Molecular Phylogenetics and Evolution 64 (2012) 342–356

Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

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

Multiple independent origins of mitochondrial control region duplications in the order Psittaciformes

Erin E. Schirtzinger [a,*], Erika S. Tavares [b,d], Lauren A. Gonzales [a], Jessica R. Eberhard [c], Cristina Y. Miyaki [d], Juan J. Sanchez [e], Alexis Hernandez [e], Heinrich Müeller [f], Gary R. Graves [g,h], Robert C. Fleischer [i], Timothy F. Wright [a]

[a] Department of Biology, New Mexico State University, Las Cruces, NM 88003, USA
[b] Department of Natural History, Royal Ontario Museum, 100 Queen's Park, Toronto, ON, Canada M5S 2C6
[c] Department of Biology and Museum of Natural Science, Louisiana State University, Baton Rouge, LA 70803, USA
[d] Departmento de Genética e Biologia Evolutiva, Universidade de São Paulo, São Paulo, SP, Brazil
[e] Instituto Nacional de Toxicologia y Ciencias Forenses, 38320 Tenerife, Canary Islands, Spain
[f] Department of Veterinary Medicine, Loro Parque Fundación, 3840 Puerto de la Cruz, Tenerife, Spain
[g] Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, PO Box 37012, MRC-116, Washington, DC 20013-7012, USA
[h] Center for Macroecology, Evolution and Climate, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark
[i] Center for Conservation and Evolutionary Genetics, Smithsonian Conservation Biology Institute, National Zoological Park, PO Box 37012, MRC 5503, Washington, DC 20013-7012, USA

ARTICLE INFO

Article history:
Received 2 May 2011
Revised 8 April 2012
Accepted 10 April 2012
Available online 20 April 2012

Keywords:
Ancestral state reconstruction
Control region
Control region duplication
Gene duplication
Mitochondrial genomes
Parrots

ABSTRACT

Mitochondrial genomes are generally thought to be under selection for compactness, due to their small size, consistent gene content, and a lack of introns or intergenic spacers. As more animal mitochondrial genomes are fully sequenced, rearrangements and partial duplications are being identified with increasing frequency, particularly in birds (Class Aves). In this study, we investigate the evolutionary history of mitochondrial control region states within the avian order Psittaciformes (parrots and cockatoos). To this aim, we reconstructed a comprehensive multi-locus phylogeny of parrots, used PCR of three diagnostic fragments to classify the mitochondrial control region state as single or duplicated, and mapped these states onto the phylogeny. We further sequenced 44 selected species to validate these inferences of control region state. Ancestral state reconstruction using a range of weighting schemes identified six independent origins of mitochondrial control region duplications within Psittaciformes. Analysis of sequence data showed that varying levels of mitochondrial gene and tRNA homology and degradation were present within a given clade exhibiting duplications. Levels of divergence between control regions within an individual varied from 0–10.9% with the differences occurring mainly between 51 and 225 nucleotides 3′ of the goose hairpin in domain I. Further investigations into the fates of duplicated mitochondrial genes, the potential costs and benefits of having a second control region, and the complex relationship between evolutionary rates, selection, and time since duplication are needed to fully explain these patterns in the mitochondrial genome.

1. Introduction

Conservation of genome size, consistent gene content, and a lack of introns or intergenic spacers in animal mitochondria are generally interpreted as evidence that mitochondrial genomes are under selection for small size (Brown et al., 1979; Quinn and Wilson, 1993; Rand, 1993). This selection regime suggests that gene duplications in the mitochondria should be very rare or quickly eliminated because smaller genomes can replicate more quickly (Attardi, 1985; Diaz et al., 2002; Gray, 1989; Rand, 2001; Selosse et al., 2001; Sogin, 1997). As more mitochondrial genomes are sequenced, however, duplications of mitochondrial genes have been identified with increasing frequency in diverse species such as birds, lizards, ostracods, fish, arthropods, and snakes (Abbott et al., 2005; Arndt and Smith, 1998; Bensch and Härlid, 2000; Black and Roehrdanz, 1998; Campbell and Barker, 1999; Desjardins and Morais, 1990; Eberhard et al., 2001; Gibb et al., 2007; Kumazawa et al., 1996, 1998; Lee and Kocher, 1995; Lee et al., 2001; Macey et al., 1997; Mindell et al., 1998; Moritz and Brown, 1987; Ogoh and Ohmiya, 2004, 2007; Quinn and Mindell, 1996; Shao and Barker, 2003; Shao et al., 2005). It is now clear that duplications do occur in the mitochondrial genome and are much more common than previously thought. However, understanding the underlying mechanisms, evolutionary dynamics, and fitness consequences of

* Corresponding author. Address: Department of Biology, New Mexico State University, PO Box 30001, MSC 3AF, Las Cruces, NM, USA. Fax: +1 575 646 4791.
E-mail address: eesem@nmsu.edu (E.E. Schirtzinger).

1055-7903/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ympev.2012.04.009

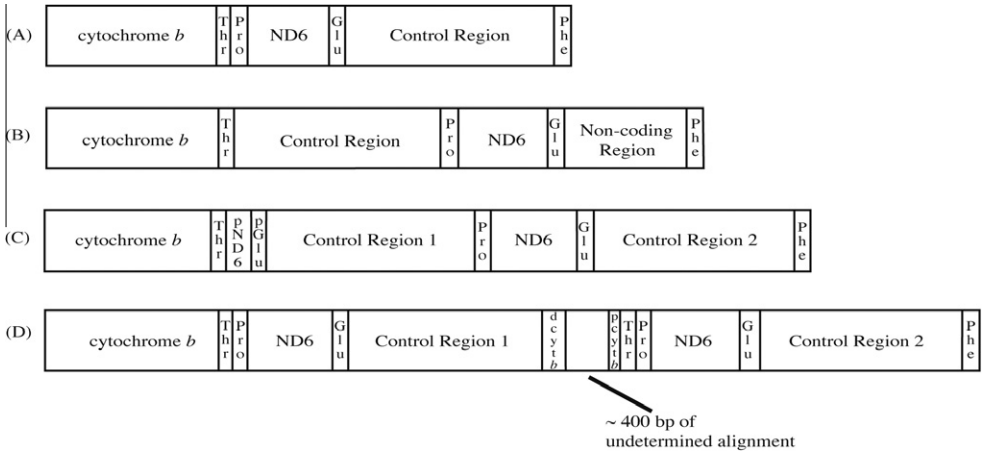

Fig. 1. (a) The typical avian mitochondrial gene order identified by Desjardins and Morais (1990). A tandem duplication followed by random loss of the control region and tRNAs can explain the rearrangement from the typical vertebrate gene order. (b) An alternative avian mitochondrial gene order identified by Mindell et al. (1998). A single tandem duplication followed by incomplete loss of the duplicated region may explain this rearrangement from the typical avian mitochondrial gene order. (c) The mitochondrial gene order identified by Eberhard et al. (2001) in *Amazona* parrots, in which two complete and putatively functional mitochondrial control regions appear to be maintained. (d) The mitochondrial gene order discovered by Abbott et al. (2005) in *Thalassarche* albatrosses in which cytochrome *b* to the control region is tandemly duplicated. Only the second copy of cytochrome *b* shows any degradation as depicted by the labels pcyt*b* and dcyt*b*, which are regions with high similarity to the full-length cytochrome *b* gene.

these duplications remains an ongoing challenge for the field of molecular evolution.

Mitochondrial duplications often occur as tandem arrays, with a gene or group of genes repeated one after the other (Campbell and Barker, 1999; Eberhard et al., 2001; Abbott et al., 2005). Several mechanisms have been proposed that would result in this type of structure. Slipped strand mispairing can frequently result in tandem duplications in the presence of repeat units or sequences that form secondary structures (Levinson and Gutman, 1987; Madsen et al., 1993; Mueller and Boore, 2005). During DNA replication, a portion of the DNA strand dissociates between two repeats forming a loop. The polymerase then reassociates at the first repeat and duplicates the looped section (Levinson and Gutman, 1987; Madsen et al., 1993; Boore, 2000). The consistent presence of repeats at either end of the junctions of mitochondrial tandem duplications in parthenogenetic lizards led Fujita et al. (2007) to conclude that slipped strand mispairing was likely the cause of duplications. Over-running the termination signal during DNA replication has also been suggested as a way to form tandem duplications (Boore, 2000; Mueller and Boore, 2005). In this case, the duplications would include the genes flanking the origin of replication, which is often seen in mitochondrial duplications (San Mauro et al., 2006). Initiation of replication at sites of secondary structure other than the origin has also been suggested to result in tandem duplications (Levinson and Gutman, 1987; Madsen et al., 1993; Stanton et al., 1994; Lunt and Hyman, 1997; Macey et al., 1997; Boore, 2000). tRNAs or other sequences that are capable of forming secondary structures often are found at the ends of duplicated regions, suggesting that these structures may cause illicit priming of mitochondrial replication (Stanton et al., 1994; San Mauro et al., 2006). Finally, unequal crossing over could potentially result in tandem duplications when two mitochondrial genomes within a single mitochondrion recombine with one genome donating its copies of a group of genes to the other genome (see Ohno, 1970; Zhang, 2003 for this mechanism in the nucleus).

Consideration of the fates of duplicated nuclear genes suggests four potential fates for duplicated mitochondrial genes: (1) nonfunctionalization in which one copy becomes a pseudogene and is eventually eliminated from the genome, (2) subfunctionalization, in which copies of a multifunctional gene can each become specialized for one of the different original functions, and will each be stably maintained within the genome because they are under selection to carry out different functions, (3) neofunctionalization, in which a duplicated gene acquires a novel function due to mutations in the regulatory region or within the gene copy, and (4) redundant maintenance, in which multiple copies of a gene are maintained through gene conversion or purifying selection because the extra copies help meet high expression demands (Force et al., 1999; Lynch et al., 2001; Rastogi and Liberles, 2005; Roth et al., 2007; Zhang, 2003). However as mitochondria are believed to be under selection for compactness, it would seem that nonfunctionalization and elimination of extra gene copies would be the most likely fate of mitochondrial duplications (Rand and Harrison, 1986). Depending upon which copy of a gene is eliminated, a new gene arrangement may arise or the original gene order may be restored (Boore, 2000). This hypothesis has come to be known as the tandem duplication/random loss model of mitochondrial genome rearrangement (Boore, 2000; Macey et al., 1997; Mindell et al., 1998; Mortiz, 1991). In this model, loss of duplicated genes is thought to occur rapidly relative to evolutionary time (Mortiz, 1991; Quinn, 1997). Therefore, residual evidence of a previous duplication such as the presence of pseudogenes may be suggestive of a relatively recent event (Mortiz, 1991; Quinn, 1997).

Four different mitochondrial genome arrangements have been identified within birds (Class Aves) that differ from that of the typical vertebrate (Fig. 1). The common avian arrangement, first identified in the chicken by Desjardins and Morais (1990), can be derived from the common vertebrate arrangement (ND6/tRNA^{Glu}/cyt *b*/tRNA^{Thr}/tRNA^{Pro}/control region) by one tandem duplication/random loss event (involving cyt *b*/tRNA^{Thr}/tRNA^{Pro}/ND6/tRNA^{Glu}/control region). Later, Mindell et al. (1998) described a second arrangement in which a non-coding region of variable length and with some similarity to the control region was found in the typical location of the control region, while the full-length control region was located after tRNA^{Thr}. A single tandem duplication/random loss event is necessary to derive this arrangement from the common avian mitochondrial arrangement (Mindell et al., 1998). This second gene order has been found in several diverse orders of birds, such as Piciformes (woodpeckers), Cuculiformes (cuckoos), Falconiformes (falcons), Passeriformes (oscines and suboscines), and Tinamiformes (tinamous) (Bensch and Härlid, 2000; Haddrath and Baker, 2001; Mindell et al., 1998). A third arrangement of mitochondrial genes was found in several species of *Amazona* parrots (Eberhard et al., 2001). In this case, one degenerate copy of the

duplicated ND6 and $tRNA^{Glu}$ was still present making the extent of the duplication more easily defined. Additionally, the second non-coding region showed high similarity with the control region and appeared to be functional. This arrangement has also been found in the osprey (*Pandion haliaetus*), ivory-billed aracari (*Pteroglossus azara*) and Philippine hornbills (*Aceros waldeni* and *Penelopides panini*) (Gibb et al., 2007; Sammler et al., 2011). The fourth arrangement was identified in *Thalassarche* albatrosses (Abbott et al., 2005). Here, the genes from cytochrome *b* to the control region were tandemly duplicated and most appeared to still be functional. However, the second copy of cytochrome *b* appeared greatly reduced in size with only portions of the 5′ and 3′ ends (designated as d-cyt *b* and p-cyt *b*) being alignable with the full-length copy. The two control regions were also easily alignable and appeared to be functional, but differed in sequence and length of domain III, with control region 1 lacking repeats at the 3′ end (Abbott et al., 2005). A similar rearrangement has also been found in the black-faced spoonbill (*Platalea minor*) (Cho et al., 2009), the ruff (*Philomachus pugnax*) (Verkuil et al., 2010), three species of booby in the genus *Sula* (Morris-Pocock et al., 2010) and two species of Philippine hornbills (Sammler et al., 2011).

Despite the many descriptions of avian mitochondrial gene arrangements that have been published, we still lack a clear understanding of when or how often mitochondrial duplications have occurred in birds. Few orders have been systematically surveyed for gene arrangements or have been paired with a well-sampled phylogeny to allow robust conclusions about the evolutionary history of mitochondrial duplications and genome rearrangements. The order Psittaciformes (parrots and cockatoos, hereafter 'parrots'), presents an excellent opportunity to identify the frequency with which mitochondrial control region duplications occur within a clade. Eberhard et al. (2001) established that several species of Neotropical parrots contained a duplicated control region, while preliminary data from other parrots suggested that these duplications were not shared by the entire order (T.F. Wright, J.R. Eberhard, unpublished data; E.S. Tavares, C.Y. Miyaki, unpublished data).

The current study seeks to address the following two questions: (1) Does the mitochondrial control region duplication, first identified in *Amazona* parrots, exist in other parrot genera? (2) If so, was there a single origin or were there multiple independent origins of these duplications? To answer these questions, we surveyed 117 parrot species by PCR for the presence of mitochondrial control region duplications and mapped these results onto a phylogeny reconstructed from mitochondrial and nuclear intron DNA sequences.

2. Materials and methods

2.1. Taxon and character sampling

For the phylogeny and survey of mitochondrial control region duplications, we added 51 new taxa to the dataset of Wright et al. (2008) for a total of 117 parrot species representing 79 of the 82 extant genera (Tables 1S and 2S). We used a stratified sampling method to determine the number of species sampled per genus such that genera with one to four species were represented by a single species or 25–100% coverage, genera with 5–11 species had two representatives (18–40% coverage), genera with 12–16 species were represented by three species (19–25% coverage) and genera with more than 17 species had four representatives (13–24% coverage). The new species included in this study were chosen based upon the accessibility of tissue or blood samples in museum or zoo collections. Samples for three genera (*Geopsittacus*, *Ognorhynchus* and *Oreopsittacus*) were unobtainable. Taxonomic nomenclature follows Forshaw (2006) for Old World species and the 2010 AOU North American and South American checklists for New World species (Chesser et al., 2010; Remsen et al., 2010). *Coccyzus americanus* (Cuculiformes), *Colius colius* (Coliiformes), *Columbina passerina* (Columbiformes), *Falco peregrinus* (Falconiformes), *Otis sunia* (Strigiformes), *Picus canus* (Piciformes), *Serinus canarius* (Passeriformes) and *Tockus flavirostris* (Coraciiformes) were included as outgroups as each has been identified as an ally or a sister group of the parrots in previous studies (Ericson et al., 2006; Fain and Houde, 2004; Hackett et al., 2008; Sibley and Ahlquist, 1990; Sorenson et al., 1999).

For the phylogenetic analyses, we sampled two mitochondrial protein-coding loci (cytochrome *c* oxidase I (COI) and nicatinamide adenosine dehydrogenase subunit 2 (ND2)) and two nuclear introns (tropomyosin intron five (TROP) and transforming growth factor beta 2 intron one (TGFB2)). These genes have proven to be informative in other phylogenetic studies of parrots (Wright et al., 2008; Joseph et al., 2012).

2.2. DNA extraction, PCR and sequencing

We extracted DNA from tissue or blood samples, performed polymerase chain reaction amplification (PCR), and sequenced the PCR products at laboratories in three locations due to legal restrictions on transporting specimens from endangered species. The laboratories were New Mexico State University (NMSU) in Las Cruces, New Mexico; the Instituto Nacional de Toxicologia y Ciencias Forenses (Spain) in Tenerife, Canary Islands, Spain; and the University of São Paulo (Brazil) in São Paulo, Brazil.

At NMSU and in Spain we extracted DNA using the DNeasy Blood and Tissue Kit (Qiagen, Valencia, CA) following the manufacturer's protocol for each tissue type. In Brazil, we extracted DNA from blood samples using a phenol/chloroform protocol (Bruford et al., 1992). At all locations we amplified the four gene regions by PCR using primers, reactions, and cycling conditions as described in Wright et al. (2008). PCR products were checked for correct size and the presence of multiple bands by electrophoresis on a 0.5–2% agarose gel and stained with ethidium bromide.

We cleaned PCR products using a QiaQuick PCR Purification Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions at NMSU and in Spain, while in Brazil we used 1 μL exonuclease I and 0.5 μL of shrimp alkaline phosphatase per 10 μL PCR reaction incubated at 37 °C for 30 min, then 80 °C for 15 min to clean PCR products. We sequenced each PCR product in both directions using the PCR primers and Big Dye v3.1 Terminator Cycle Sequencing chemistry (Applied Biosystems Inc, Foster City, CA). Each sequencing reaction at NMSU and in Spain consisted of 2 μL of BigDye, 1 μL of 5X sequencing buffer, 3.2 μL of 1uM primer, 2 μL of clean PCR product, and 11.8 μL of water. In Brazil each sequencing reaction consisted of 2 μL of BigDye, 2–4 μL of clean PCR product, and 1 μL of primer. Sequencing conditions at all locations were 25 cycles of 95 °C for 25 s, 50 °C for 5 s, and 60 °C for 4 min. We cleaned sequencing reactions at NSMU by centrifugation through Sephadex columns. Clean reactions were dried and resuspended in 20 μL of HiDi Formamide (Applied Biosystems Inc., Foster City, CA) before sequencing on an ABI 3100 Avant automated sequencer. In Brazil sequencing reactions were cleaned by isopropanol/ethanol precipitation, dried, resuspended in 1.8 μL of formamide, heated to 95 °C for 2 min, placed on ice until loaded on an ABI 377 automated sequencer. In Spain, we cleaned sequencing reactions by centrifugation through CentriSep columns in a 96 well format (Applied Biosystems Inc., Foster City, CA). The cleaned reaction was dried, resuspended in 20 μL of formamide, and sequenced on an ABI 310 automated sequencer.

2.3. Phylogenetic analysis

Raw sequences were checked for ambiguous base calls in Sequencher 4.7 (Gene Codes, Ann Arbor, MI) and combined into

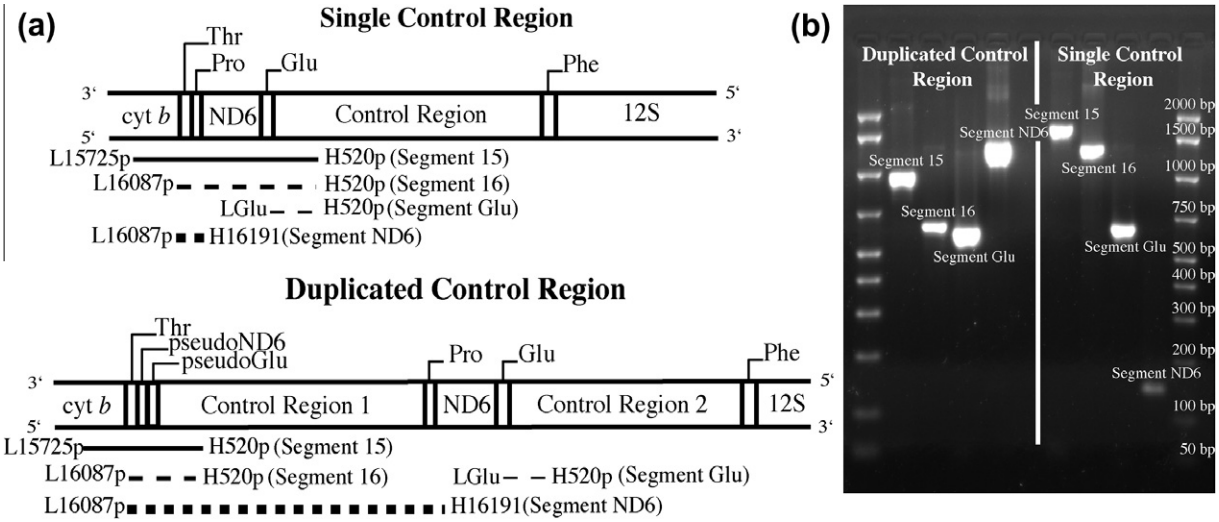

Fig. 2. Location of primer pairs and relative fragment lengths used to classify control region duplication state. (a) Schematic of fragment sizes based upon the presence of a single or duplicated mitochondrial control region. (b) A representative agarose gel of fragment sizes for a duplicated and single control region.

contigs by locus and taxon. Sequences were aligned using Clustal W with default parameters as implemented at http://www.ebi.ac.uk/Tools/clustalw and adjusted by eye. Gaps within introns were coded by the simple indel coding method (Simmons and Ochoterena, 2000) as implemented in IndelCoder 0.5 in the SeqState 1.40 program (Müller, 2005, 2006).

Maximum likelihood methods of phylogenetic reconstruction were implemented in GARLI v0.951 using the default settings (Zwickl, 2006). The General Time Reversible model with a gamma distribution of among site heterogeneity and a proportion of invariant sites was chosen, with parameter values estimated by the software and SPR branch swapping (Zwickl, 2006). To ensure that the tree was not located in a local optimum, 20 independent runs were conducted in GARLI and the tree with the highest likelihood was chosen for subsequent analyses (Zwickl, 2006). Nodal support was evaluated by 100 maximum likelihood bootstraps calculated in GARLI.

Bayesian methods of phylogeny reconstruction were implemented using MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). A separate evolutionary model was determined for each gene region in MrModelTest 2.3 under the Akaike Information Criterion (Nylander, 2004). Gaps were coded as restriction sites using the default settings. The analysis consisted of two parallel runs, each with one cold chain and three heated chains with default parameters. The mixed model analysis of the combined dataset was run for 15,000,000 generations with trees sampled every 1000 generations and a burn-in of 25%. Convergence was assumed when the average standard deviation of split frequencies was less than or equal to 0.01 and when the effective sample size (ESS) value for parameter values was greater than 200 when viewed in Tracer v1.4 (Rambaut and Drummond, 2007).

2.4. Mitochondrial control region survey

Of the 117 species included in the phylogeny 112 were surveyed at NMSU for the presence of a mitochondrial control region duplication using PCR to amplify diagnostic fragments from the regions predicted to differ in length depending on the presence or absence of a duplicated control region. We could not survey five taxa using the PCR approach (*Strigops habroptilus*, *Cyanoramphus auriceps*, *Cyanopsitta spixii*, *Enicognathus leptorhynchus* and *Triclaria malachitacea*) because tissue samples were not available at NMSU. We used the primer pairs L15725p – AAACCAGARTGATAYTTYC TMTTYGCAT (modified from Sorenson et al., 1999) and H520p – TGKSCCTGACCKAGGAACCAG (modified from Sorenson et al., 1999), L16087p – TGGYCTTGTAARCCAAARRAYGAAG (modified from Sorenson et al., 1999) and H520p and L16087p with H16191 – TCTCGDGGGGCDATTCGGGC (Sorenson et al., 1999) to amplify diagnostic Segments 15, 16 and ND6 respectively. An additional primer pair, LGlu – GCCCTGAAAARCCATCGTTG (Eberhard et al., 2001) in conjunction with H520p was used to amplify Segment Glu to verify the presence of a second control region that is usually flanked by the intact $tRNA^{Glu}$. The expected location of each primer and the genes contained within each segment are depicted in Fig. 2. PCR reaction conditions were the same as for the phylogeny except for the cycling conditions which were 94 °C for 4 min followed by 29 cycles of 94 °C for 25 s, 60 °C for 30 s on the initial cycle with a decrease of 0.4 °C each cycle, and 72 °C for 2 min. This was followed by 6 cycles of 94 °C for 25 s, 45 °C for 30, 72 °C for 2 min, with a final extension of 72 °C for 10 min. PCR products were electrophoresed on a 1.5% agarose gel and stained with ethidium bromide. Gels were imaged on a BioRad Gel Doc XR and band sizes for each segment were estimated using the BioRad Amplisize 50–2000 bp Molecular Ruler and the Band Analysis protocol in the Quantity One software (BioRad Life Sciences, Hercules, CA).

Expected sizes of the diagnostic segments used to score each species as a single or duplicated control region were as follows: Single Control Region, Segment 15 = 1579–1896 base pair (bp), Segment 16 = 1225–1542 bp, Segment ND6 = 132 bp; Duplicated Control Region, Segment 15 = 800–1250 bp, Segment 16 = 600–1000 bp, Segment ND6 $\geqslant$ 1200 bp. These segment sizes were derived from preliminary studies of parrot control region duplications and then further refined by analysis of the GenBank parrot mitochondrial sequences (Table 3S and T.F. Wright, J.R. Eberhard, E.E. Schirtzinger, unpublished data). Band sizes were expected to show some variation due to the variation in size of domains I and III of the control region (Baker and Marshall, 1997) as determined by alignment of primer H520p to parrot control region sequences available on GenBank. Another source of variation is the presence and size of intergenic spacers. This variation was evaluated by counting the base pairs between the 3′ end of one annotated gene and the 5′ end of the next annotated gene in the region of cytochrome *b* to $tRNA^{Phe}$ from three parrot mitochondrial genomes on GenBank (*Strigops habroptilus*, *Agapornis roseicollis* and *Melopsittacus undulatus*). These ranges were added to the sizes of the genes included in each segment to get the total estimated range of variation. A species was scored as having a single or duplicated control region based on the correspondence of its measured

segment sizes to the expected sizes for each segment (Fig. 1S). Taxa that did not amplify at least two diagnostic segments were classified as unscorable.

2.5. Sequencing of selected taxa

Because the expected band sizes encompass a large range, diagnostic Segment 16 and Segment Glu, from selected species were sequenced at NMSU or Brazil to confirm the status of the mitochondrial control region as classified by our PCR survey. At least one representative of each clade that contained an inferred duplicated control region was sequenced. In addition, species that were ambiguous in their classification were also sequenced. Finally, the GenBank mitochondrial sequences for *Agapornis roseicollis*, *Strigops habroptilus* (single control regions) and *Melopsittacus undulatus*, *Amazona farinosa*, *Amazona ochrocephala* and *Psittacus erithacus* (duplicated control regions) were also used as confirmation of the PCR survey results.

At NMSU, PCR products were amplified, cleaned as described above, and sent to the University of Chicago Cancer Sequencing Facility for sequencing on an ABI 3730 automated sequencer using Big Dye chemistry. In Brazil, amplifications for sequencing were performed in 10 μL reactions with 1X buffer (GE Healthcare or Biotools), 2 μM of dNTP, 1 μM of each primer, 0.5 U of *Taq* polymerase, and 20–50 nanograms (ng) of template DNA, or in 25 μL reactions with 1X buffer (Biotools), 2 μM of dNTP, 1 μM of each primer, 1 U of *Taq* polymerase, and 25–50 ng of DNA. PCR conditions were: initial denaturation 96 °C for 5 min, 30 cycles of 95 °C for 60 s, 50–54 °C for 25 s, and 65 °C for 40–80 s, with a final extension of 65 °C for 5 min. The size and quality of PCR products were verified, purified as described above or bands were excised from agarose gels and the product was isolated by centrifugation through filter tips (Axigen). Sequencing reactions were prepared, cleaned, and run as described above.

Raw sequences were proofread as previously described and combined into a consensus sequence by taxon and segment using Sequencher 4.7 (Gene Codes, Ann Arbor, MI). The location of $tRNA^{Thr}$, $tRNA^{Pro}$, ND6, and $tRNA^{Glu}$ were identified by comparison with the homologous genes from the mitochondrial genome of *Melopsittacus undulatus* (NC_009134), while the control region was identified by the presence of the goose hairpin (C_7TAC_7) near the 5′ end. The identity of pseudogenes was based on similarity with known sequences from the *Melopsittacus undulatus* mitochondrial genome or comparison with the pseudogenes defined by Eberhard et al. (2001). Functionality of tRNAs was assessed by simulation in tRNA scan-SE (Lowe and Eddy, 1997).

For each species sequenced, the gene order from $tRNA^{Thr}$ through domain I of the control region was identified by similarity with previously described avian gene orders (Abbott et al., 2005; Desjardins and Morais, 1990; Eberhard et al., 2001; Mindell et al., 1998). For those species with a duplication, we measured the length of the non-coding region (calculated as the number of nucleotides from the end of $tRNA^{Thr}$ to the goose hairpin), and the number of nucleotide differences between the two control regions (calculated as the number of differences when the two control region fragments were aligned divided by the total length of the aligned control region segment). Because the 5′ end of the control region does not have a definitive starting motif and the 3′ end of the $tRNA^{Glu}$ could not be identified for all species, the goose hairpin was used as a proxy for the beginning of the control region.

2.6. Ancestral state reconstruction

Using the classifications from the PCR survey, sequences from selected taxa and GenBank sequences, the presence of a single (0) or duplicated (1) control region was coded into a matrix and mapped onto the phylogeny. Three species (*Aprosmictus erythropterus*, *Eos reticulata*, *Pionites melanocephala*) were coded as unscorable because they either failed to amplify the diagnostic fragments by PCR or they produced ambiguous results and no confirmatory sequence was available. Ancestral state reconstructions were undertaken in Mesquite 2.01 (Maddison and Maddison, 2007) under parsimony (Fitch) and the maximum likelihood criterion. Likelihood reconstructions were undertaken using the symmetric MK1 model, which is a generalization of the Jukes–Cantor model with equal probability of changing states and the AsymmMK model, which uses different rates of change between states (Maddison and Maddison, 2007). To test the robustness of the likelihood analysis, separate reconstructions were undertaken using the AsymmMK model and rates translating to (a) gains five times as likely as losses and (b) gains 1/5 as likely as losses.

3. Results

3.1. Phylogenetic analysis

The phylogenetic dataset consisted of 117 parrot species and eight non-parrot outgroups. The mitochondrial sequences, COI and ND2, showed no insertions/deletions (indels). The COI sequences were 570 base pairs (bp) in length while the ND2 sequences were 1041 bp. The intron sequences, TROP and TGFB2, were more variable in length due to the presence of 48 and 78 indels respectively. The TROP sequences were 498–533 bp with a total of 554 aligned bp. The TGFB2 sequences were 611–630 bp with a total of 817 aligned bp. The maximum likelihood dataset consisted of 2982 concatenated bp. The Bayesian analysis was performed on the 2982 characters included in the concatenated dataset (partitioned by gene region) and the 126 coded indels for a total of 3108 characters. Table 4S includes the genetic details for each gene region and the concatenated dataset. All new sequences have been deposited in GenBank (see Table 1S).

The most likely tree ($-\ln L = -63846.569$) from 20 independent maximum likelihood (ML) analyses in GARLI of the concatenated dataset (Fig. 2S) and the consensus tree from the Bayesian analysis (Fig. 3) were broadly congruent, differing only in the placement of two genera, *Micropsitta* and *Graydidascalus*. The Bayesian analysis places *Micropsitta* with high support (posterior probability = 1) as sister to the clade consisting of *Alisterus*, *Aprosmictus*, *Polytelis*, *Eclectus*, *Geoffroyus*, *Psittacula*, *Psittinus*, *Tanygnathus* and *Prioniturus* while in the ML analysis *Micropsitta* is sister to the clade of *Alisterus*, *Aprosmictus* and *Polytelis* with poor support (ML bootstrap $\leqslant 50$). The Bayesian analysis places *Graydidascalus* as sister to *Amazona* while in the ML analysis *Graydidascalus* is sister to the clade consisting of *Amazona* and *Pionus*. However, in both analyses the position of *Graydidascalus* is poorly supported. Nucleotide substitution models, priors for the Bayesian analysis, and final estimates of parameters for each analysis are listed in Table 5S.

3.2. Mitochondrial control region survey

One hundred and twelve parrot species were surveyed for the status of its mitochondrial control region by PCR of three diagnostic segments that show variation in size when a duplicated control region is present. Table 1 reports the control region status of each species surveyed and the length of each amplicon. Segment 15 was amplifiable for 96 species, while Segment 16 and Segment ND6 were amplifiable for 108 and 110 species respectively (see Table 1). *Aprosmictus erythropterus*, *Guarouba guaroouba*, *Nandayus nenday* and *Rhynchopsitta pachyrhyncha* could not be scored by PCR at NMSU due to a lack of amplification. Five species (*Eos reticulata*, *Graydidascalus brachyurus*, *Hapalopsittaca pyrrhops*, *Pionites mela-*

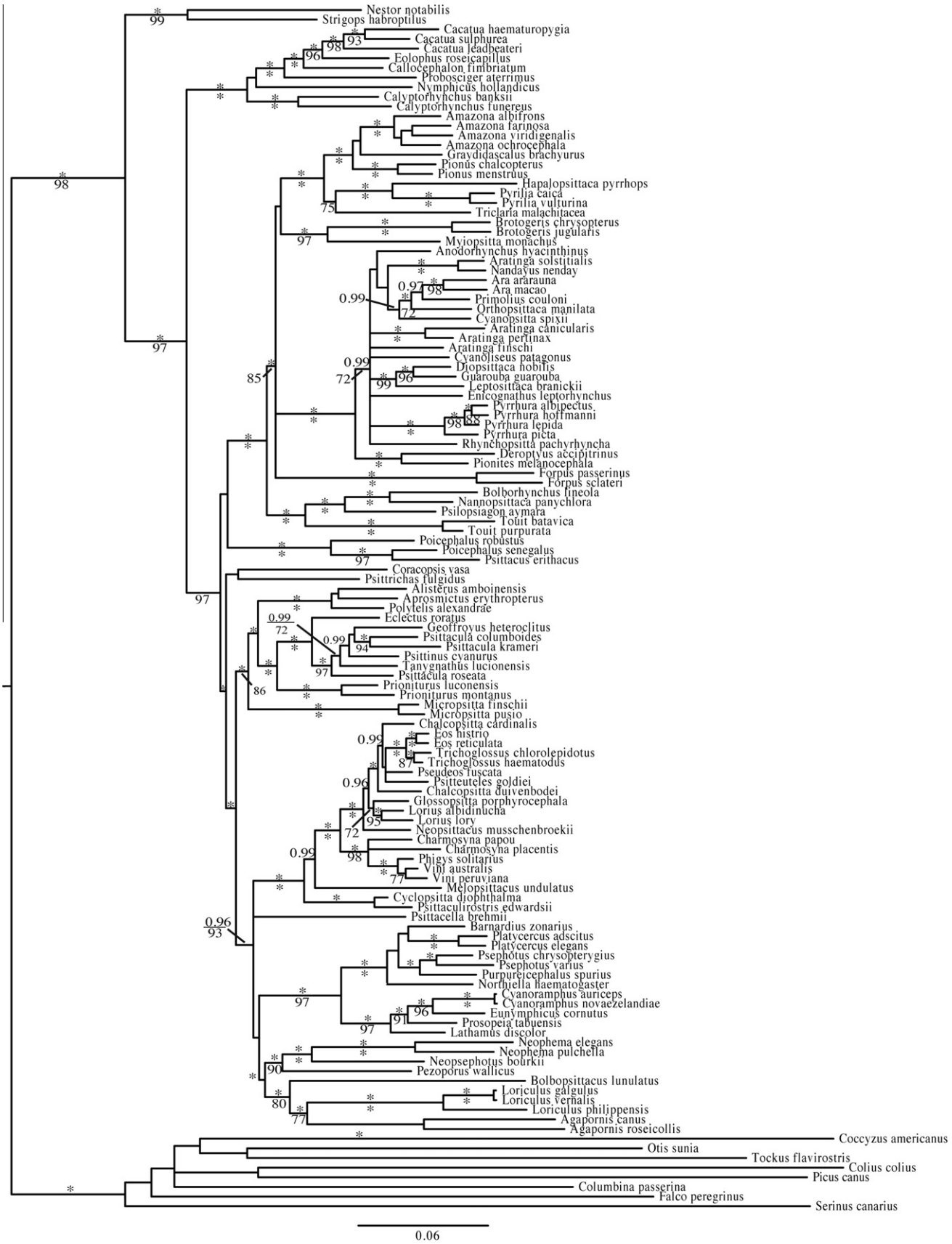

Fig. 3. The partitioned Bayesian analysis of 117 parrot species reconstructed using two mitochondrial protein-coding genes (COI and ND2), two nuclear introns (TROP and TGFB2) and coded gaps. Posterior probabilities >0.95 are given above the branches and maximum likelihood bootstraps >70 are given below the branches. Asterisks indicate a posterior probability of 1.0 or a maximum likelihood bootstrap value of 100.

Table 1
Measured PCR fragment lengths and control region status for each parrot species in the phylogeny.

		Measured fragment size (nucleotides)			
Scored by PCR	Control region status	15	16	Glu	ND6
Agapornis canus	1	1781	1394	708	141
Agapornis roseicollis	1	1912	1399	689	144
Alisterus amboinensis	1	1652	1253	488	141
Amazona albifrons	2	1042	756	609	501
Amazona farinosa	2	1134	769	618	137, 1943
Amazona ochrocephala	2	1109	779	637	1883
Amazona viridigenalis	2	1182	781	626	153, 508, 2146
Anodorhynchus hyacinthinus	1	1672	1316	645	148
Aprosmictus erythropterus	UN	*	*	*	144
Ara ararauna	1	1742	1344	627	147
Ara macao	1	1738	1302	685	152
Aratinga canicularis	1	1833	1357	648	152
Aratinga finschi	1	1658	1341	630	144
Aratinga pertinax	1	2121	1697	640	145
Aratinga solstitialis	1	*	1368	627	145
Barnardius zonarius	1	1963	1576	847	149
Bolbopsittacus lunulatus	1	*	1681	901	153
Bolborhynchus lineola	1	1580	1255	543	149
Brotogeris chrysopterus	1	1809	1376	667	149
Brotogeris jugularis	1	1717	1343	677	145
Cacatua haematuropygia	1	1627	1236	640	148
Cacatua leadbeateri	1	1722	1245	629	149
Cacatua sulphurea	1	*	1227	604	144
Callocephalon fimbriatum	1	1505	1198	606	148
Calyptorhynchus banksii	1	1543	1215	549	148
Calyptorhynchus funereus	1	1765	1320	553	150
Chalcopsitta cardinalis	2	990	656	602	1346
Chalcopsitta duivenbodei	2	943	602	680	144, 1387
Charmosyna papou	2	*	1049	625	141, 315, 1164
Charmosyna placentis	2	998	628	625	1383
Coracopsis vasa	1	1572	1237	562	143
Cyanoliseus patagonus	1	1805	1368	671	147
Cyanoramphus novaezelandiae	1	2047	1626	875	144
Cyclopsitta diophthalma	2	*	673	636	475
Deroptyus accipitrinus	2	1261	952	642	147, 2244
Diopsittaca nobilis	1	*	1294	631	155
Eclectus roratus	1	1658	1253	561	153
Eolophus roseicapillus	1	1621	1187	*	152
Eos histrio	2	1010	622	573	1465
Eos reticulata	2?	*	627	575	145
Eunymphicus cornutus	1	1928	1583	861	150
Forpus passerinus	2	1200	810	670	138, 1813
Forpus sclateri	2	1234	650	672	2980
Geoffroyus heteroclitus	1	*	1498	781	152
Glossopsitta porphyrocephala	2	1088	720	687	1440
Graydidascalus brachyurus	2?	1410	889	700	154
Guarouba guarouba	UN	*	*	*	*
Hapalopsittaca pyrrhops	2?	1321	786	663	143
Lathamus discolor	1	1962	1577	840	147
Leptosittaca branickii	1	1726	1296	620	146
Loriculus galgulus	1	1571	1269	631	146
Loriculus philippensis	1	1598	1279	640	147
Loriculus vernalis	1	1636	1295	628	143
Lorius albidinuchas	2	998	640	618	1350
Lorius lory	2	984	645	580	1472
Melopsittacus undulatus	2	968	610	585	1612
Micropsitta finschii	1	1555	1180	502	149
Micropsitta pusio	1	1505	1142	479	142
Myiopsitta monachus	1	1639	1259	596	152
Nandayus nenday	UN	*	*	*	*
Nannopsittaca panychlora	1	1620	1254	556	151
Neophema elegans	1	*	1379	693	164
Neophema pulchella	1	1673	1288	640	160
Neopsephotus bourkii	1	1657	1313	628	154
Neopsittacus musschenbroekii	2	1004	640	584	1423
Nestor notabilis	1	1615	1264	596	156
Northiella haematogaster	1	1914	1322	592	147
Nymphicus hollandicus	1	1633	1200	591	146
Orthopsittaca manilata	1	1727	1363	640	152
Pezoporus wallicus	1	1900	1500	800	155
Phigys solitarius	2	1028	646	550	155, 1524
Pionites melanocephala	2?	*	971	626	154
Pyrilia caica	2?	1136	782	638	158

Table 1 (*continued*)

Scored by PCR	Control region status	Measured fragment size (nucleotides)			
		15	16	Glu	ND6
Pyrilia vulturina	2	982	615	600	166, 2223
Pionus chalcopterus	2	1150	781	614	157, 525, 2128
Pionus menstruus	2	1167	793	640	490
Platycercus adscitus	1	1919	1521	809	149
Platycercus elegans	1	1829	1449	771	148
Poicephalus robustus	2	1058	666	560	153, 1521, 1804
Poicephalus senegalus	2	978	639	531	1549
Polytelis alexandrae	1		1265	635	145
Primolius couloni	1	1685	1313	630	147
Prioniturus luconensis	2	1231	761	500	145, 2328
Prioniturus montanus	2	1130	795	492	148
Probosciger aterrimus	1	1601	1244	594	144
Prosopeia tabuensis	1	1878	1465	758	148
Psephotus chrysopterygius	1	1872	1494	858	138
Psephotus varius	1	1924	1517	837	148
Pseudeos fuscata	2	988	619	578	1370
Psilopsiagon aymara	1	1555	1210	546	151
Psittacella brehmii	1	1640	1476	604	150
Psittacula columboides	1	*	1211	550	150
Psittacula krameri	1	*	1343	582	136
Psittacula roseata	1	1503	1198	492	144
Psittaculirostris edwardsii	2	817	620	586	391, 680
Psittacus erithacus	2	984	617	545	144, 1589
Psitteuteles goldiei	2	1341	638	599	1455
Psittinus cyanurus	1	1554	1192	486	132
Psittrichas fulgidus	2?	1244	907	535	135
Purpureicephalus spurius	1	1893	1563	849	142
Pyrrhura albipectus	1	1712	1331	626	145
Pyrrhura hoffmanni	1	1563	1297	600	142
Pyrrhura lepida	1	1665	1306	640	136
Pyrrhura picta	1	1724	1310	622	154
Rhynchopsitta pachyrhyncha	UN	*	*	638	137
Tanygnathus lucionensis	1	1585	1262	550	133
Touit batavica	1	1700	1200	500	137
Touit purpurata	1	1700	1100	500	148
Trichoglossus chlorolepidotus	2	842	664	*	143, 175, 772, 1501
Trichoglossus haematodus	2	1009	643	586	1491
Vini australis	2	1000	650	600	142. 1348
Vini peruviana	2	1000	650	600	1336
	N=112	N=96	N=108	N=107	N=110
Scored from Sequences (BR)					
Cyanopsitta spixii	1				
Enicognathus leptorhynchus	1				
Guaruba guarouba	1				
Nandayus nenday	1				
Rhynchopsitta pachyrhyncha	1				
Triclaria malachitacea	2				
	N=6				
Scored from GB Sequences or Literature					
Cyanoramphus auriceps	1	Boon 2000			
Strigops habroptilus	1	AY309456			
	N=2				

* = fragment did not amplify.
UN = unscorable by PCR.
? = ambiguous by PCR.

nocephala and *Psittrichas fulgidus*) all amplified ND6 bands indicative of a single control region while the sizes of Segment 15 and Segment 16 suggested that a duplicated control region was present. These species are listed as ambiguous in Table 1.

The band sizes for each segment are plotted in Fig. 4. Segment 15 showed a bimodal distribution, with most species clearly falling into either the single or duplicated control region size categories. However, *Gradydidascalus brachyurus* fell between the expected ranges. Segments Glu and 16 from this species were subsequently sequenced. In four taxa, Segment 15 amplicons were larger than the expected size of 1896 bp for a single control region. *Lathamus discolor* and *Barnardius zonarius* were approximately 75 bp larger than expected, while *Aratinga pertinax* and *Cyanoramphus novaezelandiae* exceeded the single control region size range by 225 bp and 150 bp. These sizes may be stochastic effects of slightly different gel conditions that affect the distance run by the size marker used to measure the bands. The other bands for these species were also larger than expected but within the range of single control region sizes.

The size distribution for Segment 16 amplicons was also bimodal with all taxa falling within the two expected range sizes. Segment ND6 showed two different patterns: a tight cluster of species from 130 to 160 bp and a scatter of species with segment sizes from 1150 bp to over 2200 bp. *Pionus menstruus* and *Amazona albifrons* had values of 450 bp and 550 bp respectively. Based upon the other band sizes, *Amazona albifrons* and *Pionus menstruus* were classified as having a duplicated control region. The presence of a duplicated control region in *Amazona albifrons* was later confirmed by sequencing.

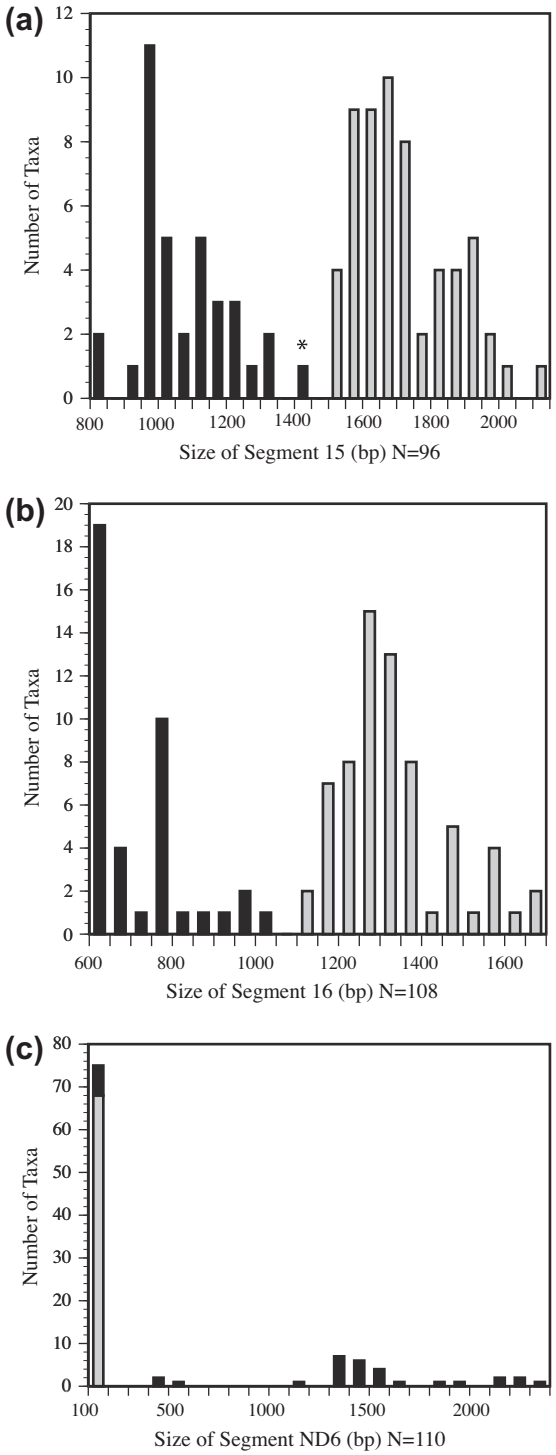

Fig. 4. Histograms of the number of species per 50 base pair bins for each amplified segment. Black bars indicate duplicate control regions. Gray bars indicate a single control region. An asterisk indicates species that fall outside of the expected fragment size ranges based on preliminary surveys of control region lengths. (a) lengths of Segment 15, (b) lengths of Segment 16, and (c) lengths of Segment ND6.

Of the 112 species surveyed by PCR, 68 were classified as having a single control region and 35 species were classified as having a duplicated control region. Four species, *Aprosmictus erythropterus*, *Guarouba guarouba*, *Nandayus nenday* and *Rhynchopsitta pachyrhyncha* were unscorable due to a lack of PCR amplification and five species (*Eos reticulata*, *Graydidascalus brachyurus*, *Hapalopsittaca pyrrhops*, *Pionites melanocephala* and *Psittrichas fulgidus*) produced ambiguous results.

3.3. Sequencing of selected taxa

In order to validate the control region classifications from the PCR survey we sequenced 44 parrot species in the phylogeny (Table 2). The common avian gene order with a single control region was confirmed in 19 species (Desjardins and Morais, 1990), while the gene order with a duplicated control region previously described by Eberhard et al. (2001) for *Amazona* and *Pionus* species was confirmed for 25 species. Of the five species that produced ambiguous results in the PCR survey, two species (*Graydidascalus brachyurus*, *Pseudoes fuscata*) were shown to have a duplicated control region by sequencing. The aberrant size of the ND6 fragment was due to a lack of retained homology in the non-coding region. Therefore, only the functional ND6 was amplified. Sequencing of *Cyanopsitta spixii*, *Enicognathus leptorhynchus*, *Guarouba guarouba*, *Nandayus nenday*, *Rhynchopsitta pachyrhyncha* and *Triclaria malachitacea* in Brazil found that only *Triclaria malachitacea* had a duplicated control region. Three species were listed as unscorable (*Aprosmictus erythropterus*, no PCR amplification) or ambiguous (*Eos reticulata* and *Pionites melanocephala*, no sequence available to confirm status).

In the set of species with duplicated control regions, non-coding regions of various sizes were found between $tRNA^{Thr}$ and the first control region. These non-coding regions were examined with tRNA-Scan (Lowe and Eddy, 1997) to determine if copies of $tRNA^{Pro}$ or $tRNA^{Glu}$ retained enough homology to be identifiable. In most cases no homology could be determined due to the extent of degeneration. The control regions were identified by the presence of conserved sequences: the goose hairpin at the 5′ end of domain I and the D-Box in domain II (Eberhard et al., 2001). The F-Box sequence of *Gallus gallus* was not identified in all species. An additional eight species (*Aratinga aurea*, *Aratinga leucophthalmus*, *Brotogeris chirri*, *Forpus xanthopterygius*, *Nannopsittaca dachillae*, *Pyrilia barrabandi*, *Pionites leucogaster*, and *Primolius auricollis*) were sequenced in Brazil for a different study (E. Tavares, C. Miyaki unpublished data), and served as additional confirmation of the distribution of control region duplications. Although these species were not included in the phylogeny, all are known from a broader phylogenetic survey to cluster with their congeners included in this study (E.E. Schirtzinger, unpublished data). GenBank sequences or published literature was used to confirm gene order for *Cyanoramphus auriceps*, *Strigops habroptilus*, *Amazona ochrocephala*, *Amazona farinosa*, *Pionus chalcopterus*, *Agapornis roseicollis*, *Melopsittacus undulatus*, and *Psittacus erithacus* (see Table 2 and Boon, 2000). The control region status of *Chalcopsitta duivenbodei* and *Charmosyna papou* could not be confirmed due to poor sequencing reactions.

Within the Neotropical parrots, a variation of the *Amazona* gene order was observed in *Deroptyus accipitrinus* and *Pionites leucogaster*. In contrast to all of the other species with control region duplications that we sequenced, *Deroptyus accipitrinus* and *Pionites leucogaster* retain a potentially functional copy of $tRNA^{Pro}$ before the non-coding region 5′ to the first control region.

3.4. Ancestral state reconstruction

To determine if mitochondrial control region duplications in parrots originated multiple times, the character states of single (0) or duplicated control region (1) were mapped onto the Bayesian tree using parsimony and maximum likelihood methods (Fig. 5). These states were assigned using our PCR classifications (68 single control region, 35 duplicated control region, three unscorable/ambiguous), sequences from selected taxa (six single control region and three duplicated control region), GenBank sequences (one single control region), or publications (one single control region). Both methods of reconstruction identified the ancestral con-

Table 2
Confirmation of PCR Classified Control Region Status by Sequencing of Selected Species or by GenBank Sequences.

Scored by PCR	Control region status	Confirmed by Sequence source	GenBank #
Agapornis rosiecollis	1	GB	EU410486
Amazona albifrons	2	NMSU	JQ341164, JQ360543
Amazona farinosa	2	GB	AF228821
Amazona ochrocephala	2	NMSU/GB	AF338819, AF338820, JQ341165, JQ360544
Anodorhynchus hyacinthinus	1	BR	EF104124
Ara ararauna	1	BR	EF104127
Aratinga solstitialis	1	BR	EF104138
Bolborhynchus lineola	1	BR	EF104137
Calyptorhynchus banksii	1	NMSU	JQ360567
Chalcopsitta cardinalis	2	NMSU	JQ341170, JQ360549
Chalcopsitta duivenbodei	2	NMSU	JQ360545
Charmosyna papou	2	NMSU	JQ341166
Charmosyna placentis	2	NMSU	JQ341167, JQ360546
Coracopsis vasa	1	NMSU	JQ341168, JQ360570
Cyanoliseus patagonus	1	BR	EF104136
Cyanopsitta spixii	1	BR	EF104128
Cyclopsitta diophthalma	2	NMSU	JQ241169, JQ36-547
Deroptyus accipitrinus	2	BR/NMSU	AF365437, JQ360548
Diopsittaca nobilis	1	BR	EF104121
Enicognathus leptorhynchus	1	BR	EF104139
Eos histrio	2	NMSU	JQ341171, JQ360550
Forpus sclateri	2	NMSU	JQ341172, JQ360551
Glossopsitta porphyrocephala	2	NMSU	JQ341173, JQ360552
Graydidascalus brachyurus	2	BR	EF104148
Guaruba guarouba	1	BR	EF104123
Lorius albidinucha	2	NMSU	JQ341174, JQ360553
Melopsittacus undulatus	2	GB	NC_009134
Micropsitta pusio	1	NMSU	JQ360568
Myiopsitta monachus	1	BR	EF104118
Nandayus nenday	1	BR	EF104131, EF104149
Neopsittacus musschenbroekii	2	NMSU	JQ341175, JQ360554
Orthopsittaca manilata	1	BR	EF104119
Phigys solitarius	2	NMSU	JQ341176, JQ360555
Pyrilia caica	2	NMSU	JQ341178, JQ360556
Pionus chalcopterus	2	NMSU/GB	AF338817, AF338818, JQ360557
Poicephalus robustus	2	NMSU	JQ360558
Prioniturus montanus	2	NMSU	JQ341180, JQ360566
Pseudeos fuscata	2	NMSU	JQ341183, JQ360561
Psittacula roseata	1	NMSU	JQ360569
Psittaculirostris edwardsii	2	NMSU	JQ341181, JQ360559
Psittacus erithacus	2	NMSU/GB	DQ335468, JQ341182, JQ360560
Psitteuteles goldiei	1	NMSU	JQ341184, JQ360562
Pyrrhura picta	1	BR	EF104130, EF104150
Rhynchopsitta pachyrhyncha	1	BR	EF104135
Trichoglossus haematodus	2	NMSU	JQ341186, JQ360564
Triclaria malachitacea	2	BR	EF104143, EF104146
Vini australis	2	NMSU	JQ341187, JQ360565
Related Species Sequenced			
Aratinga aurea	1	BR	EF104132,EF104151
Aratinga leucophthalmus	1	BR	EF104133
Brotogeris chirri	1	BR	EF104117
Primolius auricollis	1	BR	EF104126
Pionites leucogaster	2	BR	JQ749718, JQ749719
Pyrilia barrabandi	2	BR/NMSU	JQ341177, EF104141, EF104144
Forpus xanttopterygius	2	BR	EF104140, EF104147
Nannopsittaca dachillae	1	BR	EF104134
Psittaculirostris salvadorii	2	NMSU	JQ341185, JQ360563

trol region state in parrots as a single control region, and indicated that control region duplications have originated at least six times (Clades A–F in Fig. 5). No reversions from a duplicated control region to a single control region state were reconstructed by either method. Likelihood reconstructions using the symmetric rate (MK1) model and the asymmetric rate (AsymmMK) model did not affect the number of reconstructed independent origins or result in considerable differences in likelihoods of states at interior nodes. Similarly, changing the transition rate between states in the AsymmMK model did not affect our conclusions (See Fig. 3S and Table 6S for proportional likelihood values of each reconstructed state for interior nodes for each model analyzed).

3.5. Comparison of control regions

To determine if either of the two control regions had degenerated in species with a duplicated control region, the two control region fragments were aligned and nucleotide differences were calculated. These alignments found that the two control region sequences were typically highly similar, with sequence divergences ranging from 0–10.9% between the two copies within an individual. Most nucleotide differences were found between 51 and 225 nucleotides from the goose hairpin (Fig. 6). In domain I the only conserved sequence of known function is the termination-associated sequence (Baker and Marshall, 1997; Quinn, 1997; Sbisa et al., 1997). *Graydidascalus brachyurus* was not included in these calculations because neither of its control regions contained a goose hairpin.

4. Discussion

We reconstructed the phylogenetic relationships of 117 parrot species and classified their mitochondrial control region state from PCR fragment length analysis, DNA sequences or GenBank accessions to investigate the origins and distribution of mitochondrial control region duplications within the order Psittaciformes. A total of 76 parrot species were determined to have a single control region, while 38 parrot species were determined to have a duplicated control region. One species was unscorable and two species produced ambiguous results for which no sequence confirmation was available. Mapping the control region states onto the resulting phylogeny identified at least six independent origins of the duplicated control region state. Below we discuss the implications of

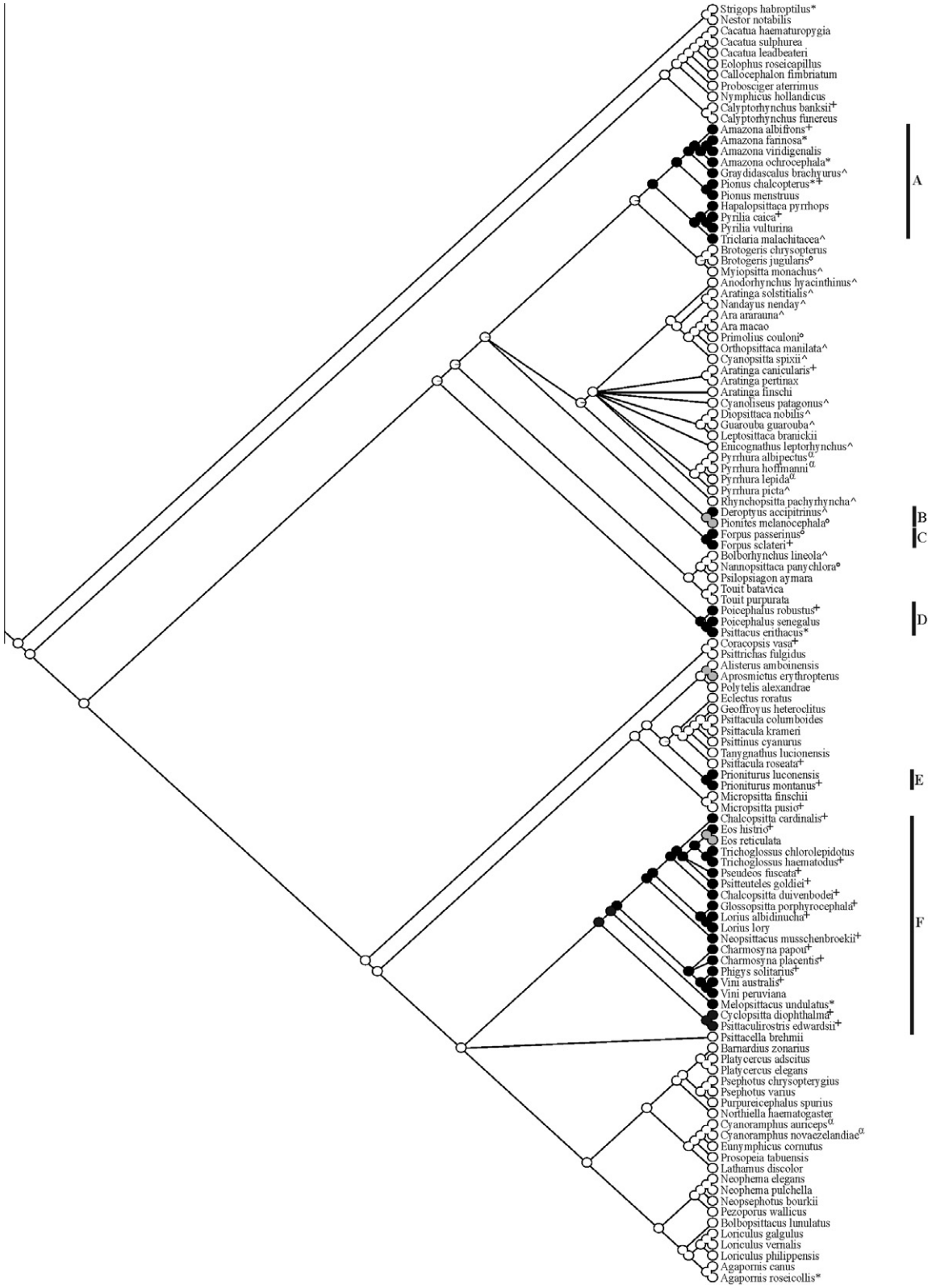

Fig. 5. Ancestral state reconstruction of the parrot mitochondrial control region duplications on the Bayesian tree under the maximum likelihood criterion using the MK1 model. White circles indicate species classified as having a single mitochondrial control region by the PCR fragment length analysis and/or sequencing. Black circles indicate species classified as having a duplicated mitochondrial control region. Gray circles indicate an unscorable state at the terminals and ambiguous ancestral states at interior nodes. The circles at nodes toward the interior of the tree are representative of the likelihood of each state at that node. ∗ = GenBank sequences, ^ = species sequenced at BR, ° = related species sequenced at BR, + = species sequenced at NMSU and α = species that have published gene orders. Letters to the right indicate clades defined by a single origin event.

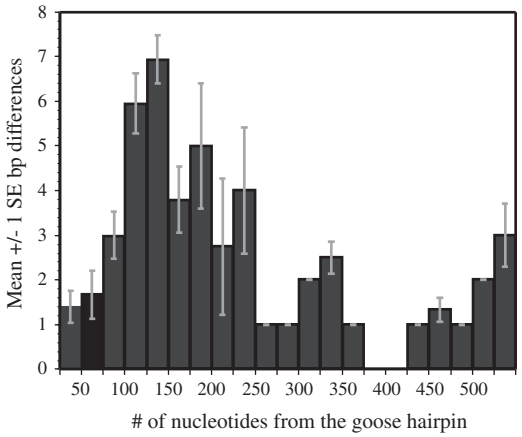

Fig. 6. Means (±1 standard error) for nucleotide differences in 25 bp non-overlapping windows of aligned duplicated control regions. The figure shows that the greatest number of nucleotide differences occurs 51–225 bp from the goose hairpin in domain I of the control region.

these results for parrot evolutionary relationships and present two alternative hypotheses for the evolution of mtDNA duplications in parrots.

4.1. Parrot phylogeny

The phylogeny reconstructed in this study is the most taxonomically comprehensive phylogeny to date of the Psittaciformes. The Bayesian tree and the maximum likelihood tree are generally well resolved and largely congruent in topology. Disagreement between the two analyses occurs solely on the location of the genus *Micropsitta* and *Graydidascalus brachyurus* (Fig. 2S and Fig. 3). The Bayesian phylogeny is broadly consistent with other published studies of parrot genus level relationships, in which a clade composed of the New Zealand endemics, *Strigops* and *Nestor*, was the sister group to all other parrots, and the Cacatuoidea (cockatoos) was the second oldest extant clade (Tavares et al., 2006; Wright et al., 2008; Schweizer et al., 2010; Joseph et al., 2012). Other well-supported parrot clades consistently recovered across various studies include the Neotropical parrots (Arinae), the African Psittacinae, the Australasian Psittaculidae and the Platycercinae from Australia, New Zealand, Oceania and Africa (de Kloet and de Kloet, 2005; Juniper and Parr, 1998; Tavares et al., 2006; Wright et al., 2008; Schweizer et al., 2010; Joseph et al., 2012). In agreement with previous studies by Tavares et al. (2006) and Wright et al. (2008), three Neotropical clades were recovered here: the parrotlets, including *Bolborhynchus*, *Nannopsittaca*, *Touit*, and *Psilopsiagon*; amazons and allies, including *Amazona*, *Pionus*, *Pyrilia*, *Triclaria* and *Graydidascalus*; and macaws and allies, including *Ara*, *Cyanopsitta*, *Aratinga*, *Orthopsittaca*, *Pyrrhura*, *Pionites* and *Anodorhynchus*.

4.2. Mitochondrial gene order in parrots

Among the 114 parrot species for which data was available there is evidence for two of the four described avian mitochondrial gene orders. The typical avian mitochondrial gene order originally described by Desjardins and Morais (1990) was inferred for 76 species by PCR, sequencing and examination of GenBank sequences. In contrast, duplicate control regions and the gene order described in *Amazona* parrots by Eberhard et al. (2001) was found in 38 of the species surveyed and/or sequenced. In this genome arrangement, a non-coding region, that in some species has apparent similarity to ND6 and $tRNA^{Glu}$, is located between $tRNA^{Thr}$ and the first control region. Two species, *Deroptyus accipitrinus* and *Pionites leucogaster*, which was not in the phylogeny, were shown by sequencing to have a variant of the *Amazona* gene order, with a functional $tRNA^{Pro}$ located between $tRNA^{Thr}$ and the first control region.

4.3. Evolutionary patterns of mtDNA control region duplications

Three striking patterns are apparent in our reconstructions of the evolution of control region duplications and sequencing of duplicate control regions in parrots. First, the mapping of control region duplication states onto the phylogeny of parrots identifies at least six independent origins of the control region duplications. Second, there were no reversions to a single control region state in any of these six clades. Third, although there was considerable interspecific variation, levels of sequence similarity between duplicated control regions within an individual were typically high (89–100% similarity in the species examined), at least for the first 400–500 nucleotides of domain I that was examined. This similarity would appear to be unusual due to the fact that this segment of the control region is often very different between species with many small insertions, deletions and mutations that are thought to occur as a result of the D-loop being single-stranded and accessible to mutagenic agents, such as reactive oxygen species (Shokolenko et al., 2007).

One hypothesis to explain these patterns is that a duplication of the control region and neighboring sequences occurred in the ancestors of these six clades and was retained in all descendent species of each clade. This hypothesis begs the question of what maintains the generally high degree of sequence similarity between the duplicated control regions within each taxon. The within-individual divergences that we observed between control region copies of 0–10.9% (Table 3) fall within the range of divergences observed in other taxa with duplicated control regions such as *Amazona* parrots, albatrosses, killifish, snakes, ticks, and ostracods (Abbott et al., 2005; Campbell and Barker, 1999; Eberhard et al., 2001; Kumazawa et al., 1996, 1998; Lee et al., 2001; Ogoh and Ohmiya, 2007; Tatarenkov and Avise, 2007). A high degree of sequence similarity is often interpreted as evidence for the maintenance of function in both duplicated control regions. Portions of the control region are under selection for the ability to bind with nuclear-encoded replication factors, and for functional control of replication and transcription (Doda et al., 1981; Gensler et al., 2001; He et al., 2007; Lee and Clayton, 1998; Schultz et al., 1998; Shadel and Clayton, 1997); such functionality may provide stabilizing selection on duplicated control regions. Alternatively, several studies of organisms with a duplicated control region have explained the high degree of similarity between the two control regions as evidence of gene conversion (Eberhard et al., 2001; Kumazawa et al., 1996, 1998; Ogoh and Ohmiya, 2007, Tatarenkov and Avise, 2007; Verkuil et al., 2010), but the molecular mechanisms responsible remain unclear. In either case, the maintenance of duplicated control regions concurrent with the elimination of duplicated mitochondrial genes and tRNAs suggests an advantage to having a second control region that overrides selection for compactness. Potential advantages include faster replication (Kumazawa et al., 1996) or protection against age-related deterioration of mitochondrial function (T.F. Wright and J.R. Eberhard, unpublished data).

An alternative hypothesis for the observed patterns is that a propensity for duplications to occur such as through replication slippage due to secondary structure in the control region was present in the common ancestor of each of the six clades, leading to repeated duplications of the region from cytochrome *b* to the control region with lineage specific degradation and deletion events passing to each of the descendent taxa (Verkuil et al., 2010; Zhuang and Cheng, 2010). Zhuang and Cheng (2010) found a similar pattern within Notothenioid fish and suggested that within each clade with control region duplications if these mutations are neutral, each descendant

Table 3
Gene order, non-coding region length and percent control region differences for parrot species with a duplicated control region.

Taxa	Source	Clade[a]	Gene order	Non-coding region length[b]	%CR differences[c]
Amazona albifrons	NMSU	A	Amazona	141	0.6
Amazona farinosa	AF338821	A	Amazona	158	0
Amazona ochrocephala	NMSU	A	Amazona	127	2.1
Pyrilia caica	NMSU	A	Amazona	155	0
Pyrilia barrabandi	BR/NMSU	A	Amazona	160	6
Pionus chalcopterus	NMSU/AF338817-18	A	Amazona	180	0.2
Triclaria malachitacea	BR	A	Amazona	112	2.5
Deroptyus accipitrinus	BR	B	Variant A	319	4.7
Pionites leucogaster	BR	B	Variant A	325	10.9
Forpus xanthopterygius	BR	C	Amazona	129	0.2
Forpus sclateri	NMSU	C	Amazona	117	5
Poicephalus robustus	NMSU	D	Amazona	129	7.7
Psittacus erithacus	NMSU	D	Amazona	88	10.8
Prioniturus montanus	NMSU	E	Amazona	360	1.9
Chalcopsitta cardinalis	NMSU	F	Amazona	84	7.9
Charmosyna placentis	NMSU	F	Amazona	45	2.7
Eos histrio	NMSU	F	Amazona	84	2.3
Glossopsitta porphyrocephala	NMSU	F	Amazona	68	3.3
Lorius albidinucha	NMSU	F	Amazona	73	3.3
Melopsittacus undulatus	NC009134	F	Amazona	58	4.6
Neopsittacus musschenbroekii	NMSU	F	Amazona	88	3.3
Phigys solitarius	NMSU	F	Amazona	32	2.7
Psitteuteles goldiei	NMSU	F	Amazona	81	2.6
Pseudeos fuscata	NMSU	F	Amazona	71	3.5
Trichoglossus haematodus	NMSU	F	Amazona	86	3.3
Vini australis	NMSU	F	Amazona	58	3.5
Cyclopsitta diophthalma	NMSU	F	Amazona	62	5.1
Psittaculirostris edwardsii	NMSU	F	Amazona	60	5.5

[a] Clades labeled in Fig. 5.
[b] Non-coding region length was calculated as the number of nucleotides from the 3′ end of $tRNA^{THR}$ to the 5′ end of the goose hairpin.
[c] Percent control region differences was calculated as the number of nucleotides that differed between the 2 aligned control regions divided by the total number of overlapping nucleotides multiplied by 100. Differences were only counted if they occurred 3′ of the goose hairpin.

species of a specific ancestor should have approximately the same amount of divergence between its two control regions. Table 3 shows that while species within the Australasian clade (labeled F in Fig. 5) exhibit similar levels of divergence between their duplicated control regions, the other clades show considerable variation among member species in the degree of divergence exhibited between their two control regions. Further investigation into patterns of sequence divergence over the entire length of the duplicated control region, coupled with functional studies of mitochondrial replication in species with and without control region duplications, should help distinguish between these alternative hypotheses for the evolution of duplicate control regions.

5. Conclusions

The presence of multiple mitochondrial gene orders within Psittaciformes supports the idea that the avian mitochondrial genome is a dynamic molecule. This study has shown that mitochondrial control region duplications have occurred many times in parrots, with ancestral state reconstructions suggesting six independent origins of the duplicated control region state and no reversions to a single control region state. Further investigations into the fates of duplicated mitochondrial genes, the potential costs and advantages of having a second control region, and the complex relationship between evolutionary rates, selection and time since duplication are needed to fully explain these patterns in the mitochondrial genome.

Acknowledgments

We thank the following institutions and people for the generous loan of tissue specimens: American Museum of Natural History (P. Sweet), Academy of Natural Sciences of Philadelphia (L. Joseph), the African Safari Zoo, the Australian National Wildlife Collection (L. Joseph), the Field Museum of Natural History (D. Willard), Fundação Crax, University of Kansas Museum of Natural History (M. Robbins), Loro Parque Fundación, Louisiana State University Museum of Natural Science (R. Brumfield), Museu Paraense Emílio Goeldi, Parque Ecológico do Tietê, the San Diego Zoological Park (N. Lamberski), Universidade Estadual Paulista (UNESP), the US National Museum of Natural History, the New Mexico State University Vertebrate Museum (P. Houde), Zoológico de Sorocaba, Zoológico de Americana, Zoológico Cyro-Gevaerd and several breeders. We particularly thank David Waugh, Sara Cappelli and Julia Scharpegge of the Loro Parque Fundación and Tania Matsumoto for generous assistance in obtaining data from critical specimens. We thank Rogério Lourenço for providing some of the sequences, Sergio L. Pereira for suggestions, and Elizabeth Moseman for many useful comments on the manuscript. Research support was provided by NIH grant S06 GM008136 (TFW), Sigma Xi Grant-in-Aid-of-Research (EES), Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (EST), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (EST) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (EST). CYM has a CNPq research productivity fellowship. This study was approved by the Institutional Animal Care and Use Committee of New Mexico State University.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2012.04.009.

References

Abbott, C.L., Double, M.C., Trueman, J.W.H., Robinson, A., Cockburn, A., 2005. An unusual source of apparent mitochondrial heteroplasmy: duplicate

mitochondrial control regions in *Thalassarche* albatrosses. Mol. Ecol. 14, 3605–3613.
Arndt, A., Smith, M.J., 1998. Mitochondrial gene rearrangement in the sea cucumber genus *Cucumaria*. Mol. Biol. Evol. 15, 1009–1016.
Attardi, G., 1985. Animal mitochondrial DNA: an extreme example of genetic economy. Int. Rev. Cyt. 93, 93–141.
Baker, A.J., Marshall, H.D., 1997. Mitochondrial control region sequences as tools for understanding evolution. In: Mindell, D.P. (Ed.), Avian Molecular Evolution and Systematics. Academic Press, San Diego, pp. 51–83.
Bensch, S., Härlid, A., 2000. Mitochondrial genomic rearrangements in songbirds. Mol. Biol. Evol. 17, 107–113.
Black, W.C., Roehrdanz, R.L., 1998. Mitochondrial gene order is not conserved in arthropods: prostriate and metastriate tick mitochondrial genomes. Mol. Biol. Evol. 15, 1772–1785.
Boon, W.M., 2000. Molecular systematics and conservation of the *Cyanoramphus* parakeet complex and the evolution of parrots. Ph.D. dissertation. Victoria University of Wellington, Wellington, Australia.
Boore, J.L., 2000. The duplication/random loss model for gene rearrangement exemplified by mitochondrial genomes of deuterostome animals. In: Sankoff, D., Nadeau, J.H. (Eds.), Comparative Genomes: Empirical and Analytical Approaches to Gene Order Dynamics, Map Alignment and the Evolution of Gene Families. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 133–148.
Brown, W.M., George Jr., M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. Proc. Natl. Acad. Sci. USA 76, 1967–1971.
Bruford, M., Hanotte, O., Brookfield, J.Y., Burke, T., 1992. Single-locus and multi-locus DNA fingerprinting. In: Hoelzel, A.R. (Ed.), Molecular Genetic Analysis of Populations: A Practical Approach. IRL Press, Oxford, UK, pp. 225–269.
Campbell, N.J.H., Barker, S.C., 1999. The novel mitochondrial gene arrangement of the cattle tick, *Boophilus microplus*: fivefold tandem repetition of a coding region. Mol. Biol. Evol. 16, 732–740.
Chesser, R.T., Banks, R.C., Barker, F.K., Cicero, C., Dunn, J.L., Kratter, A.W., Lovette, I.J., Rasmussen, P.C., Remsen Jr., J.V., Rising, J.D., Stotz, D.F., Winker, K., 2010. Fifty-first supplement to the American Ornithologists' Union check-list of North American Birds. The Auk 127, 726–744.
Cho, H-.J., Eda, M., Nishida, S., Yasukochi, Y., Chong, J-.R., Koike, H., 2009. Tandem duplication of mitochondrial DNA in the black-faced spoonbill, *Platalea minor*. Genes Genet. Syst. 84, 297–305.
de Kloet, R.S., de Kloet, S.R., 2005. The evolution of the spindlin gene in birds: sequence analysis of an intron of the spindlin W and Z gene reveals four major divisions of the Psittaciformes. Mol. Phylogenet. Evol. 36, 706–721.
Desjardins, P., Morais, R., 1990. Sequence and gene organization of the chicken mitochondrial genome: a novel gene order in higher vertebrates. J. Mol. Biol. 212, 599–634.
Diaz, F., Bayona-Bafaluy, M.P., Rana, M., Mora, M., Hao, H., Moraes, C.T., 2002. Human mitochondrial DNA with large deletions repopulates organelles faster than full-length genomes under relaxed copy number control. Nucleic Acids Res. 30, 4626–4633.
Doda, J.N., Wright, C.T., Clayton, D.A., 1981. Elongation of displacement-loop strands in human and mouse mitochondrial DNA is arrested near specific template sequences. Proc. Natl. Acad. Sci. USA 78, 6116–6120.
Eberhard, J.R., Wright, T.F., Bermingham, E., 2001. Duplication and concerted evolution of the mitochondrial control region in the parrot genus *Amazona*. Mol. Biol. Evol. 18, 1330–1342.
Ericson, P.G.P., Anderson, C.L., Britton, T., Elzanowski, A., Johansson, U.S., Källersjö, M., Ohlson, J.I., Parsons, T.J., Zuccon, D., Mayr, G., 2006. Diversification of Neoaves: integration of molecular sequence data and fossils. Biol. Lett. 2, 543–547.
Fain, M., Houde, P., 2004. Parallel radiations in the primary clades of birds. Evolution 58, 2558–2573.
Force, A., Lynch, M., Pickett, F.B., Amores, A., Yan, Y-.L., Postlethwait, J., 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics 151, 1531–1545.
Forshaw, J.M., 2006. Parrots of the World: An Identification Guide. Princeton University Press, Princeton, NY.
Fujita, M.K., Boore, J.L., Moritz, C., 2007. Multiple origins and rapid evolution of duplicated mitochondrial genes in parthenogenetic geckos (*Heteronotia binoei*; *Squamata*, *Gekkonidae*). Mol. Biol. Evol. 24, 2775–2786.
Gensler, S., Weber, K., Schmitt, W.E., Perez-Martos, A., Enriquez, J.A., Montoya, J., Wiesner, R.J., 2001. Mechanism of mammalian mitochondrial DNA replication: import of mitochondrial transcription factor A into isolated mitochondria stimulates 7S DNA synthesis. Nucleic Acids Res. 29, 3657–3663.
Gibb, G.C., Kardailsky, O., Kimball, R.T., Braun, E.L., Penny, D., 2007. Mitochondrial genomes and avian phylogeny: complex characters and resolvability without explosive radiations. Mol. Biol. Evol. 24, 269–280.
Gray, M.W., 1989. Origin and evolution of mitochondrial DNA. Annu. Rev. Cell Biol. 5, 25–50.
Hackett, S., 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.
Haddrath, O., Baker, A.J., 2001. Complete mitochondrial DNA genome sequences of extinct birds: ratite phylogenetics and the vicariance biogeography hypothesis. Proc. R. Soc. Lond. B 268, 939–945.
He, J., Mao, C-.C., Reyes, A., Sembongi, H., Di Re, M., Granycome, C., Clippingdale, A.B., Fearnley, I.M., Harbour, M., Robinson, A.J., Reichelt, S., Spelbrink, J.N., Walker, Holt, I.J., 2007. The AAA$^+$ protein ATAD3 has displacement loop binding properties and is involved in mitochondrial nucleoid organization. J. Cell Biol. 176, 141–146.
Joseph, J., Toon, A., Schirtzinger, E.E., Wright, T.F., Schodde, R., 2012. A revised nomenclature and classification for family-group taxa of parrots (Psittaciformes). Zootaxa 3205, 26–40.
Juniper, T., Parr, M., 1998. Parrots: A Guide to Parrots of the World. Yale University Press, New Haven, CT.
Kumazawa, Y., Ota, H., Nishida, M., Ozawa, T., 1996. Gene rearrangements in snake mitochondrial genomes: highly concerted evolution of control-region-like sequences duplicated and inserted into a tRNA gene cluster. Mol. Biol. Evol. 13, 1242–1254.
Kumazawa, Y., Ota, H., Nishida, M., Ozawa, T., 1998. The complete nucleotide sequence of a snake (*Dinodon semicarinatus*) mitochondrial genome with two identical control regions. Genetics 150, 313–329.
Lee, D.U., Clayton, D.A., 1998. Initiation of mitochondrial DNA replication by transcription and R-loop processing. J. Biol. Chem. 273, 30614–30621.
Lee, J.S., Miya, M., Lee, Y.S., Kim, C.G., Park, E.H., Aoki, Y., Nishida, M., 2001. The complete DNA sequence of the mitochondrial genome of the self-fertilizing fish *Rivulus marmoratus* (Cyprinodontiformes, Rivulidae) and the first description of duplication of a control region in fish. Gene 280, 1–7.
Lee, W.J., Kocher, T.D., 1995. Complete sequence of a sea lamprey (*Petromyzon marinus*) mitochondrial genome: early establishment of the vertebrate genome organization. Genetics 139, 873–887.
Levinson, G., Gutman, G.A., 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Mol. Biol. Evol. 4 (3), 203–221.
Lowe, T.M., Eddy, S.R., 1997. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. 25, 955–964.
Lunt, D.H., Hyman, B.C., 1997. Animal mitochondrial DNA recombination. Nature 387, 247.
Lynch, M., O'Hely, M., Walsh, B., Force, A., 2001. The probability of preservation of a newly arisen gene duplicate. Genetics 159, 1789–1804.
Macey, J.R., Larson, A., Ananjeva, N.B., Fang, Z., Papenfuss, T.J., 1997. Two novel gene orders and the role of light-strand replication in rearrangement of the vertebrate mitochondrial genome. Mol. Biol. Evol. 14, 91–104.
Maddison, W.P., Maddison, D.R., 2007. Mesquite: a modular system for evolutionary analysis. Version 2.01. <http://www.mesquiteproject.org>.
Madsen, C.S., Ghivizzani, S.C., Hauswirth, W.W., 1993. Protein-binding to a single termination-associated sequence in the mitochondrial DNA D-loop region. Mol. Cell. Biol. 13, 2162–2171.
Mindell, D.P., Sorenson, M.D., Dimcheff, D.E., 1998. Multiple independent origins of mitochondrial gene order in birds. Proc. Natl. Acad. Sci. USA 95, 10693–10697.
Morris-Pocock, J.A., Taylor, S.A., Birt, T.P., Friesen, V.L., 2010. Concerted evolution of duplicated mitochondrial control regions in three related seabird species. BMC Evol. Biol. 10, 14.
Mortiz, C., 1991. Evolutionary dynamics of mitochondrial DNA duplications in parthenogenetic geckos, *Heteronotia binoei*. Genetics 129, 221–230.
Moritz, C., Brown, W.M., 1987. Tandem duplications in animal mitochondrial DNAs: variation in incidence and gene content among lizards. Proc. Natl. Acad. Sci. USA 84, 7183–7187.
Mueller, R.L., Boore, J.L., 2005. Molecular mechanisms of extensive mitochondrial gene rearrangment in plethodontid salamanders. Mol. Biol. Evol. 22, 2104–2112.
Müller, K., 2005. Seqstate: primer design and sequence statistics for phylogenetic DNA datasets. Appl. Bioinf. 4, 65–69.
Müller, K., 2006. Incorporating information from length-mutational events into phylogenetic analysis. Mol. Phylogenet. Evol. 38, 667–676.
Nylander, J.A.A., 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
Ogoh, K., Ohmiya, Y., 2004. Complete mitochondrial DNA sequence of the sea-firefly, *Vargula hilgendorfii* (Crustacea, Ostracoda) with duplicate control regions. Gene 327, 131–139.
Ogoh, K., Ohmiya, Y., 2007. Concerted evolution of duplicated control regions within an ostracod mitochondrial genome. Mol. Biol. Evol. 24, 74–78.
Ohno, S., 1970. Evolution by gene duplication. Springer-Verlag, New York, NY.
Quinn, T.W., 1997. Molecular evolution of the mitochondrial genome. In: Mindell, D.P. (Ed.), Avian Molecular Evolution and Systematics. Academic Press, San Diego, CA, pp. 4–29.
Quinn, T.W., Wilson, A.C., 1993. Sequence evolution in and around the mitochondrial control region in birds. J. Mol. Evol. 37, 417–425.
Quinn, T.W., Mindell, D.P., 1996. Mitochondrial gene order adjacent to the control region in crocodile, turtle and tuatara. Mol. Phylogenet. Evol. 5, 344–351.
Rambaut, A., Drummond, A.J., 2007. Tracer v1.4. <http://www.beast.bio.ed.ac.uk/Tracer>.
Rand, D.M., 1993. Endotherms, ectotherms and mitochondrial genome-size variation. J. Mol. Evol. 37, 281–295.
Rand, D.M., 2001. The units of selection on mitochondrial DNA. Annu. Rev. Ecol. Syst. 32, 415–448.
Rand, D.M., Harrison, R.G., 1986. Mitochondrial DNA transmission genetics in crickets. Genetics 114, 955–970.
Rastogi, S., Liberles, D.A., 2005. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. BMC Evol. Biol. 5, 28.
Remsen, Jr., J.V., Cadena, C.D., Jaramillo, A., Nores, M., Pacheco, J.F., Robbins, M.B., Schulenberg, T.S., Stiles, F.G., Stotz, D.F., Zimmer, K.J., 2010. A classification of the bird species of South America Version 7. American Ornithologists' Union. <http://www.museum.lsu.edu/~Remsen/SACCBaseline.html>.

Ronquist, F., Huelsenbeck, J.P., 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.
Roth, C., Rastogi, S., Arvestad, L., Dittmar, K., Light, S., Ekman, D., Liberles, D.A., 2007. Evolution after gene duplication: models, mechanisms, sequences, systems, and organisms. J. Exp. Zool. (Mol. Dev. Evol.) 308B, 58–73.
Sammler, S., Bleidorn, C., Tiedemann, R., 2011. Full mitochondrial genomes of two endemic Philippine hornbill species (Aves: Bucerotidae) provide evidence for pervasive mitochondrial DNA recombination. BMC Genomics 12, 35.
San Mauro, D., Gower, D.J., Zardoya, R., Wilkinson, M., 2006. A hotspot of gene order rearrangement by tandem duplication and random loss in the vertebrate mitochondrial genome. Mol. Biol. Evol., 23:227–234.
Sbisa, E., Tanzariello, F., Reyes, A., Pesole, G., Saccone, C., 1997. Mammalian mitochondrial D-loop region structural analysis: identification of new conserved sequences and the functional and evolutionary implications. Gene 205, 125–140.
Schultz, R.A., Swoap, S.J., McDaniel, L.D., Zhang, B., Koon, E.C., Garry, D.J., Li, K., Williams, R.S., 1998. Differential expression of mitochondrial DNA replication factors in mammalian tissues. J. Biol. Chem. 273, 3447–3451.
Schweizer, M., Seehausen, O., Güntert, M., Hertwig, S.T., 2010. The evolutionary diversification of parrots supports a taxon pulse model with multiple trans-oceanic dispersal events and local radiations. Mol. Phylogenet. Evol. 54, 984–994.
Selosse, M-.A., Albert, B., Godelle, B., 2001. Reducing the genome size of organelles favors gene transfer to the nucleus. Trends Ecol. Evol. 16, 135–141.
Shadel, G.S., Clayton, D.A., 1997. Mitochondrial DNA maintenance in vertebrates. Ann. Rev. Biochem. 66, 409–435.
Shao, R., Barker, S.C., 2003. The highly rearranged mitochondrial genome of the plague thrips, *Thrips imaginis* (Insecta: Thysanoptera): convergence of two novel gene boundaries and an extraordinary arrangement of rRNA genes. Mol. Biol. Evol. 20, 362–370.
Shao, R., Barker, S.C., Mitani, H., Aoki, Y., Fukunaga, M., 2005. Evolution of duplicate control regions in the mitochondrial genomes of Metazoa: a case study with Australasian *Ixodes* ticks. Mol. Biol. Evol. 22, 620–629.
Shokolenko, I.N., Ledoux, S.P., Wison, G.L., 2007. Mitochondrial DNA damage and repair. In: Schaffer, S.W., Suleiman, M.-S. (Eds.), Mitochondria: The Dynamic Organelle. Springer, New York, NY, pp. 323–348.
Sibley, C., Ahlquist, J., 1990. Phylogeny and Classification of Birds. Yale University Press, New Haven, CT.
Simmons, M.P., Ochoterena, H., 2000. Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49, 369–381.
Sogin, M.L., 1997. History assignment: when was the mitochondrion founded? Curr. Opin. Genetics Dev. 7, 792–799.
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.
Tatarenkov, A., Avise, J.C., 2007. Rapid concerted evolution in animal mitochondrial DNA. Proc. R. Soc. B 264, 1795–1798.
Tavares, E.S., Baker, A.J., Pereira, S.L., Miyaki, C.Y., 2006. Phylogenetic relationships and historical biogeography of Neotropical parrots (Psittaciformes: Psittacidae: Arini) inferred from mitochondrial and nuclear DNA sequences. Syst. Biol. 55, 454–470.
Verkuil, Y.I., Piersma, T., Baker, A.J., 2010. A novel mitochondrial gene order in shorebirds (Scolopacidae, Charadriiformes). Mol. Phylogent. Evol. 57, 411–416.
Wright, T.F., Schirtzinger, E.E., Matsumoto, T., Eberhard, J.R., Graves, G.R., Sanchez, J.J., Capelli, S., Müller, H., Scharpegge, J., Chambers, G.K., Fleischer, R.C., 2008. A multilocus molecular phylogeny of the parrots (Psittaciformes): support for a Gondwanan origin during the Cretaceous. Mol. Biol. Evol. 25, 2141–2156.
Zhang, J., 2003. Evolution by gene duplication: AN update. Trends Ecol. Evol. 18, 292–298.
Zhuang, X., Cheng, C-.H., 2010. ND6 gene ‘‘lost’’ and found: evolution of mitochondrial gene rearrangement in Antarctic Notothenoids. Mol. Biol. Evol. 27, 1391–1403.
Zwickl, D.J., 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation. The University of Texas at Austin, Austin, Texas.

Figure 1S. An illustration of the criteria used to score control region state based on the comparison of measured diagnostic fragment lengths with expected fragment lengths.

Segment	Species	Measured fragment length	Single CR Segment 15 expected size	Duplicated CR Segment 15 expected size	Single CR Segment 16 expected size	Duplicated CR Segment 16 expected size	Single CR Segment ND6 expected size	Duplicated CR Segment ND6 expected size	Score
			1579-1896	800-1250	1225-1542	600-1000	132	>1200	
Segment 15	Myiopsitta monachus	1639	X						
Segment 16	Myiopsitta monachus	1266			X				
Segment ND6	Myiopsitta monachus	142					X		
									Single

Segment	Species	Measured fragment length	Single CR Segment 15 expected size	Duplicated CR Segment 15 expected size	Single CR Segment 16 expected size	Duplicated CR Segment 16 expected size	Single CR Segment ND6 expected size	Duplicated CR Segment ND6 expected size	Score
			1579-1896	800-1250	1225-1542	600-1000	132	>1200	
Segment 15	Amazona viridigenalis	1092		X					
Segment 16	Amazona viridigenalis	738				X			
Segment ND6	Amazona viridigenalis	2146						X	
									Duplicated

Figure 2S. Maximum Likelihood tree from 20 independent GARLI runs and 1000 bootstrap pseudoreplicates. Bootstrap values greater than 50 are shown. Asterisks represent a bootstrap value of 100

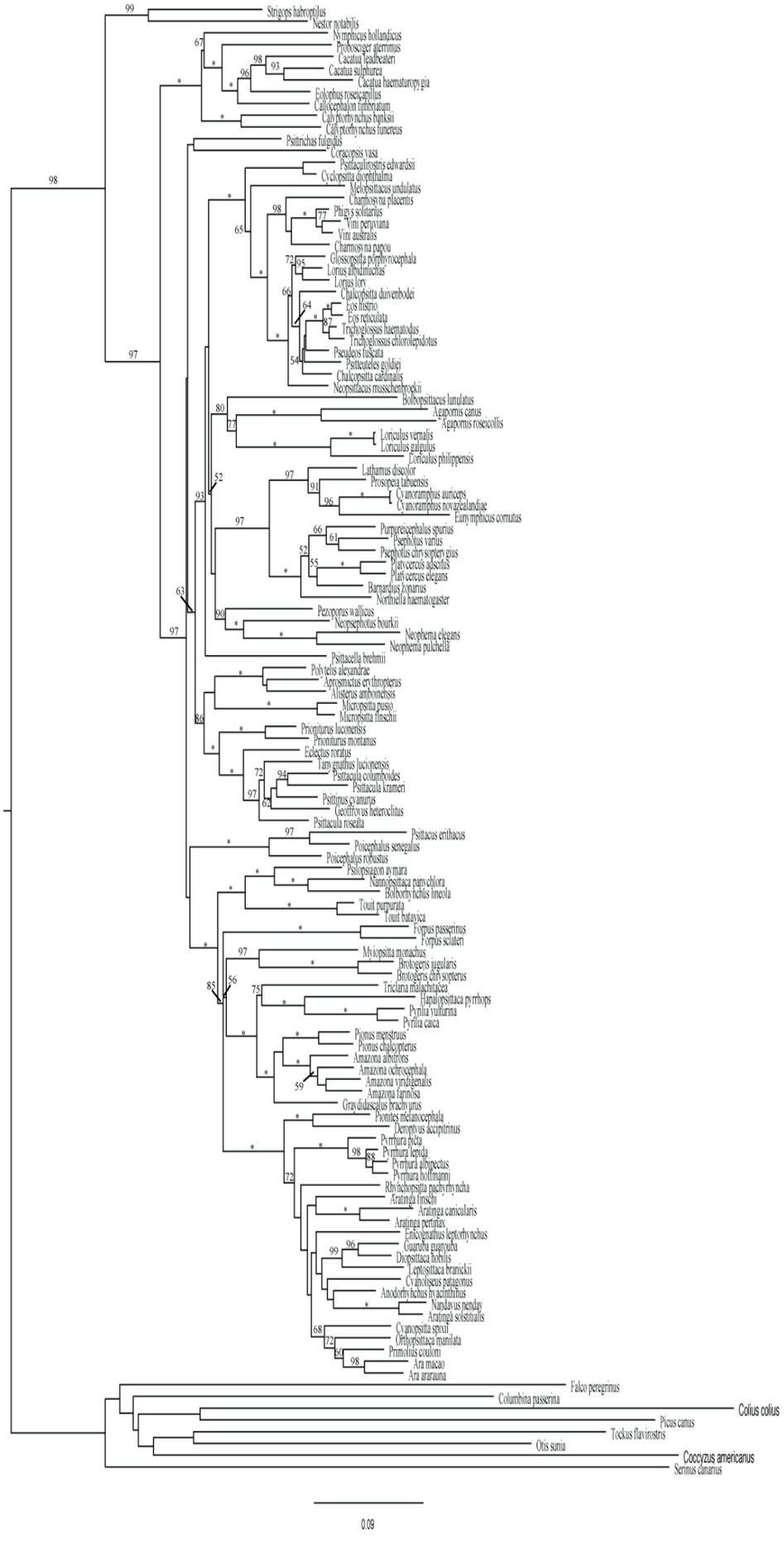

Figure 3S. Ancestral state reconstruction of mitochondrial control region states with nodes numbered to correspond with Table 6S, maximum likelihood values for each state for each model reconstructed.

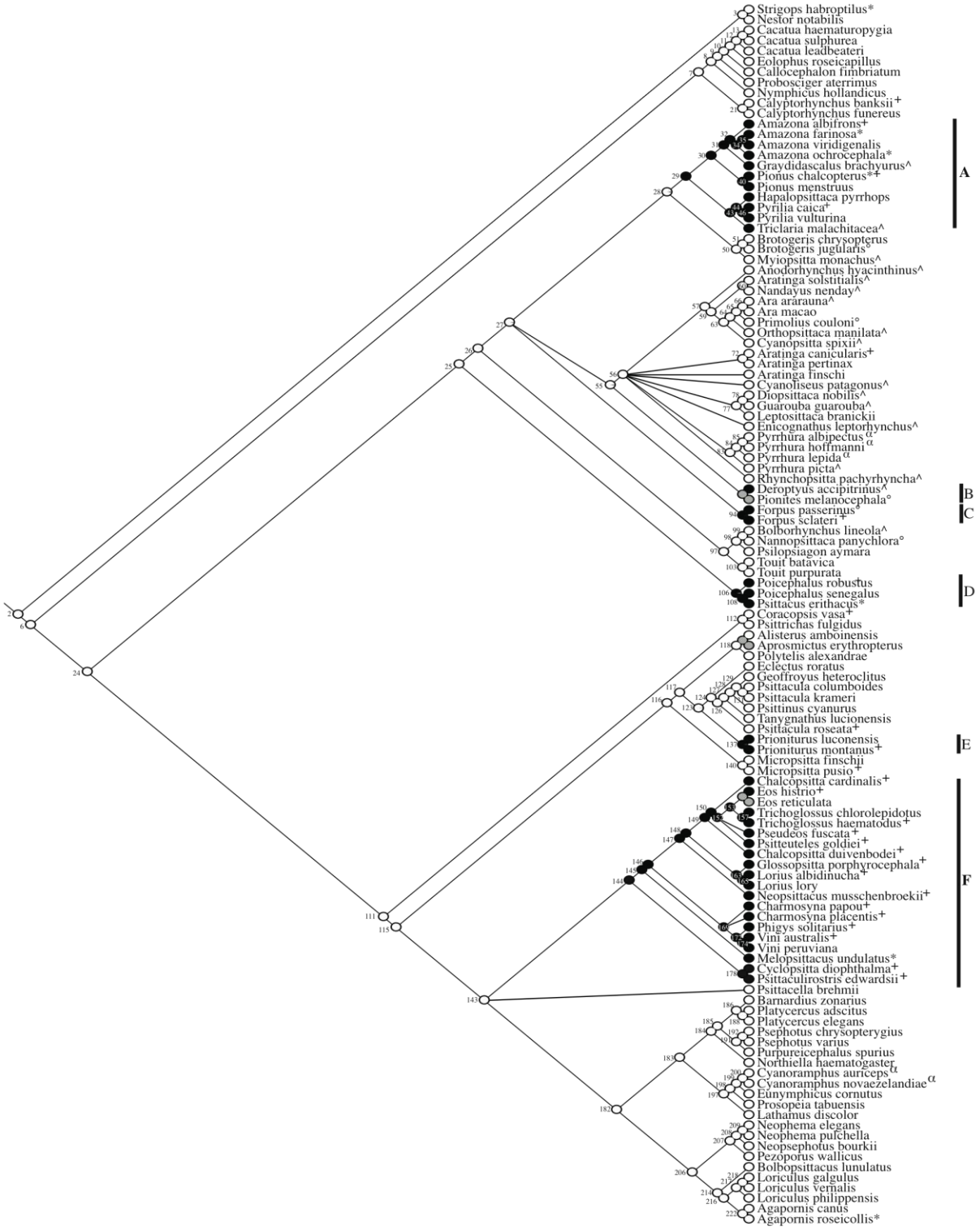

Table 1S: New sequence data and specimens included in the present study.

Taxa	Source[a]	Specimen #	Location[b]	GenBank Accn#			
				COI	ND2	TROP	TGFB2
Agapornis canus	NMNH	B08933	NMSU	HQ629749	HQ629714	HQ629670	HQ629625
Amazona albifrons	NMNH	B06567	NMSU	HQ629750	HQ629715	HQ629671	HQ629626
Amazona farinosa	NMNH	B09176	NMSU	HQ629751	HQ629716	HQ629672	HQ629627
Amazona ochrocephala	NMNH	B12937	NMSU	HQ629752	HQ629717	HQ629673	HQ629628
Anodorhynchus hyacinthinus	LSUMNS	B-13478	NMSU	GU826173	HQ270480	HQ629674	HQ629629
Ara ararauna	NMSUVM	46617	NMSU	HQ629755	HQ629720	HQ629678	HQ629630
Aratinga canicularis	AMNH	DOT 9252	NMSU	HQ629753	HQ629718	HQ629675	HQ629631
Aratinga finschi	NMNH	B00402	NMSU	HQ629754	HQ629719	HQ629676	HQ629632
Aratinga solstitialis	NMNH	B06816	NMSU	GU826185	HQ270491	HQ629677	HQ629633
Bolborhynchus lineola	LSUMNS	B-165252	NMSU	GU826187	HQ270493	HQ629679	HQ629634
Brotogeris chrysopterus	NMNH	B07096	NMSU	HQ629756	HQ629721	HQ629680	HQ629635
Cacatua haematuropygia	NMNH	B02947	NMSU	HQ629757	HQ629722	HQ629681	HQ629636
Cacatua leadbeateri	NMNH	B02881	NMSU	HQ629758	HQ629723	HQ629682	HQ629637
Callocephalon fimbriatum	ANWC	34252	NMSU	HQ629759	HQ629724	HQ629683	HQ629638
Calyptorhynchus banksii	ANWC	50042	NMSU	HQ316860	HQ316873	HQ316886	HQ316899
Chalcopsitta cardinalis	AMNH	DOT 6626	NMSU	HQ629760	HQ629725	HQ629684	HQ629639
Charmosyna placentis	AMNH	DOT 7798	NMSU	HQ629761	HQ629726	HQ629685	HQ629640
Cyanoramphus novazealandiae	AMNH	DOT 11060	NMSU	HQ316861	HQ316874	HQ316887	HQ316900
Eos histrio	AMNH	DOT 7703	NMSU	HQ629762	HQ629727	HQ629686	HQ629642
Eunymphicus cornutus	SDZ	399555	NMSU	HQ629763	HQ629728	HQ629687	HQ629643
Forpus sclateri	NMNH	B09739	NMSU	HQ629764	HQ629729	HQ629689	HQ629644
Graydidascalus brachyurus	LSUMNS	B-3626	NMSU	GU826191	HQ270497	HQ629690	HQ629645
Lathamus discolor	ANWC	34174	NMSU	HQ316862	HQ316875	HQ316888	HQ316901
Loriculus philippensis	NMNH	B03792	NMSU	HQ629765	HQ629730	HQ629691	HQ629646
Loriculus vernalis	NMNH	B02952	NMSU	HQ629766	HQ629731	HQ629692	HQ629647
Lorius lory	NMNH	B06576	NMSU	HQ629767	HQ629732	HQ629693	HQ629648
Micropsitta pusio	KUMNH	5188	NMSU	HQ629768	HQ629733	HQ629694	HQ629649

Neophema pulchella	NMNH	B06823	NMSU	HQ629769	HQ629734	HQ629695	HQ629651
Pezoporus wallicus	ANWC	45982	NMSU	HQ316865	HQ316878	HQ316891	HQ316904
Pyrilia vulturina	NMNH	B06888	NMSU	GU826193	HQ270499	HQ629696	HQ629652
Pionus chalcopterus	NMNH	B08760	NMSU	HQ629770	HQ629735	HQ629697	HQ629653
Platycercus elegans	NMNH	B06370	NMSU	HQ316866	HQ316879	HQ316892	HQ316905
Poicephalus senegalus	LSUMNS	B-3910	NMSU	HQ629771	HQ629736	HQ629698	HQ629654
Prioniturus montanus	FMNH	392241	NMSU	HQ629772	HQ629737	HQ629699	HQ629655
Primolius couloni	FMNH	395540	NMSU	GU826195	HQ270501	HQ629700	HQ629656
Psephotus chrysopterygius	LSUMNS	B-24921	NMSU	HQ629773		HQ629701	HQ629657
Psilopsiagon aymara	LSUMNS	B-1201	NMSU	HQ629774	HQ629738	HQ629703	HQ629658
Psittacella brehmii	KUMNH	4600	NMSU	HQ316870	HQ316883	HQ316896	HQ316909
Psittacula krameri	NMNH	B02548	NMSU	HQ629775	HQ629739	HQ629704	HQ629659
Psittaula roseata	AMNH	DOT 9897	NMSU	HQ629776	HQ629740	HQ629705	HQ629660
Psitteuteles goldiei	AMNH	DOT 7897	NMSU	HQ629777	HQ629741	HQ629706	HQ629661
Psittinus cyanurus	LSUMNS	B-30042	NMSU	HQ629778	HQ629742	HQ629707	HQ629662
Pyrrhura albipectus	LSUMNS	B-5958	NMSU	HQ629779	HQ629743	HQ629708	HQ629663
Pyrrhura hoffmanni	NMNH	B05272	NMSU	HQ629780	HQ629744	HQ629709	HQ629664
Pyrrhura lepida	NMNH	B07007	NMSU	GU826196	HQ270502	HQ629710	HQ629665
Touit purpurata	NMNH	B09558	NMSU	HQ629781	HQ629745	HQ629711	HQ629666
Trichoglossus chlorolepidotus	NMNH	B06422	NMSU	HQ629782	HQ629746	HQ629712	HQ629667
Triclaria malachitacea	USP	LGEMA 415	Brazil	HQ629783	HQ629747		HQ629668
Vini peruviana	AMNH	DOT 7694	NMSU	HQ6297845	HQ629748	HQ629713	HQ629669
Individual Gene Regions Added to Taxa from Wright et al. (2008)							
Psephotus varius	ANWC	32871	NMSU			HQ629702	
Forpus passerinus	NMNH	B12187	NMSU			HQ629688	
Cyanopsitta spixii	LP		Spain				HQ629641
Nandayus nenday	LP	07-23	Spain				HQ629650

[a] Source abbreviations are as follows: NMNH-US National Museum of Natural History, AMNH-American Museum

of Natural History, ANSP-Academy of Natural Science, Philadelphia, ANWC-Australian National Wildlife Collection, LP-Loro Parque, Tenerife, Canary Islands, Spain, LSUMNS-Louisiana State University Museum of Natural Science, USP-Universidade de São Paulo, Brazil, KUMNH-University of Kansas Museum of Natural History, NMSUVM-New Mexico State University Vertebrate Museum, SDZ-San Diego Zoological Park.

[b]Location indicates the laboratory in which extraction, PCR and sequencing took place. NMSU-Dept of Biology, New Mexico State University, Las Cruces, New Mexico, USA, Brazil- Departmento de Genética e Biologia Evolutiva, Universidade de São Paulo, Brazil, Spain- Instituto Nacional de Toxicologia y Ciencias Forenses, Tenerife, Canary Islands, Spain.

Table 2S: Previously published sequences included in the present study

Taxa	GenBank Accn#			
	COI	**ND2**	**TROP**	**TGFB2**
Agapornis roseicollis	EU621593	EU327596	EU665562	EU660234
Alisterus amboinensis	EU621594	EU327597	EU665563	EU660235
Amazona viridigenalis	EU621595	EU327598	EU665564	EU660236
Aprosmictus erythropterus	EU621596	EU327599	EU665565	EU660237
Ara macao	EU621598	EU327600	EU665566	EU660238
Aratinga pertinax	EU621597	EU327601	EU665567	EU660239
Barnardius zonarius	EU621599	EU327602	EU665568	EU660240
Bolbopsittacus lunulatus	EU621600	EU327603	EU665569	EU660241
Brotogeris jugularis	EU621601	EU327604	EU665570	EU660242
Cacatua sulphurea	EU621602	EU327605	EU665571	EU660243
Calyptorhynchus funereus	EU621603	EU327606	EU665572	EU660244
Chalcopsitta duivenbodei	EU621604	EU327607	EU665573	EU660245
Charmosyna papou	EU621605	EU327608	EU665574	EU660246
Coracopsis vasa	EU621608	EU327612	EU665578	EU660250
Cyanoliseus patagonus	EU621609	EU327613	EU665579	EU660251
Cyanopsitta spixii	EU621610	EU327614	EU665580	
Cyanoramphus auriceps	EU621611	EU327615	EU665581	EU660252
Cyclopsitta diophthalma	EU621612	EU327616	EU665582	EU660253
Deroptyus accipitrinus	EU621613	EU327617	EU665583	EU660254
Diopsittaca nobilis	EU621614	EU327618	EU665584	EU660255
Eclectus roratus	EU621615	EU327619	EU665585	EU660256
Enicognathus leptorhynchus	EU621616	EU327620	EU665586	EU660257
Eolophus roseicapillus	EU621617	EU327621	EU665587	EU660258
Eos reticulata	EU621618	EU327622	EU665588	EU660259
Forpus passerinus	EU621621	EU327625		EU660262
Geoffroyus heteroclitus	EU621622	EU327626	EU665591	EU660263
Glossopsitta porphyrocephala	EU621623	EU327627	EU665592	EU660264
Guarouba guarouba	EU621624	EU327628	EU665593	EU660265
Hapalopsittaca pyrrhops	EU621625	EU327629	EU665594	
Leptosittaca branickii	EU621626	EU327630	EU665595	EU660266
Loriculus galgulus	EU621627	EU327631	EU665596	EU660267
Lorius albidinuchas	EU621628	EU327632	EU665597	EU660268
Melopsittacus undulatus	EU621629	EU327633	EU665598	EU660269
Micropsitta finschii	EU621630	EU327634	EU665599	EU660270
Myiopsitta monachus	EU621631	EU327635	EU665600	EU660271
Nandayus nenday	EU621632	EU327636	EU665601	
Nannopsittaca panychlora	EU621633	EU327637	EU665602	EU660272

Neophema elegans	EU621634	EU327638	EU665603	EU660273
Neopsephotus bourkii	EU621635	EU327639	EU665604	EU660274
Neopsittacus musschenbroekii	EU621636	EU327640	EU665605	EU660275
Nestor notabilis	EU621637	EU327641	EU665606	EU660276
Northiella haematogaster	EU621638	EU327642	EU665607	EU660277
Nymphicus hollandicus	EU621639	EU327643	EU665608	EU660278
Orthopsittaca manilata	EU621640	EU327644	EU665609	EU660279
Phigys solitarius	EU621642	EU327646	EU665611	EU660281
Pionites melanocephala	EU621644	EU327648	EU665613	EU660282
Pyrilia caica	EU621645	EU327649	EU665614	EU660283
Pionus menstruus	EU621646	EU327650	EU665615	EU660284
Platycercus adscitus	EU621647	EU327651	EU665616	EU660285
Poicephalus robustus	EU621648	EU327652	EU665617	EU660286
Polytelis alexandrae	EU621649	EU327653	EU665618	EU660287
Prioniturus luconensis	EU621650	EU327654	EU665619	EU660288
Probosciger aterrimus	EU621651	EU327655	EU665620	EU660289
Prosopeia tabuensis	EU621652	EU327656	EU665621	EU660290
Psephotus varius	EU621653	EU327657		EU660291
Pseudeos fuscata	EU621654	EU327658	EU665622	EU660292
Psittacula columboides	EU621655	EU327659	EU665623	EU660293
Psittaculirostris edwardsii	EU621656	EU327660	EU665624	EU660294
Psittacus erithacus	EU621657	EU327661	EU665625	EU660295
Psittrichas fulgidus	EU621658	EU327662	EU665626	EU660296
Purpureicephalus spurius	EU621659	EU327663	EU665627	EU660297
Pyrrhura picta	EU621660	EU327664	EU665628	EU660298
Rhynchopsitta pachyrhyncha	EU621661	EU327665	EU665629	EU660299
Strigops habroptilus	EU621663	EU327667	EU665631	EU660301
Tanygnathus lucionensis	EU621664	EU327668	EU665632	EU660302
Touit batavica	EU621666	EU327670	EU665634	EU660304
Trichoglossus haematodus	EU621667	EU327671	EU665635	EU660305
Vini australis	EU621668	EU327672	EU665636	EU660306
Outgroups				
Coccyzus americanus	EU621606	EU327609	EU665575	EU660247
Colius colius		EU327610	EU665576	EU660248
Columbina passerina	EU621607	EU327611	EU665577	EU660249
Falco peregrinus	EU621620	EU327624	EU665590	EU660261
Otus sunia	EU621641	EU327645	EU665610	EU660280
Picus canus	EU621643	EU327647	EU665612	
Serinus canaria	EU621662	EU327666	EU665630	EU660300
Tockus flavirostris	EU621665	EU327669	EU665633	EU660303

Table 3S. GenBank accession numbers for sequences used to determine diagnostic fragment lengths..

Species	GenBank #
Amazona amazonica	AF338277
Amazona amazonica	AF338279
Amazona amazonica	AF338278
Amazona amazonica	AF338280
Amazona auropalliata	AF338819
Amazona auropalliata	AF338299
Amazona auropalliata	AF338300
Amazona autumnalis	AF338281
Amazona autumnalis	AF338283
Amazona autumnalis	AF338284
Amazona autumnalis	AF338282
Amazona farinosa	AF338821
Amazona farinosa	AF338285
Amazona farinosa	AF338286
Amazona farinosa	AF338287
Amazona farinosa	AF338288
Amazona ochrocephala	AF323131
Amazona ochrocephala	AF338303
Amazona ochrocephala	AF338305
Amazona ochrocephala	AF338304
Amazona ochrocephala	AF338306
Amazona oratrix	AF338820
Amazona oratrix	AF338307
Amazona oratrix	AF338309
Amazona oratrix	AF338308
Amazona oratrix	AF338310
Cyanoramphus auriceps	AF218749
Cyanoramphus auriceps	AF218752
Cyanoramphus auriceps	AF218750
Cyanoramphus cooki	AF265703
Cyanoramphus forbesi	AB179750
Cyanoramphus forbesi	AF218740
Cyanoramphus forbesi	AF265704
Cyanoramphus forbesi	AB178749
Cyanoramphus malherbi	AF218754
Cyanoramphus novaezelandiae	AF218745
Cyanoramphus saisetti	AF265706
Cyanoramphus saisetti	AF365708
Cyanoramphus unicolor	AF218739

Cyanoramphus unicolor	AF265700
Pionus chalcopterus	AF338317
Pionus chalcopterus	AF338318
Psittacus erithacus	DQ335468
Agapornis roseicollis	EU410486
Melopsittacus undulatus	NC_009134

Table 4S. Sequence characteristics for the gene regions and concatenated dataset used to reconstruct the phylogeny in this study

	COI	ND2	Trop	TGFB2	Concatenated
# characters (bp)	570.0	1041.0	554.0	817.0	2982.0
length range (bp)	570	1041	353-540	766	1716-2906
mean length (bp)	565.1	1020.3	523.4	604.0	2716.4
S.D. length	50.9	113.6	19.8	97.6	150.8
mean % divergence (range)	11.9 (0.0-19.1)	17.6 (0.4-30.0)	3.2 (0.0-13.2)	6.7 (0.0-22.3)	11.3 (0.2-23.0)
mean ti/tv (range)	3.5 (0.8-65.0)	1.9 (0.6-34.0)	6.0 (0.0-19.0)	2.5 (0.0-13.0)	2.2 (0.8-23.5)
% variable	42.8	73.0	40.3	57.0	56.6
% informative	39.6	65.7	22.7	37.2	45.2
% GC	46.5	45.0	39.7	44.8	44.2
% A	28.2	32.8	21.7	24.1	27.7
% C	29.1	36.5	21.4	21.7	28.8
% G	17.5	8.4	18.3	23.2	15.5
% T	25.3	22.2	38.6	31.1	28.0
# gap characters	0.0	0.0	48.0	78.0	N/A

Table 5S: Nucleotide substitution models and parameters for phylogenetic reconstructions of the Psittaciformes

		MrBayes					**GARLI**
Dataset		COI	ND2	Trop	TGFB2	Gaps	Concatenated
Model		Empirical GTR+I+G	Empirical GTR+I+G	Empirical GTR+G	Empirical GTR+I+G	Restriction Variable	Estimated GTR+I+G
Base Freq	Freq(A)	0.3955	0.3573	0.2209	0.2433	N/A	0.3484
	Freq(C)	0.3825	0.4628	0.1967	0.2049	N/A	0.3733
	Freq(G)	0.0714	0.0344	0.1798	0.2251	N/A	0.0792
	Freq(T)	0.1506	0.1455	0.4026	0.3267	N/A	0.1991
Rate Matrix	R(a) [A-C]	0.4624	0.0704	1.6962	1.0908	N/A	0.3024
	R(b) [A-G]	16.5334	5.8018	12.2526	5.2983	N/A	5.5010
	R(c) [A-T]	1.1307	0.167	0.5713	0.6366	N/A	0.3329
	R(d) [C-G]	0.093	0.163	1.7886	1.1683	N/A	0.2771
	R(e) [C-T]	22.0.174	1.7705	4.5985	3.9653	N/A	4.2134
	R(f) [G-T]	1	1	1	1	N/A	1
	Prop. of. Inv. Sites	0.5581	0.2316	0	0.0947	N/A	0.3057
	γ shape parameter	0.6776	0.6287	0.5172	3.5201	N/A	0.5409

Table 6S: Proportional Likelihoods of the Reconstructed States (Single or Duplicated Control Region) Under Different Maximum Likelihood Models.

	MK1 Model		Asymmetric Model					
	(rate estimated)		(rates estimated)		(rates 5:1)		(rates 1:5)	
Node	Single	Duplicated	Single	Duplicated	Single	Duplicated	Single	Duplicated
2	0.99825619	0.00174381	0.99906664	0.00093336	0.99428442	0.00571558	0.99966973	0.00033027
3	0.99938167	0.00061833	0.99980383	0.00019617	0.99797507	0.00202493	0.99992961	0.00007039
6	0.99992471	0.00007529	0.99998099	0.00001901	0.99982506	0.00017494	0.99972837	0.00027163
7	0.99999071	0.00000929	0.99999769	0.00000231	0.99997767	0.00002233	0.99999808	0.00000192
8	0.99998854	0.00001146	0.99999693	0.00000307	0.99997061	0.00002939	0.99999886	0.00000114
9	0.99998287	0.00001713	0.99999501	0.00000499	0.99995545	0.00004455	0.99999854	0.00000146
10	0.99999254	0.00000746	0.99999792	0.00000208	0.99998177	0.00001823	0.99999937	0.00000063
11	0.99999348	0.00000652	0.99999814	0.00000186	0.99998418	0.00001582	0.99999944	0.00000056
12	0.99999156	0.00000844	0.99999761	0.00000239	0.99997935	0.00002065	0.99999928	0.00000072
13	0.99998387	0.00001613	0.99999524	0.00000476	0.99995931	0.00004069	0.99999861	0.00000139
21	0.99991530	0.00008470	0.99997416	0.00002584	0.99976797	0.00023203	0.99999267	0.00000733
24	0.99990169	0.00009831	0.99998916	0.00001084	0.99997822	0.00002178	0.99723537	0.00276463
25	0.99476630	0.00523370	0.99880295	0.00119705	0.99858327	0.00141673	0.79729847	0.20270153
26	0.99511016	0.00488984	0.99950267	0.00049733	0.99979891	0.00020109	0.78738501	0.21261499
27	0.98413974	0.01586026	0.99709441	0.00290559	0.99871525	0.00128475	0.65884473	0.34115527
28	0.98021897	0.01978103	0.99516441	0.00483559	0.99549974	0.00450026	0.65706082	0.34293918
29	0.00663183	0.99336817	0.00769122	0.99230878	0.04381410	0.95618590	0.00051047	0.99948953
30	0.00006232	0.99993768	0.00007998	0.99992002	0.00205228	0.99794772	0.00000064	0.99999936
31	0.00001574	0.99998426	0.00001833	0.99998167	0.00092528	0.99907472	0.00000011	0.99999989
32	0.00000271	0.99999729	0.00000180	0.99999820	0.00005277	0.99994723	0.00000005	0.99999995
34	0.00000084	0.99999916	0.00000055	0.99999945	0.00001691	0.99998309	0.00000001	0.99999999
35	0.00000376	0.99999624	0.00000240	0.99999760	0.00004189	0.99995811	0.00000007	0.99999993
40	0.00001399	0.99998601	0.00000971	0.99999029	0.00024470	0.99975530	0.00000024	0.99999976
43	0.00203659	0.99796341	0.00250033	0.99749967	0.02745930	0.97254070	0.00003110	0.99996890
44	0.00031057	0.99968943	0.00032364	0.99967636	0.00758203	0.99241797	0.00000277	0.99999723

46	0.00001247	0.99998753	0.00000869	0.99999131	0.00023373	0.99976627	0.00000020	0.99999980
50	0.99681633	0.00318367	0.99952533	0.00047467	0.99889534	0.00110466	0.96001500	0.03998500
51	0.99993792	0.00006208	0.99998605	0.00001395	0.99988580	0.00011420	0.99979038	0.00020962
55	0.98037261	0.01962739	0.99380484	0.00619516	0.99245577	0.00754423	0.67086203	0.32913797
56	1.00000000	0.00000000	1.00000000	0.00000000	1.00000000	0.00000000	1.00000000	0.00000000
57	0.99999905	0.00000095	0.99999972	0.00000028	0.99999774	0.00000226	0.99999992	0.00000008
59	0.99999792	0.00000208	0.99999940	0.00000060	0.99999501	0.00000499	0.99999982	0.00000018
60	0.99998904	0.00001096	0.99999671	0.00000329	0.99997258	0.00002742	0.99999904	0.00000096
63	0.99999769	0.00000231	0.99999934	0.00000066	0.99999445	0.00000555	0.99999980	0.00000020
64	0.99999804	0.00000196	0.99999944	0.00000056	0.99999535	0.00000465	0.99999983	0.00000017
65	0.99999733	0.00000267	0.99999924	0.00000076	0.99999369	0.00000631	0.99999977	0.00000023
66	0.99999130	0.00000870	0.99999741	0.00000259	0.99997860	0.00002140	0.99999924	0.00000076
72	0.99997996	0.00002004	0.99999395	0.00000605	0.99994895	0.00005105	0.99999825	0.00000175
77	0.99999292	0.00000708	0.99999793	0.00000207	0.99998262	0.00001738	0.99999939	0.00000061
78	0.99999365	0.00000635	0.99999815	0.00000185	0.99998461	0.00001539	0.99999945	0.00000055
83	0.99998897	0.00001103	0.99999669	0.00000331	0.99997230	0.00002770	0.99999904	0.00000096
84	0.99999951	0.00000049	0.99999986	0.00000014	0.99999889	0.00000111	0.99999996	0.00000004
85	0.99999959	0.00000041	0.99999988	0.00000012	0.99999906	0.00000094	0.99999996	0.00000004
94	0.01094508	0.98905492	0.01174040	0.98825960	0.03102318	0.96897682	0.00119715	0.99880285
97	0.99954296	0.00045704	0.99996311	0.00003689	0.99985080	0.00014920	0.98687529	0.01312471
98	0.99993314	0.00006686	0.99998756	0.00001244	0.99989198	0.00010802	0.99938668	0.00061332
99	0.99994078	0.00005922	0.99998288	0.00001712	0.99984717	0.00015283	0.99996348	0.00003652
103	0.99996659	0.00003341	0.99999032	0.00000968	0.99991459	0.00008541	0.99995354	0.00004646
106	0.02988383	0.97011617	0.03233147	0.96766853	0.09219604	0.90780396	0.00381413	0.99618587
108	0.00114297	0.99885703	0.00131343	0.99868657	0.01041859	0.98958141	0.00002377	0.99997623
111	0.99999728	0.00000272	0.99999983	0.00000017	0.99999893	0.00000107	0.99995053	0.00004947
112	0.99995043	0.00004957	0.99998496	0.00001504	0.99986012	0.00013988	0.99997639	0.00002361
115	0.99999858	0.00000142	0.99999977	0.00000023	0.99999845	0.00000155	0.99999391	0.00000609
116	0.99997840	0.00002160	0.99999634	0.00000366	0.99997813	0.00002187	0.99998937	0.00001063
117	0.99983203	0.00016797	0.99995755	0.00004245	0.99988007	0.00011993	0.99990660	0.00009340
118	0.99993810	0.00006190	0.99998193	0.00001807	0.99984433	0.00015567	0.99999311	0.00000689

123	0.99360473	0.00639527	0.99673894	0.00326106	0.99478855	0.00521145	0.99519796	0.00480204
124	0.99984415	0.00015585	0.99995805	0.00004195	0.99987287	0.00012713	0.99991722	0.00008278
126	0.99999502	0.00000498	0.99999894	0.00000106	0.99999250	0.00000750	0.99999876	0.00000124
127	0.99999884	0.00000116	0.99999969	0.00000031	0.99999742	0.00000258	0.99999988	0.00000012
128	0.99999938	0.00000062	0.99999983	0.00000017	0.99999856	0.00000144	0.99999995	0.00000005
129	0.99999851	0.00000149	0.99999959	0.00000041	0.99999647	0.00000353	0.99999987	0.00000013
131	0.99998841	0.00001159	0.99999653	0.00000347	0.99997080	0.00002920	0.99999899	0.00000101
137	0.01727722	0.98272278	0.01896256	0.98103744	0.05519694	0.94480306	0.00276265	0.99723735
140	0.99998321	0.00001679	0.99999489	0.00000511	0.99995523	0.00004477	0.99999854	0.00000146
143	0.99988434	0.00011566	0.99996952	0.00003048	0.99992099	0.00007901	0.99993064	0.00006936
144	0.01071529	0.98928471	0.01195623	0.98804377	0.04825646	0.95174354	0.00147144	0.99852856
145	0.00164604	0.99835396	0.00198078	0.99801922	0.01849449	0.98150551	0.00004040	0.99995960
146	0.00002873	0.99997127	0.00003587	0.99996413	0.00103316	0.99896684	0.00000016	0.99999984
147	0.00000143	0.99999857	0.00000105	0.99999895	0.00003679	0.99996321	0.00000002	0.99999998
148	0.00000006	0.99999994	0.00000004	0.99999996	0.00000160	0.99999840	0.00000000	1.00000000
149	0.00000043	0.99999957	0.00000027	0.99999973	0.00000473	0.99999527	0.00000001	0.99999999
150	0.00000010	0.99999990	0.00000007	0.99999993	0.00000127	0.99999873	0.00000000	1.00000000
152	0.00000001	0.99999999	0.00000001	0.99999999	0.00000041	0.99999959	0.00000000	1.00000000
153	0.00000044	0.99999956	0.00000029	0.99999971	0.00000493	0.99999507	0.00000001	0.99999999
157	0.00000025	0.99999975	0.00000016	0.99999984	0.00000282	0.99999718	0.00000000	1.00000000
163	0.00000033	0.99999967	0.00000021	0.99999979	0.00000409	0.99999591	0.00000001	0.99999999
165	0.00000120	0.99999880	0.00000077	0.99999923	0.00001313	0.99998687	0.00000002	0.99999998
169	0.00000038	0.99999962	0.00000030	0.99999970	0.00002814	0.99997186	0.00000000	1.00000000
172	0.00000079	0.99999921	0.00000051	0.99999949	0.00000912	0.99999088	0.00000001	0.99999999
174	0.00000046	0.99999954	0.00000030	0.99999970	0.00000538	0.99999462	0.00000001	0.99999999
178	0.00005226	0.99994774	0.00005976	0.99994024	0.00075436	0.99924564	0.00000110	0.99999890
182	0.99999326	0.00000674	0.99999899	0.00000101	0.99999495	0.00000505	0.99999723	0.00000277
183	0.99995744	0.00004256	0.99998725	0.00001275	0.99988863	0.00011137	0.99999631	0.00000369
184	0.99998939	0.00001061	0.99999690	0.00000310	0.99997318	0.00002682	0.99999909	0.00000091
185	0.99999922	0.00000078	0.99999981	0.00000019	0.99999828	0.00000172	0.99999994	0.00000006
186	0.99999369	0.00000631	0.99999813	0.00000187	0.99998438	0.00001562	0.99999945	0.00000055

188	0.99999155	0.00000845	0.99999748	0.00000252	0.99997922	0.00002078	0.99999926	0.00000074
191	0.99999523	0.00000477	0.99999862	0.00000138	0.99998850	0.00001150	0.99999959	0.00000041
192	0.99999216	0.00000784	0.99999770	0.00000230	0.99998075	0.00001925	0.99999932	0.00000068
197	0.99998892	0.00001108	0.99999681	0.00000319	0.99997267	0.00002733	0.99999905	0.00000095
198	0.99999453	0.00000547	0.99999849	0.00000151	0.99998701	0.00001299	0.99999954	0.00000046
199	0.99998188	0.00001812	0.99999460	0.00000540	0.99995418	0.00004582	0.99999843	0.00000157
200	0.99999995	0.00000005	0.99999998	0.00000002	0.99999988	0.00000012	1.00000000	0.00000000
206	0.99999885	0.00000115	0.99999978	0.00000022	0.99999801	0.00000199	0.99999983	0.00000017
207	0.99998167	0.00001833	0.99999499	0.00000501	0.99995252	0.00004748	0.99999849	0.00000151
208	0.99991783	0.00008217	0.99997573	0.00002427	0.99977166	0.00022834	0.99999312	0.00000688
209	0.99982096	0.00017904	0.99994502	0.00005498	0.99948399	0.00051601	0.99998483	0.00001517
214	0.99996392	0.00003608	0.99999092	0.00000908	0.99989761	0.00010239	0.99999726	0.00000274
216	0.99991865	0.00008135	0.99997789	0.00002211	0.99976700	0.00023300	0.99999357	0.00000643
217	0.99987240	0.00012760	0.99996076	0.00003924	0.99963543	0.00036457	0.99998915	0.00001085
218	0.99999992	0.00000008	0.99999998	0.00000002	0.99999982	0.00000018	0.99999999	0.00000001
222	0.99951874	0.00048126	0.99984866	0.00015134	0.99847180	0.00152820	0.99995989	0.00004011