Molecular Phylogenetics and Evolution 56 (2010) 536–542

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Comparative molecular evolution and phylogenetic utility of 3′-UTRs and introns in Galliformes

Amber J. Bonilla [1], Edward L. Braun, Rebecca T. Kimball *

Department of Biology, The University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Article history:
Received 12 June 2009
Revised 11 March 2010
Accepted 6 April 2010
Available online 14 April 2010

Keywords:
Game birds
Untranslated regions
Evolutionary rates
Base composition
Tree distances
Non-coding DNA
Isochore

ABSTRACT

Many phylogenetic studies have used nuclear introns or coding exons, but few have included untranslated regions (UTRs). Here we compare the phylogenetic utility and patterns of molecular evolution for 3′-UTRs and introns from five unlinked loci in Galliformes (Aves). 3′-UTRs evolved at slower rates and exhibited greater spatial clustering of sites with similar evolutionary rates than associated introns, though they exhibited similar overall model complexities. Base compositions differed between the 3′-UTR and associated intron for two of five loci, suggesting that base composition was not exclusively driven by isochore structure. Phylogenies estimated using individual and concatenated 3′-UTRs were more similar to an independent reference tree than those based upon introns, though all phylogenies were largely congruent. However, some 3′-UTRs were difficult to amplify, potentially limiting their use in phylogenetic studies. This study suggests that 3′-UTRs and introns can exhibit distinct patterns of molecular evolution and that they provide useful phylogenetic signal.

1. Introduction

Over the past several decades taxonomy and systematics have seen an explosive growth in the availability of molecular data (reviewed by Delsuc et al. (2005)). For vertebrates, mitochondrial DNA (mtDNA) has been an extremely popular phylogenetic tool due to its high evolutionary rate and rapid coalescence, which leads to excellent power (Braun and Kimball, 2001) and a high probability of matching the species tree (Moore, 1995). Additionally, sets of reliable mtDNA PCR primers are available (e.g., Sorenson et al., 1999). However, mtDNA provides a single gene tree that may differ from the species tree (Maddison, 1997) due to either lineage sorting (though the rapid coalescence of mtDNA will minimize this) or past hybridization. Furthermore, mtDNA can also show biased patterns of evolution (Braun and Kimball, 2002; Delsuc et al., 2003; Phillips and Penny, 2003). Finally, the length of the mitochondrial genome limits the amount of mtDNA data that can be obtained.

Nuclear sequence data have proved useful for phylogenetic analyses in vertebrates (e.g., Matthee et al., 2007; Hackett et al., 2008; Harshman et al., 2008; Wiens et al., 2008) with most studies focusing on coding regions or introns rather than intergenic regions or non-coding RNAs. Introns evolve more rapidly than coding exons, leading to a higher probability of accumulating substitutions that unite groups of interest (Chojnowski et al., 2008), though a high rate of evolution may increase homoplasy (e.g., Yang, 1998). Introns can also be amplified using primers anchored in conserved exon regions (Palumbi and Baker, 1994) and sets of vetted primers similar to those available for mtDNA are now available for some groups (e.g., Kimball et al., 2009). However, introns exhibit length variation that makes them more difficult to align than coding regions (Hackett et al., 2008; Harshman et al., 2008; Pratt et al., 2009). Thus, the ideal type(s) of data for vertebrate systematics remains unclear.

One type of region that has not been as extensively examined in phylogenetic studies is the untranslated regions (UTRs) of messenger RNAs (mRNAs). UTRs have some functional regions, such as those necessary for regulation of translation (e.g., Mazumder et al., 2003) and microRNA interaction (e.g., Xie et al., 2005). UTRs are enriched for putative regulatory regions called pyknons (Rigoutsos et al., 2006). Although there are functional regions within UTRs, they are non-coding and so should exhibit less overall functional constraint than coding exons. Consistent with this, UTRs may be easier to align than introns (Harshman et al., 2003) but evolve at higher rates than coding exons (e.g., Murphy et al., 2001; Cooper et al., 2003). These features suggest that UTRs may be very useful for phylogenetics. Indeed, UTRs that have been examined for phylogenetics evolve more slowly than introns (Harshman et al., 2003; Whittall et al., 2006; Kimball et al.,

* Corresponding author. Address: Department of Biology, University of Florida, P.O. Box 118525, Gainesville, FL 32611, USA. Fax: +1 352 392 3704.
E-mail address: rkimball@ufl.edu (R.T. Kimball).
[1] Present address: Hinman Box 7560, Dartmouth College, Hanover, NH 03755, USA.

1055-7903/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2010.04.006

2009), yet still exhibit sufficient signal for phylogenetic reconstruction (Harshman et al., 2003; Whittall et al., 2006) while potentially exhibiting less homoplasy (Harshman et al., 2003). 3′-UTRs in particular have a mean length of >500 base pairs (bp) (e.g., Mazumder et al., 2003), which is large enough to yield many variable sites but short enough to be readily amplified and sequenced using standard technology.

The avian order Galliformes includes many of the best-studied avian species in the fields of genetics, physiology, and development, like the chicken (*Gallus gallus*), guineafowl (*Numida meleagris*), quail (*Coturnix coturnix*), and turkey (*Meleagris gallopavo*). The phylogeny of the Galliformes has long been problematic (reviewed by Sibley and Ahlquist (1990)). While progress has been made recently using various types of data (Kimball et al., 1999, 2006; Armstrong et al., 2001; Crowe et al., 2006; Cox et al., 2007; Kriegs et al., 2007; Kaiser et al., 2007; Kimball and Braun, 2008), there are still open questions and several areas of conflict. In this study we sequenced 3′-UTRs and introns from five unlinked genes in 23 galliform species to determine the characteristics of 3′-UTR evolution, explored the ability of 3′-UTRs to resolve phylogenetic questions, and compared these results to introns at the same locus.

2. Methods

2.1. Amplification, sequencing, and alignment

Twenty-three galliform species used in previous studies (e.g., Kimball et al., 1999; Armstrong et al., 2001; Cox et al., 2007; Kimball and Braun, 2008) were selected, including representatives from all families and major lineages (Appendix A: Table S1). Although grouse and ptarmigan have been placed into their own family (Tetraonidae) in some treatments, they are currently considered a subfamily (Tetraoninae) of the Phasianidae (AOU, 1998) and multigene studies (Kimball and Braun, 2008) strongly support a close relationship between *M. gallopavo* (the turkey) and members of this subfamily. Given the support for this close relationship we only included *Meleagris* in this study. Five nuclear loci that are unlinked in the chicken genome were targeted for amplification (Appendix A: Table S2). Primers for 3′-UTR amplification (Appendix A: Table S3) were designed based upon alignments of chicken mRNAs with assembled American alligator (*Alligator mississippiensis*) expressed sequence tags (ESTs; from Chojnowski et al. (2007)), and/or with mammalian mRNAs. In some cases, we used preliminary sequences to design additional primers. Primers for introns are listed along with the 3′-UTR primers (Table S3) or they were taken from previous publications (Cox et al., 2007; Kimball et al., 2009).

PCR amplification used standard PCR conditions and amplified products were cleaned by PEG:NaCl (20%:2.5 M) precipitation. Cleaned PCR products were sequenced in the forward and reverse direction using the same primers used in PCR amplification. Cycle sequencing was performed using ABI BigDye® Terminator v.3.1 according to manufacturer's recommendations (with the exception of reducing the reaction volumes) and sequence reactions were analyzed using an ABI Prism™ 3100-Avant genetic analyzer (PE Applied Biosystems).

Roughly 25% of the PCR products exhibited length heterozygosity within the 3′-UTR. We cloned these samples using the pGEM T-Easy vector (Promega Corp.) following manufacturer protocols. Plasmids with the target DNA were prepared using the Eppendorf Perfectprep Plasmid Mini kit. Plasmids were sequenced in both directions using the same protocol used for PCR products.

We assembled sequence reads into double-stranded contiguous sequences using Sequencher 4.1 (Gene Codes Corp.). Sequences were aligned across taxa using ClustalX (Thompson et al., 1997) and the resulting alignment was optimized by eye in MacClade 4.0 (Maddison and Maddison, 2000). Unpublished sequences used in this study have been deposited in GenBank (Accession Nos. FJ881696–FJ881859).

2.2. Molecular evolution and phylogenetic analyses

We determined best-fit models of evolution for all regions (single 3′-UTRs and introns, as well as concatenated datasets) using Modeltest 3.7 (Posada and Crandall, 1998). Parameter estimates were those determined by Modeltest, except for transition/transversion (ti/tv) ratios that were calculated using Eq. (1) in Kumar (1996). Since the sequenced regions of EEF2 and HMGN2 contained two introns, model parameters and base composition were assessed independently for each intron.

We used a sliding window analysis to assess patterns of spatial variation across each 3′-UTR and intron and to identify highly conserved regions. Maximum likelihood (ML) estimates of the rates of evolution for each site in the alignments were estimated using PAML (Yang, 1997). Since rates at each site are drawn from a Γ distribution with a mean of one, sites that evolved at the mean rate had a rate of one. The sliding windows used to calculate rates were 20 bp wide with a 10 bp overlap. The standard deviation of these windows was used as a metric of spatial variation in evolutionary rate, with low values indicating similar rates of evolution among windows along the length of the region and high values indicating high heterogeneity in the evolutionary rates among windows. To determine whether the standard deviation was greater than expected (i.e., whether there was more spatial heterogeneity in evolutionary rates than expected), we repeated the analysis for 100 distinct permutations of the sites to establish an expected value for each region.

The ML tree was estimated using PAUP* (version 4.0b10; Swofford, 2003), with a heuristic search using 10 random addition sequence replicates and the model and parameters estimated from Modeltest. ML analyses were performed on the concatenated dataset as well as each individual partition. Since the goal was to obtain estimates of the phylogeny for each data type and locus, we combined the two HMGN2 introns and the two EEF2 introns into a single partition (excluding the short, intervening exon) for the phylogenetic comparisons. ML bootstrap analysis was performed with 100 replicates using GARLI (Zwickl, 2006) using the concatenated intron plus 3′-UTR dataset as well as using an intron-only and 3′-UTR-only dataset. The tree was rooted between the megapodes (Megapodiidae) and the remaining galliforms based upon the robust support for this position of the root provided by recent studies (Crowe et al., 2006; Hackett et al., 2008).

To compare trees generated using ML searches of the different data partitions, we used Robinson and Foulds (1981) distances (RF distances) from a reference phylogeny. To obtain the reference phylogeny, we used an ML search and appropriate models as described above with the loci included in Kimball and Braun (2008), which used two mitochondrial regions and four unlinked nuclear introns for a total of 5533 bp. Since two species in this study were absent from the Kimball and Braun (2008) dataset, we collected sequence data for the Kimball and Braun (2008) loci for the additional taxa, so the reference tree contained all species included in this study. Since the PCBD1 gene was included in both this study and Kimball and Braun (2008), we excluded this locus of Kimball and Braun (2008) in generating the reference tree, so we obtained a reference tree that was independent of the 3′-UTR and intron data presented here (this independent reference tree was actually topologically identical to one that included PCBD1 along with other loci from Kimball and Braun (2008)). To estimate homoplasy, we calculated the retention index for each partition using the reference tree. Evolutionary rates for each partition were determined by estimating branch lengths using the reference tree and

the best model for each partition, then summing branch lengths and normalizing to the length of a total evidence tree for the same taxa. For partitions with missing taxa, the appropriate branch lengths were subtracted from the length of the total evidence tree so data was comparable across partitions.

3. Results

3.1. Patterns of molecular evolution

The final alignments contained 6702 bp from the five unlinked loci, with 3109 bp of UTR data and 3593 of intron data (coding DNA was excluded from these calculations and all analyses). The lengths of the 3′-UTR alignments ranged from 398 bp (ALDOB) to 780 bp (HMGN2), while the intron alignments ranged from 510 bp (ALDOB) to 1166 bp (CRYAA). We identified two microinversions (short inverted segments relative to the remaining taxa) and excluded them from phylogenetic analyses. We found it difficult to reliably amplify the EEF2 3′-UTR region, even after primers were redesigned based on initial sequences, and we obtained sequence data from only 16 taxa successfully. Since key lineages were not represented, we excluded the EEF2 locus from phylogenetic analyses. While 25% of the dataset exhibited length heterozygosity that required cloning, about half of these samples were from

Table 1
Comparison of the 3′-UTR and intron data for each locus. Best-fit models, shape parameter (α) of the gamma distribution, empirical %GC, and transition/transversion ratios are presented. Values for the two introns in HMGN2 and EEF2 are both presented, with the value from the first intron above that of the second.

Gene	Model		α (Γ distribution)		GC (%)		ti/tv	
	UTR	Intron	UTR	Intron	UTR	Intron	UTR	Intron
ALDOB	TVM + Γ	TVM + Γ	2.55	3.96	37.1	37.3	2.0	2.1
CRYAA	K81 + Γ	TVM + Γ	0.62	2.41	49.7	40.4	3.6	2.0
PCBD1	GTR + Γ	TVMef + Γ	1.09	2.21	47.2	50.9	3.1	2.3
HMGN2	GTR + Γ	TVM + Γ/HKY + Γ	0.38	1.45/1.21	40.9	47.8/46.2	3.0	3.1/2.7
EEF2	TVM + Γ	GTR + Γ/TrN + Γ	0.45	1.15/0.80	48.5	47.0/50.9	2.2	2.6/2.4

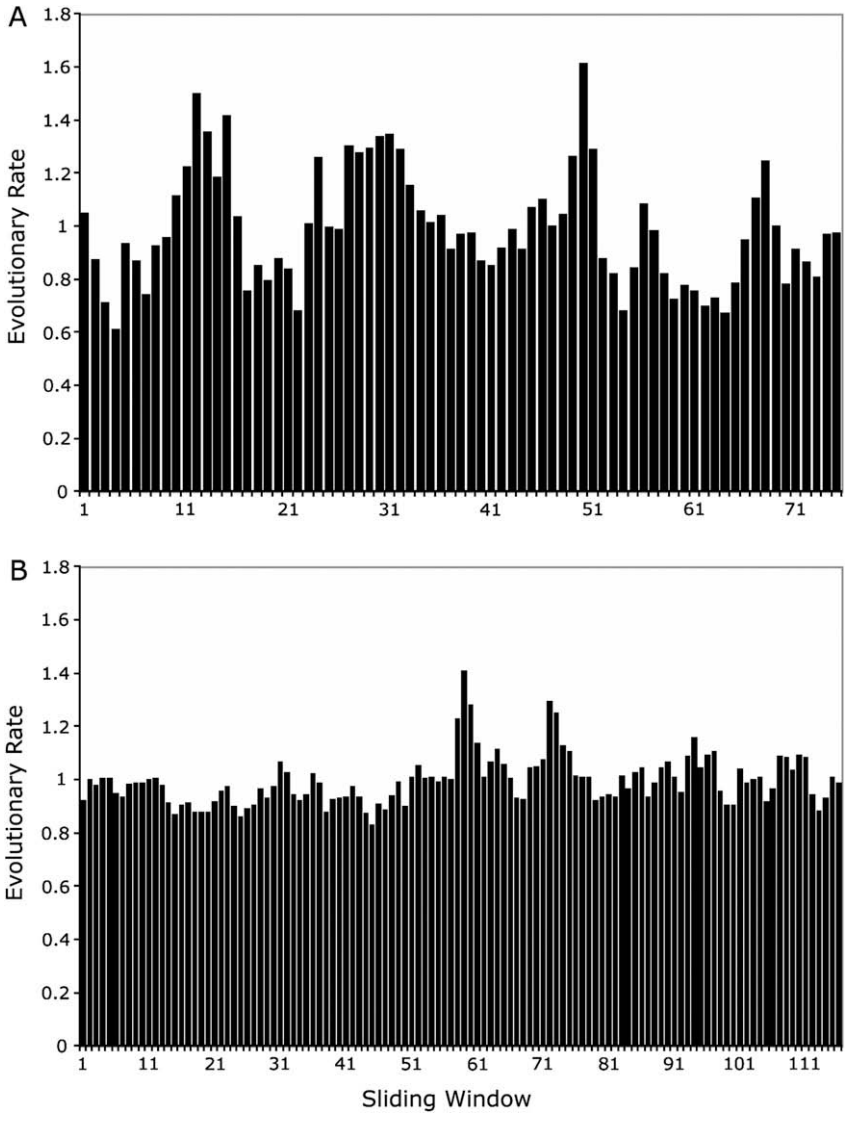

Fig. 1. Sliding window analysis of the evolutionary rates of the CRYAA 3′-UTR (a) and the CRYAA intron (b).

EEF2 amplicons. With the exception of PCBD1 (which was not obtained from *Megapodius layardi*) we obtained sequences of all other loci from all taxa.

The best-fitting models of evolution were generally parameter-rich (GTR + Γ and TVM + Γ), with the exception of the CRYAA 3′-UTR (K81 + Γ, three parameters) and the PCBD1 intron (TVMef + Γ, five parameters). Thus, 3′-UTRs did not differ from introns in average model complexity (Table 1). All loci exhibited moderate site-to-site rate heterogeneity that could be modeled using a Γ distribution (in no case did the best model include invariant sites as well as a Γ distribution). The shape parameters of the Γ distribution were universally lower for the 3′-UTRs than the associated introns, indicating greater among-sites rate heterogeneity in 3′-UTRs.

The loci varied in base composition, though most partitions were AT-biased (Table 1). For three loci, the base composition of the 3′-UTRs and introns were similar. However, the CRYAA 3′-UTR and intron sequences differed by nearly 10% GC content and the 3′-UTR and introns of the HMGN2 locus differed by up to 8%. In the former case the 3′-UTR had greater %GC, while in the latter case the introns had greater %GC. The maximum ti/tv ratio differences were found at the CRYAA locus: the CRYAA 3′-UTR exhibited a ratio of 3.6 whereas the corresponding intron exhibited a ratio of 2.0. For other loci, estimates of the ti/tv ratio were similar.

If the 3′-UTRs contained regulatory elements, we would expect them to exhibit specific regions of high conservation. We quantified the relative clustering of slowly evolving and more rapidly evolving sites by examining the standard deviation of evolutionary rates in a sliding window analysis (Fig. 1). All 3′-UTRs showed greater spatial clustering of rates than expected by chance (Table 2). The 3′-UTRs also showed greater spatial clustering of evolutionary rates than the introns at all loci (particularly for the HMGN2 and CRYAA loci). We would expect this if 3′-UTRs have regions subjected to greater constraints than introns. Two introns, CRYAA and one of the two HMGN2 introns, also showed greater spatial clustering of rates than expected by chance, although these introns still exhibited less spatial clustering of sites with specific rates than their associated 3′-UTRs.

3.2. Phylogenetic reconstruction

The phylogeny obtained from the combined intron and 3′-UTR dataset (Fig. 2) was largely similar to published galliform phylogenies (e.g., Crowe et al., 2006; Cox et al., 2007; Kimball and Braun, 2008). Most nodes found in the combined tree were also present in the 3′-UTR-only and intron-only phylogenies (Fig. 2); the 3′-UTR data appeared to perform well, with fewer conflicting nodes than the intron-only phylogeny. Although the 3′-UTR data appeared to support many of the same nodes (Table 2), bootstrap support for some of the more basal nodes was lower than in the combined or intron phylogenies.

The nodes with the lowest bootstrap support (and greatest conflict) values had short internodes (Fig. 3). Rates of evolution of specific taxa estimated using the 3′-UTR sequences (Fig. 3) were similar to those obtained using the intron-only data (not shown). Both types of data suggested that there has been an overall accelerated rate of evolution in the families Phasianidae and Odontophoridae relative to Megapodiidae, Cracidae, and Numididae. Likewise, taxa within these groups were characterized by specific rates (e.g., *Coturnix japonica* is a relatively long branch) and these rate differences were consistent between the 3′-UTR and intron partitions.

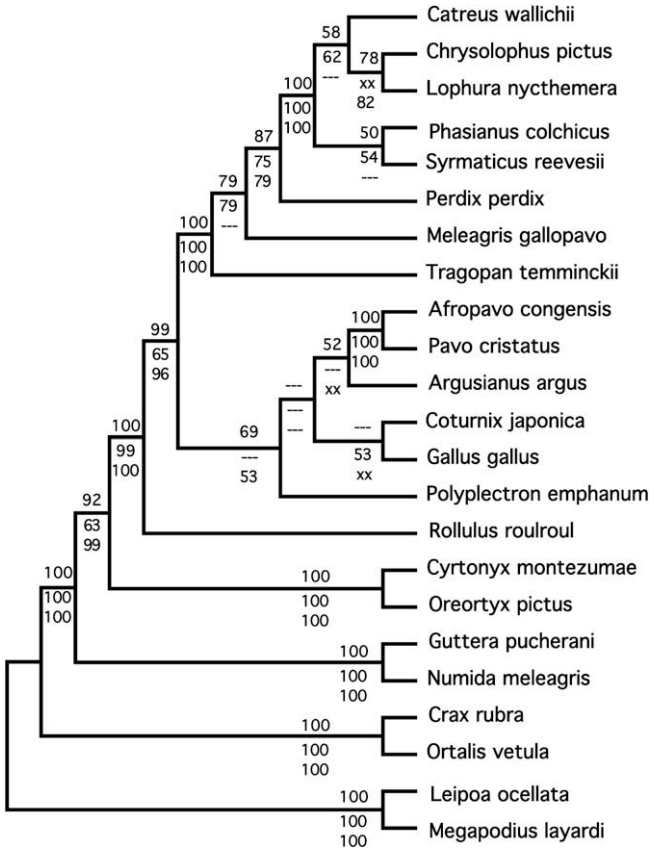

Fig. 2. Maximum likelihood estimate of phylogeny of the combined UTR and intron data. At each node, the top value is the ML bootstrap value from the combined data; the middle value is from ML bootstrap analysis of the UTR data only; and the lower value is from ML bootstrap analysis of the intron data. - - - indicates nodes with less than 50% bootstrap support; ×× indicates nodes conflicting with this topology that have at least 50% bootstrap support.

We used tree distances to quantitatively examine differences among the 3′-UTR, intron and combined phylogenies. We calculated RF distances by comparing the trees generated in this study with a reference tree for the same taxa. The RF distances (Table 3) revealed that the 3′-UTR tree had fewer nodes that were incongruent with the reference tree than the intron tree. This was true when all 3′-UTRs were combined as well as for analyses of individual 3′-UTRs.

Table 2
Standard deviation of the evolutionary rate sliding window analysis. For comparison, the maximum value from 100 permutations is also included. Loci whose standard deviation exceeded that of the maximum permuted value are indicated by an *. Values for the two introns in HMGN2 and EEF2 are both presented.

Gene	Standard deviation		Permutation maximum		Standard deviation
	UTR	Intron	UTR	Intron	UTR–Intron
ALDOB	0.042*	0.0326	0.037	0.037	0.009
CRYAA	0.216*	0.093*	0.179	0.067	0.122
PCBD1	0.154*	0.073	0.141	0.080	0.081
HMGN2	0.294*	0.126/0.174*	0.229	0.147/0.171	0.168/0.144
EEF2	0.443*	0.106/0.247*	0.270	0.137/0.222	0337/0.197

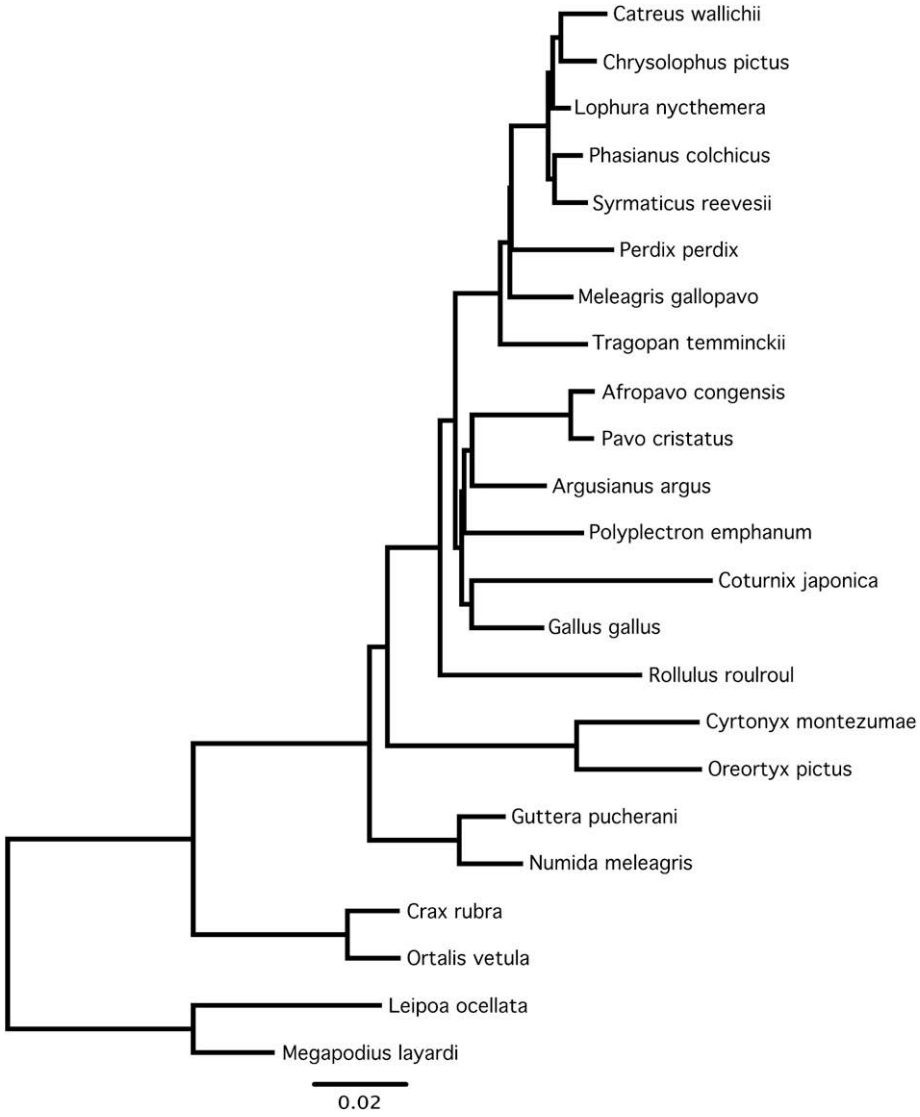

Fig. 3. Maximum likelihood phylogram based upon the 3′-UTR data.

Table 3
RF distances to the reference tree, retention indices given the reference tree topology, and evolutionary rates of different partitions. When two maximum likelihood trees were found, the RF distance is an average of the two trees. The two introns in HMGN2 were analyzed together.

Region	RF distance to reference tree		Retention index		Evolutionary rate	
	3′-UTR	Intron	3′-UTR	Intron	3′-UTR	Intron
3′-UTR + Intron	8		0.73		–	
All four loci	6	14	0.72	0.73	–	–
ALDOB	13	15.5	0.79	0.76	0.46	0.77
CRYAA	11	14	0.70	0.80	0.49	0.60
PCBD1	13	17	0.67	0.68	0.79	0.79
HMGN2	13.5	11.5	0.77	0.69	0.37	1.36

Higher rates of evolution should result in greater numbers of variable sites, and hence a greater power to resolve phylogenies (Braun and Kimball, 2001). Although sequences exhibiting higher rates of evolution have the potential to be misleading (see Yang, 1998), the simulations conducted by Chojnowski et al. (2008) suggest that sequences evolving at a typical intronic rate are not problematic within birds. The 3′-UTRs we examined did evolve at a lower rate than the introns in the same locus (Table 3), with HMGN2 being the slowest. This is consistent with the observation that only the HMGN2 3′-UTR showed significant similarity to the human genome in BLASTn (Altschul et al., 1997) searches when the chicken 3′-UTR sequence was used as a query. Despite our expectation that the more slowly evolving 3′-UTRs would exhibit lower power that introns, the 3′-UTR tree showed higher congruence with the reference tree than the intron tree. Only for HMGN2 did the intron partition perform better than the 3′-UTR, possibly due to the very low rate of evolution of this 3′-UTR that may have limited its power to resolve phylogenies. We expect high rates of

evolution to increase homoplasy, potentially explaining the improved performance of the 3′-UTRs relative to the introns. However, the retention index was very similar between 3′-UTRs and introns (Table 3), suggesting similar levels of homoplasy in both types of data.

4. Discussion

The analysis of nuclear sequence data is advantageous because it allows collection of data from multiple unlinked loci, which is important to obtain an estimate of a species tree rather than an individual gene tree (Maddison, 1997). Within birds, many studies have used introns (e.g., Fain and Houde, 2004; Hackett et al., 2008) or nuclear coding regions (e.g., Groth and Barrowclough, 1999; Barker et al., 2004). A few avian studies have included 3′-UTRs (Chubb, 2004; Hackett et al., 2008; Harshman et al., 2008), but these studies did not explicitly compare the patterns of 3′-UTR molecular evolution to introns or explore the relative phylogenetic utility of 3′-UTRs and introns. In mammals, 3′-UTRs exhibit less homoplasy than coding exons (Murphy et al., 2001); in crocodilians the 3′-UTR of a single locus (MYC) exhibited less homoplasy than either the intron or a coding exon (Harshman et al., 2003). Both of these studies suggest that 3′-UTRs may hold great potential for limiting the effect of homoplasy in phylogenetic studies.

There are consistent differences between 3′-UTRs and introns with respect to the among-sites rate heterogeneity, since the shape parameter of the Γ distribution was always lower for the 3′-UTRs. This result is consistent with the existence of a larger number of sites subject to constraint in the 3′-UTRs. Likewise, the standard deviation of the window rate averages was consistently higher for the 3′-UTRs, indicating that there is spatial clustering of sites with similar evolutionary rates as we would expect if there are functional segments within the 3′-UTRs. In contrast, the introns have a more even distribution of changes across sites with lower levels of spatial clustering than occurs in the 3′-UTRs.

In the absence of functional constraints, both 3′-UTRs and introns should exhibit similar patterns of molecular evolution within a locus. The isochore structure typical of archosaurs would be expected to result in similar base compositions within a locus (Chojnowski et al., 2007), and this should be evident even if there are small regions of functional constraint within 3′-UTRs. Base composition also correlates with other evolutionary parameters (e.g., Webster et al., 2006), so we would expect the patterns of molecular evolution to be similar within loci for sites not directly subject to constraint. However, we found differences between the 3′-UTRs and introns for most loci. In some cases loci differed in base composition, while in other cases loci differed in the best-fitting model of evolution. In contrast to the loci that differed between the 3′-UTR and intron, ALDOB exhibited very similar patterns for both data types. Overall, our results suggest that the two types of data generally exhibited different patterns of evolution, though we could not identify consistent differences among the five loci we examined other than the greater spatial variation in the 3′-UTRs.

Phylogenies generated from the 3′-UTR, intron, and combined datasets are largely congruent among each other and with other galliform phylogenies based on a variety of data types (Crowe et al., 2006; Cox et al., 2007; Kaiser et al., 2007; Kimball and Braun, 2008; Kriegs et al., 2007). However, there are some differences between the 3′-UTR and intron phylogenies. For instance, the position of *C. japonica* has differed among studies (Crowe et al., 2006; Cox et al., 2007; Kimball and Braun, 2008; Kriegs et al., 2007; Kaiser et al., 2007). Other relationships have been controversial or poorly supported, such as the relationship between *Chrysolophus pictus* and *Lophura nycthemera* or that between *Argusianus argus* and the clade containing *Afropavo congensis* and *Pavo cristatus* (Crowe et al., 2006; Kimball et al., 2001; Kimball and Braun, 2008). These small differences suggest that 3′-UTRs, although exhibiting different patterns of molecular evolution than introns, generally recover the same phylogenetic relationships.

Our data suggest that 3′-UTRs typically have better phylogenetic signal and slightly lower homoplasy than nuclear introns. These results are consistent with observations made using a single locus in crocodilians (Harshman et al., 2003), but the increased number of loci examined here suggests a more general phenomenon. Many 3′-UTRs are long enough to provide a substantial number of variable sites, yet short enough to amplify readily (see also Whittall et al., 2006). Like many introns, where it is possible to anchor PCR primers in conserved regions (the "exon-primed, intron-crossing" or "EPIC" strategy; Palumbi and Baker, 1994), 3′-UTR forward primers can be anchored in the last coding exons of genes. Reverse primers for 3′-UTRs typically must be anchored in a conserved non-coding sequence such as the polyadenylation sequence. While the core polyadenylation signal is short and AT-rich, it is typically embedded in a longer conserved region that can be used for a reverse primer. However, only two 3′-UTRs (ALDOB and HMGN2) amplified in other avian orders, and on average the 3′-UTRs were more difficult to amplify than their associated introns in Galliformes, suggesting limited conservation of some priming sites. These results are similar to Whittall et al. (2006), where amplification success of 3′-UTRs decreased outside of the target genus. The difficulty associated with the amplification of 3′-UTRs from diverse organisms is likely to be the primary drawback to the use of 3′-UTRs in phylogenetic studies. Nonetheless, when robust primers are available (which can be established by conducting test PCRs in appropriately chosen target taxa) the desirable characteristics of 3′-UTRs indicate that they should be considered as a phylogenetic tool in vertebrates in addition to (and possibly even instead of) nuclear introns.

Acknowledgments

This manuscript was greatly improved by helpful suggestions from the members of the Kimball-Braun lab and two anonymous reviewers. Many thanks to the Louisiana State University Museum of Natural History for providing additional tissues samples for *Argusianus argus* (LSUMNS B-13314) and *Rollulus roulroul* (LSUMNS B-24971). This research was supported by a grant from the National Science Foundation (DEB-0228682) to R.T.K., E.L.B., and D.W. Steadman.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ympev.2010.04.006.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

American Ornithologists' Union, 1998. Check-list of North American Birds, seventh ed. American Ornithologists' Union, Washington, DC.

Armstrong, M.H., Braun, E.L., Kimball, R.T., 2001. Phylogenetic utility of avian ovomucoid intron G: a comparison of nuclear and mitochondrial phylogenies in the Galliformes. Auk 118, 799–804.

Barker, F.K., Cibois, A., Schikler, P., Feinstein, J., Cracraft, J., 2004. Phylogeny and diversification of the largest avian radiation. Proc. Natl. Acad. Sci. USA 101, 11040–11045.

Braun, E.L., Kimball, R.T., 2001. Polytomies, the power of phylogenetic inference, the stochastic nature of molecular evolution: a comment on Walsh et al. (1999). Evolution 55, 1261–1263.

Braun, E.L., Kimball, R.T., 2002. Examining basal avian divergences with mitochondrial sequences: model complexity, taxon sampling, and sequence length. Syst. Biol. 51, 614–625.

Chojnowski, J.L., Franklin, J., Katsu, Y., Iguchi, T., Guillette, L.J., Kimball, R.T., Braun, E.L., 2007. Patterns of vertebrate isochore evolution revealed by comparison of expressed mammalian, avian, and crocodilian genes. J. Mol. Evol. 65, 259–266.
Chojnowski, J.L., Kimball, R.T., Braun, E.L., 2008. Introns outperform exons in analyses of basal avian phylogeny using clathrin heavy chain genes. Gene 410, 89–96.
Chubb, A.L., 2004. New nuclear evidence for the oldest divergence among neognath birds: the phylogenetic utility of ZENK (i). Mol. Phylogenet. Evol. 30, 140–151.
Cooper, G.M., Brudno, M., Program, N.C.S., Green, E.D., Batzoglou, S., Sidow, A., 2003. Quantitative estimates of sequence divergence for comparative analyses of mammalian genomes. Genome Res. 13, 813–820.
Cox, W.A., Kimball, R.T., Braun, E.L., 2007. Phylogenetic position of the New World quail (Odontophoridae): eight nuclear loci and three mitochondrial regions contradict morphology and the Sibley–Ahlquist tapestry. Auk 124, 71–84.
Crowe, T.M., Bowie, R.C.K., Bloomer, P., Mandiwana, T.G., Hedderson, T.A.J., Randi, E., Pereira, S.L., Wakeling, J., 2006. Phylogenetics, biogeography and classification of, and character evolution in gamebirds (Aves: Galliformes): effects of character exclusion, data partitioning and missing data. Cladistics 22, 1–38.
Delsuc, F., Brinkmann, H., Philippe, H., 2005. Phylogenomics and the reconstruction of the tree of life. The evolution of isochores. Nat. Rev. Genet. 6, 361–375.
Delsuc, F., Phillips, M.J., Penny, D., 2003. Comment on "Hexapod origins: monophyletic or paraphyletic?". Science 301, 1482.
Fain, M.G., Houde, P., 2004. Parallel radiations in the primary clades of birds. Evolution 58, 2558–2573.
Groth, J.G., Barrowclough, G.F., 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol. Phylogenet. Evol. 12, 115–123.
Hackett, S.J., Kimball, R.T., Reddy, S., Bowie, R.C.K., Braun, E.L., Braun, M.J., Chojnowski, J.L., Cox, W.A., Han, K.-L., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., Yuri, T., 2008. A phylogenomic study of birds reveals their evolutionary history. Science 320, 1763–1768.
Harshman, J., Braun, E.L., Braun, M.J., Huddleston, C.J., Bowie, R.C.K., Chojnowski, J.L., Hackett, S.J., Han, K.-L., Kimball, R.T., Marks, B.D., Miglia, K.J., Moore, W.S., Reddy, S., Sheldon, F.H., Steadman, D.W., Steppan, S.J., Witt, C.C., Yuri, T., 2008. Phylogenomic evidence for multiple losses of flight in ratite birds. Proc. Natl. Acad. Sci. USA 105, 13462–13467.
Harshman, J., Huddleston, C.J., Bollback, J.P., Parsons, T.J., Braun, M.J., 2003. True and false gharials: a nuclear gene phylogeny of Crocodylia. Syst. Biol. 52, 386–402.
Kaiser, V.B., van Tuinen, M., Ellegren, H., 2007. Insertion events of *CR1* retrotransposable elements elucidate the phylogenetic branching order in galliform birds. Mol. Biol. Evol. 24, 338–347.
Kimball, R.T., Braun, E.L., 2008. A multigene phylogeny of Galliformes supports a single origin of erectile ability in non-feathered facial traits. J. Avian Biol. 39, 438–445.
Kimball, R.T., Braun, E.L., Bowie, R.C.K., Braun, M.J., Chojnowski, J.L., Hackett, S., Han, K.-L., Harshman, J., Heimer-Torres, V., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Reddy, S., Sheldon, F.H., Smith, J.V., Witt, C.C., Yuri, T., 2009. A set of resources to amplify nuclear regions across the avian genome. Mol. Phylogenet. Evol. 50, 654–660.
Kimball, R.T., Braun, E.L., Ligon, J.D., Lucchini, V., Randi, E., 2001. A molecular phylogeny of the peacock-pheasants (Galliformes: *Polyplectron* spp.) indicates loss and reduction of ornamental traits and display behaviours. Biol. J. Linn. Soc. 73, 187–198.
Kimball, R.T., Braun, E.L., Ligon, J.D., Randi, E., Lucchini, V., 2006. Using molecular phylogenetics to interpret evolutionary changes in morphology and behavior in the Phasianidae. Acta Zool. Sin. 52 (Suppl.), 362–365.
Kimball, R.T., Braun, E.L., Zwartjes, P.W., Crowe, T.M., Ligon, J.D., 1999. A molecular phylogeny of the pheasants and partridges suggests that these lineages are not monophyletic. Mol. Phylogenet. Evol. 11, 38–54.
Kriegs, J.O., Matzke, A.J.M., Churakov, G., Kuritzin, A., Mayr, G., Brosius, J., Schmitz, J., 2007. Waves of genomic hitchhikers shed light on the evolution of gamebirds (Aves: Galliformes). BMC Evol. Biol. 7, 190.
Kumar, S., 1996. Patterns of nucleotide substitution in mitochondrial protein coding genes of vertebrates. Genetics 143, 537–548.
Maddison, D., Maddison, W., 2000. MacClade 4: Analysis of Phylogeny and Character Evolution. Sinauer Associates, Inc., Sunderland, MA.
Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.
Matthee, C.A., Eick, G., Willows-Munro, S., Montgelard, C., Pardini, A.T., Robinson, T.J., 2007. Indel evolution of mammalian introns and the utility of non-coding nuclear markers in eutherian phylogenetics. Mol. Phylogenet. Evol. 42, 827–837.
Mazumder, B., Seshadri, V., Fox, P.L., 2003. Translational control by the 3′-UTR: the ends specify the means. TIBS 28, 91–98.
Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. Evolution 49, 718–726.
Murphy, W.J., Eizirik, E., Johnson, W.E., Zhang, Y.P., Ryder, O.A., O'Brien, S.J., 2001. Molecular phylogenetics and the origins of placental mammals. Nature 409, 614–618.
Palumbi, S.R., Baker, C.S., 1994. Contrasting population-structure from nuclear intron sequences and mtDNA of humpback whales. Mol. Biol. Evol. 11, 426–435.
Phillips, M.J., Penny, D., 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. Mol. Phylogenet. Evol. 28, 171–185.
Posada, D., Crandall, K.A., 1998. MODELTEST: testing the model of DNA substitution. Bioinformatics 14, 817–818.
Pratt, R.C., Gibb, G.C., Morgan-Richards, M., Phillips, M.J., Hendy, M.D., Penny, D., 2009. Toward resolving deep Neoaves phylogeny: data, signal enhancement, and priors. Mol. Biol. Evol. 26, 313–326.
Rigoutsos, I., Huynh, T., Miranda, K., Tsirigos, A., McHardy, A., Platt, D., 2006. Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. Proc. Natl. Acad. Sci. USA 103, 6605–6610.
Robinson, D.F., Foulds, L.R., 1981. Comparison of phylogenetic trees. Math. Biosci. 53, 131–147.
Sibley, C.G., Ahlquist, J.E., 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven.
Sorenson, M.D., Ast, J.C., Dimcheff, D.E., Yuri, T., Mindell, D.P., 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Mol. Phylogenet. Evol. 12, 105–114.
Swofford, D.L., 2003. PAUP*: Phylogenetic Analysis using Parsimony (* And Other Methods), version 4.0. Sinauer Associates, Inc., Sunderland, MA.
Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 25, 4876–4882.
Webster, M.T., Axelsson, E., Ellegren, H., 2006. Strong regional biases in nucleotide substitution in the chicken genome. Mol. Biol. Evol. 23, 1203–1216.
Whittall, J.B., Medina-Marino, A., Zimmer, E.A., Hodges, S.A., 2006. Generating single-copy nuclear gene data for a recent adaptive radiation. Mol. Phylogenet. Evol. 39, 124–134.
Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D., Sites, J.W., Townsend, T.M., Reeder, T.W., 2008. Branch length, support, and congruence: testing the phylogenomic approach with 20 nuclear loci in snakes. Syst. Biol. 57, 420–431.
Xie, X., Lu, J., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., Kellis, M., 2005. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 434, 338–345.
Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. Biosci. 13, 555–556.
Yang, Z., 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. 47, 125–133.
Zwickl, D.J., 2006. Genetic Algorithm Approaches for the Phylogenetic Analysis of Large Biological Sequence Datasets under the Maximum Likelihood Criterion. University of Texas, Austin.