Syst. Biol. 54(1):35–55, 2005

ISSN: 1063-5157 print / 1076-836X online
DOI: 10.1080/10635150590910249

Is Homoplasy or Lineage Sorting the Source of Incongruent mtDNA and Nuclear Gene Trees in the Stiff-Tailed Ducks (*Nomonyx-Oxyura*)?

KEVIN G. MCCRACKEN[1] AND MICHAEL D. SORENSON[2]

[1]*Institute of Arctic Biology, Department of Biology and Wildlife, and University of Alaska Museum, University of Alaska Fairbanks, Fairbanks, Alaska 99775, USA; E-mail: fnkgm@uaf.edu*
[2]*Department of Biology, Boston University, Boston, Massachusetts 02215, USA*

Abstract.—We evaluated the potential effects of homoplasy, ancestral polymorphism, and hybridization as obstacles to resolving phylogenetic relationships within *Nomonyx-Oxyura* stiff-tailed ducks (Oxyurinae; subtribe Oxyurina). Mitochondrial DNA (mtDNA) control region sequences from 94 individuals supported monophyly of mtDNA haplotypes for each of the six species and provided no evidence of extant incomplete lineage sorting or inter-specific hybridization. The ruddy ducks (*O. j. jamaicensis, O. j. andina, O. j. ferruginea*) are each others' closest relatives, but the lack of shared haplotypes between *O. j. jamaicensis* and *O. j. ferruginea* suggests long-standing historical isolation. In contrast, *O. j. andina* shares haplotypes with *O. j. jamaicensis* and *O. j. ferruginea*, which supports Todd's (1979) and Fjeldså's (1986) hypothesis that *O. j. andina* is an intergrade or hybrid subspecies of *O. j. jamaicensis* and *O. j. ferruginea*. Control region data and a much larger data set composed of ~8800 base pairs of mitochondrial and nuclear sequence for each species indicate that the two New World species, *O. vittata* and *O. jamaicensis*, branch basally within *Oxyura*. A clade of three Old World species (*O. australis*, *O. maccoa*, *O. leucocephala*) is well supported, but different loci and also different characters within the mtDNA data support three different resolutions of the Old World clade, yielding an essentially unresolved trichotomy. Fundamentally different factors limited the resolution of the mtDNA and nuclear gene trees. Gene trees for most nuclear loci were unresolved due to slow rates of mutation and a lack of informative variation, whereas uncertain resolution of the mtDNA gene tree was due to homoplasy. Within the mtDNA, approximately equal numbers of characters supported each of three possible resolutions. Parametric and nonparametric bootstrap analyses suggest that resolution of the mtDNA tree based on ~4,300 bp per taxon is uncertain but that complete mtDNA sequences would yield a fully resolved gene tree. A short internode separating *O. leucocephala* from (*O. australis*, *O. maccoa*) in the best mtDNA tree combined with long terminal branches and substantial rate variation among nucleotide sites allowed the small number of changes occurring on the internode to be obscured by homoplasy in a significant portion of simulated data sets. Although most nuclear loci were uninformative, two loci supported a resolution of the Old World clade (*O. maccoa, O. leucocephala*) that is incongruent with the best mtDNA tree. Thus, incongruence between nuclear and mtDNA trees may be due to random sorting of ancestral lineages during the short internode, homoplasy in the mtDNA data, or both. The *Oxyura* trichotomy represents a difficult though likely common problem in molecular systematics. Given a short internode, the mtDNA tree has a greater chance of being congruent with the history of speciation because its effective population size (N_e) is one-quarter that of any nuclear locus, but its resolution is more likely to be obscured by homoplasy. In contrast, gene trees for more slowly evolving nuclear loci will be difficult to resolve due to a lack of substitutions during the internode, and when resolved are more likely to be incongruent with the species history due to the stochastic effects of lineage sorting. We suggest that researchers consider first whether independent gene trees are adequately resolved and then whether those trees are congruent with the species history. In the case of *Oxyura*, the answer to both questions may be no. Complete mtDNA sequences combined with data from a very large number of nuclear loci may be the only way to resolve such trichotomies. [Ancestral polymorphism; homoplasy; lineage sorting; *Nomonyx*; *Oxyura*; Oxyurinae; stiff-tailed ducks; waterfowl.]

Resumen.—Evaluamos los efectos potenciales de homoplasía, polimorfismo ancestral, e hibridación como obstáculos para resolver las relaciones filogenéticas dentro de los patos buceadores correspondientes a *Nomonyx-Oxyura* (Oxyurinae; subtribu Oxyurina). Las secuencias de las regiones de control de ADN mitocondrial (ADNmt) de 94 individuos, respaldaron la monofilia de haplotipos de ADNmt para cada una de las seis especies y no presentaron ninguna evidencia de selección incompleta latente de los lineajes o de hibridación inter-específica. Los patos buceadores (*O. j. jamaicensis, O. j. andina, O. j. ferruginea*) son entre ellos, sus parientes más cercanos, pero la falta de haplotipos compartidos entre *O. j. jamaicensis* y *O. j. ferruginea* sugieren un largo periodo de aislamiento. Sin embargo, *O. j. andina* comparte haplotipos con *O. j. jamaicensis* y *O. j. ferruginea*, lo que respalda la hipótesis de Todd (1979) y Fjeldså (1986) que *O. j. andina* está integrada o es una subespecie híbrido de *O. j. jamaicensis* y *O. j. ferruginea*. Datos de la región de control y una base de datos mucho más amplia de ~8,800 pares de bases de secuencias mitocondriales y nucleares para cada especie indican que las dos especies del nuevo mundo, *O. vittata* y *O. jamaicensis*, ramifican en la base del árbol filogenético de *Oxyura*. La agrupación de las tres especies del viejo mundo (*O. australis*, *O. maccoa*, *O. leucocephala*) está bien sustentada, sin embargo si se usa diferentes loci y también si se usa diferentes caracteres dentro de los datos de ANDmt se encuentra respaldo para tres soluciones diferentes de este grupo del viejo mundo, dando como resultado una tricotomía esencialmente no resuelta. Hay factores fundamentalmente diferentes que limitan la resolución del ADNmt y de los árboles genéticos nucleares. Los árboles genéticos para la mayoría de los loci nucleares no se resolvieron debido a la baja tasa de mutación y la falta de variación informativa, mientras que la falta de resolución del árbol genético basado en ADNmt fue por causa de homoplasía. Dentro del ADNmt, un número aproximadamente igual de caracteres respaldan cada uno de las tres resoluciones posibles. Un análisis estadístico paramétrico y no-paramétrico de "bootstrap" sugiere que la resolución del árbol basado en ~4,300 pares de bases de ADNmt por taxón es incierto pero que las secuencias completas de ADNmt darían como resultado un árbol genético completamente resuelto. La distancia inter-nodal corta entre *O. leucocephala* y (*O. australis*, *O. maccoa*) en el mejor árbol de ADNmt, además de las ramas terminales largas y de una tasa de variación significativa entre nucleótidos, permitió que el pequeño número de cambios que ocurrieron en el inter-nodo sean ensombrecidos por homplasía en una porción significativa de las bases de datos simuladas. Aún cuando la mayoría de los loci nucleares fueron no informativos, dos loci respaldaron la resolución del grupo del viejo mundo (*O. maccoa, O. leucocephala*) el que es inconsistente con el mejor árbol de ADNmt. Por lo tanto, la inconsistencia entre ADN nuclear y mitocondrial puede deberse a selección al azar de lineajes ancestrales durante el inter-nodo corto, homoplasía en los datos

de ADNmt, o ambos. La tricotomía de *Oxyura* representa un problema difícil pero es probablemente común en la sistemática molecular. Debido al inter-nodo cort, el árbol de ADNmt tiene una probabilidad más grande de coincidir con la historia de especiación porque el tamaño efectivo de su población (N_e) es un cuarto del tamaño de cualquier locus nuclear, pero su resolución es más probable de ser ensombrecida por homoplasía. En contraste, los árboles genéticos de los evolutivamente lentos loci nucleares van a ser difíciles de resolver debido a la falta de substituciones durante el inter-nodo, y cuando son resuletos, es más probable que sean inconsistentes con la historia de la especie debido a eventos estocásticos de la selección de lineajes. Sugerimos que los investigadores resuelvan independientemente primero si los árboles genéticos están adecuadamente resueltos y luego si esos árboles son consistentes con la historia de la especie. En el caso de *Oxyura*, la respuesta a ambos dilemmas podría ser que no. Secuencias de ADNmt completas combinadas con datos con un número muy grande de loci nucleares podría ser el único camino para resolver tricotomías como estas. [Polimorfismo ancestral, homoplasía, selección de lineajes, *Nomonyx*; *Oxyura*; Oxyurinae; partos buceadores, aves acuáticas.] *Traducido por Thomas H. Valqui.*

Lack of resolution in a phylogenetic tree is represented as a polytomy in which three or more descendant lineages diverge from a single node. Systematists generally seek fully resolved trees, which may yield stronger inferences about character evolution or biogeographic history, and usually view polytomies as reflecting uncertainty about relationships. Such "soft" polytomies may be due to insufficient data and often can be resolved by adding more or different kinds of characters (Maddison, 1989; DeSalle et al., 1994). In contrast, a "hard" polytomy represents the simultaneous origin of three or more species (or gene lineages) from a common ancestor and has no bifurcating solution. Instances in which intervals between successive branching events are too short to accumulate informative variation also are effectively hard polytomies (Hoelzer and Melnick, 1994). Operationally, hard polytomies can be identified by internal branch lengths that do not differ significantly from zero (e.g., Walsh et al., 1999). The internal branch of a soft polytomy, however, might also meet this criterion if limited data are available. Thus, one test to discriminate between soft and hard polytomies is to add data and determine if significant internal branch lengths are obtained.

Resolution of a soft polytomy in a species-level phylogeny can be complicated further by incongruence of individual gene trees with the history of speciation (Tajima, 1983; Neigel and Avise, 1986; Avise et al., 1990). Consider, for example, a clade of three species (A, B, C) in which the divergence of B and C follows shortly after the divergence of A from the ancestor of B and C, such that ancestral polymorphism is retained in all three of the descendant lineages. If these events occurred very recently, incomplete lineage sorting may be detected with population level sampling. Over time, however, genetic drift will lead to the monophyly of alleles (or haplotypes) at each locus in each species, but due to the differential assortment of ancestral polymorphisms, species B and C will share closely related alleles at only some loci. By chance, species A and B will share alleles at another set of loci, and species A and C will share alleles at a third set of loci. If the internode (T) is very short (essentially a hard polytomy) and the effective population size (N_e) is large (e.g., $2N_e > T$; Avise, 2000), and the speciation events occurred far enough in the past for lineage sorting to have completed, approximately 33% of gene trees from independent, polymorphic loci should support each of the three possible clades. As the length of the internode increases, however, an increasing proportion of gene trees should be congruent with the species history. Thus, a second test of soft versus hard polytomies is to survey multiple, unlinked loci.

Synapomorphies also can be obscured by homoplasy, which increases with time since speciation and rates of mutation. Homoplasy may be inconsequential for resolving polytomies in cases of very recent speciation, including those where retained ancestral polymorphism is still extant. As the time depth of a soft polytomy increases, however, homoplasy will tend to obscure relationships and increase the length of nucleotide sequence needed to resolve a given gene tree, regardless of whether that gene tree matches the species tree or not. For soft polytomies that occurred farther in the past, homoplasy may overwrite phylogenetic signal, such that each gene tree effectively becomes a hard polytomy with internal branch lengths that do not differ significantly from zero in a statistical framework.

Lineage sorting and homoplasy thus differ in their effects because the potential to resolve a polytomy depends not only on the length of the internode between speciation events but also on the overall time depth of the polytomy (e.g., O'hUigin et al., 2002). A soft polytomy that is one million years old now may be resolved with sufficient data, but 100 million years hence the same polytomy might meet all the criteria for a hard polytomy due to the cumulative effects of homoplasy. Ancestral polymorphism per se eventually disappears as each descendant species becomes fixed for a different set of alleles or haplotypes, but the mismatches between gene trees and species trees that ancestral polymorphism generates persist. We consider the above issues in attempting to resolve the species-level relationships of stiff-tailed ducks (subfamily Oxyurinae).

Stiff-Tailed Ducks

Stiff-tailed ducks traditionally have been classified as a monophyletic group of eight waterfowl species and four genera, three of which are monotypic: (1) black-headed duck (*Heteronetta atricapilla*), (2) masked duck (*Nomonyx dominicus*), (3) five or six *Oxyura* species, and (4) musk duck (*Biziura lobata*) (Delacour and Mayr, 1945; Raikow, 1970; Livezey, 1986, 1995; Johnsgard and Carbonell, 1996). Livezey (1997) restricted the term "stiff-tailed ducks" (subtribe Oxyurina) to *Nomonyx*, *Oxyura*, and *Biziura*, which along with *Heteronetta* compose the tribe Oxyurini ("stiff-tailed ducks and allies"). However, Harshman (1996), Sraml et al. (1996), and McCracken et al. (1999) demonstrated that *Biziura* is not

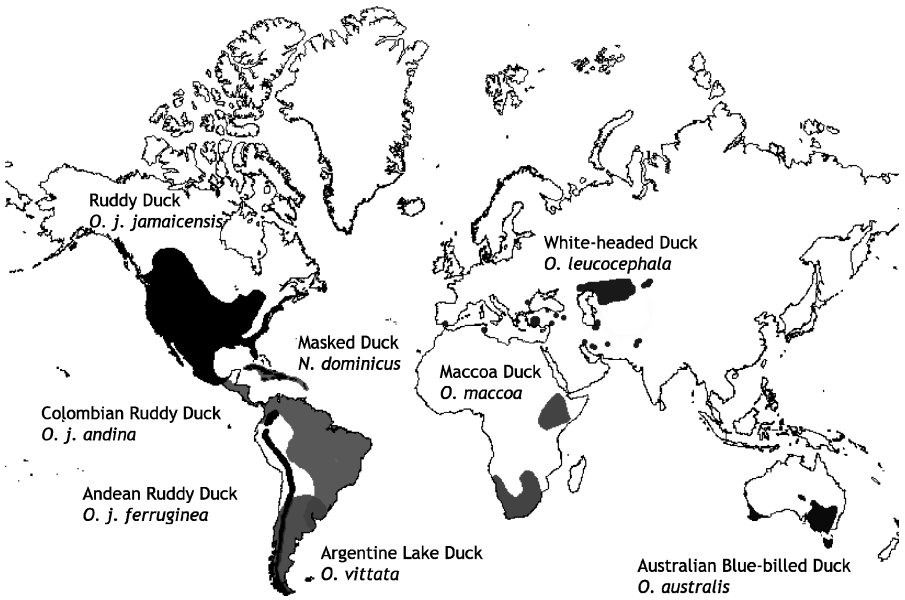

FIGURE 1. Geographic ranges of *Nomonyx-Oxyura* stiff-tailed ducks.

a close relative of the other stiff-tailed ducks or their allies.

Excluding *Biziura*, six species of stiff-tailed ducks (*Nomonyx*, *Oxyura*) currently are recognized. *Nomonyx dominicus* inhabits tropical and subtropical wetlands of the New World and is the sister group of *Oxyura* (McCracken et al., 1999, Sorenson et al., in preparation; but see Livezey, 1995). *Oxyura* species are broadly distributed (Fig. 1), with Argentine lake duck (*O. vittata*) and ruddy duck (*O. jamaicensis*) in the New World; and white-headed duck (*O. leucocephala*), Australian blue-billed duck (*O. australis*), and maccoa duck (*O. maccoa*) in the Old World. The South American ruddy duck subspecies, *O. j. andina* and *O. j. ferruginea*, are endemic to the Andes, *O. j. andina* in Colombia and *O. j. ferruginea* from southern Colombia south to Tierra del Fuego.

Despite three previous morphological and molecular studies (Raikow, 1970; Livezey, 1995; McCracken et al., 1999), species-level relationships within *Oxyura* remain poorly resolved. McCracken et al. (1999) found strong support for *Nomonyx* as the sister group of *Oxyura*, but a small cytochrome *b* data set (1045 bp) and short internal branches (McCracken et al., 1999: fig. 2, 3) yielded weak support for relationships within *Oxyura*. Behavioral and morphological character data (McCracken et al., 1999: fig. 9) also provided no conclusive support. Parsimony analysis of geographical range data supported a clade of Old World species (*O. australis*, *O. maccoa*, *O. leucocephala*) with the New World species (*Nomonyx*, *O. vittata*, *O. jamaicensis*) branching basally, but McCracken et al. (1999) considered the short internodes within *Oxyura* essentially unresolved.

As outlined above, poor resolution of *Oxyura* phylogeny may stem from insufficient character data, the presence of hard polytomies, retained ancestral polymorphism, incongruence among data partitions due to the stochastic nature of the lineage sorting process, or homoplasy. Current or historical hybridization also might complicate resolution of the phylogeny. For example, four *Nomonyx-Oxyura* stiff-tailed ducks are sympatric in South America (Fig. 1). In Colombia, where the ranges of *O. j. andina* and *O. j. ferruginea* abut, *O. j. andina* shows variable amounts of white cheek spotting (Adams and Slavid, 1984; Fjeldså, 1986). Todd (1979) believed *O. j. andina* to be an intergrade, and Fjeldså (1986) suggested that it might represent a relict Pleistocene hybrid population of *O. j. jamaicensis* and *O. j. ferruginea*, with males approaching one or the other phenotype. Livezey (1995) classified *O. j. andina* as a subspecies of *O. jamaicensis*, but classified *O. j. ferruginea* as a separate species. This study is the first to assay genetic diversity of *O. j. jamaicensis, O. j. andina, and O. j. ferruginea*. Hybridization might also occur elsewhere in South America. *O. j. ferruginea* and *O. vittata* occur sympatrically in an ecotone at the base of the southern Andes, *O. j. ferruginea* and *Nomonyx* occur sympatrically in the central Andes, and *Nomonyx* and *O. vittata* occur sympatrically in the lowland subtropics of South America, as do *Nomony*x and *O. j. jamaicensis* in Central America and the Caribbean (Johnsgard, 1965, 1978; Johnsgard and Carbonell, 1996). Elsewhere, geographic ranges of *Oxyura* species are widely separated, except in Europe where *O. j. jamaicensis* and *O. leucocephala* now hybridize in a human-mediated contact zone caused by the escape of captive *O. j. jamaicensis* from North America (Green and Hughes, 1996; Hughes et al., 1999).

To evaluate the potential effects of hybridization, homoplasy, ancestral polymorphism, and lineage sorting on our efforts to resolve the relationships of stiff-tailed ducks, we collected tissues from 94 individuals representing all *Nomonyx* and *Oxyura* species (including *O. j. andina* and *O. j. ferruginea*) and studied their

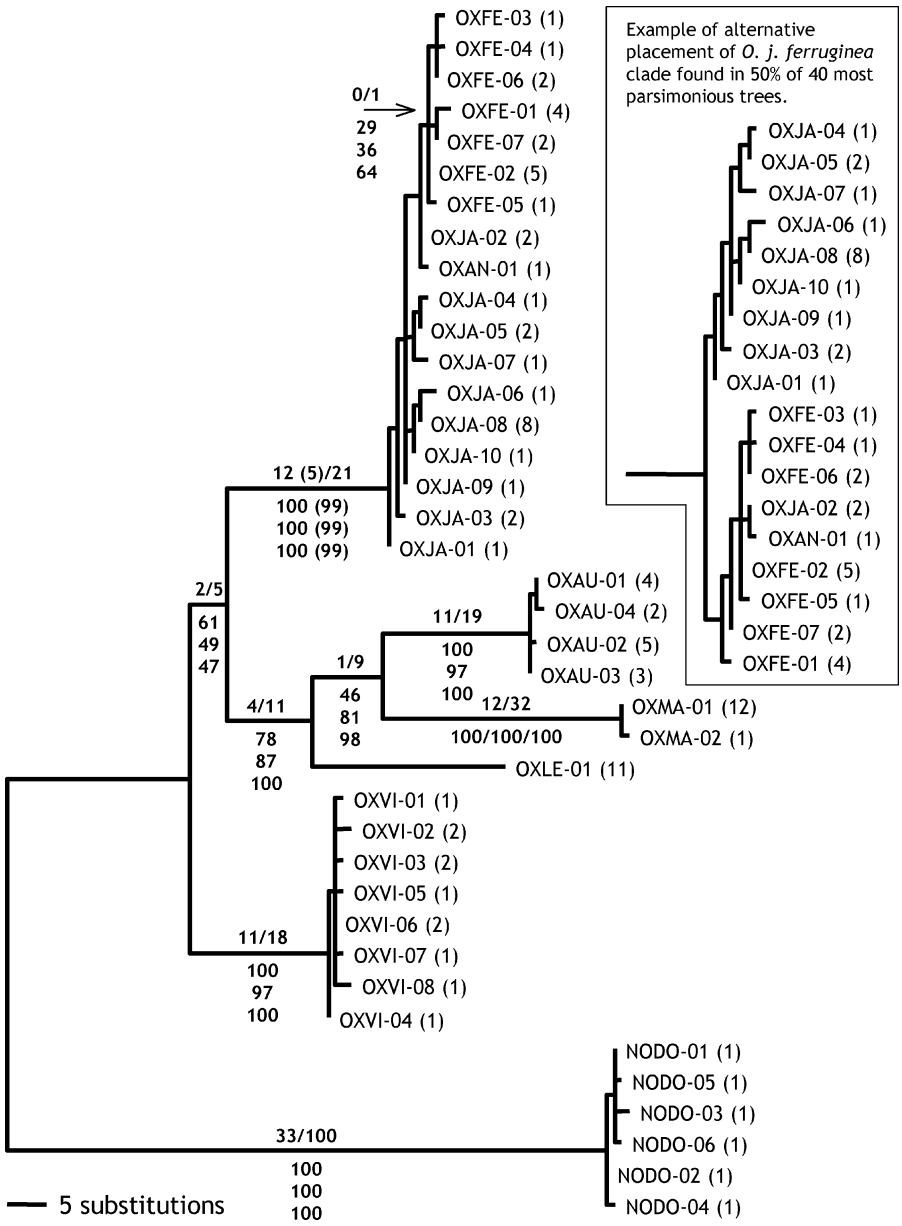

FIGURE 2. One of 40 most parsimonious trees and two equally likely maximum likelihood trees derived from 39 haplotypes and 924 aligned positions of the mtDNA control region and phenylalanine t-RNA genes (length = 279, CI = 0.81, RI = 0.95, −lnL = 2510.41, $\alpha = 0.21$). Decay indices/parsimony branch lengths (above branches) and parsimony bootstrap values/maximum likelihood bootstrap values/Bayesian posterior probabilities (below branches; listed top to bottom) indicate support for clades. Numbers of individuals possessing each haplotype are listed in parentheses. If partial sequences from *O. j. jamaicensis*, *O. j. ferruginea*, and *O. j. andina* museum skins are included in the analysis, the decay index for the ruddy duck clade is 5 steps, and bootstrap values are 99% (alternative decay index and bootstrap values are listed in parentheses).

phylogenetic relationships at two hierarchical levels. (1) We sequenced most of the mitochondrial (mtDNA) control region and part of the phenylalanine t-RNA gene from all 94 samples to test species monophyly and compare genetic divergence within and among species and among subspecies of *O. jamaicensis*. (2) We also sequenced 12S rRNA, ND2, part of ND5, cytochrome *b*, and a number of t-RNA genes from the mtDNA as well as introns and exons from 10 different nuclear loci from one representative of each *Nomonyx-Oxyura* species. We sought to determine if polytomies within *Oxyura* are soft or hard, whether character conflict occurs primarily within or between data partitions, and how much data would be required to resolve the relationships of this group. We also used parametric bootstrapping to evaluate the relationship between data set size and support for alternative arrangements within *Oxyura*.

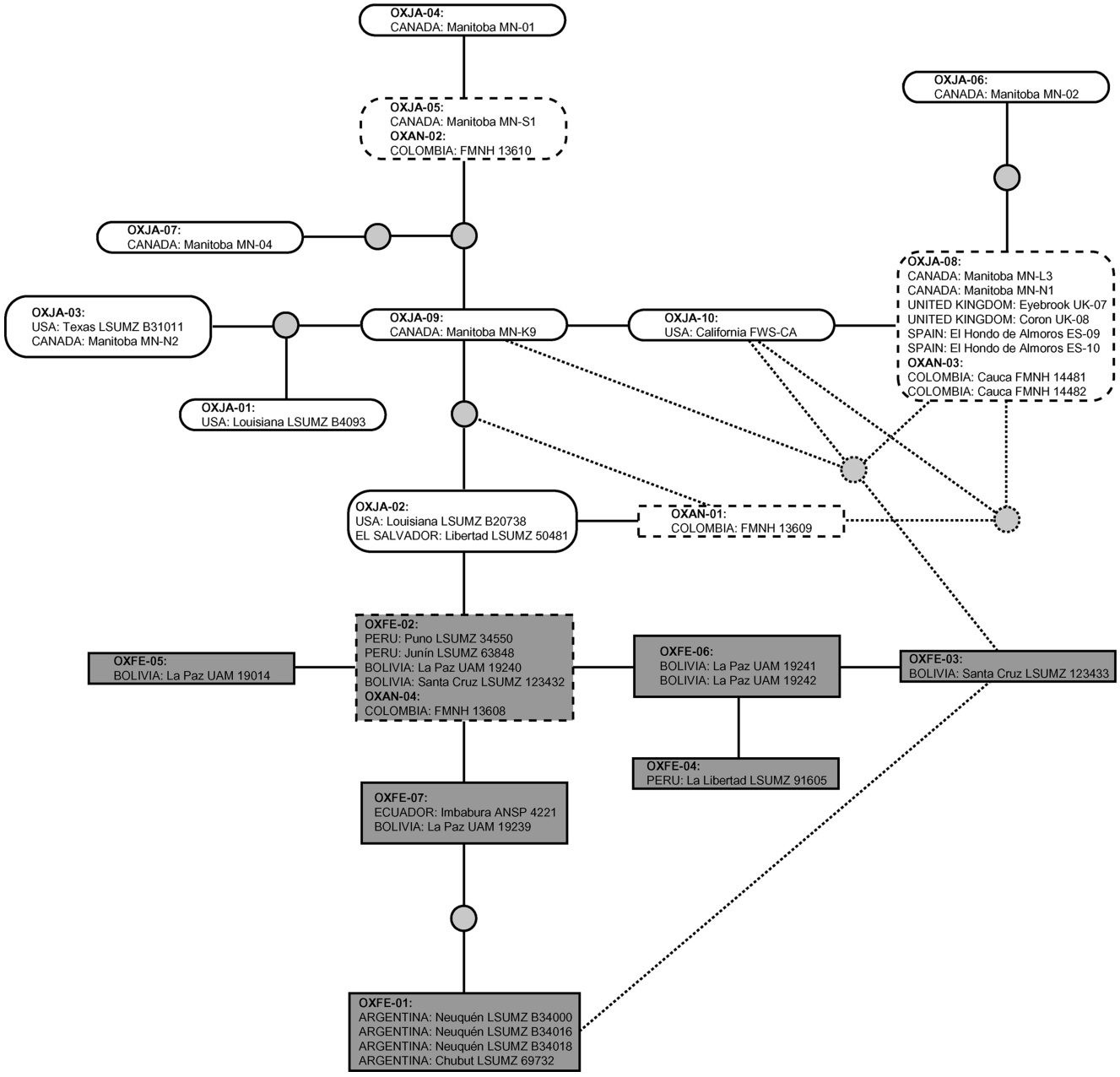

FIGURE 3. Unrooted statistical parsimony tree pruned from Figure 2 illustrating phylogenetic relationships of 18 mtDNA control region and phenylalanine t-RNA haplotypes collected from 37 ruddy ducks, including 17 *O. j. jamaicensis*, 15 *O. j. ferruginea*, and 5 *O. j. andina* (length = 23, CI = 0.78, RI = 0.83). Dashed lines and circles connecting haplotypes show alternative branches and inferred haplotypes included in the 95% set of unrooted parsimony networks. *O. j. jamaicensis* haplotypes are illustrated in white rounded rectangles, and *O. j. ferruginea* (dark gray) and *O. j. andina* (dark gray or white) haplotypes are illustrated in square rectangles. Haplotypes occurring in *O. j. andina* and shared with *O. j. jamaicensis* or *O. j. ferruginea* are illustrated in rectangles with dashed line borders. Each branch represents one step, and solid circles indicate intermediate ancestral haplotypes not sampled.

MATERIALS AND METHODS

Collection, Amplification, and Sequencing

Collecting information for 94 *Nomonyx-Oxyura* genetic samples is provided in Appendix 1. DNA was isolated from muscle, blood, or the base of one or two feather quills using DNeasy Tissue Kits (QIAGEN, Valencia, California). Thirty microliters of 100 mg/mL DTT (dithiothreitol) was added to the digestion buffer to dissolve feather quills. We amplified and sequenced most of the mtDNA control region and part of the phenylalanine t-RNA gene (corresponding to bp 79 to 1250 in the chicken [*Gallus gallus*] mitochondrial genome; Desjardins and Morais, 1990). Our sequences span most of domain I (5′ variable region) and all of domains II (conserved central domain) and III (3′ variable region) of the avian control region (e.g., Marshall and Baker, 1997). Primer sequences used in this study are listed

in Appendix 2. We amplified fresh tissues using the overlapping primer pairs L78–H774 and L736–H1251. We used L78–H493, L432–H774, and NODOCRL1–NODO12SH to amplify the same region from one *Nomonyx* museum skin (LSUMZ 123431) and L78–H493 (or L78–OXJACRH1 and OXJACRAL1–H493) to amplify 5′ control region sequences from three *Nomonyx*, one *O. j. jamaicensis*, five *O. j. andina*, and four *O. j. ferruginea* museum skins.

We also sequenced 12S rRNA, ND2, part of ND5, cytochrome *b*, and parts of the adjacent phenylalanine, valine, methionine, tryptophan, and threonine t-RNA genes (chicken mtDNA genome positions 1268 to 2293, 5217 to 6312, and 14771 to 16063; Desjardins and Morais, 1990) from one individual of each species included in the control region data set (but not including *O. j. andina* or *O. j. ferruginea*). Each region was amplified and sequenced in two or more overlapping fragments (see Appendix 2 for primers used). To this data set we also added introns and exons from ten nuclear loci: chromosome Z chromo-ATPase/helicase DNA-binding protein (CHD-Z), α-enolase (α-ENOL), glyceraldehyde-3-phosphate dehydrogenase (GAPDH), hemoglobin α-A (hemoglobin α-A), lactate dehydrogenase-B (LDHB), lamin A (lamin-A), laminin receptor precursor/p40 (LRP/p40), myelin proteolipid protein (MPP), and phosphoenolpyruvate carboxykinase introns 3 and 9 (PEPCK3, PEPCK9). Hemoglobin α-A sequences included three exons and two introns; all other nuclear sequences are introns with adjacent portions of the flanking exons. Our choice of nuclear loci follows previous studies (e.g., Friesen et al., 1997; Primmer et al., 2002; Pacheco et al., 2002), but we designed our own primers for each locus (Appendix 2) based on sequences for chicken and other vertebrates in GenBank. The LDHB sequence we obtained for *Nomonyx* was only 118 bp, and translation of the sequence indicated complete deletion of intron 3 in this species.

PCR reactions were carried out in GeneAmp PCR System 2400 and 9600 (Applied Biosystems, Foster City, California) and DNA Engine DYAD (MJ Research, Waltham, Massachusetts) thermalcyclers using 50-μL reactions containing 2 μL template DNA, 2.5 μL of each primer (10 μM), 5 μL of 10 μM dNTPs, 5 μL of 25 mM $MgCl_2$, 5 μL of 10$\times$ PCR buffer, and 0.25 μL Taq Polymerase. We used AmpliTaq Gold PCR Master Mix (Applied Biosystems, Foster City, California) for the 10 nuclear loci. Thermal cycling typically was as follows: 7 min preheat at 94°C, followed by 45 cycles of 20 s at 94°C, 20 s at 52°C, 1 min at 72°C, and a final extension of 7 min at 72°C. We used nested PCR for the PEPCK introns. Product from an initial 20 μl PCR using the outer primers PEPCK3.F–PEPCK3.R and GTP1601.F–GTP1793.R was used as the template for a second PCR using the inner primers PEPCK3.F2–PEPCK3.R and PEPCK9.F–PEPCK9.R (Appendix 2). PEPCK3.F2 was paired with the same reverse primer as in the first PCR.

PCR products were electrophoresed in 1% to 2% agarose, excised from the gel, and purified using QIAquick Gel Extraction Kits (QIAGEN, Valencia, California). Both strands of each product were cycle-sequenced in 10-μL reactions using BigDye Terminator Cycle Sequencing Kits diluted fourfold, followed by electrophoresis on ABI 377 or ABI 3100 automated DNA-sequencers (Applied Biosystems, Foster City, California). Sequences from opposite strands were reconciled and verified for accuracy using Sequencher 3.1 (Gene Codes, Ann Arbor, Michigan) and Sequence Navigator 1.0.1 (Applied Biosystems, Foster City, California). Sequences are archived in GenBank: AY747698–AY747869.

Phylogenetic Analysis of Control Region Data

Control region sequences varied in length among species due to insertions and deletions of nucleotides. We aligned control region sequences using direct optimization as implemented in POY 2.7 (Gladstein and Wheeler, 2000). Unlike other alignment algorithms such as Clustal W (Thompson et al., 1994), direct optimization considers sequence alignment and tree topology using a single optimality criterion (i.e., minimizing the number of steps on a tree), and, as such, is a logically consistent approach to phylogenetic analysis of variable length sequences (Wheeler, 1996).

Alignment of control region sequences was performed in two steps. We first used POY 2.7 in an analysis including each unique *Oxyura* haplotype using 100 random addition replicates (each limited to five trees), equal weights for all changes, tree bisection and reconnection (TBR) branch-swapping, and an insertion-deletion cost equal to one. Using the implied alignment generated by POY as a guide, we next optimized the alignment for the shortest tree found by POY using methods described by Sorenson and Payne (2001). Two long and unambiguous insertions observed in *Nomonyx* (see results) were uninformative with respect to relationships within *Oxyura* and were therefore removed from all subsequent analyses. Remaining, generally single-base indels were in included in parsimony analyses by treating gaps as a fifth character state. We used heuristic tree searches with random addition of taxa in PAUP* 4.0 (Swofford, 1998) to find all equally parsimonious trees for the optimized static alignment. Given uncertain resolution of the many closely related haplotypes within species, we reduced the number of trees found by collapsing all branches with a minimum length of zero. Decay indices (Bremer, 1988) were determined using TreeRot (Sorenson, 1999). The control region alignment is available for download in nexus format online at http://mercury.bio.uaf.edu/~kevin_mccracken/alignments/ and http://systematicbiology.org.

We used hierarchical likelihood ratio tests as implemented in Modeltest 3.06 (Posada and Crandall, 1998) to determine the appropriate model of sequence evolution, which for the control region was the Hasegawa et al. (1985) model with gamma-distributed rate variation among sites ($\alpha = 0.20$, ti/tv ratio = 3.41, A = 0.2781, C = 0.3111, G = 0.1513). With these parameter values fixed, we completed 100 full heuristic searches to find maximum likelihood tree(s) for 39 unique control region

haplotypes. We used bootstraping to measure support for clades (Felsenstein, 1985) under both parsimony (1000 replicates, each with 100 random addition replicates limited to 5 trees each) and maximum likelihood (100 replicates with TBR branch-swapping) criteria. Posterior branch probabilities were obtained using MrBayes 3.0 (Huelsenbeck and Ronquist, 2001). We used the HKY+G model and completed two replicate analyses each with four Markov chains of 2.5 million generations. Trees were sampled every 1000 generations, and posterior branch probabilities were calculated after excluding the first 0.5 million generations. Finally, we used statistical parsimony as implemented in TCS (Clement et al., 2000) to illustrate the 95% set of unrooted trees for *O. j. jamaicensis*, *O. j. andina*, and *O. ferruginea*. Maximum likelihood pairwise genetic distances within and between each species were calculated using PAUP* 4.0 (Swofford, 1998).

Phylogenetic Analysis of Additional mtDNA and Nuclear Data

Sequence data from one representative of each *Nomonyx-Oxyura* species were assembled to generate a larger mtDNA plus nuclear data set that included portions of twenty genes from ten different linkage groups: mtDNA (control region, tRNA-Phe, 12S rRNA, tRNA-Val, tRNA-Met, ND2, tRNA-Trp, ND5, cytochrome *b*, tRNA-Thr), CHD-Z, α-ENOL, GAPDH, hemoglobin α-A, LDHB, lamin-A, LRP/p40, MPP, and PEPCK introns 3 and 9. Alignments of all gene regions except the control region (see above) were unambiguous and made by eye. Multiple base indels in the control region (69 bp and 123 bp), LDHB (466 bp, 8 bp, and 3 bp), LRP/p40 (4 bp), and MPP (8 bp) were coded as missing data, and new binary characters for each unique gap (0 = absent, 1 = present) were added to the end of the data matrix. All other gaps (1 to 3 bp in length) were coded as a fifth character state. A total of 8608 characters were derived from 8801 aligned positions. Double peaks in nuclear gene sequences, reflecting heterozygous positions, were coded with IUPAC degeneracy codes and were treated as polymorphisms. All trees were rooted with *Nomonyx* as the outgroup.

We used an exhaustive tree search and unweighted parsimony to find the most parsimonious tree(s) for the combined data set and explore suboptimal trees. We identified characters supporting three possible branching patterns among the Old World *Oxyura* species, *O. australis*, *O. maccoa*, and *O. leucocephala*, and compared these characters in terms of sequence position, substitution types, numbers of steps, and consistency indices.

We also completed maximum likelihood and Bayesian analyses on the combined mitochondrial data set (4335 characters) using both a single model of sequence evolution fitted to the entire data set or mixed models in which separate parameters were estimated for each of several different data partitions (see results). The fit of mixed models was assessed using the Akaike information criterion (AIC) (Akaike, 1973), and alternative tree topologies were compared using the approximately unbiased test of Shimodaira (2002). As described for the control region analysis above, the Bayesian analyses were replicated and each was run for 2.5 million generations, of which the first 0.5 million were discarded before determining posterior branch probabilities.

Finally, we used parametric bootstrapping (e.g., Saitou and Nei, 1986) to measure the power of the mtDNA data set to resolve relationships among three Old World *Oxyura* species in the mtDNA haplotype tree if, as indicated by our analysis, a short internode separates *O. leucocephala* from *O. maccoa* and *O. australis*. Using 4070 bp of aligned mtDNA sequence for which there were no gaps or missing data in any taxa, we estimated maximum likelihood parameters for the GTR+I+G model on the most parsimonious mtDNA tree using PAUP* 4.0 (Swofford, 1998). We then used Seq-Gen 1.2.6 (Rambaut and Grassly, 1997) to simulate the evolution of sequences on the same tree with branch lengths and model parameters estimated from the empirical data. We generated 500 data sets for each of four sequence lengths: 2 kb, 4 kb, 8 kb, and 16 kb. Simulated data were analyzed using maximum likelihood under the same model of sequence evolution, and the proportion of replicate data sets yielding each of the nodes in the original tree was determined. We also determined the "expected" distribution of parsimony bootstrap values for each of the three possible resolutions of the trichotomy by conducting a complete bootstrap analysis with 100 replicates and branch and bound tree searches for each of the simulated data sets.

RESULTS

Control Region Sequence Diversity

Excluding within-individual length variation in *Nomonyx*, we found 39 unique haplotypes composed of 900 to 1097 nucleotides in 94 individuals. Two large nucleotide insertions (69 bp at positions 152 and 15 to 123 bp at positions 951) occurred in *Nomonyx*. Four *Nomonyx* individuals collected in Argentina apparently possessed multiple control region copies that varied in the length of the insertion at position 951 ranging from 15 to 73 bp. Multiple bands generated by PCR amplification of this region were isolated from agarose gels and sequenced, yielding six clean sequences from four individuals; three had 15-bp insertions at position 951, and three had 73-bp insertions. Although PCR artifact is a possible explanation for this length variation, the same results were produced repeatedly using four different overlapping primer combinations (L78–NODO12SH, L78–H2294, NODOCRL1–NODO12SH, NODOCRL1–H2294). Additional length heteroplasmy near position 951 might very well exist in some *Nomonyx* as we were not able to sequence all of the bands generated by PCR amplification of this region. Excluding these two insertions in *Nomonyx*, the alignment was 924 positions, of which 207 (22.4%) varied among haplotypes and 188 (20.3%) were parsimony informative. Among these, transitions occurred at 131 (63.3%) positions, transversions occurred at 57 (27.7%) positions, and indels occurred at 37 (17.9%) positions. As in other birds, the distribution

TABLE 1. Number of alignment positions, variable and informative positions, %GC, consistency index, rescaled consistency index, best-fit models, uncorrected percent sequence divergence, and transition:transversion ratios for mtDNA loci.

Locus	Alignment positions	No. variable/ informative positions[a]	%GC	CI	RC	Best fit model (AIC)	Uncorrected % sequence divergence[a]	Ti/tv ratio[b]
Control region, tRNA-Phe	924	113/30 (196/48)	46.4	0.87	0.40	K81uf+I	3.2–6.7 (11.1–12.3)	2.40
tRNA-Phe, 12S rRNA, tRNA-Val	1033	51/13 (82/23)	48.7	0.83	0.25	GTR+I	1.4–2.3 (4.2–5.1)	7.49
tRNA-Met, ND2, tRNA-Trp	1093	122/28 (189/58)	48.5	0.79	0.22	TrN+I+G	3.8–6.0 (11.0–11.5)	13.73
ND5, cytochrome b, tRNA-Thr	1285	140/34 (196/55)	48.4	0.81	0.27	TrN+I	4.1–6.0 (9.5–9.8)	7.52

[a]Ingroup (ingroup and outgroup).
[b]Estimates for HKY85 model.

of variable sites across the stiff-tailed duck control region showed considerable heterogeneity: 35.6% of substitutions and indels occurred within the first 341 base pairs of the 5′ end of the alignment, and 61.2% occurred within 558 base pairs of the 3′ end. Only 3.2% of substitutions and indels occurred in the intervening 217 base pairs. These divisions correspond to domains I, II, and III of Marshall and Baker (1997). Four conserved sequenced motifs reported for birds (Desjardins and Morais, 1990; Quinn and Wilson, 1993; Marshall and Baker, 1997; Crochet and Desmerais, 2000) also were observed in our sequences: F box at positions 227 to 254; D box at positions 333 to 357; C box at positions 382 to 408; and CSB-1 at positions 690 to 716 in the 924-position alignment excluding the two large gap regions in *Nomonyx*. One to 10 unique haplotypes were observed within each species or subspecies. Intra- and interspecific pairwise genetic distances are summarized in Table 1. Variation within species, identical sequences in overlapping regions amplified using different combinations of primers, and matching sequences from muscle, feather, and blood samples support the conclusion that our sequences are not nuclear copies (see Sorenson and Quinn, 1998).

Control Region Haplotype Relationships

Optimization alignment and unweighted parsimony analysis of 39 unique *Nomonyx-Oxyura* haplotypes produced 40 most parsimonious trees (length = 279, CI = 0.81, RI = 0.95). Interspecific relationships were as follows in 100% of the 40 most parsimonious trees as well as in two maximum likelihood trees and the Bayesian consensus tree: *Nomonyx* (*O. vittata* (*O. jamaicensis* (*O. leucocephala* (*O. australis*, *O. maccoa*)))). One of two equally likely trees (lnL = −2510.42) is illustrated in Figure 2 (the two maximum likelihood trees were identical, respectively, to 2 of the 40 most parsimonious trees). Monophyly of *Nomonyx* and monophyly of each *Oxyura* species was strongly supported (97% to 100% bootstrap values or posterior branch probabilities). There is no evidence of incomplete lineage sorting or hybridization among species: internal branches between species were long compared to branches within species. Branch support for interspecific relationships was more varied, but the control region data yielded generally higher support indices than cytochrome *b* (see McCracken et al., 1999: fig. 2) as well as a more stable topology. In the present analysis, three Old World species (*O. maccoa, O. australis*, *O. leucocephala*) form a clade nested within New World lineages (*Nomonyx*, *O. vittata*, *O. jamaicensis*), reconfirming a New World origin for this group.

Oxyura j. jamaicensis, O. j. andina, *and* O. j. ferruginea

Two different patterns of *O. jamaicensis* haplotype relationships occurred in the 40 most parsimonious trees. Twenty of 40 trees (50%), including both maximum likelihood trees (Fig. 2), included a monophyletic *O. j. ferruginea* nested within *O. jamaicensis*. Bootstrap support for *O. j. ferruginea* monophyly was low (29%). The remaining 20 trees (50%) had one *O. j. jamaicensis* haplotype (OXJA-02) and one *O. j. andina* haplotype (OXAN-01) nested within the *O. j. ferruginea* clade. *O. j. jamaicensis* was paraphyletic in all 40 trees. An unrooted parsimony tree showing the 95% statistical parsimony set (Clement et al., 2000) of possible haplotype relationships within *O. jamaicensis* is illustrated in Figure 3. Uncertainty about relationships, reflected by multiple connections between six haplotypes in this network, is due in part to partial sequence data for museum specimens. Two results are worth noting: (1) Four *O. j. ferruginea* from southern Argentina are identical and differ by two nucleotide substitutions from eleven *O. j. ferruginea* collected in Ecuador, Perú, and Bolivia; and (2) five *O. j. andina* collected from Colombia show four different haplotypes. Three *O. j. andina* haplotypes are identical to *O. j. jamaicensis* haplotypes, one is identical to an *O. j. ferruginea* haplotype found in Perú and Bolivia, and one is unique.

Combined Analyses of Additional mtDNA and Nuclear Data

To further test interspecific relationships within *Oxyura*, we combined control region data with three additional mtDNA regions. Sequence variation in cytochrome *b* and ND2 were comparable to the control region in terms of number of variable and informative sites and average genetic distances between species, whereas 12S rRNA showed lower levels of variation (Table 1). Consistency and rescaled consistency indices were highest for the control region. This may be due in part to somewhat lower rate variation among sites in

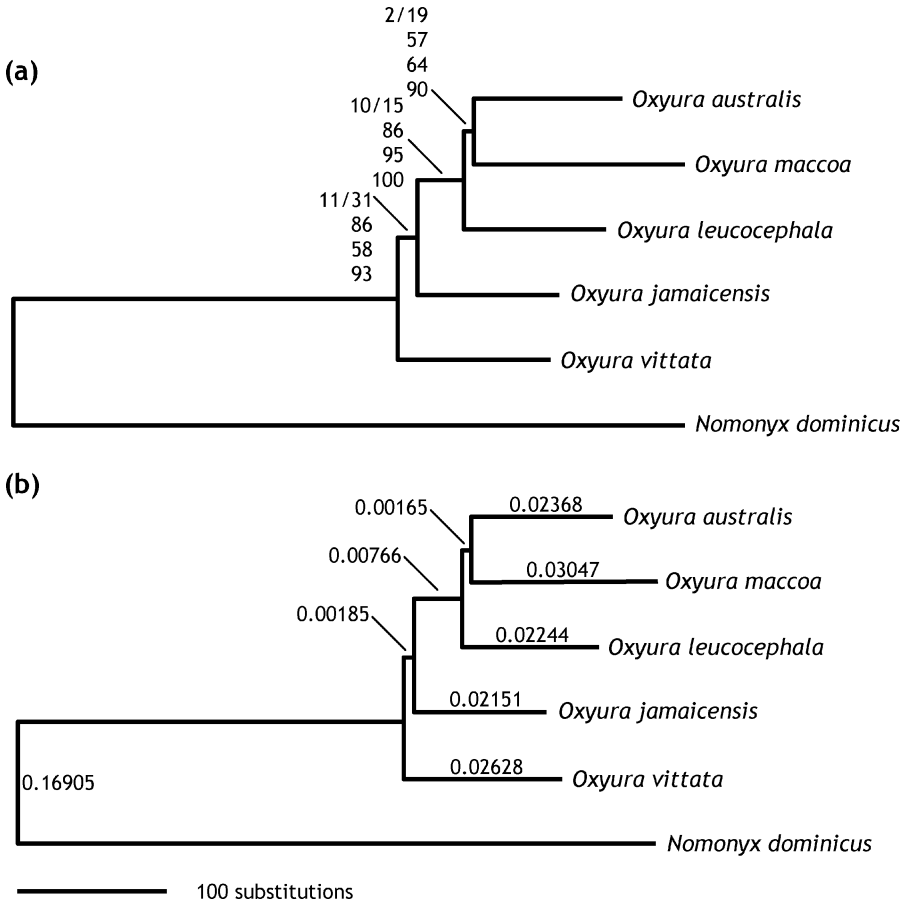

FIGURE 4. (a) *Nomonyx-Oxyura* gene tree based on combined data analysis of 4335 mtDNA positions (lnL = −9600.54). The same tree was found with parsimony, maximum likelihood, and Bayesian analyses. Decay indices/parsimony branch lengths, parsimony bootstrap values, maximum likelihood bootstrap values, and Bayesian posterior probabilities (listed top to bottom) indicate support for nodes. Maximum likelihood branch lengths are illustrated using the same scale as (b). (b) Maximum likelihood tree with branch lengths used for parametric bootstrap analyses of 4070 mtDNA positions for which there were no gaps or missing data in any taxa (lnL = −9077.89). Parameter estimates used to simulate sequence evolution and analyze the resulting data sets were set as follows. Base frequencies: A = 0.29, C = 0.32, G = 0.16, T = 0.23. Relative transformation rates: A-C = 3.17, A-G = 50.36, A-T = 1.16, C-G = 1.37, C-T = 28.84, G-T = 1.00. Proportion of invariant sites: I = 0.73. Gamma shape parameter for the rate distribution: $\alpha = 1.72$.

the control region but perhaps more significantly due to a lower transition:transversion ratio than in the other genes (Table 1). With higher transition:transversion ratios, multiple substitutions at a site are more likely to result in homoplasy.

Combined analysis of the entire mtDNA data set resulted in the same species-level topology as in the control region analysis but with increased support indices at most nodes, suggesting generally congruent phylogenetic signal among data partitions (compare Fig. 2 and Fig. 4a; lnL = −9600.54; Tables 2, 3). Decay indices increased from 2, 4, and 1 to 11, 10, and 2, respectively, for nodes within *Oxyura*. The maximum likelihood bootstrap value and Bayesian posterior branch probability were slightly lower in the combined analysis only for the node uniting *O. australis* and *O. maccoa*. This sister relationship and the same overall tree topology was supported in a combined parsimony analysis of the three protein-coding genes (ND2 and ND5/cytochrome *b*), whereas the 12S rRNA data partition supported a sister relationship between *O. leucocephala* and *O. australis*. Mixed models with parameters estimated independently for different data partitions provided a better fit to the data as measured by the AIC but did not change phylogenetic inferences. The same topology found in the combined parsimony analysis was found in both maximum likelihood and Bayesian analyses using mixed models (Figs. 4, 5). Although all mtDNA data partitions contributed to the two basal nodes in the *Oxyura* phylogeny, likelihood scores suggest that the control region data were most discriminating with respect to relationships of the three Old World *Oxyura*. The approximately unbiased test of Shimodaira (2002) indicated that no data partitions or mixed models discriminated significantly among the three possible resolutions of the trichotomy, except for a marginally significant rejection of *O. maccoa* and *O. leucocephala* for the control region data (Table 2).

Combined data analysis of 8608 characters from 20 mtDNA and nuclear genes yielded 196 informative characters and two equally parsimonious trees

TABLE 2. Log-likelihoods for best-fit models selected using AIC and partitioned by codon position and locus for three possible resolutions of the *Oxyura* trichotomy.

	Best-fit model	(AU, MA)	(MA, LE)	(LE, AU)	Parameters	AIC	*P*-values[a]
1st position	HKY+I	**1388.17**	**1388.17**	**1388.17**	14	2804.35	0.44
2nd position	HKY	**1038.19**	**1038.19**	**1038.19**	13	2102.38	—
3rd position	TrN+G	**2271.24**	**2271.24**	**2271.24**	15	4572.49	0.46
Control region	K81uf+G	**2205.47**	2210.34	2209.41	15	4440.94	0.04
12S/tRNAs	GTR+I	2057.97	2057.65	**2057.23**	18	4150.46	0.12
ND2	TrN+I+G	**2376.94**	2376.95	2376.95	16	4785.89	0.50
ND5/cyt *b*	TrN+I	**2815.60**	**2815.60**	**2815.60**	15	5661.20	0.48
Total mtDNA data combined	TIM+I	**9600.54**	9601.90	9601.61	16	19233.07	0.17
Mixed 1 (CR, 12S, ND2, ND5/cyt *b*)		**9455.99**	9460.54	9459.19	64	19039.98	0.08
Mixed 2 (1, 2, 3, CR, 12S)		**8961.05**	8965.59	8964.24	75	**18070.61**	0.07

[a]Minimum *P*-values for the approximately unbiased test of Shimodaira (2002). For the control region, only (MA, LE) is excluded; (AU, MA) and (LE, AU) do not differ significantly ($P > 0.20$). Too few variable sites were available to perform the test for 2nd positions.

(length = 1013, CI = 0.84, RI = 0.34; Fig. 5). In one tree, *O. australis* is sister to a clade composed of *O. maccoa* and *O. leucocephala*, whereas the other tree has *O. leucocephala* sister to *O. maccoa* plus *O. australis*. The next most parsimonious tree was two steps longer (length = 1015, CI = 0.84, RI =0.33; tree not shown) and included the third possible resolution of the above trichotomy. All three trees had the same basal relationships: (*Nomonyx* (*O. vittata*, (*O. jamaicensis*, (*O. australis*, *O. maccoa*, *O. leucocephala*)))).

Bootstrap support for the three possible resolutions of the Old World clade differed between the combined analysis and separate analyses of mtDNA and nuclear data (Table 3). In the combined analysis, none of the three clades was well supported: 41% for (*O. maccoa*, *O. leucocephala*), 35% for (*O. australis*, *O. maccoa*), and 20% for (*O. australis*, *O. leucocephala*). Based on mtDNA only, bootstrap support for these three alternative clades was 57% for (*O. australis*, *O. maccoa*), 35% for (*O. australis*, *O. leucocephala*), and 0% for (*O. maccoa*, *O. leucocephala*). In contrast, the nuclear data set yielded strong bootstrap support (92%) for (*O. maccoa*, *O. leucocephala*) and essentially zero support for the two alternative clades. This result reflects primarily the lack of characters supporting the alternative clades.

TABLE 3. Parsimony and maximum likelihood bootstrap values and posterior probabilities for *Oxyura* clades supported by different mtDNA data partitions.

	(JA, LE, AU, MA)	(MA, AU, LE)	(AU, MA)[a]	(MA, LE)[a]	(LE, AU)[a]
Control region—39 taxa					
Parsimony	61	78	46	0	**50**
ML full heuristic	49	87	**81**	0	8
Bayesian	47	100	**98**	0	1
Six taxa—parsimony					
Control region	51	68	28	2	**68**
12S/tRNAs	71	32	0	23	**55**
ND2	17	0	0	0	0
ND5/cyt *b*	31	64	19	**47**	10
All	86	86	**57**	0	35
All non-CR	68	46	23	**29**	9
Six taxa—maximum likelihood					
CR	47	84	**70**	0	18
All	58	95	**64**	0	31
All non-CR	18	49	8	0	**23**
Six taxa—Bayesian mixed models					
Codon positions + CR + 12S	93	100	**90**	0	9
CR + 12S + ND2 + cyt *b*	97	100	**92**	1	8
Codon positions + 12S	63	72	13	27	**36**
12S + ND2 + cyt *b*	60	63	19	17	**35**

[a]Bootstrap values in bold type indicate maximum value for three possible resolutions of the *Oxyura* trichotomy for each analysis.

In parsimony analyses with heterozygous positions treated as polymorphisms, five characters in three different introns supported (*O. maccoa*, *O. leucocephala*), whereas no nuclear characters supported either of the alternative resolutions of the Old World trichotomy. For the α-ENOL intron, however, the inference changes when the data for *O. jamaicensis* are coded as two separate sequences that differ at a single position in the diploid sequence. When this is done, *O. jamaicensis* shares one allele with *O. vittata* and *O. australis* and another allele with *O. maccoa*. The sequence in *O. leucocephala* is unique but shares a derived change with *O. maccoa* and one of the two *O. jamaicensis* alleles. In total, only 14 parsimony informative characters were obtained from 4274 aligned positions of nuclear sequence, and five loci (LDHB, LRP/p40, lamin-A, GAPDH, PEPCK3) included no variation that was informative about relationships at any level within *Oxyura*. At 8 of 10 nuclear loci, two or more *Oxyura* species shared identical sequences, suggesting that retention of ancestral sequences was due primarily to low mutation rates. Only PEPCK9 and MPP yielded fully resolved gene trees within *Oxyura*. These trees agreed in supporting (*O. maccoa*, *O. leucocephala*) but differed in other respects (Fig. 6).

Table 4 lists individual characters that support each of the three possible Old World clades. Interestingly, character support for (*O. maccoa*, *O. leucocephala*) was distributed among eight genes, including three nuclear introns, whereas character support for the other two clades was distributed among three and four genes, all mitochondrial. Thirteen of 29 characters (45%) supporting (*O. australis*, *O. maccoa*) or (*O. australis*, *O. leucocephala*) were in the control region, whereas only 1 character supporting (*O. maccoa*, *O. leucocephala*) was in the control region.

Nearly equal numbers of mitochondrial characters supported (*O. australis*, *O. maccoa*) ($n = 15$) and (*O.

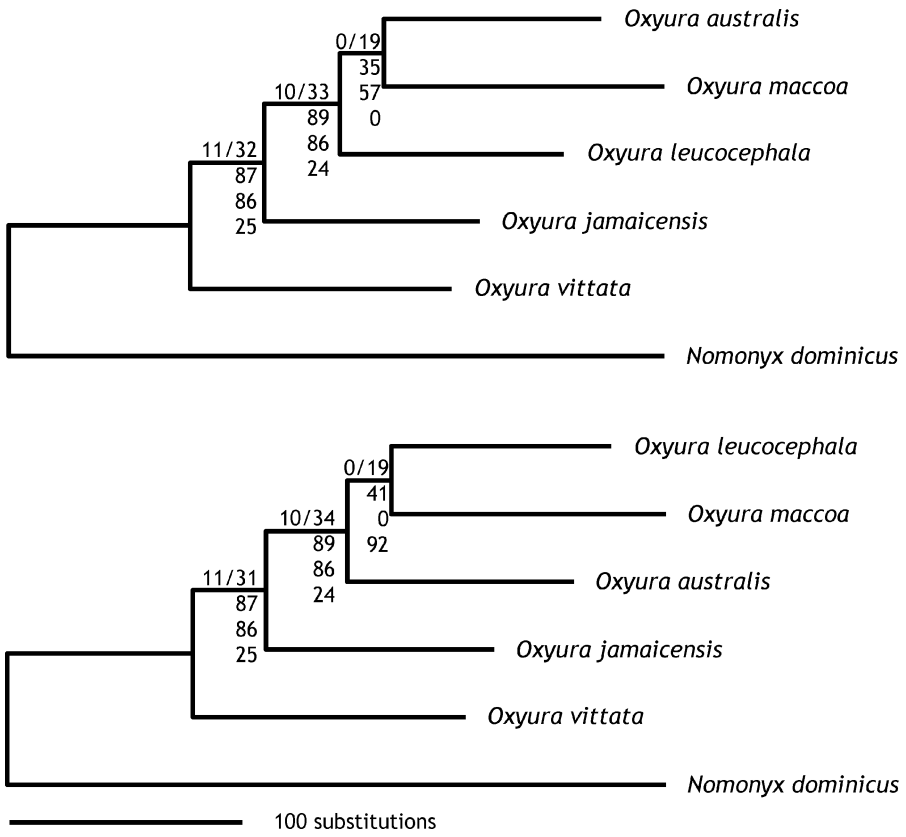

FIGURE 5. Two most parsimonious trees based on combined mtDNA and nuclear data analysis of 8608 characters showing two possible arrangements of *O. australis*, *O. maccoa*, *O. leucocephala* (length = 1013, CI = 0.84, RI = 0.34). Decay indices/branch lengths (above branches) and bootstrap values for combined, mtDNA only, and nuclear DNA data only (below branches; listed top to bottom) indicate support for clades.

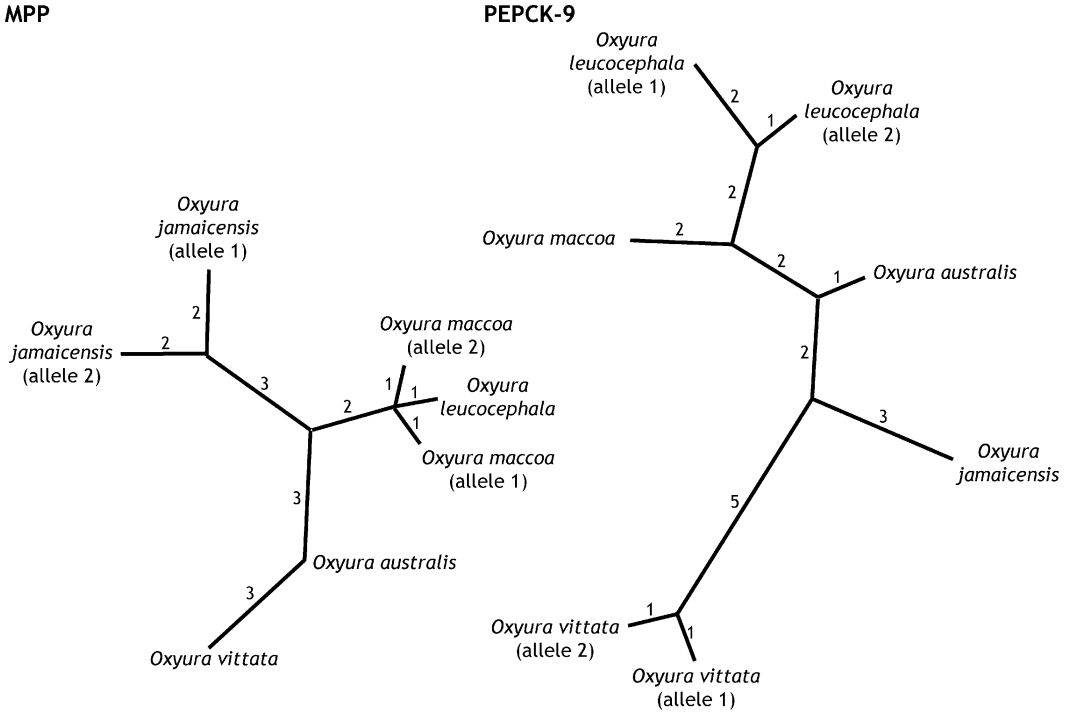

FIGURE 6. Unrooted parsimony trees for MPP and PEPCK-9 introns.

TABLE 4. Synapomorphic character state changes in the *Nomonyx-Oxyura* combined data set supporting three possible Old World stiff-tailed duck branching patterns.

Position[a]	Locus	Substitution	Type	Steps	Consistency index
(MA, LE)					
94	MPP	A → G	TI	1	1[b]
163		G → A	TI	1	1[b]
215		C → G	TV	1	1[b]
1334	α-ENOL	G → A	TI	2	1
4067	PEPCK9	C → T	TI	1	1
4475	12S rRNA	A → G	TI	1	1
5820	ND2	A → G	TI, 3rd pos.	2	1
5844		A → G	TI, 3rd pos.	2	0.5
6406	ND5	A → G	TI, 3rd pos.	1	1
7004	Cytochrome *b*	A → G	TI, 3rd pos.	1	1
7007		A → G	TI, 3rd pos.	2	0.5
7307		G → A	TI, 3rd pos.	2	0.5
7427		A → G	TI, 3rd pos.	1	1
7605		C → T (Arg → Cys)	TI, 1st pos.	2	0.5
8597 (912)	Control region	A → G	TI	1	1
Mean ± SD					0.86 ± 0.25
(AU, MA)					
5637	ND2	A → G	TI, 3rd pos.	2	0.5
5805		A → G	TI, 3rd pos.	1	1
5877		T → C	TI, 3rd pos.	2	0.5
5883		T → C	TI, 3rd pos.	2	1
6054		C → T	TI, 3rd pos.	2	0.5
6941	Cytochrome *b*	A → G	TI, 3rd pos.	1	1
7163		C → T	TI, 3rd pos.	2	0.5
7265		T → C	TI, 3rd pos.	1	1
7517		C → T	TI, 3rd pos.	2	0.5
7769 (84)	Control region	G → A	TI	2	0.5
8258 (573*)		C → A	TV, stem	1	1
8259 (574*)		T → C	TI, loop	1	1
8378 (693*)		A → G	TI, stem	2	0.5
8472 (787*)		C → T	TI, stem	2	1
8481 (796*)		G → A	TI, loop/stem	2	0.5
Mean ± SD					0.73 ± 0.25
(AU, LE)					
4402	12S rRNA	C → T	TI, loop	1	1
4995		C → T	TI	1	1
5583	ND2	T → C	TI, 3rd pos.	2	0.5
6234		T → C	TI, 3rd pos.	2	0.5
6558	Cytochrome *b*	A → G (Ile → Val)	TI, 1st pos.	2	0.5
7358		T → C	TI, 3rd pos.	2	0.5
7505		A → G	TI, 3rd pos.	1	1
7922 (237)	Control region	T → C	TI	2	0.5
7927 (242)		C → T	TI	1	1
7937 (252)		T → –	Insertion	1	1
8412 (727*)		C → T	TI, loop	1	1
8449 (764*)		A → G	Ti, stem	1	1
8509 (824)		A → G	TI	1	1
8770 (1085)		T → –	Insertion	1	1
Mean ± SD					0.82 ± 0.25

[a]Underlined positions show no homoplasy. Numbers in parentheses correspond to positions in the control region alignment, which numbered 1116 positions with the two *Nomonyx* insertions and 924 positions without these insertions.

[b]Not informative in the context of the six taxon analysis because data for *Nomonyx* are missing.

australis, *O. leucocephala*) ($n = 14$), whereas somewhat fewer characters ($n = 10$) supported (*O. maccoa*, *O. leucocephala*). A single transversion in the control region supported (*O. australis*, *O. maccoa*); all other changes were transitions. In protein-coding genes, all but two characters were in third codon positions. The overall consistency of characters supporting the three alternative clades was on average slightly lower for the set of characters supporting (*O. australis*, *O. maccoa*). This issue was somewhat difficult to evaluate, however, in that five of the seven control region characters supporting (*O. australis*, *O. leucocephala*) were in portions of the control region that vary in length among taxa, such that homologous nucleotide positions in somewhat more divergent taxa (e.g., *Stictonetta*, *Heteronetta*) could not be identified. This prevented us from obtaining better estimates of character consistency by evaluating each character across a broader sample of waterfowl taxa.

TABLE 5. Parametric bootstrap values showing the proportion of simulated 2 kb, 4 kb, 8 kb, and 16 kb mtDNA data sets in which the clade was present in the maximum likelihood tree (Fig. 4b).

	Size of simulated data set			
Clade	2 kb	4 kb	8 kb	16 kb
(JA, AU, MA, LE)	40.6	54.9	59.0	73.3
(AU, MA, LE)	89.9	97.8	99.8	100
(AU, MA)	58.3	72.0	86.2	97.2
(AU, LE)	13.1	11.4	6.7	0.6
(MA, LE)	15.3	11.0	5.6	1.9

Parametric Bootstrap Analysis of the mtDNA Data

The mtDNA maximum likelihood tree used to estimate branch lengths for the parametric bootstrap analyses is shown in Figure 4b. If this is the true haplotype tree and the actual branch lengths were as reconstructed, there is only a 72% chance of recovering the correct topology for the three Old World taxa with 4 kb of sequence data (Table 5). The proportion of data sets in which incorrect nodes are recovered declines with the size of the simulated data set until with 16-kb sequences less than 3% of data sets yield an incorrect resolution of the Old World trichotomy (Table 5). Although the (*O. australis*, *O. maccoa*) node is present in most trees derived from the simulated data sets, standard bootstrap values for this node vary greatly among data sets (Fig. 7). Strong bootstrap support for the correct node ($\geq$70%) increases steadily with data set size (9%, 20%, 33%, and 64% of 2-, 4-, 8-, and 16-kb data sets, respectively). Greater than 50% bootstrap support for an incorrect resolution of the trichotomy, however, is obtained in about 15% of the 2- to 8-kb data sets and exceeds 70% in 3% to 4% of data sets. With 16 kb, however, less than 1% of data sets yield $\geq$70% support for an incorrect node.

DISCUSSION

Systematic Relationships and Historical Biogeography

Monophyly of all *Nomonyx-Oxyura* species was robustly supported by control region data with bootstrap values ranging from 97% to 100% (Fig. 2). Internal branches between species and branches leading to each species clade were long compared to branches within each species. We found no evidence in the mtDNA data of incomplete lineage sorting or inter-specific hybridization based on the individuals sampled in this study.

Analyses of the control region and combined data sets (Figs. 2, 4, 5) agree on two points: *Nomonyx* is substantially diverged from *Oxyura*, and the two New World *Oxyura* species, *O. vittata* and *O. jamaicensis*, are basal to the three Old World species, *O. australis*, *O. maccoa*, and *O. leucocephala*. The ancestral geographic range of *Nomonyx-Oxyura* probably was lowland South America because most basal stiff-tailed ducks, *Nomonyx* and *O. vittata* (and also *Heteronetta*), are endemic to lowland South America, Central America, or the Caribbean.

The ruddy ducks, *O. j. jamaicensis*, *O. j. andina*, *and O. j. ferruginea*, are each others' closest relatives (Figs. 2, 3). Like most authors, we consider these three taxa subspecies based on their close genetic, morphological, and behavioral similarities (Johnsgard, 1965; Todd, 1979; Adams and Slavid, 1984; Fjeldså, 1986; Johnsgard and Carbonell, 1996; McCracken et al., 1999). In contrast to Livezey (1995), who placed *O. ferruginea* as the sister group of *O. vittata* and *O. australis*, our analyses show clearly that it is very closely related to other ruddy ducks. Monophyly of *O. j. ferruginea* mtDNA haplotypes also was supported, albeit with weak support. The lack of shared haplotypes between *O. j. jamaicensis* and *O. j. ferruginea*, however, suggests a strong degree of historical isolation between these subspecies. The *O. j. ferruginea* haplotype group was either nested within a larger clade of *O. j. jamaicensis* haplotypes or together with one *O. j. jamaicensis* and *O. j. andina* haplotype formed a sister group to all other *O. jamaicensis* haplotypes—both results were found among the set of 40 most parsimonious control region trees (Figs. 2, 3). Three *O. j. andina* had haplotypes also found in *O. j. jamaicensis*, whereas one was identical to a *O. j. ferruginea* haplotype, and one was unique, supporting Todd's (1979) and Fjeldså's (1986) hypothesis that *O. j. andina* is either an intergrade subspecies or hybrid population of *O. j. jamaicensis* and *O. j. ferruginea*, with males approaching one or the other in phenotype and genotype. A further test of this hypothesis would be to genotype *O. j. andina* samples for single nucleotide polymorphisms (SNPs) or other nuclear markers that distinguish *O. j. jamaicensis* and *O. j. ferruginea* to determine if the *O. j. andina* genome represents the admixture of *O. j. jamaicensis* and *O. j. ferruginea* genes. Likewise, more *O. j. ferruginea* samples are required from Patagonia and Tierra del Fuego to determine if southern populations of *O. j. ferruginea* are distinct from populations inhabiting the northern Andes. Our data indicate that the four individuals we sampled from Argentina's Patagonian provinces of Neuquén and Chubut differ from other *O. j. ferruginea* by two nucleotide differences (Fig. 3).

Two biogeographic hypotheses for *O. jamaicensis* are potentially consistent with our results. First, the evolution of *O. jamaicensis* from ancestral *Oxyura* might have followed the colonization of the Northern Hemisphere and establishment of a migratory population. Much more recently (given the limited divergence among *O. jamaicensis* subspecies) and perhaps in response to Pleistocene climate change, wintering *O. j. jamaicensis* from North America might have established local breeding populations in the Northern Central Andes, extended their range southwards, and subsequently colonized most of the Andean Cordillera and adjacent steppe and coastal habitat "top-down" from north to south and from high elevation to low elevation. Southern populations of *O. j. ferruginea* subsequently diverged as *O. j. jamaicensis* resident in the northern Andes expanded their range southwards and evolved phenotypic differences such as the all-black head, larger body size, and

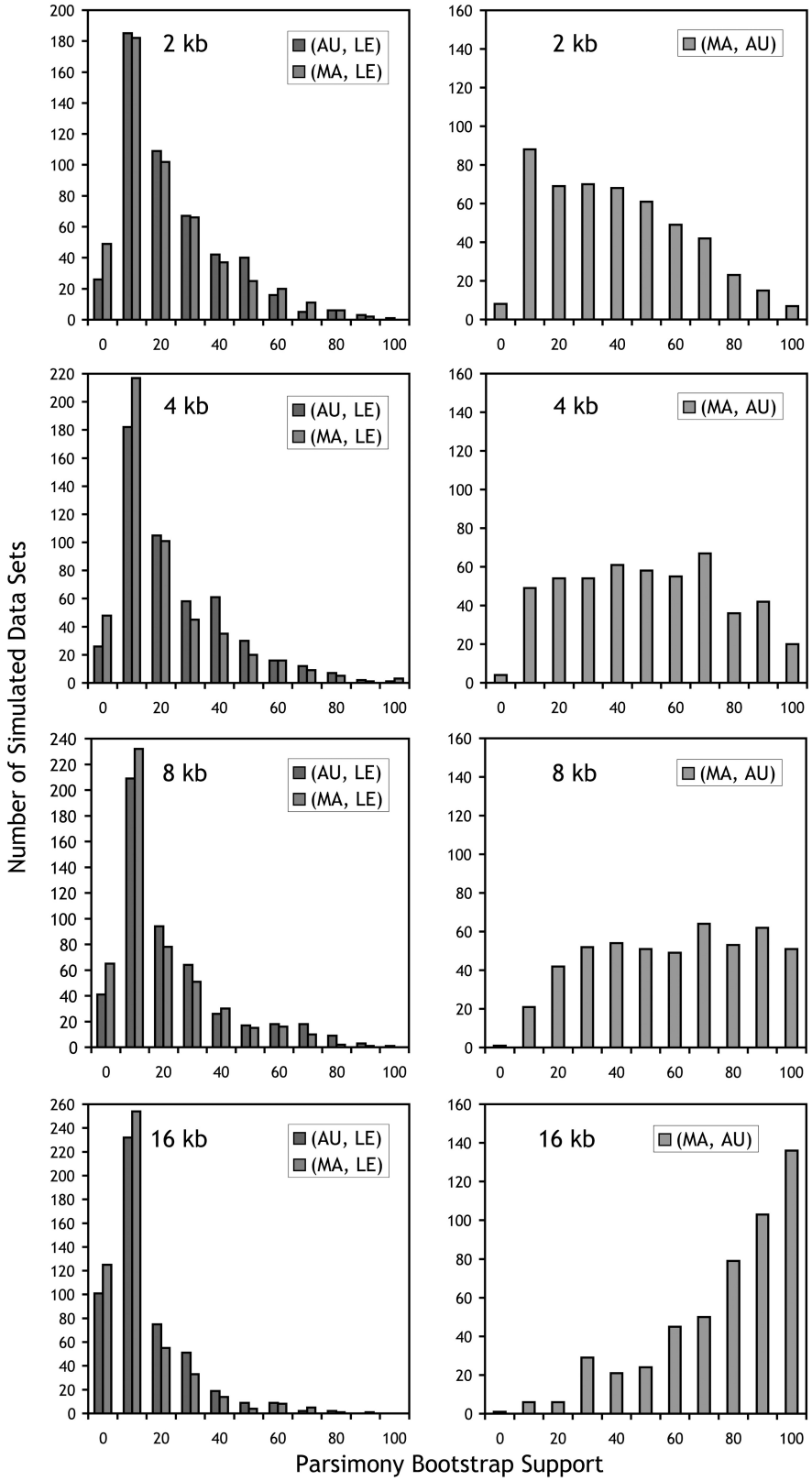

FIGURE 7. Distribution of unweighted parsimony bootstrap values for three alternative resolutions of a trichotomy involving three Old World *Oxyura* species: AU = *O. australis*; MA = *O. maccoa*; LE = *O. leucocephala*. The distributions are based on bootstrap analysis of 500 simulated data sets for 2 kb, 4 kb, 8 kb, and 16 kb sequence lengths. Sequence evolution was simulated on the tree shown in Figure 4b, in which *O. maccoa* and *O. australis* are sister taxa. The maximum bootstrap value in each category is indicated (e.g., the number of replicates with values between 70 and 80 is plotted at 80).

reduced spotting on the breast (Johnsgard, 1965; Fjeldså, 1986; Johnsgard and Carbonell, 1996). This hypothesis is consistent with the somewhat greater haplotype diversity found in *O. j. jamaicensis* and the nesting of *O. j. ferruginea* haplotypes within a larger clade of *O. jamaicensis* haplotypes (Figs. 2, 3), the result found in 50% of the most parsimonious trees.

Alternatively, ancestral *O. jamaicensis* might have evolved in the Southern Hemisphere and only recently colonized the Northern Hemisphere. The discovery of additional haplotype diversity and deeper divergences among *O. j. ferruginea* populations in South America would support this hypothesis. In either model, subsequent hybridization of the two subspecies in Colombia, where the winter range of *O. j. jamaicensis* abuts the range of *O. j. ferruginea*, likely explains the intermediate phenotypic traits (Adams and Slavid, 1984) and divergent haplotypes present in a hybrid *O. j. andina* population.

The above models have different implications for the route by which ancestral *Oxyura* reached the Old World. If *O. jamaicensis* colonized the Northern Hemisphere early in its history, then the common ancestor of *O. leucocephala*, *O. maccoa*, and *O. australis* might also have been a migratory species in Eurasia. If *O. jamaicensis* was restricted to the Southern Hemisphere until recently, as in the alternative model, then the common ancestor of Old World *Oxyura* might have originated in Africa or perhaps Australia. The almost contemporaneous divergence of the three Old World species from a common ancestor is consistent with either an "out-of-Africa" hypothesis, in which an ancestral *O. maccoa* population colonizes both Eurasia and Australia at about the same time, or a "north-to-south" hypothesis similar to that proposed above for *O. jamaicensis*, in which a historically widespread and migratory *O. leucocephala* ancestor in Eurasia colonizes Africa and Australia at about the same time. Distinguishing between these scenarios is complicated by the difficulty of resolving the relationships among these three species (see below). Nonetheless, we can conclude unambiguously that the divergence of *O. j. jamaicensis* and *O. j. ferruginea* is far more recent than the divergence of ancestral *O. jamaicensis* from ancestral *Oxyura*.

Conflicting mtDNA and Nuclear Characters

Control region data suggest that Eurasian *O. leucocephala* is basal to an Australian–African clade (*O. australis*, *O. maccoa*), but this clade is not well supported (46% to 81% parsimony and likelihood bootstrap; Fig. 2, Tables 2, 3). Increasing the size of the data set to 8801 positions from 20 different mtDNA and nuclear loci (including the mtDNA control region) yielded only 196 informative characters and also failed to resolve the *Oxyura* trichotomy (Fig. 5, Table 3). Two alternative rearrangements, (*O. maccoa*, *O. leucocephala*) and (*O. australis*, *O. maccoa*) were equally parsimonious and had similar parsimony bootstrap support, 41% and 35%. The mtDNA data alone yielded 57% to 64% parsimony and maximum likelihood bootstrap support for (*O. australis*, *O. maccoa*), and 31% to 35% for (*O. australis*, *O. leucocephala*), and essentially no support for (*O. maccoa*, *O. leucocephala*). In contrast, nuclear data yielded 92% parsimony bootstrap support for (*O. maccoa*, *O. leucocephala*) but no support for the other two possible resolutions of the trichotomy.

Analysis of characters supporting the three possible resolutions of this trichotomy indicate that different loci and different characters within the same loci support different clades (Table 4). Within the mtDNA control region, six characters support (*O. australis*, *O. maccoa*) and seven characters support (*O. australis*, *O. leucocephala*), but only one character supports (*O. maccoa*, *O. leucocephala*). The protein-coding genes ND2, ND5, and cytochrome *b*, also show similar patterns of conflicting character support, with eight, nine, and five characters supporting the three clades. One character from the noncoding 12S rRNA gene supports (*O. maccoa*, *O. leucocephala*), and two characters support (*O. australis*, *O. leucocephala*). In contrast, the two introns (MPP, PEPCK9; Fig. 6) with synapomorphies relevant to the *Oxyura* trichotomy support (*O. maccoa*, *O. leucocephala*). Given limited homoplasy in the nuclear data, the nuclear topologies can probably be regarded as reasonable accurate gene trees that are in conflict with the results of the mtDNA analysis, particularly with respect to the relationships of these three species. Combining these conflicting data from different linkage groups into a single analysis resulted in reduced levels of bootstrap support.

The Old World Oxyura Trichotomy

Despite collecting more than 8000 base pairs of sequence data per taxon, we are left with an unresolved trichotomy among the three Old World *Oxyura* species. Two distinct questions can be considered in attempting to resolve this kind of trichotomy, and both appear to be relevant to the present example: (1) Have individual gene trees been adequately resolved? (2) Are the individual gene trees congruent with the species history?

In the present example, different factors limited the resolution of the mtDNA and nuclear gene trees, respectively. Gene trees for most nuclear loci were unresolved primarily due to a lack of informative variation. Two or more species shared identical sequences at eight of ten nuclear loci. For many loci, retention of ancestral sequences does not necessarily reflect incomplete lineage sorting during the internodes of the *Oxyura* phylogeny, but may simply result from a lack of informative substitutions. If the mutation rate is sufficiently low, such that no substitutions accumulated during one or more internodes, an unresolved gene tree and shared ancestral sequences may be observed even if the lineage sorting process was completed during each internode in a phylogeny. The lack of variation, however, makes it impossible to recover the genealogical history. For three loci, however, we observed patterns consistent with retained ancestral polymorphism or the persistent effects of differential lineage sorting in the past. For example, *O. jamaicensis* shares one allele for α-ENOL with *O. leucocephala* and another with *O. vittata* and *O. australis*. All species have unique alleles for MPP, but the MPP gene

tree is incongruent with the clade of three Old World species that is strongly supported by the mtDNA data and two nuclear loci (hemoglobin α-A, PEPCK9; Fig. 6).

In contrast to the nuclear DNA results, uncertain resolution of the mtDNA gene tree is due primarily to homoplasy. Nearly equal numbers of characters support each of three possible resolutions of a trichotomy among three Old World species (Table 4). A very short internode separating *O. leucocephala* from (*O. australis*, *O. maccoa*) combined with relatively long terminal branches apparently allowed the small number of changes occurring on the internode to be largely obscured by homoplasy, with conflicting characters generated either by parallel changes on two of the three terminal branches or by reversals of changes occurring on the longer branch leading to the most recent common ancestor of these three species.

The parametric bootstrap analyses provide an estimate of the probability of recovering the (*O. australis*, *O. maccoa*) clade if these two species are indeed sister taxa in the mtDNA haplotype tree and if the internode separating them from *O. leucocephala* has the length estimated from the empirical data. The analysis is conducted under the assumption that sequences evolve perfectly, but stochastically, under a known model of sequence evolution that is also used to reconstruct the phylogeny, providing an estimate of the power of the analysis. Failure to reconstruct the expected tree even under these conditions stems from the stochasticity of sequence evolution. Given the short internode and relatively long terminal branches, some replicate data sets by chance accumulate few if any changes on the internode and some of these are overwritten by homoplasy. Given the branch lengths estimated from the empirical data, there is only a 72% chance of recovering the correct resolution of the trichotomy in a maximum likelihood analysis of 4-kb sequences (Table 5), whereas parsimony bootstrap support for the correct resolution of the trichotomy ranges almost uniformly from 10 to 100% (Fig. 7). Expected bootstrap values for the incorrect resolutions of the trichotomy are substantially lower, but values greater than 50% occur in up to 15% of replicates for data sets up to 8 kb. The probability of finding bootstrap support greater than 70% for an incorrect resolution of the trichotomy, however, is low even for data sets of only 2 kb. With 16-kb sequences, much better discrimination between alternative hypotheses is achieved in the simulated data (Fig. 7), suggesting that complete mtDNA sequences for these taxa might allow a more definitive resolution of the Old World *Oxyura* trichotomy. This assumes, however, that the true length of the internode is greater than or equal to that modeled in our analyses.

In addition, it should be noted that the parametric bootstrapping analysis uses a single model of sequence evolution to both simulate and analyze data. Additional complexities affect the evolution of real sequence data, including heterogeneity in the substitution process among genes and individual nucleotide positions as well as shifts in base composition among taxa. Considering parsimony informative sites in the mtDNA only, *Oxyura* species differed slightly in base composition, although *O. maccoa* and *O. leucocephala* are more similar to each other in base composition than either is to *O. australis*, suggesting that directional evolution of base composition is not responsible for the mtDNA result (i.e., sister relationship of *O. australis* and *O. maccoa*).

We conclude that resolution of the mtDNA tree in favor of (*O. australis*, *O. maccoa*) and 64% bootstrap support for this clade (based on the 4070 bp used to estimate parameters for the parametric bootstrap analysis) could occur by chance even if the true mtDNA tree included either a (*O. australis*, *O. leucocephala*) or (*O. maccoa*, *O. leucocephala*) clade. Nonetheless, (*O. australis*, *O. maccoa*) is the clade most strongly supported by the empirical data and must be viewed as the best current hypothesis for mtDNA haplotype relationships. Additional mtDNA sequence data are needed to determine if this trichotomy in the *Oxyura* mtDNA tree is a hard or soft polytomy.

The above discussion considers only the resolution of the mtDNA gene tree. In the case of *Oxyura*, there is uncertainty in both the mtDNA tree and gene trees for most nuclear loci, but for different reasons. Because each gene tree is but one manifestation of the lineage sorting process through a single history of population divergence, we must also consider whether individual gene trees are congruent with the species history. If branching events occur in rapid succession relative to a large N_e, then trees derived from independent loci are likely to be incongruent, due to the lasting effects of ancestral polymorphism and differential lineage sorting. Thus, even if the mtDNA tree is resolved in favor of (*O. australis*, *O. maccoa*), it may be difficult to say with any confidence that the *O. australis* and *O. maccoa* species lineages share a more recent common ancestor than either species lineage does with *O. leucocephala*.

In other words, congruence of the mtDNA gene tree and the species history depends on both internode length and N_e during this interval. In contrast, recovery of the correct mtDNA gene tree depends only on internode length. However, if the internode separating the three Old World *Oxyura* was not quite zero, the probability that the mtDNA gene tree matches the species history should be greater than for nuclear loci becaue mtDNA N_e is one quarter the size of nuclear N_e and therefore completes the lineage sorting process more rapidly (e.g., Moore 1995).

Old Ancestral Polymorphism or Homoplasy?

At the time the Old World stiff-tailed ducks diverged, ancestral lineages probably showed substantial mtDNA polymorphism as extant lineages do today (Figs. 2, 3), and for some time afterwards, ancestral lineages likely persisted in a state of incomplete lineage sorting (Avise et al., 1990; Avise, 1994; Moore, 1995). However, our data indicate that extant *Oxyura* lineages no longer show incomplete lineage sorting in the mtDNA (Fig. 2). Moreover, conflicting characters occur within individual mtDNA genes, in addition to between different linkage groups (mtDNA versus nuclear DNA). Ancestral

polymorphism per se, if it occurred, no longer exists in the *Oxyura* mtDNA, but it may have lasting effects through the stochastic outcome of the lineage sorting process in independent genetic loci. If ancestral polymorphism generated mismatches between different gene trees (mtDNA versus nuclear gene trees), which gene tree correctly reflects the species tree?

If the true species tree is (*O. australis, O. maccoa*), then the mtDNA data yielded the correct tree, and two introns are incongruent (presumably due to the differential sorting of ancestral polymorphisms). If (O. *maccoa, O. leucocephala*) is the true species tree, then we are faced with an incongruent mtDNA tree (either because of old ancestral polymorphism and stochastic lineage sorting or incorrect resolution of the mtDNA tree due to homoplasy) plus two nuclear loci congruent with the species history. Alternatively, the Old World *Oxyura* trichotomy might very well be a hard polytomy in the species history. If so, any particular gene tree for a given linkage group should have an equal chance of supporting any of three possible resolutions. There also should be no a priori reason to expect difficulty in resolving each gene tree, unless the rate of evolution is so low that few changes have accumulated, as is generally the case for the nuclear introns and exons we sequenced. Our difficulty in resolving the mtDNA gene tree is clearly related to homoplasy. However, parametric bootstrapping suggests that complete mtDNA sequences would yield a fully resolved tree. It is the addition of nuclear data, which strongly support (*O. maccoa, O. leucocephala*; 92% bootstrap support), that turns the combined mtDNA and nuclear data tree into a trichotomy.

By concatenating mtDNA and nuclear data, we combined independent markers that potentially sorted in different patterns through the species history. Thus, one point of view is that each linkage group effectively provides a single measure of species history and what is important is how many loci support each resolution. If *Oxyura* is indeed a hard polytomy, 33% of gene trees based on independent, polymorphic loci should support each of the three possible clades. Thus, one test would be to sequence many additional polymorphic introns or exons. However, our sequencing effort yielded little more than one informative single nucleotide polymorphism (SNP) for every other locus. Several tens of kb of nuclear sequence from 50 to 100 loci might be required to perform such a test.

Stochastic homoplasy generally is expected to increase cumulatively with elapsed time and mutation rate (Takahata, 1995; O'hUigin et al., 2002). Bootstrap support, on the other hand, results from the interaction between depth of divergence and internode length. For *Oxyura*, this pattern is evident in our data. Within the Old World *Oxyura* clade, terminal branches are long, but the internodes separating the lineages are short (Figs. 2, 4, 5). Parametric bootstrap analysis, likewise, suggests that the levels of homoplasy we observed in the mtDNA data partition do not differ from those that would be expected given the short internodes and long terminal branches we observed in *Oxyura*. For example, the average number characters supporting the three possible resolutions of the trichotmy in 500 simulated data sets (14, 10, 10) are strikingly similar to the numbers of conflicting characters observed in the empirical data (15, 9, 11; numbers differ slightly from Table 4 due to exclusion from this analysis of positions with gaps in one or more taxa).

Similar patterns of conflicting characters also might have resulted if recombination occurred between sites as polymorphisms sorted differentially in related species. However, recombination in mtDNA is rare, and our mtDNA data show no apparent recombination axes. Nucleotide positions supporting the three possible resolutions of the trichotomy are distributed relatively uniformly throughout the control region, 12S rRNA, ND2, ND5, and cytochrome *b* genes, and the same pattern also occurs within genes (Table 4). No conflicting characters, likewise, were found within individual nuclear loci (Table 4), suggesting that the small size of the products we amplified (mean = 427 bp) and the recent history of *Oxyura* limited the potential influence of recombination at these loci.

Our results lead to the conclusion that sequences of tens of thousands of base pairs probably would be needed to adequately resolve the *Oxyura* phylogeny. Complete mtDNA genomes probably would produce a resolved mtDNA gene tree, but the mtDNA gene tree might not reflect the species tree. The root cause of our inability to fully resolve the mtDNA tree with 4 kb clearly is homoplasy. The effects of homoplasy we identified in our study apply generally and ultimately explain why many trees derived from mtDNA data show poor resolution and low bootstrap support. Although the problems of reconstructing short internodes for taxa with relatively long terminal branches are well known, these issues are more often considered in studies of highly divergent taxa. Our findings suggest that even among closely related species, substantial rate variation among sites can generate enough homoplasy to obscure relatively recent speciation events. Future efforts to resolve short internodes separating taxa with long terminal braches, such as those observed in *Oxyura*, might be better served by sequencing and scoring dozens (if not hundreds) of nuclear loci from a variety of different polymorphic linkage groups. Generally slower rates of mutation and fixation, however, will make the identification of sufficiently variable nuclear loci challenging.

ACKNOWLEDGMENTS

We thank Yanina Arzamendia, Daniel Blanco, Bob Brua, Guillermo Cao, Raúl Cardón, Tamar Cassidy, Sonia Chavarra, Claudio Chehébar, Les Christidis, Raúl Clarke, Mike Christie, Donna Dittmann, Martin Funes, Sergio Goldfeder, Andy Green, Baz Hughes, Anthony MaGuire, Rodolfo Miatello, Manuel Nores, David Paton, Carmen Quiroga, Van Remsen, Kevin Shaw, Esther Signer, Alejandro del Valle, and Heidi Weingartz for helping to locate or provide stiff-tailed duck tissues. Melissa Cunningham, Sergio Goldfeder, Mike Grogan, Jess Hemmings, Gordon Jarrell, Bill Johnson, Kevin Johnson, Pamela McCracken, Rob Wilson, and Jason Weckstein assisted in the field, and Rich Olsen loaned a shotgun. Collections were made possible by Cape Nature Conservation, Centro de Ecología Aplicada y Dirección Provincial Recursos

Faunisticos y Areas Naturales Protegidas Neuquén, Colección Boliviana de Fauna, Delegación Regional Patagonia Administración de Parques Nacionales, Dirección de Fauna y Flora Silvestre Argentina, Dirección de Fauna Santa Cruz, Dirección de Ordenamiento Ambiental–Área Técnica de Fauna Cordoba, Dirección Provincial de Recursos Naturales y Medioambiente Jujuy, Dirección de Recursos Naturales y Gestion Ambiental Corrientes, Ministerio de la Producción Chubut, Secretaría de Estado de Producción Río Negro, Secretaría de Medioambiente y Desarrollo Sustentable Salta, Estacion Biologica de Doñana, Field Museum of Natural History, Louisiana Cooperative Fish and Wildlife Research Unit, Louisiana State University Agricultural Center, College of Agriculture, Museum of Natural Science, and School of Forestry, Wildlife, and Fisheries, South Australia National Parks and Wildlife, Texas Parks and Wildlife, Transvaal Museum, Treehaven Wildfowl Trust, and Wildfowl and Wetlands Trust. Allan Baker, George Barrowclough, Michael Braun, and Chris Simon provided helpful comments on the manuscript. Laboratory costs were funded by the Institute of Arctic Biology at the University of Alaska Fairbanks, Alaska EPSCoR (NSF EPS-0092040), a grant from the Frank M. Chapman Fund at the American Museum of Natural History to K.G.M., and an NSF grant to M.D.S.

REFERENCES

Adams, J., and E. R. Slavid. 1984. Cheek plumage in the Colombian ruddy duck *Oxyura jamaicensis*. Ibis 126:405-407.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* Proceedings of Second International Symposium on Information Theory (B. N. Petrov and F. Csaki, eds.). Akademia Kiado, Budapest.

Avise, J. C. 1994. Molecular markers, natural history, and evolution. Chapman and Hall, New York.

Avise, J. C. 2000. Phylogeography: The history and formation of species. Harvard University Press, Cambridge, Massachusetts.

Avise, J. C., C. D. Ankney, and W. S. Nelson. 1990. Mitochondrial gene trees and the evolutionary relationship between mallard and black ducks. Evolution 44:109–1119.

Bremer, K. 1988. The limits of amino-acid sequence data in angiosperm phylogenetic reconstruction. Evolution 42:795–803.

Clement M., D. Posada, K. A, Crandall. 2000. TCS: A computer program to estimate gene genealogies. Mol. Ecol. 9:1657–1660.

Crochet, P. A., and E. Desmerais. 2000. Slow rate of evolution in the mitochondrial control region of gulls (Aves: Laridae). Mol. Biol. Evol. 17:1797–1806.

Delacour, J., and E. Mayr. 1945. The family Anatidae. Wilson Bull. 56:3–55.

DeSalle, R., R. Absher, and G. Amato. 1994. Speciation and phylogenetic resolution. TREE 9:297–298.

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

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 39:783–791.

Fieldså, J. 1986. Color variation in the ruddy duck (*Oxyura jamaicensis andina*). Wilson Bull. 98:592–594.

Friesen, V. L, B. C. Congdon, H. E. Walsh, and T. P. Birt. 1997. Intron variation in marbled murrelets detected using analyses of single-stranded conformational polymorphisms. Mol. Ecol. 6:1047–1058.

Gladstein, D., and W. Wheeler. 2000. POY: Phylogeny reconstruction via direct optimization of DNA data, version 2.0. American Museum of Natural History, New York, New York.

Green, A. J., and B. Hughes. 1996. Action plan for the white-headed duck *Oxyura leucocephala*. Pages 119–146 *in* Globally threatened birds in Europe (B. Heredia, L. Rose, and M Painter, eds.). Council of Europe Publishing, Strasbourg, France.

Harshman, J. 1996. Phylogeny, evolutionary rates, and ducks. Ph.D. dissertation, University of Chicago, Chicago, Illinois.

Hasegawa, M. H. Kishino, and T. Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. J. Mol. Evol. 22:160–174.

Hoelzer, G. A., and D. J. Melnick. 1994. Patterns of speciation and limits to phylogenetic resolution. TREE 9:104–107.

Huelsenbeck, J. P., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

Hughes, B., J. Criado, S. Delany, U. Gallo-Orsi, A. J. Green, M. Grussu, C. Perennou, and J. A. Torres. 1999. The status of the North American ruddy duck *Oxyura jamaicensis* in the Western Palearctic: Towards an action plan for eradication. Council of Europe Publication T-PVS/Birds (99) 9. Council of Europe Publishing, Strasbourg, France.

Johnsgard, P. A. 1965. Handbook of waterfowl behavior. Cornell University Press, Ithaca, New York.

Johnsgard, P. A. 1978. Ducks, geese, and swans of the world University of Nebraska Press, Lincoln, Nebraska.

Johnsgard, P. A., and M. Carbonell. 1996. Ruddy ducks and other stifftails: Their behavior and biology. University of Oklahoma Press, Norman, Oklahoma.

Livezey, B. C. 1986. A phylogenetic analysis of recent anseriform genera using morphological characters. Auk 103:737–754.

Livezey, B. C. 1995. Phylogeny and comparative ecology of stifftailed ducks (Anatidae: Oxyurini). Wilson Bull. 107:214–234.

Livezey, B. C. 1997. A phylogenetic classification of waterfowl (Aves: Anseriformes), including selected fossil species. Ann. Carnegie Mus. 66:457– 496.

Maddison, W. 1989. Reconstructing character evolution on polytomous cladograms. Cladistics 5:365–377.

Marshall, H. D., and A. J. Baker. 1997. Structural conservation and variation in the mitochondrial control region of fringilline finches (*Fringilla* spp.) and in the greenfinch (*Carduelis chloris*). Mol. Biol. Evol. 14:173–184.

McCracken, K. G., J. Harshman, D. A. McClellan, and A. D. Afton. 1999. Date set incongruence and correlated character evolution: An example of functional convergence in the hind-limbs of stifftail diving ducks. Syst. Biol. 48:683–714.

Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: Mitochondrial gene trees versus nuclear-gene trees. Evolution 49:718–726.

Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation. Pages 515–534 *in* Evolutionary processes and theory (E. Nevo and S. Karlin, eds.). Academic Press, New York.

O'hUigin, C., Y. Satta, N. Takahata, and J. Klein. 2002. Contribution of homoplasy and of ancestral polymorphism to the evolution of genes in Anthropoid primates. Mol. Biol. Evol. 19:1501–1513.

Pacheco, N. M., B. C. Congdon, and V. L. Friesen. 2002. The utility of nuclear introns for investigating hybridization and genetic introgression: A case study involving *Brachyramphus* murrelets. Conserv. Genet. 3:175–182.

Posada, D., and K. A. Crandall. 1998. Modeltest: Testing the model of DNA substitution. Bioinformatics 14:817–818.

Primmer, C. R., T. Borge, J. Lindell, and G. P. Sætre. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: High nucleotide diversity revealed in the avian genome. Mol. Ecol. 11:603–612.

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

Raikow, R. J. 1970. Evolution of diving adaptations in the stiff-tailed ducks. Univ. Calif. Publ. Zool. 94:1–52.

Rambaut, A., and N. C. Grassly. 1997. Seq-Gen: An application for the Monte Carlo simulation of DNA sequence evolution along phylogenetic trees. Comput. Appl. Biosci. 13:235–238.

Saitou, N., and M. Nei. 1986. The number of nucleotides required to determine the branching order of three species, with special reference to the human-chimpanzee-gorilla divergence. J. Mol. Evol. 24:189–204.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. Syst. Biol. 51:492–508.

Sorenson, M. D. 1999. TreeRot, version 2. Boston University, Boston, Massachusetts.

Sorenson, M. D., J. C. Ast, D. E. Dimcheff, T. Yuri, and D. P. Mindell. 1999. Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Mol. Phylogenet. Evol. 12:105–114.

Sorenson, M. D., and R. C. Fleischer. 1996. Multiple independent transpositions of mitochondrial DNA control region sequences to the nucleus. Proc. Nat. Acad. Sci. USA 93:15239–15243.

Sorenson, M. D., and R. B. Payne. 2001. A single, ancient origin of obligate brood parasitism in African finches: Implications for host-parasite coevolution. Evolution 55:2550–2567.
Sorenson, M. D., and T. W. Quinn, T. W. 1998. Numts: A challenge for avian systematics and population biology. Auk 115:214–221.
Sraml, M., L. Christidis, S. Easteal, P. Horn, and C. Collet. 1996. Molecular relationships within Australian waterfowl (Anseriformes). Aust. J. Zool. 44:47–58.
Swofford, D. L. 1998. PAUP*: Phylogenetic analysis using parsimony (*and other methods), version 4. Sinauer Associates, Sunderland, Massachusetts.
Tajima, F. 1983. Evolutionary relationships of DNA sequences in finite populations. Genetics 105:437–460.
Takahata, N. 1995. *Mhc* diversity and selection. Immunol. Rev. 143:225–247.
Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties, and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
Todd, F. S. 1979. Waterfowl: Ducks, geese, and swans of the world. Sea World Press, Sand Diego, California.
Walsh, H. E., M. G. Kidd, T. Moum, and V. L. Friesen. 1999. Polytomies and the power of phylogenetic inference. Evolution 53:932–937.
Wheeler, W. C. 1996. Optimization alignment: The end of multiple sequence alignment in phylogenetics? Cladistics 12:1–9.

First submitted 26 March 2003; reviews returned 4 December 2003; final acceptance 10 September 2004
Associate Editor: Allan Baker

APPENDIX 1. Stiff-tailed duck species and sources of genetic material. Specimen information is grouped by common name (Latin name) and includes: catalog number,[a] haplotype, and specimen locality information.

Masked Duck (*Nomonyx dominicus*):

LSUMZ 127446,[b] NODO-06, USA: Louisiana, Vermillion Parish, 4 mi SW Ester, 6 January 1973; LSUMZ 98966,[b] NODO-07, PERU: Dpto. Amazonas, Caterpiza on Quebrada Caterpiza, E bank tributary of Río Santiago, 656 m, February 1980; LSUMZ 123430,[b] NODO-08, BOLIVIA: Dpto. Santa Cruz, Prov. Cordillera, 10 km E Gutierrez, Laguna Caucaya, 875 m, 14 April 1984; LSUMZ 123431,[b] NODO-01, BOLIVIA: Dpto. Santa Cruz, Prov. Cordillera, 10 km E Gutierrez, Laguna Caucaya, 875 m, 14 April 1984; UAM 14601, NODO-05A, ARGENTINA: Corrientes, S Bella Vista, 28° 41′ 33″ S, 59° 00′ 33″ W, 20 October 2001 (KGM 476); UAM 14784, NODO-02A, NODO-02B, ARGENTINA: Corrientes, N La Concepción, 29° 40′ 02″ S, 59° 21′ 45″ W, 58 m, 18 October 2001 (KGM 464); UAM 19510, NODO-03B, ARGENTINA: Corrientes, N La Concepción, 29° 40′ 02″ S, 59° 21′ 45″ W, 58 m, 18 October 2001 (KGM 465); UAM 14785, NODO-04A, NODO-04B, ARGENTINA: Corrientes, N La Concepción, 29° 40′ 02″ S, 59° 21′ 45″ W, 58 m, 18 October 2001 (KGM 466)

Argentine Lake Duck (*Oxyura vittata*):

UAM 17536, OXVI-07, ARGENTINA: Cordoba, Arroyo Chucul, ca. Laguna La Felipa, 33° 12′ 15″ S, 63° 43′ 44″ W, 28 April 2001 (KGM 305); UAM 17535, OXVI-06, ARGENTINA: Cordoba, Arroyo Chucul, ca. Laguna La Felipa, 33° 12′ 05″ S, 63° 37′ 32″ W, 29 April 2001 (KGM 314); UAM 17534, OXVI-08, ARGENTINA: Cordoba, Arroyo Chucul, ca. Laguna La Felipa, 33° 12′ 41″ S, 63° 34′ 14″ W, 30 April 2001 (KGM 317); LSUMZ B34015, OXVI-01, ARGENTINA: Río Negro, 5 km NW General Conesa, 22 November 1998 (#1); LSUMZ B34017, OXVI-02, ARGENTINA: Río Negro, 5 km NW General Conesa, 22 November 1998 (#2); LSUMZ B34019, OXVI-03, ARGENTINA: Río Negro, 15 km W General Conesa, 22 November 1998 (#3); LSUMZ B34025, OXVI-04, ARGENTINA: Río Negro, 15 km W General Conesa, 22 November 1998 (#4); LSUMZ B34034, OXVI-02, ARGENTINA: Río Negro, 30 km W General Conesa, 22 November 1998 (#5); LSUMZ B34036, OXVI-05, ARGENTINA: Río Negro, 25 km W General Conesa, 22 November 1998 (#6); LSUMZ B34037, OXVI-06, ARGENTINA: Río Negro, 25 km W General Conesa, 22 November 1998 (#7); SHW-MDS, OXVI-03, CAPTIVE: Sylvan Heights Waterfowl, USA

Ruddy Duck (*Oxyura j. jamaicensis*):

MN-K9, OXJA-09, CANADA: Manitoba, Minnedosa, 11 May 1995 (FWS 936-01302); MN-N1, OXJA-08, CANADA: Manitoba, Minnedosa, 12 May 1995 (FWS 936-01303); MN-N2, OXJA-03, CANADA: Manitoba, Minnedosa, 14 May 1995 (FWS 936-01304); MN-L3, OXJA-08, CANADA: Manitoba, Minnedosa, 19 May 1995 (FWS 936-01309); MN-S1, OXJA-05, CANADA: Manitoba, Minnedosa, 21 May 1995 (FWS 936-01313); MN-01, OXJA-04, CANADA: Manitoba, Minnedosa, 2 July 1995; MN-02, OXJA-06, CANADA: Manitoba, Minnedosa, 3 July 1995; MN-04, OXJA-07, CANADA: Manitoba, Minnedosa, 6 June 1995; FWS-CA, OXJA-10, USA: California, Siskiyou County; LSUMZ 155831, OXJA-01, USA: Louisiana, St. Tammany Parish, 1.5 mi W causeway, 30 January 1987 (B4093); LSUMZ 154354, OXJA-02, USA: Louisiana, La Salle Parish, Catahoula Lake, 9 December 1991 (B20738); LSUMZ 165151, OXJA-03, USA: Texas, Jefferson Co., J. D. Murphree WMA, Salt Bayou Unit, 6 February 1999 (B31011); LSUMZ 50481,[b] OXJA-02, EL SALVADOR: Dpto. Libertad, Laguna de Chanmico, 18 June 1964; UK-07, OXJA-08, UNITED KINGDOM: Eyebrook; UK-08, OXJA-08, UNITED KINGDOM: Coron; ES-09, OXJA-08, SPAIN: El Hondo de Almoros, 15 January 1996; ES-10, OXJA-08, SPAIN: El Hondo de Almoros, 18 December 1995

Colombian Ruddy Duck (*Oxyura j. andina*):

FMNH 13608,[b] OXAN-04,[c] COLOMBIA; FMNH 13609,[b] OXAN-01, COLOMBIA; FMNH 13610,[b] OXAN-02,[d] COLOMBIA; FMNH 414481,[b] OXAN-03,[e] COLOMBIA: Dpto. Cauca, Páramo de Puracé, Río Magdalena side, 28 January 1940; FMNH 414482,[b] OXAN-03,[e] COLOMBIA: Dpto. Cauca, Páramo de Puracé, Río Magdalena side, 28 January 1940

Andean Ruddy Duck (*Oxyura j. ferruginea*):

ANSP 4221, OXFE-07, ECUADOR: Imbabura, N Ibara, Laguna de Yaguarcocha, 2280 m, 11 June 1992; LSUMZ 34550,[b] OXFE-02, PERU: Dpto. Puno, Puno, Lago Titicaca, 13 November 1964; LSUMZ 63848,[b] OXFE-02, PERU: Dpto. Junín, Paca, 3250 m, 9 June 1968; LSUMZ 91605,[b] OXFE-04, PERU: Dpto. La Libertad, Quebrada la Caldera, 10 km NE Tayabamba, 3575 m, 15 August 1979; UAM 19014, OXFE-05, BOLIVIA: Dpto. La Paz, Cohani, Lago Titicaca, 16° 21′ 02″ S, 68° 37′ 48″ W, 3840 m, 28 October 2001 (KGM 497); UAM 19239, OXFE-07, BOLIVIA: Dpto. La Paz, SE Taraco, Lago Titicaca, 16° 27′ 52″ S, 68° 53′ 29″ W, 3840 m, 10 November 2001 (KGM 553); UAM 19240, OXFE-02, BOLIVIA: Dpto. La Paz, SE Taraco, Lago Titicaca, 16° 27′ 52″ S, 68° 53′ 29″ W, 3840 m, 10 November 2001 (KGM 554); UAM 19241, OXFE-06, BOLIVIA: Dpto. La Paz, SE Taraco, Lago Titicaca, 16° 27′ 52″ S, 68° 53′ 29″ W, 3840 m, 10 November 2001 (KGM 555); UAM 19242, OXFE-06, BOLIVIA: Dpto. La Paz, SE Taraco, 16 Lago Titicaca, 16° 27′ 52″ S, 68° 53′ 29″ W, 3840 m, 10 November 2001 (KGM 556); LSUMZ 123432,[b] OXFE-02, BOLIVIA: Dpto. Santa Cruz, Prov. Cordillera, 10 km E Gutierrez, Laguna Caucaya, 875 m, 6 April 1984; LSUMZ 123433,[b] OXFE-03, BOLIVIA: Dpto. Santa Cruz, Prov. Cordillera, 10 km E Gutierrez, Laguna Caucaya, 875 m, 6 April 1984; LSUMZ B34016, OXFE-01, ARGENTINA: Neuquén, Río Alumine, ca. Lago Pulmari, 12 November 1998; LSUMZ B34018, OXFE-01, ARGENTINA: Neuquén, Río Alumine, ca. Lago Pulmari, 12 November 1998; LSUMZ B34000, OXFE-01, ARGENTINA: Neuquén, Río Colló Curá, R.N. 40, 40° 12′ 45″ S, 70° 38′ 58″ W, 18 November 1998; LSUMZ 69732,[b] OXFE-01, ARGENTINA: Chubut, El Hoyo, 5 May 1969

(Continued on the next page)

APPENDIX 1. (CONTINUED).

Australian Blue-Billed Duck (*Oxyura australis*):

KGM SA-01, OXAU-01, AUSTRALIA: South Australia, Cape Gantheaume Conservation Park, Murray Lagoon, 22 September 1995; KGM SA-02, 05, OXAU-01, AUSTRALIA: South Australia, Cape Gantheaume Conservation Park, Murray Lagoon, 3 October 1997; KGM SA-03, 04, 06, 07, 11, OXAU-02, AUSTRALIA: South Australia, Cape Gantheaume Conservation Park, Murray Lagoon, 3 October 1997; KGM SA-08, 09, 13, OXAU-03, AUSTRALIA: South Australia, Cape Gantheaume Conservation Park, Murray Lagoon, 3 October 1997; KGM SA-10, 12, OXAU-04, AUSTRALIA: South Australia, Cape Gantheaume Conservation Park, Murray Lagoon, 3 October 1997; MV 1991, OXAU-01, AUSTRALIA: Victoria

White-Headed Duck (*Oxyura leucocephala*):

EBD-02, 03, 04, 05, 07, 08, 11, 12, 13, 15, OXLE-01, CAPTIVE: Estacion Biologica de Doñana, Spain; SHW-MDS , OXLE-01, CAPTIVE: Sylvan Heights Waterfowl, USA

Maccoa Duck (*Oxyura maccoa*):

TM 77818, OXMA-02, SOUTH AFRICA: Gauteng, Randburg, 18 February 1999; TM 77819, OXMA-01, SOUTH AFRICA: Gauteng, Randburg, 18 February 1999; TWT-01, 02, 04, 05, OXMA-01, CAPTIVE: Treehaven Wildfowl Trust, South Africa; WWT-LC0305, SC5119, SC5395, SC5396, SC5399, SC5534, OXMA-01, CAPTIVE: Wildfowl and Wetlands Trust, UK; SHW-MDS , OXMA-01, CAPTIVE: Sylvan Heights Waterfowl, USA

[a]Collections and waterfowl aviary abbreviations: Estacion Biologica de Doñana (EBD); Field Museum of Natural History (FMNH); Kevin G. McCracken Field Catalog (KGM); Louisiana State University Museum of Natural Science (LSUMZ); Bob Brua Manitoba Field Catalog (MN); Museum of Victoria (MV); Sylvan Heights Waterfowl (SHW); Transvaal Museum (TM); Treehaven Wildfowl Trust (TWT); United States Fish and Wildlife Service (FWS); University of Alaska Museum (UAM); and Wildfowl and Wetlands Trust (WWT). Estacion Biologica de Doñana, Sylvan Heights Waterfowl, Treehaven Wildfowl Trust, and Wildfowl and Wetlands Trust samples came from captives; all other samples are from wild ducks.

[b]DNA was extracted from the base of a feather quill plucked from a museum skin.

[c]*O. j. andina* L78–H493 sequence identical to OXFE-02, OXFE-05, and OXFE-07 haplotypes.

[d]*O. j. andina* L78–H493 sequence identical to OXJA-05 haplotype.

[e]*O. j. andina* L78–H493 sequence identical to OXJA-08, OXJA-09, and OXJA-10 haplotypes.

APPENDIX 2. Primers used to amplify and sequence mtDNA and nuclear DNA from *Nomonyx* and *Oxyura* stiff-tailed ducks.

Locus	Primer	5′ to 3′ primer sequence	Reference
Control region, tRNA-Phe	L78	GTTATTTGGTTATGCATATCGTG	Sorenson and Fleischer, 1996
	OXJACRL1	CCAACACCATTAACATGAATGC	
	L432	GTACACCTCACGTGAAATCAG	
	OXJACRH1	RGARRTRYTGCRGGTTATGTCG	
	H493	AAAATGTGAGGAGGGCGAGG	Sorenson and Fleischer, 1996
	NODOCRL1	TGGGCACACCTGCGTCATC	
	L736	ATCTAAGCCTGGACACACCTG	
	H774	CCATATACGCCAACCGTCTC	Sorenson et al., 1999
	H1251	TCTTGGCATCTTCAGTGCCRTGC	Sorenson et al., 1999
	NODO12SH	GCGTTTATGTACCATCTTGGCG	
tRNA-Phe, 12S rRNA, tRNA-Val	L1267	YAAAGCATGRCACTGAAGHYG	Sorenson et al., 1999
	L1753	AAACTGGGATTAGATACCCCACTAT	Sorenson et al., 1999
	H1858	TCGATTATAGAACAGGCTCCTCTAG	Sorenson et al., 1999
	H2294	TYTCAGGYGTARGCTGARTGCTT	Sorenson et al., 1999
tRNA-Met, ND2, tRNA-Trp	L5216	GGCCCATACCCCGRAAATG	
	L5219	CCCATACCCCGAAAATGATG	Sorenson et al., 1999
	L5758	GGCTGAATRGGMCTNAAYCARAC	Sorenson et al., 1999
	H5766	RGAKGAGAARGCYAGGATYTTKCG	
	H6313	CTCTTATTTAAGGCTTTGAAGGC	Sorenson et al., 1999
ND5, cytochrome b, tRNA-Thr	L14770	AATAGGMCCMGAAGGMCTNGC	
	L14996	AAYATYTCWGYHTGATGAAAYTTYGG	
	H15295	CCTCAGAAKGATATYTGNCCTCAKGG	Sorenson et al., 1999
	L15413	GGGGGWTTYTCMGTNGAYAAYCC	
	H15646	GGNGTRAAGTTTTCTGGGTCNCC	
	H16064	CTTCANTYTTTGGYTTACAAGRCC	Sorenson et al., 1999
Chromosome Z chromo-ATPase/helicase/DNA-binding protein (CHD-Z)	CHD-Z.1272F	TCCAGAATATCTTCTGCTCC	
	CHD-Z.1237R	GAGAAACTGTGCAAAACAG	
Alpha enolase (α-ENOL)	α-ENOL.8F	GACTTCAAATCYCCYGATGAYCCCAG	
	α-ENOL.9R	CCAGTCRTCYTGGTCAAADGGRTCYTC	
Glyceraldehyde-3-phosphate dehydrogenase (GAPDH)	GAPDH.11F	TCCACCTTTGAYGCGGGTGCTGG	
	GAPDH.12R	CAAGTCCACAACACGGTTGCTGTATCC	
Hemoglobin α-A	Hα-A.F1	GGGCACCCGTGCTGGGGGCTGCCAAC	
	Hα-A.R2	TAACGGTACTTGGCAGTMAG	
Lactate dehydrogenase-B (LDHB)	LDHB.3F	GAAGAYAARCTNAARGGRGAAATGATGGA	
	LDBH.4R	CTCTGRACCAGGTTRAGRCGACTCTC	
Lamin A (Lamin-A)	LAMIN-A.3F	AAGAAGCARCTNCAGGATGAGATGCT	
	LAMIN-A.4R	GCCCRTTRTCRATCTCCACCAG	

APPENDIX 2.

Locus	Primer	5′ to 3′ primer sequence	Reference
Laminin receptor precursor/p40 (LRP/p40)	LRP/p40.5F	GGCCTGATGTGGTGGATGCTGGC	
	LRP/p40.6R	GCTTTCTCAGCAGCAGCCTGCTC	
Myelin proteolipid protein (MPP)	MPP.4F	TACATYTACTTYAAYACCTGGACCACCTG	
	MPP.5R	AGATGGAGAGNAGGTTGGAGCCACA	
Phosphoenolpyruvate carboxykinase (PEPCK3)	PEPCK3.F	GGTCGCTGGATGTCAGAAGAGG	
	PEPCK3.F2	TCAATACCAGATTCCCAGGCTGC	
	PEPCK3.R	CCATGCTGAAGGGGATGACATAC	
Phosphoenolpyruvate carboxykinase (PEPCK9)	GTP1601.F	ACGAGGCCTTTAACTGGCAGCA	S. E. Stanley (personal communication)
	PEPCK9.F	GGAGCAGCCATGAGATCTGAAGC	
	PEPCK9.R	GTGCCATGCTAAGCCAGTGGG	
	GTP1793.R	CTTGGCTGTCTTTCCGGAACC	S. E. Stanley (personal communication)