


Received 7 February 2000, accepted 20 January 2001. Associate Editor: J. L. Dickinson


Phylogenetic Utility of Avian Ovomucoid Intron G: A Comparison of Nuclear and Mitochondrial Phylogenies in Galliformes

MARGARET H. ARMSTRONG,¹ EDWARD L. BRAUN,¹,² AND REBECCA T. KIMBALL¹,³,4

¹Department of Biology, University of New Mexico, Albuquerque, New Mexico 87131, USA; ²Department of Plant Biology, Ohio State University, 1735 Neil Avenue, Columbus, Ohio 43210, USA; and ³Department of Evolution, Ecology, and Organismal Biology, Ohio State University, 1735 Neil Avenue, Columbus, Ohio 43210, USA.

ABSTRACT.—A novel nuclear marker, the avian ovomucoid intron G (OVOG) was sequenced from 19 galliform taxa. Results of the phylogenetic analyses using OVOG were compared to those obtained using the mitochondrial cytochrome b (cytb) gene to determine the phylogenetic utility of OVOG. OVOG appeared to have strong phylogenetic signal for reconstructing relationships among genera and families, and the only difference between OVOG and cytb was in the placement of the New World quail (Odontophoridae). Genetic distances estimated using OVOG are approximately half of those estimated using cytb, although that relationship was not linear. OVOG exhibited patterns of nucleotide substitution very different from cytb, with OVOG having little base compositional bias, a relatively low transition–transversion ratio, and little among-site rate heterogeneity.

Mitochondrial DNA (mtDNA) sequences are commonly used to estimate vertebrate phylogenies. MtDNA markers evolve rapidly, making investigation among closely related species possible, yet also contain enough slowly evolving sites to resolve deeper relationships. Although mtDNA phylogenies are likely to be correct in many cases (Moore 1995), use of mtDNA sequences can be problematic. MtDNA rarely undergoes recombination (Wolstenholme 1992), so problems due to lineage sorting or introgression cannot be detected. Nuclear pseudogenes of mtDNA sequences can also confound phylogenetic estimation (Sorenson and Quinn 1998). Therefore, it is useful to compare mtDNA phylogenies with nuclear gene phylogenies to control for such problems.

¹ Address correspondence to this author. E-mail: kimball.23@osu.edu
TABLE 1. Species examined and accession number of sequence data. Accession numbers available from GenBank.

<table>
<thead>
<tr>
<th>Group</th>
<th>Species</th>
<th>Common name</th>
<th>cytb</th>
<th>OVOG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cracids</td>
<td>Crax pauxi</td>
<td>Helmeted Currassow</td>
<td>AF068190</td>
<td>AF170973</td>
</tr>
<tr>
<td></td>
<td>Ortalis vetula</td>
<td>Plain Chachalaca</td>
<td>L08384</td>
<td>AF170974</td>
</tr>
<tr>
<td>New World quail</td>
<td>Cyrtonyx montezumae</td>
<td>Montezuma Quail</td>
<td>AF068192</td>
<td>AF170976</td>
</tr>
<tr>
<td></td>
<td>Oreortyx pictus</td>
<td>Mountain Quail</td>
<td>AF252860</td>
<td>AF170977</td>
</tr>
<tr>
<td>Guineafowl</td>
<td>Numida meleagris</td>
<td>Helmeted Guineafowl</td>
<td>L08383</td>
<td>AF170975</td>
</tr>
<tr>
<td>Turkeys</td>
<td>Meleagris gallopavo</td>
<td>Wild Turkey</td>
<td>L08381</td>
<td>AF170984</td>
</tr>
<tr>
<td>Grouse</td>
<td>Falcipennis canadensis</td>
<td>Spruce Grouse</td>
<td>AF170992</td>
<td>AF170986</td>
</tr>
<tr>
<td></td>
<td>Tympanuchus phasianellus</td>
<td>Sharp-tailed Grouse</td>
<td>AF068191</td>
<td>AF170985</td>
</tr>
<tr>
<td></td>
<td>Alectoris chukar</td>
<td>Chukar</td>
<td>L08378</td>
<td>AF170987</td>
</tr>
<tr>
<td>Partridges</td>
<td>Alectoris rufa</td>
<td>Red-legged Partridge</td>
<td>Z48775</td>
<td>AF170988</td>
</tr>
<tr>
<td></td>
<td>Bambusiscus tharacica</td>
<td>Chinese Bamboo Partridge</td>
<td>AF028790</td>
<td>AF170978</td>
</tr>
<tr>
<td>Pheasants</td>
<td>Afropavo congensis</td>
<td>Congo Peafowl</td>
<td>AF13760</td>
<td>AF170991</td>
</tr>
<tr>
<td></td>
<td>Catreus wallichi</td>
<td>Cheer Pheasant</td>
<td>AF028792</td>
<td>AF170980</td>
</tr>
<tr>
<td></td>
<td>Crossoptilon crossoptilon</td>
<td>White-eared Pheasant</td>
<td>AF028794</td>
<td>AF170981</td>
</tr>
<tr>
<td></td>
<td>Gallus gallus</td>
<td>Red Junglefowl</td>
<td>AF028795</td>
<td>AF170979</td>
</tr>
<tr>
<td></td>
<td>Pavo cristatus</td>
<td>Common Peafowl</td>
<td>L08379</td>
<td>AF170990</td>
</tr>
<tr>
<td></td>
<td>Pavo muticus</td>
<td>Green Peafowl</td>
<td>AF013763</td>
<td>AF170989</td>
</tr>
<tr>
<td></td>
<td>Pucrasia macrolopha</td>
<td>Koklass Pheasant</td>
<td>AF028800</td>
<td>AF170983</td>
</tr>
</tbody>
</table>

* Samples were obtained from aviaries and zoos that maintain captive-reared stock of legal and documented origin, except for Falcipennis (collected by J. Hugelin in Alaska) and Tympanuchus (collected by H.B. Tordoff in North Dakota).

Nuclear intron sequences appear to evolve more rapidly than nuclear exons due to the relative lack of constraint upon their sequence (Li 1997). That rapid evolution makes nuclear introns attractive candidates for comparison with mtDNA phylogenies. However, relatively few phylogenies using nuclear intron data have been published in birds (but see Prychitko and Moore 1997, 2000; Johnson and Clayton 2000; all of which use [3-fibrinogen intron 7).

Ovomucoid is a single-copy gene that acts as a storage protein and protease inhibitor (Laskowski et al. 1987a), and it was one of the earliest genes for which intron and exon structure was determined (Stein et al. 1980). A large set of partial amino acid sequences exists for avian ovomucoids (Laskowski et al. 1987b, 1990; Apostle et al. 1993), providing data that can be used to develop primers for phylogenetic studies of birds. Here we compare the molecular evolution of intron G in domain III of ovomucoid (OVOG) with that of the mitochondrial cytochrome-b (cytb) gene in Galliformes. These comparisons indicate that OVOG is useful for avian phylogenetics.

**Methods.**—OVOG primers were designed by locating conserved amino acid sequences surrounding the intron (using Laskowski et al. 1987b, 1990; Apostle et al. 1993) and back translating from those sequences. Nucleotide data from the chicken ovomucoid (Stein et al. 1980) was used to establish the nucleotides at 3rd codon positions; one site near the 3' end of each primer was left degenerate to ensure template matching (OVOG Forward: 5'-CAAGACTACGGCAACAARTG-3'; OVOG Reverse: 5'-GGCCTAAGTGAGGTCCTCCTT-3'). DNA was extracted as described elsewhere (Kimball et al. 1997), and most samples were identical to those used in our previous research (Kimball et al. 1997, 1999). OVOG was amplified by PCR using standard protocols, and the products cleaned using QiAquick™ PCR Purification columns (Qiagen, Inc., Valencia, California). Sequencing reactions (one-quarter or one-half volume) were performed using Thermosequenase dye, terminator (Amersham Life Science, Piscataway, New Jersey), or BigDye™ Terminator (Perkin-Elmer Applied Biosystems, Forest City, California) chemistry and the primers used for PCR amplification. Sequence reactions were analyzed using an ABI 310 or 377 automated DNA sequencer. Sequence chromatographs were edited manually and assembled into double-stranded contigs. A few ambiguous sites, probably reflecting heterozygosity, were evident, and those sites were coded using the proper IUPAC symbols. Cytb sequences were published or sequenced for this study as described previously (Kimball et al. 1999). Species and sequence accession numbers are in Table 1.

OVOG sequences were aligned using ClustalW (Thompson et al. 1994), and optimized by eye. The aligned data have been submitted to GenBank. Gaps were treated as missing data, recoded as described by Bruns et al. (1992), or alignment positions with gaps were deleted. Unless otherwise stated, all results reported correspond to the gap positions deleted alignment. On the basis of Sibley and Ahlquist (1990) and Kornegay et al. (1993), phylogenetic trees...
were rooted to the Cracidae (Ortalis vetula and Crax pauxi).

Analyses were performed using PAUP* 4.0b4a (Swofford 1999), unless otherwise noted. The most parsimonious trees were identified using equally weighted parsimony with 100 random addition sequence replicates. The incongruence length difference test was conducted using 1,000 replicates and heuristic searches with 10 random addition sequence replicates.

Neighbor joining of maximum likelihood distances were also conducted. Parameters (the transition-transversion \([s/v]\) ratio or substitution model rate matrix, the proportion of invariant sites \([\text{inv}]\), and the shape parameter \([\alpha]\) of a \(\Gamma\)-distribution) were estimated using MODELTEST 3.04 (Posada and Crandall 1998). The best fitting models, based upon the hierarchical likelihood ratio test implemented in MODELTEST, were used. Synonymous and nonsynonymous \(p\)-distances were calculated using MEGA 1.02 (Kumar et al. 1993).

Reliability of specific groupings in parsimony and distance analyses was assessed using 1,000 bootstrap replicates. For parsimony analyses, 10 random addition sequence replicates were used per bootstrap replicate. Clades with <50% bootstrap support were collapsed.

Results and discussion.—The OVOG alignment was quite variable, containing 505 sites, of which 182 (36%) were variable and 112 (22%) were parsimony informative. After removal of gaps, 388 sites remained, with 151 (39%) variable and 97 (25%) parsimony informative. A total of 126 equally parsimonious trees requiring 209 steps were found.

For OVOG, the best fitting model was HKY+\(\Gamma\), with a slight transition bias \((s/v \text{ ratio } = 2.33)\) and relatively weak among-site rate heterogeneity \((\alpha = 1.18)\). The biological significance of the observed among-site rate variation is unclear, because few sites in intron sequences are expected to be constrained. Although absolute conservation of the 5' and 3' splice sites for OVOG (see Li 1997 for description) was observed, exclusion of the conserved nucleotides near the 5' and 3' splice sites did not alter the model selected (it remained HKY+\(\Gamma\)) and had little affect upon the shape parameter \((\alpha = 1.23)\). Thus, OVOG exhibits currently unappreciated sources of among-site rate variation.

The cytb alignment contained 1,143 sites with no insertions or deletions (see Kimball et al. 1999). As reported previously (Kornegay et al. 1993, Randi 1996, Kimball et al. 1997, 1999), cytb is highly variable, with 473 variable (41%) and 375 parsimony informative sites (33%). A single most parsimonious tree requiring 1,536 steps was identified.
Fig. 2. Bootstrap consensus trees of OVOG (left) and cyt(b) (right). Values represent percentage of 1,000 bootstrap replicates for equally-weighted parsimony (above) and neighbor-joining of maximum likelihood distances (below). Branches were collapsed when both bootstrap values were <50%.

In contrast to OVOG, cyt(b) is characterized by a more complex model of evolution (also see Kimball et al. 1999), exhibiting the best fit to the GTR+Γ+inv model. More than half of cyt(b) sites appear to be invariant (inv = 0.51), and the remaining sites exhibit moderate among-site rate variation (α = 0.73). As expected, there is a bias toward transition substitutions, with the rate parameter for the most slowly accumulating (A–G) transitions being 3.8x greater than that of the most rapidly accumulating (A–T) transversions. There was also more than two-fold variation in rates among different classes of transitions and transversions (data not shown).

The base composition of cyt(b) is highly biased (Table 2), which can lower the amount of detectable variation in an alignment (e.g. Meyer 1994). More interestingly, the base composition of cyt(b) variable sites differed from that of all cyt(b) sites (Table 2). In contrast, the base composition of the OVOG alignment, after removing sites aligned with gap positions, was similar to an equal base composition. Likewise, there was little difference between the variable sites and all sites in OVOG (Table 2).

Genetic distances in OVOG comparisons were about 50% that of cyt(b), but that relationship is not linear (Fig. 1a). Between closely related taxa, the degree of sequence divergence in cyt(b) increases more rapidly than OVOG as cyt(b) accumulates synonymous mutations; between more distant taxa, sequence divergence continues to increase in OVOG,
but cytb sequence divergence slows as synonymous mutations appear to reach saturation and only non-synonymous substitutions continue to accumulate (Fig. 1b). Thus, the evolution of OVOG is characterized by accumulation of mutations in a manner substantially more linear relative to time than cytb. A comparison of the β-fibrinogen intron 7 (Prychitko and Moore 1997, 2000; Johnson and Clayton 2000) suggests that, relative to cytb, rate of sequence divergence in OVOG is at least twice that of the β-fibrinogen intron, assuming similar rates of cytb evolution in galliforms, woodpeckers (Picidae), and doves (Columbiformes).

Trees constructed using OVOG were compared to those constructed using cytb (Fig. 2). OVOG and cytb represented congruent data partitions on the basis of the incongruence length difference test \((P = 0.945)\), suggesting both partitions support similar estimates of galliform phylogeny. However, the consistency indices \((CI)\) excluding uninformative characters were substantially higher for OVOG \((CI = 0.780)\) than for cytb \((CI = 0.408)\).

Gaps (insertions or deletions) in OVOG provided modest phylogenetic information, with most parsimony informative gaps uniting groups that were clearly monophyletic on the basis of other analyses. We found that gap treatment had little effect on topology or levels of bootstrap support (data not shown).

All nodes well supported in cytb analyses (those with >70% bootstrap; see Hillis and Bull 1993) were also well supported in OVOG analyses (Fig. 2). OVOG strongly supported a close relationship between Afropavo and Pavo (Kimball et al. 1997) and between Gallus and Bambusicola (Kimball et al. 1999). Relative to cytb, one additional well-supported node was evident in analyses using OVOG (Fig. 2). That node involved the New World quail (Cyrtonyx and Oreortyx), which have been difficult to place within the galliforms (Crowe 1988, Kornegay et al. 1993). Analyses of morphological data (reviewed by Crowe 1988) and OVOG (Fig. 2) place the New World quail in a derived position relative to the guineafowl, such as Numida. However, DNA–DNA hybridization place the New World quail in a basal position (Sibley and Ahlquist 1990), and the New World quail are currently placed in their own family (Odontophoridae) distinct from the other phasians (which include the guineafowl; American Ornithologists’ Union 1997).

Deep coalescence at the OVOG locus is unlikely to explain the differences between estimates of phylogeny using this marker and those obtained using cytb or DNA–DNA hybridization. Inspection of phylogenograms from Sibley and Ahlquist (1990), Kornegay et al. (1993), Kimball et al. (1999), and OVOG (data not shown) coupled with consideration of the galliform fossil record (reviewed by Kornegay et al. 1993, Kimball et al. 1997) suggest that the internodes between the relevant groups are too long to be attributed to intraspecific coalescence times. A complete understanding of the relationship between the New World quail and other galliforms may require data from additional nuclear loci.

Overall, OVOG appears to be a reliable marker for phylogenetic studies. It provided novel information that may change previous conclusions regarding placement of the New World quail, and further emphasizes the value of examining multiple markers, including both nuclear and mitochondrial, in avian phylogenetics.

Acknowledgments.—We thank L. Bean and J. Parkhani for laboratory assistance. Additional samples were generously provided by J. Hagelin and H. B. Tordoff. Much of this work was conducted in the National Science Foundation-supported (RIMI program grant HRD-9550649) Molecular Biology Facility at the University of New Mexico.

LITERATURE CITED


July 2001] Short Communications 803


Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: Molecular evolutionary genetics analysis, version 1.01. Pennsylvania State University, University Park.


Received 17 April 2000, accepted 23 January 2001. Associate Editor: F. Sheldon