# Phylogenetic Relationships Within the Alcidae (Charadriiformes: Aves) Inferred from Total Molecular Evidence

*V. L. Friesen,* A. J. Baker,* and J. F. Piatt†*

*Department of Ornithology, Royal Ontario Museum and Department of Zoology, University of Toronto; and †National Biological Survey, Alaska Fish and Wildlife Research Center, United States Department of the Interior

The Alcidae is a unique assemblage of Northern Hemisphere seabirds that forage by "flying" underwater. Despite obvious affinities among the species, their evolutionary relationships are unclear. We analyzed nucleotide sequences of 1,045 base pairs of the mitochondrial cytochrome *b* gene and allelic profiles for 37 allozyme loci in all 22 extant species. Trees were constructed on independent and combined data sets using maximum parsimony and distance methods that correct for superimposed changes. Alternative methods of analysis produced only minor differences in relationships that were supported strongly by bootstrapping or standard error tests. Combining sequence and allozyme data into a single analysis provided the greatest number of relationships receiving strong support. Addition of published morphological and ecological data did not improve support for any additional relationships. All analyses grouped species into six distinct lineages: (1) the dovekie (*Alle alle*) and auks, (2) guillemots, (3) brachyramphine murrelets, (4) synthliboramphine murrelets, (5) true auklets, and (6) the rhinoceros auklet (*Cerorhinca monocerata*) and puffins. The two murres (genus *Uria*) were sister taxa, and the black guillemot (*Cepphus grylle*) was basal to the other guillemots. The Asian subspecies of the marbled murrelet (*Brachyramphus marmoratus perdix*) was the most divergent brachyramphine murrelet, and two distinct lineages occurred within the synthliboramphine murrelets. Cassin's auklet (*Ptychoramphus aleuticus*) and the rhinoceros auklet were basal to the other auklets and puffins, respectively, and the Atlantic (*Fratercula arctica*) and horned (*Fratercula corniculata*) puffins were sister taxa. Several relationships among tribes, among the dovekie and auks, and among the auklets could not be resolved but resembled "star" phylogenies indicative of adaptive radiations at different depths within the trees.

## Introduction

Accurate information about evolutionary relationships among organisms is essential for addressing many biological problems, yet insights into these relationships can be difficult to obtain. For example, phylogenies based on morphological data may be affected by selection, epistasis, and polygenic effects, whereas hypotheses based on molecular characters may be obscured by superimposed changes. Other difficulties in reconstructing phylogenies include uncertainties surrounding the accuracy of alternative tree-building algorithms (e.g., Stewart 1993; Hillis, Huelsenbeck, and Cunningham 1994; Huelsenbeck 1995) and the desirability of combining qualitatively different data sets (e.g., Helm-Bychowski and Cracraft 1993; Jones, Kluge, and Wolf 1993; Hillis 1995; Miyamoto and Fitch 1995). Many of these problems have been addressed on a theoretical basis (e.g., Huelsenbeck and Hillis 1993; Hillis, Huelsenbeck, and Cunningham 1994; Huelsenbeck 1995), but few empirical comparisons have been performed.

The Alcidae is an excellent group for assessing the merits of alternative methods of phylogeny reconstruction. The family constitutes a unique group of 22 species of pursuit-diving seabirds that are currently classified into seven tribes and placed within the order Charadriiformes (American Ornithologists' Union 1983; see table 1 for scientific names and classifications). Phylogenetic relationships among many species are unclear, and several hypotheses have been proposed (reviewed in Strauch 1985; Sibley and Ahlquist 1990). In 1985, Strauch published a compatability analysis of 33 morphological and ecological characters from all Recent species. Watada et al. (1987) constructed a protein tree based on genetic distances at 24 allozyme loci, but their study included only 12 species. Reliabilities of the phylogenetic hypotheses were not evaluated in either of these studies. Moum et al. (1994) compared nucleotide sequences of two mitochondrial genes, but they lacked four taxa and were unable to construct a fully bifurcating tree.

In the present study, our objectives were to clarify relationships among the alcids and to examine empirically the effects of character weighting and combination of independent data sets on the resolution of phylogenetic relationships. We analyzed nucleotide sequences of 1,045 base pairs (bp) of the mitochondrial cytochrome *b* gene and allelic profiles for 37 allozyme loci from representatives of all extant species of the family. Published data on morphology and ecology (Strauch 1985) also were combined with genetic data (present study and Moum et al. 1994) to examine whether additional characters would help to resolve phylogenetic relationships.

Key words: Alcidae, allozymes, congruence, cytochrome *b*, evolution, marbled murrelet, phylogenetics, total evidence.

Present address and address for correspondence: Vicki Friesen, Department of Biology, Queen's University, Kingston, Ontario K7L 3N6, Canada. Email: friesenv@biology.queensu.ca.

*Mol. Biol. Evol.* 13(2):359–367. 1996
© 1996 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

**Table 1**
**Samples Used for Phylogenetic Analysis of the Alcidae**

| Tribe | Scientific Name | Common Name | Number[a] | Sampling Location |
|---|---|---|---|---|
| Allini (dovekie) | *Alle alle alle* | Dovekie | 1/1 | Newfoundland |
| | *A. a. polaris* | | 1/0 | Alaska |
| Alcini (auks) | *Uria aalge inornata* | Common murre | 2/6 | Alaska |
| | *U. a. aalge* | | 0/3 | Newfoundland |
| | *U. lomvia arra* | Thick-billed murre | 2/8 | Alaska |
| | *Alca torda torda* | Razorbill | 2/0 | Newfoundland, Norway |
| | | | 0/3 | Quebec |
| Cephini (guillemots) | *Cepphus grylle arcticus* | Black guillemot | 2/2 | Newfoundland |
| | *C. columba adianta* | Pigeon guillemot | 2/5 | Alaska |
| | *C. carbo* | Spectacled guillemot | 2/1 | Japan |
| Brachyramphini (brachyramphine murrelets) | *Brachyramphus marmoratus marmoratus* | Marbled murrelet | 2/5 | Alaska |
| | *B. m. perdix* | Long-billed murrelet | 2/4 | Japan |
| | *B. brevirostris* | Kittlitz's murrelet | 2/4 | Alaska |
| Synthliboramphini (synthliboramphine murrelets) | *Synthliboramphus hypoleucus scrippsi* | Xantus' murrelet | 2/0 | California |
| | *S. craveri* | Craveri's murrelet | 2/2 | California |
| | *S. antiquus* | Ancient murrelet | 2/5 | Alaska |
| | *S. wumizusume* | Japanese murrelet | 2/2 | Japan |
| Aethini (true auklets) | *Ptychoramphus aleuticus aleuticus* | Cassin's auklet | 2/5 | Alaska |
| | *Cyclorrhynchus psittacula* | Parakeet auklet | 2/5 | Alaska |
| | *Aethia pusilla* | Least auklet | 2/3 | Alaska |
| | *A. pygmaea* | Whiskered auklet | 2/2 | Alaska |
| | *A. cristatella* | Crested auklet | 2/5 | Alaska |
| Fraterculini (puffins) | *Cerorhinca monocerata* | Rhinoceros auklet | 2/2 | Alaska |
| | *Fratercula cirrhata* | Tufted puffin | 2/5 | Alaska |
| | *F. arctica arctica* | Atlantic puffin | 2/3 | Newfoundland |
| | *F. corniculata* | Horned puffin | 2/5 | Alaska |

[a] The first number represents the number of samples used for sequence analysis of cytochrome *b*; the second number represents the number of samples used for allozyme electrophoresis.

## Materials and Methods

### Molecular Analyses

Fresh tissue (heart, liver, and/or muscle) was collected from representatives of most extant species of alcids (table 1). The only exception was the Xantus' murrelet, for which a small piece of skin was cut from each of two museum specimens (Royal Ontario Museum collections). Tissue also was obtained from two red-billed gulls (*Larus novaehollandiae,* Laridae, Charadriiformes) and two redwinged pratincoles (*Glareola pratincola,* Glareolidae, Charadriiformes) for outgroup rooting. Whenever possible, specimens were obtained from geographically distant sites to obtain a broad representation of intraspecific variation.

DNA was extracted from tissue by protease digestion and phenol extraction, and a 1,045-bp fragment of cytochrome *b* was amplified using the primers L14841 (Kocher et al. 1989) and H16065 (located in the t-RNA$^{thr}$; unpublished data) and standard protocols (Kocher et al. 1989). Amplification products were subjected to electrophoresis in 2% agarose gels, and the amplified DNA was excised and purified using GeneClean® II kits (Bio-101, Vista, California) according to the manufacturer's directions. Double-stranded sequencing was conducted on purified DNA using Sequenase® kits (U.S. Biochemicals, Cleveland, Ohio) with one of the two amplification primers or four internal primers (unpublished data). Sequences were confirmed both by sequencing complementary strands (approximately 40% of base pairs) and by obtaining complete sequences for at least two representatives of each species (table 1).

Electrophoretic variation at 37 presumptive nuclear loci (table 2) was analyzed in all taxa except the Xantus' murrelet and pratincole (table 1) using standard electrophoretic conditions (Baker, Lynch, and Edwards 1985).

### Phylogenetic Analyses of Sequence Data

A neighbor-joining tree (Saitou and Nei 1987) was constructed from genetic distances among cytochrome *b* sequences using Kimura's (1980) two parameter correction for multiple hits in MEGA (version 1.0; Kumar, Tamura, and Nei 1993), with the gull and pratincole designated as outgroup taxa. Support for relationships was assessed both through bootstrap analysis and through standard error tests (Rzhetsky and Nei 1992, 1993). Although support of less than 95% is not considered significant for most statistical tests, bootstrap analysis is thought to be highly conservative (Jones, Kluge, and Wolf 1993; Rodrigo 1993). In the present study, rela-

**Table 2**
**Most Common Allele for 23 Variable Allozyme Loci for 22 Alcids and a Gull**

| Taxon | Locus[a] | | | | |
|---|---|---|---|---|---|
| Dovekie | ABBEJ | ACAAA | ABAAE | BIAAD | AAB |
| Common murre | ABBBA | ABAAA | ABBBD | BBAAD | ABB |
| Thick-billed murre | ABBBA | AIAAA | ABBAD | BBAAC | AAB |
| Razorbill | ABBBA | AFAAA | ABABC | GHAAJ | AAB |
| Black guillemot | ABAAF | ADAAA | ABCBD | BCBAA | AAC |
| Pigeon guillemot | ABAAA | ADAAA | ABCBD | BCBAA | AAC |
| Spectacled guillemot | ABAAA | ADAAA | ABCBD | BCBAA | AAC |
| Marbled murrelet | ABAAC | AJAEA | ABDBD | CAAAH | BAC |
| Long-billed murrelet | ABAAC | AJAAA | DBGBD | CJGAH | BAC |
| Kittlitz's murrelet | ABAAC | AKAFA | ABABD | CAFAG | AAC |
| Craveri's murrelet | ACAAI | AJAGA | ABAFB | BAAAG | AAC |
| Ancient murrelet | ACACI | AJAGB | ABEFD | BEAAF | AAC |
| Japanese murrelet | ACACI | AJAGA | CBEHF | BEHAH | AAC |
| Cassin's auklet | AAAAB | AEBFA | AAAED | DADAA | AED |
| Parakeet auklet | AEAAB | AHADA | BADCD | BBDAA | ACD |
| Least auklet | BAAAB | AEBAA | AAAGD | BADAA | ADD |
| Whiskered auklet | AAAAB | AEBAA | AAAED | BADAA | CDD |
| Crested auklet | AAAAB | BEBHA | AAAED | BADAA | CFD |
| Rhinoceros auklet | AAAAB | AEBAA | AAAED | BADAA | CDD |
| Tufted puffin | AEAAA | AHADA | BADDD | BBDAA | ACC |
| Atlantic puffin | ADAAA | AGADA | BADCD | BBAAB | ACD |
| Horned puffin | AEAAA | AGADA | BADCD | BBAAB | ACD |
| Red-billed gull | AABAG | AAAAC | AAAAD | AAAAA | AAA |

NOTES.—Monomorphic loci include *Ak-2* (enzyme commission number 2.7.4.3), *Ca-1,* (4.2.1.1), *Ca-2, Eap-1* (3.1.3.2), *Eap-2, Est-4* (3.1.1.1), *Got-2* (2.6.1.1), *GP-1, GP-2, Idh-2* (1.1.1.42), *Ldh-1* (1.1.1.27), *Mdh-1* (1.1.1.37), *Pep-D* (3.4.11), and *Sod-2* (1.15.1.1).

[a] Loci in order of presentation: *Acon* (4.2.1.3), *Ada* (3.5.4.4), *Ak-1, Ck-2* (2.7.3.2), *Est-1, Gda* (3.5.4.3), *Glud* (1.4.1.3), *Got-1, GP-3,* αGpd (1.1.1.8), *Idh-1, Ldh-2, Mdh-2, Me* (1.1.1.40), *Mpi* (5.3.1.8), *Pep-A, Pep-B, 6-Pgd* (1.1.1.44), *Pgi* (5.3.1.9), *Pgm-1* (2.7.5.1), *Pgm-2, Sdh* (1.1.1.14), and *Sod-1.*

tionships with bootstrap support ≥90% always were associated with values ≥95% in standard error tests (see Results). Such groups were considered to be "strongly supported" or "clearly indicated."

Maximum parsimony analysis of sequence data was conducted on PAUP (version 3.1.1; Swofford and Begle 1993) using the heuristic search algorithm, with the gull and pratincole designated as outgroup taxa. Starting trees were generated using the "closest" addition option, and "tree bisection–reconstruction" was used for branch-swapping. Ten replicates were performed using the "random" addition option to check for more parsimonious trees. Reliabilities of phylogenetic relationships were evaluated using both bootstrap analysis (Felsenstein 1985) and Bremer indices (Bremer 1988). Bremer indices were calculated by searching for trees up to 10 steps longer than the most parsimonious tree and assigning the value of $S$ to nodes that collapsed in the consensus of trees $\leq S$ steps longer than the most parsimonious tree.

Because nucleotide sequences are subject to superimposed changes that may obscure the phylogenetic signal, various methods have been proposed to minimize the influence of homoplasies on deep branches. For example, classes of substitutions that occur with low frequencies (e.g., G ↔ T transversions) are less likely to be affected by multiple hits than are those that occur with high frequencies (e.g., C ↔ T transitions), so may be assigned a greater weight during phylogenetic reconstruction. Edwards, Arctander, and Wilson (1991) weighted transversions 20 times transitions, and Knight and Mindell (1993) suggested assigning weights that are inversely proportional to the observed ("empirically derived") frequencies for each class of substitution. In the present study, we compared results of several weighting schemes. For the simplest analysis, all substitutions were given equal weights. We then tried weighting transversions four times transitions (the inverse of the average transition bias in the present data set; see Results) or 20 times transitions (Edwards, Arctander, and Wilson 1991). Finally, we used an empirically derived scheme based on estimates of the instantaneous frequencies of each class of substitution. Specifically, we assumed that G ↔ T transversions (the rarest substitutions) accumulate approximately linearly with time. For each of the six classes of substitution we then conducted a regression of $\log(N + 1)$ (where $N$ is the number of substitutions) against $\log(\text{G} \leftrightarrow \text{T} + 1)$. (The value 1 was added to $N$ to prevent class frequencies of 0, and data were log-transformed to correct for saturation.) Instantaneous frequencies were estimated as $(10^a - 1)$, where $a$ is the intercept of the regression equation. Each class of substitution was given a weight inversely proportional to its instantaneous frequency (table 3). Several other methods of calculating empirical weights also were investigated (e.g., following the above approach but using the most common type of transversion as the independent variable or using the inverse of the mean frequency of each class of substitution for all pairs of taxa with one G ↔ T transversion): results were essentially the same for all methods.

**Table 3**
**Observed Frequencies of Nucleotide Substitutions (above Diagonal) and Estimated Instantaneous Frequencies of Substitutions (below Diagonal)**

| | Base | | | |
|---|---|---|---|---|
| Base | A | C | G | T |
| A | — | 20 | 20 | 5.3 |
| C | 13 | — | 1.6 | 66 |
| G | 19 | 1 | — | 0.8 |
| T | 3 | 60 | 1 | — |

## Phylogenetic Analyses of Allozyme Data

Maximum likelihood analysis of allozyme data was conducted on allele frequencies (Felsenstein 1981) using

PHYLIP (version 3.41; Felsenstein 1989). Parsimony analysis was performed by treating loci as unordered characters, with the character state defined as the most common allele for each taxon (Mindell, Sites, and Graur 1989).

### Analyses of Taxonomic Congruence and Total Molecular Evidence

Several authors have argued that, when strong support for a phylogenetic relationship cannot be obtained from a single analysis, the reliability of the relationship can be evaluated from its presence or absence in alternative analyses (e.g., Helm-Bychowski and Cracraft 1993; Hillis 1995). We compared results of analyses of cytochrome *b* and allozyme data to find relationships that occurred in all topologies, whether or not they were supported strongly in either analysis.

To increase the number of phylogenetically informative characters and thus maximize resolution, several authors have suggested combining data from several sources into a single phylogenetic analysis (e.g., Eernisse and Kluge 1993; Helm-Bychowski and Cracraft 1993; Jones, Kluge, and Wolf 1993). The rationale is that, despite differences in character types, each character provides equal information about phylogenetic relationships. We therefore conducted maximum parsimony analysis on pooled data from cytochrome *b* and allozymes. Searches were performed as for independent data sets except that transversions and allozyme characters were given equal weights, and transitions were allotted one-fourth the weight of other characters to reduce the influence of superimposed changes on deeper branches. In a final set of parsimony analyses, we combined the present molecular data with nucleotide sequence data from Moum et al. (1994) and/or morphological and ecological data from Strauch (1985), using the alternative weighting schemes outlined above.

## Results and Discussion

### Variation in Nucleotide Sequences and Allozymes

No indels were found within cytochrome *b* sequences of the 25 taxa analyzed in the present study. Base composition was strongly biased: C occurred at a mean (±SE) of 33.5% ± 0.15% of sites, whereas G occurred at only 12.7% ± 0.09% of sites; A and T were represented roughly equally, occurring at 28.4% ± 0.15% and 25.6% ± 0.18% of sites, respectively. These biases are similar to those found in other species of birds (e.g., Kocher et al. 1989; Edwards, Arctander, and Wilson 1991; Kornegay et al. 1993; unpublished data). Of 1,045 nucleotide sites, 384 were variable. Variable regions corresponded roughly to transmembrane domains of cytochrome *b*, and conserved regions matched blocks where redox centers are thought to occur (Howell

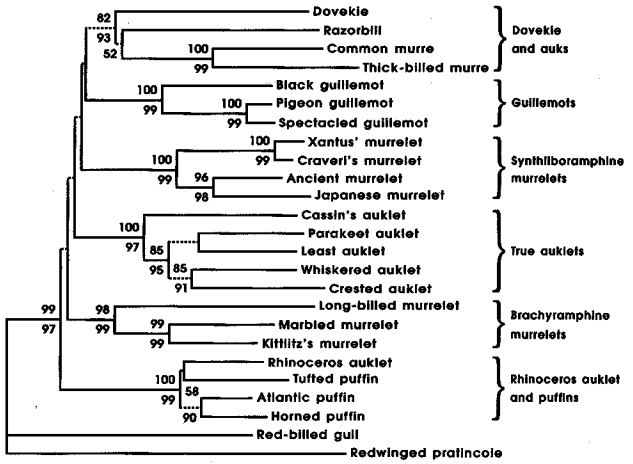


FIG. 1.—Phylogenetic tree derived from neighbor joining on genetic distances among cytochrome *b* sequences for the Alcidae. Numbers above branches indicate the number of times a group occurred in 100 bootstrap replications; numbers below branches indicate support from standard error tests. Values of support less than 50% are not shown. Dashed lines indicate branches with <95% support from standard error tests.

1989). A total of 315 sites were phylogenetically informative, with 120 involving transversions. In pairwise comparisons of sequences, transitions outnumbered transversions by a mean ratio of 4.2:1 (SE = 0.29, range = 1.4–52): C ↔ T transitions were the most common substitutions, and G ↔ T transversions were the least common (table 3). Only pairwise comparisons involving the pratincole were approaching a transition:transversion ratio of 0.5 indicative of saturation. Genetic distances among cytochrome *b* sequences, using Kimura's (1980) two-parameter correction, averaged 0.117 (SE = 0.0014) and ranged from 0.012 to 0.170. Distances within tribes averaged 0.062 (SE = 0.023), whereas distances among tribes averaged 0.118 (SE = 0.118).

Twenty-three of 37 allozyme loci were variable, and 18 were phylogenetically informative (table 2).

### Phylogenetic Analyses of Sequence Data

The neighbor-joining tree based on Kimura-corrected distances provided strong support for a monophyletic relationship among the 23 alcid sequences (fig. 1). Within the Alcidae, sequences grouped into six major lineages: (1) the dovekie and auks, (2) guillemots, (3) brachyramphine murrelets, (4) synthliboramphine murrelets, (5) true auklets, and (6) the rhinoceros auklet and puffins. All major groups except the dovekie and auks were supported by values of at least 95% from standard error tests and 90% from bootstrap analysis. Several intermediate and terminal nodes also were supported well: (1) the two species of murres formed sister taxa, (2) synthliboramphine murrelets constituted two species pairs, (3) the pigeon and spectacled guillemots were sis-

**Table 4**
**Indices of Support Produced by Alternative Methods of Weighting Substitutions in Maximum Parsimony Analysis**

| Monophyletic Relationship | Weighting Scheme | | | |
|---|---|---|---|---|
| | 1:1 | 1:4 | 1:20 | Empirical |
| Alcidae | 100/+ | 98/+ | 92/+ | 96/+ |
| All species except puffins | <50/1 | — | — | — |
| All species except brachyramphine murrelets | — | <50/2 | 59/+ | — |
| All species except guillemots | — | — | — | <50/+ |
| Dovekie, auks, guillemots, all murrelets | <50/1 | — | — | — |
| Dovekie, auks, guillemots, synthliboramphine murrelets | <50/1 | <50/2 | 73/2 | — |
| Dovekie, auks, guillemots | — | — | — | — |
| Dovekie, auks, synthliboramphine murrelets | <50/1 | 68/4 | <50/+ | 57/+ |
| Auklets, puffins, brachyramphine murrelets | — | — | — | <50/+ |
| Auklets, puffins | — | 74/7 | 80/+ | 57/+ |
| Dovekie, auks | 90/3 | 95/+ | 87/+ | 91/+ |
| Dovekie, razorbill | <50/0 | — | — | — |
| Dovekie, murres | — | — | 57/+ | <50/+ |
| Auks | <50/0 | <50/2 | — | — |
| Common & thick-billed murres | 100/+ | 100/+ | 86/+ | 97/+ |
| Guillemots | 100/10 | 100/+ | 100/+ | 100/+ |
| Synthliboramphine murrelets | 100/+ | 100/+ | 100/+ | 100/+ |
| Pigeon & spectacled guillemots | 100/+ | 100/+ | 100/+ | 100/+ |
| Xantus' & Craveri's murrelets | 100/+ | 100/+ | 100/+ | 100/+ |
| Ancient & Japanese murrelets | 96/2 | 95/9 | 96/+ | 96/+ |
| Brachyramphine murrelets | 97/9 | 99/+ | 99/+ | 99/+ |
| Marbled & Kittlitz's murrelets | 100/9 | 98/+ | 95/+ | 97/+ |
| Auklets | 100/+ | 100/+ | 100/+ | 100/+ |
| Parakeet, least, whiskered & crested auklets | 83/2 | 97/+ | 97/+ | 94/+ |
| Parakeet & least auklets | 58/2 | 60/3 | 68/3 | <50/+ |
| Whiskered & crested auklets | 77/1 | 77/1 | 58/1 | 76/+ |
| Puffins | 100/+ | 100/+ | 100/+ | 97/+ |
| Rhinoceros auklet, Atlantic & horned puffins | — | — | — | <50/+ |
| Tufted, Atlantic & horned puffins | <50/1 | 53/2 | 56/1 | — |
| Atlantic & horned puffins | 67/3 | 70/4 | 81/3 | <50/+ |
| Number of relationships receiving strong support | 12 | 13 | 11 | 13 |

NOTES.—1:1 = all substitutions weighted equally; 1:4 = transversions weighted four times transitions; 1:20 = transversions weighted 20 times transitions; empirical = empirically derived weighting scheme. The first number for each analysis is the number of times a given relationship occurred in 100 bootstrap replications. The second number is the Bremer support index, ranging from 1 to 10; values >10 are denoted +. — = relationship not present.

ter taxa, (4) Cassin's auklet was basal to the other true auklets, and (5) the marbled and Kittlitz's murrelets grouped together. Most relationships among the six lineages, among the auklets, and among the puffins were not resolved in this analysis in that they received poor support.

Maximum parsimony analysis in which all substitutions were weighted equally produced two trees that were virtually identical to the neighbor-joining tree and that differed from each other only in the position of the razorbill relative to the dovekie and murres (length = 1,014 steps, consistency index = 0.41, mean ± SD length of 1,000 random trees = 1517 ± 30.5; table 4). This analysis provided strong support for a monophyletic relationship among the alcid sequences and grouped sequences into six major clades identical to those produced by the neighbor-joining method. Each of the major groups was supported by bootstrap values of at least 90%, as well as by generally high Bremer indices. Most intermediate and terminal branches also were similar to those delineated by neighbor-joining, although a basal position for Cassin's auklet within the true auklets was not supported well. None of the relationships that differed between topologies produced by neighbor-joining versus maximum parsimony was supported well by either method. The number of relationships receiving strong support also was similar between the two results.

Alternative schemes for weighting nucleotide substitutions resulted in only minor differences in tree topologies (table 4), with all differences involving branches that received poor support in one or more analyses; relationships that were indicated clearly in all models were identical. The tree in which transversions were

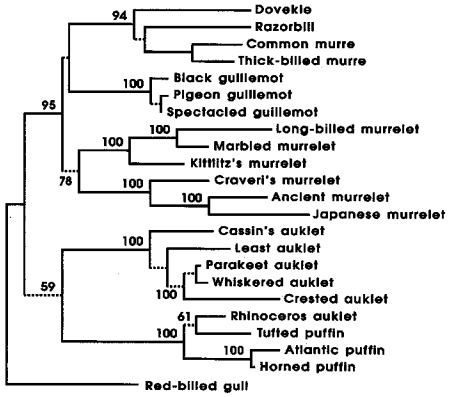


FIG. 2.—Maximum likelihood analysis of allozyme allele frequencies for the Alcidae. Dashed lines indicate branches with lengths that do not differ significantly from 0. Numbers indicate the percentage of 441 shortest trees from maximum parsimony analysis that support the node to the right; percentages <50% are not shown.

weighted 20 times transitions (length = 4,491 steps, mean length 1,000 random trees = 8,332 ± 226 steps) did not provide strong support for any relationships that were not indicated clearly in other analyses and provided only poor support for two relationships that were supported strongly by the other schemes. Trees that resulted either from weighting transversions four times transitions (length = 1,591 steps, mean length of 1,000 random trees = 2,617 ± 60.0 steps) or from the empirically derived scheme (length = 2,591 steps, mean length 1,000 random trees = 4,277 ± 226 steps) supported all clades that were indicated clearly in other schemes and also provided strong support for one additional relationship: a basal position for Cassin's auklet among the true auklets. The latter two weighting schemes therefore were the most informative in that they provided strong support for the largest number of clades.

## Phylogenetic Analyses of Allozyme Data

Maximum likelihood analysis of allozyme data (fig. 2) provided strong support for a monophyletic relationship among the alcids, as well as for the six major lineages indicated by sequence data. Four sister-taxa relationships (common and thick-billed murres, marbled and long-billed murrelets, ancient and Japanese murrelets, and Atlantic and horned puffins), as well as a monophyletic relationship among the dovekie, auks, guillemots, and murrelets, also were clearly indicated.

Maximum parsimony analysis of allozyme data resulted in 441 shortest trees (length 88 steps, mean length 1,000 random trees = 174 ± 6.8 steps; fig. 2). Because of the poor level of resolution, bootstrap and Bremer support values were not derived. However, all relationships except one that occurred in 95% or more of the most parsimonious trees (monophyly of crested, whiskered, and parakeet auklets) also were supported well by maximum likelihood analysis. Similarly, all relationships except one that were indicated clearly by maximum likelihood analysis (monophyly of the murres) occurred in at least 95% of the most parsimonious trees. Thus, phylogenetic analysis of allozyme data provided strong support for as many relationships as did analysis of cytochrome *b* sequences, despite the smaller number of informative characters in the allozyme data set. Neither data set provided clear solutions to the relationships among the tribes, among the auks and dovekie, or among the auklets.

## Taxonomic Congruence

Without exception, all relationships that received strong support in at least one analysis of cytochrome *b* sequences were present in all reconstructions based on sequence data (table 4). Three sister-group relationships (parakeet and least auklets, whiskered and crested auklets, and Atlantic and horned puffins) never obtained high bootstrap values but were present in all topologies; however, the first two of these were not present in the allozyme tree and thus they should be regarded with caution.

Although results for cytochrome *b* sequences and allozymes differed in several respects, with one exception (the brachyramphine murrelets), relationships that were indicated clearly by both data sets were identical; differences occurred almost entirely in groups that were supported poorly in one or both analyses. Thus, results were almost completely compatible. Strict congruence among all topologies for both sequence and allozyme data (figs. 1 and 2 and table 4) did not support any relationships that were not supported strongly by either data set alone.

## Total Molecular Evidence

Because results from analyses of nucleotide sequences and allozymes were generally compatible, we pooled data from the two sources in an attempt to resolve phylogenetic relationships further. Maximum parsimony analysis based on total molecular evidence resulted in three trees that differed only in the position of the dovekie relative to the razorbill and of the guillemots relative to the other tribes (fig. 3). This analysis provided strong support for most relationships that were indicated clearly by sequence data and/or allozymes alone, including one that was strongly supported by cytochrome *b* sequences but not allozyme data (monophyly of the dovekie and auks) and one that was supported by allozymes but not sequences (a sister-species relationship between the Atlantic and horned puffins). Monophyly of the dovekie, auks, guillemots, and murrelets received strong support from allozymes alone but not from the combined data set, and whereas a sister-taxa relationship

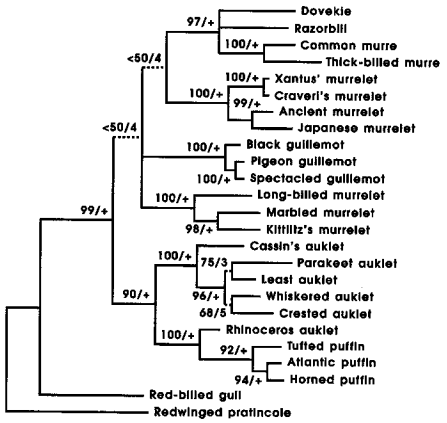


FIG. 3.—Consensus of three most parsimonious trees obtained using total evidence from cytochrome *b* nucleotide sequences and allozyme data for the Alcidae, with transversions and allozyme characters weighted four times transitions (length = 1,999 steps, mean length of 1,000 random trees = 3,318 steps, SD = 76.0). Branch lengths are proportional to numbers of supporting synapomorphies. Dashed lines indicate branches with <90% bootstrap support. The first number for each node is the number of times the group to the right occurred in 100 bootstrap replications; bootstrap frequencies less then 50% are not shown.The second number at each node is the Bremer support index; + indicates support >10 steps.

between the marbled and long-billed murrelets was indicated clearly by allozymes, both total evidence and sequence data alone provided strong support for a sister-taxa relationship between the marbled and Kittlitz's murrelets. The latter discrepancy probably results from the relatively small number of phylogenetically informative characters in the allozyme data set (18) compared with the nucleotide sequences (315). Two relationships that were suggested by both data sets but were not supported strongly by either (monophyly of Atlantic, horned, and tufted puffins, and a sister-group relationship between auklets and puffins) were clearly indicated by total evidence. Thus, analysis of total evidence was the most informative approach in that it resulted in the largest number of strongly supported phylogenetic relationships. Most deep branches, as well as relationships between the dovekie and auks and among the auklets, remained unresolved.

Results of the present study are similar in many respects to those of Moum et al. (1994), who compared nucleotide sequences of the mitochondrial NADH dehydrogenase subunit 6 and 12S rRNA genes among most extant alcids. They also reported strong support for six major lineages, as well as for sister-taxa relationships between the murres, between ancient and Japanese murrelets, and between Atlantic and horned puffins. As in the present study, they could not determine most relationships among the major lineages, among the auklets, or among the dovekie and auks. In an attempt to clarify these ambiguities, we conducted maximum parsimony

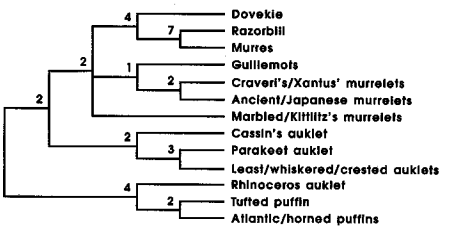


FIG. 4.—Phylogeny of the Alcidae proposed by Strauch (1985; extant species only). Numbers above branches are numbers of characters supporting the node to the right.

analyses (including alternative weighting schemes) using total evidence from cytochrome *b*, NADH6, 12S rRNA (total = 2,051 bp), and allozymes: this much larger data set did not provide strong support for any relationships that were not indicated clearly by allozymes and cytochrome *b* sequences alone.

## Comparison with Results of Morphological Analyses

Strauch (1985) conducted a compatibility analysis of the Alcidae using 33 morphological and ecological characters (fig. 4). Our results agree with his conclusions in many respects, including supporting six major lineages within the Alcidae, basal positions for the Cassin's and rhinoceros auklets within the auklets and puffins, respectively, and sister-taxa relationships between (1) the two species of murres, (2) Craveri's and Xantus' murrelets, (3) ancient and Japanese murrelets, and (4) Atlantic and horned puffins. One relationship that was ambiguous in the present study (monophyly of the auks) was supported by seven synapomorphies in Strauch's analysis, and one relationship that was strongly supported by molecular data (a sister-taxa relationship between pigeon and spectacled guillemots) was unresolved in Strauch's phylogeny. Although several differences exist between results of the two studies (especially in the relationships among the major lineages; compare figs. 3 and 4), all differences involved nodes either with low bootstrap values in the present study or with fewer than three synapomorphies in Strauch's topology. Results of the two studies therefore were entirely compatible. Taxonomic congruence between molecular and morphological analyses did not support any relationships that were not already clearly indicated by either data set alone.

Inclusion of Strauch's (1985) data on morphology and ecology in an analysis of total evidence (including various weighting schemes) did not increase the total number of groups receiving strong support, although sometimes the sister-group relationships changed. For example, a scheme in which transversions, allozymes, and morphological characters were given equal weights and transitions were assigned one-fourth the weight of other characters provided 98% bootstrap support for a monophyletic relationship among the auks but reduced

bootstrap support to 77% for a basal position of Cassin's auklet within the Aethini (tree not shown).

## Conclusions

Analyses of mitochondrial nucleotide sequences, allozymes, morphology, and ecology, including alternative methods of phylogeny reconstruction, demonstrated generally minor differences in the numbers and identities of relationships receiving strong support. Results from different data sets were almost completely compatible, suggesting that they reflect true phylogenetic relationships. Combination of independent data sets into a single analysis was the most sensitive approach in that it generated the largest number of relationships receiving strong support, but several relationships among the tribes, among the auklets, and among the dovekie and auks remained unresolved. Difficulties in determining the sister-group relationships among the tribes, in conjunction with the very short lengths of the deep branches (e.g., figs. 1 and 2), are typical of "star" phylogenies and suggest an ancient and relatively rapid basal radiation within the Alcidae. Fossil evidence also suggests an ancient radiation: despite a relatively strong fossil record (Olson 1985), alcids are represented by only two genera (*Miocepphus* and *Alcodes*) until the late Miocene, when representatives of all six tribes appear together. Thus, the six types of alcids (as well as a seventh, the extinct flightless mancallines) apparently arose sufficiently long ago and over such a short period of time that most phylogenetically informative characters have been obscured by superimposed changes. Secondary radiations appear to have occurred within the dovekie and auks as well as within the auklets. Resolution of these polytomies, if possible, may require much larger data sets and/or qualitatively different data, such as nuclear sequences or karyotypic data.

## Sequence Availability

Sequences have been deposited in GENBANK (accession numbers U37087, U37104, and U37286-308).

## Acknowledgments

We thank G. Chapdelaine (Canadian Wildlife Service, Ste.-Foy), D. Reynolds (Museum of Natural Science, Louisiana State University), S. Rohwer, G. Voelker, and C. Wood (The Burke Museum of Natural History, University of Washington), and K. Tucker (Canadian Wildlife Service, Lewisport) for tissue donations. J. Pitocchelli (American Museum of Natural History) and crew of the USFWS Research Vessel "M/V Tiglax" helped with tissue collections. T. P. Birt, O. P. Haddrath, H. D. Marshall, B. Millen, and C. A. Toline provided laboratory assistance, computer expertise, and/or helpful discussions. M. K. Peck helped with tissue collections and paid exceptional attention to allozyme electrophoresis. T. Moum shared previously unpublished sequence data. R. Honeycutt and an anonymous referee provided helpful comments on a previous version of the manuscript. Funding was provided by the Natural Sciences and Engineering Research Council (operating grant A0200 to A.J.B. and a postdoctoral fellowship to V.L.F.).

## LITERATURE CITED

AMERICAN ORNITHOLOGISTS' UNION. 1983. Check-list of North American birds, 6th ed. American Ornithologists' Union, Lawrence, Kans.

BAKER, A. J., A. M. LYNCH, and C. E. EDWARDS. 1985. Biochemical genetic studies of shorebirds: methods and applications. Wader Study Group Bull. **44**:34–39.

BREMER, K. 1988. The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution **42**:795–803.

EDWARDS, S. V., P. ARCTANDER, and A. C. WILSON. 1991. Mitochondrial resolution of a deep branch in the genealogical tree for perching birds. Proc. R. Soc. Lond. B **243**:99–107.

EERNISSE, D. J., and A. G. KLUGE. 1993. Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. Mol. Biol. Evol. **10**:1170–1195.

FELSENSTEIN, J. 1981. Evolutionary trees from gene frequencies and quantitative characters: finding maximum likelihood estimates. Evolution **35**:1229–1242.

———. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution **39**:783–791.

———. 1989. PHYLIP—phylogeny inference package (version 3.2). Cladistics **5**:164–166.

HELM-BYCHOWSKI, K., and J. CRACRAFT. 1993. Recovering phylogenetic signal from DNA sequences: relationships within the corvine assemblage (Class Aves) as inferred from complete sequences of the mitochondrial DNA cytochrome *b* gene. Mol. Biol. Evol. **10**:1196–1214.

HILLIS, D. M. 1995. Approaches for assessing phylogenetic accuracy. Syst. Biol. **44**:3–16.

HILLIS, D. M., J. P. HUELSENBECK, and C. W. CUNNINGHAM. 1994. Application and accuracy of molecular phylogenies. Science **264**:671–677.

HOWELL, N. 1989. Evolutionary conservation of protein regions in the proton-motive cytochrome *b* and their possible roles in redox catalysis. J. Mol. Evol. **29**:157–169.

HUELSENBECK, J. P. 1995. Performance of phylogenetic methods in simulation. 1995. Syst. Biol. **44**:17–48.

HUELSENBECK, J. P., and D. W. HILLIS. 1993. Success of phylogenetic methods in the four-taxon case. Syst. Biol. **42**: 247–264.

JONES, T. R., A. G. KLUGE, and A. J. WOLF. 1993. When theories and methodologies clash: a phylogenetic reanalysis of the North American ambystomatid salamanders (Caudata: Ambystomatidae). Syst. Zool. **42**:92–102.

KIMURA, M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J. Mol. Evol. **16**:111–120.

KNIGHT, A., and D. P. MINDELL. 1993. Substitution bias, weighting of DNA sequence evolution, and the phylogenetic position of Fea's viper. Syst. Biol. **42**:18–31.

KOCHER, T. D., W. K. THOMAS, A. MEYER, S. V. EDWARDS, S. PÄÄBO, F. X. VILLABLANCA, and A. C. WILSON. 1989. Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. Proc. Natl. Acad. Sci. USA **86**:6196–6200.

KORNEGAY, J. R., T. D. KOCHER, L. A. WILLIAMS, and A. C. WILSON. 1993. Pathways of lysozyme evolution inferred from the sequences of cytochrome *b* evolution in birds. J. Mol. Evol. **37**:367–379.

KUMAR, S., K. TAMURA, and M. NEI. 1993. MEGA: molecular evolutionary genetics analysis, version 1.0. Pennsylvania State University, University Park.

MINDELL, D. P., J. W. SITES, and D. GRAUR. 1989. Speciation evolution: a phylogenetic test with allozymes in *Sceloporus* (Reptilia). Cladistics **5**:49–61.

MIYAMOTO, M. M., and W. M. FITCH. 1995. Testing species phylogenies and phylogenetic methods with congruence. Syst. Biol. **44**:64–76.

MOUM, T. , S. JOHANSEN, K. E. ERIKSTAD, and J. F. PIATT. 1994. Phylogeny and evolution of the auks (subfamily Alcinae) based on mitochondrial DNA sequences. Proc. Natl. Acad. Sci. USA **91**:7912–7916.

OLSON, S. L. 1985. A selective synopsis of the fossil record of birds. Pp. 79–238 *in* D. S. FARNER, J. R. KING, and K. C. PARKES, eds. Avian biology. Vol. **8**. Academic Press, New York.

RODRIGO, A. G. 1993. Calibrating the bootstrap test of monophyly. Int. J. Parasitol. **23**:507–514.

RZHETSKY, A., and M. NEI. 1992. A simple method for estimating and testing minimum-evolution trees. Mol. Biol. Evol. **9**:945–967.

———. 1993. Theoretical foundation of the minimum-evolution method of phylogenetic inference. Mol. Biol. Evol. **10**: 1073–1095.

SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. **4**:406–425.

SIBLEY, C. G., and J. E. AHLQUIST. 1990. Phylogeny and classification of birds. Yale University Press, New Haven, Conn.

STEWART, C.-B. 1993. The powers and pitfalls of parsimony. Nature **361**:603–607.

STRAUCH, J. G., JR. 1985. The phylogeny of the Alcidae. Auk **102**:520–539.

SWOFFORD, D. L., and D. P. BEGLE. 1993. PAUP: phylogenetic analysis using parsimony, version 3.1. Illinois Natural History Survey, Champaign.

WATADA, M., R. KAKIZAWA, N. KURODA, and S. UTIDA. 1987. Genetic differentiation and phylogenetic relationships of an avian family, Alcidae (auks). J. Yamashna Inst. Ornithol. **19**:79–88.

RODNEY L. HONEYCUTT, reviewing editor

Accepted October 18, 1995