Rates and Patterns of Mitochondrial DNA Sequence Evolution in Fringilline Finches (*Fringilla* spp.) and the Greenfinch (*Carduelis chloris*)

H. Dawn Marshall and Allan J. Baker

Centre for Biodiversity and Conservation Biology, Royal Ontario Museum; and Department of Zoology, University of Toronto

Rates and patterns of evolution in partial sequences of five mitochondrial genes (cytochrome *b,* ATPase 6, NADH dehydrogenase subunit 5, tRNA^{Glu}, and the control region) were compared among taxa in the passerine bird genera *Fringilla* and *Carduelis*. Rates of divergence do not vary significantly among genes, even in comparisons with the control region. Rate variation among lineages is significant only for the control region and NADH dehydrogenase subunit 5, and patterns of variation are consistent with the expectations of neutral theory. Base composition is biased in all genes but is stationary among lineages, and there is evidence for directional mutation pressure only in the control region. Despite these similarities, patterns of substitution differ among genes, consistent with alternative regimes of selective constraint. Rates of nonsynonymous substitution are higher in NADH dehydrogenase subunit 5 than in other protein-coding genes, and transitions exist in elevated proportions relative to transversions. Transitions appear to accumulate linearly with time in tRNA^{Glu}, and despite exhibiting the highest overall rate of divergence among species, there are no transversional changes in this gene. Finally, for resolving phylogenetic relationships among *Fringilla* taxa, the combined protein-coding data are broadly similar to those of the control region in terms of phylogenetic informativeness and statistical support.

Introduction

Relative simplicity of characterization combined with a high rate of evolution and an often uniparental mode of inheritance have made the mitochondrial genome a uniquely well studied molecule from a variety of evolutionary perspectives. With complete sequences available from some 44 organisms, the organization and function of mitochondrial DNA (mtDNA) genes are largely understood, and sequence variation within them provides a rich source of genetic markers for phylogenetic and population studies. However, certain aspects of sequence evolution remain problematic for these applications. For example, variation in evolutionary rate among sequences can substantially affect the accuracy of phylogenetic inference, as in the "long-branch attraction" phenomenon (Hendy and Penny 1989). Additionally, strong bias in base composition compromises some methods used to correct sequence divergence for multiple substitutions (Saccone et al. 1990).

Underlying its popularity as a genetic marker is the widespread belief that mtDNA variation conforms to a neutral model, thereby exemplifying the "molecular-clock" hypothesis that substitutions between amino acid or (by extension) nucleotide sequences occur approximately linearly with time (Zuckerkandl and Pauling 1965). However, even small amounts of positive selection can influence the evolutionary history of the entire mitochondrial genome; moreover, many studies never attempt to test the assumptions of the neutral model (Ballard and Kreitman 1995). One such assumption is that of mutation rate constancy, under which substitution rates are determined by selective constraint. A potential problem for evolutionary studies arises when unknown patterns of selective constraint, especially in conjunction with high rates of substitution, confound attempts to account for multiple substitutions (Kocher and Wilson 1991). Furthermore, mutation rates are thought to vary in different physiological environments (e.g., Martin and Palumbi 1993), and substitution rate may vary among lineages with different generation times (Li, Tanimura, and Sharp 1987).

The avian genus *Fringilla* comprises three closely related species, one of which (the common chaffinch, *F. coelebs*) is widely distributed throughout Europe, northern Africa, and the Atlantic islands, and comprises several morphologically and genetically distinct subspecies (Baker et al. 1990). The blue chaffinch (*F. teydea*) of the Canary Islands is the sister species to the common chaffinch, and these two form a sister group to the brambling (*F. montifringilla*) of Eurasia (Baker and Marshall 1997). Along with the closely related greenfinch (*Carduelis chloris*), these species provide an opportunity to compare sequences from a range of taxonomic levels, an advantage when attempting to understand the dynamics of mtDNA evolution (Thomas and Beckenbach 1989). Herein, we examine rates and patterns of substitution in several mitochondrial genes in this group of finches to test the assumption of selective neutrality, to assess differences in patterns of selective constraint among genes and rate variation among lineages, and to evaluate the suitability of mtDNA sequences for phylogenetic studies of fringilline finches.

Materials and Methods
Mitochondrial Genes Examined

We amplified and sequenced 641 bp of cytochrome *b* (cytb), 342 bp of NADH dehydrogenase subunit 5 (nd5), 300 bp of ATPase 6 (atp6), 918–932 bp of the control region (cr), and the tRNA^{Glu} gene (tglu; 71 bp). Cytb was amplified and sequenced using two overlapping sets of internal primers, L14841 and H15149 (Kocher et al. 1989) and L15063 and H15506 (unpublished data), which together provide 56.1% of the gene 98 bp

Table 1
Specimens Examined

Taxonomic Designation	Collection Locality	Genes Sequenced
Common chaffinches		
1. *Fringilla coelebs moreletti*	Terceira, Azores	All
2. *F. c. moreletti*.	Pico, Azores	All
3. *F. c. moreletti*.	Santa Maria, Azores	cytb, atp6, nd5
4. *Fringilla coelebs maderensis* . . .	Madeira Island	All
5. *F. c. maderensis*.	Madeira Island	cytb, atp6, nd5, tglu
6. *F. c. maderensis*.	Madeira Island	cr
7. *Fringilla coelebs canariensis* . . .	Tenerife, Canary Islands	All
8. *F. c. canariensis*	Gran Canaria, Canary Islands	All
9. *F. c. canariensis*	Gomera, Canary Islands	cytb, atp6, nd5
10. *Fringilla coelebs palmae*.	La Palma, Canary Islands	All
11. *F. c. palmae*	Hierro, Canary Islands	All
12. *Fringilla coelebs coelebs*	Segovia, Spain	All
13. *F. c. coelebs*.	Pitlochry, U.K.	cytb, atp6, nd5
14. *F. c. coelebs*.	Setubal, Portugal	cytb, atp6, nd5, tglu
15. *F. c. coelebs*.	Asiago, Italy	cr
16. *Fringilla coelebs africana*.	Rabat, Morocco	All
17. *F. c. africana*	Agadir, Morocco	cytb, atp6, nd5, tglu
18. *F. c. africana*	Nefza, Tunisia	cr
Blue chaffinches		
19. *Fringilla teydea*	Tenerife, Canary Islands	cytb, atp6, nd5
20. *F. teydea*.	Tenerife, Canary Islands	cytb, atp6, nd5
21. *F. teydea*.	Tenerife, Canary Islands	tglu, cr
22. *F. teydea*.	Tenerife, Canary Islands	tglu, cr
Bramblings		
23. *Fringilla monitifringilla*.	Tromsø, Norway	cytb, atp6, nd5
24. *F. monitifringilla*	Tromsø, Norway	All
25. *F. monitifringilla*	Tromsø, Norway	tglu, cr
Greenfinches		
26. *Carduelis chloris*	Agadir, Morocco	cytb, atp6, nd5
27. *C. chloris*	Uppsala, Sweden	cytb, nd5, tglu, cr
28. *C. chloris*	Uppsala, Sweden	atp6, nd5, tglu, cr

NOTE.—cytb = cytochrome *b*; atp6 = ATPase 6; nd5 = NADH dehydrogenase subunit 5; tglu = tRNA^{Glu}; cr = control region.

downstream of the 5′ end. The nd5 primers (L12301: 5′-AGGAGCAATCCGTTGGTCTTAGG-3′, and H12766: 5′-GACATGATTCCTACTCCTTCTCA-3′) target a relatively small portion (19%) of this approximately 1,800-bp gene, located downstream of the tRNA^{Leu}, in which the forward primer was placed. The atp6 primers (L8552 and H8957; unpublished data) are situated within the gene and yield sequence for approximately half (43.9%) of it, closer to the 5′ end. The alphanumerical designation of each primer indicates whether it occurs in the heavy or light strand (H or L) and the position of its 3′ nucleotide in the human mtDNA sequence (Anderson et al. 1981). Primers used to amplify and sequence the control region and tglu primers are described in Baker and Marshall (1997). The control region sequences used here include the hypervariable portion of domain I, the 3′ portion of the conserved central domain, and all of the third domain described in Marshall and Baker (1997), although for a different suite of specimens.

Collection of Samples and DNA

A total of 28 specimens representing six subspecies of common chaffinch, the blue chaffinch, the brambling, and the greenfinch were used in this study. Taxa, collection locality, and genes sequenced for each specimen are given in table 1. At least two individuals from each taxon were sequenced for each gene, although the particular individuals vary among genes. Genomic DNA was extracted from liver, heart, or spleen using standard proteinase K–phenol–chloroform methods (Sambrook, Fritsch, and Maniatis 1989).

Polymerase Chain Reaction (PCR) and Sequencing

Double-stranded amplification reactions contained 10 mM Tris-HCl (pH 8.3), 1.5 mM MgCl$_2$, 50 mM KCl, 50 μM of each dNTP, 0.4 μM of each primer, and 1 U *Taq* DNA polymerase (Boehringer Mannheim) in a 25-μl volume. Amplification was achieved with a thermal cycle of 93°C for 30 s, 48–50°C for 30 s, and 72°C for 60 s, repeated 35 times. Products were purified using agarose separation followed by binding to glass beads (Gene Clean; BIO 101), and were sequenced using either the Sequenase 2.0 dideoxy sequencing kit (United States Biochemical) and ^{35}S radiolabel or the Ampli-Cycle cycle sequencing kit (Perkin Elmer) and ^{33}P radiolabel, according to the manufacturer's instructions. At least two individuals per taxon were sequenced, and portions of each gene were sequenced from both strands.

Sequence Analysis

Sequences were compared using the computer programs ESEE (Cabot and Beckenbach 1989) and MEGA

(Kumar, Tamura, and Nei 1993). With the exception of the control region, sequences were aligned manually; the alignment of the control region is described in Marshall and Baker (1997). To describe the evolution of mtDNA in finches, we first evaluated conformity of the data to expectations under neutral theory using statistical tests. McDonald and Kreitman's (1991) test was used to assess the distribution of synonymous and replacement changes within and between species for the combined protein-coding data. Genes were combined for this test to increase the number of variable sites and, therefore, the potential to reject the null hypothesis. This was considered valid because mitochondrial genes belong to a single linkage group. Tajima's (1989) test was applied to samples of *F. coelebs* and *F. montifringilla* control region sequences in a previous study (Marshall and Baker 1997).

Compositional evolution was evaluated by calculating nucleotide bias, AT skew, and GC skew for each gene individually and at each codon position, and amino acid composition bias was estimated for the three protein-coding genes by the formulae described in Kocher et al. (1995). As an estimate of directional mutation pressure, we evaluated deviations from 50% GC content at third positions and in the two noncoding genes with the χ^2 test. To determine whether sequences are compositionally stationary, compositional divergence was calculated for among-species pairwise comparisons of each gene according to the method of Gillespie (1986). To evaluate patterns of substitution, transition-to-transversion (ts/tv) ratios were estimated using the computer program PAML (Yang 1995). In addition, pairwise estimates of transition and transversion differences for each gene were calculated and plotted against third-position transversions from the three protein-coding genes combined. For coding genes, we confined this analysis to third-position changes, as they were most frequent. Third-position transversions were chosen to represent time, as they are thought to be the least saturated component of the sequence data (Kocher et al. 1995).

To compare rates of evolution among mitochondrial genes, pairwise estimates of divergence were calculated among species and among individuals within *F. coelebs* at all sites for each gene, and separately at synonymous and nonsynonymous sites for the three protein-coding genes. The Nei and Gojobori (1986) algorithm, which includes a Jukes and Cantor (1969) correction for multiple substitutions, was considered suitable for this purpose, despite certain shortcomings described by Muse (1996), as pairwise divergences were generally not high (except in comparisons with *C. chloris*). The significance of differences in divergences among genes was assessed using the Kolmogorov-Smirnov two-sample one-tailed test (D; as in Bielawski and Gold 1996), which does not assume a particular distribution, and/or the *t*-test with infinite degrees of freedom (as suggested in Kumar, Tamura, and Nei 1993).

Rate variation among lineages was evaluated for each gene and for the combined protein-coding sequence data using the likelihood values of the best tree produced with and without the imposition of equal branch lengths (Fel-senstein 1988), assuming that the tree topology is the same under these two criteria. Differences in log-likelihoods for these trees were compared using the *G*-test of goodness of fit (Sokal and Rohlf 1981). The programs DNAML and DNAMLK implemented in Phylip 3.5 (Felsenstein 1993) were used to obtain the likelihood values. Ts/tv parameters were estimated as already described. Felsenstein's test was chosen because it does not depend on being able to accurately assess the phylogenetic relationships among the taxa. As a result, it identifies overall rate heterogeneity in the data set.

To assess the utility of these mitochondrial genes for phylogeny estimation, we constructed separate neighbor-joining trees (Saitou and Nei 1987) using (1) the control-region sequences and (2) the protein-coding genes. We used the Kimura two-parameter method (Kimura 1980) to estimate distances, as recommended by Kumar, Tamura, and Nei (1993) for situations in which distances are small ($d < 0.3$) and/or ts/tv ratios are low (ts/tv < 2). Branch length confidence probabilities and bootstrap confidence levels (500 replicates) were calculated for each tree.

Results and Discussion
Sequences

A total of almost 2.3 kb was sequenced for one to three individuals from each taxon (table 1; GenBank accession numbers AF002879–AF002978). The cr sequences correspond to positions 110–411 and 611–1252 presented for these species in Marshall and Baker (1997). Approximately equal numbers of variable sites were identified for the protein-coding genes combined as for the noncoding tglu and cr sequences. In the former, 251 variable sites (89 potentially phylogenetically informative) were found, corresponding to 18.1%, 18.7%, and 23.1% of the total sequence for cytb, atp6, and nd5, respectively. As expected, most (192) are third-position changes; there are 44 first-position and 15 second-position variable sites. In the tRNA[Glu] gene, 21 sites are variable, while the control region contains 226 variable positions (96 potentially phylogenetically informative), accounting for approximately 24%–25% of the sequence obtained. Consistent with a mitochondrial rather than a nuclear origin, all protein-coding sequences translate appropriately with the mitochondrial code and the tglu and cr sequences contain expected putative structural features (Marshall and Baker 1997).

According to the McDonald-Kreitman (1991) test, if sequence variation is neutral, the ratio of synonymous to replacement changes should be the same for within-species polymorphism as for between-species divergence. Here, 63 silent and 13 replacement sites were found within *F. coelebs* (15 individuals in six subspecies), and 39 silent and 6 replacement sites were found between *F. coelebs* and *F. teydea* (consensus sequence). According to the *G*-test of independence (Sokal and Rohlf 1981), these ratios are sufficiently similar to be consistent with the expectation under neutrality ($G = 0.31$, $P = 0.60$). The fact that *F. coelebs* individuals were not sampled from a randomly mating population

Table 2
Compositional Divergence Estimates (*d*) from Pairwise Comparisons of Finch Species

Species Pair	cytb	atp6	nd5	tglu	cr
Fringilla coelebs–Fringilla teydea	0.00791	0.0314	0.0901	0.0749	0.288
F. coelebs–Fringilla montifringilla	0.0882	0.0876	0.0772	0.0187	0.0800
F. teydea–F. montifringilla	0.0949	0.0238	0.00606	0.0940	0.0741
F. coelebs–Carduelis chloris	0.258	0.332	0.254	0.648	0.380
F. teydea–C. chloris	0.286	0.456	0.0911	0.499	0.371
F. montifringilla–C. chloris	0.105	0.591	0.132	0.857	0.306

NOTE.—$d = \Sigma(p_i - q_i)^2/\Sigma[2P_i(1 - p_i)/n]$, where the frequencies of the four bases (I = G, A, T, C) in the two sequences being compared are p_i and q_i, $P_i = (p_i - q_i)/2$; and n is the number of nucleotides in the comparison. cytb = cytochrome *b*; atp6 = ATPase; nd5 = NADH dehydrogenase subunit 5; tglu = tRNA^{Glu}; cr = control region.

does not invalidate the result, as the test does not make this assumption. The test requires that the within-species sample of alleles have the same common ancestor, that the mutation rate not vary between species, and that the species be closely related, all features of these data. The acceptance of the null hypothesis of neutrality must nevertheless be considered conservative due to the fragmentary nature of the genes used. None of the other commonly used statistical tests of neutrality are appropriate for the coding-region data presented herein. However, Tajima's (1989) tests applied to samples of *F. coelebs* and *F. montifringilla* cr sequences in a previous study (Marshall and Baker 1997) were consistent with neutrality.

Sequence Composition and Patterns of Substitution

Nucleotide and amino acid composition contain information related to patterns of selective constraint and, thus, are important descriptive features of genes. Here, composition bias is strongest in the atp6 gene (0.269) and weakest in tglu (0.094), while in the protein-coding genes it is strongest at the third codon position (0.497–0.563) and weakest at the first (0.027–0.294). The composite bias at third positions and in the noncoding genes reflects an elevated proportion of A and reduced G, resulting in positive AT (0.558–0.765) and negative GC (0.896–0.942) skew values at this position. First positions are characterized by slight positive AT skew (0.0208–0.290) and, except in cytb, moderate negative GC skew (0.333–0.587), while second positions exhibit moderate negative AT (0.334–0.424) and GC (0.228–0.600) skew. Only in the control region was the GC content significantly different from 50% (χ^2 = 16.18, $P < 0.0001$). There was no evidence that base composition varies among lineages for any gene in finches, as compositional divergence estimates (table 2) were less than one (Gillespie 1986). Amino acid composition and bias are broadly similar in the three protein-coding genes (0.28–0.37).

Patterns of composition bias comprising reduced G and variation in bias among codon positions have also been reported in other vertebrates including mammals (Irwin, Kocher, and Wilson 1991), fishes (Kocher et al. 1995; Bielawski and Gold 1996), and sharks (Martin 1995). Such patterns likely reflect the reduced-guanine feature of vertebrate mitochondrial DNA as well as selective constraints. Composition bias at third positions in cytb is thought to be higher in birds and mammals

than in sharks and fishes (Cantatore et al. 1994) due to higher AT skew in birds and mammals. Consistent with this, the values reported here are slightly higher than the averages for birds (0.45–0.03 for nucleotide bias and 0.49–0.03 for AT skew; Cantatore et al. 1994). Significant directional mutation pressure has been reported for the cytb gene in other organisms including muscovy ducks (Jermiin et al. 1994), and Jukes and Bhushan (1986) reported a high correlation between G+C content of the control region and that at silent positions in mtDNA. This trend was not observed here, indicating a possible functional constraint acting on the control region of finches.

Patterns of substitution (the relative probabilities of change from one nucleotide to another) provide another indication of the spectrum of mutations tolerated by a gene. Substitution bias is prevalent in each gene among the four finch species and within *F. coelebs*. For example, ts/tv ratios range from 1.42 (among species for the cr gene) to 14.53 (third positions in cytb within *F. coelebs*), substantially exceeding the unbiased expectation of 1/2. Differences in the accumulation of transitions and transversions among genes are apparent from plots of pairwise transition and transversion divergence estimates against total third-position transversions (fig. 1). In all genes, transversions appear to accumulate linearly. The maximum transversion difference is about 9% in the cr gene and between 12% and 14% at third positions of the coding genes. In contrast, transitions accumulate more slowly after third-position transversion differences reach 3%–5%, with saturation occurring most quickly in the nd5 gene. An exception is the tglu gene, which shows no evidence of transition saturation even at the maximum observed difference of approximately 25%. Maximum observed transition differences in all genes are higher than transversion differences, especially in cytb where third-position transition differences approach 20%. Differences in the accumulation of transitions among genes indicate different proportions of sites that are free to vary (Wills 1995) as well as different rates of substitution, as discussed below.

Evolutionary Rates of Mitochondrial Genes

Relative rates of substitution of the different partial genes were inferred from average pairwise divergence estimates, and similar patterns were observed in among-species comparisons relative to within-species comparisons (fig. 2). In both sets of comparisons, protein-cod-
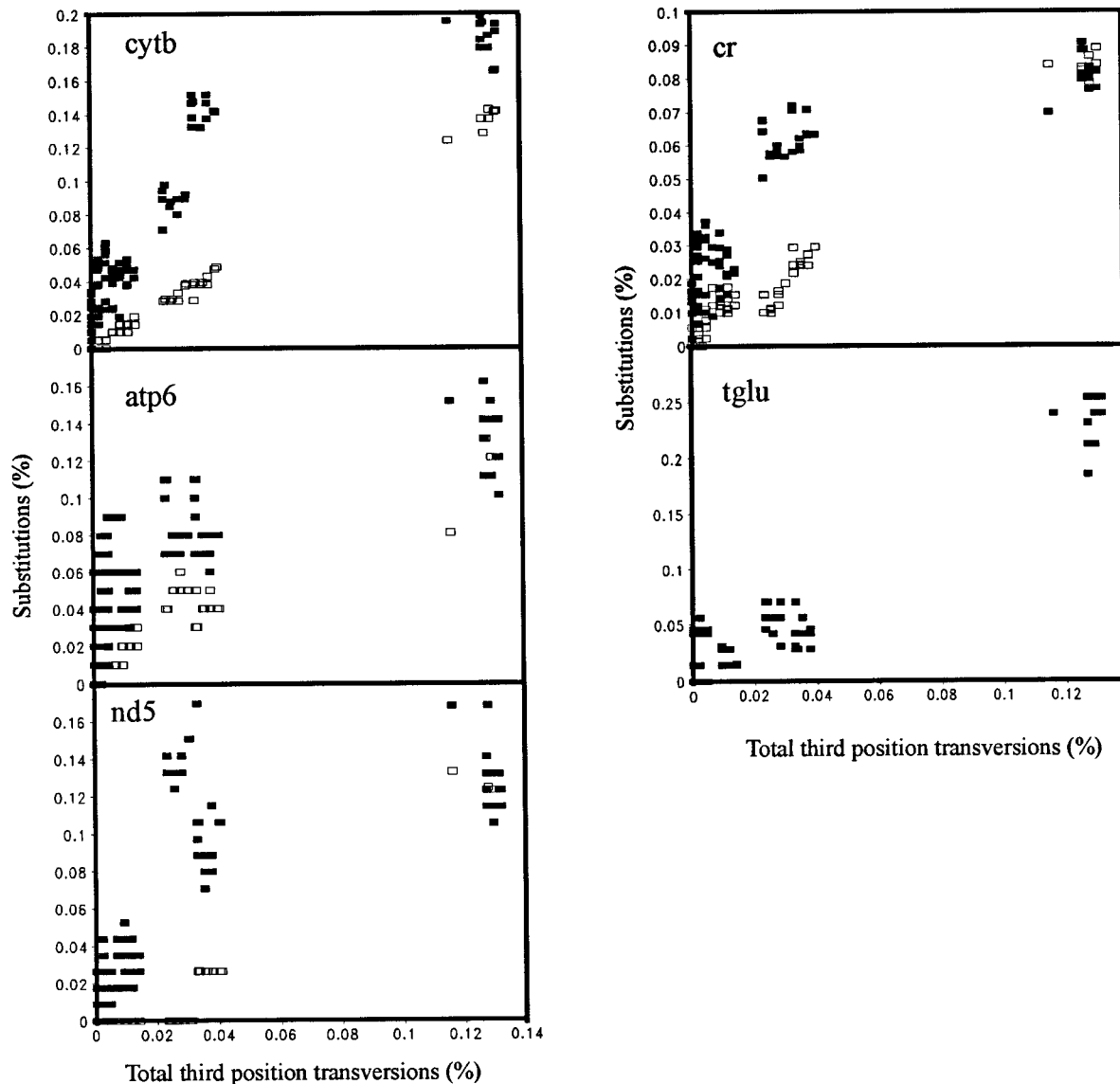
FIG. 1.—Pairwise estimates of transition and transversion divergences plotted against third-position transversions for each of the five genes. Filled squares indicate transitions, and empty ones indicate transversions.

ing gene divergences are lower (although not significantly so) than those in the noncoding genes. However, the average protein-coding synonymous divergence is significantly higher ($t = 4.80$, $P < 0.0001$; $D = 5/6$, $P = 0.025$) than the average cr divergence among species, while nonsynonymous differences in protein-coding genes are significantly lower than cr rates both within ($t = 2.08$, $P = 0.04$) and among ($t = 2.08$, $P = 0.04$; $D = 1$, $P = 0.0005$) species. Most rate differences among protein-coding genes are insignificant, with the exception that nonsynonymous differences are signficantly higher in nd5 among species (nd5 vs. atp6: $t = 3.5$, $P = 0.0003$; $D = 1$, $P = 0.0005$; nd5 vs. cytb: $t = 3.25$, $P = 0.001$; $D = 5/6$, $P = 0.025$). Significant rate variation among lineages was found among species in the cr ($G = 6.8$, $P = 0.03$) and nd5 ($G = 12.7$, $P = 0.002$) genes, and within species in the nd5 ($G = 28.8$, $P = 0.007$) and atp6 ($G = 27.6$, $P = 0.01$) genes. For the combined protein-coding sequences, significant rate variation was found among species ($G = 11.8$, $P = 0.025$), but not within *F. coelebs*.

With respect to rate variation among genes, the results for finches suggest that the intensity of selective constraint is weaker for nd5 and the cr than for the other protein-coding genes and that patterns of selective constraint may vary among genes. Rate variation among lineages only occurs for certain genes, so we suspect that it is also related to different selective constraints in different environments, as opposed to differences in metabolic rate or generation time. The species examined are all small passerines with similar weight-specific metabolic rates and generation times, but they occupy a range of habitats from hot and arid, to subtropical or temperate with high rainfall, to boreal. In support of the selective-constraint hypothesis, the nd5 gene of finches not only contains an unusually large number of replace-
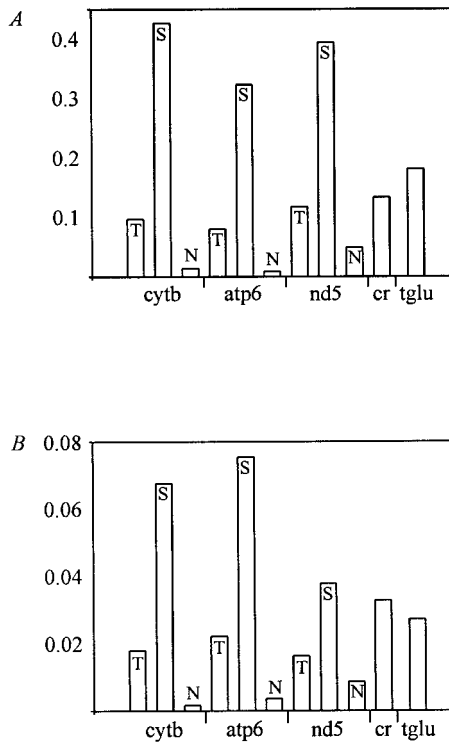
FIG. 2.—Histograms of the relative divergences among genes for (*A*) among-species and (*B*) within–*Fringilla coelebs* comparisons. T = total; S = synonymous; N = nonsynonymous.

ment substitutions, but also exhibits the most significant rate variation among lineages, consistent with relaxed constraint in this gene. Therefore, deviations from conformity to a strict "molecular clock" in this group of finches probably do not reflect mutation rate inconstancy and are not in violation of this assumption of neutrality.

Among other studies addressing rate variation among genes, Zhu et al. (1994) also found similar levels of divergence between cytb and cr sequences among rainbow fishes, while Bielawski and Gold (1996) reported significantly lower divergences in ND4L relative to cytb among closely related cyprinid species. In addition, numerous cases of rate variation in mtDNA genes among lineages of vertebrates have been documented. For example, Martin and Palumbi (1992) related lower substitution rates for sharks than for mammals in two mitochondrial genes to reduced metabolic rates in sharks. Mindell et al. (1996) and Cantatore et al. (1994) suggested that slower rates for birds and fish than for mammals involve body temperature and other physiological or life history traits, or different patterns of selective constraint in different environments. Finally, Krajewski and King (1996) postulated that significant rate variation among cytb sequences of cranes reflected a generation time effect. The assumption of mutation rate constancy for mitochondrial genes may therefore be violated for certain species.

Phylogenetic Utility and Signal of Mitochondrial Genes

The goal of this study was to evaluate features of mitochondrial DNA sequences, such as base composi-

tion bias and rate variation among lineages, which may confound phylogenetic inferences in fringilline finches. The results obtained must be interpreted with caution, given that partial genes were examined. However, no departures from compositional homogeneity, such as those reported in sharks (Martin 1995) and cichlid fishes (Kocher et al. 1995), were observed. There was no evidence for positive selection, nor for variation in substitution rate among genes or lineages due to forces other than different intensities and patterns of selective constraint. Rate variation among lineages was apparent for certain genes, but this can be circumvented by choosing a phylogenetic method which does not assume equal rates. Patterns of selective constraint do clearly vary among genes. For example, nd5 contains an excess of replacement substitutions relative to other protein-coding genes, while cytb exhibits high third-position transition differences. Tglu appears to evolve quickly, yet transitions are unsaturated in this gene, and the cr contains the lowest proportion of transitions relative to transversions of any gene. In rigorous phylogenetic analyses, such as will be performed in a subsequent study, these patterns will need to be addressed by the incorporation of weighting schemes and rate parameters. In summary, though, these sequences appear to be suitable for phylogeny reconstruction.

In addition to examining relevant aspects of the data, we are also interested in how the genes compare in their ability to resolve the relationships among fringilline finches. Because only small portions of certain genes were examined, we decided to compare combined protein-coding sequences with cr sequences. We feel that combining genes is justified, because mitochondrial genes are linked and thus share a phylogenetic history, and because the major difference among protein-coding genes appears to concern the intensity of selective constraint, which can ultimately be accounted for by a weighting scheme. Superficially, both the cr gene and the combined protein-coding genes show similar phylogenetic potential. Interestingly, the neighbor-joining phylogenies recovered from the two data sets (fig. 3) are different. Both trees share the same specific relationships ((*F. coelebs, F. teydea*) *F. montifringilla*) reported previously for cr data (Marshall and Baker 1997). In the protein-coding tree (fig. 3*A*), individuals from the Canary Islands form a clade whose sister group contains Madeiran birds, and island subspecies form a separate group from continental subspecies. However, the cr tree exhibits a within–*F. coelebs* topology in which the island subspecies are polyphyletic (fig. 3*B*). Variation also occurs in the arrangement of the different Canaries haplotypes (*F. c. canariensis* and *F. c. palmae* individuals) and the different continental haplotypes (*F. c. coelebs* and *F. c. africana* individuals). Neither using a parsimony method nor limiting the analysis to the same number of lineages results in the same topology for both data sets (results not shown). Similarly, trees constructed from single genes do not provide better support for poorly supported nodes (results not shown).

Topological differences between the two data sets are reconcilable on examination of statistical support.

A

B

Fig. 3.—Neighbor-joining trees constructed separately from (*A*) the combined protein-coding data and (*B*) the control region data. Bootstrap values and branch length confidence probabilities (in brackets) are given near the appropriate node.

Clades appearing in both trees tend to be well supported both by bootstrap confidence levels and by branch length confidence probabilities, and values for these are similar in the two data sets. Topologies that differ between the analyses are quite poorly supported; for instance, the clade grouping all island subspecies has a bootstrap value of 54 for the protein-coding data, and the inclusion of the continental subspecies with two of the island subspecies in the cr tree is supported by only 42 bootstrap replicates. Similarly, support values are very low for some groupings within the Canaries clade, and within the continental clade. Thus, neither data set is capable of satisfactorily resolving certain relationships. This may indicate a paucity of variation, resolvable by collecting more data. However, combining the protein-coding and cr data does not result in greatly improved statistical support (data not shown). Alternatively, the explanation may be that irresolvable nodes result from an evolutionary polytomy, consistent with a rapid wave of colonization. The relationships among fringilline taxa will be addressed more thoroughly in a sub-

sequent study pertaining to the colonization of the Atlantic islands by chaffinches.

The phylogenetic utility of the combined protein-coding genes therefore compares favorably with that of the cr among closely related taxa such as subspecies, and the cr actually performs well for resolving species-level splits. The cr is usually described as hypervariable (see Simon 1991 for a review), but this may result in part because only the more variable portions are usually targeted for analysis. Alternatively, because the taxa studied here are relatively recently diverged, large differences among genes may not have had time to accumulate. Nevertheless, this unexpected result underscores the importance of evaluating rates and patterns of mtDNA evolution prior to employing it as a phylogenetic marker.

Acknowledgments

LITERATURE CITED

ANDERSON, S., A. T. BANKIER, B. G. BARRELL et al. (14 co-authors). 1981. Sequence and organization of the human mitochondrial genome. Nature **290**:457–465.

BAKER, A. J., M. D. DENNISON, A. LYNCH, and G. LE GRAND. 1990. Genetic divergence in peripherally isolated populations of chaffinches in the Atlantic islands. Evolution **44**:981–999.

BAKER, A. J., and H. D. MARSHALL. 1997. Mitochondrial control-region sequences as tools for understanding evolution of avian taxa. Pp. 49–80 in D. P. MINDELL, ed. Avian molecular systematics and evolution. Academic Press, New York.

BALLARD, J. W. O., and M. KREITMAN. 1995. Is mitochondrial DNA a neutral marker? TREE **10**:485–488.

BIELAWSKI, J. P., and J. R. GOLD. 1996. Unequal synonymous substitution rates within and between two protein-coding mitochondrial genes. Mol. Biol. Evol. **13**:889–892.

CABOT, E. L., and A. T. BECKENBACH. 1989. Simultaneous editing of multiple nucleic acid and protein sequences with ESEE. Comput. Appl. Biosci. **5**:233–234.

CANTATORE, P., M. ROBERTI, G. PESOLE, A. LUDOVICO, F. MILELLA, M. N. GADALETA, and C. SACCONE. 1994. Evolutionary analysis of cytochrome *b* sequences in some perciformes: evidence for a slower rate of evolution than in mammals. J. Mol. Evol. **39**:589–597.

FELSENSTEIN, J. 1988. Phylogenies from molecular sequences: inferences and reliability. Annu. Rev. Genet. **22**:521–565.

———. 1993. Phylogeny inference package (PHYLIP). Version 3.5. Distributed by the author, Department of Genetics, University of Washington, Seattle.

GILLESPIE, J. H. 1986. Variability in evolutionary rates of DNA. Genetics **113**:1077–1091.

HENDY, M. D., and D. PENNY. 1989. A framework for the quantitative study of evolutionary trees. Syst. Zool. **38**:297–309.

IRWIN, D. M., T. D. KOCHER, and A. C. WILSON. 1991. Evolution of the cytochrome *b* gene in mammals. J. Mol. Evol. **29**:157–169.

JERMIIN, L. S., D. GRAUR, R. M. LOWE, and R. H. CROZIER. 1994. Analysis of directional mutation pressure and nucleotide content in mitochondrial cytochrome *b* genes. J. Mol. Evol. **39**:160–173.

JUKES, T. H., and V. BHUSHAN. 1986. Silent nucleotide substitutions and G+C content of some mitochondrial and bacterial genes. J. Mol. Evol. **24**:39–44.

JUKES, T. H., and C. R. CANTOR. 1969. Evolution of protein molecules. Pp. 21–132 in H. N. MUNRO, ed. Mammalian protein metabolism. Academic Press, New York.

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.

KOCHER, T. D., J. A. CONROY, K. R. MCKAYE, J. R. STAUFFER, and S. F. LOCKWOOD. 1995. Evolution of NADH dehydrogenase subunit 2 in east African cichlid fish. Mol. Phylogenet. Evol. **4**:420–432.

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.

KOCHER, T. D., and A. C. WILSON. 1991. Sequence evolution of mitochondrial DNA in humans and chimpanzees: control region and a protein-coding region. Pp. 391–413 in S. OSAWA and H. HONJO, eds. Evolution of life: fossils, molecules, and culture. Springer-Verlag, Tokyo.

KRAJEWSKI, C., and D. G. KING. 1996. Molecular divergence and phylogeny: rates and patterns of cytochrome *b* evolution in cranes. Mol. Biol. Evol. **13**:21–30.

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

LI, W.-H., M. TANIMURA, and P. M. SHARP. 1987. An evaluation of the molecular clock hypothesis using mammalian DNA sequences. J. Mol. Evol. **25**:330–342.

MCDONALD, J. H., and M. KREITMAN. 1991. Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature **351**:652–654.

MARSHALL, H. D., and A. J. BAKER. 1997. Structural conservation and variation in the control region of fringilline finches (*Fringilla* spp.) and the greenfinch (*Carduelis chloris*). Mol. Biol. Evol. **14**:173–184.

MARTIN, A. 1995. Mitochondrial DNA sequence evolution in sharks: rates, patterns, and phylogenetic inferences. Mol. Biol. Evol. **12**:1114–1123.

MARTIN, A., and S. PALUMBI. 1992. Rates of mitochondrial DNA evolution in sharks are slow compared with mammals. Nature **357**:153–155.

———. 1993. Protein evolution in different cellular environments: cytochrome *b* in sharks and mammals. Mol. Biol. Evol. **10**:873–891.

MINDELL, D. P., A. KNIGHT, C. BAER, and C. J. HUDDLESTON. 1996. Slow rates of molecular evolution in birds and the metabolic rate and body temperature hypotheses. Mol. Biol. Evol. **13**:422–426.

MUSE, S. V. 1996. Estimating synonymous and nonsynonymous substitution rates. Mol. Biol. Evol. **13**:105–114.

NEI, M., and T. GOJOBORI. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. **3**:418–426.

SACCONE, C., C. LANAVE, G. PESOLE, and G. PREPARATA. 1990. Influence of base composition on quantitative estimates of gene composition. Methods Enzymol. **183**:570–583.

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

SAMBROOK, J., E. F. FRITSCH, and T. MANIATIS. 1989. Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

SIMON, C. 1991. Molecular systematics at the species boundary: exploiting conserved and variable regions of the mitochondrial genome of animals via direct sequencing from amplified DNA. Pp. 33–71 in G. M. HEWITT, A. W. B. JOHNSTON, and J. P. W. YOUNG, eds. Molecular techniques in taxonomy, NATO ASI series, vol. 57. Springer-Verlag, Berlin.

SOKAL, R. R., and F. J. ROHLF. 1981. Biometry. Freeman, San Francisco.

TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**:585–595.

THOMAS, W. K., and A. T. BECKENBACH. 1989. Variation in salmonid mitochondrial DNA: evolutionary constraints and mechanisms of substitution. J. Mol. Evol. **29**:233–245.

WILLS, C. 1995. When did Eve live? An evolutionary detective story. Evolution **49**:593–607.

YANG, Z. 1995. Phylogenetic analysis by maximum likelihood (PAML). Version 1.1. Institute of Molecular Evolutionary Genetics, the Pennsylvania State University, University Park.

ZHU, D., B. G. M. JAMIESON, A. HUGALL, and C. MORITZ. 1994. Sequence evolution and phylogenetic signal in control-region and cytochrome *b* sequences of rainbow fishes (Melanotaeniidae). Mol. Biol. Evol. **11**:672–683.

ZUCKERKANDL, E., and L. PAULING. 1965. Evolutionary divergence and convergence in proteins. Pp. 97–166 *in* V. BRYSON and H. J. VOGEL, eds. Evolving genes and proteins. Academic Press, New York.

RODNEY L. HONEYCUTT, reviewing editor