Insertion Events of *CR1* Retrotransposable Elements Elucidate the Phylogenetic Branching Order in Galliform Birds

Vera B. Kaiser, Marcel van Tuinen,† and Hans Ellegren**

*Department of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, Norbyvägen, Uppsala, Sweden; and †Department of Ornithology and Mammalogy, California Academy of Sciences

Using standard phylogenetic methods, it can be hard to resolve the order in which speciation events took place when new lineages evolved in the distant past and within a short time frame. As an example, phylogenies of galliform birds (including well-known species such as chicken, turkey, and quail) usually show low bootstrap support values at short internal branches, reflecting the rapid diversification of these birds in the Eocene. However, given the key role of chicken and related poultry species in agricultural, evolutionary, general biological and disease studies, it is important to know their internal relationships. Recently, insertion patterns of transposable elements such as long and short interspersed nuclear element markers have proved powerful in revealing branching orders of difficult phylogenies.

Here we decipher the order of speciation events in a group of 27 galliform species based on insertion events of chicken repeat 1 (*CR1*) transposable elements. Forty-four *CR1* marker loci were identified from the draft sequence of the chicken genome, and from turkey BAC clone sequence, and the presence or absence of markers across species was investigated via electrophoretic size separation of amplification products and subsequent confirmation by DNA sequencing. Thirty markers proved possible to type with electrophoresis of which 20 were phylogenetically informative. The distribution of these repeat elements supported a single homoplasy-free cladogram, which confirmed that megapodes, cracids, New World quail, and guinea fowl form outgroups to Phasianidae and that quails, pheasants, and partridges are each polyphyletic groups. Importantly, we show that chicken is an outgroup to turkey and quail, an observation which does not have significant support from previous DNA sequence– and DNA-DNA hybridization–based trees and has important implications for evolutionary studies based on sequence or karyotype data from galliforms. We discuss the potential and limitations of using a genome-based retrotransposon approach in resolving problematic phylogenies among birds.

Introduction

The order Galliformes evolved approximately 90 MYA (van Tuinen and Dyke 2004) and includes some well-known birds, such as chicken, turkey, quail, grouse, partridge, peacock, and pheasant. The order is divided into 5 families: Megapodiidae (megapodes), Cracidae (cracids), Odontophoridae (New World quails), Numididae (guinea fowl), and Phasianidae (pheasants, Old World quail, turkey, chicken, and allies). Given the importance of chicken and other galliforms in agriculture, as models in studies of development (Brown et al. 2003), virology, oncogenesis, and immunology (Cooper et al. 1966; Stehelin et al. 1976), as well as in a variety of evolutionary studies including biogeography, sexual selection, and karyotype evolution (von Schantz et al. 1989; Pizzari et al. 2003; Shibusawa et al. 2004b), the phylogenetic relationships among galliforms have been thoroughly investigated. Moreover, a number of approaches have been used, in itself reflecting the technological developments of molecular biology from which systematics research has benefited: immunological reactions (Jolles 1976; Prager and Wilson 1976), protein sequencing (Henderson et al. 1981), allozymes (Gutierrez et al. 1983), restriction fragment length polymorphisms (Helm-Bychowski and Wilson 1986; Ellsworth et al. 1995), DNA-DNA hybridization (Sibley and Ahlquist 1990), and mitochondrial (Kimball et al. 1999; Dimcheff et al. 2002; Bush and Strobeck 2003) and, more recently, nuclear sequence data (Armstrong et al. 2001; Smith et al. 2005). Despite these extensive efforts, the branching order among the 5 families remains problematic, in particular, the relative position of guinea fowl and New World quails within Galliformes (Johnsgard 1986; Armstrong et al. 2001; Dyke et al. 2003; van Tuinen and Dyke 2004); the current understanding is that the Phasianidae, containing all the familiar species mentioned above, forms the most derived group that evolved about 40–50 MYA (van Tuinen and Dyke 2004; Pereira and Baker 2006). Within the Phasianidae, branching orders of the subfamilies Tetraoninae (grouse), Meleagridinae (turkeys), Perdicinae (partridges, quail), Phasianinae (pheasants, jungle fowl), and Pavoninae (peacock, peacock pheasants) are unresolved. Internal nodes of mitochondrial and nuclear DNA phylogenies are generally very short, suggesting a rapid diversification of phasianid subfamilies (Kimball et al. 1999; Dimcheff et al. 2002; Shibusawa et al. 2004a; Smith et al. 2005). In particular, bootstrap support values on branches separating the lineages of chicken, turkey, and quail are too low to yield confidence in the phylogenetic relationship among the 3 lineages.

Using standard phylogenetic methods that rely on models of sequence evolution, it can be hard to resolve tree topologies when new lineages evolved in the distant past and within a short time frame: back mutations, rate inconsistencies, and base composition biases are well-known factors that can contribute to such problems. However, insertion events of transposable elements have been recently proved as powerful tools in phylogenetics, not least in cases when standard tree reconstruction methods produced inconsistent results (e.g., Shedlock and Okada 2000; Okada et al. 2003). Examples of the use of transposable elements in phylogenetic reconstruction are available for several vertebrate groups, including salmons (Murata et al. 1993), whales (Shimamura et al. 1997), Lake Malawi cichlids (Takahashi et al. 2001; Terai et al. 2003), turtles (Sasaki et al. 2004), eels (Kajikawa et al. 2005), Afrotheria (Nishihara et al. 2005), and primates (Ray et al. 2005).

Key words: retrotransposon, cladogram, phylogenetics, birds, genomics.

E-mail: hans.ellegren@ebc.uu.se.

Mol. Biol. Evol. 24(1):338–347. 2007
doi:10.1093/molbev/msl164
Advance Access publication October 31, 2006

© The Author 2006. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved. For permissions, please e-mail: journals.permissions@oxfordjournals.org

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

However, the application of transposable elements in avian systematics has been limited until very recently (St John et al. 2005; Watanabe et al. 2006).

LINEs and SINEs (long and short interspersed nuclear elements, respectively) are thought to be exclusively transmitted vertically from parent to offspring and inserted "unidirectionally" (no character reversal) into the host genome, that is, once an insertion event has taken place, the retrotransposon remains in that locus (Nikaido et al. 1999; Okada 1991; but see van de Lagemaat et al. 2005), unless a chromosomal segment containing the repeat becomes deleted. As retrotransposition is thought to be more or less random with respect to the region of insertion, insertions at exactly the same location appear unlikely. Thus, as opposed to reversible changes in DNA sequence composition, insertion events of retrotransposons have been claimed to be homoplasy-free phylogenetic markers with minimal noise from reversal or parallel events: species, which share an insertion at a particular locus are grouped together on the tree and all species that lack the insertion are considered basal to this clade.

Chicken repeat 1 (*CR1*) elements are LINEs, which presumably arose prior to the common ancestor of birds and reptiles (Vandergon and Reitman 1994; Wicker et al. 2005). A complete *CR1* element comprises roughly 4.5 kb and contains 2 open reading frames, ORF1 and ORF2. ORF2 encodes a reverse transcriptase, whereas the functional role of ORF1 is less clear, although an esterase domain and a zinc-finger motif have been recognized (Kapitonov and Jurka 2003). Most of the estimated 200,000 *CR1* copies in the chicken genome are shorter than 400 bp (ICGSC 2004), presumably owing to severe 5′ truncations arising from premature reverse transcription termination during the integration process ("dead-on-arrival"; Haas et al. 1997). *CR1* elements are believed to propagate via the "Master gene model" in which a limited number of master elements produce new daughter elements that are incorporated into the genome intermittently (Vandergon and Reitman 1994; Shedlock and Okada 2000). Once a daughter element has been inserted, it probably becomes silent and evolves without functional constraint. At least 11 subfamilies of *CR1* elements have been identified (ICGSC 2004), each of which is derived from a different master element that was active at a certain time during avian evolution. It is not clear whether *CR1* elements are still actively spreading in the chicken genome; only 1 functional copy was identified in the draft genome sequence with the limited number of additional full-length copies containing frameshift or stop codon mutations. Clearly, the majority of *CR1* repeats are relics of ancient retrotransposition events.

In this study we address the branching order within Galliformes and clarify the position of chicken relative to other model galliform species such as turkey and quail. We use a genomic approach to retrotransposon-based phylogenetics by first surveying the draft chicken genome sequence for potentially informative *CR1* markers and then complement these with additional loci identified from turkey BAC clone sequences. By screening for the presence or absence of *CR1* markers across a panel of 20 phasianid and several nonphasianid species, we demonstrate monophyly of Phasianidae and that, within Phasianidae, turkeys and Old World quail are more derived than the chicken lineage. This result supports one phylogenetic interpretation of the cytochrome b gene (Kimball et al. 1999) but not alternative interpretations based on cytochrome b and other genes or on morphology.

Methods

Searching for Candidate Marker Loci

Candidate loci for phylogenetic inference were detected via 2 different computational approaches. First, the draft chicken genome sequence (www.ensembl.org/Gallus_gallus/index.html) and the program REPEATMASKER (www.repeatmasker.org) were used to search for intronic *CR1* elements with an estimated divergence from the subfamily consensus sequence (ICGSC 2004) approximately consistent with the minimum time when galliform diversification took place (about 40 MYA for crown Phasianidae; van Tuinen and Dyke 2004; see Axelsson et al. (2004) for galliform substitution rates). The rationale for using intron-specific *CR1* elements as markers was the expectation that cross-species amplification success would be increased when primers from flanking exons were used. To facilitate polymerase chain reaction (PCR) amplification with exonic primers, we only considered elements in introns of less than 1 kb.

Because derived branches of a tree cannot be resolved using insertion sites of a basal lineage, we also used alignments of chicken genome sequence and turkey BAC clone sequence available in GenBank to search for loci at which turkey carried a *CR1* element but chicken did not, that is, for insertions that occurred on the lineage leading to turkey after its split from the chicken lineage. The BLAT function at http://www.genome.ucsc.edu was used to confirm that these insertions were in fact *CR1* elements.

PCR primers were designed using Primer3 at http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi and are available as Supplementary Material online (Table S1). Turkey-derived primers were exclusively targeted against noncoding DNA. However, to increase the chance of successful cross-species amplification in the other galliform species, particularly conserved regions were identified from the turkey-chicken alignments.

Laboratory Work

Our set of galliform taxa included 20 phasianid species spanning all subfamilies, including Phasianinae (*Gallus gallus* [chicken], *Gallus varius* [green jungle fowl, CAS 85707], *Phasianus colchicus* [ring-necked pheasant], *Tragopan satyra* [satyr tragopan, CAS 92236], *Crossoptilon crossoptilon* [white-eared pheasant, CAS 86072]), Pavoninae (*Pavo cristatus* [Indian peafowl, CAS92906], *Polyplectron chalcurum* [bronze-tailed peacock pheasant, CAS 89862]), Perdicinae (*Coturnix coturnix* [common quail], *Coturnix japonica* [japanese quail], *Alectoris rufa* [red-legged partridge], *Bambusicola thoracica* [Chinese bamboo partridge, CAS 89821], *Rollulus roulroul* [crested wood partridge, CAS 89902], *Francolinus pondicerianus* [grey francolin, CAS 87894]), Tetraoninae (*Bonasa bonasia* [hazel grouse], *Centrocercus urophasianus* [sage grouse], *Tetrao tetrix* [black grouse], *Lagopus lagopus* [willow grouse], *Lagopus mutus* [rock ptarmigan]), and Meleagridinae (*Meleagris gallopavo* [wild turkey], *Meleagris ocellata* [ocellated turkey, CAS 85834]), as well as 1 megapode

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

species (*Alectura lathami* [brush turkey]), 1 cracid species (*Crax mitu* [razor-billed curassow, CAS 86070]), 3 numidid species (*Numida meleagris* [helmeted guinea fowl], *Acryllium vulturinum* [vulturine guinea fowl, CAS 85748], *Guttera pucherani* [crested guinea fowl, CAS 86157]), 2 odontophorid species (*Colinus virginianus* [northern bobwhite, MVZ 180367], *Callipepla californica* [California quail, CAS 90626]), and 1 anseriform species (*Anser anser* [domestic goose]). Ten microliters of PCRs contained 0.1 μl deoxynucleoside triphosphates (20 mM), 1 μl $MgCl_2$ (25 mM), 0.3 μl of each of the 2 primers (10 μM), 0.1 μl Ampli *Taq* Gold (Applied Biosystems, Foster City, CA) (5 U/μl), 1 μl 10× Gold Buffer, and 2.5 μl template DNA (10 ng/μl). The thermal cycling involved 35 repeats of denaturation at 95 °C for 30 s, annealing at 55 °C for 45 s, and primer extension at 72 °C for 1 min. A 2-μl aliquot of each PCR product was run in a 1.25% low–melting point agarose gel for genotyping.

Crucial marker loci were sequenced to confirm the presence or absence of a *CR1* insertion. For this purpose, the PCR products were cleaned by adding 1 μl ExoSAP-IT (Amersham Biosciences, Uppsala, Sweden) to every 2.5 μl of PCR product and by incubating the reaction for 15 min at 37 °C and for 15 min at 80 °C. Five microliters of cleaned PCR product was used as a template in DYEnamic cycle sequencing reactions containing 4 μl DYEnamic sequencing premix (Amersham Biosciences) and 1 μl sequencing primer (10 μM). Sequencing cycles (29) consisted of denaturation for 20 s at 95 °C, annealing for 15 s at 50 °C, and primer extension for 1 min at 60 °C. Excess dye terminators were removed from the sequencing reactions by gel filtration, and the products were run on a MegaBACE capillary sequencing instrument (Amersham Biosciences). DNA chromatograms were edited and base calls checked using SEQUENCHER 4.2.2 (Gene Codes, Ann Arbor, MI). All sequences have been deposited in GenBank under the accession numbers EF115225-EF115285.

Reconstruction of Species Cladogram

If the size difference between PCR fragments matched the expected insertion size based on sequence data for the marker in question, we inferred that a *CR1* element was inserted in those species that showed the longer fragment (see fig. 1 for schematic view of methods). A matrix containing all species and loci was constructed designating “1” and “0” as presence or absence of an insertion and “–” as an unsuccessful PCR amplification (Bashir et al. 2005). The relationship between *CR1* subfamily divergence and the position of *CR1* subfamily members in our cladogram was examined based on the average divergence of *CR1* copies to the subfamily-specific consensus sequence of a subfamily (ICGSC 2004). Turkey-derived *CR1* elements were aligned to all known *CR1* consensus sequences using ClustalW and the respective subfamily identified.

Results

Forty-four out of 48 candidate *CR1* loci could be successfully amplified in all or several of the galliform species. However, only 30 loci were included in subsequent analyses (table 1) as some loci showed length variation that could not be ascribed to the presence or absence of *CR1* elements

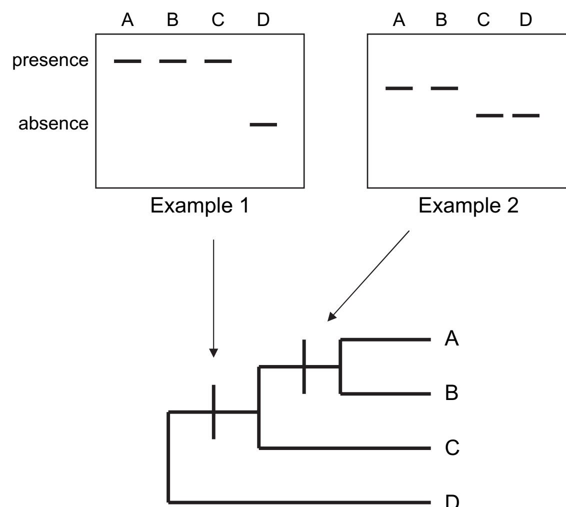

FIG. 1.—Schematic view of methods. (*a*) An intronic *CR1* element is amplified by PCR using primers located in flanking exons. (*b*) Electrophoretic separation of 2 marker loci (L1 and L2) amplified in species A–D reveals the presence (+) or absence (−) of an insertion. (*c*) A cladogram is reconstructed: species D lacks both insertions and is basal to species C, which carries an insertion only at locus 2. Species A and B show insertions at both loci and thus form the most derived group.

(i.e., likely harboring other types of insertions or deletions). Although spurious length variation lending a marker to be excluded from the phylogenetic analysis was typically in the size range of ±50–100 bp, 6 loci showed amplification products significantly shorter (500–600 bp) than predicted. Unexpectedly, all these cases involved a single lineage, the *Coturnix* quails. Sequencing revealed that 5 of these loci had a large deletion covering the whole *CR1* element in quails, making genotyping impossible; however, at one locus, part of the *CR1* insertion was still detectable. Although not within the aim of this study, these observations warrant further investigation as they suggested an unusually high rate of deletion mutations in the quail lineage.

Amplification success was similar for chicken-derived (14 out of 23 giving specific amplification) and turkey-derived (16 out of 25) markers. This may be seen as somewhat surprising given that chicken-derived markers had exon-specific primers, whereas turkey-derived markers had primers in anonymous noncoding DNA. However, because the latter were obtained from chicken-turkey alignments, we could target regions of noncoding DNA that appeared particularly conserved in galliform evolution. Notably, the frequency of amplification of the 30 loci that were included in the matrix decreased markedly toward the base of Galliformes (6/30 anseriforms; 9/30 megapodes; 15/30 cracids).

Twenty loci were phylogenetically informative (markers present in some species, absent in others, see fig. 2), and a further 6 loci showed lineage-specific insertion in turkeys (turkey-derived markers) (table 2). The remaining markers (4 loci) gave similarly sized products in all species in which they amplified, implying either that the *CR1* insertion occurred before the divergence of Phasianidae and the outgroup species or simply that the time of the insertion event within Galliformes was not possible to pinpoint because species from the clades that diverged before insertion events would have failed to amplify.

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

Table 1
Information on Location (Start and End Position in the Chicken Genome Assembly), ID of the Associated Gene, and Length of the Repeat Element for *CR1* Markers Used in This Study

Marker	Chromosome	Start	End	Ensembl ID	Intron	*CR1* length(bp)
4755	11	1954532	1954901	ENSGALT00000004755	13	369
24741	1	79395407	79395738	ENSGALT00000024741	1	331
19948	7	31194158	31194283	ENSGALT00000019948	17	125
14195	3	5417327	5417457	ENSGALT00000014195	8	130
25774	4	88642677	88642800	ENSGALT00000025774	11	123
27300	1	138581968	138582102	ENSGALT00000027300	10	134
6825	21	5137725	5137828	ENSGALT00000006825	11	103
8684	8	12433202	12433417	ENSGALT00000008684	17	215
26384	2	145623928	145623991	ENSGALT00000026384	14	63
5407	23	4638714	4638903	ENSGALT00000005407	2	189
14297	14	16871280	16871326	ENSGALT00000014297	15	46
20769	1	51410989	51411113	ENSGALT00000020769	2	124
27549	1	164661837	164661990	ENSGALT00000027549	4	153
19495	2	43445791	43445866	ENSGALT00000019495	4	75
455-141	2	21783918				187
458-170	1	17785762				142
458-569	2	21999046				381
464-577	4	3731432				326
465-524	4	3746849				316
515-493	2	16177644				150
464-741	4	3746850				317
515-280	2	16400239				317
515-302	2	16420771				104
574-779	1	17883656				860
580-666	3	39031248				590
580-812	3	39045157				80
580-129	3	39090772				510
588-636	17	2315309				158
588-887	17	2317603				202
588-147	17	2449820				120

NOTE.—For markers derived from turkey, chromosome number is taken as given by the orthologous sequence in chicken. For these markers, the insertion site is given as the position −1 of the site in the chicken genome. No end position is given for turkey-derived markers.

Overall, the information provided by the total set of markers was consistent with a single cladogram shown in fig. 3, that is, no markers revealed contradictory results. Four markers were shared among all Phasianidae but absent in other species, confirming that megapodes, cracids, guinea fowl, and New World quail form outgroup lineages to Phasianidae. Another 2 markers showed that cracids and megapodes form more primitive lineages than the remaining galliform families. With the exception of 1 enigmatic species, the crested wood partridge, the chicken formed the basal offshoot among phasianid species together with green jungle fowl, francolins, and the Chinese bamboo partridge, an observation supported by 2 independent *CR1* insertion events (2 elements from different subfamilies inserted in different chromosomes). Seven markers showed that the lineage leading to *A. rufa* (red-legged partridge) and *Coturnix* (Old World quail) is basal to grouses, pheasant, white-eared pheasant, tragopan, and turkeys, whereas one marker showed that the core phasianine lineage (excluding *Gallus*) is basal to grouses and turkeys. In addition, one marker highlighted that, compared with the relative phylogenetic position of chicken, the pavonine (peacock) species are more derived within Phasianidae. Furthermore, 2 markers showed the close relationship of *F. pondicerianus* and *B. thoracica* to *Gallus,* and one marker showed that francolins are primitive within this clade.

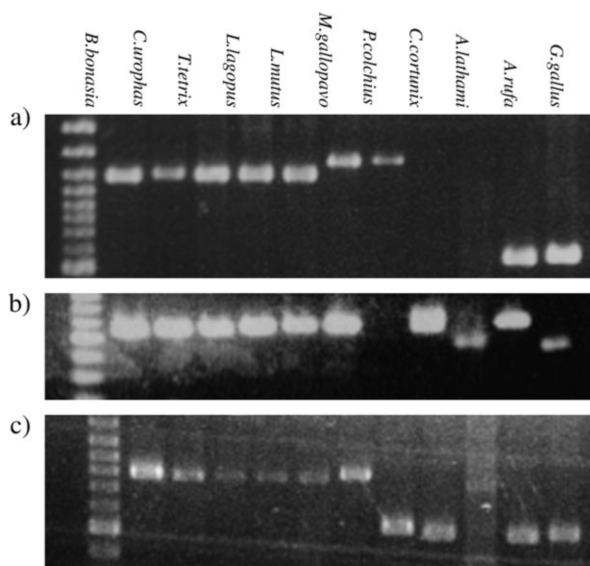

FIG. 2.—PCR amplification of 3 *CR1* marker loci in 11 galliform species. Size differences between fragments of orthologous loci largely corresponded to the lengths of *CR1* elements: (*a*) 860 bp, (*b*) 142 bp, and (*c*) 317 bp.

DNA sequencing was subsequently used to confirm the results from genotyping of phylogenetically informative loci. We obtained sequence data for 18 markers, which in all cases confirmed the lack or presence of a *CR1* element in representative taxa as interpreted by agarose electrophoresis (see supplementary fig. S1 for alignments, Supplementary

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

Table 2
Matrix of Genotype Data Used for Cladogram Construction

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Anser	0	0	0	–	×	–	–	–	–	–	–	–	–	–	0	0	–	–	–	–	–	–	–	–	–	–	–	–	–	–
Alectura	–	0	0	0	–	–	–	–	0	0	–	–	–	0	–	0	–	–	–	–	–	–	–	–	–	0	0	–	–	–
Crax	0	0	0	0	×	0	–	–	–	0	×	–	–	0	0	0	–	–	0	0	–	0	0	–	–	–	–	–	–	–
Colinus	0	0	0	0	×	×	×	×	×	0	×	0	0	0	0	0	0	–	–	–	–	0	0	–	–	0	0	–	0	–
Callipepla	0	0	0	0	×	×	×	×	–	–	–	0	0	0	0	–	–	–	–	–	–	0	0	–	–	0	–	–	–	–
Numida	0	0	0	0	×	×	–	–	×	0	–	0	0	0	0	–	–	–	–	0	–	0	–	–	0	0	0	–	0	–
Acryllium	0	0	0	0	×	×	–	×	×	0	×	0	0	0	0	–	–	–	0	–	–	0	–	–	0	0	0	–	0	0
Guttera	0	0	0	0	×	–	–	×	–	–	×	–	–	0	0	0	–	–	–	–	0	–	–	–	0	0	–	–	0	–
Rollulus	0	0	×	0	–	×	×	×	×	0	×	–	×	×	0	0	0	0	–	0	0	0	–	–	0	0	0	–	0	–
Francolinus	×	0	×	×	×	×	×	×	×	×	×	–	×	×	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bambusicola	×	×	–	×	×	–	×	–	×	×	×	–	×	×	0	0	0	0	0	0	0	–	–	0	0	0	0	–	0	0
Gallus gallus[a]	×	×	×	×	×	×	×	×	×	×	×	×	×	×	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Gallus varius	×	×	×	×	×	×	×	×	×	×	×	×	×	×	0	0	0	0	0	0	0	0	0	0	0	0	0	–	0	0
Pavo	×	0	×	0	×	×	×	×	×	0	×	×	×	×	0	0	0	0	0	0	0	0	0	0	0	0	0	0	–	0
Polyplectron	×	0	×	0	×	×	×	×	×	0	×	×	–	–	0	0	0	0	0	–	0	0	0	0	0	0	0	–	×	0
Alectoris	×	0	×	0	–	×	×	–	×	0	×	–	–	×	0	×	–	0	0	0	0	0	0	0	0	0	0	–	–	0
Coturnix coturnix	×	0	–	0	×	×	–	–	×	0	–	–	×	×	0	×	–	0	–	0	0	–	–	–	–	0	–	0	×	0
Coturnix japonica	×	0	×	0	×	×	×	×	×	0	×	×	×	×	0	×	–	0	0	0	0	–	–	0	–	0	0	0	×	0
Phasianus	×	0	×	0	×	–	×	×	×	0	×	×	×	–	×	–	0	0	0	×	0	0	–	×	0	×	–	×	×	×
Crossoptilon	×	0	×	0	×	×	×	×	×	0	×	×	×	×	×	×	0	0	0	×	0	0	0	×	0	×	×	×	×	×
Tragopan	×	0	×	0	×	×	×	×	×	0	×	×	×	×	×	×	0	0	0	×	0	–	0	×	0	–	×	×	×	×
Bonasa	–	0	×	0	×	×	×	×	×	0	×	×	×	–	×	×	0	0	0	–	×	0	–	×	0	×	–	–	×	×
Centrocercus	×	0	×	0	×	×	×	×	×	0	×	×	×	×	×	×	0	0	0	×	×	0	–	×	0	×	×	×	×	×
Tetrao	×	0	×	0	×	×	×	×	–	0	×	×	×	×	×	×	0	0	–	×	×	0	–	×	0	×	×	×	×	–
Lagopus lagopus	×	0	×	0	×	×	×	–	–	0	×	×	×	–	×	×	0	0	0	×	×	0	–	×	0	×	×	×	×	×
Lagopus mutus	–	0	×	0	×	×	×	–	–	0	×	×	×	–	×	×	0	0	0	×	×	0	–	×	0	×	×	×	×	×
Meleagris gallopavo[a]	×	0	×	0	×	×	×	×	×	0	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×
Meleagris ocellata	×	0	×	0	×	×	×	×	×	0	×	–	×	–	×	×	×	×	×	×	×	–	×	×	×	×	–	×	×	×

NOTE.—× = insertion present; 0 = insertion absent; – = no data obtained. 1 = 4755; 2 = 24741; 3 = 19948; 4 = 14195; 5 = 25774; 6 = 27300; 7 = 6825; 8 = 8684; 9 = 26384; 10 = 5407; 11 = 14297; 12 = 20769; 13 = 13948; 14 = 19495; 15 = 455-141; 16 = 458-170; 17 = 458-569; 18 = 464-577; 19 = 465-524; 20 = 515-493; 21 = 464-741; 22 = 515-280; 23 = 515-302; 24 = 574-779; 25 = 580-666; 26 = 580-812; 27 = 580-129; 28 = 588-636; 29 = 588-887; 30 = 588-147.

[a] *Gallus gallus* and *Meleagris gallopavo* verified for, respectively, 11 and 2 individuals.

Material online). Moreover, to confirm the basal position of the chicken we tested for polymorphism by genotyping all informative markers in 10 chicken from equally many different chicken breeds. This confirmed the results from the initial analysis; all chicken either had or had not a particular repeat insertion.

When we compared the average divergence of a *CR1* element from its master sequence within a subfamily versus the position of the insertion in the galliform tree, no clear pattern could be observed (table 3). *CR1* elements belonging to subfamilies of 4–15% divergence were observed on basal branches as well as on more derived parts of the tree, that is, the position in the tree was not associated with the ranked divergence value from the respective master sequence (Spearman rank correlation $R = 0.12$; $N = 13$; $P > 0.1$) (chicken-specific insertions and those insertions that could not be ascribed to a subfamily were excluded from this test).

Similar to previous sequence-based studies (e.g., Dimcheff et al. 2002; Smith et al. 2005), phylogenetic analysis of a 7-kb alignment encompassing the concatenation of *CR1* 5′ and 3′ flanking regions of the 18 sequence loci (using the TN + G + I model) revealed low bootstrap support and short internodes for the divergences leading to chicken, peacock, Old World quail, and turkey (supplementary fig. S2, Supplementary Material online). These internodes measured less than one-tenth the length of the relative time of divergence among all Phasianidae. Based on available divergence times for crown Phasianidae (van Tuinen and Dyke 2004; Pereira and Baker 2006), the rapid diversification leading to chicken, quail, and turkey may have occurred over a time span less than 5 Myr. The resulting phylogeny (Neighbor-Joining tree) based on this 7-kb data set showed significant (>90%) bootstrap support for a grouse + turkey clade, a grouse/turkey + core phasianine clade, phasianid monophyly, and a sister-group relationship between Phasianidae and New World quails. However, despite based on relatively large amount of sequence data, the tree did not exactly recover the *CR1* cladogram and failed to find the basal position of chicken.

Discussion

St John et al. (2005) used the insertion pattern of a single *CR1* locus to clarify the phylogenetic position of 2 goose species, and Watanabe et al. (2006) used *CR1* elements for phylogenetic analysis of 5 penguin species but, to our knowledge, the present report constitutes the first large-scale study in avian systematics where a suite of transposable elements has been used to infer the tree topology among a wide range of species and families. It also forms one of only few studies using the approach of whole-genome sequence survey for the identification of candidate markers, as opposed to performing anonymous library screenings for new markers (Okada et al. 2003).

Galliform Phylogenetics

The cladogram we obtained for 27 galliform species contained 10 branching points, with 5 internodes supported

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

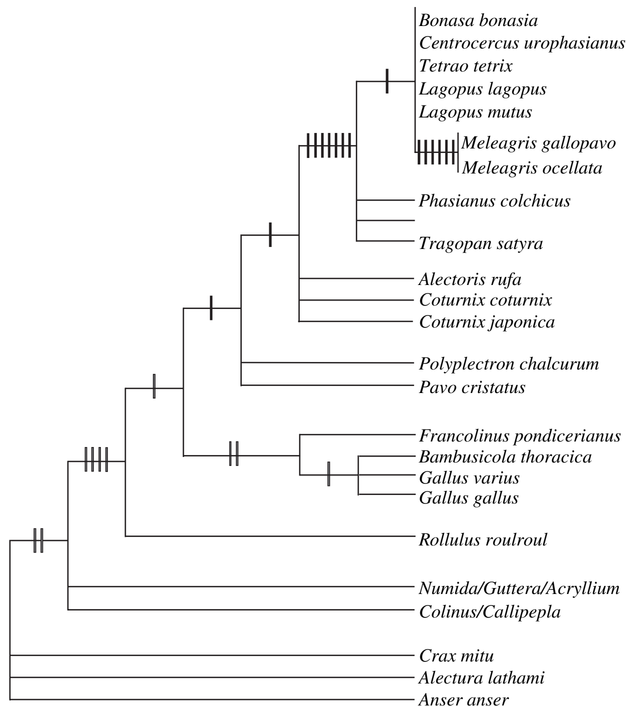

FIG. 3.—Cladogram constructed on the basis of insertion pattern of *CR1* retrotransposons in Galliformes. Each bar indicates a *CR1* insertion shared among the species derived from that branch. Branches 1–5: chicken-derived markers; branches 6–10: turkey-derived markers. Branch numbers are referred to in table 3.

Table 3
Position on the Cladogram (See fig. 3) for Individual *CR1* Copies, Subfamily Affiliation, and Mean Divergence within Subfamilies

Branch	Marker	Subfamily	Average divergence (%)
1	27300		
1	26384	E	12
2	13948		
2	19495	C	7
2	19948	C	7
2	20769		
3	4755		
4	5407	H	4
4	14195	G	7
5	24741	C	7
6	588-887	Y	10
7	458-170	C	7
8	580-812	Outgroup to F, G, H	
8	574-779		
8	588-636	Outgroup to F, G, H, X	
8	515-493	C	7
8	455-141	H	4
8	588-147	C	7
8	580-129		
9	464-741	X	15
10	464-577	B(2)	4
10	458-569	Outgroup to F, G, H, X	
10	580-666	Outgroup to F, G, H, X	
10	465-524	X	15
10	515-302	Y(2)	10
10	515-280	G	7

NOTE.—Some turkey-derived markers could not be assigned to a particular *CR1* subfamily but formed outgroup to known subfamilies in chicken.

by two or more clade-specific *CR1* insertions. The topology corroborates previous notions from diverse data sets in that megapodes and cracids are distantly related to the monophyletic Phasianidae. In support with molecular but not morphology-based phylogenies (Johnsgard 1986; Dyke et al. 2003), the cladogram also places the New World quail outside of Phasianidae, hence confirming the polyphyletic nature of the "quails." The branching pattern within Phasianidae is to a large extent in line with recent phylogenetic trees based on mitochondrial or nuclear sequences (see, e.g., Dimcheff et al. 2002; Dyke et al. 2003, Smith et al. 2005), showing that, though supported by a single insertion only, the pheasants and tragopan form an outgroup to the turkeys (Meleagrinae) and the clade containing grouse and ptarmigan (Tetraoninae). The close relationship among jungle fowl, Chinese bamboo partridge, and francolin and between peafowl and peacock pheasant is also consistently recovered with mitochondrial sequence data (Kimball et al. 1999; Dimcheff et al. 2002; Bush and Strobeck 2003). Our topology does not recover monophyly of phasianines (pheasants) because jungle fowl does not group with pheasant, white-eared pheasant, or tragopan. The partridges group in several places on the cladogram: the red-legged partridge with Old World quail, the Chinese bamboo partridge with jungle fowl (based on a single insertion), and the crested wood partridge (*R. roulroul*) at the base of Phasianidae. Polyphyly of partridges and pheasants has been suggested previously from phylogenetic study of the cytochrome b gene (Kimball et al. 1999). Interestingly, *Rollulus* has never been included in molecular phylogenetic studies of Galliformes. The *CR1* results thus reveal the presence of a primitive phasianid lineage unknown before. Because the relict crested wood partridge is currently registered as near threatened on the International Union for the Conservation of Nature and Natural Resources red list of threatened species, further study is warranted to investigate whether other yet unsampled perdicine species are part of this primitive lineage. No phylogenetic resolution was found for the branching order of New World quail and guinea fowl relative to Phasianidae; this resolution will likely come from development of additional *CR1* markers active at deeper times than used in the present study.

Most phylogenetic uncertainty among galliform molecular studies exists in the relative placement of chicken, quail, turkey, and peafowl: our data provide resolution among these lineages and show that chicken is basal to peafowl, turkey, and Old World quail. Previous studies have proposed several hypotheses: most studies placed Old World quails (*Coturnix*) at the base of phasianids, with chicken (and francolins) branching off from the remaining genera shortly after the Old World quail split (Dimcheff et al. 2002; Smith et al. 2005; but see Shibusawa et al. 2004a). A second study placed turkeys and grouse at the base of phasianids (Sibley and Ahlquist 1990). Based on amino acid analysis of cytochrome b, the pavonine peacocks are placed basally with turkeys as outgroup to a jungle fowl–Old World quail clade (Kimball et al. 1999). Nucleotide analysis of cytochrome b instead supports jungle fowl and francolins as the initial phasianid divergence. Importantly, these studies show little bootstrap support for any

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

of these branching orders. If one accepts retrotransposition events as unidirectional, essentially homoplasy free and generally identical by descent (see below), this means that the phylogenetic position of chicken within Galliformes is now firmly placed near the base of Phasianidae. One important consequence is that comparative genomic studies using sequence or karyotype data from chicken, turkey, and quail—the latter 2 representing the second-most well-studied galliform species in agricultural and evolutionary research (Kikuchi et al. 2005; Reed et al. 2005)—should use chicken as the outgroup, not quail.

Substitution rate heterogeneity among lineages can affect both the resolution and branch lengths of sequence-based trees (van Tuinen and Dyke 2004). We have recently found evidence for a higher rate of neutral sequence evolution in the *Coturnix* quail than in the chicken and turkey lineages, possibly due to a generation time effect (Berlin et al. 2006). It is possible that this can at least in part explain the discrepancy between the relative positions of chicken and *Coturnix* quail in the *CR1*-based cladogram versus sequence-based trees. Albeit based on a limited number of loci, it is noteworthy in this respect that we observed an unusually high rate of genomic deletions in *Coturnix* quail compared with the other galliform species.

Strength and Weakness of Using Retrotransposons as Phylogenetic Markers

Some caveats of using insertion events as phylogenetic markers do exist. Even though our cladogram seems to reflect a distinct order of branching events, it does not exclude the scenario of incomplete lineage sorting accompanied with a rapid radiation in the early evolution of phasianid birds. Like for nucleotide substitutions (Poe and Chubb 2004), incomplete lineage sorting of genomic segments containing recently inserted and still polymorphic *CR1* elements is possible in theory (Tachida and Iizuka 1993) and has been found in at least one retrotransposon study (Nikaido et al. 2006). The likelihood for this is more related to the length than to the depth of internodes, meaning that even for quite old divergences, gene trees may not necessarily be congruent with species trees (Degnan and Rosenberg 2006). Everything else being equal, internodes supported by two or more retrotransposition events might be viewed as more congruent with species trees than those supported by singletons (but see Degnan and Rosenberg 2006). The short lengths of basal phasianid internodes in trees based on sequence data (including the *CR1* flanking sequences analyzed here) indicate rapid diversification and also illustrate the usefulness of combining retrotransposition-based cladograms with trees based on DNA sequence evolution to investigate the timescales of evolution involved (Nikaido et al. 1999).

It has been suggested that *CR1* repeats and other mobile elements will become important tools for solving systematic relationships among birds (Edwards et al. 2005). Our study provides proof of principle for this suggestion and illustrates both the strength and the limitations of using transposable elements in avian systematics. On the positive side, our results clearly indicate that *CR1* elements can resolve particularly difficult topologies where standard sequence-based methods have revealed conflicting results. Moreover, our study also shows the utility of the chicken genome sequence, providing a basically unlimited and easily accessible source of informative markers. Although the avian genome shows a significantly lower density of interspersed repeats than the human genome—constituting $\approx$10% versus 45% of the genome, respectively—there are still 200,000 *CR1* copies in chicken (ICGSC 2004). A further benefit of the chicken genome sequence is that the wealth of potentially useful markers means that loci with particularly conserved flanking exons can be selected for PCR amplification. Finally, we note that extensive taxon sampling should be less critical in this cladistic approach than with standard phylogenetic methods (Hillis et al. 2003).

The fact that the chicken-derived markers were not able to clarify the branching order among the majority of other Phasianidae species illustrates the limitation of using mobile elements as phylogenetic markers. Specifically, this approach cannot resolve topologies of clades branching off the focal clade (i.e., the lineage leading to the species in which markers were originally identified) either before or after the insertion event. Fortunately, we could obtain additional markers based on BAC clone sequences from the turkey genome and thus make inference on 5 further branching points within the clade that split off from the chicken lineage early in phasianid evolution. Ideally, one would have liked to start by identifying insertion sites that arose between the most basal and the most derived species in the whole tree in order to be able to resolve the order of all internal branching points. Of course, this approach implies an a priori assumption of what the most derived group might be, but it does not introduce any bias into the process of tree reconstruction.

Using agarose gel electrophoresis to score for the presence or absence of *CR1* markers is in itself associated with both advantages and disadvantages. The method is obviously rapid and cheap. On the other hand, length variation of amplification products due to mutational events other than *CR1* insertion introduces noise when genotyping is based on fragment size analysis alone. To the price of reducing the number of potentially informative markers, such noise can of course be avoided by only using loci that show discrete length variation consistent with the presence or absence of a *CR1* element of known size; this is the approach we used, leading to a reduction from 44 to 30 markers. However, using DNA sequencing or Southern blotting for scoring *CR1* repeats circumvents this problem.

A final cautionary note relates to the often-cited characteristic of retrotransposons generally being homoplasy free. Homoplasious similarity of SINEs has been demonstrated (Pecon-Slattery et al. 2004), and in theory, other such examples may have been missed in earlier studies based on scoring of PCR product lengths. It may be useful in this context to distinguish between homoplasy in reality and homoplasy in scoring. Accepting the former to be a rare phenomenon, the support from two or more insertions at a particular branch would definitely strengthen the inferred phylogeny. Moreover, some retrotransposons are severely truncated upon insertion at a chromosomal site. Even if one accepts the possibility that 2 independent insertions could

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

happen at the same position during different time points in evolution, the combined probability that these insertions would also be of more or less exactly the same size should be very low. In the case of avian *CR1* elements, most inserted elements vary in size from 50 to 500 bp (ICGSC 2004).

Homoplasy in scoring seems more of a realistic concern when genotyping is based on sizing alone, although the independent insertion of 2 repeat elements of different lengths should be possible to detect with this approach. As discussed above, hybridization to Southern blots using a repeat probe would confirm the presence on an insertion and DNA sequencing offers a formal demonstration of the sequence contained within amplified fragments, as done in this study. Southern blot hybridization and/or DNA sequencing also provide means for excluding the possibility of character loss from small- to medium-size deletion events (larger deletions would make PCR amplification impossible).

The Future of *CR1* Elements in Avian Systematics

Although the analysis of completely sequenced mitochondrial DNA genomes and, more recently, an increasing use of nuclear sequences have been helpful to avian phylogenetics (e.g., Groth and Barrowclough 1999; van Tuinen et al. 2000, 2001; Ericson et al. 2002; Garcia-Moreno et al. 2003; Fain and Houde 2004; Poe and Chubb 2004; Slack et al. 2006), many groups and relationships remain problematic (Cracraft et al. 2004). Examples include the higher level tree topology of major Neoavian orders and the relationships within groups such as ratites and passerines. Available data suggest that ratites and tinamous (Paleognathae) represent the most basal clade among extant bird lineages and that Anseriformes and Galliformes are sister lineages, together forming the most basal clade (Galloanserae) among Neognathae (nonratite birds) (Groth and Barrowclough 1999; van Tuinen et al. 2000). We anticipate that *CR1* markers derived from the chicken genome will find a useful application for obtaining independent support for these conclusions.

Chicken-derived markers shall also be able to reveal deeper family relationships within the Galliformes. However, if the above-mentioned basal structure of the avian tree is correct, chicken-derived *CR1* elements cannot be used for cladogram construction within other major avian clades. This might become less of a problem in the future given the steady accumulation of large-scale sequence data based on BAC sequencing from several bird species (Edwards et al. 2005; see also Hess et al. 2000; Gasper et al. 2001), including emu (*Dromaius novaehollandiae*), zebra finch (*Taeniopygia guttata*), and California condor (*Gymnogyps californianus*) (Thomas et al. 2003), each being representatives of orders with problematic phylogenies. At the time of writing, about 10 Mb of BAC clone sequence from bird species other than chicken is available in GenBank, containing an estimated >1000 *CR1* elements. Moreover, the sequencing of the zebra finch genome is soon to be completed, which should provide for an inexhaustible source of markers for a second bird species. It should also be noted that, once extensive amounts of orthologous sequence data from several bird species are available, in silico approaches for scoring the distribution of repeat elements shall become possible (cf. Bashir et al. 2005).

The time range over which *CR1* markers will be useful in avian systematics depends on the temporal distribution of their activity. Mean divergence within subfamilies ranges between 3% and 18% (ICGSC 2004). Assuming that mean divergence reflects the time of bursts of transposition events and using a nuclear rate estimate as in Axelsson et al. (2004), this would indicate that different subfamilies were active 20–110 MYA. This is a suitable time span both for studying basal and more recent divergences among extant bird lineages. We did not find any correlation between the mean divergence within a particular subfamily and the relative position of markers in the cladogram. Several explanations for this finding are possible, including the relatively short time span of resolved nodes and the wide range of divergence estimates of individual repeat elements within a subfamily. Divergence estimates of repeats may be more important when zooming in on more ancient events, in which case elements which show low divergence compared with the master consensus should be avoided. Rate heterogeneity in the spread of subfamilies of transposons (i.e., variation in temporal activity) seems to occur among all repeat types and in all organisms (e.g., IHGSC 2001, MGSC 2002). In itself, it represents more of a benefit than a problem to phylogenetic studies because it allows for choosing markers that are more likely to be informative at a particular time point of evolution than is the case for any random DNA sequence.

In summary, our study has shown the utility of retrotransposon markers in avian systematics and has also revisited the phylogenetic position of chicken within Galliformes. More generally, this study, together with a number of recent reports using short insertion and deletion mutations in unique sequence (Groth and Barrowclough 1999; Ericson and Johansson 2003; Prychitko and Moore 2003; Fain and Houde 2004; Irestedt et al. 2004), illustrates the usefulness of cladistic approaches even in days when DNA sequence data accumulate faster than ever.

Supplementary Material

Supplementary Figure S1, Supplementary Figure S2, and Supplementary Table S1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Matthew Webster for providing data on chicken *CR1* elements, Mikael Brandström for computational help, Jack Dumbacher and Maureen Flannery at the California Academy of Sciences for providing DNA samples from several galliform species, and Carla Cicero at the University of California Berkeley Museum of Vertebrate Zoology for access to DNA of a second American quail species. We also thank N. Okada for discussion and 2 anonymous reviewers for useful comments on an earlier version of the manuscript. The study was supported by the Swedish Research Council.

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

Literature Cited

Armstrong MH, Braun EL, Kimball RT. 2001. The phylogenetic utility of avian ovomucoid intron G: a comparison of nuclear and mitochondrial phylogenies in Galliformes. Auk. 118: 799–804.

Axelsson E, Smith NGC, Sundström H, Berlin S, Ellegren H. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. Mol Biol Evol. 21:1538–1547.

Bashir A, Ye C, Price AL, Bafna V. 2005. Orthologous repeats and mammalian phylogenetic inference. Genome Res. 15: 998–1006.

Berlin S, Brandström M, Backström N, Axelsson E, Smith NGC, Ellegren H. Forthcoming 2006. Substitution rate heterogeneity and the male mutation bias. J Mol Evol. 2262–2323.

Brown WR, Hubbard SJ, Tickle C, Wilson SA. 2003. The chicken as a model for large-scale analysis of vertebrate gene function. Nature Rev Genet. 4:87–98.

Bush KL, Strobeck C. 2003. Phylogenetic relationships of the Phasianidae reveals possible non-pheasant taxa. J Hered. 94:472–489.

Cooper MD, Raymond DA, Peterson RD, South MA, Good RA. 1966. The functions of the thymus system and the bursa system in the chicken. J Exp Med. 123:75–102.

Cracraft J, Barker FK, Braun M. et al. (14 co-authors) 2004. Phylogenetic relationships among modern birds (Neornithes). In: Cracraft J, Donoghue MJ, editors. Assembling the tree of life. New York: Oxford University Press. p. 468–489.

Degnan JH, Rosenberg NA. 2006. Discordance of species trees with their most likely gene trees. PloS Genet. 2:e68.

Dimcheff DE, Drovetski SV, Mindell DP. 2002. Phylogeny of Tetraoninae and other galliform birds using mitochondrial 12S and ND2 genes. Mol Phylogenet Evol. 24: 203–215.

Dyke GJ, Gulas BE, Crowe TM. 2003. Suprageneric relationships of galliform birds (Aves: Galliformes): a cladistic analysis of morphological characters. Zool J Linn Soc. 137:227–244.

Edwards SV, Bryan Jennings W, Shedlock AM. 2005. Phylogenetics of modern birds in the era of genomics. Proc R Soc Lond Ser B. 272:979–992.

Ellsworth DL, Honeycutt RL, Silvy NJ. 1995. Phylogenetic relationships among North American grouse inferred from restriction endonuclease analysis of mitochondrial DNA. Condor. 97:492–502.

Ericson PGP, Johansson US. 2003. Phylogeny of Passerida (Aves: Passeriformes) based on nuclear and mitochondrial sequence data. Mol Phylogenet Evol. 29:126–138.

Ericson PG, Christidis L, Cooper A, Irestedt M, Jackson J, Johansson US, Norman JA. 2002. A Gondwanan origin of passerine birds supported by DNA sequences of endemic New Zealand wrens. Proc R Soc Lond B. 269:235–241.

Fain MG, Houde P. 2004. Parallel radiations in the primary clades of birds. Evolution. 58:2558–2573.

Garcia-Moreno J, Sorenson MD, Mindell DP. 2003. Congruent avian phylogenies inferred from mitochondrial and nuclear DNA sequences. J Mol Evol. 56:1–11.

Gasper J, Shiina T, Inoko H, Edwards SV. 2001. Songbird genomics: analysis of 45 kb upstream of a polymorphic Mhc class II gene in red-winged blackbirds (*Agelaius phoeniceus*). Genomics. 75:26–34.

Groth JG, Barrowclough GF. 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol Phylogenet Evol. 12:115–123.

Gutierrez RJ, Zink RM, Yang SY. 1983. Genic variation, systematic biogeographic relationships of some galliform birds. Auk. 100:33–47.

Haas NB, Grabowski JM, Sivitz AB, Burch JBE. 1997. Chicken repeat 1 (*CR1*) elements, which define an ancient family of vertebrate non-LTR retrotransposons, contain two closely spaced open reading frames. Gene. 197:305–309.

Helm-Bychowski KM, Wilson AC. 1986. Rates of nuclear DNA evolution in pheasant-like birds: evidence from restriction maps. Proc Natl Acad Sci USA. 83:688–692.

Henderson JY, Moir AJ, Fothergill LA, Fothergill JE. 1981. Sequences of sixteen phosphoserine peptides from ovalbumins of eight species. Eur J Biochem. 114:439–450.

Hess CM, Gasper J, Hoekstra H, Hill C, Edwards SV. 2000. MHC class II pseudogene and genomic signature of a 32-kb cosmid in the house finch (*Carpodacus mexicanus*). Genome Res. 10:613–623.

Hillis DM, Pollock DD, McGuire JA, Zwickl DJ. 2003. Is sparse taxon sampling a problem for phylogenetic inference. Syst Biol. 52:124–126.

[ICGSC] International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature. 432:695–777.

[IHGSC] International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of the human genome. Nature. 409:860–921.

Irestedt M, Fjeldså J, Nylander JAA, Ericson PGP. 2004. Phylogenetic relationships of typical antbirds (Thamnophilidae) and test of incongruence based on Bayes factors. BMC Evol Biol. 4:23.

Johnsgard PA. 1986. The pheasants of the world. Oxford: Oxford University Press.

Jolles J, Schoentgen F, Jolles P, Prager EM, Wilson AC. 1976. Amino acid sequence and immunological properties of chachalaca egg white lysozyme. J Mol Evol. 8:59–78.

Kajikawa M, Ichiyanagi K, Tanaka N, Okada N. 2005. Isolation and characterization of active LINE and SINEs from the eel. Mol Biol Evol. 22:673–682.

Kapitonov VV, Jurka J. 2003. The esterase and PHD domains in CR1-like non-LTR retrotransposons. Mol Biol Evol. 20: 38–46.

Kikuchi S, Fujima D, Sasazaki S, Tsuji S, Mizutani M, Fujiwara A, Mannen H. 2005. Construction of a genetic linkage map of Japanese quail (*Coturnix japonica*) based on AFLP and microsatellite markers. Anim Genet. 36:227–231.

Kimball RT, Braun EL, Zwartjes PW, Crowe TM, Ligon JD. 1999. A molecular phylogeny of the pheasants and partridges suggests that these lineages are not monophyletic. Mol Phylogenet Evol. 11:38–54.

[MGSC] Mouse Genome Sequencing Consortium. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature. 420:520–562.

Murata S, Takasaki N, Saitoh M, Okada N. 1993. Determination of the phylogenetic relationships among Pacific salmonids by using short interspersed elements (SINEs) as temporal landmarks of evolution. Proc Natl Acad Sci USA. 90:6995–6999.

Nikaido M, Hamilton H, Makino H, Sasaki T, Takahashi K, Goto M, Kanda N, Pastene LA, Okada N; SMBE Tri-National Young Investigators. 2006. Proceedings of the SMBE Tri-National Young Investigators' Workshop 2005. Baleen whale phylogeny and past extensive radiation event revealed by SINE insertion analysis. Mol Biol Evol. 23: 866–873.

Nikaido M, Rooney AP, Okada N. 1999. Phylogenetic relationships among cetartiodactyls based on insertions of short and long interspersed elements: hippopotamuses are the closest extant relatives of whales. Proc Natl Acad Sci USA. 96:10261–10266.

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024

Nishihara H, Satta Y, Nikaido M, Thewissen JGM, Stanhope MJ, Okada N. 2005. A retrosposon analysis of Afrotherian phylogeny. Mol Biol Evol. 22:1823–1833.
Okada N. 1991. SINEs: short interspersed repeated elements of the eukaryotic genome. Trends Ecol Evol. 6:358–361.
Okada N, Shedlock A, Nikaido M. 2003. Retroposon mapping in molecular systematics. In: Miller WJ, Capy P, editors. Methods in molecular biology. Totowa (NJ): Humana Press. p. 189–226.
Pecon-Slattery J, Pearks Wilkerson AJ, Murphy WJ, O'Brien SJ. 2004. Phylogenetic assessment of introns and SINEs within the Y chromosome using the cat family Felidae as a species tree. Mol Biol Evol. 21:2299–2309.
Pereira SL, Baker AJ. 2006. A molecular time-scale for galliform birds accounting for uncertainty in time estimates and heterogeneity of rates of DNA substitutions across lineages and sites. Mol Phylogenet Evol. 38:499–509.
Pizzari T, Cornwallis CK, Lovlie H, Jakobsson S, Birkhead TR. 2003. Sophisticated sperm allocation in male fowl. Nature. 426:70–74.
Poe S, Chubb AL. 2004. Birds in a bush: five genes indicate explosive evolution of avian orders. Evolution. 58:404–415.
Prager EM, Wilson AC. 1976. Congruency of phylogenies derived from different proteins. A molecular analysis of the phylogenetic position of cracid birds. J Mol Evol. 9:45–57.
Prychitko TM, Moore WS. 2003. Alignment and phylogenetic analysis of b-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). Mol Phylogenet Evol. 8:193–204.
Ray DA, Xing J, Hedges DJ, et al. (13 co-authors). 2005. *Alu* insertion loci and platyrrhine primate phylogeny. Mol Phylogenet Evol. 35:117–126.
Reed KM, Chaves LD, Hall MK, Knutson TP, Harry DE. 2005. A comparative genetic map of the turkey genome. Cytogenet Genome Res. 111:118–127.
Sasaki T, Takahashi K, Nikaido M, Miura S, Yasukawa Y, Okada N. 2004. First application of the SINE (short interspersed repetitive element) method to infer phylogenetic relationships in reptiles: an example from the turtle superfamily Testudinoidea. Mol Biol Evol. 21:705–715.
Shedlock AM, Okada N. 2000. SINE insertions: powerful tools for molecular systematics. BioEssays. 22:148–160.
Shibusawa M, Nishibori M, Nishida-Umehara C, Tsudzuki M, Masabanda J, Griffin DK, Matsuda Y. 2004a. Karyotypic evolution in the Galliformes: an examination of the process of karyotypic evolution by comparison of the molecular cytogenetic findings with the molecular phylogeny. Cytogenet Gen Res. 106:111–119.
Shibusawa M, Nishida-Umehara C, Tsudzuki M, Masabanda J, Griffin DK, Matsuda Y. 2004b. A comparative karyological study of the blue-breasted quail (*Cotunix chinensis*, Phasianidae) and California quail (*Callipepla californica*, Odontophoridae). Cytogenet Gen Res. 106:82–90.
Shimamura M, Yasue H, Oshima K, Abe H, Kato H, Kishiro T, Goto M, Munechika I, Okada N. 1997. Molecular evidence from retroposons that whales form a clade within even-toed ungulates. Nature. 388:666–670.
Sibley GC, Ahlquist JE. 1990. Phylogeny and classification of birds: a study in molecular evolution. New Haven (CT): Yale University Press.
Slack KE, Jones CM, Tatsuro A, Harrison GL, Fordyce RE, Arnason U, Penny D. 2006. Early penguin fossils, plus mitochondrial genomes, calibrate avian evolution. Mol Biol Evol. 23:1144–1155.
Smith EJ, Shi L, Tu Z. 2005. *Gallus gallus* agrecan gene-based phylogenetic analysis of selected avian groups. Genetica. 124:23–32.
Stehelin D, Varmus HE, Bishop JM, Vogt PK. 1976. DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA. Nature. 260:170–173.
St John J, Cotter JP, Quinn TW. 2005. A recent chicken repeat 1 retrotransposition confirms the Coscoroba–Cape Barren goose clade. Mol Phylogenet Evol. 37:83–90.
Tachida H, Iizuka M. 1993. A population genetic study of the evolution of SINEs. I. Polymorphism with regard to the presence or absence of an element. Genetics. 133:1023–1030.
Takahashi K, Nishida M, Yuma M, Okada N. 2001. Retroposition of the AFC family of SINEs (short interspersed repetitive elements) before and during the adaptive radiation of cichlid fishes in Lake Malawi and related inferences about phylogeny. J Mol Evol. 53:496–507.
Terai Y, Takahashi K, Nishida M, Sato T, Okada N. 2003. Using SINEs to probe ancient explosive speciation: hidden radiation of African cichlids? Mol Biol Evol. 20:924–930.
Thomas JW, Touchman JW, Blakesley RW, et al. (71 co-authors). 2003. Comparative analyses of multi-species sequences from targeted genomic regions. Nature. 424:788–793.
Vandergon TL, Reitman M. 1994. Evolution of chicken repeat 1 (*CR1*) elements: evidence for ancient subfamilies and multiple progenitors. Mol Biol Evol. 11:886–898.
van de Lagemaat LN, Gagnier L, Medstrand P, Mager DL. 2005. Genomic deletions and precise removal of transposable elements mediated by short identical DNA segments in primates. Genome Res. 15:1243–1249.
van Tuinen M, Butvill DB, Kirsch JAW, Hedges SB. 2001. Convergence and divergence in the evolution of aquatic birds. Proc R Soc Lond Ser B. 268:1345–1350.
van Tuinen M, Dyke GJ. 2004. Calibration of galliform molecular clocks using multiple fossils and genetic partitions. Mol Phylogenet Evol. 30:74–86.
van Tuinen MC, Sibley G, Hedges SB. 2000. The early history of modern birds inferred from DNA sequence of mitochondrial and nuclear ribosomal genes. Mol Biol Evol. 17:451–457.
von Schantz T, Göransson G, Andersson G, Fröberg I, Grahn M, Helgee A, Wittzeel H. 1989. Female choice selects for a viability-based male trait in pheasant. Nature. 337:166–169.
Watanabe M, Nikaido M, Tsuda TT, Inoko H, Mindell DP, Murata K, Okada N. 2006. The rise and fall of the CR1 subfamily in the lineage leading to penguins. Gene. 365:57–66.
Wicker T, Robertson JS, Schulze SR, et al. (11 co-author) 2005. The repetitive landscape of the chicken genome. Genome Res. 15:126–136.

Scott Edwards, Associate Editor

Accepted October 24, 2006

Downloaded from https://academic.oup.com/mbe/article/24/1/338/1076832 by guest on 24 April 2024