para estas dos aproximaciones, ambas serían más simples si todos los estudios filogenéticos pusieran a dispo-

---

[1] E-mail: martyn.kennedy@bio.gla.ac.uk

sición tanto sus juegos de datos como los árboles que generan a través de bases de datos como TREEBASE.

RELATIONSHIPS WITHIN THE tubenose seabirds (Procellariiformes) are of general interest to many biologists, at least in part because they are a diverse and wide-ranging group, and because they are commonly found in most oceanic regions of the world. Nunn and Stanley (1998) generated a phylogeny for 85 species of tubenose seabirds for their discussion of effects of body size on rates of molecular evolution. By being such a diverse group (e.g. showing such a great range in body size), these birds allow comparative questions like that of Nunn and Stanley (1998) to be readily addressed when a phylogeny is available. As well as ranging greatly in size, the tubenose seabirds are particularly interesting because they are behaviorally and ecologically very diverse, and they provide a model system for investigating cospeciation (e.g. Paterson et al. 2000).

In addition to Nunn and Stanley's (1998) phylogeny for tubenose seabirds, there are several other phylogenies available for that group. Some of those phylogenies include taxa not present in Nunn and Stanley's (1998) study, whereas some also disagree with the relationships found in their study. The storm petrels (Hydrobatidae), for example, unexpectedly do not form a monophyletic group in Nunn and Stanley's (1998) phylogeny. Given that there are a number of phylogenies available for that group, precisely what the best estimate of the phylogeny is remains uncertain.

Within the last few decades, there has been a dramatic increase in number of studies using phylogenies to address a wide range of issues. Those issues include, for example, behavior (e.g. Zyskowski and Prum 1999), biogeography (e.g. Kennedy and Spencer 2000), coevolution (e.g. Paterson et al. 2000), genetic systems (e.g. Cruickshank and Thomas 1999), language (e.g. Gray and Jordan 2000), rates of molecular evolution (e.g. Johnson and Sorenson 1998), speciation (e.g. Friesen and Anderson 1997), and taxonomy (e.g. Kennedy et al. 1999). Because biologists are becoming more convinced of the utility of taking a phylogenetic approach to questions they wish to address, robust hypotheses about phylogenetic relationships for the taxa of interest are required.

Even with ongoing advances in molecular technology, phylogenetic hypotheses do not exist for most of the world's taxa. When phylogenies do exist for a group, they will often not include all taxa of interest to the researcher. A single phylogeny for the taxa of interest is not always available, and studies have sometimes had to combine two or more phylogenies to obtain a tree that contains all the taxa (Sanderson et al. 1998). Kennedy et al. (1996), for example, had to combine four different source trees (two generated from morphological data and two from DNA–DNA hybridization data) to investigate homology of pelecaniform behaviors by mapping them onto the best estimate of that group's phylogeny.

A tree that results from combination of multiple source-tree topologies has been termed a ''supertree'' (Sanderson et al. 1998). Intuitively, the ideal way of combining several source trees would appear to be to combine all source phylogenies' data matrices into a single ''supermatrix'' that could then be analyzed to estimate the phylogeny. Sanderson et al. (1998) note, however, that the supermatrix approach will often not be tenable because of the cost involved with filling in gaps in the data as well as difficulties associated with combining some types of data. They point out, for example, that DNA–DNA hybridization data would not be able to be included in a supermatrix, and that homologizing characters would become increasingly difficult as the size of the matrix increased as more distantly taxa were added (Sanderson et al. 1998).

The supertree approach offers an alternative to the supermatrix approach. One method for constructing supertrees is matrix representation with parsimony (MRP; Baum 1992, Ragan 1992). Matrix representation with parsimony converts topologies of individual source trees into a data matrix (for a general explanation, see Sanderson et al. 1998). Once matrices for each of the source trees are combined, supertrees can be found using parsimony analysis. Because matrices are derived from the source trees' topologies, MRP allows different data types (e.g. sequences, morphology, behavior, allozymes, DNA–DNA hybridization) to be combined (Bininda-Emonds and Bryant 1998).

Whereas some studies have had to cobble multiple source trees together (e.g. Kennedy et al. 1996), MRP uses an explicit optimality criterion. As well as providing an optimal solution (or solutions) to combining multiple source trees, MRP allows different coding schemes and different weights to be applied to certain ''characters'' (see Sanderson et al. 1998). Nodes from source trees (which are effectively characters in the MRP matrix), for example, could be differentially weighted depending on the level of support that the nodes received in source trees. Whereas supertree and consensus approaches have limitations (Steel et al. 2000), and there is some debate over how to best construct phylogenetic supertrees (e.g. see Sanderson et al. 1998, Bininda-Emonds et al. 1999), MRP is currently the most commonly used method in construction of large supertrees.

Constructing supertrees from multiple source trees is not necessarily a trivial exercise. Bininda-Emonds et al. (1999), for example, used 177 literature sources to construct supertrees for the 271 extant species of Carnivora. Finding all source trees for a large number of taxa is potentially an extremely time consuming task, without considering that source trees then need to be coded into an appropriate single-matrix format, and then analyzed. Finding source trees would be an easier task if all studies that generate phylogenies were submitted to a phylogeny database such as TREEBASE (see Acknowledgments). If all trees for a group were available in NEXUS file format (Maddison et al. 1997) from TREEBASE, source trees could then simply be combined into a MRP matrix using software such as RADCON (Thorley and Page 2000). RADCON will take all source trees and output them as a single MRP matrix written in NEXUS format ready for parsimony analysis. An alternative (and supplementary) approach to searching phylogeny databases is to search the Web of Science (see Acknowledgments) or some other literature database for relevant journal references and thus search for any articles that contain phylogenies for the group of interest.

In this article, we compare results of a MRP analysis of several phylogenies available for tubenose seabirds with a supermatrix approach using available sequence data. Because several phylogenies available for members of the procellariiforms were generated from the same type of data (i.e. mitochondrial DNA sequence), that group offers an ideal opportunity to compare supertree and supermatrix approaches. A robust, comprehensive, and conservative estimate of what is currently known about the phylogeny of that group of birds will also allow future comparative analyses to be readily performed.

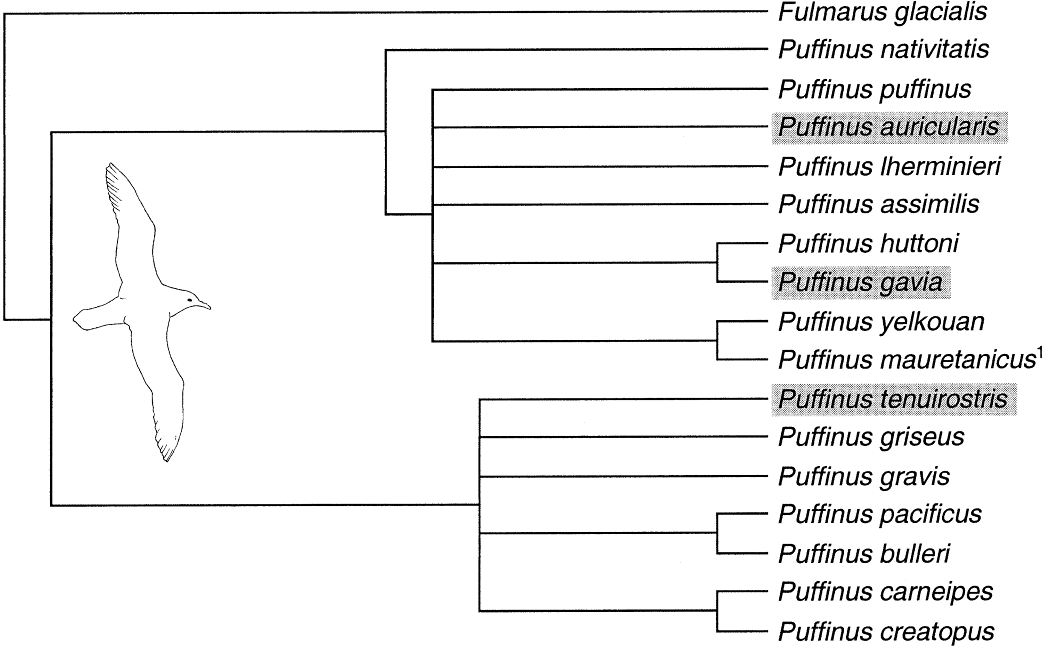
METHODS

Data.—We combined topological information from seven different studies (see Fig. 1) using MRP. In addition to our preexisting knowledge of the literature (e.g. Sibley and Ahlquist 1990), we searched the Web of Science in order to comprehensively survey the literature for procellariiform phylogenies. Our initial search criteria were the terms phylogen* and seabird* (the asterisk wildcard allows any word with that root to be found). After obtaining the list of studies that included derivatives of both of those terms, we investigated appropriateness of each of the individual studies using their abstracts. For each individual study, we also surveyed studies that cited it

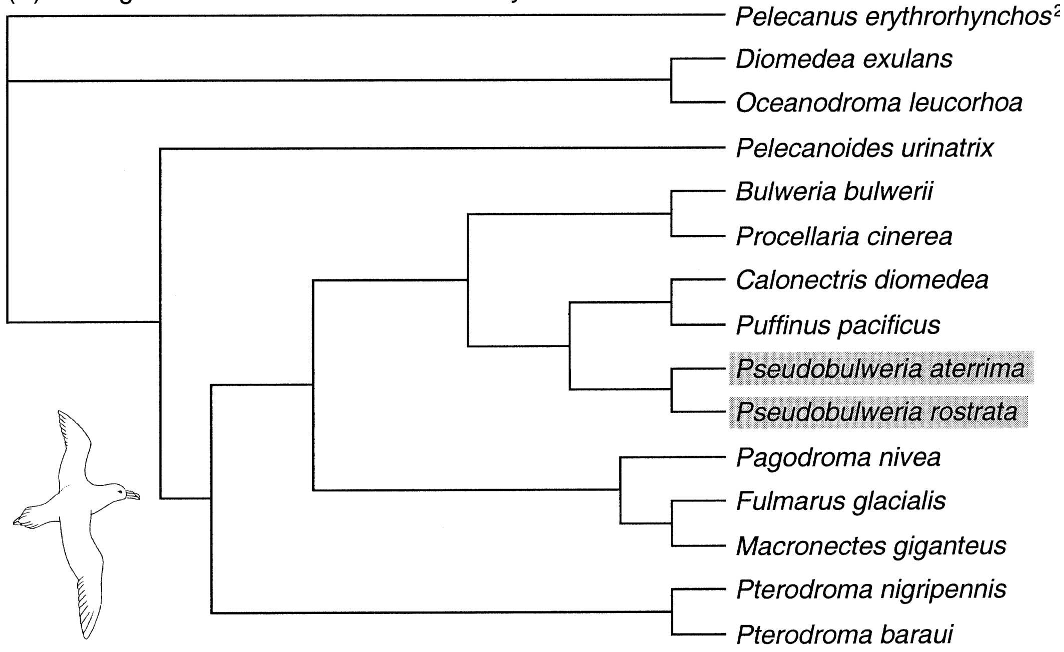
→

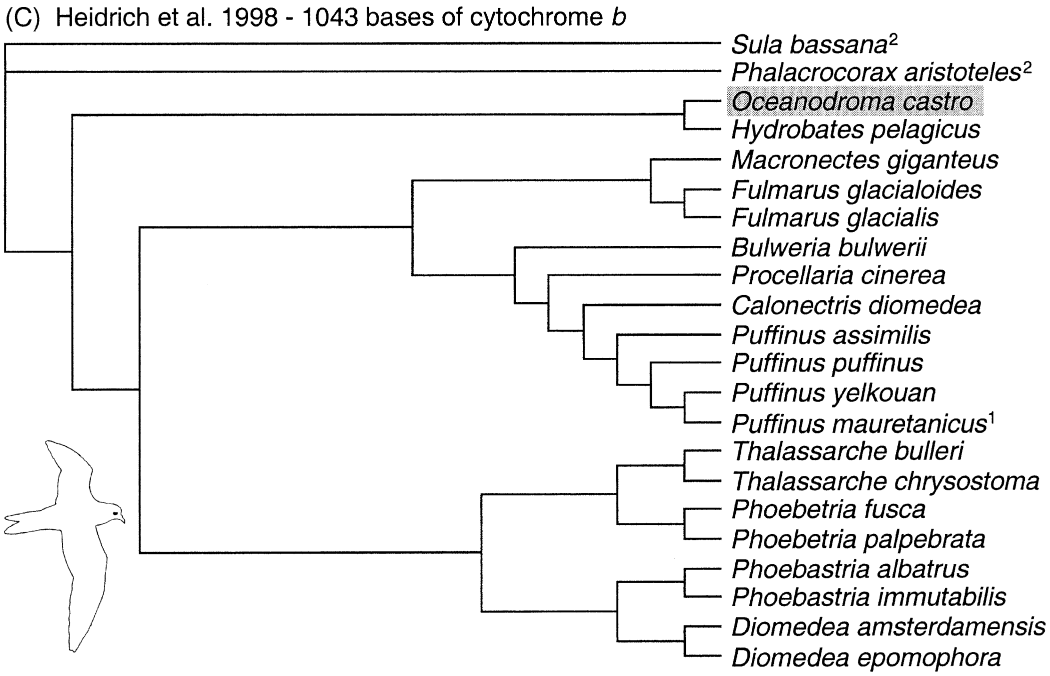
FIG. 1. The topologies of the source trees: (A) Austin (1996)—307 bases of cytochrome *b*; (B) Bretagnolle et al. (1998)— 496 bases of cytochrome *b*; (C) Heidrich et al. (1998)— 1,043 bases of cytochrome *b*; (D) Imber (1985)—morphology and life history information; (E) Nunn and Stanley (1998)— 1,143 bases of cytochrome *b*; (F) Paterson et al. (1995)— 12S rRNA, isozyme, and behavioral life history data; and (G) Sibley and Ahlquist (1990)—DNA–DNA hybridization data. We have followed the classification of Sibley and Munroe (1990) for consistency, but have used four genera accepted in the source phylogenies: *Pseudobulweria* is *Pterodroma* in Sibley and Munroe (1990); *Phoebastria* and *Thalassarche* are *Diomedea* in Sibley and Munroe (1990); and *Halocyptena* is *Oceanodroma* in Sibley and Munroe (1990). The taxa highlighted by a shaded background are unique to that source tree. [1]Not classified in Sibley and Munroe (1990). [2]Not used in the analysis. [3]No species name was associated with the genus in Sibley and Ahlquist's (1990) tree. Because for the MRP analysis the taxa labels must all contain both generic and specific names, the first species of the genus named in Sibley and Munroe's (1990) classification (see Appendix) that occurred in another of the source trees was allocated to the genus.

(A) Austin 1996 - 307 bases of cytochrome *b*



(B) Bretagnolle et al. 1998 - 496 bases of cytochrome *b*




(C) Heidrich et al. 1998 - 1043 bases of cytochrome *b*

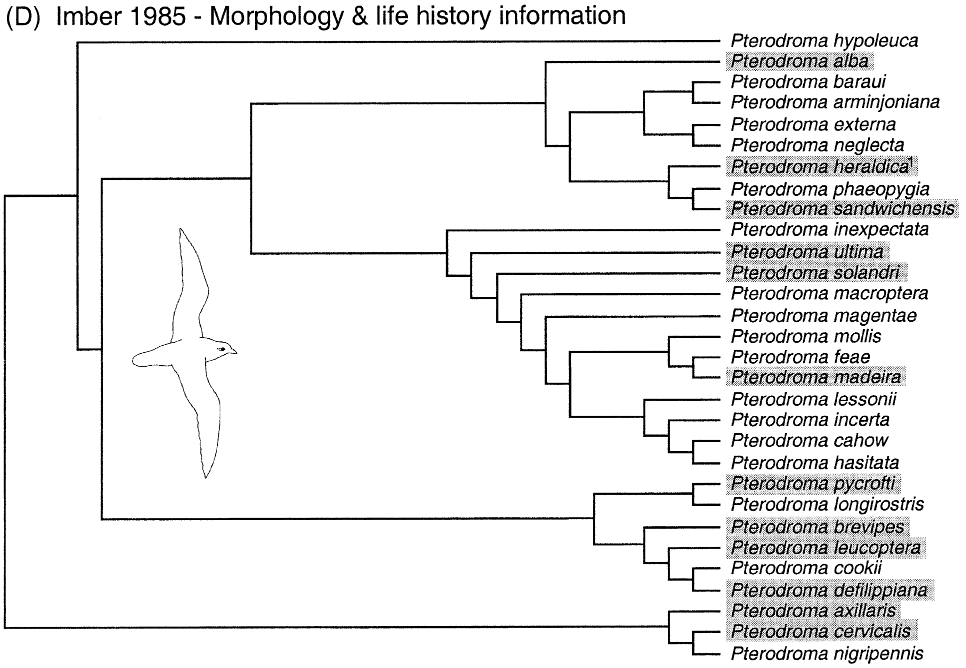

(D) Imber 1985 - Morphology & life history information

FIG. 1. Continued.

(E) Nunn and Stanley 1998 - 1143 bases of cytochrome *b*

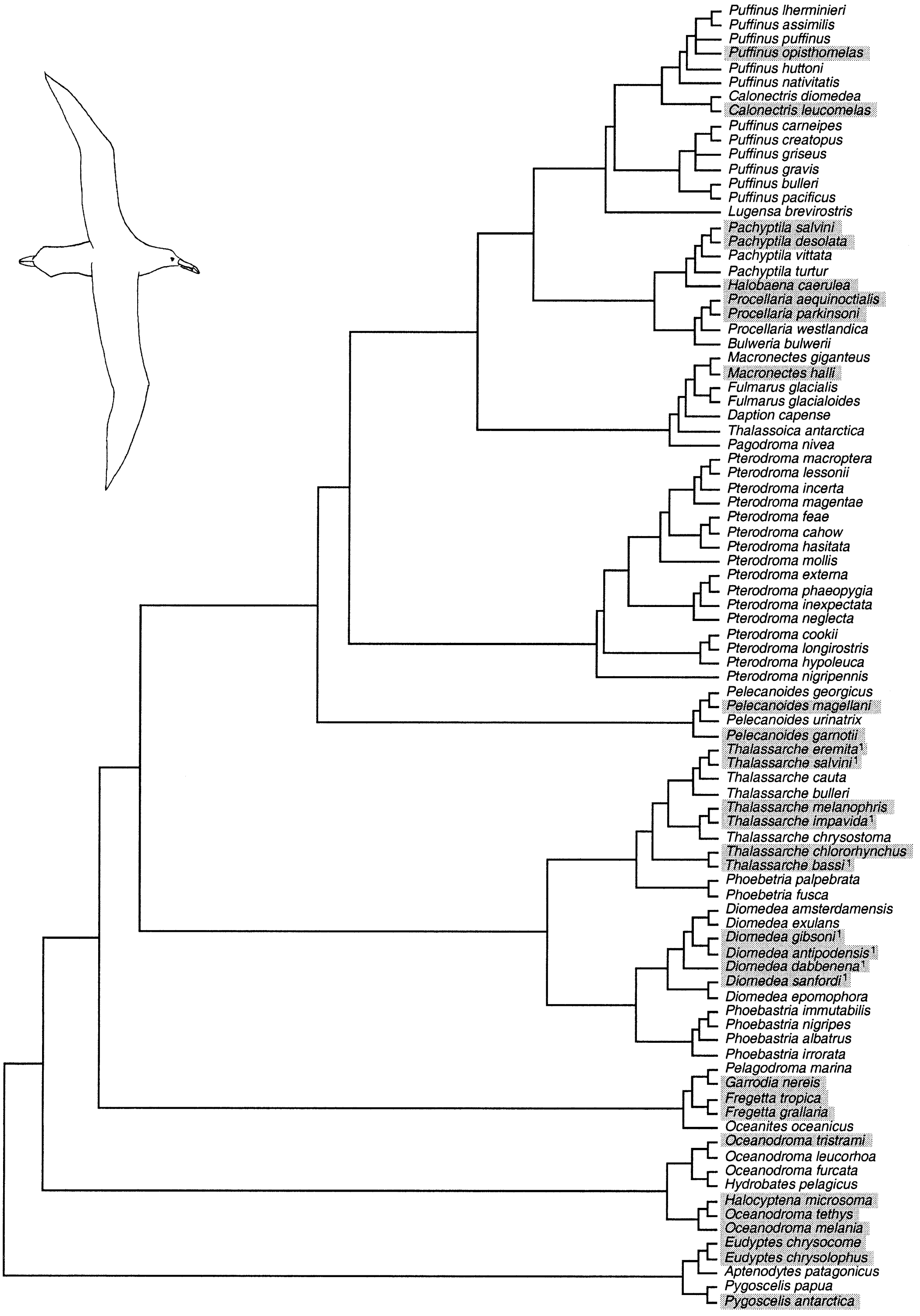
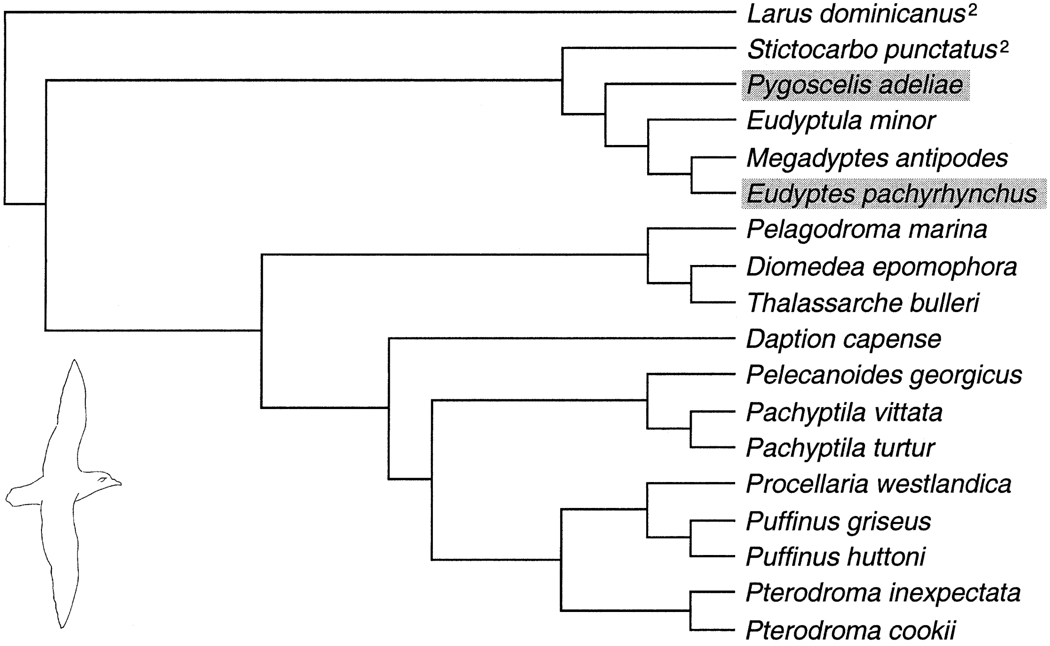


FIG. 1. Continued.

(F) Paterson et al. 1995 - 12SrRNA, isozyme & behavioral life history data



(G) Sibley and Ahlquist 1990 - DNA-DNA hybridization data

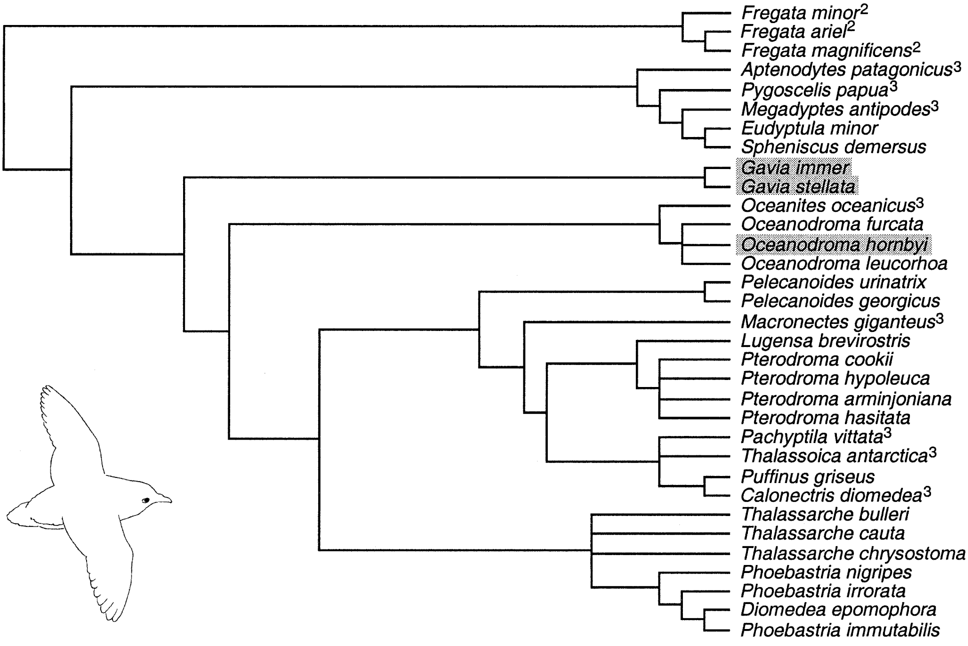


FIG. 1. Continued.

and all studies cited in it. Following that procedure we found the journal reference for all but one of the source trees. We continued with that procedure using the combination of each of phylogen*, clado*, taxonom*, and cladistic* with the terms seabird*, Procellariiformes, diomedeidae, procellariidae, hydrobatidae, pelecanoididae, albatross*, petrel*, shearwater*, storm-petrel*, and diving-petrel*. The combination of taxonom* and petrel* provided the reference to the study (Imber 1985) that the other searches had not found.

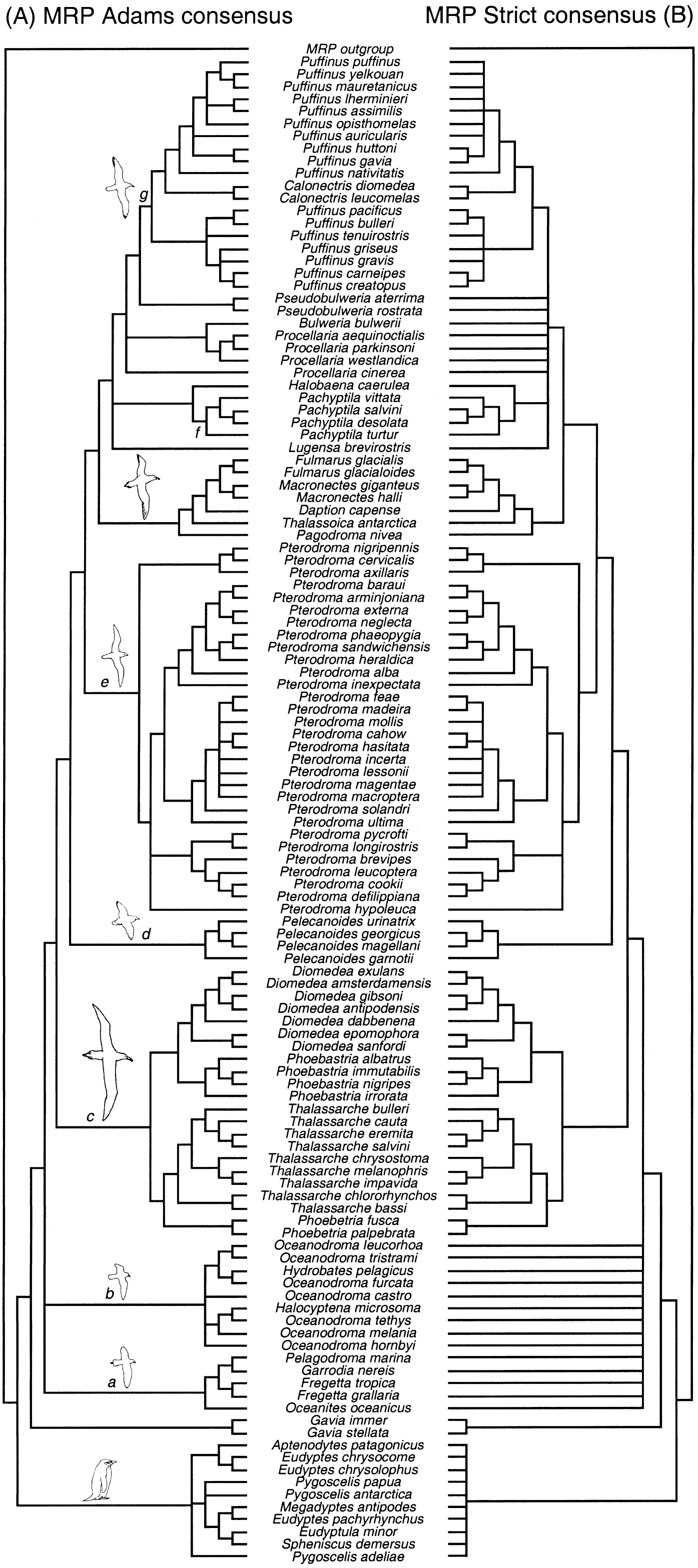
Some studies found were not used as source trees because they were by the same authors and the taxa and data used overlapped between studies (i.e. Paterson et al. 1993, 1995; Heidrich et al. 1996, Nunn et al. 1996, Heidrich et al. 1998, Nunn and Stanley 1998). In this situation, we used the more comprehensive of the studies as our source trees. So that each study had a single taxon for each species, we pruned the additional subspecies or representatives of the same species from the source trees. The source trees were generated using TREEVIEW (Page 1996) to provide NEXUS formatted files that could be translated into a MRP matrix in RADCON. Because of the reduced number of terminal taxa in some of our source trees, they are not necessarily identical to those published in the original papers, but represent abridged versions of the originals' topologies.

When the authors presented more than one topology, we attempted to use the topology that they indicated was the best estimate of their phylogeny. For example, from Austin (1996) we used what the author described as a conservative hypothesis for the evolutionary relationships of the *Puffinus* shearwaters (figure 5 in Austin 1996; Fig. 1A). Austin (1996) used a 307 base-pair (bp) fragment of cytochrome *b* to construct his phylogeny. We used the maximum parsimony (MP) tree found by Bretagnolle et al. (1998, figure 3; Fig. 1B) using a 496 bp fragment of cytochrome *b* When each species is collapsed to a single representative, the topologies of MP and maximum likelihood (ML) trees of Heidrich et al. (1998, figure 2B, C; Fig. 1C) are equivalent. Heidrich et al. (1998) used a 1,043 bp of cytochrome *b* to generate their phylogeny. Imber (1985, figure 8) presented a single topology (Fig. 1D) for the Gadfly Petrels (*Pterodroma*) based on morphological and life-history information (the characters are discussed in his paper, but there is no explicit data set or criterion for construction of the phylogeny). Nunn and Stanley (1998, figure 2) presented a consensus of their four equally weighted MP trees (Fig. 1E) which were generated from whole (1,143 bp) cytochrome *b* sequences. Paterson et al. (1995) used a combination of mitochondrial 12S ribosomal RNA gene-sequence data, isozyme information, and behavioral life-history data to construct their best estimate of phylogeny (figure 4 in Paterson et al. 1995; Fig. 1F). Finally, Sibley and Ahlquist (1990, figure 368) used UPGMA on DNA–DNA hybridization data to generate a hypothesis for the evolutionary history of some of the seabirds (Fig. 1G). The procellariiform taxa, outgroups, their common names, and which source tree(s) contain them are shown in the Appendix.

Because several source studies used mtDNA to construct their trees, it is possible to compare results of our MRP analysis with results from analysis of a character supermatrix for a large proportion of the taxa. To construct that character supermatrix, we downloaded the mtDNA data available for this group from GenBank. The majority of sequences for this combined data set come from the whole cytochrome-*b*-based study of Nunn and Stanley (1998; 90 taxa, this data set is in TREEBASE, study accession number S351 and matrix accession number M468). Some additional taxa were added using the partial cytochrome *b* sequences of Austin (1996; *Puffinus auricularis*, *Pu. gavia*, *Pu. mauretanicus*, and *P. yelkouan*) and Bretagnolle et al. (1998; *Pseudobulweria aterrima*, *Ps. rostrata*, *Pterodroma baraui*). We used the largest fragment of cytochrome *b* available to us for each taxon. Any additional taxon for which sequences originated in Heidrich et al. (1998) could not be included because their sequence does not appear to have been submitted to GenBank/EMBL. Heidrich et al. (1998) did, however, use an unpublished sequence from GenBank (*Pu. tenuirostris*, Nunn and Zino, accession number U74352) and a sequence from Nunn et al. (1996; *Procellaria cinerea*, accession number U48940) that we have included in our mtDNA data set.

Sequence data from another gene was added by including the 12S rRNA data of Paterson et al. (1995), giving the mtDNA data set a total of 1,509 characters. The aligned 12S data was obtained from the author. In addition to 12 taxa for which there are cytochrome *b* data, the Paterson et al. (1995) data set includes an additional four penguin species for which no cytochrome *b* data is available.

Analysis.—The source-tree topologies were combined and converted into a NEXUS file with a matrix suitable for parsimony analysis using the ''component coding'' option of MRP Supertree Consensus in RADCON (Thorley and Page 2000). The resulting MRP data set had 122 taxa and 188 ''characters.'' In a MRP data set, the characters represent topologies of source trees, where each node from a source tree provides one character to the matrix (Sanderson et al. 1998). That is, for each of the source trees, the taxa present in the group derived from any given node are coded as 1 for that node, whereas the taxa not in that group are coded as 0 for that node. All other taxa (those present in one or more of the other source trees, but not the one being coded) are coded as missing for that node. RADCON thus codes nodes for each of the source trees for all taxa and combines those into a single matrix. This data matrix was analyzed using PAUP* 4.0b4a (Swofford 2000). Characters were equally weighted and the starting-tree



(A) MRP Adams consensus

MRP Strict consensus (B)

for branch swapping (TBR) was obtained for the heuristic search by stepwise addition using a random-addition sequence. We defined the outgroup as the hypothetical taxon RADCON constructs for this use. This taxon (called ''MRP outgroup'') is added when the MRP matrix is constructed and all of its characters are zeros (i.e. it shares no nodes with any of the other taxa). The MRP data set is available in TREEBASE.

The combined mtDNA data set has 103 taxa. We used the partition-homogeneity test (Farris et al. 1995) to investigate whether the cytochrome-*b* and 12S rRNA sequence can be analyzed as a single data set. Only those 12 taxa for which we had data for both genes were able to be used in this analysis. We defined penguins as the outgroup for all mtDNA analyses. To find the appropriate value to relatively weight transitions and transversions, we evaluated different ML models for the data set using MODELTEST (Posoda and Crandall 1998). The best-fitting model that estimates the transition to transversion ratio was the HKY85 (Hasegawa et al. 1985) with both proportion of invariable sites and gamma shape parameter being estimated. For this model, the transition to transversion ratio was estimated at 10:1. We thus weighted transversions 10:1 over transitions for all our subsequent analyses and the starting-tree for branch swapping (TBR) was obtained for the heuristic search by stepwise addition using 100 random-addition sequences. For reasons discussed below, the same analyses were also conducted with the four penguin species from the 12S rRNA data set excluded from our mtDNA data set. For this version of the data set, we investigated the strength of the associations implied by the sequence data using bootstrap analysis (Felsenstein 1985). For the bootstrap analysis 1,000 replicates were performed using a fast-heuristic search. The combined mtDNA data set is available in TreeBASE.

RESULTS

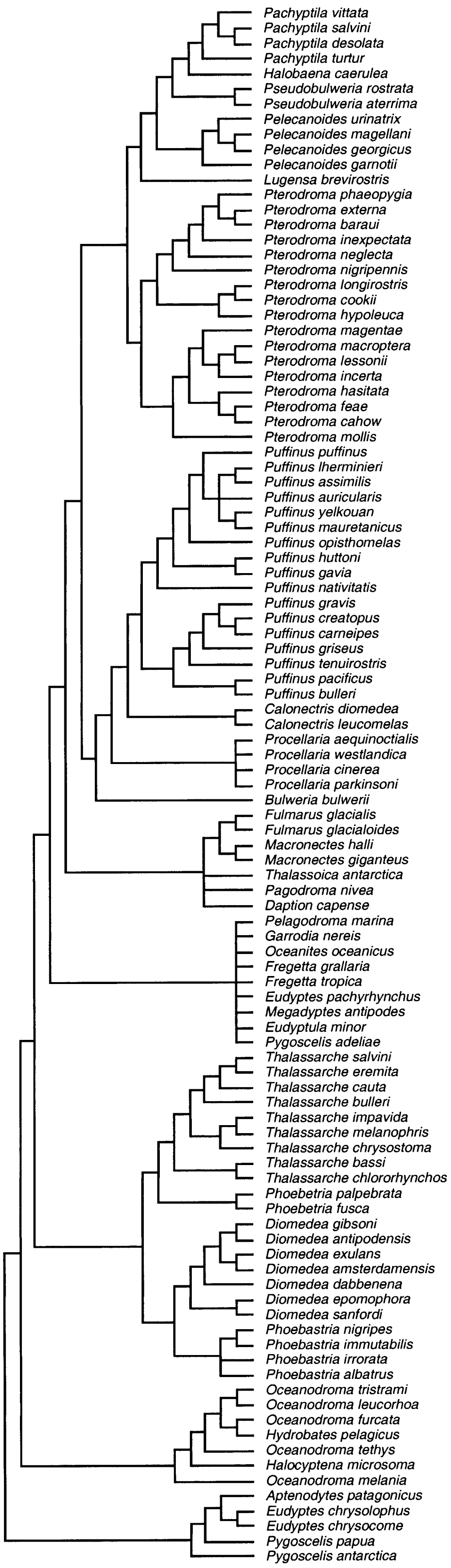
Matrix representation with parsimony.—After exploratory searches suggested that there were at least 100,000 equally parsimonious trees, we set the maximum number of trees to 10,000. We stopped 10 searches with different random-addition sequences at this number of MP trees in an attempt to ensure that we had found the optimal trees. We also ran additional exploratory searches with other random-addition sequences, but found no trees shorter than 214 steps. We compared each sample of 10,000 MP trees to ensure they each gave representative consensus trees. Nine of our 10 different random-addition sequence searches were stopped at 10,000 trees of 214 steps each, whereas the other search found 10,000 suboptimal trees. Each of the optimal sets of 10,000 MP trees were summarized as both Adams and strict-consensus trees. All of the strict-consensus trees were identical (see Fig. 2B), but there was some variation in the Adams consensus trees. Five of the Adams consensus trees were identical and this topology was the most conservative (i.e. resolved the fewest branches) of the three topologies and thus we present that topology here (see Fig. 2A).

Mitochondrial DNA.—The partition-homogeneity test showed that there was no significant difference in phylogenetic signal between the cytochrome-*b* and 12S sequences (1,000 replicates, $P = 0.69$). We ran a series of exploratory searches with different random-addition sequences and found no trees shorter than 8,636 steps. A completed heuristic search found 1,458 trees of 8,636 steps and those are summarized as a strict (Fig. 3) consensus tree.

The unusual result from the mtDNA supermatrix of the nonmonophyly of the penguins (see Fig. 3) is caused by none of the penguin species having both cytochrome *b* and 12S rRNA sequences. The four penguin species from the Paterson et al. (1995) 12S rRNA-based study grouped with one of the storm-petrel groups (see Fig. 3) within the procellariiforms. Initial exploratory equally weighted analyses had also placed these penguins within procellariiforms, and it appeared that their position within the trees may have been constrained by the unexpected result (Fig. 1F) of the 12S sequence grouping the White-faced Storm-Petrel (*Pelagodroma marina*) with the albatrosses (*Diomedea epomophora* and *Thalassarche bulleri*). In further equally weighted analyses, we removed the 12S sequence for the White-faced Storm-Petrel and found that the problematic penguins were placed more basally in the tree

←

FIG. 2. The consensus MRP trees: (A) Adams consensus, (B) strict consensus. Each of the most parsimonious MRP trees had a length of 214, a CI of 0.8785, and a RI of 0.9681. The italicized letters associated with branches of the Adams consensus tree refer to groups mentioned in the text.



(though not necessarily with the other penguins). Thus, the 12S sequence of the White-faced Storm-Petrel appears to have pulled the penguins of Paterson et al. (1995) inside of the albatrosses, and thus away from the other penguins. Because none of the penguins have sequence for both cytochrome *b* and 12S, the mtDNA supermatrix is only able to place the penguins of Paterson et al. (1995) in a general area of the tree that is loosely constrained by positions of the other taxa for which there is 12S data. Furthermore, this inability to resolve the position of these penguins generates a great number of additional optimal tree topologies. Hence, we excluded the penguins that have only 12S sequence from the subsequent supermatrix analyses.

With the penguins of Paterson et al. (1995) excluded from the analysis, we again ran a series of exploratory searches with different random-addition sequences and found no trees shorter than 8,464 steps. A completed heuristic search found 18 trees of 8,464 steps, and those are summarized as a strict (Fig. 4) consensus tree. The bootstrap values (above 50%) are shown on the branches of the strict-consensus tree. Five branches not shown in the consensus tree were supported by the bootstrap analysis: *Pygoscelis antarctica* and *Py. papua* (60% support), *Oceanodroma tethys* and *Halocyptena microsoma* (80% support), *Oc. tethys* + *Halocyptena microsoma* and *Oc. melania* (66% support), *Phoebastria albatrus* and *Ph. irrorata* (53% support), and *Pagodroma nivea* and *Thalassoica antarctica* (60% support).

DISCUSSION

As a combined summary of existing knowledge of evolutionary relationships of procellariiforms, both types of consensus tree for the MRP analysis are well resolved. Given the large number of equally optimal MRP trees generated from the combination of several source trees, both Adams and strict-consensus trees (Fig. 2) represent conservative estimates

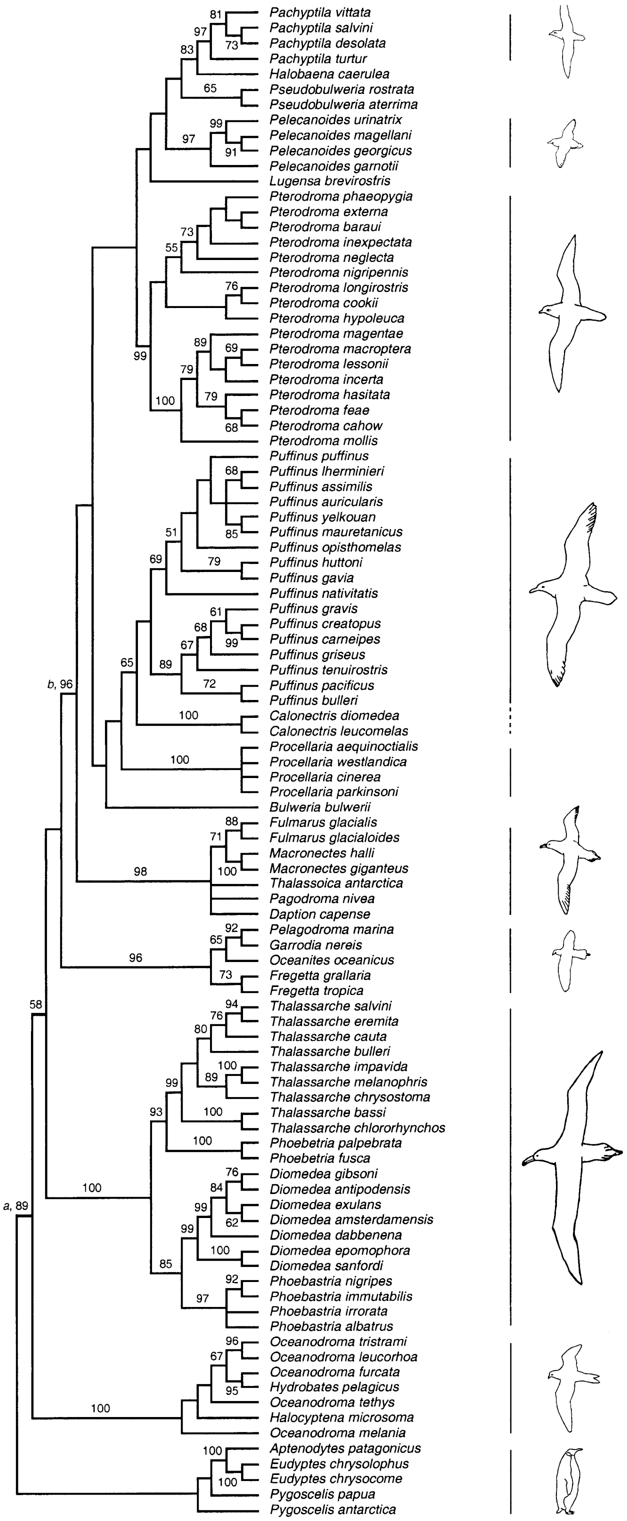
←

FIG. 3. The strict consensus of 1,458 MP trees for the mtDNA supermatrix with transversion weighted 10:1 over transitions. Each of the MP trees had a tree length of 8,636, a CI of 0.3654, and a RI of 0.8129.

of what is currently known about the phylogeny for this group. No unexpected, or what may be perceived to be wrong, groups were recovered for the procellariiforms in this analysis—although there are examples of circumstances when that may occur (e.g. see Bininda-Emonds and Bryant 1998). The Adams and strict-consensus trees (shown in Fig. 2) differ only in their level of resolution. The strict-consensus tree collapses any branch that does not exist in all 10,000 of the optimal trees, whereas the Adams consensus tree retains the general structure common to all of the MRP trees. Adams consensus places a taxon that is difficult to resolve (i.e. occurs in different places in the different optimal trees) as a polytomy at the most basal node from which it is derived in all optimal trees. Thus the polytomy indicates that the taxon is a member of that group in all optimal trees, but that it cannot be placed more precisely. Adams consensus thus retains more informative structure than the strict-consensus tree by summarizing similarity of optimal trees and indicating which groups the difficult-to-place taxa fall within. Because the Adams consensus provides a more resolved estimate of what is known about evolutionary relationships of the procellariiforms than the strict consensus tree, it provides a more useful starting point for any future comparative analyses that may be conducted with that group. One possibility, for example, is that taxa identified as difficult to place by Adams consensus could be removed from a subsequent analysis in order to obtain a more resolved phylogeny for any future comparative analyses.

Most of the expected major groups were recovered by the MRP analysis, but monophyly of storm-petrels (labeled *a* and *b* Fig. 2A) remains unresolved. This result is not surprising given that the source tree that provides almost all information about storm-petrels separated them into two groups (Fig. 1E; Nunn and Stanley 1998). The only other information about storm-petrels in the source trees supports their monophyly, but is a single node from the Sibley and Ahlquist (1990) tree (grouping *Oceanites oceanicus* with three *Oceanodroma* species). As expected groups such as the albatrosses (*c* Fig. 2A), diving petrels (*d* Fig. 2A), gadfly petrels (*e* Fig. 2A), prions (*f* Fig. 2A), and shearwaters (*g* Fig. 2A) were all monophyletic. The position of some taxa, however, is difficult to resolve. The differences between Adams and strict-consensus trees make it apparent which taxa are difficult to place. The position of the Grey Petrel (*Procellaria cineria*), for example, is difficult to resolve (Fig. 2). The source trees suggest that it could either be sister taxon to Bulwer's Petrel (*Bulweria bulweria*, Fig. 1B; Bretagnolle et al. 1998) or sister taxon to shearwaters (Fig. 1C; Heidrich et al. 1998). Contradictory evidence provided by those two source trees and lack of a source tree that includes the Grey Petrel and other *Procellaria* species makes it impossible to resolve its position. Relationships within the other *Procellaria* species and with the *Pseudobulweria* species, however, are stable and resolved (if the position of the Grey Petrel is not considered) which is better represented in the Adams consensus tree than the strict-consensus tree.

We would not, however, unconditionally recommend accepting all relationships shown in the Adams consensus tree that are unresolved in the strict-consensus tree. Relationships within the outgroup penguins in the Adams consensus tree, for example, show that combining multiple source trees can lead to incorrect relationships. The nonmonophyly of *Eudyptes* and *Pygoscelis* in the Adams consensus tree is an artifact of the degree of overlap between the penguin species in the source trees, rather than being a novel result. The strict-consensus tree shows that the structure within the penguins is difficult to resolve. Of the three source trees that include penguins, two have five species (Fig. 1E, G; Sibley and Ahlquist 1990, Nunn and Stanley 1998), whereas the other has four penguin species (our Fig. 1F; Paterson et al. 1995). Two of these source trees share no penguin species (Fig. 1E and 1F), whereas the third (Fig. 1G) contains two penguin species from each of the other two studies (Fig. 5). The lack of monophyly in *Eudyptes*, for example, is due to the three species occurring in the two source trees in which penguins do not overlap (and thus they group with other penguin taxa found in their source trees rather than with one another). Thus, some caution needs to be applied to the use of the Adams consensus tree rather than strict-consensus tree, and areas of disagreement between the two techniques should be evaluated to ensure that the increased resolution afforded by the Adams consensus tree is not spurious. Our finding, like that for storm-


81
Pachyptila vittata
97
Pachyptila salvini
73
Pachyptila desolata
83
Pachyptila turtur
Halobaena caerulea
65
Pseudobulweria rostrata
Pseudobulweria aterrima
99
Pelecanoides urinatrix
97
Pelecanoides magellani
91
Pelecanoides georgicus
Pelecanoides garnotii
Lugensa brevirostris
Pterodroma phaeopygia
Pterodroma externa
73
Pterodroma baraui
55
Pterodroma inexpectata
Pterodroma neglecta
Pterodroma nigripennis
76
Pterodroma longirostris
Pterodroma cookii
Pterodroma hypoleuca
Pterodroma magentae
89
69
Pterodroma macroptera
99
Pterodroma lessonii
79
Pterodroma incerta
79
Pterodroma hasitata
100
Pterodroma feae
68
Pterodroma cahow
Pterodroma mollis
Puffinus puffinus
68
Puffinus lherminieri
Puffinus assimilis
Puffinus auricularis
Puffinus yelkouan
51
85
Puffinus mauretanicus
Puffinus opisthomelas
69
79
Puffinus huttoni
Puffinus gavia
Puffinus nativitatis
61
Puffinus gravis
68
Puffinus creatopus
67
99
Puffinus carneipes
Puffinus griseus
65
89
Puffinus tenuirostris
b, 96
72
Puffinus pacificus
Puffinus bulleri
100
Calonectris diomedea
Calonectris leucomelas
Procellaria aequinoctialis
100
Procellaria westlandica
Procellaria cinerea
Procellaria parkinsoni
Bulweria bulwerii
88
Fulmarus glacialis
71
Fulmarus glacialoides
Macronectes halli
98
100
Macronectes giganteus
Thalassoica antarctica
Pagodroma nivea
Daption capense
92
Pelagodroma marina
65
Garrodia nereis
96
Oceanites oceanicus
73
Fregetta grallaria
Fregetta tropica
58
94
Thalassarche salvini
76
Thalassarche eremita
80
Thalassarche cauta
Thalassarche bulleri
100
Thalassarche impavida
99
89
Thalassarche melanophris
Thalassarche chrysostoma
93
100
Thalassarche bassi
Thalassarche chlororhynchos
100
Phoebetria palpebrata
Phoebetria fusca
76
Diomedea gibsoni
100
84
Diomedea antipodensis
a, 89
99
Diomedea exulans
62
Diomedea amsterdamensis
99
Diomedea dabbenena
100
Diomedea epomophora
85
Diomedea sanfordi
92
Phoebastria nigripes
97
Phoebastria immutabilis
Phoebastria irrorata
Phoebastria albatrus
96
Oceanodroma tristrami
67
Oceanodroma leucorhoa
Oceanodroma furcata
95
Hydrobates pelagicus
Oceanodroma tethys
100
Halocyptena microsoma
Oceanodroma melania
100
Aptenodytes patagonicus
Eudyptes chrysolophus
100
Eudyptes chrysocome
Pygoscelis papua
Pygoscelis antarctica

petrels, suggests that the phylogenetic relationships of penguins require more study (though we did not explicitly sample studies including penguin taxa and thus have not comprehensively surveyed the penguin literature).

It is possible to evaluate the fit between source trees and MRP consensus trees and to find whether any of the individual source trees is particularly at odds with relationships suggested by combination of source trees. In Table 1, we show both consistency index (CI) and retention index (RI) for the all MRP characters, and for those characters derived from each of the source trees individually for both the strict-consensus tree and one of the 10,000 MRP trees (chosen at random). The CI and RI of all data combined is similar to that found for each of the source trees for both the strict-consensus tree and the single MRP tree, with none of the studies appearing to be a substantially worse fit than any of the others. Because the single MRP tree is more resolved than the strict-consensus tree, its values are consistently higher. All source trees are generally consistent with the supertree and thus, the MRP trees are a good representation of all source trees.

Combining topologies of the source trees allows us to construct a phylogeny containing 122 taxa from a wide range of data types. Although 90 of those taxa and 86 of the 188 MRP characters come from a single study (Nunn and Stanley 1998), that study does not dominate the supertree construction as much as it does the supermatrix analysis. Two-thirds of Nunn and Stanley's (1998) taxa occur in one or more other source tree, thus relative positions of most of their taxa are subject to the constraint of where the other source trees place them. Nunn and Stanley's (1998) study is no more consistent with the strict-consensus MRP tree than the other source trees are, suggesting that rather than dominating the MRP analysis, it is a unifying large study that allows for greater overlap between available source trees. For the character supermatrix approach to combining the seabird source studies, the Nunn and Stanley (1998) study provides over 90% of the bases in the supermatrix (1,143 bases for 90 taxa). Thus, for this study on seabirds, the supermatrix approach is dominated more by a single large study than the supertree approach. (We do not wish to imply that being dominated by a single study is necessarily a problem, rather that any errors in the dominant study are more likely to be maintained.)

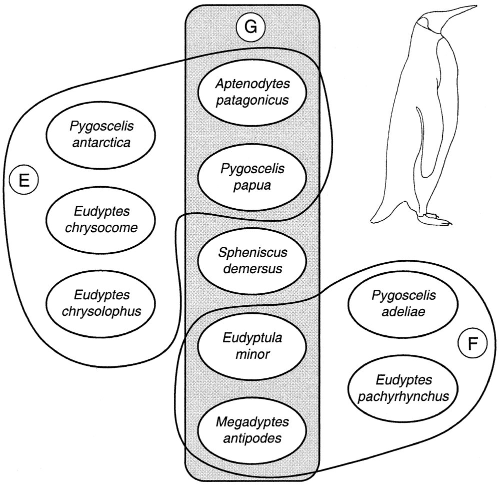


FIG. 5. The overlap between penguin species in source trees (labeled as in Fig. 1 as E, F, and G). Note that the two studies using mitochondrial genes (E and F) have no taxa in common.

Combining sequence data from different genes for multiple studies can potentially generate misleading relationships within resultant phylogenies. Supermatrix approaches may be affected by introduction of large amounts of missing data (e.g. we have 12S sequence for only 16 of 103 taxa) as well as not having overlapping sequence for some groups of taxa. Our initial supermatrix analysis included penguins from both Nunn and Stanley (1998) and Paterson et al. (1995) source trees. Those studies, however, did not contain any penguin species in common (see Fig. 5), and thus our analysis gave an extraordinary result by grouping penguins present in the Paterson et al. (1995) study as sister taxa to one of the storm-petrel subfamilies (the Hydrobatinae; see Fig. 3). As noted

---

←

FIG. 4. The strict consensus of 18 MP trees for the mtDNA supermatrix with penguins from the Paterson et al. (1995) study excluded (i.e. 99 taxa in total) and transversion weighted 10:1 over transitions. Each of the MP trees had a tree length of 8,464, a CI of 0.3640, and a RI of 0.8146.

TABLE 1. Measures of fit for the source tree data on the strict consensus of the 10,000 MRP trees and on one of the most parsimonious MRP trees (i.e. a fully resolved tree).

Tree	Number of MRP characters	Strict consensus tree		A single MRP tree	
		CI	RI	CI	RI
Combined	188	0.66	0.88	0.88	0.97
Austin (1996)	8	0.89	0.96	1.00	1.00
Bretagnolle et al. (1998)	12	0.71	0.81	0.86	0.92
Heidrich et al. (1987)	18	0.82	0.93	1.00	1.00
Imber (1985)	28	0.76	0.92	0.80	0.94
Nunn and Stanley (1998)	86	0.61	0.88	0.94	0.99
Paterson et al. (1995)	14	0.64	0.79	0.78	0.90
Sibley and Ahlquist (1990)	22	0.63	0.86	0.76	0.92

earlier, including penguins from the smaller study in the supermatrix leads to penguins from that study being placed in a part of the tree that is constrained by positions of other taxa for which there is 12S data available. Excluding penguin 12S sequences (and thus four penguin species) removes the major cause of the lack of resolution from the DNA data set (i.e. number of MP trees is reduced from 1,458 to 18). This result shows that, as with degree of overlap in source trees for supertree construction, caution needs to be taken when deciding whether taxa from different studies can validly be combined in a supermatrix approach.

With exclusion of the four penguin taxa from the Paterson et al. (1995) study, the final combined DNA data set is further dominated by the Nunn and Stanley (1998) study. As with our supertree analysis and Nunn and Stanley's (1998) study, results of the supermatrix analysis also suggest that the storm-petrels are not monophyletic. In this analysis, however, the albatrosses (see Fig. 4) separate the two subfamilies of the Hydrobatidae. Equally weighted parsimony analysis with only Nunn and Stanley's (1998) data placed both subfamilies of storm-petrels outside of albatrosses (see Fig. 1E), whereas other exploratory analyses using equal weights and the additional cytochrome-*b* sequence data placed them both nonmonophyletically inside the albatrosses. Bootstrap analysis of the final combined DNA data set weakly supports separating the two subfamilies by the albatrosses (i.e. 58% bootstrap support; Fig. 4), and none of the sequence-based analyses we performed place storm-petrels as a monophyletic group. Although storm-petrels may not be monophyletic (as noted earlier, that group requires more extensive study), it is unlikely that they are separated by albatrosses as suggested by our weighted analysis (but not by any of our other exploratory sequence analyses). This result may again be due to the 12S sequence grouping the White-faced Storm-Petrel with albatrosses, and suggests that combining genes with relatively little taxonomic overlap needs to be approached with caution. The MRP supertree, by incorporating more lines of (albeit equivocal) evidence, thus provides a more conservative estimate of relationships of storm-petrels by indicating a lack of resolution for the two subfamilies.

Apart from the node separating penguins from the procellariiforms (89%) and the node separating penguins (*a* Fig. 4), storm-petrels and albatrosses from the remaining procellariiforms (96%; *b* Fig. 4) there is a general lack of bootstrap support for the deeper branches in our weighted parsimony analysis. This finding suggests that genes for which sequence data is available are evolving too rapidly to robustly resolve relationships within that group of birds. Several groups, however, are well supported by bootstrapping. Relationships within albatrosses are generally well supported, and monophyly of several groups including storm-petrel subfamilies, the group including fulmars and giant petrels, *Procellaria* species, Gadfly petrels, diving petrels, and prions, are well supported (see Fig. 4). Thus, in this study the supertree approach provides a less-resolved phylogeny than the sequence supermatrix because of its relative conservatism (i.e. the supermatrix approach provides resolution that is not well supported).

We were unable to improve resolution of the supertree by weighting nodes from source trees according to their relative levels of support. If all source studies had been able to provide a comparable measure of support for their nodes, we could have weighted the support for each node appropriately in our matrix. Weighting nodes based on their bootstrap support, for example, may improve resolution of a supertree (see Bininda-Emonds and Bryant 1998). Because our source trees were generated in several different ways, however, they do not have comparable support values.

Given the constraints of the two approaches that have been discussed above, we suggest that the MRP supertree presented here represents the best current estimate of relationships of procellariiforms. The supertree combines information provided by more studies than are available for the supermatrix approach, and sequence data currently available appear to be evolving too rapidly to robustly resolve the more basal relationships of that group. The supertree thus provides the best available starting point for future studies addressing questions about procellariiforms from a phylogenetic perspective. The Adams consensus tree provides a better starting point for comparative studies because it is more resolved than the strict-consensus tree (but polytomies within an Adams consensus tree need to be interpreted with caution). We recommend, however, that any areas of disagreement between the Adams and strict-consensus trees should be evaluated to ensure that relationships presented in the Adams consensus tree are not artifacts of a lack of overlap between source trees.

Possible uses for such a supertree include coevolutionary studies. Paterson et al. (2000), for example, recently used a pruned version of the Paterson et al. (1995) tree (Fig. 1F) to investigate seabird and louse coevolution. Although several branches of their phylogeny are not supported by bootstrap analysis (figure 3A in Paterson et al. 2000), with the exception of the placement of *Pelagodroma marina* as sister taxon to *Diomedea epomophora,* branches in their tree are consistent with our supertree. Although a single difference, that unusual relationship potentially affects the results of Paterson et al. (2000) and could alter the conclusions they draw about the coevolutionary history of seabirds and their feather lice.

As more phylogenies become available for different groups, it will become more common for researchers to combine these phylogenies in some way in order to either address questions at a different level or in a more comprehensive way. In the foreseeable future, because data available from the individual studies are unlikely to be readily compatible (e.g. combining sequence data can be difficult when there is little overlap between taxa sequenced for different genes), supertree approaches may be more generally used than supermatrix approaches. In order for source trees to be more easily found and for supertrees to be more readily constructed, it would aid researchers if all potential source trees were submitted to a phylogeny database (such as TREEBASE). Searching GenBank and the Web of Science and other databases can lead to studies being missed as possible source trees if they could not be found (book chapters would be particularly susceptible to this). Once the literature for source trees have been found, it is possible that errors may be introduced to the supertree matrix if source trees are not regenerated accurately. To have the data set and resulting phylogenies in a phylogeny database (see Cohen et al. 1998) would eliminate those possible sources of error. As with sequence data, alignments, data sets, and the trees that are generated from them should consistently be made available through a phylogeny database.

ACKNOWLEDGMENTS

We thank Jennifer Bell for assistance with this study. Kevin Johnson, Adrian Paterson, Rick Prum, and an anonymous reviewer provided comments during preparation of the manuscript. Peter Harrison generously gave permission for us to reproduce figures from his book (Harrison 1987), which Jane Paterson kindly redrew. Adrian Paterson provided a copy of their aligned 12S data set. This project was supported by Natural Environment Research Council grant GR3/11075 and the Wolfson Foundation. M.K. was supported by a New Zealand Science and Technology Postdoctoral Fellowship.

TREEBASE is available online at www.treebase.org. Web of Science is available at www.webofscience.com.

LITERATURE CITED

AUSTIN, J. J. 1996. Molecular phylogenetics of *Puffinus* shearwaters: Preliminary evidence from mitochondrial cytochrome *b* gene sequences. Molecular Phylogenetics and Evolution 6:77–88.

BAUM, B. R. 1992. Combining trees as a way of combining datasets for phylogenetic inference, and the desirability of combining gene trees. Taxon 41:3–10.

BININDA-EMONDS, O. R. P., AND H. N. BRYANT. 1998. Properties of matrix representation with parsimony analyses. Systematic Biology 47:497–508.

BININDA-EMONDS, O. R. P., J. L. GITTLEMAN, AND A. PURVIS. 1999. Building large trees by combining phylogenetic information: A complete phylogeny of the extant Carnivora (Mammalia). Biological Reviews 74:143–175.

BRETAGNOLLE, V., C. ATTIE, AND E. PASQUET. 1998. Cytochrome-*b* evidence for validity and phylogenetic relationships of *Pseudobulweria* and *Bulweria* (Procellariidae). Auk 115:188–195.

COHEN, B. L., J. A. SHEPS, AND M. WILKINSON. 1998. Archiving molecular phylogenetic alignments as NEXUS files. Systematic Biology 47:495–496.

CRUICKSHANK, R. H., AND R. H. THOMAS. 1999. Evolution of haplodiploidy in dermanyssine mites (Acari: Mesostigmata). Evolution 53:1796–1803.

FARRIS, J. S., M. KÄLLERSJÖ, A. G. KLUGE, AND C. BULT. 1995. Constructing a significance test for incongruence. Systematic Biology 44:570–572.

FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

FRIESEN, V. L., AND D. J. ANDERSON. 1997. Phylogeny and evolution of the Sulidae (Aves: Pelecaniformes): A test of alternative modes of speciation. Molecular Phylogenetics and Evolution 7:252–260.

GRAY, R. D., AND F. M. JORDAN. 2000. Language trees support the express-train sequence of Austronesian expansion. Nature 405:1052–1055.

HARRISON, P. 1987. Seabirds of the World: A Photographic Guide. Christopher Helm, London.

HASEGAWA, M., H. KISHINO, AND T. YANO. 1985. Dating of the human–ape splitting by the molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22:160–174.

HEIDRICH, P., J. AMENGUAL, AND M. WINK. 1998. Phylogenetic relationships in Mediterranean and North Atlantic shearwaters (Aves: Procellariidae) based on nucleotide sequences of mtDNA. Biochemical Systematics and Ecology 26:145–170.

HEIDRICH, P., D. RISTOW, AND M. WINK. 1996. Molecular differentiation of Cory's and Manx shearwaters (*Calonectris diomedea*, *Puffinus puffinus*, *Puffinus yelkouan*) and the Herring Gull complex (*Larus argentatus*, *L. fuscus*, *L. cachinnans*). Journal für Ornithologie 137:281–294.

IMBER, M. J. 1985. Origins, phylogeny and taxonomy of the gadfly petrels *Pterodroma* spp. Ibis 127: 197–229.

JOHNSON, K. P., AND M. D. SORENSON. 1998. Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). Molecular Phylogenetics and Evolution 10:82–94.

KENNEDY, M., A. M. PATERSON, J. C. MORALES, S. PARSONS, A. P. WINNINGTON, AND H. G. SPENCER. 1999. The long and short of it: Branch lengths and the problem of placing the New Zealand short-tailed bat, *Mystacina*. Molecular Phylogenetics and Evolution 13:405–416.

KENNEDY, M., AND H. G. SPENCER. 2000. Phylogeny, biogeography, and taxonomy of Australasian teals. Auk 117:154–163.

KENNEDY, M., H. G. SPENCER, AND R. D. GRAY. 1996. Hop, step and gape: Do the social displays of the Pelecaniformes reflect phylogeny? Animal Behaviour 51:273–291. (An erratum to this paper was published in Animal Behaviour, 1996, 51: 1197.)

MADDISON, D. R., D. L. SWOFFORD, AND W. P. MADDISON. 1997. NEXUS: An extensible file format for systematic information. Systematic Biology 46:590–621.

MARCHANT, S., AND P. J. HIGGINS. 1990. Handbook of Australian, New Zealand and Antarctic Birds, vol. 1. Ratites to Ducks. Oxford University Press, Oxford.

NUNN, G. B., J. COOPER, P. JOUVENTIN, C. J. R. ROBERTSON, AND G. G. ROBERTSON. 1996. Evolutionary relationships among extant albatrosses (Procellariiformes: Diomedeidae) established from complete cytochrome-*b* gene sequences. Auk 113:784–801.

NUNN, G. B., AND S. E. STANLEY. 1998. Body size effects and rates of cytochrome *b* evolution in tube-nosed seabirds. Molecular Biology and Evolution 15:1360–1371.

PAGE, R. D. M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. Computer Applications in the Biosciences 12:357–358.

PATERSON, A. M., R. D. GRAY, AND G. P. WALLIS. 1993. Parasites, petrels and penguins: Does louse phylogeny reflect seabird phylogeny? International Journal of Parasitology 23:515–526.

PATERSON, A. M., G. P. WALLIS, AND R. D. GRAY. 1995. Penguins, petrels and parsimony: Does cladistic analysis of behavior reflect seabird phylogeny? Evolution 49:974–989.

PATERSON, A. M., G. P. WALLIS, L. J. WALLIS, AND R. D. GRAY. 2000. Seabird and louse coevolution: Complex histories revealed by 12S rRNA sequences and reconciliation analyses. Systematic Biology 49:383–399.

POSODA, D., AND K. A. CRANDALL. 1998. MODELTEST: Testing the model of DNA substitution. Bioinformatics 14:817–818.

RAGAN, M. A. 1992. Phylogenetic inference based on matrix representation of trees. Molecular Phylogenetics and Evolution 1:53–58.

SANDERSON, M. J., A. PURVIS, AND C. HENZE. 1998. Phylogenetic supertrees: Assembling the trees of life. Trends in Ecology and Evolution 13:105–109.

SIBLEY, C. G., AND J. E. AHLQUIST. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven, Connecticut.

SIBLEY, C. G., AND B. L. MUNROE. 1990. Distribution and Taxonomy of the Birds of the World. Yale University Press, New Haven, Connecticut.

STEEL, M., A. W. M. DRESS, AND S. BOCKER. 2000. Simple but fundamental limitations on supertree and consensus tree methods. Systematic Biology 49:363–368.

SWOFFORD, D. L. 2000. PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4. Sinauer Associates, Sunderland, Massachusetts.

THORLEY, J. L., AND R. D. M. PAGE. 2000. RADCON: Phylogenetic tree comparison and consensus. Bioinformatics 16:486–487.

ZYSKOWSKI, K., AND R. O. PRUM. 1999. Phylogenetic analysis of the nest architecture of neotropical ovenbirds (Furnariidae). Auk 116:891–911.

Associate Editor: R. Prum

APPENDIX. Classification for the taxa used in this study (adapted from Sibley and Munroe 1990).

Procellariidae (petrels and allies)	Common names	Source study
Macronectes giganteus	Antarctic Giant-Petrel	B, H, N, S[1]
Macronectes halli	Hall's Giant-Petrel	N
Fulmarus glacialis	Northern Fulmar	A, B, H, N
Fulmarus glacialoides	Southern Fulmar	H, N
Thalassoica antarctica	Antarctic Petrel	N, S[1]
Daption capense	Cape Petrel	N, P
Pagodroma nivea	Snow Petrel	B, N
Lugensa[2] *brevirostris*	Kerguelen Petrel	N, S
Pseudobulweria[2] *aterrima*	Mascarene Petrel	B
Pseudobulweria[2] *rostrata*	Tahiti Petrel	B
Pterodroma axillaris	Chatham Islands Petrel	I
Pterodroma nigripennis	Black-winged Petrel	B, I, N
Pterodroma cervicalis	White-necked Petrel	I
Pterodroma inexpectata	Mottled Petrel	I, N, P
Pterodroma hypoleuca	Bonin Petrel	I, N, S
Pterodroma leucoptera	Gould's Petrel	I
Pterodroma cookii	Cook's Petrel	I, N, P, S
Pterodroma pycrofti	Pycroft's Petrel	I
Pterodroma brevipes	Collared Petrel	I
Pterodroma defilippiana	Defilippe's Petrel	I
Pterodroma longirostris	Stejneger's Petrel	I, N
Pterodroma alba	Phoenix Petrel	I
Pterodroma arminjoniana	Herald Petrel	I, S
Pterodroma heraldica[3,4]		I
Pterodroma sandwichensis	Hawaiian Petrel	I
Pterodroma phaeopygia	Galapagos Petrel	I, N
Pterodroma neglecta	Kermadec Petrel	I, N
Pterodroma externa	Juan Fernandez Petrel	I, N
Pterodroma baraui	Barau's Petrel	B, I
Pterodroma ultima	Murphy's Petrel	I
Pterodroma solandri	Providence Petrel	I
Pterodroma macroptera	Great-winged Petrel	I, N
Pterodroma magentae	Magenta Petrel	I, N
Pterodroma lessonii	White-headed Petrel	I, N
Pterodroma madeira	Madeira Petrel	I
Pterodroma feae	Cape Verde Petrel	I, N
Pterodroma mollis	Soft-plumaged Petrel	I, N
Pterodroma incerta	Atlantic Petrel	I, N
Pterodroma cahow	Bermuda Petrel	I, N
Pterodroma hasitata	Black-capped Petrel	I, N, S
Halobaena caerulea	Blue Petrel	N
Pachyptila vittata	Broad-billed Prion	N, P, S[1]
Pachyptila salvini	Medium-billed Prion	N
Pachyptila desolata	Antarctic Prion	N
Pachyptila turtur	Fairy Prion	N, P
Bulweria bulwerii	Bulwer's Petrel	B, H, N
Procellaria aequinoctialis	White-chinned Petrel	N
Procellaria parkinsoni	Black Petrel	N
Procellaria westlandica	Westland Petrel	N, P
Procellaria cinerea	Grey Petrel	B, H
Calonectris diomedea	Cory's Shearwater	B, H, N, S[1]
Calonectris leucomelas	Streaked Shearwater	N
Puffinus pacificus	Wedge-tailed Shearwater	A, B, N
Puffinus bulleri	Buller's Shearwater	A, N
Puffinus carneipes	Flesh-footed Shearwater	A, N
Puffinus creatopus	Pink-footed Shearwater	A, N
Puffinus gravis	Great Shearwater	A, N
Puffinus griseus	Sooty Shearwater	A, N, P, S
Puffinus tenuirostris	Short-tailed Shearwater	A

APPENDIX. Continued.

Procellariidae (petrels and allies)	Common names	Source study
Puffinus nativitatis	Christmas Island Shearwater	A, N
Puffinus puffinus	Manx Shearwater	A, H, N
Puffinus yelkouan	Mediterranean Shearwater	A, H
Puffinus mauretanicus[1]		A, H
Puffinus auricularis	Townsend's Shearwater	A
Puffinus opisthomelas	Black-vented Shearwater	N
Puffinus gavia	Fulttering Shearwater	A
Puffinus huttoni	Hutton's Shearwater	A, N, P
Puffinus lherminieri	Audubon's Shearwater	A, N
Puffinus assimilis	Little Shearwater	A, H, N
Pelecanoides garnotii	Peruvian Diving-Petrel	N
Pelecanoides magellani	Magellanic Diving-Petrel	N
Pelecanoides georgicus	South Georgia Diving-Petrel	N, P, S
Pelecanoides urinatrix	Common Diving-Petrel	B, N, S
Diomedea exulans	Wandering Albatross	B, N
Diomedea amsterdamensis	Amsterdam Island Albatross	H, N
Diomedea epomophora	Royal Albatross	H, N, P, S
Diomedea dabbenena[3]		N
Diomedea gibsoni[3]		N
Diomedea antipodensis[3]		N
Diomedea sanfordi[3]		N
Phoebastria[5] *irrorata*	Waved Albatross	N, S
Phoebastria[5] *albatrus*	Short-tailed Albatross	H, N
Phoebastria[5] *nigripes*	Black-footed Albatross	N, S
Phoebastria[5] *immutabilis*	Laysan Albatross	H, N, S
Thalassarche[5] *melanophris*	Black-browed Albatross	N
Thalassarche[5] *cauta*	Shy Albatross	N, S
Thalassarche[5] *chrysostoma*	Grey-headed Albatross	H, N, S
Thalassarche[5] *chlororhynchos*	Yellow-nosed Albatross	N
Thalassarche[5] *bulleri*	Buller's Albatross	H, N, P, S
Thalassarche bassi[3]		N
Thalassarche eremita[3]		N
Thalassarche salvini[3]		N
Thalassarche impavida[3]		N
Phoebetria fusca	Sooty Albatross	H, N
Phoebetria palpebrata	Light-mantled Albatross	H, N
Oceanites oceanicus	Wilson's Storm-Petrel	N, S[1]
Garrodia nereis	Grey-backed Storm-Petrel	N
Pelagodroma marina	White-faced Storm-Petrel	N, P
Fregetta tropica	Black-bellied Storm-Petrel	N
Fregetta grallaria	White-bellied Storm-Petrel	N
Hydrobates pelagicus	European Storm-Petrel	H, N
Halocyptena[6] *microsoma*	Least Storm-Petrel	N
Oceanodroma tethys	Wedge-rumped Storm-Petrel	N
Oceanodroma castro	Band-rumped Storm-Petrel	H
Oceanodroma leucorhoa	Leach's Storm-Petrel	B, N, S
Oceanodroma tristrami	Tristram's Storm-Petrel	N
Oceanodroma melania	Black Storm-Petrel	N
Oceanodroma hornbyi	Ringed Storm-Petrel	S
Oceanodroma furcata	Fork-tailed Storm-Petrel	N, S
	Gaviidae (Loons)	
Gavia stellata	Red-throated Loon	S
Gavia immer	Common Loon	S
	Spheniscidae (Penguins)	
Aptenodytes patagonicus	King Penguin	N, S[1]
Pygoscelis papua	Gentoo Penguin	N, S[1]
Pygoscelis adeliae	Adelie Penguin	P
Pygoscelis antarctica	Chinstrap Penguin	N
Eudyptes chrysocome	Rockhopper Penguin	N

APPENDIX. Continued.

Procellariidae (petrels and allies)	Common names	Source study
Eudyptes pachyrhynchus	Fiordland Penguin	P
Eudyptes chrysolophus	Macaroni Penguin	N
Megadyptes antipodes	Yellow-eyed Penguin	P, S[1]
Eudyptula minor	Little Penguin	P, S
Spheniscus demersus	Jackass Penguin	S
Taxa in Sibley and Munroe's (1990) classification not included in any of the source studies		
Pterodroma becki	Beck's Petrel	
Pterodroma macgillivrayi	Fiji Petrel	
Pachyptila belcheri	Slender-billed Prion	
Pachyptila crassirostris	Fulmar Prion	
Bulweria fallax	Jouanin's Petrel	
Puffinus persicus	Persian Shearwater	
Puffinus bannermani	Bannerman's Shearwater	
Puffinus heinrothi	Heinroth's Shearwater	
Oceanites gracilis	White-vented Storm-Petrel	
Oceanodroma monorhis	Swinhoe's Storm-Petrel	
Oceanodroma macrodactyla	Guadalupe Storm-Petrel	
Nesofregetta fuliginosa	Polynesian Storm-Petrel	
Oceanodroma markhami	Markham's Storm-Petrel	
Oceanodroma matsudairae	Matsudaira's Storm-Petrel	
Oceanodroma homochroa	Ashy Storm-Petrel	
Gavia arctica	Arctic Loon	
Gavia pacifica	Pacific Loon	
Gavia adamsii	Yellow-billed Loon	
Aptenodytes forsteri	Emperor Penguin	
Eudyptes robustus	Snares Penguin	
Eudyptes sclateri	Erect-crested Penguin	
Eudyptes schlegeli	Royal Penguin	
Spheniscus humboldti	Humboldt Penguin	
Spheniscus magellanicus	Magellanic Penguin	
Spheniscus mendiculus	Galapagos Penguin	

[1] This species name was given to the genus in Sibley and Ahlquist's (1990) tree.

[2] *Pterodroma* in Sibley and Munroe (1990).

[3] Not classified in Sibley and Munroe (1990).

[4] *Pterodroma heraldica* was not classified in Sibley and Munroe (1990). It is sometimes treated as a subspecies of *Pterodroma arminjoniana*, hence Sibley and Munroe (1990) gave it the common name of the Herald Petrel, which can be substituted for the Trinidade Petrel when *Pterodroma heraldica* is classified and given that common name (see Marchant and Higgins 1990).

[5] *Diomedea* in Sibley and Munroe (1990).

[6] *Oceanodroma* in Sibley and Munroe (1990).