Molecular Phylogenetics and Evolution
Vol. 14, No. 1, January, pp. 141–151, 2000
Article ID mpev.1999.0682, available online at http://www.idealibrary.com on IDEAL®

Nuclear and Mitochondrial Genes Contain Similar Phylogenetic Signal for Pigeons and Doves (Aves: Columbiformes)

Kevin P. Johnson and Dale H. Clayton

Department of Biology, University of Utah, Salt Lake City, Utah 84112-0840

Received February 2, 1999; revised May 10, 1999

Molecular systematic studies generally assume that gene trees are reasonable estimates of species trees. We tested the validity of this assumption in the pigeons and doves (Aves: Columbiformes) by comparing phylogenies derived from nuclear (β-fibrinogen intron 7) and mitochondrial (cytochrome *b*) genes. Trees derived from the two genes when analyzed separately contained many nodes in common. A partition homogeneity test revealed no significant incongruence between trees derived from the two genes; so, we combined nuclear and mitochondrial data in subsequent phylogenetic analyses. The resulting tree, which was highly resolved and generally well supported, contained a strong biogeographic component. The rate of nucleotide substitution for the nuclear intron was approximately six times slower than that of cytochrome *b*. This resulted in a much higher consistency index for trees derived from the intron because of the low level of multiple substitution. However, the degree of resolution and support for trees reconstructed from the two genes was similar. We also examined the transition and transversion substitution rates for the genes. Third position transversions for cytochrome *b* accumulated linearly with intron divergence, suggesting low levels of multiple substitution for third position transversions. © 2000 Academic Press

Key Words: **cytochrome *b*; β-fibrinogen intron 7; Columbidae; gene trees; partition homogeneity test.**

INTRODUCTION

A major assumption of molecular systematics is that gene trees accurately reflect species trees. However, the stochastic nature of lineage sorting can yield gene trees that differ from the species tree (Pamilo and Nei, 1988; Avise 1989, 1994). It is unlikely, however, that two gene trees based on genes from different linkage groups would differ from the species tree in the same way. It is thus possible to check for lineage sorting that is incongruent with the species tree by comparing independent gene trees. If the two gene trees are the same, lineage sorting differences can be inferred to have been rare.

Mitochondrial genes have smaller effective population sizes than nuclear genes, making them less subject to lineage sorting that is incongruent with the species tree (Moore, 1995). However, mitochondrial genes are not independent of one another because the mitochondrial genome is a single linkage unit. Because nuclear genes sort independently of the mitochondrial genome, comparisons of nuclear and mitochondrial trees allow for potential identification of incongruent sorting events.

Mitochondrial genes have been used extensively in molecular phylogenetic studies of birds (many examples in Mindell, 1997). In contrast, nuclear genes have been used much less extensively (Caspers *et al.,* 1994, 1997; Prychitko and Moore, 1997). Prychitko and Moore (1997) found that nuclear and mitochondrial genes recovered identical phylogenies for a group of five woodpecker species (Aves: Picidae). A study in plants, however, found significant incongruence between trees derived from nuclear genes and trees derived from chloroplast genes (Mason-Gamer and Kellogg, 1996). In this paper, we compare nuclear and mitochondrial genes for a much larger number of species to evaluate further the likelihood of lineage sorting differences between independent gene trees.

The usefulness of different genes for reconstructing a phylogeny can be influenced by the substitution properties of those genes. Cytochrome *b* and other mitochondrial protein coding genes generally have high rates of substitution at third positions, relative to first and second positions (Holmquist *et al.,* 1983; Johnson and Sorenson, 1998). In addition, transition substitutions are generally more frequent than transversions in mitochondrial genes. These properties often result in multiple substitutions at third positions and possibly reduced levels of support for phylogenies derived from cytochrome *b* (Meyer, 1994). In contrast, the potential for homoplasy in nuclear introns is less because of slower mutation rates and a lack of strong selective constraints (Prychitko and Moore, 1997).

Another goal of this study was to compare the effects of substitution properties of mitochondrial and nuclear genes on phylogenetic resolution and support. We reconstructed and compared trees based on mitochondrial

1055-7903/00 $35.00

cytochrome *b* (cyt *b*) and nuclear β-fibrinogen intron 7 (FIB7) for 32 species of pigeons and doves (Aves: Columbiformes). Extant members of the Columbiformes are usually placed in a single family: Columbidae (Goodwin, 1983). Del Hoyo *et al.* (1997) recognized five subfamilies within Columbidae: Columbinae (181 species), Otidiphabinae (1 species), Gourinae (3 species), Didunculinae (1 species), and Treroninae (123 species). Phylogeny and taxonomy of this group have not been subjected to a cladistic treatment; so, we used cladistic analysis of our DNA sequence data to assess previous noncladistic classifications.

METHODS

DNA Sequencing

We obtained tissue and feather samples from 32 species of Columbidae representing the subfamilies Columbinae, Gourinae, and Treroninae (Table 1). We did not have material from the remaining two columbid subfamilies Otidiphabinae and Didunculinae, which are monotypic. The swiftlet *Aerodramus salanganus* (Apodiformes) was used as an outgroup to root the columbid tree.

TABLE 1

Columbid Species Sequenced

Name	Source
Aerodramus salanganus	DHC01, Sabah, Malaysia
Claravis pretiosa[1]	KUMNH, B85, Paraguay
Scardafella inca[1]	Tuscon, AZ
Scardafella squammata[1]	FMNH, SML88-153
Columbina minuta[1]	FMNH, DFS92-210
Columbina passerina[1]	KUMNH, B1755, Missouri
Columbina picui[1]	KUMNH, B153, Paraguay
Metriopelia ceciliae[1]	LSUMZ, B-23851, captive
Columba leucocephala[1]	KUMNH, B1718, Florida
Columba oenops[1]	FMNH, AJB-556
Columba plumbea[1]	FMNH, ATP86-136
Columba subvinacea[1]	FMNH, SML-1045
Macropygia phasianella[1]	FMNH, TPG-1005
Columba livia[1]	Salt Lake City, UT
Streptopelia chinensis[1]	FMNH, DW-4712
Geotrygon montana[1]	KUMNH, B995, Peru
Leptotila megalura[1]	FMNH, PS-002, Argentina
Leptotila rufaxilla[1]	KUMNH, B793, Peru
Zenaida meloda[1]	LSUMZ, B-5236, Peru
Zenaida auriculata[1]	FMNH, PS-003, Argentina
Zenaida galapagoensis[1]	Tracy Aviary, captive
Zenaida graysoni[1]	LSUMZ, B-23847, captive
Zenaida macroura[1]	Phoenix, AZ
Zenaida aurita[1]	FMNH, SML87-062
Ducula bicolor[3]	LSUMZ, B-19214, captive
Phapitreron amethystina[1]	FMNH, ATP92-109
Oena capensis[1]	FMNH, SMG-4180
Ptilinopus leclancheri[3]	FMNH, TPG-990
Goura cristata[2]	KMNH, B1588, captive
Treron vernans[3]	LSUMZ, B-20696, captive
Geopelia cuneata[1]	KMNH, B1586, captive
Leucosarcia melanoleuca[1]	LSUMZ, B-20539, captive
Phaps chalcoptera[1]	Tracy Aviary, captive

Note. Nomenclature follows del Hoyo *et al.* (1997). Subfamilies: [1]Columbinae, [2]Gourinae, [3]Treroninae.

We extracted tissue samples using a Quiaquick Tissue Kit (Qiagen). For feather extractions, we removed 1 to 2 mm from the tip of the feather shaft and included 30 μl of 10% dithiothreitol (final concentration 1.3%) in the digestion protocol (Johnson and Sorenson, 1998). We amplified a portion of cyt *b* and nuclear fibrinogen intron 7 from total genomic extracts using PCR. We used the primers L14841 (Kocher *et al.,* 1989) and H4a (Harshman, 1996) to amplify cyt *b* and the primers FIB-BI7L and FIB-BI7U (Prychitko and Moore, 1997) to amplify FIB7. We used a Perkin–Elmer Thermal Cycler 9700 with the following reaction conditions: 94°C for 2 min, 35 cycles of 94°C for 30 s, 46°C for 30 s, 72°C for 30 s, and 72°C for 7 min. We purified PCR products using a Qiagen PCR Purification kit. We performed cycle DNA Sequencing with *Taq* FS DNA polymerase using either ABI dRhodamine dye terminators or ABI Prism BigDye Terminators (Perkin–Elmer). In sequencing reactions, we used the primers L14841, H4a, H15299 (Kocher *et al.,* 1989), and L15517 (Johnson and Sorenson, 1998) for cyt *b* and FIB-BI7L, FIB-BI7U, FIB-DOVEF (5′-TTT CTC TTT CCT CAT GAC CC-3′, this study), and FIB-DOVER (5′-CAT AAT GGG TCA TGA GGA AAG AG-3′, this study) for FIB7. We collected and analyzed DNA sequence data (GenBank Accession Nos. AF182648–AF182713) using an ABI Prism 377 automated DNA sequencer (PE Applied Biosystems). We aligned and reconciled complementary chromatograms using Sequencher 3.1 (GeneCodes). We also used Sequencher to align sequences across species; final alignments were done by eye. We treated indels as ambiguous in the phylogenetic analysis.

Analysis

We determined the number of variable and phylogenetically informative characters for both genes using PAUP* (Swofford, 1998). We compared the proportion of sites that were variable and phylogenetically informative using the z statistic approximation (Milton and Arnold, 1990). We calculated pairwise sequence divergences for both genes independently using PAUP* (Swofford, 1998). To compare relative rates of substitution between the two genes, we plotted sequence divergences from pairwise comparisons for cyt *b* against those for FIB7. We fit third order regressions to these data and took the derivative at the origin to derive a heuristic estimate of the relative rates of substitution. An alternative estimate of the relative rates of substitution can be obtained by comparing the relative tree length (rescaled for the number of characters). We computed an estimate of relative rates using this measure for comparison with the estimate using pair-

wise comparisons. We also examined the average number of changes per site for each codon position of cyt *b* and over all positions of FIB7. Of sites that changed for each of these site types, we calculated the fraction of sites which showed multiple substitutions as well as the average number of substitutions for each site showing at least one substitution. We also calculated these values for cyt *b* third position transitions and transversions.

We also examined percentage sequence divergence using transitions and transversions only for FIB7 and for third positions of cyt *b*. To evaluate the "native" ratio (unbiased by multiple substitution) of transition versus transversion substitutions for both genes, we plotted transitions against transversions for FIB7 and third position transitions against third position transversions for cyt *b* (Sturmbauer and Meyer, 1992). To derive a heuristic estimate of transition to transversion ratio for both genes, we fit first (FIB7) or third (cyt *b*) order regressions through the origin and took the first derivative at the origin to determine the initial slope. We did not perform statistical tests because of the nonindependence of the many pairwise comparisons and use these regressions only heuristically to estimate approximate relative rates. We evaluated the potential of third position transitions and transversions in cyt *b* for multiple substitutions by plotting percentage transversion and transition divergence at third positions of cyt *b* against overall divergence for FIB7 in pairwise comparisons. Since FIB7 showed little evidence for multiple substitution (see Results), we were able to use FIB7 divergence as a linear approximation of divergence.

To compare phylogenies derived from nuclear (FIB7) and mitochondrial (cyt *b*) gene regions, we first conducted parsimony searches with the two genes independently using PAUP* (Swofford, 1998). We compared tree to tree symmetric difference distances (Penny and Hendy, 1985) between trees derived from separate analyses of the two genes and between the gene trees and the combined tree. To determine the sensitivity of tree topologies derived from cyt *b* to weighting, we progressively increased weighting of transversions over transitions using 2:1, 3:1, 4:1, 5:1, and 10:1 weighting of transversions. We also compared tree to tree distances resulting from these weighted searches to determine if transversion weighting made cyt *b* trees more similar to the FIB7 trees (Johnson and Sorenson, 1998).

We conducted bootstrap (Felsenstein, 1985) replicate searches with 1000 "fast" replicates (PAUP*; Swofford, 1998). We also computed and compared rescaled consistency indices for trees derived from both genes to determine the overall level of homoplasy present in each data set. We conducted a partition homogeneity test (Farris *et al.,* 1994, 1995; Swofford, 1998) using the unweighted data to compare the similarity of phylogenetic signal in the two data sets. Since this test indicated no evidence for phylogenetic incongruence between genes (see Results), we combined the two gene regions into one data set. Using the combined data set, we constructed parsimony trees using PAUP* (Swofford, 1998) with all characters unordered.

We also used maximum likelihood analysis of the combined data set to evaluate the sensitivity of tree topology to method of analysis. To select a likelihood model, we used likelihood ratio tests (Kishino and Hasegawa, 1989) to determine the simplest model that could not be rejected in favor of a more complex model (i.e., a model with more parameters). We considered the addition of additional substitution types, rate heterogeneity, and a fraction of invariant sites, using the combined parsimony tree with PAUP* (Swofford, 1998). Simpler models could be rejected in favor of a general time reversible model with six substitution types, incorporating rate heterogeneity and a fraction of invariant sites. We used this model with quartet puzzling in PAUP* to determine a likelihood tree under this model.

RESULTS

Sequence Evolution

Alignment of cyt *b* sequences was straightforward because, as expected, no insertions or deletions were present. This was not the case for β-fibrinogen intron 7, the length of which varied from 476 bp in *Claravis pretiosa* to 1141 bp in *Ptilinopus leclancheri.* Barring *C. pretiosa,* all other sequences exceeded 900 bp. Lengths of indel regions inferred from aligned sequences of FIB7 varied from 1 to 695 bp. Because indels were relatively infrequent, alignment was not compromised.

For cyt *b*, 564 of 1045 bp (54.0%) were variable and 381 of these (36.5%) were phylogenetically informative. FIB7 contained a smaller fraction of variable sites (505 of 1179, 42.8%, $P < 0.001$), and a much smaller fraction of these (167, 14.2%) was phylogenetically informative ($P < 0.0001$). Sequence divergence between columbiform species ranged between 0.97 and 17.07% for cyt *b* and between 0.27 and 7.30% for FIB7.

Comparison of pairwise divergences indicated that cyt *b* evolves approximately 5.6 times faster than FIB7 (Fig. 1). The estimate of relative rates using total tree lengths was 3.6; this value was lower than the pairwise estimate presumably because of unrecovered multiple substitutions in the cyt *b* tree. The fraction of sites (with at least one substitution) that was multiple in FIB7 was 31.7%, lower than all three codon positions for cyt *b* (see below). Pairwise comparisons also suggested that cyt *b* is more subject to multiple substitutions than FIB7 because the cyt *b* sequence divergence leveled out relative to FIB7 (Fig. 1). In addition, the number of reconstructed changes per site was 1.91 for cyt *b*, compared to 0.53 for FIB7. Most of this multiple substitution can be accounted for by third position transitions (Fig. 2) which, when plotted against FIB7

distance in pairwise comparisons, showed a dramatic leveling off compared to third position transversions.

Transitions at first positions also showed evidence of multiple substitution when plotted against overall divergence for FIB7 (not shown); however, leveling was only evident in pairwise comparisons to the outgroup. First position transversions and second position transitions showed no evidence of leveling when plotted either against total cyt *b* divergence or against FIB7 divergence (not shown). Second position transversions showed very little accumulation (not shown); so, it is difficult to interpret the potential of these substitutions for multiple substitution, but presumably it is minor. The fraction of second position sites (showing at least one substitution) that experienced multiple substitution was 44.8%, lower than first (70.5%) and third (91.4%) sites.

The substitution rate of third position transitions in cyt *b* was approximately six times greater than that for third position transversions (not shown). Transitions showed a much higher potential for multiple substitution, with the average number of changes for cyt *b* third position transitions equal to 3.7 compared to 1.2 for third position transversions. Of third positions for cyt *b* which show a transition or transversion, the fraction of transition changes that are multiple sites is 85.7%, while the fraction of transversion changes that are at multiple sites is 70.1%. Multiple substitutions in cyt *b* third position transitions are apparent at about 8% cyt *b* sequence divergence (Fig. 1). In contrast, FIB7 exhibited a much lower transition/transversion substitution ratio of approximately 1.5 (plot not shown). Neither transitions nor transversions in FIB7 appeared to be subject to increased multiple substitutions at high divergences (plot not shown). These analyses indicate that transversions should be weighted over transitions in phylogenetic analyses using cyt *b* but not weighted in analyses of FIB7.

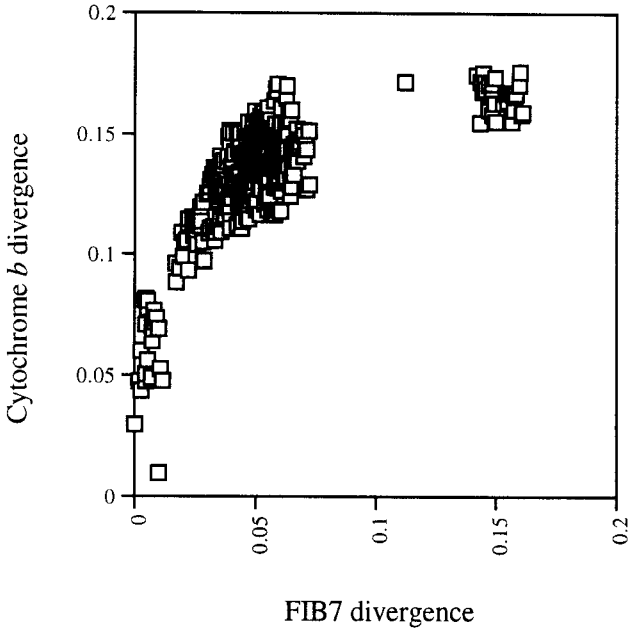

FIG. 1. Plot of overall pairwise divergences in cyt *b* against those for FIB7. The single outlying point is the comparison of *Claravis pretiosa* with the outgroup.

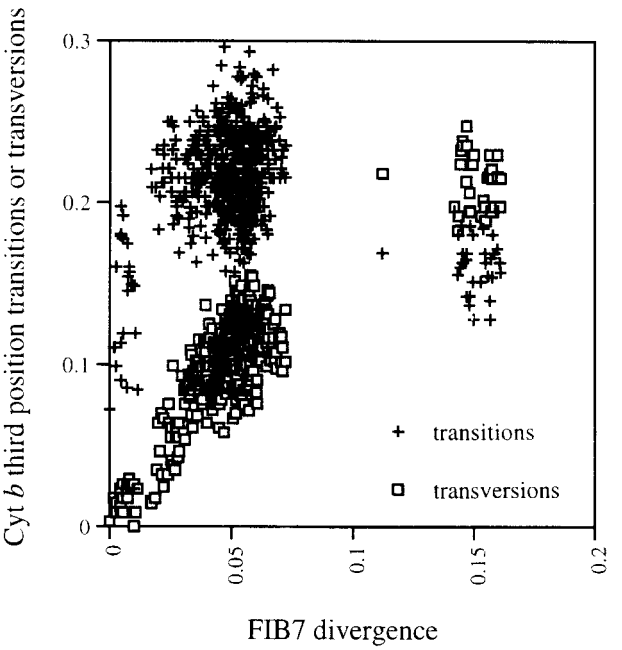

FIG. 2. Plot of pairwise third position transition and transversion divergences in cyt *b* against overall divergences for FIB7. The two outlying points are comparisons of *Claravis pretiosa* with the outgroup.

Phylogenetic Relationships

Phylogenies derived from unweighted parsimony for cyt *b* (Fig. 3) were very similar to those derived from FIB7 (Fig. 4) (symmetric difference distances: range 26–36, mean 32.33). Trees based on FIB7 had more branches supported in >50% of bootstrap replicates (16 nodes) than did trees based on cyt *b* (13 nodes). In no case did a node with >50% bootstrap support for one gene conflict with a node supported in over 50% of bootstrap replicates for the other gene. The rescaled consistency index (RC) for the FIB7 tree (0.668) was much higher than that for the cyt b tree (0.154). The RC for cyt *b* transversions only (0.253) was higher than for the entire gene (0.154). Interestingly, the RC for cyt *b* transversions was higher over the unweighted combined topology (0.257) than over the unweighted cyt *b* tree (0.253). In addition, cyt *b* trees with 5:1 weighting of transversions were more similar to trees derived from FIB7 (symmetric difference distances: range 24–30, mean 27.52) than were unweighted cyt *b* trees (range 26–36, mean 32.33). Topologies derived from cyt *b* changed as weighting of transversions increased. However, trees derived from 5:1 or higher weighting

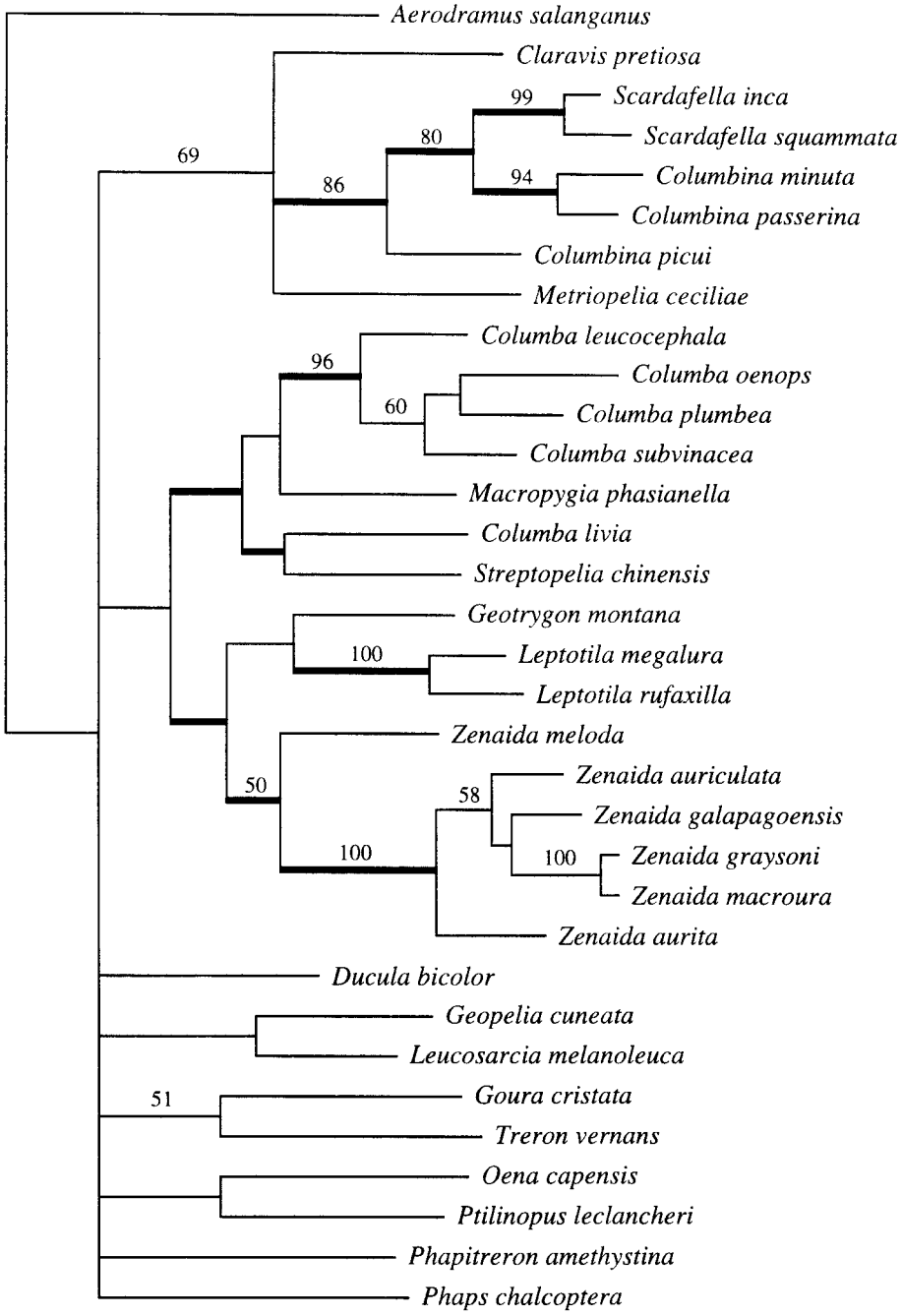

FIG. 3. Strict consensus of three trees (length = 1967, RC = 0.154) resulting from an unweighted parsimony search of bases of the mitochondrial cyt *b* gene. Branch lengths are proportional to reconstructed changes. Numbers above branches indicate nodes supported in >50% of bootstrap replicates. Bold branches are those also recovered in an unweighted analysis of the FIB7 gene (Fig. 4).

did not change (not shown). All nodes with >50% bootstrap support for the unweighted data were present in trees across all cyt *b* transversion weighting schemes (2:1 to 10:1) except for the node uniting *Goura* and *Treron* in the unweighted cyt *b* tree.

The partition homogeneity test (Farris *et al.,* 1994, 1995) revealed that the cyt *b* and FIB7 data sets could be considered samples of the same phylogenetic history ($P = 0.51$; Bull *et al.,* 1993). Because phylogenies derived from the two genes were not significantly incongruent, it is reasonable to combine data from the two genes in one analysis (Bull *et al.,* 1993). When we did this, fewer trees resulted (consensus Fig. 5) and these trees were more highly resolved and better supported. A total of 18 nodes had >50% bootstrap support. Eleven branches in the combined tree were common to analy-

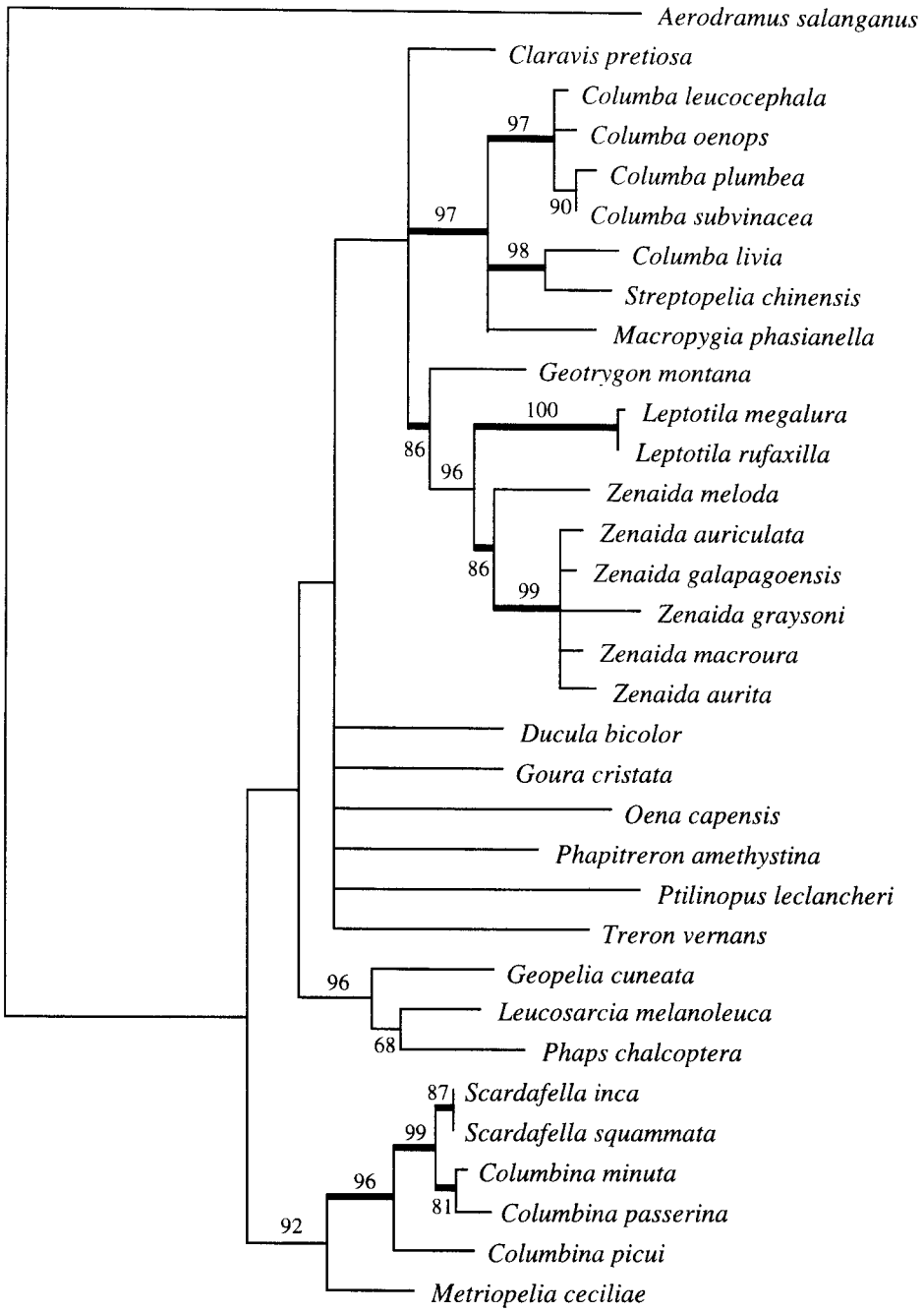

FIG. 4. Strict consensus of 1000 trees (length = 611, RC = 0.668) resulting from an unweighted parsimony search (maxtrees set to 3000) of bases of the nuclear FIB7 gene. Branch lengths are proportional to reconstructed changes. Numbers above branches indicate nodes supported in >50% of bootstrap replicates. Bold branches are those also recovered in an unweighted analysis of the cyt *b* gene (Fig. 3).

ses of cyt *b* and FIB7, as well as all cyt *b* transversion weighting schemes. The combined tree contained all of the nodes that were present in >50% of bootstrap replicates of the unweighted cyt *b* data. All but 2 of the 16 nodes with >50% bootstrap support in the FIB7 data appeared in the combined data consensus tree. Of nodes receiving >50% bootstrap support in the unweighted cyt *b* tree, only 1 showed a decrease in bootstrap support in the combined tree. In contrast, 8 nodes supported at >50% in the FIB7 analysis showed a decrease in bootstrap support in the combined tree. The combined tree was more similar to the unweighted cyt *b* tree (symmetric difference distances: range 14–16, mean 14.67) than it was to the FIB7 tree (symmetric difference distances: range 20–30, mean 25.64). Maximum likelihood analysis of the combined data (Fig. 6)

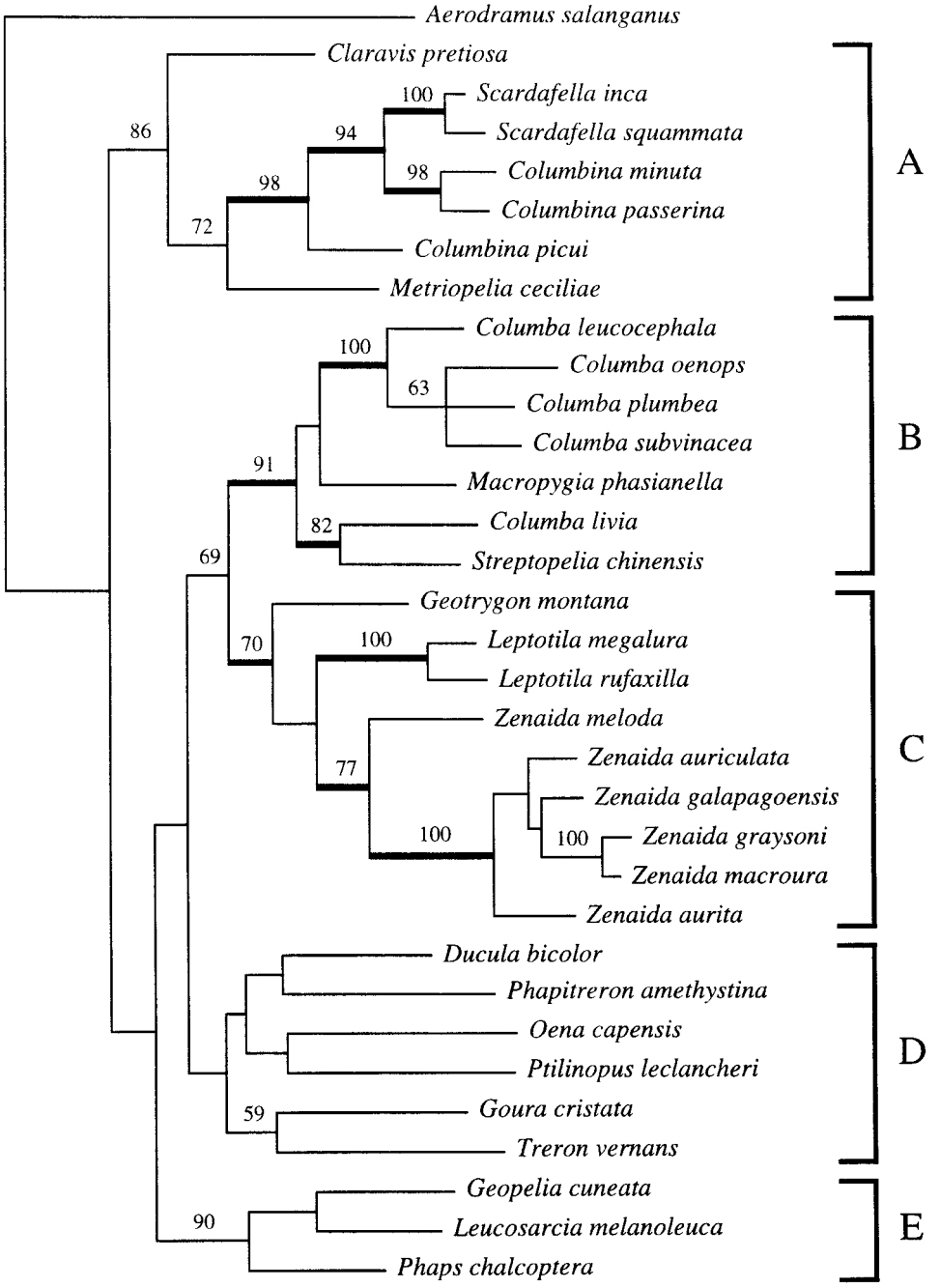

FIG. 5. Strict consensus of two trees (length = 2606, RC = 0.235) resulting from an unweighted parsimony search of the combined bases of the mitochondrial cyt *b* and nuclear FIB7 genes. Branch lengths are proportional to reconstructed changes. Numbers above branches indicate nodes supported in >50% of bootstrap replicates. Bold branches are those also recovered in all analyses of the cyt *b* gene (Fig. 3 and transversion weighted trees [not shown]) and the unweighted analysis of the FIB7 gene (Fig. 4). A–E indicate clades discussed in the text.

produced a tree that is very similar to the combined parsimony tree.

The two large subfamilies within Columbidae (Columbinae and Treroniae) were not monophyletic in any of the analyses. Rather, members of these two groups were interspersed among five major clades (Fig. 5). These clades generally sort into biogeographic zones. One well-supported group (A), containing the small New World ground doves in the genera *Columbina, Scardafella, Claravis,* and *Metriopelia,* is sister to the other four groups in the combined tree (as well as in many other analyses). A separate clade (C) of mid-sized, New World doves contains the genera *Geotrygon, Leptotila,* and *Zenaida,* all of which belong to the subfamily

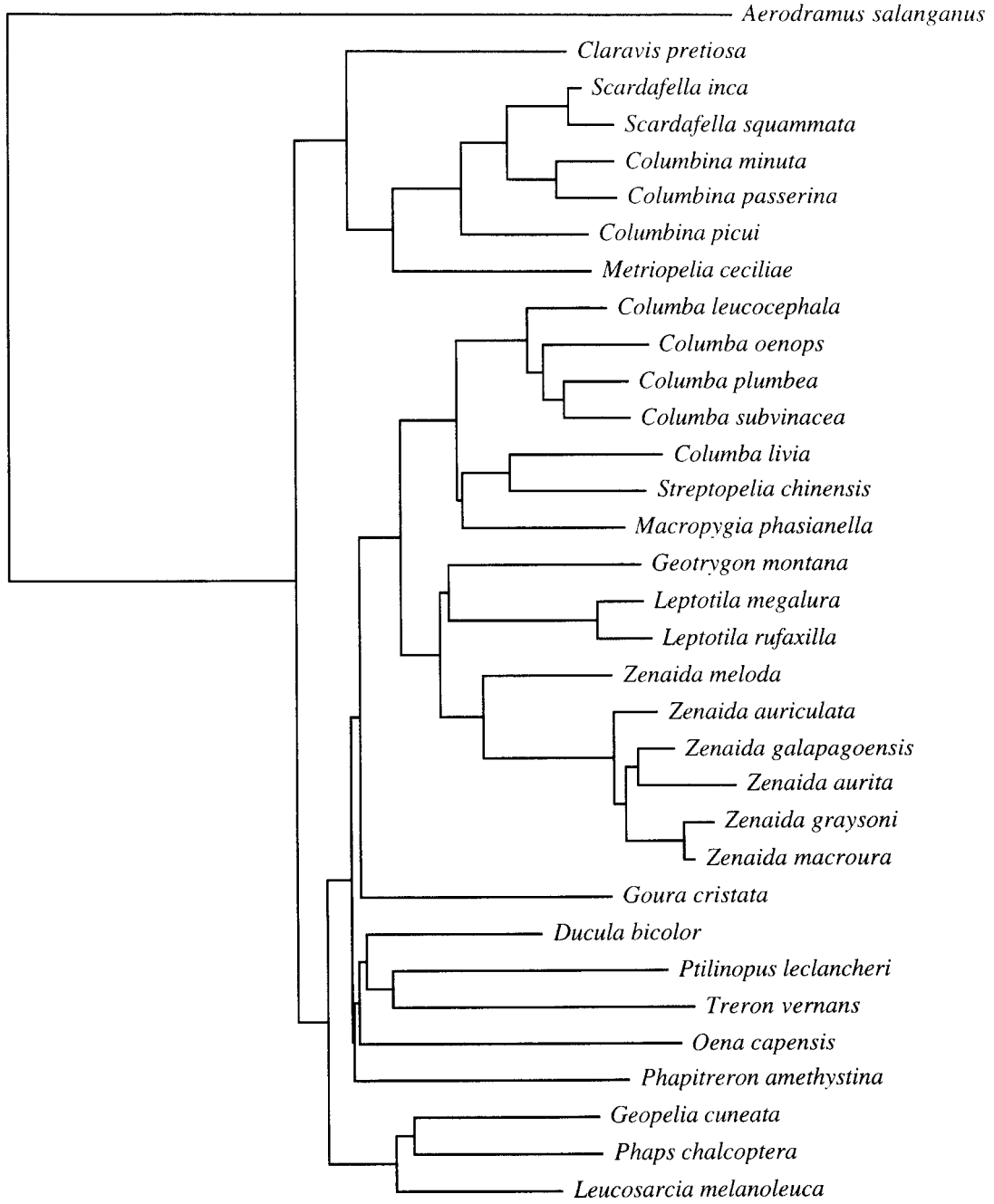

FIG. 6. Tree resulting from maximum likelihood quartet puzzling searches using a likelihood model with six substitution types, rate heterogeneity (shape parameter = 0.56, four rate categories), and 36.5% invariant sites.

Columbinae. Another strongly supported group (B) contains members of the subfamily Columbinae, including the Old World genera *Macropygia* and *Streptopelia,* as well as New and Old World representatives of the worldwide genus *Columba.* Another Old World group (D), with poorly supported monophyly, contains representatives of the subfamilies Treroninae (*Ducula, Ptilinopus,* and *Treron*), Columbinae (*Phapitreron* and *Oena*), and Gourinae (*Goura*). Finally, another well-supported group (E) is composed of the Australasian genera *Phaps, Leucosarcia,* and *Geopelia.*

In the case of five genera, we sequenced more than one species. Three of these genera, *Zenaida, Leptotila,* and *Scardafella,* were monophyletic in all trees. The remaining two genera, *Columba* and *Columbina,* were paraphyletic. The four species of New World *Columba* in this study formed a clade to the exclusion of the single Old World *Columba* (*livia*) in this study. The

genus *Scardafella* was imbedded within *Columbina,* resulting in paraphyly of *Columbina.*

DISCUSSION

Sequence Evolution

Patterns and rates of sequence substitution differ dramatically between cyt *b* (a mitochondrial protein coding gene) and FIB7 (a nuclear intron) in pigeons and doves (Columbiformes). Cyt *b* shows a five- to sixfold higher substitution rate than FIB7, and this is likely due to the higher mutation rate typical of vertebrate mitochondrial DNA (Brown *et al.,* 1979). The higher rate of substitution in cyt *b* (especially at third sites) makes this gene more prone to multiple substitutions. Interestingly, transversions at third positions in columbiform cyt *b* seem largely unaffected by multiple substitution, while third position transitions are highly subject to homoplasy. The difference between transitions and transversions in multiple substitution frequency may reflect a difference in rates, given that third position transitions accumulate approximately six times faster than transversions. These differences in rates correspond to differences between rescaled consistency indices, with that of FIB7 being much higher than that of cyt *b*. These observations reflect the tradeoff between variation and homoplasy in phylogenetic studies. Sites and substitution types with high rates, such as third position transitions, have the potential to resolve short internodes and recent speciation events. On the other hand, fast sites are the most subject to multiple substitution, making older nodes difficult to resolve. In contrast, short and/or recent nodes are more difficult to resolve with slow sites and substitution types; however, these sites show very little homoplasy.

Another striking difference between cyt *b* and FIB7 is the "native" (excluding multiple substitutions) ratio of transitions to transversions: approximately 6:1 for cyt *b* and 1.5:1 for FIB7. The expected ratio without mutation or substitution bias is 0.5:1 because there are twice as many ways to have a transversion as a transition (Li, 1997). Presumably the 1.5:1 ratio for FIB7 reflects the mutation bias (Li, 1997) for nuclear DNA. This assumes that substitution bias (due to selection) against transversions is minimal in nuclear introns. In contrast, the high ratio (6:1) for third sites in cyt *b* likely reflects both a mitochondrial mutation bias as well as a substitution bias (because many transversions result in amino acid substitutions, while no third position transitions do so). Currently, it is unclear whether the transition mutation bias is the same between nuclear and mitochondrial DNA in birds. Examination of regions of mtDNA (such as the control region), where transversions are less likely to be at a selective disadvantage, may help to determine the mutation bias for mtDNA.

Differences between transversion and transition accumulation in cyt *b* have implications for phylogeny reconstruction. Because transitions are much more subject to multiple substitution in cyt *b*, they should be downweighted. In addition, the rescaled consistency index for transversions over the unweighted cyt *b* tree (Fig. 3) is actually lower than it is over the unweighted combined tree (Fig. 5). This result, which suggests that transversions are more consistent with nuclear data than they are with cyt *b* transitions, argues for weighting transversions in analyses of cyt *b*. Other studies show that transversions within a gene are more consistent with combined evidence than they are with transitions within that same gene (Johnson and Sorenson, 1998, 1999).

While other studies have reported that bootstrap proportions generally increase as gene regions are combined (Johnson and Sorenson, 1998; Johnson and Lanyon, 1999), we found that many bootstrap values for nodes in the FIB7 tree decreased in the combined tree. This decrease occurred even though there was no evidence for significant conflict between cyt *b* and FIB7. Rather, it was probably due to the combination of a data set with very little homoplasy (FIB7) and one with a high level of homoplasy (cyt *b*). Bootstrap proportions underestimate the confidence level of a node as homoplasy increases (Zharkikh and Li, 1992, 1995). Thus, since the overall homoplasy of the combined data is higher (due to the addition of cyt *b*), the bootstrap values for nodes are expected to decline even though the confidence level for a node may actually have increased. This phenomenon warrants theoretical treatment with simulations.

Phylogenetic Relationships

Given the large number of taxa involved in this study, it is somewhat surprising that the trees from FIB7 and cyt *b* do not show significant incongruence. We might have predicted that differences in lineage sorting between nuclear and mitochondrial genes would have occurred at least once in the history of these 32 columbiform taxa but we found no evidence for this. However, given that a difference in lineage sorting between genes is most likely to occur when speciation events happen in rapid succession, it may actually be that lineage sorting differences will be difficult to detect. The difficulty of detecting incongruence resulting from lineage sorting differences may be higher for older nodes than for recent speciation events.

Many portions of the trees derived from the two genes showed agreement, and in places where nodes differed, these differences were poorly supported. The lack of resolution in both trees involved mainly a group of poorly sampled Old World genera. Groups for which taxa were more thoroughly sampled in this study (e.g., New World genera), tended to have better resolved and more strongly supported nodes in both trees. We sug-

gest that more thorough sampling of Old World genera would increase resolution and support for these nodes.

Despite the fact that the RC for the FIB7 tree was four times greater than that for cyt *b*, the number of nodes showing >50% bootstrap support decreased only from 16 (FIB7) to 13 (cyt *b*). Part of this may be due to a lack of resolution within the genus *Zenaida* for the FIB7 gene, presumably because FIB7 is not evolving fast enough to resolve relationships among these closely related species. The number of nodes with very high bootstrap support (>90%) was higher for FIB7 (11) than for cyt *b* (6), and this presumably more strongly reflects the underlying homoplasy in each gene.

An analysis of nuclear and mitochondrial gene trees in woodpeckers produced identical trees (Prychitko and Moore, 1997). However, in plants, it appears that trees between nuclear genes and organelles may often be in significant conflict (Mason-Gamer and Kellogg, 1996). In our analysis, phylogenies from nuclear and mitochondrial genes were very similar, with no evidence of significant conflict. While a phylogeny from a single gene region (or linkage group) may not reflect the species phylogeny (because of hybridization or lineage sorting problems), it is very unlikely that two unlinked gene regions would result in the same topology unless both genes tracked phylogenetic history. Lack of significant incongruence between FIB7 and cyt *b* also indicates that phylogenies from single gene regions may often be reasonable estimates of the organismal phylogeny (at least for nodes with high levels of bootstrap support).

The phylogeny derived from combined gene regions shows considerable biogeographic structuring and supports the monophyly of several genera. The two genera which are paraphyletic (*Columba* and *Columbina*) have previously been identified by morphological taxonomists as problematic (Goodwin, 1983). If paraphyly of these two genera persists upon further study, *Columba* should be split into two genera and *Scardafella* should be merged with *Columbina.*

Early noncladistic morphological and molecular work on the genus *Columba* has provided conflicting results. Some authors split New and Old World species into two distinct groups, which may or may not be each other's nearest relatives (morphology: Ridgway, 1916; Verheyen, 1957; behavior: del Hoyo *et al.,* 1997; molecular data: Cumley and Irwin, 1944; Irwin, 1953; Irwin and Miller, 1961), while others suggest merging New and Old World species (morphology: Peters, 1937; Boetticher, 1954; molecular data: Corbin, 1968). This study is consistent with the former interpretation, with the Old World *Columba* being closer to the Old World genus *Streptopelia* than they are to the New World *Columba.* It is unclear from our sampling of taxa what the sister taxon to New World *Columba* is but these species fall within a clade (B) containing *Macropygia, Streptopelia,* and Old World *Columba.*

Just as New and Old world species of *Columba* seem to be separated phylogenetically and biogeographically, biogeographic patterns for other columbid taxa emerge in the phylogeny. All three strictly Australian genera (*Geopelia, Leucosarcia,* and *Phaps*) form a clade (E) and this is consistent with traditional taxonomy (del Hoyo *et al.,* 1997). Other Old World genera form a clade (D) in the combined tree with representatives throughout Africa, Eurasia, and the South Pacific. Further work is needed on this Old World group to determine the limits of this clade.

One surprising result is that the small New World ground doves (*Claravis, Metriopelia, Columbina,* and *Scardafella*) are not closely related either to the mid-sized New World doves (*Geotrygon, Leptotila,* and *Zenaida*) or to the New World pigeons (*Columba*). Instead, the small New World ground doves are the sister group to all other species of Columbiformes. Traditionally, these small ground doves have been considered closely related to, albeit distinct form, other New World genera (Goodwin, 1983; del Hoyo *et al.,* 1997).

In general, this study shows that phylogenies of Columbiformes derived from two independent genes (nuclear and mitochondrial) are largely congruent. This congruence exists despite large differences in rates and substitution patterns between gene regions. These phylogenies are also consistent with much of the previous taxonomic classification of this group. We suggest that a combination of relatively short segments of nuclear and mitochondrial genes can provide considerable resolution and confidence concerning relationships among avian genera.

ACKNOWLEDGMENTS

We thank the DNA Sequencing Facility at the University of Utah, supported in part by NCI Grant 5p30CA42014. D. Drown provided assistance with statistical analyses. We thank D. L. Swofford, who gave us permission to publish results using a prerelease version of PAUP*. Two anonymous reviewers provided valuable comments on the manuscript. The following institutions generously provided samples for use in this study: Field Museum of Natural History, University of Kansas Museum of Natural History, Louisiana State University Museum of Natural Science, and the Tracy Aviary. This work was supported by NSF-CAREER award DEB-9703003 to D.H.C.

REFERENCES

Avise, J. C. (1989). Gene trees and organismal histories: A phylogenetic approach to population biology. *Evolution* **43:** 1192–1208.

Avise, J. C. (1994). "Molecular Markers, Natural History and Evolution," Chapman & Hall, New York.

Boetticher, H. von. (1954). Die Taubengattung *Columba* L. *Zool. Anz.* **153:** 49–64.

Brown, W. M., George, M., Jr., and Wilson, A. C. (1979). Rapid evolution of animal mitochondrial DNA. *PNAS* **76:** 1967–1971.

Bull, J. J., Huelsenbeck, J. P., Cunningham, C. W., Swofford, D. L., and Waddell, P. J. (1993). Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* **42:** 384–397.

Caspers, G. J., De Weerd, D. U., Wattel, J., and De Jong, W. W. (1997). Alpha-Crystallin sequences support a galliform/anseriform clade. *Mol. Phylogenet. Evol.* **7:** 185–188.

Caspers, G. J., Wattel, J., and De Jong, W. W. (1994). Alpha-A-crystallin sequences group tinamou with ratites. *Mol. Biol. Evol.* **11:** 711–713.

Corbin, K. W. (1968). Taxonomic relationships of some *Columba* species. *Condor* **70:** 1–13.

Cumley, R. W., and Irwin, M. R. (1944). The correlation between antigenic composition and geographic range in the Old or New World of some species of *Columba*. *Am. Nat.* **78:** 238–256.

del Hoyo, J., Elliott, A., and Sargatal, J., Eds. (1997). "Handbook of the Birds of the World. Vol. 4. Sandgrouse to Cuckoos." Lynx Edicions, Barcelona.

Farris, J. S., Kallersjo, M., Kluge, A. G., and Bult, C. (1994). Testing significance of congruence. *Cladistics* **10:** 315–320.

Farris, J. S., Kallersjo, M., Kluge, A. G., and Bult, C. (1995). Constructing a significance test for incongruence. *Syst. Biol.* **44:** 570–572.

Felsenstein, J. (1987). Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39:** 783–791.

Goodwin, D. (1983). "Pigeons and Doves of the World." Cornell Univ. Press, Ithaca, NY.

Harshman, J. (1996). "Phylogeny, Evolutionary Rates, and Ducks," Ph. D. thesis, Univ. of Chicago.

Holmquist, R., Goodman, M., Conroy, T., and Czelusniak, J. (1983). The spatial distribution of fixed mutations within genes coding for proteins. *J. Mol. Evol.* **19:** 437–448.

Irwin, M. R. (1953). Evolutionary patterns of antigenic substances of the blood corpuscles in Columbidae. *Evolution* **7:** 31–50.

Irwin, M. R., and Miller, W. J. (1961). Interrelationships and evolutionary patterns of cellular antigens in Columbidae. *Evolution* **15:** 30–43.

Johnson, K. P., and Lanyon, S. M. (1999). Molecular systematics of the grackles and allies and the effect of additional sequence (cyt *b* and ND2). *Auk* **116:**759–768.

Johnson, K. P., and Sorenson, M. D. (1998). Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome *b* and ND2) in the dabbling ducks (Tribe: Anatini). *Mol. Phylogenet. Evol.* **10:** 82–94.

Johnson, K. P., and Sorsenson, M. D. (1999). Phylogeny and biogeography of the dabbling ducks (Genus: *Anas*): A comparison of molecular and morphological evidence. *Auk* **116:**792–805.

Kishino, H., and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Homonoidea. *J. Mol. Evol.* **29:** 170–179.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., and Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: Amplification and sequencing with conserved primers. *PNAS* **86:** 6196–6200.

Li, W-H. (1997). "Molecular Evolution," Sinauer, Sunderland, MA.

Mason-Gamer, R. J., and Kellogg, E. A. (1996). Testing for phylogenetic conflict among molecular data sets in the tribe Triticeae (Gramineae). *Syst. Biol.* **45:** 524–545.

Meyer, A. (1994). Shortcomings of the cytochrome *b* gene as a molecular marker. *TREE* **9:** 278–280.

Milton, J. S., and Arnold, J. C. (1990). "Introduction to Probability and Statistics," 2nd ed., McGraw–Hill, New York.

Mindell, D. P. (1997). "Avian Molecular Evolution and Systematics," Academic Press, San Diego.

Moore, W. S. (1995). Inferring phylogenies from mtDNA variation: Mitochondrial gene trees versus nuclear gene trees. *Evolution* **49:** 718–726.

Pamilo, P., and Nei, M. (1988). Relationships between gene trees and species trees. *Mol. Biol. Evol.* **5:** 568–583.

Penny, D., and Hendy, M. D. (1985). The use of tree comparison metrics. *Syst. Zool.* **34:** 75–82.

Peters, J. L. (1937). "Check-List of Birds of the World," Vol. 3, Harvard Univ. Press, Cambridge, MA.

Prychitko, T. M., and Moore, W. S. (1997). The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). *Mol. Phylogenet. Evol.* **8:** 193–204.

Ridgway, R. (1916). The birds of North and Middle America. *Bull. U.S. Natl. Mus.* No. 50, **7:** 275–506.

Sturmbauer, C., and Meyer, A. (1992). Genetic divergence, speciation, and morphological stasis in a lineage of African cichlid fishes. *Nature* **358:** 578–581.

Swofford, D. L. (1998). PAUP*: Phylogenetic analysis using parsimony, version 4.0. Sinauer, Sunderland, MA.

Verheyen, R. (1957). Analyse du potential morphologique et project de classification des Columbiformes (Wetmore 1934). *Bull. Inst. Sci. Belg.* **33:** 1–42.

Zarkikh, A., and Li, W-H. (1992). Statistical properties of bootstrap estimation of phylogenetic variability from nucleotide sequences. I. Four taxa with a molecular clock. *Mol. Biol. Evol.* **9:** 1119–1147.

Zarkikh, A., and Li, W-H. (1995). Estimation of confidence in phylogeny: The complete-and-partial bootstrap technique. *Mol. Phylogenet. Evol.* **4:** 44–63.