The Condor 115(2):316–329

MITOCHONDRIAL AND NEXT-GENERATION SEQUENCE DATA USED TO INFER PHYLOGENETIC RELATIONSHIPS AND SPECIES LIMITS IN THE CLAPPER/KING RAIL COMPLEX

JAMES M. MALEY[1] AND ROBB T. BRUMFIELD

Museum of Natural Science, 119 Foster Hall, Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803

Abstract. The family Rallidae is a distinct, species-rich group of birds, many of which are exceptional at long-distance colonization. Six of the ten species in the genus *Rallus* are distributed in the Americas. Among these, *R. longirostris* and *R. elegans* have an interwoven taxonomic history that reflects their weak phenotypic differentiation, the multiplicity of allopatric, morphologically distinct populations, a long zone of secondary contact in the eastern U.S. along which hybridization occurs, and apparent ecological "replacement" of each other in marshes of varying salinity. We used mitochondrial and nuclear gene sequences, the latter generated from next-generation sequencing, to infer phylogenetic relationships in the complex from a sample of 70 individuals collected throughout their distribution. Average levels of mitochondrial and nuclear divergence were relatively low (<2%) both within and between the species. The complex is composed of three distinct biogeographic groups: (1) eastern North America and the Caribbean, (2) South America, and (3) western North America, including Mexico. Our results indicate that *R. elegans* as currently recognized is paraphyletic, with birds of the highlands of Mexico sister to *R. longirostris* of California. *Rallus elegans* of eastern North America and Cuba is sister to *R. longirostris* from eastern North America and the Caribbean. This paraphyly, along with the reproductive isolation of the ecologically divergent *R. elegans* and *R. longirostris* in the eastern United States, supports splitting the complex into five morphologically and genetically distinct species.

Key words: *ancient DNA, divergence, next-generation sequencing, phylogeny,* Rallus longirostris, Rallus elegans, *species limits.*

Datos Mitocondriales y de la Próxima Generación Usados para Inferir Relaciones Filogenéticas y Límites de Especies en el Complejo *R. longirostris*/*R. elegans*

Resumen. La familia Rallidae es un grupo de aves distintivo y rico en especies, muchas de las cuales son excepcionales para la colonización a larga distancia. Seis de las nueve especies en el género *Rallus* están distribuidas en las Américas. Entre estas, *R. longirostris* y *R. elegans* tienen una historia taxonómica entretejida que refleja su débil diferenciación fenotípica, la multiplicidad de poblaciones alopátricas morfológicamente distintivas, una larga zona de contacto secundario en el este de EEUU a lo largo de la cual aparece hibridación y el aparente reemplazo ecológico de una por otra en marismas de salinidad variable. Empleamos secuencias de genes mitocondriales y nucleares, las últimas generadas a partir de la secuenciación de la próxima generación, para inferir las relaciones filogenéticas en el complejo a partir de una muestra de 70 individuos colectados a través de sus distribuciones. Los niveles medios de divergencia mitocondrial y nuclear fueron relativamente bajos (<2%) dentro y entre las especies. El complejo está compuesto por tres grupos biogeográficos distintivos: (1) este de América del Norte y el Caribe, (2) Sud América, y (3) oeste de América del Norte, incluyendo México. Nuestros resultados indican que *R. elegans*, como se la reconoce actualmente, es parafilética, con aves de las tierras altas de México hermanas de *R. longirostris* de California. *R. elegans* del este de América del Norte North y Cuba es hermana de *R. longirostris* del este de América del Norte y el Caribe. Esta parafilia, junto con el aislamiento reproductivo de las ecológicamente divergentes *R. elegans* y *R. longirostris* en el este de Estados Unidos, apoyan la separación del complejo en cinco especies morfológicamente y genéticamente distintivas.

INTRODUCTION

Members of the family Rallidae are characterized by exceptional long-distance dispersal and colonization, which is thought to contribute to their high species richness, especially on islands (Steadman 1995). The genus *Rallus* comprises ten species (Tavares et al. 2010). Six of the ten species occur in the New World, the other four in Africa, Madagascar, and Eurasia. Delimitation of species and subspecies in the *R. longirostris*/*R. elegans* complex is difficult because of their weak phenotypic differentiation. The two species differ

Manuscript received 26 August 2011; accepted 17 October 2012.
[1]E-mail: jmaley1@uwyo.edu

The Condor, Vol. 115, Number 2, pages 316–329. ISSN 0010-5422, electronic ISSN 1938-5422. DOI: 10.1525/cond.2013.110138

primarily in size and plumage color, the latter through a palette of reds, browns, and grays. There is also considerable morphological variation between and within populations. Introgression may play a role in the variation found in eastern North America. Additionally, occupancy of different habitats and the multiplicity of phenotypically distinct allopatric populations have confounded definitions of species limits in the group (Olson 1997). There has only been one published attempt to distinguish the two species in eastern North America genetically, and the results were not conclusive (Avise and Zink 1988).

Rallus longirostris includes three distinct groups of subspecies: six subspecies in the nominate *longirostris* group of the Pacific and Atlantic coasts of South America, eleven subspecies in the *crepitans* group of eastern North America, the Caribbean, and the Yucatan Peninsula, and six subspecies in the *obsoletus* group of western North America and the Gulf of California (Fig. 1, Olson 1997). *Rallus elegans* currently comprises three subspecies (Fig. 1): nominate *elegans* Audubon, 1834 of eastern North America, *ramsdeni* Ripley, 1913 of Cuba, and *tenuirostris* Ridgway, 1874 of the highlands of central Mexico.

All three main groups of *R. longirostris* breed in saltmarsh and mangrove habitats, their breast plumage ranges from dull silvery to dull rufous, and they are generally smaller than all subspecies of *R. elegans*. Some populations of the *R. l. obsoletus* group occur in freshwater habitats seasonally, and some populations of *R. e. elegans* occur seasonally in saltwater habitats. For example, *R. l. yumanensis* breeds in freshwater marshes along the Colorado River and in the Salton Sink (Fig. 1) but inhabits saltmarshes of the Gulf of California during its nonbreeding season (Tomlinson and Todd 1973, Banks and Tomlinson 1974). In contrast, *R. e. elegans* breeds in freshwater marshes, is brighter rufescent on the breast, and

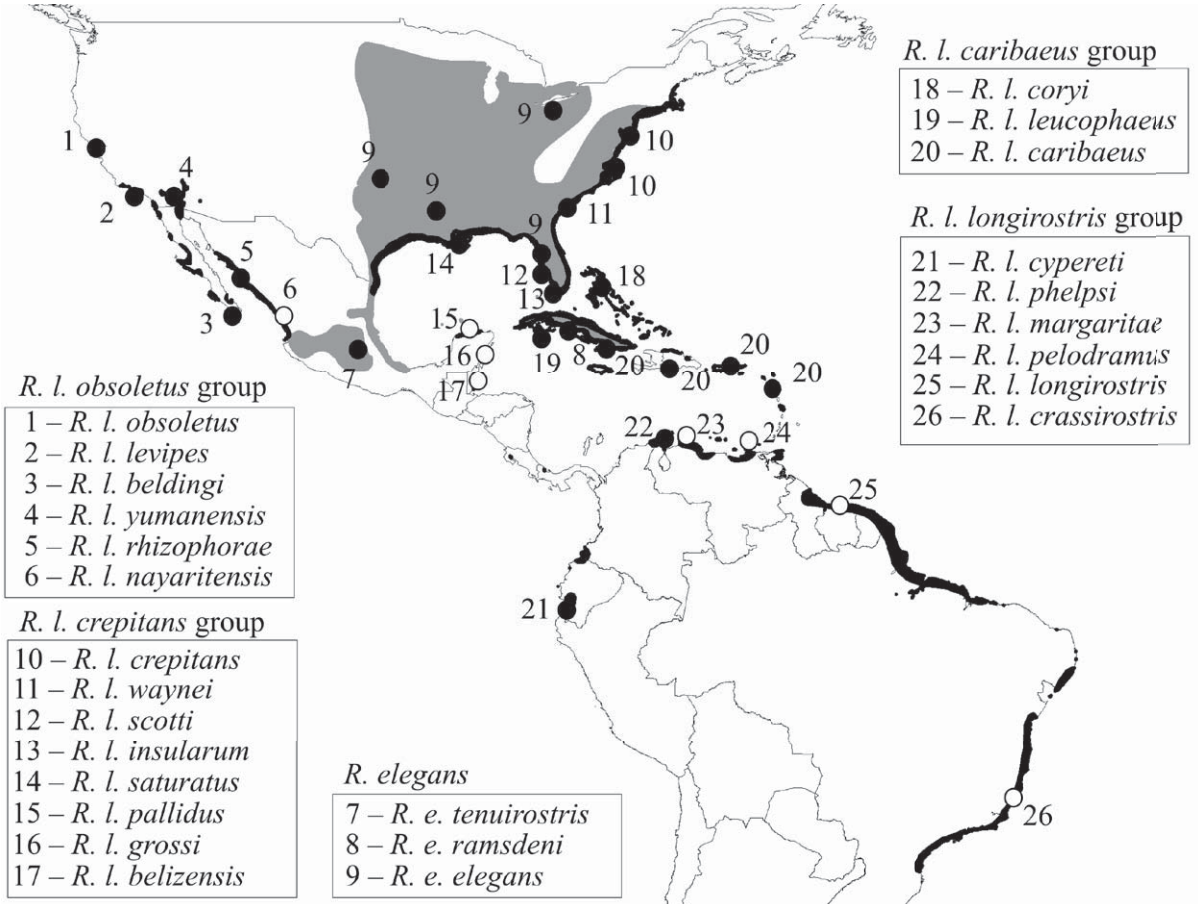

FIGURE 1. Distribution of sampling sites and approximate geographic distribution of *Rallus elegans* (gray) and *R. longirostris* (black), with points denoting subspecies (Taylor 1996). The gray and black distribution underlying the samples was downloaded from InfoNatura (Ridgely et al. 2005). White-filled circles represent random points along the distribution of subspecies that we were unable to sample; black-filled circles represent approximate locations of sampling. Subspecific identifiers are organized by group within the complex.

is larger than *R. longirostris* of all three groups (Olson 1997). *Rallus e. elegans* can be found in saltmarshes during migration and the boreal winter but apparently vacates saltmarshes during breeding (Meanley 1992).

Species limits in the group are of conservation interest because both species are hunted legally in the southeastern United States despite the endangered status of some migratory populations of *R. e. elegans* that presumably spend the nonbreeding season in the region. Additionally, the taxonomic status of the *R. l. obsoletus* group of the western United States and western Mexico is of particular interest because these populations are highly endangered (Eddleman and Conway 1998). The taxonomic affinity of this subspecies group has long been questioned because individuals are bright rufous ventrally like *R. e. elegans* but primarily inhabit saltwater marshes like most populations of *R. longirostris* (Olson 1997). Genetic relationships of *R. e. elegans* breeding in the eastern United States are of interest because migratory populations in the northern portion of the species' range are of conservation concern due to habitat loss (Cooper 2008). These birds may also be hunted during the nonbreeding season in southeastern North America, but no genetic studies have addressed whether migratory and nonmigratory populations are distinct, or where migratory populations spend the nonbreeding season. A review of the literature on the complex led Ripley (1977) to lump both species into *R. longirostris*, primarily on the basis of reports of extensive hybridization between *R. l. crepitans* and *R. e. elegans* in brackish marshes of the eastern U. S. (Meanley and Wetherbee 1962). Here we use DNA sequence data from the mitochondrion and four nuclear loci to infer a phylogeny of the subspecies of *R. longirostris* and *R. elegans*. We developed the nuclear loci by a reduced-representation library method and next-generation sequencing to attempt to develop phylogenetically informative markers. For some populations lacking tissues we sampled prepared museum specimens and sequenced mitochondrial DNA (mtDNA).

METHODS

SAMPLING

Through collecting and loans from other institutions we were able to include samples from throughout the distribution of the species complex (Table 1). We were unable to include samples from three of the six species of *Rallus* in the Americas (*R. antarcticus, wetmorei,* and *semiplumbeus*) because of the unavailability of any recent specimens or tissues and because they are highly restricted and/or critically endangered. In total, we sampled 14 of the 21 recognized subspecies of *R. longirostris* and all three subspecies of *R. elegans* (Fig. 1, Table 1). For populations lacking fresh tissue samples we scraped toe pads of older specimens housed at Louisiana State University's Museum of Natural Science (LSUMNS). These 16 samples ranged in collection date from 1937 to 1967 (Table 1). Each sample was cut with a new, sterile scalpel blade and sterilized scissors and forceps. We included two samples of *R. limicola* as an outgroup. All further references to "*R. l.*" in this paper indicate *R. longirostris*, not *R. limicola*.

We extracted total DNA from fresh tissue samples with a DNEasy Tissue Kit (Qiagen, Valencia, CA), following the manufacturer's protocols. DNA concentration was quantified with a NanoDrop spectrophotometer. Because of low yield, we concentrated extractions from feather samples with a heated Speedvac (Savant). For the toe-pad samples, we lysed a portion of each sample for 3 weeks in a facility for ancient DNA in Sabrina Taylor's lab at LSU's School of Renewable Natural Resources. These samples were extracted with two positive and two negative controls. Following lysis, we extracted ancient DNA with a DNEasy Tissue Kit (Qiagen) by following the instructions provided by the manufacturer. These extractions were then concentrated with a heated Speedvac.

MITOCHONDRIAL SEQUENCES

Using PCR amplicons from the primers L5215 (Hackett 1996) and H6313 (Johnson and Sorenson 1998), we sequenced eight samples of Louisiana birds (four of each species) for the protein-coding NADH subunit 2 (ND2). From these sequences we used the online resource Primer3 (Rozen and Skaletsky 2000) to design primers for a 787-base-pair (bp) portion of the ND2 gene (RallusND2F and RallusND2R, Table 2). For the toe-pad samples we used Primer3 to design 15 internal primers that allowed us to sequence eight small, overlapping portions (87–108 bp) of ND2 totaling 681 bp when concatenated.

All PCR amplifications were 10 μL in volume and contained ~50 ng of DNA, 1X Standard *Taq* Buffer, 0.13 mM $MgCl_2$, 0.5 μM of each primer, 0.1 mM dNTPs, 0.1 mg mL^{-1} bovine serum albumen (BSA), 0.25 units *Taq* DNA Polymerase (New England Biolabs), and 4.95 μL water. All PCRs were run as follows: 2 min at 94 °C, followed by 35 cycles of 94 °C for 30 sec, 51 °C for 30 sec (for ND2) or 60 °C (for nuclear [nu] DNA), and 72 °C for 1 min, followed by 72 °C for 10 min. After confirming amplification by running PCR products on 1% agarose gels, we cleaned the remaining PCR products with 10 μL ExoSAP reactions containing 5 μL of PCR product, 4.25 μL water, 1.25 units Antarctic Phosphatase, 5 units ExoI, and 1X Antarctic Phosphatase Buffer. For amplification of ancient DNA we followed the same protocol but substituted 0.1 mg mL^{-1} RSA (Rabbit Serum Albumin) for BSA and 0.25 units HotStar*Taq* (Qiagen) for *Taq*. We then used a Big Dye Terminator Cycle-Sequencing Kit, version 3.1 (Applied Biosystems, Foster City, CA), for cycle-sequencing reactions. All cycle-sequencing reactions included 1 μL of cleaned PCR product, 0.93 X Sequencing Buffer, 0.3 μL of 1X Big Dye 3.1, and 2×10^{-5} μmoles of forward or reverse primer, combined with distilled water to make a final volume of 7 μL. These reactions were run as follows: 2 cycles of 96 °C for 20 sec, 50 °C for 15 sec, and 60 °C for 4 min, followed

TABLE 1. Samples of *Rallus* used in this study. Groups refer to those pictured in Fig. 1. Voucher numbers are provided when available. Abbreviated IDs correspond to the tips in Fig. 2.

Species, group, subspecies, and voucher number[a]	Locality	Date	ID	Type
longirostris				
obsoletus				
obsoletus				
UWBM 87934	California: Marin Co.; Santa Venetia Marsh Open Space Preserve	7 Feb 2007	CALO6	Toe pad
levipes				
SDNHM 50948	California: San Diego Co.; San Diego, Mission Valley	29 Sep 2004	CALO2	Tissue
SDNHM 50958	California: San Diego Co.; Carlsbad	10 Jul 2004	CALO3	Tissue
SDNHM 51596	California: San Diego Co.; San Elijo Lagoon	14 Nov 2006	CALO4	Tissue
SDNHM 51955	California: San Diego Co.; San Elijo Lagoon	10 Mar 2008	CALO5	Tissue
beldingi				
UWBM 82691	Mexico: Baja California Sur; Municipio La Paz, La Paz	2 Sep 2006	MXLO2	Tissue
rhizophorae				
LSUMZ B-71222	Mexico: Sonora; Agiabampo	16 May 1937	MXLO1	Toe pad
yumanensis				
LACM 107400	California: Imperial Co.	19 Apr 1993	CALO1	Tissue
crepitans				
crepitans				
ANSP 22095	New Jersey: Cape May Co.; Dennis Township	9 Sep 1999	NJLO1	Tissue
ANSP 22479	New Jersey: Ocean Co.; Tuckerton Marsh	7 Oct 2000	NJLO2	Tissue
ANSP 25897	New Jersey: Cape May Co.; Nummy Island	26 May 2002	NJLO3	Tissue
ANSP 25898	New Jersey: Cape May Co.; Nummy Island	26 May 2002	NJLO4	Tissue
MBM 6009	North Carolina: Carteret Co.; Atlantic Beach	2 Nov 1998	NCLO1	Tissue
MBM 6010	North Carolina: Carteret Co.; Atlantic Beach	7 Sep 1998	NCLO2	Tissue
waynei				
KU 94320	Georgia: Glynn Co.; Brunswick, Blythe Island	12 Dec 2000	GALO1	Tissue
KU 94322	Georgia: Glynn Co.; Brunswick, Blythe Island	12 Dec 2000	GALO2	Tissue
KU 94324	Georgia: Glynn Co.; Brunswick, Blythe Island	12 Dec 2000	GALO3	Tissue
KU 94323	Georgia: Glynn Co.; Brunswick, Blythe Island	12 Dec 2000	GALO4	Tissue
KU 94321	Georgia: Glynn Co.; Brunswick, Blythe Island	12 Dec 2000	GALO5	Tissue
scottii				
KU 112081	Florida: Lee Co.; Sanibel	9 Mar 2007	FLLO2	Tissue
insularum				
FMNH 385736	Florida; Monroe Co.; Upper Keys	20 Nov 1994	FLLO1	Tissue
saturatus				
LSUMZ B-63400	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	9 Apr 2009	LALO1	Tissue
LSUMZ B-63401	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	9 Apr 2009	LALO2	Tissue
LSUMZ B-63402	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	3 Jun 2009	LALO3	Tissue
LSUMZ B-63403	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	3 Jun 2009	LALO4	Tissue
LSUMZ B-63404	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	4 Jun 2009	LALO5	Tissue
LSUMZ B-63405	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	4 Jun 2009	LALO6	Tissue
LSUMZ B-63406	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	4 Jun 2009	LALO7	Tissue
LSUMZ B-63407	Louisiana: Jefferson Par.; Grand Isle, ~1.6 km W Grand Isle	4 Jun 2009	LALO8	Tissue
LSUMZ B-63474	Louisiana: Lafourche Par.; 9 km E Port Fourchon	25 Jun 2009	LALO9	Tissue
LSUMZ B-63475	Louisiana: Lafourche Par.; 9 km E Port Fourchon	25 Jun 2009	LALO10	Tissue
caribaeus				
caribaeus				
LSUMZ B-71223	Puerto Rico: Mayagüez; Guanajibo, Caño Corazones	18 Dec 1942	CBLO1	Toe pad
LSUMZ B-71224	Cuba: Las Villas; Cayo Baíia de Cádiz	27 Jun 1958	CULO1	Toe pad
LSUMZ B-71225	Cuba: Las Villas; Laguna de San Mateo	8 Jul 1960	CULO2	Toe pad
LSUMZ B-71226	Cuba: Las Villas; Laguna de San Mateo	8 Jul 1960	CULO3	Toe pad
coryi				
LSUMZ B-71227	Bahamas: San Salvador; lake 1.6 km E Cockburn Town	29 Dec 1963	CBLO2	Toe pad
leucophaeus				
LSUMZ B-71228	Cuba: Isle of Pines; Siguanea Bay	9 Jul 1958	CULO4	Toe pad

(Continued)

TABLE 1. Continued.

Species, group, subspecies, and voucher number[a]	Locality	Date	ID	Type
LSUMZ B-71230	British Virgin Islands: Beef Island; western end	12 Aug 1964	CBLO3	Toe pad
LSUMZ B-71231	Antigua: St. John Par.; The Flashes	24 Mar 1962	CBLO4	Toe pad
KU 95142	Dominican Republic: Monte Cristi National Park	28 Mar 2003	CBLO5	Tissue
longirostris				
cypereti				
LSUMZ B-66005	Peru: Depto. Tumbes; ~29 km NE Tumbes	1 Jun 2009	PELO1	Tissue
LSUMZ B-66008	Peru: Depto. Tumbes; ~29 km NE Tumbes	1 Jun 2009	PELO2	Tissue
LSUMZ B-67817	Peru: Depto. Tumbes; ~2 km S Santuario Nacional los Manglares de Tumbes	26 Jul 2009	PELO3	Tissue
LSUMZ B-67818	Peru: Depto. Tumbes; ~2 km S Santuario Nacional los Manglares de Tumbes	26 Jul 2009	PELO4	Tissue
LSUMZ B-67819	Peru: Depto. Tumbes; ~2 km S Santuario Nacional los Manglares de Tumbes	26 Jul 2009	PELO5	Tissue
LSUMZ B-67820	Peru: Depto. Tumbes; ~2 km S Santuario Nacional los Manglares de Tumbes	26 Jul 2009	PELO6	Tissue
phelpsi				
UMMZ 620	Venezuela: Zulia; Ancon de Iturre, ~1 km NE of town	11 Feb 1989	VZLO1	Tissue
elegans				
elegans				
elegans				
LSUMZ B-49394	Florida: Lee Co.; Fort Meyers	9 Jul 2002	FLEL1	Tissue
LSUMZ B-63464	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL1	Tissue
LSUMZ B-63465	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL2	Tissue
LSUMZ B-63466	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL3	Tissue
LSUMZ B-63467	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL4	Tissue
LSUMZ B-63468	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL5	Tissue
LSUMZ B-63469	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	23 Jun 2009	LAEL6	Tissue
LSUMZ B-63470	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	24 Jun 2009	LAEL7	Tissue
LSUMZ B-63471	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	24 Jun 2009	LAEL8	Tissue
LSUMZ B-63472	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	24 Jun 2009	LAEL9	Tissue
LSUMZ B-63473	Louisiana: St. Landry Par.; 6 km NNW Church Point, Dusty Rd.	24 Jun 2009	LAEL10	Tissue
ramsdeni				
LSUMZ B-71232	Cuba: Las Villas; El Jibaro, Sancti Spiritus	14 May 1967	CUEL1	Toe pad
tenuirostris				
LSUMZ B-71233	Mexico; San Luis Potosi; Media Luna	28 Jun 1952	MXEL1	Toe pad
LSUMZ B-71234	Mexico; San Luis Potosi; Villa de Reyes Reg, Laguna de las Rusias, 1840 m	6 Aug 1947	MXEL2	Toe pad
LSUMZ B-71235	Mexico: Michoacan, near Cumuato	21 Nov 1959	MXEL3	Toe pad
LSUMZ B-71236	Mexico: Michoacan, near Cumuato	10 Nov 1959	MXEL4	Toe pad
LSUMZ B-71237	Mexico: Michoacan, near Cumuato	21 Nov 1959	MXEL5	Toe pad
elegans				
LSUMZ B-71305	Ohio: Ottawa Co.; Ottawa National Wildlife Refuge	3 May 2010	OHEL1	Feather
LSUMZ B-71308	Oklahoma: McCurtain Co.; Red Slough Wildlife Management Area	11 Apr 2010	OKEL1	Feather
LSUMZ B-71309	Oklahoma: McCurtain Co.; Red Slough Wildlife Management Area	23 Mar 2010	OKEL2	Feather
FMNH 461050	Illinois: Cook Co.; Evanston	5 Jun 2008	ILEL1	Tissue
limicola				
limicola				
LSUMZ B-25122	Louisiana: Lafayette Par.; Potter's Road, 2.4 km W Hwy. 719	2 Jan 1995	*limicola*	Tissue
LSUMZ B-46858	Louisiana: Iberia Par.; 7.8 km S Lydia, Hwy. 83	26 Oct 2005	*limicola*	Tissue

[a]LSUMZ = Louisiana Museum of Natural Science; FMNH = Field Museum of Natural History; AR = University of Arkansas; UWBM = Burke Museum of Natural History and Culture, University of Washington; KU = University of Kansas Natural History Museum; ANSP = Academy of Natural Sciences of Philadelphia; MBM = Marjorie Barrick Museum of Natural History, University of Nevada, Las Vegas; SDNHM = San Diego Natural History Museum; UMMZ = University of Michigan Museum of Zoology; LACM = Natural History Museum of Los Angeles County.

TABLE 2. Primers used to amplify mitochondrial and nuclear loci in *Rallus* and in the AFLP method to reduce the genome. All primers were designed with Primer3 (Rozen and Skaletsky 2000), except AFLP primers (Vos et al. 1995).

Primer	Locus	Sequence (5′ to 3′)
RallusND2F	ND2–external	CCCACATGCAAGCCTAATCT
RallusND2IF2	ND2–internal	CCTTATCTCAAAATCCCACCA
RallusND2IF3	ND2–internal	TCCACACTCCTACTATTTTCCAG
RallusND2IF4	ND2–internal	CCAACAGCCTCCATCCTACT
RallusND2IF5	ND2–internal	CAGAAGTACTACAGGGCACATCC
RallusND2IF6	ND2–internal	CCCATTCACTTAACCCAACC
RallusND2IF7	ND2–internal	CCAGACTCAACTCCGAAAAA
RallusND2IF8	ND2–internal	GCCTTTTACCTCTACTGCTTAACAA
RallusND2IR1	ND2–internal	ATTTGATTGCTGCTTCGGTAG
RallusND2IR2	ND2–internal	CCCATTGTCCTGTGTGTCAG
RallusND2IR3	ND2–internal	GGGAATCAAAAGTGGAATGG
RallusND2IR4	ND2–internal	TGGAGGGAGTTTTATTATTGTGG
RallusND2IR5	ND2–internal	TGCTGAAAGGATGGCTAGTG
RallusND2IR6	ND2–internal	TCCCAGGTGAGAAATAGACGA
RallusND2IR7	ND2–internal	TGTTGATGGTGAGGAAAATGG
RallusND2IR8	ND2–internal	CCTGCAAGGGATAGCAGTGT
RallusND2R	ND2–external	GGTTGGGTGGAAGTGTGATT
F139	139–external	CTGGGAGTGAGGGCAGAG
R139	139–external	CCTTGCAAGTTGATGGATGA
F472	472–external	GGTATGCAGCTCACCACAGA
R472	472–external	GCAAGAAAGAGGCTGTGTCC
F1166	1166–external	CCTGTGCTCCTGCTGATGTA
R1166	1166–external	CAACCAGATGTTGCAGGAGA
F1766	1766–external	GCAGCAAACTAGGAGCAGTACA
R1766	1766–external	CTTCAGCCTCCCATCAAGG
EcoRI-F adaptor	N/A	CTCGTAGACTGCGTACC
EcoRI-R adaptor	N/A	AATTGGTACGCAGTCTAC
MseI-F adaptor	N/A	GACGATGAGTCCTGAG
MseI-R adaptor	N/A	TACTCAGGACTCAT
EcoRI primer	N/A	GACTGCGTACCAATTC
MseI primer	N/A	GATGAGTCCTGAGTAA
EcoRI fusion primer[a]	N/A	CGTATCGCCTCCCTCGCGCCATCAGXXXXXXXXXXGACTGCGTACCAATTC
MseI fusion primer[b]	N/A	CTATGCGCCTTGCCAGCCCGCTCAGGATGAGTCCTGAGTAAN

[a]X denotes different barcode sequence per individual.
[b]N indicates selective base; primer is biotinylated on the 5′ end.

by 24 cycles of 96 °C for 12 sec, 50 °C for 15 sec, and 60 °C for 4 min. Cycle-sequence products were then purified with a Sephadex G-50 Fine (Sigma-Aldrich., St. Louis, MO) matrix in a 96-well filter plate. Cleaned products were run on a 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA), and sequences were aligned and checked for quality with Sequencher version 4.7 (Gene Codes Corp., Ann Arbor, MI).

NUCLEAR DNA SEQUENCES

Using four Louisiana samples of each species (LALO1–4; LAEL1–4; Table 1) we amplified and sequenced 11 nuclear markers, *ADAMTS6*, *SLC30A5* (Backstrom et al. 2006), *ALDOB4* (Sætre et al. 2003), *ARNTL*, *PER2*, *VIM* (Kimball et al. 2009), *BF5* (Weissbach et al. 1991), *BF7* (Prychitko and Moore 1997), *G3PDH* (Fjeldså et al. 2003), *MYO2* (Slade et al. 1993), and *VLDLR9* (Borge et al. 2005), and attempted to amplify an additional nine nuDNA markers, *AQP5*, *ARVPA*, *BRM15*, *CLTC*, *CRYAA*, *IPOII*, *MSHR*, *PPWDI*, and *RHO1* (International Chicken Genome Sequencing Consortium 2004). We found no informative single-nucleotide polymorphisms (SNPs) among the 11 loci. Some loci were completely invariant, while others had one to three unique SNPs found only in one individual. To develop informative nuDNA markers with SNPs in which the two species differed in frequency by >0.7, we used an AFLP-based genome-reduction technique combined with 454 pyrosequencing (McCormack et al. 2012). Because we used this method to develop markers for analysis of the hybrid zone in Louisiana, we took next-generation data only from members of the *R. l. crepitans* group and nominate *R. elegans* from Louisiana.

We prepared samples from 20 individuals, 10 of each species collected in Louisiana (LALO1–LALO10; LAEL1–LAEL10; Table 1), to be run on a quarter plate of a 454 pyrosequencing run. We standardized initial DNA concentrations at 100 ng μL^{-1}. Then, in a single step, 250 ng DNA was digested and adaptors (Table 2; Vos et al. 1995) were ligated on the resulting sticky ends for 2 hr at 37 °C in an 11-μL volume reaction containing 3.47 μL water, 2.5 μL DNA template, 1.1 μL T4 ligase buffer, 0.05 M NaCl, 5 units EcoRI, 1 unit MseI, 0.05 mg mL^{-1} BSA, 5 units T4 ligase (New England Biolabs), and 0.91 μM adaptor for both MseI and EcoRI (all concentrations are final). We made each adaptor beforehand by heating equal volumes of two complementary pieces of DNA to 95 °C and allowing them to cool slowly to room temperature. We then conducted a round of PCR in a 20-μL reaction containing 4 μL of a 10-fold dilution of the digest-ligation product, 10.08 μL water, 1.5 mM $MgCl_2$, 1X Phusion buffer, 0.2 mM dNTPs, 0.3 μM of each adaptor-specific primer (Table 2; Vos et al. 1995), and 0.4 units Phusion High-fidelity DNA Polymerase (Finnzymes, Woburn, MA). The PCR protocol was 2 min at 72 °C followed by 15 cycles of 98 °C for 30 sec, 56 °C for 30 sec, and 72 °C for 2 min, followed by 72 °C for 10 min. For each individual, resulting PCR products were visualized on a 0.8% SeaKem GTG Agarose (Lonza, Rockland, ME) gel. A section of the resulting smear of DNA fragments (400–550 bp) was excised. During this step DNA from each individual was separated by a well containing Quick-Load 100-bp DNA ladder (New England Biolabs) to ensure alignment and to avoid cross-contamination. These samples were column purified with a QIAquick gel-extraction kit (Qiagen), eluted with 50 μL volume, and then subjected to another round of PCR, this time with longer "fusion primers" (Table 2). These primers include the complementary adaptor-priming sites and overhanging DNA sequence containing the necessary binding sites for emulsion PCR and individual-identifying barcodes (20 different primers, each with a different barcode sequence). The MseI fusion primer was biotinylated on the 5′ end and included one selective base to reduce the genome further. The primer was biotinylated to aid in removing nontargeted short fragments and to remove any EcoRI–EcoRI fragments. This PCR was 10 μL in volume and contained 2.5 μL of eluted PCR product, 3.94 μL water, 1.5 mM $MgCl_2$, 1X Phusion buffer, 0.2 mM dNTPs, 0.3 μM of MseI reverse fusion primer (Table 2), and 0.2 units Phusion Polymerase. To each reaction, we separately added 0.3 μM of EcoRI forward fusion primer with indexes (Table 2). For this PCR, we used a touchdown profile beginning with 98 °C for 2 min, then 12 cycles of 98 °C for 30 sec, 65 °C for 30 sec (reducing temperature by 0.7 °C in each cycle), 72 °C for 2 min, then 10 cycles of 98 °C for 30 sec, 56° C for 30 sec, and 72 °C for 2 min, followed by 10 min at 72 °C.

Conversion of the raw 454 output into loci and alleles for each individual was performed with the computer pipeline PRGmatic version 1.4 (Hird et al. 2011). The data were processed and screened for fixed or nearly fixed polymorphic sites. We identified 34 loci containing SNPs for which allele-frequency differences between the two species were high (>0.70; mean = 0.90, SD = 0.08). We chose four of these loci that contained between one and four fixed or nearly fixed SNPs (frequency > 0.94; mean = 0.96) between *R. elegans* and *R. l. saturatus* for Sanger sequencing. We found no evidence of paralogous gene copies in that there were no variable positions with more than two nucleotides and all of the loci aligned with avian autosomal intron sequences when searched with BLAST (Table 3; International Chicken Genome Sequencing Consortium 2004, Warren et al. 2010). We also assessed paralogy by using PRGmatic version 1.4 to calculate heterozygosities observed and expected within a species. The observed heterozygosity of paralogous loci should be much greater than expected, but we found that the observed heterozygosity of all four loci was equal or slightly higher than expected. We used Primer3 to design primers directly from 454 reads (Rozen and Skaletsky 2000); the loci are named for the contig order from

TABLE 3. BLAST results for the four nuclear genes used in this study. Sequences were compared to the two available avian genomic reference sequences (International Chicken Genome Sequencing Consortium 2004), Warren et al. 2010). All four align to intron sequences and do not align closely to anything but bird sequence. Only alignments with E values lower than 1.00E-05 are shown.

	Gallus gallus			*Taeniopygia guttata*		
Locus	Chromosome	Feature	*E* value	Chromosome	Feature	*E* value
139	26	54845 bp at 3′ side: green sensitive cone opsin	1.0×10^{-8}	none	none	none
472	13	similar to macrophage colony-stimulating factor I receptor	1.0×10^{-62}	13	similar to colony-stimulating factor 1	2.0×10^{-61}
1166	none	none	none	19	237690 bp at 5′ side: ligase III, DNA, ATP-dependent, 68044 bp at 3′ side: transmembrane protein 132E	1.0×10^{-24}
1766	24	centromere/kinetochore protein zw10	2.0×10^{-21}	24	centromere/kinetochore protein zw10	1.0×10^{-21}

the data-assembly process (139, 472, 1166, 1766; Table 2). See above for PCR conditions for nuclear loci.

We aligned DNA sequences with Sequencher 4.7 (GeneCodes) and detected no indels in any sequence. We coded heterozygous sites in the nuDNA sequences with standard IUPAC ambiguity codes.

PHYLOGENETIC ANALYSIS

We phased nuDNA haplotypes in each of the four datasets (see below) with PHASE 2.1.1 (Stephens et al. 2001, Stephens and Donnelly 2003, Stephens and Scheet 2005) after preparing the datasets with the online tool SeqPHASE (Flot 2010). The PHASE runs included a threshold of 0.9 probabilities for acceptance. We constructed median-joining haplotype networks for each locus (including mtDNA) with the program Network version 4.6 (Bandelt et al. 1999).

We first identified the best-fit finite-sites model of sequence evolution for each of the loci by using MrModelTest (Nylander 2004) with PAUP* version 4.0b10 (Swofford 1998) under the Akaike information criterion (AIC). Using these optimal models we then estimated gene trees and species trees with BEAST version 1.6.1 (Drummond and Rambaut 2007) and *BEAST (Heled and Drummond 2010), respectively. We estimated a gene tree and a species tree for the full mtDNA dataset and estimated species trees for the mtDNA and nuDNA datasets and the nuDNA datasets alone. We define the first dataset as the full mtDNA dataset, which includes all of the samples sequenced from fresh tissue and feathers, as well as those sequenced from toe pads. We refer to the second dataset as mtDNA and nuDNA; it consists of the mtDNA sequences from all samples excluding the samples from toe pads and the nuDNA sequences from all four loci from the same samples. The third dataset, nuDNA only, is just the four nuDNA loci sequences from all samples with fresh tissue. We estimated clock rates for each locus in each dataset, used a strict clock, and kept the substitution rates at 1. The three analyses were run three times for 10^7 generations, the scaling factors being optimized successively with each run. Following optimization we ran each analysis for 200 million generations to ensure effective sample sizes reached at least 200 for each parameter. We checked for convergence of parameters in Tracer version 1.5 (Rambaut and Drummond 2007) and found the Maximum Clade Credibility tree with TreeAnnotater version 1.6.1 (Drummond and Rambaut 2007), discarding the first 10% of trees. Species trees were delimited on the basis of major lineages recovered in the mtDNA gene tree.

We estimated the maximum-likelihood tree for the full mtDNA dataset with Garli version 1.0 (Zwickl 2006). We also performed a maximum-likelihood bootstrap analysis with 100 replicates for the full mtDNA dataset with the program Garli version 0.951 (Zwickl 2006). We conducted partition-homogeneity tests for all possible combinations of the five loci, treating each as a separate partition, with 100 replicates in PAUP* version 4.0b10 (Swofford 1998). We found significant heterogeneity in all tests excluding one or two loci (P = 0.01 or 0.02) and significant heterogeneity for seven out of ten pairs of loci (P = 0.01–0.04). Because we found significant heterogeneity for the majority of locus combinations, we did not run any analyses using concatenations.

RESULTS

The final mtDNA alignment included 787 bp of ND2 for the 54 DNA samples extracted from fresh tissue or feathers (Table 1) and 681 bp for 15 of the 16 toe-pad samples. We were able to sequence only 656 bp from one individual (CBLO2, Table 1). For the 54 samples extracted from fresh tissue we aligned 817 bp of nuDNA sequences (n = 216; 2 alleles per individual for four loci from 54 individuals) from the four autosomal introns. Within the complex, levels of uncorrected divergence in the mtDNA data were relatively low, ranging from 0.3% to 2.1% (Table 4). Jukes–Cantor-corrected distances for synonymous sites calculated with DnaSP version 5 (Librado and Rozas 2009) only showed higher levels of divergence, ranging from 0.6% to 6.0% (Table 4).

TABLE 4. Mitochondrial genetic distances between groups of samples of Clapper and King Rails. Jukes–Cantor-corrected distances for synonymous sites only are above the diagonal. Below the diagonal are uncorrected (p) distances.

	R. e. elegans	*R. e. ramsdeni*	*R. l. crepitans*	*R. l. caribaeus* Cuba	*R. l. caribaeus* Dominican Republic	*R. l. cypereti*	*R. l. phelpsi*	*R. l. obsoletus*	*R. e. tenuirostris*
R. e. elegans	—	0.012	0.035	0.029	0.017	0.035	0.035	0.054	0.047
R. e. ramsdeni	0.006	—	0.023	0.017	0.006	0.023	0.035	0.041	0.035
R. l. crepitans	0.01	0.012	—	0.006	0.017	0.035	0.035	0.053	0.047
R. l. caribaeus Cuba	0.01	0.01	0.001	—	0.012	0.029	0.029	0.06	0.054
R. l. caribaeus Dominican Republic	0.006	0.007	0.006	0.006	—	0.029	0.029	0.047	0.041
R. l. cypereti	0.01	0.01	0.013	0.012	0.011	—	0.012	0.041	0.035
R. l. phelpsi	0.009	0.01	0.011	0.012	0.01	0.004	—	0.041	0.035
R. l. obsoletus	0.015	0.016	0.018	0.021	0.017	0.013	0.011	—	0.006
R. e. tenuirostris	0.013	0.013	0.016	0.018	0.015	0.012	0.012	0.003	—
R. limicola	0.156	0.155	0.161	0.163	0.16	0.152	0.155	0.156	0.157

Of the 216 nuDNA sequences generated, most were homozygotes (n = 167). With the exception of one SNP in one individual we were able to phase most (n = 93) nuDNA haplotypes with a probability of 1.0 and the rest (n = 4) with a probability equal to or greater than 0.93. Only one SNP was called below the 0.9 threshold, an allele with a phase probability of 0.67 that was detected in one individual. Networks constructed with or without this allele were essentially identical, so we used the allele with a probability of 0.67 in all subsequent analyses.

The optimal model of sequence evolution for the full mtDNA dataset was GTR + Γ. The full mtDNA dataset refers to sequences from all extracted samples, including samples from toe pads. We found the same model of sequence evolution by using AIC whether the sequences from toe pads were included or excluded. Using AIC, we selected different models for each of the nuDNA loci (*139*, JC + I; *472*, K80; *1166*, F81; *1766*, K80 + I).

Bayesian phylogenetic analyses of the full mtDNA data set in Beast version 1.6.1 (Drummond and Rambaut 2007) resulted in several well-supported nodes in the gene tree (Fig. 2). Maximum-likelihood analyses yielded a very similar topology. However, bootstrap support for many clades was low (Fig. 2). We used *BEAST (Heled and Drummond 2010), as implemented in BEAST version 1.6.1 (Drummond and Rambaut 2007), to estimate a species tree for the full mtDNA dataset (Fig. 2), for the mtDNA and nuDNA datasets combined, and for the nuDNA datasets combined. Both trees including nuDNA data had the same topology, which was different from the full mtDNA species tree. Nodal probability values were

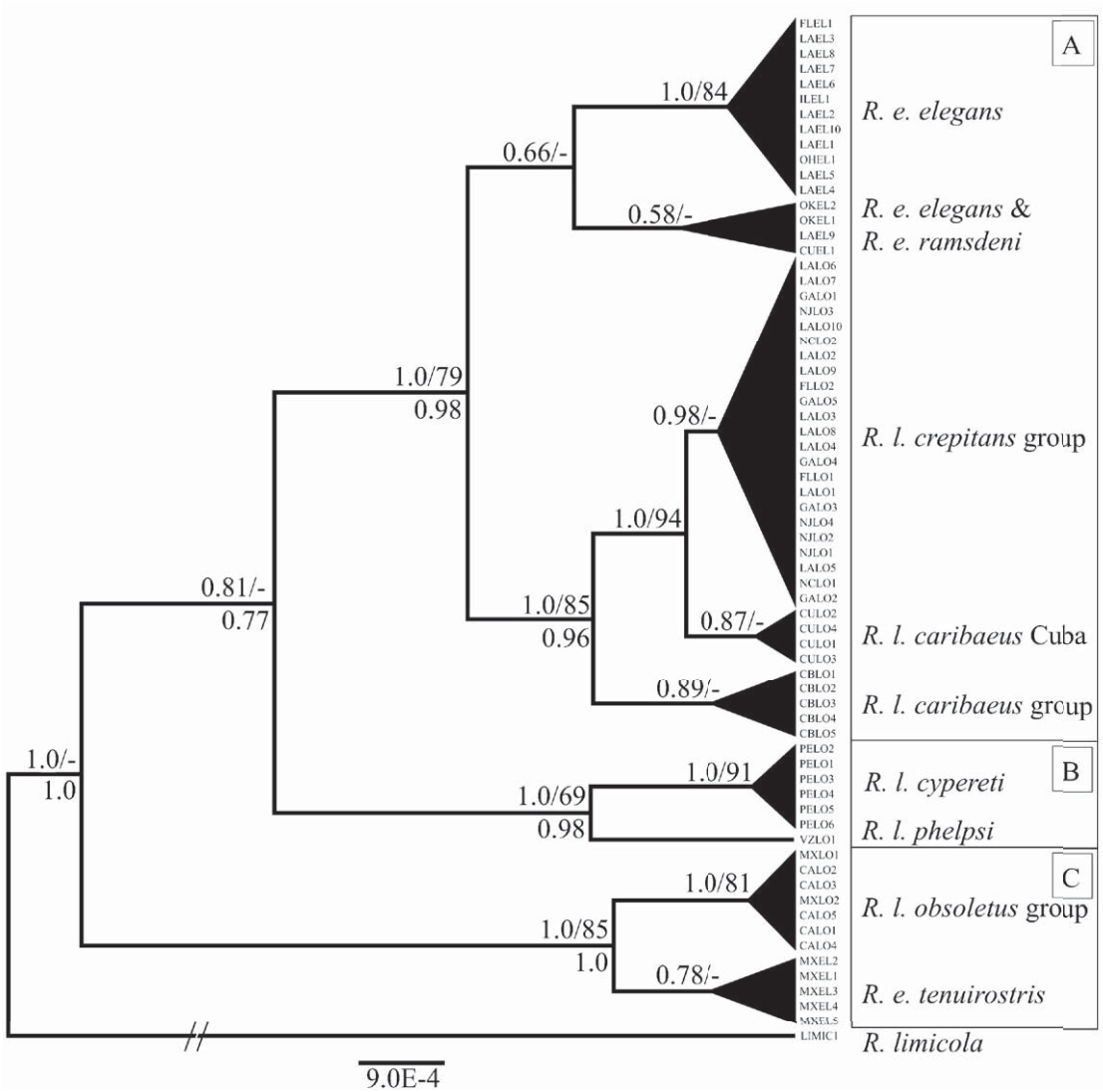

FIGURE 2. Maximum Clade Credibility gene tree of ND2 inferred in Beast (Drummond and Rambaut 2007). The labels above nodes are the posterior probability followed by the bootstrap support value (if greater than 65) for that node. The labels below nodes are the posterior probability for that node in the estimate of the species tree; this label is not included if the value was below 0.95. Each of the three major clades is outlined and labeled by geography, with clade A comprising eastern North American and Caribbean birds, clade B comprising South American birds, and clade C comprising birds of western North America, including Mexico.

low (<0.70) throughout the trees, with the exception of a sister relationship between *R. l. caribaeus* and the *R. l. crepitans* group (posterior probability = 0.99), so we do not present these trees here.

The entire species complex was well supported as distinct from *R. limicola* (Fig. 2), not surprisingly given the deep mtDNA divergence between the complex and *R. limicola* (15.2–16.3%; Table 4). There are three main lineages in the mtDNA topology, consisting of an eastern North American and Caribbean clade (A; Fig. 2), a South American clade (B; Fig. 2), and a western North American clade (C; Fig. 2). Clade A, comprising *R. e. elegans*, *ramsdeni*, and members of the *R. l. caribaeus* and *crepitans* groups, was well supported. Within this clade, a subset of *R. e. elegans* had high support as a distinct clade, as did a clade including members of the *R. l. caribaeus* and *crepitans* groups. A clade comprised solely of members of the *R. l. crepitans* group also had high support (Fig. 2). Clade B included only the South American birds (*R. l. cypereti* and *phelpsi*, both members of the *R. l. longirostris* group) and was well supported in Bayesian analyses (posterior probability = 1.0), and received moderate bootstrap support (69%). Also, clade C included members of the *R. l. obsoletus* group and *R. e. tenuirostris* and was well supported in both Bayesian and likelihood analyses (Fig. 2). The mitochondrial haplotype network corroborated the above clades (Fig. 3). *R. e. tenuirostris*, whose weakly differentiated haplotypes grouped together but did not form a well-supported clade in the phylogenetic analyses, clustered together in the mtDNA haplotype network (Fig. 3E).

The nuDNA provided limited phylogenetic information, with little, if any, divergence between subspecies or species groups. The two eastern North American mtDNA lineages, *R. l. crepitans* group and *R. e. elegans*, generally had two or more haplotypes distinct from one another, with some haplotype sharing (Fig. 3A–D). This is expected because we chose the nuclear markers on the basis of the distinctiveness of the two groups. The sole member of the *R. l. caribaeus* group for which we had fresh tissue always shared nuclear haplotypes with members of the *R. l. crepitans* group (Fig. 3), consistent with the close relationship observed in the mtDNA data (Fig. 2). Members of the *R. l. obsoletus* group shared haplotypes with either *R. e. elegans* (Fig. 3B, D) or the *R. l. crepitans* group (Fig. 3A, C). The South American birds either had distinct haplotypes (Fig. 3A, B) or shared them with either *R. e. elegans* (Fig. 3D) or the *R. l. crepitans* group (Fig. 3C). Haplotypes of *R. limicola* were distinct from those of the *R. longirostris* complex at three loci, but, unexpectedly, they shared a haplotype with several members of the complex (Fig. 3C).

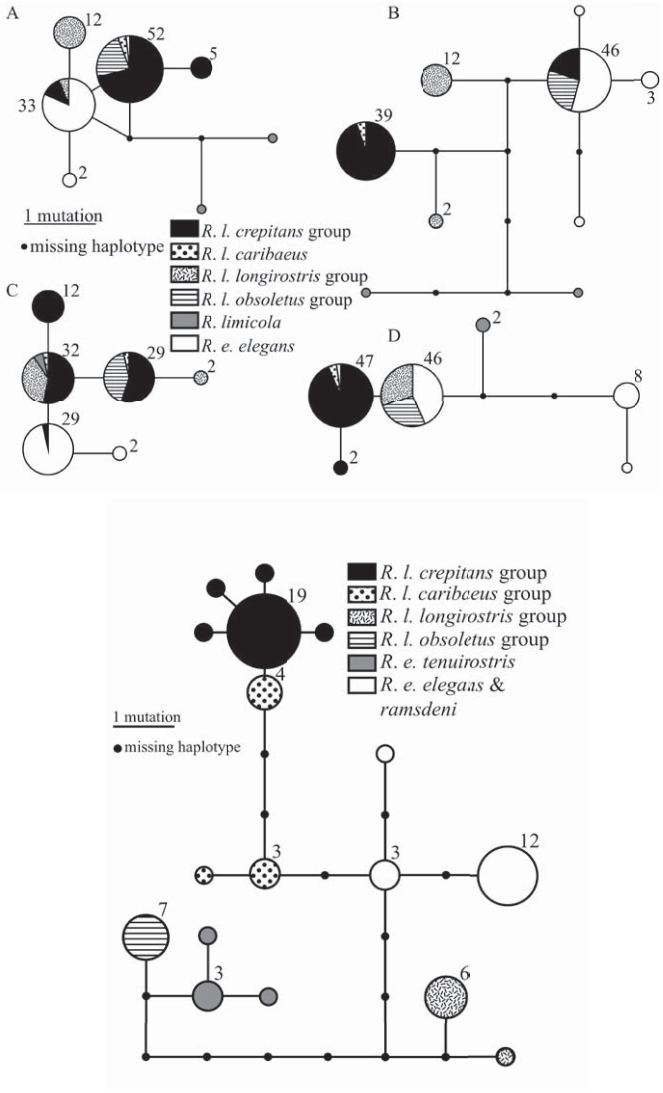

FIGURE 3. Median-joining haplotype networks inferred with Network version 4.6 (Bandelt et al. 1999) from the mitochondrial data and from data for each of the four nuclear loci generated by next-generation sequencing. Circles are sized proportionally to the number of alleles tallied for each haplotype. Numbers indicate the tally for a given haplotype, except when the tally was one. (A) Locus 139, (B) locus 472, (C) locus 1166, (D) locus 1766, (E) mtDNA. Samples are labeled by subspecies groups defined in the text and Figure 1.

DISCUSSION

Our primary findings are that *R. elegans*, as currently recognized, is paraphyletic with respect to *R. longirostris*, and species limits in the group correspond roughly to geography instead of current species designations. For example, the *R. l. obsoletus* subspecies group, endemic to California and northwestern Mexico, is sister to *R. e. tenuirostris* of the highlands of Mexico, instead of to either *R. l. crepitans* or *R. e. elegans* of eastern North America as implied in previous taxonomic treatments (Hellmayr and Conover 1942, Ripley 1977, Olson 1997). Additionally, the lineages of the *R. l. crepitans* group and *R. e. elegans*, which are known to hybridize in eastern North America (Olson 1997), are in the same clade. This

pattern of hybridization between members of these same two lineages apparently also occurs in Cuba (Olson 1997). This clade also includes birds of the Caribbean.

Despite the shallow mtDNA divergence between the *R. l. crepitans* group and nominate *R. elegans*, by using next-generation sequencing we were able to find fixed nuclear differences. These results do not reflect an overall pattern across the genome because, during preparation of the genomic library (see Methods), we biased the selection of nuclear loci to those in which *R. elegans* and *R. l. saturatus* are most divergent. That the set of loci recovered by this method provided very limited phylogenetic resolution across the tree suggests that researchers should include all divergent lineages when developing markers with this method (McCormack et al. 2013).

In terms of taxonomic recommendations there are two possibilities for the complex. Following a previous taxonomic evaluation of the group (Ripley 1977), both species could be combined under *R. longirostris*. We argue that this treatment obscures the fact that two members of the complex are in extensive secondary contact in eastern North America and Cuba but have not fused despite hybridizing (Olson 1997). Our morphological and genetic characterization of the hybrid zone in Louisiana (Maley 2012) found that it is very narrow (~4.2 km wide), with selection against hybrids acting to maintain it. These data suggest there is strong, albeit incomplete, reproductive isolation between these species in Louisiana. Extending these results to the remaining taxa and considering the different levels of morphological, ecological, and genetic divergence between previously identified subspecies groups suggests at least five species could be recognized in this complex. This treatment would be consistent with recent genetic analyses of other members of the Rallidae showing similar levels of divergence (Tavares et al. 2010, Goodman et al. 2011). The most divergent clade within the complex, according to mtDNA data, represents a pair of subspecies groups from both currently recognized species (*R. l. obsoletus* group and *R. e. tenuirostris*). This pair shares the same pattern observed in the birds of eastern North America, in which individuals of one group are smaller and found primarily in saltmarshes, while those of the other are larger, brighter, and found in freshwater habitats (Olson 1997).

We propose species status for five members of the complex described below. These taxonomic recommendations are based primarily on two factors: (1) that each of the species represents a morphologically distinct group within the complex and (2) that the species is a clade in the mtDNA gene tree or forms a cluster in the mtDNA haplotype network. We propose restricting *R. longirostris* to the nominate form plus the subspecies *phelpsi* Wetmore, 1941, *margaritae* Zimmer and Phelps, 1944, *pelodramus* Oberholser, 1937, *cypereti* Taczanowski, 1877, and *crassirostris* Lawrence, 1871. Found in South America and adjacent islands, these birds are relatively small, dull-breasted, robust-billed, and restricted to mangroves so could be called the "Mangrove Rail" (Fig. 4; Eddleman and Conway 1998). We retain as species the South American taxa *R. antarcticus, R. wetmorei*, and *R. semiplumbeus*, none of which were sampled in this study. The second species would be *R. tenuirostris* Ridgway, 1874 the "Mexican Rail," inhabiting the highland freshwater marshes of Mexico. These birds are large, very bright rufous ventrally, and have diffuse flank banding (Fig. 4; Meanley 1992). The third species would be *R. obsoletus*, including the critically endangered populations that occur along the Pacific coast of North America, for which we suggest the English name "Ridgway's Rail." This species includes the subspecies *levipes* Bangs, 1899, *beldingi* Ridgway, 1882, *yumanensis* Dickey, 1923, *rhizophorae* Dickey, 1930, and *nayaritensis* McLellan, 1927. This group is characterized by its relatively small body size (although larger than South American birds), bright rufous breast, and occurrence in saltmarshes (Fig. 4; Eddleman and Conway 1998). All populations within this proposed species have very bright rufous breasts. If our taxonomic recommendations are adopted, members of the *R. l. obsoletus* group would be elevated from the status of endangered subspecies to endangered species in the United States. The fourth species we propose is the King Rail, *R. elegans*, comprising two subspecies, *R. e. elegans* and *R. e. ramsdeni*. These birds breed in freshwater marshes of eastern North America and Cuba, are rufous-breasted, and relatively large (Fig. 4; Meanley 1992). The fifth species is the Clapper Rail, *R. crepitans*, comprising the eastern North American group of subspecies *R. l. crepitans* Gmelin, 1789, *waynei* Brewster, 1899, *scotti* Sennett, 1888, *insularum* Brooks, 1920, and *saturatus* Ridgway, 1880, as well as the birds of the Caribbean and Yucatan, including *R. l. caribaeus* Ridgway, 1880, *pallidus* Nelson, 1905, *grossi* Paynter, 1950, *belizensis* Oberholser, 1937, *leucophaeus* Todd, 1913, and *coryi* Maynard, 1887. These birds are intermediate in size, and the breast spans a range of colors from very dull, silvery gray, to dull rufous (Fig. 4). They breed in saltmarshes and saltmeadows of *Spartina* spp. along the Atlantic and Gulf of Mexico coasts of North America, as well as in mangroves in the Yucatan Peninsula, extreme southern United States, and Caribbean (Eddleman and Conway 1998). Under these recommendations the linear classification in the South American checklist would be unchanged. For the North American checklist the linear classification of species of the *Rallus longirostris/elegans* complex would be

R. obsoletus
R. tenuirostris
R. elegans
R. crepitans

Ecological speciation between populations breeding in saltwater and freshwater habitats may be a recurrent theme in the history of the diversification of *Rallus*. In Louisiana, introgression between *R. elegans* and *R. crepitans* is limited despite extensive hybridization in brackish marshes (Maley 2012). Elsewhere along the Atlantic coast of the United States and in Cuba the situation may be similar. Although

FIGURE 4. Specimens of representatives of each major group of subspecies ventrally and in profile. The subspecies, from left to right, are *R. e. tenuirostris* (LSUMZ 39438; *R. elegans* group), *R. l. yumanensis* (LSUMZ 39435; *R. l. obsoletus* group), *R. l. cypereti* (B-67819; *R. l. longirostris* group), *R. l. coryi* (LSUMZ 141611 *R. l. caribaeus* group), *R. l. saturatus* (B-63400; *R. l. crepitans* group), and *R. e. elegans* (B-63513; *R. elegans* group). All specimens are males in adult plumage. Photographs courtesy of R. E. Gibbons.

reproductive isolation has not been assessed directly, the closest relative of *R. tenuirostris* from the highlands of Mexico is *R. obsoletus*, which breeds primarily in saltmarshes. Likewise, *R. semiplumbeus* of the eastern Andes of Colombia may be closely related to coastal populations of *R. longirostris*, a hypothesis that needs to be assessed genetically. Collectively, these results and patterns suggest that inland *R. obsoletus yumanensis* may be in an early stage of ecological speciation. This merits further study and conservation attention.

Because of extensive habitat loss, *R. elegans* is of federal conservation concern and is treated as endangered in several states (Cooper 2008). We found little evidence of genetic differentiation between migratory and nonmigratory populations in the eastern United States. Two migratory birds breeding in Oklahoma did have an mtDNA haplotype (Fig. 2) which was found otherwise in only one bird from Louisiana, suggesting some structure within *R. e. elegans* in the U. S. may exist. This warrants further study.

ACKNOWLEDGMENTS

We thank the following institutions, curators, collections managers, staff, and collectors for providing valuable tissue specimens critical to this study: Academy of Natural Sciences of Philadelphia (N. Rice), Field Museum of Natural History (J. Bates, S. Hackett, D. Willard), Marjorie Barrick Museum of Natural History (J. Klicka), Natural History Museum of Los Angeles County (K. Garrett), San Diego Natural History Museum (K. Burns, P. Unitt), University of Arkansas (D. Krementz), University of Kansas Natural History Museum (R. Moyle, T. Peterson, M. Robbins), University of Michigan Museum of Zoology (J. Hinshaw), University of Washington, Burke Museum of Natural History and Culture (S. Birks). J. Saucier provided invaluable field assistance in collecting specimens for this study. We thank E. Derryberry, B. Smith, J. V. Remsen, M. Hellberg, F. Sheldon, F. Galvez, R. Shaw, and the Brumfield lab for discussion. For computational support we also thank S. Houston and the University of Alaska, Fairbanks Life Science Informatics Portal, which is supported by grant RR016466 from the National Center for Research Resources, a component of the National Institutes of Health. Collecting permits in Louisiana were granted by United States Fish and Wildlife Service and the Louisiana Department of Wildlife and Fisheries. We thank J. Huner, L. Richard Jr., and P. Smith Jr., for access to private land for collecting. This research was supported by grants DEB-0841729, DEB-0956909, and DEB-1110624 from the National Science Foundation.

LITERATURE CITED

Avise, J. C., and R. M. Zink. 1988. Molecular genetic divergence between avian sibling species: King and Clapper rails, Long-billed and Short-billed dowitchers, Boat-tailed and Great-tailed grackles, and Tufted and Black-crested titmice. Auk 105: 516–528.

Backstrom, N., M. Brandstrom, L. Gustafsson, A. Qvarnstrom, H. Cheng, and H. Ellegren. 2006. Genetic mapping in a natural population of Collared Flycatchers (*Ficedula albicollis*): conserved synteny but gene order rearrangements on the avian Z-chromosome. Genetics 174:377–386.

Bandelt, H.-J., P. Forster, and A. Röhl. 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular Biology and Evolution 16:37–48.

Banks, R. C., and R. E. Tomlinson. 1974. Taxonomic status of certain Clapper Rails of southwestern United States and northwestern Mexico. Wilson Bulletin 86:325–335.

Borge, T., M. T. Webster, G. Andersson, and G. P. Sætre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes in two *Ficedula* flycatcher species. Genetics 171:1861–1873.

Cooper, T. R. 2008. King Rail conservation plan, version 1. U. S. Fish aand Widlife Service, Fort Snelling, MN.

Drummond, A. J., and A. Rambaut. 2007. Beast: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

EDDLEMAN, W. R., AND C. J. CONWAY. 1998. Clapper Rail (*Rallus longirostris*), no. 340. *In* A. Poole and F. Gill [EDS.], The Birds of North America, Inc., Philadelphia.

FJELDSÅ, J., D. ZUCCON, M. IRESTEDT, U. S. JOHANSSON, AND P. G. P. ERICSON. 2003. *Sapayoa aenigma*: a New World representative of 'Old World suboscines.' Proceedings of the Royal Society B 270:S238–S241.

FLOT, J.-F. 2010. SeqPhase: a web tool for interconverting phase input/output files and fasta sequence alignments. Molecular Ecology Resources 10:162–166.

GOODMAN, S. M., M. J. RAHERILALAO, AND N. L. BLOCK. 2011. Patterns of morphological and genetic variation in the *Mentocrex kioloides* complex (Aves: Gruiformes: Rallidae) from Madagascar, with the description of a new species. Zootaxa 2776: 49–60.

HACKETT, S. J. 1996. Molecular phylogenetics and biogeography of tanagers in the genus *Ramphocelus* (Aves). Molecular Phylogenetics and Evolution 33:368–382.

HELED, J., AND A. J. DRUMMOND. 2010. Bayesian inference of species trees from multilocus data. Molecular Biology and Evolution 27:570–580.

HELLMAYR, C. E., AND H. B. CONOVER. 1942. Catalogue of birds of the Americas, part 1, no. 1. Field Museum of Natural History Zoological Series 13.

HIRD, S. M., R. T. BRUMFIELD, AND B. C. CARSTENS. 2011. PRGmatic: an efficient pipeline for collating genome-enriched second generation sequencing data using a "provisional-reference genome." Molecular Ecology Resources 11:743–748.

INTERNATIONAL CHICKEN GENOME SEQUENCING CONSORTIUM. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695–716.

JOHNSON, K. P., AND M. D. SORENSON. 1998. Comparing molecular evolution in two mitochondrial protein coding genes (cytochrome b and ND2) in the dabbling ducks (tribe Anatini). Molecular Phylogenetics and Evolution 10:82–94.

KIMBALL, R. T., E. L. BRAUN, F. K. BARKER, R. C. K. BOWIE, M. J. BRAUN, J. L. CHOJNOWSKI, S. J. HACKETT, K. L. HAN, J. HARSHMAN, V. HEIMER-TORRES, W. HOLZNAGEL, C. J. HUDDLESTON, B. D. MARKS, K. J. MIGLIA, W. S. MOORE, S. REDDY, F. H. SHELDON, J. V. SMITH, C. C. WITT, AND T. YURI. 2009. A well-tested set of primers to amplify regions spread across the avian genome. Molecular Phylogenetics and Evolution 50:654–660.

LIBRADO, P., AND J. ROZAS. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25:1451–1452.

MALEY, J. M. 2012. Ecological speciation of King Rails (*Rallus elegans*) and Clapper Rails (*Rallus longirostris*). Ph.D. dissertation, Louisiana State University, Baton Rouge, LA.

MCCORMACK, J. E., J. M. MALEY, S. HIRD, E. P. DERRYBERRY, G. R. GRAVES, AND R. T. BRUMFIELD. 2012. Next-generation sequencing reveals phylogeographic structure and a species tree for recent bird divergences. Molecular Phylogenetics and Evolution 62:397–406.

MCCORMACK, J. E., S. M. HIRD, A. J. ZELLMER, B. C. CARSTENS, AND R. T. BRUMFIELD. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Molecular Phylogenetics and Evolution 66:526–538.

MEANLEY, B. 1992. King Rail (*Rallus elegans*), no 3. *In* A. Poole, P. Stettenheim, and F. Gill [EDS.], The Birds of North America. Academy of Natural Sciences, Philadelphia.

MEANLEY, B., AND D. K. WETHERBEE. 1962. Ecological notes on mixed populations of King Rails and Clapper Rails in Delaware Bay marshes. Auk 79:453–457.

NYLANDER, J. A. A. 2004. MrModeltest, version 2. Program distributed by the author. Evolutionary Biology Centre, Uppsala, Sweden.

OLSON, S. L. 1997. Towards a less imperfect understanding of the systematics and biogeography of the Clapper and King rail complex (*Rallus longirostris* and *R. elegans*), p. 93–111. *In* R. W. Dickerman [ED.], The era of Allan R. Phillips: a festschrift. Horizon Communications, Albuquerque, NM.

PRYCHITKO, T. M., AND W. S. MOORE. 1997. The utility of DNA sequences of an intron from the β-fibrinogen gene in phylogenetic analysis of woodpeckers (Aves: Picidae). Molecular Phylogenetics and Evolution 8:193–204.

RAMBAUT, A., AND A. J. DRUMMOND. 2007. Tracer, version 1.5.

RIDGELY, R. S., T. F. ALLNUTT, T. BROOKS, D. K. MCNICOL, D. W. MEHLMAN, B. E. YOUNG, AND J. R. ZOOK. 2005. Digital distribution maps of the birds of the Western Hemisphere, version 2.1.

RIPLEY, S. D. 1977. Rails of the world: a monograph of the family Rallidae. M. F. Feheley, Toronto.

ROZEN, S., AND H. J. SKALETSKY. 2000. Primer3 on the WWW for general users and for biologist programmers, p. 365–386. *In* S. Krawetz and S. Misener [EDS.], Bioinformatics methods and protocols: methods in molecular biology. Humana Press, Totowa, NJ.

SÆTRE, G. P., T. BORGE, K. LINDROOS, J. HAAVIE, B. C. SHELDON, C. PRIMMER, AND A. C. SYVANEN. 2003. Sex chromosome evolution and speciation in *Ficedula* flycatchers. Proceedings of the Royal Society of London B 270:53–59.

SLADE, R. W., C. MORITZ, A. HEIDEMAN, AND P. T. HALE. 1993. Rapid assessment of single-copy nuclear DNA variation in diverse species. Molecular Ecology 2:359–373.

STEADMAN, D. W. 1995. Prehistoric extinctions of Pacific island birds: biodiversity meets zooarchaeology. Science 267: 1123–1131.

STEPHENS, M., AND P. DONNELLY. 2003. A comparison of Bayesian methods for haplotype reconstruction from population genotype data. American Journal of Human Genetics 73: 1162–1169.

STEPHENS, M., AND P. SCHEET. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. American Journal of Human Genetics 76:449–462.

STEPHENS, M., N. J. SMITH, AND P. DONNELLY. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics 68:978–989.

SWOFFORD, D. L. 1998. PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4d64. Sinauer, Sunderland, MA.

TAVARES, E. S., G. H. J. DE KROON, AND A. J. BAKER. 2010. Phylogenetic and coalescent analysis of three loci suggest that the Water Rail is divisible into two species, *Rallus aquaticus* and *R. indicus*. BMC Evolutionary Biology 10:226.

TAYLOR, P. B. 1996. Family Rallidae (rails, gallinules and coots), p. 108–209. *In* J. del Hoyo, A. Elliot and J. Sargatal [EDS.], Handbook of the birds of the world. Lynx Edicions, Barcelona.

TOMLINSON, R. E., AND R. L. TODD. 1973. Distribution of two western Clapper Rail races as determined by responses to taped calls. Condor 75:177–183.

VOS, P., R. HOGERS, M. BLEEKER, M. REIJANS, T. VAN DE LEE, M. HORNES, A. FRIJTERS, J. POT, J. PELEMAN, AND M. KUIPER. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Research 11:4407–4414.

WARREN, W. C., D. F. CLAYTON, H. ELLEGREN, A. P. ARNOLD, L. W. HILLIER, A. KUNSTNER, S. SEARLE, S. WHITE, A. J. VILELLA, S. FAIRLEY, A. HEGER, L. KONG, C. P. PONTING, E. D. JARVIS, C. V. MELLO, P. MINX, P. LOVELL, T. A. F. VELHO, M. FERRIS, C. N. BALAKRISHNAN, S. SINHA, C. BLATTI, S. E. LONDON, Y. LI, Y.-C. LIN, J. GEORGE, J. SWEEDLER, B. SOUTHEY,

P. Gunaratne, M. Watson, K. Nam, N. Backstrom, L. Smeds, B. Nabholz, Y. Itoh, O. Whitney, A. R. Pfenning, J. Howard, M. Volker, B. M. Skinner, D. K. Griffin, L. Ye, W. M. McLaren, P. Flicek, V. Quesada, G. Velasco, C. Lopez-Otin, X. S. Puente, T. Olender, D. Lancet, A. F. A. Smit, R. Hubley, M. K. Konkel, J. A. Walker, M. A. Batzer, W. Gu, D. D. Pollock, L. Chen, Z. Cheng, E. E. Eichler, J. Stapley, J. Slate, R. Ekblom, T. Birkhead, T. Burke, D. Burt, C. Scharff, I. Adam, H. Richard, M. Sultan, A. Soldatov, H. Lehrach, S. V. Edwards, S.-P. Yang, X. Li, T. Graves, L. Fulton, J. Nelson, A. Chinwalla, S. Hou, E. R. Mardis, and R. K. Wilson. 2010. The genome of a songbird. Nature 464: 757–762.

Weissbach, L., C. Oddoux, R. Procyk, and G. Grieninger. 1991. The β chain of chicken fibrinogen contains an atypical thrombin cleavage site. Biochemistry 30:3290–3294.

Zwickl, D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, University of Texas, Austin, TX.