The Complete Sequence of the Mitochondrial Genome of *Buteo buteo* (Aves, Accipitridae) Indicates an Early Split in the Phylogeny of Raptors

Elisabeth Haring, Luise Kruckenhauser,* Anita Gamauf,* Martin J. Riesing,* and Wilhelm Pinsker†*

*Zoological Department, Museum of Natural History, Vienna, Austria; and †Institute of Medical Biology, University of Vienna, Austria

The complete sequence of the mitochondrial (mt) genome of *Buteo buteo* was determined. Its gene content and nucleotide composition are typical for avian genomes. Due to expanded noncoding sequences, *Buteo* possesses the longest mt genome sequenced so far (18,674 bp). The gene order comprising the control region and neighboring genes is identical to that of *Falco peregrinus*, suggesting that the corresponding rearrangement occurred before the falconid/accipitrid split. Phylogenetic analyses performed with the mt sequence of *Buteo* and nine other mt genomes suggest that for investigations at higher taxonomic levels (e.g., avian orders), concatenated rRNA and tRNA gene sequences are more informative than protein gene sequences with respect to resolution and bootstrap support. Phylogenetic analyses indicate an early split between Accipitridae and Falconidae, which, according to molecular dating of other avian divergence times, can be assumed to have taken place in the late Cretaceous 65–83 MYA.

Introduction

Among birds of prey, buzzards and hawks of the genus *Buteo* are an extremely successful group that is widely distributed, being absent only in Australia, Antarctica, and most parts of the oriental region. The genus currently comprises between 25 (Brown and Amadon 1968) and 27 species (del Hoyo, Elliott, and Sargatal 1994), thus representing about 7% of the species in the family Accipitridae. In the present study, we focus on the common buzzard *Buteo buteo* (Linnaeus, 1758), which is the most abundant raptor species in many parts of Europe (Bijlsma 1997).

The two largest families of birds of prey are the Accipitridae and the Falconidae. The accipitrids, known colloquially as hawks, kites, harriers, vultures, and eagles, are rather similar in their basic morphological structures, although they show great diversity in size, shape, flying ability, ecology, and predatory habits. The falconids resemble the accipitrids in some characteristics, such as a powerful hooked bill, a fleshy cere straddling the bill, heavy bony brow ridges, and a crop (to store freshly eaten food). The differences between the two families have been summarized by Olsen (1995) in a survey of 25 anatomical and behavioral traits.

According to traditional morphological classifications (e.g., Brown and Amadon 1968; Storer 1971; Stresemann and Amadon 1979; Cracraft 1981), Accipitridae and Falconidae belong to the order Falconiformes. However, based on detailed morphological studies of several families, Jollie (1976, 1977) concluded that falconids and accipitrids are not closely related. According to his interpretation, the order Falconiformes is polyphyletic, especially with respect to the inclusion of New World vultures (Cathartidae); that view is supported by studies of several behavioral traits (König 1982) and, more recently, by molecular analyses (e.g., Sibley and Ahlquist 1990; Seibold and Helbig 1995; Wink et al. 1998). Sibley and Ahlquist (1990) estimated the overall genomic similarity by DNA-DNA hybridization and proposed a new classification of birds in which the New World vultures appear as close relatives of the storks (Ciconiidae). In their classification, the falconiform taxa are placed within an expanded order Ciconiiformes, in which they include the infraorders Falconides (including Falconidae and Accipitridae) and Ciconiides (including the family Ciconiidae with the subfamilies Cathartinae and Ciconiinae). Sequence analyses of the *cytochrome b* gene (*cyt b*) (Avise, Nelson, and Sibley 1994; Seibold and Helbig 1995; Wink et al. 1998) support this taxonomic position for New World vultures, but the relationships of the Ciconiidae with respect to Accipitridae and Cathartidae have not been resolved unambiguously. So far, no other genetic markers (mitochondrial or nuclear) have been employed to elucidate the phylogenetic relationships of these families.

Analyses of complete mitochondrial (mt) genomes provide not only sequence data for phylogenetic studies, but also information about structural genomic rearrangements which may serve as additional markers. Sequencing of the first complete vertebrate mt genomes suggested that gene content and gene order are highly conserved, but subsequent sequence data have demonstrated that the gene order in vertebrates is not uniform (for review, see Quinn 1997). A major rearrangement within the mt genome of chickens and other galliform birds has been described by Desjardins and Morais (1990). It comprises the *cyt b* gene, the NADH dehydrogenase subunit 6 gene (*nd6*), and several tRNA genes. Subsequently, this particular gene order has been found in several other avian species, suggesting that this rearrangement could have occurred at the base of the avian branch and thus might be shared by all recent bird species. However, the hypothesis of a universal gene order characteristic for all birds was refuted by the discovery of yet another rearrangement of the mt genes in *Falco peregrinus*, as well as in birds of four additional orders (Mindell, Sorenson, and Dimcheff 1998). In an inves-

Key words: *Buteo buteo*, mitochondrial genome, avian phylogeny, gene order, control and pseudo control regions.

Address for correspondence and reprints: Elisabeth Haring, 1. Zoological Department, Museum of Natural History Vienna, Burgring 7, A-1014 Vienna, Austria. E-mail: elisabeth.haring@nhm-wien.ac.at.

Mol. Biol. Evol. 18(10):1892–1904. 2001

Table 1
PCR Primers Used

Primer	Sequence (5′–3′)	Binding Site[a]	Reference[b]
12S-3+	CCCAAGACGACCCTAGCGC	305	ps
16S-3−	CTTTTCTATCGCCTATACTAAG	1,214	ps
12S-2+	AAGTCGTAACAAGGTAAGTGT	998	ps
16S-2−	ATCCCTGGGGTAGCTTGGTCC	2,407	ps
16S-1+	CGAGAAGACCCTATGGAGCTT	2,172	K, L2738
nd2-2−	TGATGAGAAGGCTAGGATTTTTCG	4,543	S, H5766
nd2-1+	GGATGAATAGGACTCAACCAGAC	4,490	S, H5758
cox2-1−	TCGTAAGCTCAATATCATTGGTGTCC	7,404	S, H8628
cox2-3+	ATCCTACCAGCAATTGTCCTC	7,277	ps
atp6-2−	ATGTTTTCTTGTTAGGTATAGG	8,673	ps
atp6-1+	ACGTCTTCGTCCTCCTACTAAG	8,631	ps
nd4-4−	GCTTTCTAGGCATAGTAGGGC	10,060	ps
nd3-1+	CAAGGAGGACTAGAGTGAGCAG	9,854	ps
nd4-5−	ATGGTTAGTTCTGCCATTAGG	11,387	ps
nd4-3+	ACCAACTACGAGCGGACACACAG	11,239	ps
nd5-2−	ATGATTCCCACTCCTTCTCAGCC	12,285	ps
nd5-3+	AATTCGCAACATGATACATAGC	12,141	ps
cytb-2−	TGTACGTTTCGGCATGTGTGGGC	13,887	ps
cytb-3+	ACTACCCTAGCCTTCTCGTC	13,841	ps
cytb-4−	TAGGTGAGGGAAGCTAGTTG	14,739	ps
cytb-1+	ACCCATTCATCATCATTGGC	14,700	H
nd6-3−	CGGTTGGATTTTAGTGGTGTTGC	16,770	ps

[a] The binding sites (5′ position) within the mitochondrial genome of *Buteo* are given.
[b] K = Knight and Mindell (1993); S = Sorenson et al. (1999); H = Haring et al. (1999); ps = present study.

tigation of 137 species, representing 13 orders, Mindell, Sorenson, and Dimcheff (1998) hypothesize that this novel arrangement, which includes the control region (CR) and surrounding sections, must have originated independently four times in avian evolution. In addition to the CR, these species possess a second noncoding (nc) region, probably generated through a duplication process. The same arrangement was recently detected in warblers of the genus *Phylloscopus* (Bensch and Härlid 2000). Partial sequence analysis of the *B. buteo* mt genome (Haring et al. 1999) revealed the existence of a noncoding section corresponding to the nc region of *F. peregrinus*, which was designated a pseudo control region (ΨCR). Although the position of the functional CR in *B. buteo* was not determined, this finding suggested that *F. peregrinus* and *B. buteo* might share the same gene order.

In this paper, we report the complete sequence of the mt genome of *B. buteo*. For sequence comparisons, we used the previously published complete mt genomes of nine other bird species, along with that of *Alligator mississippiensis* as an outgroup, to address the following questions: (1) Do Falconidae and Accipitridae share the same rearrangement within their mt genomes? (2) What are the phylogenetic positions of the genera *Buteo, Ciconia*, and *Falco* within the avian tree? (3) Are subsections as useful as complete mt genomes for resolving phylogenies? (Up to now, the majority of avian molecular phylogenies have been based on sequence data of the mitochondrial *cyt b* gene.) (4) What is the degree of divergence between *Buteo* and *Falco*, and can it be related to other splits in the phylogeny of birds to estimate their approximate divergence time?

Materials and Methods

DNA Extraction

DNA from *B. buteo* was extracted from the blood (stored in EDTA buffer) of a single specimen (specimen b.but-2, common buzzard, *B. buteo buteo*; Haring et al. 1999) by overnight incubation at 37°C in extraction buffer (10 mM Tris-HCl [pH 8.0], 10 mM EDTA, 50 mM NaCl, 40 mM dithiothreitol, 1% SDS, 0.5 mg/ml proteinase K). DNA was purified by two phenol/chloroform/isoamylalcohol (25:24:1) extractions and one chloroform/isoamylalcohol (1:1) extraction, followed by precipitation with 1/10 volumes 3 M NaAc, 3 × volumes EtOH.

PCR Amplification

PCR was carried out with an Eppendorf Thermocycler in a volume of 25 μl containing 1 unit Dynazyme DNA polymerase (Finnzymes OY), 1 μM of each primer, 0.2 mM of each dNTP, and 50–200 ng template DNA. The solutions were heated to 95°C (2 min) and then put through 30 reaction cycles: 95°C (10 s), annealing temperature (10 s), and 72°C (1 min/1 kb expected length), followed by a final extension at 72°C (10 min). Primers used for PCR amplification are listed in table 1. The 11 PCR fragments that were subsequently cloned and sequenced are depicted in figure 1. These overlapping fragments cover the whole mt genome of *B. buteo* except the region comprising a part of the *nd*6 gene, $tRNA^{Glu}$, ΨCR, $tRNA^{Phe}$, and a part of the *12S* rRNA gene (primer pair nd6-1+/12S-1−) determined in a previous study from the same individual of *B. buteo* (GenBank accession number AF202186). The 11 fragments sequenced in the present study were obtained us-

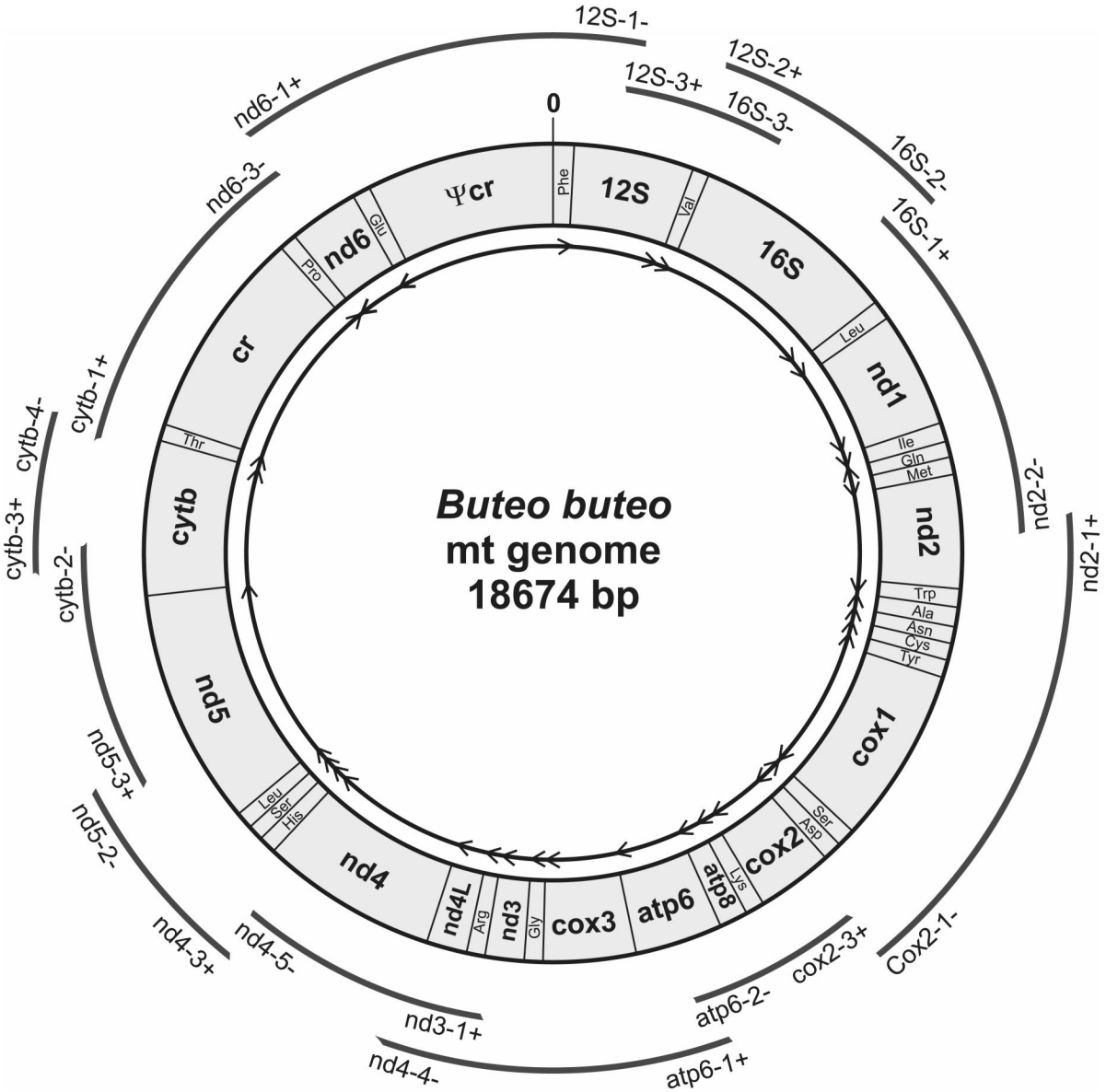

FIG. 1.—Mitochondrial genome of *Buteo*. Overlapping PCR fragments are depicted with their corresponding primers. Arrows at the inner circle indicate orientation of genes. For designation of tRNAs, the three-letter code for amino acids is used. Position 0 of the complete sequence (table 2) is indicated.

ing the following primer combinations (fig. 1): 12S3+/16S3− (910 bp), 12S-2+/16S-2− (1,410 bp), 16S-1+/nd2-2− (2,372 bp), nd2-1+/cox2-1− (2,913 bp), cox2-3+/atp6-2− (1,397 bp), atp6-1+/nd4-4− (1,430 bp), nd3-1+/nd4-5− (1,534 bp), nd4-3+/nd5-2− (1,047 bp), nd5-3+/cytb-2− (1,747 bp), cytb-3+/cytb-4− (899 bp), cytb-1+/nd6-3− (2,071 bp).

Cloning and Sequencing

PCR products were extracted from agarose gels using the QIAquick Gel Extraction Kit (QIAGEN) and cloned (TOPO TA Cloning Kit, Invitrogen). Sequencing of the clones (both directions, M13 primers) was performed by MWG-Biotech (Ebersberg, Germany) with a Li-Cor Sequencer. Due to overlapping clones and overlap of regions sequenced with internal primers, the whole mitochondrial genome was read at least twice (some sections four times). Thus, the complete sequence was determined without any ambiguities.

Sequence Analysis

Alignments of DNA and amino acid sequences were produced with the program CLUSTAL X (Thompson et al. 1997) and improved manually. Distance-based (the neighbor-joining [NJ] algorithm; Saitou and Nei 1987) and maximum-parsimony (MP) methods were used to infer the phylogenetic relationships. All dendrograms were calculated with the software package PAUP (test version 4.0b3a; Swofford 1998) using a heuristic search with the tree bisection-reconnection (TBR) algorithm and a random-taxon-addition sequence. Gaps were treated as ''missing,'' and all characters were

weighted equally. The robustness of trees was tested with bootstrap analysis (100 replications). Relative-rate tests were performed with the program Phyltest (Kumar 1996), which provides the two-cluster test of Takezaki, Rzhetsky, and Nei (1995). In this test, the constancy of substitution rates is examined for two lineages with a given outgroup lineage. The program allows inclusion of multiple sequences in each of the lineages. If L_a and L_b are the averages of observed numbers of substitutions per site from the common ancestor of clusters A and B, then $L_a = L_b$ is the null hypothesis under rate constancy ($\delta = L_a - L_b = 0$). Because the variance of δ can be estimated, the deviation of δ from 0 can be tested by a two-tailed normal deviate test. For the tests, new alignments including only the sequences used were created, and Kimura (1980) two-parameter distances were used for the calculations.

DNA Sequences

For phylogenetic comparisons, the complete mt genome sequences of the following avian species were retrieved from GenBank: *Gallus gallus* f. *dom.*, NC001323 (Desjardins and Morais 1990); *Struthio camelus,* Y12025 (Härlid, Janke, and Arnason 1997); *Rhea americana,* NC000846 (Härlid, Janke, and Arnason 1998); *Vidua chalybeata,* NC000880; *Aythya americana,* AF090337; *Falco peregrinus,* NC000878; *Smithornis sharpei,* NC000879 (Mindell, Sorenson, and Dimcheff 1998; Mindell et al. 1999); *Corvus frugilegus,* NC002069 (Härlid and Arnason 1999); *Ciconia ciconia,* AB026818 (Yamamotu, unpublished). The sequence of *Alligator mississippiensis,* Y13113 (Janke and Arnason 1997), was used as an outgroup for the rooting of phylogenetic trees.

Designation of Sequences

Although the sequences of the various avian mt genomes used in this study are derived from defined taxonomic entities (species/subspecies), e.g., *B. b. buteo,* the phylogenetic reconstructions were made at the genus level or at even higher taxonomic levels. Therefore, for the sake of brevity, we refer to the taxa only by their genus names in the text. In the tables and figures, the following abbreviations are used: All, *A. mississippiensis*; Ayt, *A. americana*; But, *B. buteo*; Cic, *C. ciconia*; Cor, *C. frugilegus*; Fal, *F. peregrinus*; Gal, *G. gallus*; Rhe, *R. americana*; Smi, *S. sharpei*; Str, *S. camelus*; Vid, *V. chalybeata.*

Results

Genome Organization

The complete mitochondrial genome of *Buteo* is 18,674 bp in size. In comparison with other complete avian mt genomes, which range from 16,591 bp (*Struthio*) to 18,068 bp (*Falco*), *Buteo* has the largest mt genome sequenced so far. This considerable difference in size is due to an expansion of noncoding sequences. The three bird species with the largest portions of noncoding mt DNA (*Buteo,* 17.1%; *Falco,* 14.1%; *Smithornis,* 10.7%) possess, in addition to the CR, a second nc region (ΨCR). The gene content of the *Buteo* mt genome is typical of vertebrates (fig. 1 and table 2), consisting of genes for 13 proteins (12 H-strand-encoded, 1 L-strand-encoded), 22 tRNAs, and 2 rRNAs.

The nucleotide composition of the *Buteo* mt genome (H-strand) is similar to that of other avian species (A = 31.1%, C = 31.7%, G = 13.2%, T = 24.0%). The A+T content of 55.1% is within the range for avian mt genomes (51.6%–55.7%). The usage of translation initiation and termination signals in comparison with other bird species is given in table 3. The most common start codon is ATG. In *Buteo,* nonstandard start codons are found in the *cox1* and *nd5* genes (GTG) and in the *nd3* gene (ATC). The unusual ATC start codon in the *nd3* gene is also found in *Smithornis.* As in the mt genomes of the other birds and *Alligator,* TAA is the most frequent stop codon in *Buteo*; TAG and AGG are used twice. In *cox3* and *nd4,* a terminal T probably serves as the stop signal after it is completed to UAA by posttranscriptional polyadenylation (Ojala, Montoya, and Attardi 1981). *Buteo* and *Falco* differ in two start codons and in three stop codons. With respect to the length of intergenic spacers and overlaps, *Buteo* (67-bp spacers) has a rather compact genome compared with *Falco* (101-bp spacers) (table 4).

Gene Order and Noncoding Regions

In contrast to the standard avian gene order (e.g., *Gallus*), the CR of *Buteo* is located between $tRNA^{Thr}$ and $tRNA^{Glu}$. Between $tRNA^{Glu}$ and $tRNA^{Phe}$, at the position of the CR in the majority of the birds analyzed so far, another nc section is found in *Buteo* which was designated a ΨCR by Haring et al. (1999). It consists of a short nonrepetitive part (23.2% of its length) followed by 48-bp tandem repeats (76.8%). A similar array of nc sections (designated CR and nc) has been described for the mt genome of *F. peregrinus* (Mindell, Sorenson, and Dimcheff 1998). Thus, gene order and content are identical in *Buteo* and *Falco* but differ from the standard gene order of many other birds (including *Ciconia, Gallus, Aythya, Rhea, Struthio, Vidua,* and *Corvus*) with respect to CR and ΨCR (table 5). The gene order of *Buteo* and *Falco* is similar to that of *Smithornis* (Mindell, Sorenson, and Dimcheff 1998) and *Phylloscopus* (Bensch and Härlid 2000), but the ΨCR regions of these species lack any repetitive parts. An appreciable degree of sequence identity is found between *Buteo* and *Falco* in the nonrepetitive sections (fig. 2) of the CR (71.4%, 691 bp; uncorrected, gaps ≥5 excluded) and the ΨCR (58.2%, 189 bp; gaps ≥5 excluded). No sequence similarity was detectable in an intragenomic comparison between CR and ΨCR of *Buteo.* Some structural features of the nc sequences and their tandem-repetitive sections are compared in table 6. The majority of birds listed in this table have tandem repeats within their CRs. The CR of *Buteo* contains seven different types of repeats; that of *Falco* contains only two. The mt genome of *Buteo* has the largest percentage of tandem repetitive sequences (9.6%). We could find no sequence relationships be-

Table 2
Organization of the Mitochondrial Genome of *Buteo buteo*

Gene/Region	Start Position	End Position	Start Codon	End Codon	bp	aa	Strand[a]
tRNA-Phe	1	70	—	—	70	—	H
12SrRNA	71	1,042	—	—	972	—	H
tRNA-Val	1,043	1,113	—	—	71	—	H
16SrRNA	1,114	2,711	—	—	1,598	—	H
tRNA-Leu-UUR	2,712	2,785	—	—	74	—	H
nd1	2,795	3,772	ATG	AGG	978	325	H
tRNA-Ile	3,771	3,842	—	—	72	—	H
tRNA-Gln	3,926	3,856	—	—	71	—	L
tRNA-Met	3,926	3,994	—	—	69	—	H
nd2	3,995	5,035	ATG	TAG	1,041	346	H
tRNA-Trp	5,034	5,105	—	—	72	—	H
tRNA-Ala	5,175	5,107	—	—	69	—	L
tRNA-Asn	5,250	5,178	—	—	73	—	L
tRNA-Cys	5,319	5,253	—	—	67	—	L
tRNA-Tyr	5,389	5,319	—	—	71	—	L
cox1	5,391	6,941	GTG	AGG	1,551	516	H
tRNA-Ser-UCN	7,006	6,933	—	—	74	—	L
tRNA-Asp	7,011	7,079	—	—	69	—	H
cox2	7,082	7,765	ATG	TAA	684	227	H
tRNA-Lys	7,767	7,834	—	—	68	—	H
atp8	7,836	8,003	ATG	TAA	168	55	H
atp6	7,994	8,677	ATG	TAA	684	227	H
cox3	8,677	9,460	ATG	T++	784	261	H
tRNA-Gly	9,461	9,529	—	—	69	—	H
nd3	9,530	9,880	ATC	TAA	351	116	H
tRNA-Arg	9,885	9,954	—	—	70	—	H
nd4L	9,956	10,252	ATG	TAA	297	98	H
nd4	10,246	11,623	ATG	T++	1,378	459	H
tRNA-His	11,624	11,693	—	—	70	—	H
tRNA-Ser-AGY	11,694	11,759	—	—	66	—	H
tRNA-Leu-CUN	11,760	11,830	—	—	71	—	H
nd5	11,831	13,648	GTG	TAA	1,818	605	H
cytb	13,664	14,806	ATG	TAA	1,143	380	H
tRNA-Thr	14,809	14,878	—	—	70	—	H
CR	14,879	16,550	—	—	1,672	—	—
tRNA-Pro	16,620	16,551	—	—	70	—	L
nd6	17,145	16,627	ATG	TAG	519	172	L
tRNA-Glu	17,149	17,219	—	—	71	—	L
ΨCR	17,220	18,674	—	—	1,455	—	—

[a] H = heavy strand; L = light strand.

Table 3
Usage of Start and Stop Codons

	But	Fal	Gal	Ayt	Smi	Vid	Cor	Rhe	Str	All
Start										
ATG	10	11	12	10	11	12	12	10	9	7
GTG	2	1	1	3	—	—	1	2	3	1
ATC	1	—	—	—	1	—	—	—	—	1
ATA	—	1	—	—	1	—	—	1	1	3
ATT	—	—	—	—	—	1	—	—	—	1
Stop										
TAA	7	8	9	7	8	7	8	6	6	7
TAG	2	1	1	2	1	1	—	3	2	1
T−−	2	2	2	2	2	2	2	2	2	3
AGG	2	1	1	2	2	1	2	2	2	2
AGA	—	1	—	—	—	2	1	—	1	—

NOTE.—All = *Alligator mississippiensis*; Ayt = *Aythya americana*; But = *Buteo buteo*; Cic = *Ciconia ciconia*; Cor = *Corvus frugilegus*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Rhe = *Rhea americana*; Smi = *Smithornis sharpei*; Str = *Struthio camelus*; Vid = *Vidua chalybeata.*

tween the repeat units in the CR and the ΨCR within or between the mt genomes of *Buteo* and *Falco.* In both species, the largest repetitive block is located in the ΨCR (table 6), but the repeat units differ considerably in length (48 vs. 27 bp). We sequenced the ΨCRs of two other accipitrids (*Aquila chrysaetos* and *Haliaeetus albicilla*) which had the same gene order as *Buteo* and *Falco.* Although the repeat units in these species resemble that of *Buteo* in length (*Aquila*, 49 bp; *Haliaeetus*, 48 bp), there is no apparent relatedness among them at the sequence level.

An alignment of the nonrepetitive parts of the CRs of *Buteo*, *Falco*, and *Ciconia* (not shown) revealed several conserved sections (fig. 2). Some of them show similarities to conserved sequence blocks (CSBs) of possible functional importance (CSB1, E box, D box) which have previously been described by Walberg and Clayton (1981), Baker and Marshall (1997), Randi and Lucchini (1998), and Clayton (1991). The section designated ''C stretch'' is located close to the 5′ end and corresponds to the ''goose hairpin'' in other avian species (Quinn and Wilson 1993; Randi and Lucchini 1998), which con-

Table 4
Spacer and Overlaps

Genes	But	Fal	Gal	Ayt	Smi	Vid	Cor	Rhe	Str	All
tRNA-Leu-UUR	9	15	9	4	10	8	9	11	9	‡
nd1	−2	16	−2	−2	21	6	8	8	−2	−1
tRNA-Ile	13	9	5	8	11	5	5	14	11	−2
tRNA-Gln	−1	−1	−1	−1	12	−1	−1	−1	−1	−1
tRNA-Met	‡	1	‡	‡	‡	‡	‡	‡	‡	4
nd2	−2	−1	−2	−2	1	−1	2	−2	−2	‡
tRNA-Trp	1	10	6	3	1	1	1	1	‡	3
tRNA-Ala	2	10	3	2	6	9	9	2	4	1
tRNA-Asn	2	2	1	‡	6	‡	2	3	1	4
tRNA-Cys	−1	−1	−1	−1	−1	−1	−1	−1	−1	‡
tRNA-Tyr	1	1	1	1	1	1	1	1	1	1
cox1	−9	−9	−9	−9	−9	−9	−9	−9	−9	−5
tRNA-Ser-UCN	4	2	2	2	2	5	5	2	3	6
tRNA-Asp	2	6	1	1	3	11	8	1	1	‡
cox2	1	1	1	1	1	11	1	‡	‡	‡
tRNA-Lys	1	1	1	1	1	1	1	1	1	2
atp8	−10	−10	−10	−10	−10	−10	−10	−10	−10	−4
atp6	−1	−1	−1	−1	9	5	7	−1	−1	−1
cox3	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
tRNA-Gly	‡	‡	‡	‡	‡	‡	‡	‡	‡	‡
nd3	4	1	1	1	4	1	1	2	1	29
tRNA-Arg	1	1	‡	‡	1	1	1	‡	‡	2
nd4L	−7	−7	−7	−7	−7	−7	−7	−7	−7	−7
nd4	‡	‡	‡	‡	‡	‡	‡	1	1	23
tRNA-His	‡	‡	‡	‡	1	‡	‡	‡	‡	7
tRNA-Ser-AGY	‡	−1	‡	−1	−1	−1	−1	−1	−1	11
tRNA-Leu-CUN	‡	‡	‡	‡	‡	‡	‡	‡	‡	2
nd5	15	5	4	−1	9	8	10	12	9	−5
cytb	2	2	3	2	4	3	3	5	2	32
tRNA-Pro	6	15	6	10	5	26	9	15	10	8
nd6	3	3	2	‡	‡	1	1	3	3	‡
Total of spacers	67	101	46	36	109	103	84	82	57	135
Total of overlaps	−33	−31	−33	−35	−28	−30	−29	−32	−34	−26

NOTE.—The lengths of spacers between adjacent tRNA or protein-coding genes are given in base pairs. The numbers relate to the spacers at the 3′ ends of the respective genes. Negative numbers represent overlaps; ‡ indicates exact fit. Spacers flanked by rRNA genes or noncoding regions were excluded because their limits could not be determined unambiguously. All = *Alligator mississippiensis*; Ayt = *Aythya americana*; But = *Buteo buteo*; Cic = *Ciconia ciconia*; Cor = *Corvus frugilegus*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Rhe = *Rhea americana*; Smi = *Smithornis sharpei*; Str = *Struthio camelus*; Vid = *Vidua chalybeata.*

sists of a stem of seven complementary C's/G's and a loop containing a TCCC motif that may be involved in H-strand termination (Dufresne, Mignotte, and Guéride 1996). Yet, in the CRs of *Buteo, Falco,* and *Ciconia,* the C stretch is not followed by a G stretch, and thus the motif lacks the ability to form a hairpin structure. The predicted TCCC stop motif within the C stretch is found in *Buteo* and *Falco,* whereas in *Ciconia* the corresponding motif is ACCC. Downstream of the C stretch, a 41-bp section corresponding to the consensus sequence of the mammalian ETAS1 (extended termination-associated sequence) described by Sbisà et al. (1997) is found in all three species, but there are no sequences capable of forming hairpin structures. The *Buteo* ETAS1 contains two termination-associated sequence (TAS) motifs, TACAT and TATAT, whereas in *Falco* and *Ciconia* such motifs are absent. Five additional CSBs are found in the three species: (1) a 34-bp section that includes a sequence similar to the E box, although this motif is less conserved than the rest of the section; (2) a 25-bp stretch with high similarity to the D box; (3) a 40-bp section, designated CSB-a, which has no similarity to other described CSBs; (4) a 31-bp section, designated CSB-b, with no similarity to other CSBs (in *Buteo,* the section surrounding CSB-b is duplicated [97-bp repeat; table 6], and therefore this motif

Table 5
Different Gene Orders in Bird and the Alligator

But	His	Ser	Leu	ND5	Cyb	Thr	CR	Pro	ND6	Glu	ΨCR	Phe
Fal	His	Ser	Leu	ND5	Cyb	Thr	CR	Pro	ND6	Glu	nc	Phe
Smi	His	Ser	Leu	ND5	Cyb	Thr	CR	Pro	ND6	Glu	nc	Phe
Gal	His	Ser	Leu	ND5	Cyb	Thr	—	Pro	ND6	Glu	CR	Phe
All	Ser	His	Leu	ND5	ND6	Glu	Cyb	Thr	Pro	—	Phe	CR

NOTE.—Only the rearranged section is depicted. Genes coding for tRNAs are named according to their corresponding amino acids. All = *Alligator mississippiensis*; But = *Buteo buteo*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Smi = *Smithornis sharpei.*

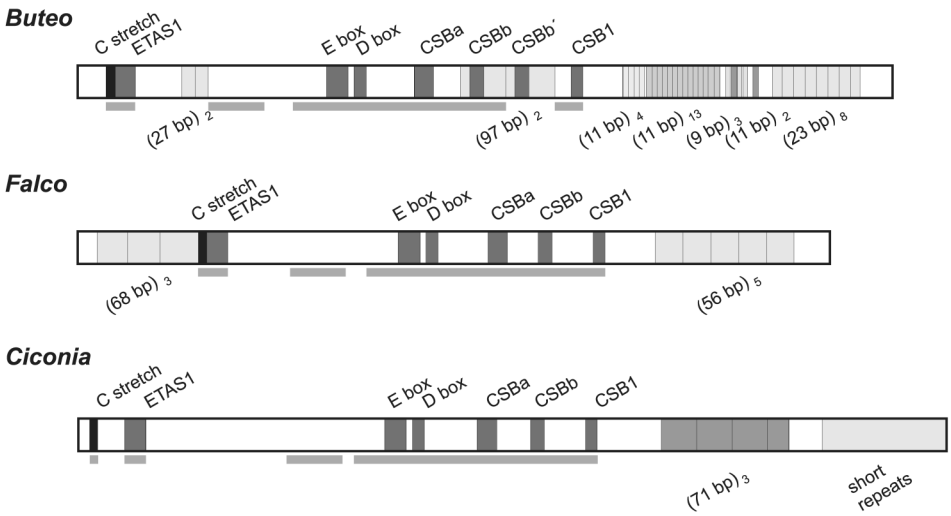

FIG. 2.—Comparison of the control regions (CRs) of *Buteo, Falco,* and *Ciconia.* Conserved sequence blocks (CSBs) are depicted in black, and repetitive parts are depicted in gray. Designations of CSBs (above CR) are described in the text; designations of repeats (below CR) correspond to table 5. Nonrepetitive sections used for the calculation of sequence divergence between *Buteo, Falco,* and *Ciconia* are marked by gray bars.

is present in two copies); and (5) CSB1, which is well conserved. The following motifs conserved in other avian species cannot be identified in *Buteo, Falco,* or *Ciconia*: CSB2 and CSB3 (Walberg and Clayton 1981), a polyC sequence upstream of CSB1 assumed to represent the origin of H-strand replication (Walberg and Clayton 1981), and the putative bidirectional promoter (L'Abbé et al. 1991).

Phylogenetic Analyses

The phylogenetic analysis was carried out using the coding regions of *Buteo* and the nine bird species listed above and *Alligator*. Each gene was aligned separately after excluding nonmatching positions at the length-variable 3′ ends. Although the *Alligator* sequence was considerably diverged from the avian sequences, a reasonable alignment could be achieved. In a distance matrix (not shown) calculated from the concatenated coding sequences (15,742 bp), the distances (HKY85; Hasegawa, Kisgino, and Yano 1985) between all ingroup taxa and *Alligator,* respectively, are rather homogenous, indicating no major differences in substitution rates. The distances among ingroup taxa range from 18.1% to 27.2%. The distance between *Buteo* and *Falco* (21.3%) is on the same order of magnitude as that between the two struthioniform species *Rhea* and *Struthio* (19.6%) and between the two galloanserine species *Gallus* and *Aythya* (22.3%). Concerning the phylogenetic relationships of *Buteo,* it is remarkable that with 11 out of 14 genes compared, *Buteo* appears more closely related to *Ciconia* than to *Falco* (table 7), an affiliation that does not conform with classical taxonomy. The divergence of the total coding sequence is more extensive between *Buteo* and *Falco* than between *Buteo* and *Ciconia* ($\chi^2 = 34.86$, df = 1, $P < 0.001$).

Table 7 shows the uncorrected distances (DNA and proteins) between pairs of related species calculated for different sections of the mt genome. In general, for each region, the values observed for the six species pairs are in the same range. In the *Buteo*/*Falco* comparison, the most conserved regions are the tRNAs (with respect to 16S rRNA: $\chi^2 = 7.55$, df = 1, $P < 0.01$; with respect to *cox1*: $\chi^2 = 6.72$, df = 1, $P < 0.01$). Among the protein-coding genes, *cox(1–3) + cyt b* are more conserved than *nd(1–6) + atp(8+6)* ($\chi^2 = 46.63$, df = 1, $P < 0.001$). Among amino acid sequences, COX(1–3) + CYTb are also more conserved than ND(1–6) + ATP(8+6) ($\chi^2 = 93.18$, df = 1, $P < 0.001$).

An MP dendrogram based on the complete coding sequence (15,742 bp) is depicted in figure 3. Three species pairs are supported by high bootstrap values: *Gallus*/*Aythya*, *Rhea*/*Struthio*, and *Corvus*/*Vidua*. In the cluster of the three ciconiiform species, *Buteo* and *Falco* appear as sister taxa, although with weak bootstrap support. The passeriform split is at the base of the tree, yet Passeriformes do not form a monophyletic group, since *Smithornis* branches off as the most basal bird taxon. The corresponding NJ dendrogram (not shown) resembles the MP dendrogram with two exceptions: *Buteo* clusters with *Ciconia* (bootstrap value = 97), and *Smithornis* is not the sister group to all other ingroup taxa. Instead, there is a trichotomy of the three lineages (*Smithornis*/*Corvus–Vidua*/remaining avian taxa). In both the MP and the NJ dendrograms, the passeriform taxa are placed at the base of the birds, and the two pairs *Gallus*/*Aythya* and *Rhea*/*Struthio* are grouped as one clade that is the sister group of the cluster *Buteo*/*Falco*/*Ciconia*. Dendrograms were also calculated from deduced protein sequences and from DNA sequences using transversions only (data not shown). However, in general, better resolution and stronger bootstrap support were obtained with DNA sequences that included all substitutions, so those were used for subsequent phylogenetic analysis.

To estimate in detail which clusterings were supported by different sections of the mt genome, MP analyses were performed for each gene separately, as well

Table 6
Comparison of Genome Size and Noncoding (NC) Sections

	Total bp	NC %	Spacer bp	CR		ΨCR		CR Tandem Repeats	ΨCR Tandem Repeats
				Length (bp)	Tan-rep %	Length (bp)	Tan-rep %		
But	18,674	17.1	67	1,672	40.3	1,455	76.8	2×27, 2×97, 4×11, 13×11(+8), 3×9, 2×11, 8×23	23×48(+13)
Fal	18,068	14.1	101	1,510	32.8	941	82.0	3×68, 5×56(+11)	28×27(+24)
Cic	17,347	10.7	76	1,779	28.7	—	—	3×71(+43), 255 bp short repeats	—
Gal	16,775	7.6	46	1,228	11.7	—	—	2×53(8-bp spacer), 2×19	—
Ayt	16,616	6.7	46	1,066	—	—	—	—	—
Smi	17,344	10.7	109	1,454	20.9	301	0.0	5×56(+24)	—
Vid	16,895	8.3	115	1,295	—	—	—	—	—
Cor	16,932	8.5	95	1,337	—	—	—	—	—
Rhe	16,714	7.6	91	1,172	11.7	—	—	2×26(+13), 18×4	—
Str	16,591	6.9	112	1,031	—	—	—	—	—
All	16,646	6.9	153	988	37.6	—	—	3×22 + 14×21	—

Note.—Tan-rep = tandem repetitive section; 13×11(+8) = 13 repeats of 11 bp (+ one incomplete repeat of 8 bp). All = *Alligator mississippiensis*; Ayt = *Aythya americana*; But = *Buteo buteo*; Cic = *Ciconia ciconia*; Cor = *Corvus frugilegus*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Rhe = *Rhea americana*; Smi = *Smithornis sharpei*; Str = *Struthio camelus*; Vid = *Vidua chalybeata*.

as for combinations of genes (e.g., rRNAs + tRNAs; *nd1–6*; complete coding sequence). The three species pairs observed in the majority of these dendrograms are *Rhea*/*Struthio*, *Corvus*/*Vidua*, and *Gallus*/*Aythya*, which is in accordance with traditional classification. However, the position of *Smithornis* varied considerably. Often, it appeared as the sister taxon of the two other passeriform species *Corvus*/*Vidua*, as expected, but in some cases it clustered with other clades or else branched off at the base of the ingroup taxa. The results are summarized in table 8, where bootstrap values >50 supporting different species pairs are given. One gene, *cox3*, does not support any of the splits; two others, *nd3* and *nd6*, support only one, namely, *Rhea*/*Struthio* and *Corvus*/*Vidua*, respectively. Two genes (*12S* and *nd5*) support the *Buteo*/*Ciconia* clade, but with comparatively low bootstrap values.

Altogether, the relationships among the ciconiiform species *Buteo*/*Falco*/*Ciconia* are not clearly resolved. To test whether the use of *Alligator* as an outgroup taxon might influence the branching pattern of these three species, MP dendrograms with different outgroup taxa (or sets of taxa) were calculated. First, the Alligator sequence was omitted, and the clade *Smithornis*/*Vidua*/*Corvus* was used instead as an outgroup. Trees were calculated from the total coding sequence as well as from a combination of rRNAs and tRNAs. In these dendrograms, the *Buteo*/*Falco* cluster is supported by high bootstrap values. When the set of ingroup taxa is reduced to *Buteo*/*Falco*/*Ciconia* and one of the three species *Gallus*, *Rhea*, or *Vidua* is used as outgroup, the *Buteo*/*Falco* clade appears in all trees derived from tRNA and/or rRNA genes (bootstrap values = 56–88). In the trees obtained from the total coding sequence, *Buteo* clusters either with *Falco* or with *Ciconia*. Relative-rate tests (Takezaki, Rzhetsky, and Nei 1995) were carried out to test whether these discrepancies were due to differences in substitution rates among the *Buteo*/*Falco*/*Ciconia* lineages. Since the use of a closely related outgroup makes the statistical test more accurate and powerful, we used *Gallus*, a taxon of the sister clade, as an outgroup. First, tests were performed with two of the three species forming group A and the third representing group B. In these tests, the relative-rate test *Z*-statistic ($Z = 0.055$) rejected rate constancy ($P < 0.05$) only for the constellation *Falco*+*Buteo*/*Ciconia*. When only two of the three species were tested, rate constancy was always rejected at the 5% level. *Falco* had the highest rate, followed by *Buteo* and *Ciconia*. The *Z* values for the three tests were 2.69829 (*Buteo*/*Falco*), 3.04896 (*Buteo*/*Ciconia*), and 5.55873 (*Falco*/*Ciconia*).

Discussion

Gene Order and Noncoding Regions

The sequence data revealed that *Buteo* and *Falco* share the same mt gene order, one which is different from that of most other bird species. This difference could have arisen in the common ancestor of Falconidae and Accipitridae prior to their split, or it could have arisen independently, as has been postulated for other

Table 7
Sequence Divergences of Mitochondrial Genes

Gene	But/Fal	But/Cic	Cor/Vid	Rhe/Str	Gal/Ayt	Birds/All	bp
12SrRNA	16.6	12.9	11.1	11.3	14.6	29.9	1,036
16SrRNA	15.7	15.2	8.0	13.0	17.4	28.1	1,717
tRNAs	12.4	10.8	8.7	8.4	11.7	28.8	1,597
atp(8+6)	22.4	21.4	19.8	18.9	21.2	35.5	842
cox1	14.6	12.2	13.1	16.6	16.3	22.1	1,551
cox2	16.8	13.9	19.7	16.7	16.7	32.9	690
cox3	16.4	13.1	17.1	18.2	14.5	25.2	784
cytb	17.0	15.5	17.3	16.2	17.1	31.9	1,146
nd1	18.4	19.4	19.3	20.9	23.2	32.4	983
nd2	24.6	21.2	22.5	22.0	25.1	45.3	1,041
nd3	14.8	14.8	20.2	16.5	22.8	33.0	352
nd(4L+4)	22.2	15.9	18.2	20.6	22.4	39.1	1,671
nd5	21.3	16.9	19.0	19.0	23.9	40.2	1,818
nd6	24.5	25.2	25.9	27.1	23.4	51.5	528
Total coding	18.3	15.8	16.8	17.8	19.1	33.3	15,742
Protein	**But/Fal**	**But/Cic**	**Cor/Vid**	**Rhe/Str**	**Gal/Ayt**	**Bird/All**	**aa**
ATP(8+6)	19.1	17.4	14.9	11.0	14.9	38.5	282
COX(1–3)	6.8	3.7	6.7	2.9	4.9	20.7	1,007
CYTb	8.2	7.4	12.4	6.3	8.4	32.4	380
ND(1–6)	18.5	14.9	15.0	13.0	19.0	46.8	2,127
All proteins	14.4	11.4	12.5	9.5	13.9	37.8	3,796

NOTE.—Pairwise distances (uncorrected) determined from different genes and proteins are given for species pairs supposed to cluster according to traditional taxonomy and for the pairs *Buteo*/*Falco* and *Buteo*/*Ciconia.* Birds/All = average distance between birds and Alligator. Total coding = all RNA and protein-coding sequences. All = *Alligator mississippiensis*; Ayt = *Aythya americana*; But = *Buteo buteo*; Cic = *Ciconia ciconia*; Cor = *Corvus frugilegus*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Rhe = *Rhea americana*; Smi = *Smithornis sharpei*; Str = *Struthio camelus*; Vid = *Vidua chalybeata.*

bird lineages. As a possible mechanism for the rearrangement in *Falco*, a tandem duplication of the entire region including the CR and $tRNA^{Glu}$ with subsequent deletion of duplicated sequences except parts of the CR has been assumed (Mindell, Sorenson, and Dimcheff 1998). While in *Falco* the sequence similarity between CR and ΨCR corroborates this hypothesis, in *Buteo* no apparent sequence homology can be detected between the CR and the ΨCR. The CRs of *Buteo* and *Falco* can be well aligned, at least in the conserved parts, whereas no reasonable alignment could be achieved between the ΨCR of *Buteo* and the corresponding ΨCR region of *Falco*. Therefore, it cannot be decided from sequence divergence alone whether the similar rearrangements in *Buteo* and *Falco* are a shared derived character, nor can we trace unequivocally the events (duplication or transposition) that caused the rearrangements. Nevertheless, a strong argument in favor of a common origin comes from structural similarities between the ΨCRs of *Buteo* and *Falco.* Both sequences are composed of a short nonrepetitive section followed by a long stretch of tandem repeats. The same structure has been found within the ΨCRs of two other accipitrids, *A. chrysaetos* (Masuda et al. 1998) and *H. albicilla* (present study). In contrast, *Smithornis* (Mindell, Sorenson, and Dimcheff 1998) and *Phylloscopus* (Bensch and Härlid 2000), two passeriform species that have presumably acquired the same gene order independently, do not possess any repetitive sequences in the region corresponding to the ΨCR in *Buteo.* The sequences of the ΨCR sections of other species (belonging to the Cuculiformes, Piciformes, and Passeriformes) with rearranged gene orders (described by Mindell, Sorenson, and Dimcheff 1998) have not yet been published. Thus, an array of tandem repeats in the ΨCR section of the mt genome has been detected so far only in falconid and accipitrid species.

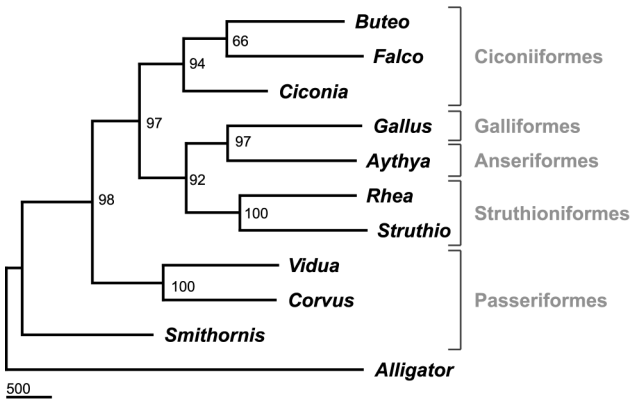

FIG. 3.—Single most-parsimonious tree (tree length = 21,916, consistency index = 0.588) for complete mt coding sequences of 10 avian taxa (15,742 bp) generated with a heuristic search using the tree bisection-reconnection (TBR) algorithm and a random-taxon-addition sequence. Gaps were treated as ''missing,'' and all characters were weighted equally. Alligator was used as the outgroup. Bootstrap values (%, 100 replicates) are given at the nodes. Branch lengths are proportional to nucleotide substitutions. The bar indicates 500 substitutions.

Relationships Among *Buteo*, *Falco,* and *Ciconia*

Although we included all avian mt genomes available, our main interest in this study was the group *Buteo*/*Falco*/*Ciconia.* Nevertheless, despite the huge amount of DNA sequence, the relationships within this triad were not clearly resolved. Whereas in the distance-based dendrogram of the complete coding sequence *Buteo* clusters with *Ciconia*, the MP analysis yielded the

Table 8
Comparison of Bootstrap Values

Gene	But/Fal	But/Cic	Cor/Vid	Rhe/Str	Gal/Ayt	bp
12SrRNA	—	54	100	92	—	1,036
16SrRNA	92	—	100	—	52	1,717
tRNAs	88	—	99	98	79	1,597
atp8+6	—	—	65	95	—	842
cox1	—	—	93	54	—	1,551
cox2	—	—	82	76	—	690
cox3	—	—	—	—	—	784
cytb	—	—	68	68	—	1,146
nd1	—		92	60	—	983
nd2	—	—	97	90	—	1,041
nd3	—	—	—	94	—	352
nd4L+4	—	—	100	97	—	1,671
nd5	—	79	97	98	—	1,818
nd6	—	—	95	—	—	528
RNAs(r+t)	96	—	100	99	68	4,350
cox-(1–3)	—	—	99	99	79	3,025
nd-(1–6)	—	—	100	100	—	6,408
Total coding	66	—	100	100	97	15,742

NOTE.—The bootstrap values (100 replicates) are shown for nodes joining species pairs that are supposed to cluster according to traditional taxonomy (*Corvus*/*Vidua*, *Rhea*/*Struthio*, *Gallus*/*Aythya*), as well as the pairs *Buteo*/*Falco* and *Buteo*/*Ciconia*. The maximum-parsimony dendrograms were calculated for single genes, groups of genes, and the total coding sequence. Total coding = all RNA and protein-coding sequences. All = *Alligator mississippiensis*; Ayt = *Aythya americana*; But = *Buteo buteo*; Cic = *Ciconia ciconia*; Cor = *Corvus frugilegus*; Fal = *Falco peregrinus*; Gal = *Gallus gallus*; Rhe = *Rhea americana*; Smi = *Smithornis sharpei*; Str = *Struthio camelus*; Vid = *Vidua chalybeata*.

expected group *Buteo*/*Falco*. Saturation effects are not the only explanation for the poor resolution, because the topology of the rest of the tree (with the exception of the position of *Smithornis*) seems to be unambiguous. It is more likely that unequal substitution rates among *Buteo*, *Falco*, and *Ciconia* are the reason for the conflicting topologies. Another possibility might be that the radiation of the lineages leading to *Buteo*, *Falco*, and *Ciconia* occurred within a relatively short time frame. The fossil evidence for the first appearance of the three lineages is rather vague: the earliest storklike fossil, *Palaeoepphippiorhynchus*, stems from the early Oligocene (37–26 MYA) of Fayum, Egypt (Olson 1985). The oldest accipitrid fossils are also from early Oligocene deposits in France. These are thought to be *Buteo*-like (Newton and Olsen 1990), but according to Feduccia (1996), these remains are in need of revision. Falconids have been reported from 55 MYA (del Hoyo, Elliott, and Sargatal 1994) and 48 MYA (Peters 1991), respectively, but from fragmentary fossils. The first well-documented falconids, however, were described from the Eo-Oligocene in France and England (Peters 1991). Therefore, from the fossil record, it is not possible to decide which of the three lineages split first. With respect to anatomical traits, storks differ from accipitrids and falconids mainly in numerous characteristics of skeletal bone structure, skull formation, and the arrangement of some muscles (Rea 1983). One exceptional behavioral trait which storks share with cathartids, but not with accipitrids and falconids, is that storks keep cool by squirting their legs with urine. Although the distance-based algorithms favor a topology with *Buteo* and *Ciconia* as sister taxa, a closer relationship between *Buteo* and *Falco* is suggested by the following arguments: In the CR, the most variable part of the mt genome, sequence similarity is stronger between *Buteo* and *Falco* (71.4%, 691 bp; uncorrected, gaps ≥5 excluded) than between either of them and *Ciconia* (*Buteo*/*Ciconia*: 66.8%; *Falco*/*Ciconia*: 69.0%, 691 bp; uncorrected, gaps ≥5 excluded). Furthermore, *Buteo* and *Falco* share, in contrast to *Ciconia*, the same derived rearrangement of the CR. Thus, the topology of the ciconiiform clade in the MP dendrogram is not only in accordance with the classical taxonomic view, but is also corroborated by structural features of the mt genome.

Usefulness of Markers

Investigations of mt genes of various vertebrates by Russo, Takezaki, and Nei (1996) indicated that amino acid sequences were more informative than nucleotide sequences for reconstructing reliable trees. Our results do not conform with this, since resolution, as well as bootstrap support, decreases when amino acid sequences are used. Moreover, *nd5*, *cyt b*, and *nd4* do not appear to be the most appropriate genes in our analyses. Nevertheless, the two studies differ with respect to taxonomic level: Russo, Takezaki, and Nei (1996) analyzed vertebrate phylogeny, whereas we focused on avian evolution only. Moreover, we also included RNA sequences for our comparisons.

With respect to usefulness of marker genes, the results of our MP analyses of nucleotide sequences can be interpreted as follows: In general, dendrograms derived from single mt genes have low resolution and bootstrap support for expected species pairs. Whereas most genes support the species pairs *Corvus*/*Vidua* and *Rhea*/*Struthio*, the clades *Gallus*/*Aythya* and *Buteo*/*Falco* are found in the *16S* and tRNA trees only. Therefore, none of the protein-coding genes can be recommended as reliable markers for phylogenetic studies at this taxonomic level. Among the concatenated protein-coding sequenc-

es, *cox(1–3)* (although only half as long) appears better than *nd(1–6).* Concatenated rRNAs plus tRNAs (4,350 bp) resolve all four expected species pairs and are thus as good as the complete coding sequence (15,742 bp). In a dendrogram calculated from the complete coding sequence except *nd5* and *12S,* the two genes supporting the *Buteo*/*Ciconia* clustering, the bootstrap value of the *Buteo*/*Falco* clade rose to 92, while the rest of the tree remained unchanged. A tree based on only *16S* and tRNAs yielded high bootstrap support in the *Buteo*/*Falco*/*Ciconia* cluster but lower values for the other nodes. To summarize, the concatenated RNA genes (rRNAs, tRNAs) appear to be the favorable combination, although bootstrap support is lower than that for the total sequence.

Basal Relationships and Dating of Splits

The phylogeny presented in this study, which is based on MP analyses of mt genomes, is in accordance with several other studies based on whole mt genomes (Härlid, Janke, and Arnason 1997, 1998; Härlid and Arnason 1999; Mindell et al. 1999; Waddell et al. 1999) and nuclear sequences (Stapel et al. 1984; Caspers et al. 1997). It is also compatible with the results of Griffiths (1994), based on syringeal morphology. In comparison to the dendrogram of Mindell et al. (1999), three additional taxa were included in the present study: *Buteo, Ciconia,* and *Corvus.* Whereas in the maximum-likelihood tree of Mindell et al. (1999) *Falco* clusters with *Smithornis,* in our trees *Smithornis* never clusters with ciconiiform species. Instead, the Ciconiiformes appear as a stable monophyletic group that is the sister group of the clade Galloanserinae/Ratitae. The Passeriformes split at the base of the avian tree, although in some of the dendrograms they do not appear to be a monophyletic group. Instead, *Smithornis* (a representative of the suboscines, which are considered the most basal passeriform group) splits off as the most basal lineage of the dendrogram. Our mt-based phylogeny contradicts those derived from other studies of mt as well as nuclear genes, in which ratites appear at the base of the avian tree (Groth and Barrowclough 1999; van Tuinen, Sibley, and Hedges 2000). This incongruity is not necessarily due to marker selection (nuclear/mitochondrial). One reason for it might be that the studies differ in taxon composition. For example, in the dendrogram presented by Groth and Barrowclough (1999) based on the nuclear gene *RAG-1,* no ciconiiform and no suboscine species are represented. On the other hand, sequences of the mt genomes of birds included in the *RAG-1* study (cranes, loons, penguins, hemipodes, shorebirds, rollers) have not been published so far. Another reason for the incongruity might be the different lengths of the marker sequences (e.g., 3-kb *RAG* vs. 15,742-bp mt genes). In our MP dendrogram, the branches of the passeriform taxa appear shorter. This may be due either to the fact that the *Alligator* sequence is not a suitable outgroup to root the tree or to a slower substitution rate in the Passeriformes. Thus, we cannot rule out the possibility that the basal position of Passeriformes might be caused by long-branch attraction of the ciconiiform and galloanserine clusters. Altogether, a reasonable comparison between nuclear and mt-derived phylogenies will be possible only when (1) representatives of more avian groups (with both mt and nuclear genes) have been analyzed and (2) more nuclear sequence data (different genes) are available.

Unfortunately, the fossil record allows neither corroboration nor falsification of our data concerning the position of passeriformes. When the reports about the first appearances of avian groups are compared, no clear conclusions about the succession of splits or divergence times of the various lineages can be drawn. For example, according to Feduccia (1996, p. 166), the oldest fossils of Ratitae are from the Paleocene (65–53 MYA), and according to Houde and Haubold (1987), the oldest ostrich fossils are from the early Eocene (53–37 MYA), the epoch from which the oldest putative passeriform fossils (Boles 1995) as well as falconids (see above) are also dated. As the oldest neognathous fossils are from the Mesozoic (Olson 1992), there is at least no support from the fossil record for the hypothesis that the Paleognathae represent the most basal lineage.

Faced with the incomplete fossil record for birds of prey, the dating of divergences has to rely on a molecular approach. The following molecular datings are available. For the *Rhea*/*Gallus* split, Härlid, Janke, and Arnason (1997) calculated 80–90 MYA, and Waddel et al. (1999) calculated 92 MYA. For the divergence of *Aythya*/*Gallus,* 68 MYA has been estimated (Waddel et al.1999). According to our HKY85 distances (22.3% for *Aythya*/*Gallus,* 23.7% for Galloanserinae/Struthioniformes), these two splits should be closer. From the two different reference points, the divergence of *Buteo*/*Falco* (21.3%) can be estimated to have occurred in the late Cretaceous, either at 72–83 MYA or at 65 MYA.

Supplementary Material

The complete sequence of the mt genome of *B. buteo* is registered under GenBank accession number AF380305. For the analysis of the ΨCR, the following partial sequences have been determined: ΨCR of *H. albicilla,* AY034150; repeat units of ΨCR of *A. chrysaetos,* AY034151. Alignments of avian mt sequences can be viewed under ''Sequences'' at our web site at http://www.nhm-wien.ac.at/NHM/1Zoo/first_zoological_department/web/chemsyst/cuhp_24e.html.

Acknowledgments

We are grateful to Werner Mayer for critical discussions and comments on the manuscript and to two anonymous referees for numerous valuable suggestions. This work was supported by the Austrian Science Foundation (FWF) (project number P14069-BIO).

LITERATURE CITED

AVISE, J. C., W. S. NELSON, and C. G. SIBLEY. 1994. DNA sequence support for a close phylogenetic relationship between some storks and New World vultures. Proc. Natl. Acad. Sci. USA **91**:5173–5177.

BAKER, A. J., and H. D. MARSHALL. 1997. Mitochondrial control region sequences as tools for understanding evolution. Pp. 51–82 *in* D. P. MINDELL, ed. Avian molecular evolution and systematics. Academic Press, London.

BENSCH, S., and A. HÄRLID. 2000. Mitochondrial genomic rearrangements in songbirds. Mol. Biol. Evol. **17**:107–113.

BIJLSMA, R. G. 1997. *Buteo buteo*—buzzard. Pp. 160–161 *in* W. J. M. HAGEMEIJER and M. J. BLAIR, eds. The EBCC atlas of European breeding birds. T. & A. D. Poyser, London.

BOLES, W. E. 1995. The worlds oldest songbird. Nature **374**: 21–22.

BROWN, L. H., and D. AMADON. 1968. Eagles, hawks and falcons of the world. Wellfleet Press, Secaucus, N.J.

CASPERS, G.-J., D. U. DE WEERD, J. WATTEL, and W. W. DE JONG. 1997. α-crystallin sequences support a galliform/anseriform clade. Mol. Phylogenet. Evol. **7**:185–188.

CLAYTON, D. A. 1991. Replication and transcription of vertebrate mitochondrial DNA. Annu. Rev. Cell Biol. **7**:453–478.

CRACRAFT, J. 1981. Toward a phylogenetic classification of the recent birds of the world (class Aves). Auk **98**:681–714.

DEL HOYO, J., A. ELLIOT, and J. SARGATAL. 1994. Handbook of the birds of the world. Vol. 2. Lynx Edicions, Barcelona.

DESJARDINS, P., and R. MORAIS. 1990. Sequence and gene organization of the chicken mitochondrial genome. A novel gene order in higher vertebrates. J. Mol. Biol. **212**:599–634.

DUFRESNE, C., F. MIGNOTTE, and M. GUÉRIDE. 1996. The presence of tandem repeats and the initiation of replication in rabbit mitochondrial DNA. Eur. J. Biochem. **235**:593–600.

FEDUCCIA, A. 1996. The origin and evolution of birds. Yale University Press, New Haven, Conn.

GRIFFITHS, C. S. 1994. Monophyly of the Falconiformes based on syringeal morphology. Auk **111**:787–805.

GROTH, J. G., and G. F. BARROWCLOUGH. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol. Phylogenet. Evol. **12**:115–123.

HARING, E., M. J. RIESING, W. PINSKER, and A. GAMAUF. 1999. Evolution of a pseudo-control region in the mitochondrial genome of Palearctic buzzards (genus *Buteo*). J. Zool. Syst. Evol. Res. **37**:185–194.

HÄRLID, A., and U. ARNASON. 1999. Analyses of mitochondrial DNA nest ratite birds within the Neognathae supporting a neotenous origin of ratite morphological characters. Proc. R. Soc. Lond. B Biol. Sci. **266**:305–309.

HÄRLID, A., A. JANKE, and U. ARNASON. 1997. The mtDNA sequence of the ostrich and the divergence between paleognathous and neognathous birds. Mol. Biol. Evol. **14**:754–761.

———. 1998. The complete mitochondrial genome of *Rhea americana* and early avian divergences. J. Mol. Evol. **46**: 669–679.

HASEGAWA, M., H. KISHINO, and T.-A. YANO. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. **22**:160–174.

HOUDE, P., and H. HAUBOLD. 1987. *Palaeotis weigelti* restudied: a small Eocene ostrich (Aves: Struthioniformes). Paleovertebrata **17**:27–42.

JANKE, A., and U. ARNASON. 1997. The complete mitochondrial genome of *Alligator mississippiensis* and the separation between recent archosauria (birds and crocodiles). Mol. Biol. Evol. **14**:1266–1272.

JOLLIE, M. T. 1976. A contribution to the morphology and phylogeny of the Falconiformes, part 1. Evol. Theory **1**:285–298.

———. 1977. A contribution to the morphology and phylogeny of the Falconiformes, part 2–3. Evol. Theory **2**:15–300.

KIMURA, M. 1980. A simple method for estimating evolutionary rates 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 phylogenetic position of Fea's viper. Syst. Biol. **42**:18–31.

KÖNIG, C. 1982. Zur systematischen Stellung der Neuweltgeier (Cathartidae). J. Ornithol. **123**:259–267.

KUMAR, S. 1996. PHYLTEST: a program for testing phylogenetic hypotheses. Version 2.0. Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park.

L'ABBÉ, D. L., J.-F. DUHAIME, B. F. LANG, and R. MORAIS. 1991. The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. J. Biol. Chem. **266**:10844–10850.

MASUDA, R., M. NORO, N. KUROSE, C. NISHIDA-UMEHARA, H. TAKECHI, T. YAMAZAKI, M. KOSUGE, and M. C. YOSHIDA. 1998. Genetic characteristics of endangered Japanese golden eagles (*Aquila chrysaetos japonica*) based on mitochondrial DNA D-loop sequences and karyotypes. Zoo Biol. **17**: 111–121.

MINDELL, D. P., M. D. SORENSON, and D. E. DIMCHEFF. 1998. Multiple independent origins of mitochondrial gene order in birds. Proc. Natl. Acad. Sci. USA **95**:10693–10697.

MINDELL, D. P., M. D. SORENSON, D. E. DIMCHEFF, M. HASEGAWA, J. C. AST, and T. YURI. 1999. Interordinal relationships of birds and other reptiles based on whole mitochondrial genomes. Syst. Biol. **48**:138–152.

NEWTON, I., and P. OLSEN. 1990. Birds of prey. Merehurst Press, London.

OJALA, D., J. MONTOYA, and G. ATTARDI. 1981. tRNA punctuation model of RNA processing in human mitochondria. Nature **290**:470–474.

OLSEN, S. L. 1992. *Neogaeornis wetzeli* Lambrecht, a cretaceous loon from Chile (Aves: Gaviidae). J. Vertebr. Paleontol. **12**:122–124.

OLSEN, P. D. 1995. Australian birds of prey. Helm, Melbourne.

OLSON, S. L. 1985. 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.

PETERS, D. S. 1991. Zoogeographical relationships of the Eocene avifauna from Messel (Germany). Acta XXth Intl. Ornithol. Congr. **1**:572–577.

QUINN, T. W. 1997. Molecular evolution of the mitochondrial genome. Pp. 4–28 *in* D. P. MINDELL, ed. Avian molecular evolution and systematics. Academic Press, London.

QUINN, T. W., and A. C. WILSON. 1993. Sequence evolution in and around the mitochondrial control region in birds. J. Mol. Evol. **37**:417–425.

RANDI, E., and V. LUCCHINI. 1998. Organization and evolution of the mitochondrial DNA control region in the avian genus *Alectoris.* J. Mol. Evol. **47**:449–462.

REA, A. M. 1983. Cathartid affinities: a brief overview. Pp. 26–54 *in* S. R. WILBUR and J. A. JACKSON, eds. Vulture biology and management. University of California Press, Berkeley.

RUSSO, C. A. M., N. TAKEZAKI, and M. NEI. 1996. Efficiencies of different genes and different tree-building methods in recovering a known vertebrate phylogeny. Mol. Biol. Evol. **13**:525–536.

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

SBISÀ E., F. TANZARIELLO, A. REYES, G. PESOLE, and C. SACCONE. 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequenc-

es and their functional and evolutionary implications. Gene **205**:125–140.

SEIBOLD, I., and A. J. HELBIG. 1995. Evolutionary history of New and Old World vultures inferred from nucleotide sequences of the mitochondrial cytochrome *b* gene. Philos. Trans. R. Soc. Lond. B Biol. Sci. **350**:163–178.

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

SORENSON, M. D., J. C. AST, D. E. DIMCHEFF, T. YURI, and D. P. MINDELL. 1999. Primers for PCR-based approach to mitochondrial genome sequencing in birds and vertebrates. Mol. Phylogenet. Evol. **12**:105–114.

STAPEL, O., J. A. M. LEUNISSEN, M. VERSTEEG, J. WATTEL, and W. W. DE JONG. 1984. Ratites as oldest offshoot of avian stem-evidence from α-crystallin A sequences. Nature **311**:257–259.

STORER, R. W. 1971. Classification of birds. Pp. 1–19 *in* D. S. FARNER and J. KING, eds. Avian biology. Academic Press, New York.

STRESEMANN, E., and D. AMADON. 1979. Falconiformes. Pp. 271–424 *in* E. MAYR and G. W. COTTRELL, eds. Checklist of the birds of the world. Vol. 1. Harvard University Press, Cambridge, Mass.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b3a. Sinauer, Sunderland, Mass.

TAKEZAKI, N., A. RZHETSKY, and M. NEI. 1995. Phylogenetic test of the molecular clock and linearized trees. Mol. Biol. Evol. **12**:823–833.

THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, and D. G. HIGGINS. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**:4876–4882.

VAN TUINEN, M., C. G. SIBLEY, and S. B. HEDGES. 2000. The early history of modern birds inferred from DNA sequences of nuclear and mitochondrial ribosomal genes. Mol. Biol. Evol. **17**:451–457.

WADDELL, P. J., Y. CAO, M. HASEGAWA, and D. P. MINDELL. 1999. Assessing the cretaceous superordinal divergence times within birds and placental mammals by using whole mitochondrial protein sequences and an extended statistical framework. Syst. Biol. **48**:119–137.

WALBERG, M. W., and D. A. CLAYTON. 1981. Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. Nucleic Acids Res. **9**:5411–5421.

WINK, M., I. SEIBOLD, F. LOTFIKHAH, and W. BEDNAREK. 1998. Molecular systematics of holarctic raptors (order Falconiformes). Pp. 29–47 *in* R. D CHANCELLOR, B.-U. MEYBURG, and J. J. FERRERO, eds. Holarctic birds of prey. IGRAMEX, Calamonte, Spain.

ELIZABETH KELLOGG, reviewing editor

Accepted June 12, 2001