

## Congruence between nuclear and mitochondrial DNA: Combination of multiple nuclear introns resolves a well-supported phylogeny of New World orioles (*Icterus*)

Frode Jacobsen \*, Nicholas R. Friedman, Kevin E. Omland

Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, MD 21250, USA

### ARTICLE INFO

#### Article history:

Received 20 November 2009

Revised 25 March 2010

Accepted 29 March 2010

Available online 2 April 2010

#### Keywords:

Phylogenetics

Concatenation

Cytosuclear congruence

Z chromosome

Introns

*Icterus*

### ABSTRACT

Darwin's vision of a "Tree of Life" showing evolutionary relationships among all extant species seems an increasingly feasible goal, at least for vertebrate animals. However, virtually all published molecular phylogenies for closely related animals are based on a single locus – maternally inherited mitochondrial DNA. New approaches using multiple nuclear loci are needed to test published trees and better resolve the twigs of the entire tree of life. Here we use New World orioles (*Icterus*) to test an approach based on combined analysis of six independent Z chromosome introns. Combined analysis of multiple introns using traditional phylogenetic methods resolved a well-supported species phylogeny of New World orioles. In fact, all major lineages of orioles and several sub-clades that are well-supported by previously published mtDNA data are also strongly supported by the combined nuclear Z-intron tree. The male-biased Z-intron tree presented here is overwhelmingly congruent with the female-exclusive mtDNA tree. A slow rate of mutation relative to mtDNA resulted in generally poorly resolved gene trees when intron loci were analyzed separately. However, strong phylogenetic signal for all but the most recent divergences emerged once multiple loci were concatenated and analyzed in combination. Although there clearly are conditions under which concatenation analysis of nuclear DNA can be misleading, the congruence between mitochondrial and nuclear estimates of the *Icterus* phylogeny suggests that concatenation remains a powerful tool for inferring phylogenetic relationships for all but very recent divergences.

© 2010 Elsevier Inc. All rights reserved.

### 1. Introduction

Mitochondrial DNA has long been the locus of preference for determining phylogenetic relationships between closely related animal species. The great utility of mitochondrial DNA (mtDNA) comes from its rapid mutation rate and short coalescence time relative to autosomal nuclear DNA, making it particularly well suited for addressing lower-level taxonomic questions. However, inferring evolutionary relationships solely based on evidence from one gene tree is inadequate, as it provides only one independent estimate of the species tree (Moore, 1995). Despite its favorable attributes, mtDNA is not exempt from confounding genealogical processes such as incomplete lineage sorting of retained ancestral polymorphisms, introgression, and different maternal lineages (Funk and Omland, 2003; Ballard and Rand, 2005). A general lack of recombination (Rokas et al., 2003a; Berlin et al., 2004; but see Tsaousis et al., 2005) also means that sampling more mitochondrial loci will not help overcome these challenges.

Nuclear genes are currently utilized extensively to test mtDNA derived phylogenies of a wide range of organisms (e.g., Sota and Sasabe, 2006; Benavides et al., 2007; Hackett et al., 2008). A single nuclear gene tree might provide a reliable estimate of higher-level phylogenies, where population sizes typically have been small relative to the branch lengths and tend to generate gene trees that overall are congruent with mtDNA trees at deeper phylogenetic levels. However, in order to accurately infer recent nodes in the "Tree of Life" it is necessary to combine data from multiple independent loci such as nuclear introns, non-coding portions of protein-coding genes generally assumed to be selectively neutral (Irimia and Roy, 2008).

However, it remains uncertain how well nuclear introns work at the species-level (but see Allen and Omland, 2003; Peters et al., 2005; Barker et al., 2008; Carling and Brumfield, 2008b). Their much slower mutation rate relative to mtDNA limits the amount of information available for phylogenetic inference and the level of resolution that can be obtained along short internodes and terminal branches. Moreover, their generally much larger effective population size ( $N_e$ ) relative to mtDNA extends the coalescence process (sometimes even by several million years) before reaching reciprocal monophyly (Hudson and Turelli, 2003). Random retention and extinction of ancestral alleles and incomplete lineage sorting (deep coalescence) causes large variance in the coalescent

\* Corresponding author. Address: Department of Biological Sciences, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA. Fax: +1 410 455 3875.

E-mail address: [frode1@umbc.edu](mailto:frode1@umbc.edu) (F. Jacobsen).

process that could lead to incongruent gene trees and incongruence between gene trees and the underlying species tree at early stages of divergence (Pamilo and Nei, 1988; Takahata, 1989; Carstens and Knowles, 2007; Knowles, 2009).

So how can we infer phylogenies of recently diverged taxa given the challenges of nuclear introns? Other rapidly mutating nuclear markers such as microsatellites and AFLPs have provided robust phylogenies of closely related taxa (e.g., Petren et al., 1999; Sullivan et al., 2004; Mendelson and Shaw, 2005; Mendelson and Simons, 2006), but are generally less favored for phylogenetic inference due to more questionable homology and homoplasy compared to sequence data (Jarne and Lagoda, 1996; Meudt and Clarke, 2007). Furthermore, these hyper variable markers suffer from the same slow rate of coalescence experienced by any autosomal nuclear locus evolving by neutral genetic drift (Zink and Barrowclough, 2008).

Concatenation has long been the leading method for inference of multilocus data sets (e.g., Delsuc et al., 2003; Rokas et al., 2003b; Stepan et al., 2005; Pagès et al., 2008; Wiens et al., 2008; Parra et al., 2009). However, recent studies have shown that for simulated data, the most probable tree is sometimes not identical to the true species tree and that concatenation of highly incongruent genealogies can produce inconsistent results (Degnan and Rosenberg, 2006; Kubatko and Degnan, 2007; Liu and Edwards, 2009).

What is truly needed, regardless of method of inference, are faster sorting nuclear loci that may provide resolution near the tips of a species tree. For animal taxa with chromosomal sex determination, loci linked to the sex chromosomes may offer just that (Ezaz et al., 2006). Like several other animal taxa (e.g., Mank et al., 2006; Traut et al., 2007; Nakamura, 2009), birds display reversed heterogamety wherein females are the heterogametic sex (ZW) and males are the homogametic sex (ZZ) (Fridolfsson et al., 1998). Loci linked to the avian Z chromosome are therefore expected to have a slightly smaller ( $\frac{3}{4}$ )  $N_e$  than autosomal loci and are thus also likely to sort slightly faster. Furthermore, Z-linked loci are expected to recombine less frequently than autosomal loci since females carry a single copy of the Z chromosome. While maternally inherited W-linked genes should trace the history of mitochondrial genes, Z-linked genes that spend  $\frac{2}{3}$  of their evolutionary time in the male genome may offer a more robust way to test mtDNA derived phylogenies. In addition, thanks to the fully sequenced chicken (*Gallus gallus*) genome, partially assembled zebra finch (*Taeniopygia guttata*) genome and the recently completed linkage mapping of the collared flycatcher (*Ficedula albicollis*) genome (Backström et al., 2006, 2008), we now know that there is extensive synteny and gene order conservation across deep avian lineages. This knowledge is invaluable to the study of non-model organisms, as we now have a huge number of conserved loci “known” to be linked to the Z chromosome to pick from. Several Z-linked loci have been utilized in comparative studies and population genetic studies (e.g., Sætre et al., 2002; Borge et al., 2005; Carling and Brumfield, 2008b), but the potential power of these loci for testing phylogenies has yet to be demonstrated. Their smaller  $N_e$  and thus expected faster sorting rate may render them much better suited at resolving species-level relationships of organisms with similar sex determination.

This study seeks to test the robustness of a well-resolved mtDNA phylogeny of New World orioles (Omland et al., 1999; Sturge et al., 2009) using multiple Z-linked introns. The genus *Icterus* constitutes a monophyletic group within the New World blackbirds (family Icteridae: Lanyon and Omland, 1999), and consists of roughly 30 species with more than 60 sub-species described to date (Clements, 2007). The genus contains three well-supported lineages of orioles, labeled clades A, B, and C by Omland et al. (1999). Our lab has relied on the mitochondrial phylogeny of *Icterus* for many years, as the New World orioles have been the focus of extensive research on reconstructing the evolution of

behaviors such as migration and song (Kondo, 2006; Price et al., 2007) as well as plumage traits such as pigmentation and sexual dichromatism (Hofmann et al., 2006, 2007a, 2007b, 2008a,b; Friedman et al., 2009). Due to the potential pitfalls of over-reliance on mtDNA evidence, it has been a long-standing goal to rigorously test the phylogenetic relationships within *Icterus* using a multilocus approach. Previous intron work on *Icterus* demonstrated that nuclear introns can provide some support for deeper relationships (clades A, B, and C) within the genus (Allen and Omland, 2003), but it remains to be demonstrated to what degree the challenges presented by slow mutation, (incomplete) lineage sorting and potential introgressive hybridization can be overcome. Here, we show that traditional combined analysis of multiple sex-linked intron loci can resolve all but the most recent divergences and produce a robust, concordant estimate of the *Icterus* phylogeny.

## 2. Materials and methods

### 2.1. Taxon sampling

The study included a total of 32 individuals. Single representatives were selected for each of 29/30 recognized species of *Icterus* orioles (Clements, 2007). The only missing species, the monotypic South American Campo troupial (*Icterus jamacaii*), was excluded due to a lack of fresh voucher tissue. For the same reason, this taxon was not included in the previously published mitochondrial phylogeny (Omland et al., 1999). To root the nuclear *Icterus* trees, the study also included single representatives of genera from the other major clades within the family Icteridae (*Agelaius*, *Caciccus*, and *Sturnella*) (Lanyon and Omland, 1999). Voucher sample details are listed in Appendix A.

### 2.2. Data collection

Genomic DNA was extracted from muscle tissue using Qiagen DNeasy extraction kits (Qiagen). The six Z-linked loci included in the study are ADAMTS6 metalloproteinase with thrombospondin type 1, motif 6 intron 5 (*ADAM-5*), aldolase-B fructose-biphosphate intron 5 (*ALDO-5*), Brahma protein intron 15 (*BRM-15*), chromohelicase-DNA binding protein intron 18 (also known as locus *CHD1Zb*; *CHD-18*), muscle skeleton receptor tyrosine kinase intron 4 (*MUSK-4*), and solute carrier family 30 (zinc transporter), member 5 intron 9 (*SLC-9*). With the exception of *ALDO-5*, all six targeted intronic regions were amplified using exon-primed intron-crossing (EPIC) primer combinations available in the literature (Table 1). Amplification of *ALDO-5* was achieved using modified primers designed specifically for *Icterus* (Table 1). In order to enhance the quality of the sequence reads, all locus-specific primers were end-tagged with flanking oligos corresponding to universal M13 sequencing primers. All 15  $\mu$ L PCR mixes included 10 $\times$  Taq buffer w/KCl, 0.25 mM MgCl<sub>2</sub>, 0.26  $\mu$ M of each primer, 0.2 mM of each dNTP, and 0.75 U Taq DNA polymerase. All loci were amplified on a PTC200 (MJ Research) thermal cycler under the following profile: initial denaturing at 95 °C for 5 min, followed by 30 cycles of 95 °C for 30 s, primer-specific annealing temperature ( $T_a$ ; see Table 1) for 30 s, 72 °C for 45 s, followed by a final extension at 72 °C for 10 min. Three microliters of PCR products was electrophoresed in 1% agarose TAE gels to confirm amplification success and the absence of contamination (i.e., no amplification of negative control). PCR products were purified by digesting unincorporated nucleotides and primers using a diluted (1:9) mixture of exonuclease and shrimp alkaline phosphatase (ExoSAP-IT; USB). Five microliters of ExoSAP-IT mixture was added to each remaining 12  $\mu$ L of PCR product and incubated for 30 min at 37 °C, followed by 15 min at 80 °C. Purified reactions for all loci were then sequenced

**Table 1**

Loci included in study, primer sequences and sources, locus-specific annealing temperatures ( $T_a$ ), locus length, best-fitted substitution models, and number of parsimony-informative sites (PI) for ingroup taxa.

Gene region	Primer sequence (5' → 3')	Primer source	$T_a$ (°C)	Length (bp)	Substitution model	PI
ADAM-5	F: GGAGAGAATGGATTTCGCC R: TGATTCCAGTCTAGGAAACG	Backström et al. (2006)	55	596–599	HKY + I	10
ALDO-5	F: ATCCAGGAGAATGCCAACAC R: CTCCACAATGGGTACCAAGC	Friesen et al. (1999) and Kondo et al. (2008)	60	249–258	K80	6
BRM-15	F: AGCACCTTGAACAGTGGTT R: TACTTTATGGAGACGACGGA	Borge et al. (2005)	51	357–364	HKY	10
CHD-18	F: ACATACAGGCTCTACTCCT R: CCCCTTCAGGTCTTTAAAA	Ellegren and Fridolfsson (1997)	55	320–330	GTR + $\Gamma$	11
MUSK-3	F: CTTCCATGCACTACAATGGGAAA R: CTCTGAACATTGTGGATCTCAA	Designed by F. K. Barker (see Clark and Witt, 2006)	50	556–565	GTR + $\Gamma$	16
SLC-9	F: TCTGGAGGAGGGGTAGTGAG R: AGGAGAATAGCCAATAAGGG	(Backström et al., 2006)	56	406–416	HKY + I + $\Gamma$	13
zDNA combined				2540	GTR + $\Gamma$	66
mtDNA combined				2005	GTR + I + $\Gamma$	475

using the universal primers M13F-29/M13R and BigDye 3.1 cycle sequencing chemistry (Applied Biosystems) following the manufacturer's recommended protocol and cycling conditions. Cycle sequencing products were purified using the manufacturers' recommended ethanol + EDTA precipitation protocol and analyzed on an Applied Biosystems 3100 automated capillary sequencer. Resulting DNA chromatograms and base calls were checked and edited in Sequencher 4.2 (Gene Codes). Heterozygous sites confirmed by the presence of double peaks in both complementary DNA strands were coded as polymorphisms using the appropriate IUPAC standard ambiguity codes, gaps were coded as (-), and missing data were coded as (?). All sequences are deposited in GenBank (Accession Nos. GU972812–GU973001).

### 2.3. Data analysis

Edited sequence contigs for all six loci were aligned in ClustalX (Thompson et al., 1997) using gap opening and gap extension penalties of 10 and 5, respectively. The resulting near-optimal alignments were subsequently refined by eye. An ambiguously aligned region (17 bp) of single nucleotide repeats was excluded from locus *SLC-9* prior to analysis.

The intron partitions were analyzed both separately and combined (by concatenation) using maximum parsimony (MP), maximum likelihood (ML), and Bayesian approaches (BA). Alignment gaps were initially treated as missing data. To provide a better comparison of phylogenetic estimates, a subset of the original mtDNA data set from Omland et al. (1999) and recently added sequence data (Sturge et al., 2009) corresponding to the focal 32 taxa was re-analyzed using MP, ML, and BA. Equally-weighted maximum parsimony heuristic searches were conducted in PAUP\* (Swofford, 2002), with ten random addition sequence replicates and tree bisection–reconnection (TBR) branch-swapping. Separate MP analyses were performed with the exclusion and inclusion of indel characters. Potentially informative indel (insertion–deletion) characters identified by alignment gaps were coded as binary characters according to the method of modified complex indel coding (Müller, 2006) implemented in the software SeqState 1.4 (Müller, 2005). Nodal support for MP trees was assessed via heuristic bootstrap searches (Felsenstein, 1985) with ten random addition sequence replicates per 1000 bootstrap replicates. Due to indeterminate search lengths for some of the intron partitions, we restricted each search to a maximum of 1,000,000 rearrangements per replicate for the separate analyses.

Maximum likelihood searches were conducted in GARLI 0.96b7 (Zwickl, 2006) by implementing the best fitting substitution mod-

els identified for individual intron partitions, the combined nuclear data set, and the mtDNA data set (Omland et al., 1999; Sturge et al., 2009) using MrModelTest (Nylander et al., 2004) under the Akaike information criterion (see Table 1). Two independent searches were performed for each intron partition and the combined data set using different random starting seeds. Alignment gaps (indels) were treated as missing data for these analyses. As for MP analyses, nodal support was evaluated using non-parametric bootstrap searches with 1000 pseudo-replicates. ML majority-rule bootstrap consensus trees were assembled by importing the GARLI tree files into PAUP\*. MP and ML bootstrap proportions of  $\geq 70\%$  were considered to provide strong nodal support.

Bayesian tree searches were conducted in MrBayes 3.1 (Huelsenbeck and Ronquist, 2001; Ronquist and Huelsenbeck, 2003). Here, individual intron partitions were allowed to evolve independently by specifying separate substitution models (as determined by MrModeltest: see Table 1) for all partitions and unlinking the parameter estimates across partitions. Since MrBayes only allows designation of a single outgroup taxon in its current implementation, *Sturnella* was selected as outgroup for the Bayesian analysis. The mtDNA data was treated as a single partition in both ML and Bayesian analyses. The Markov chain Monte Carlo (MCMC) algorithm was run for  $1 \times 10^8$  generations for both Z-intron DNA and mtDNA data sets, sampling trees every  $1 \times 10^4$  generations, using four Markov chains with default heating values. Stationarity of the MCMC and the posterior probability estimates obtained from two independent runs from random starting points was assessed by examination of trace plots of log likelihoods ( $\ln L$ ) using TRACER 1.3 (Rambaut and Drummond, 2004). Trees sampled before reaching stationarity were discarded as burnin, and a majority-rule consensus tree (excluding burnin trees) was assembled using MrBayes (Ronquist and Huelsenbeck, 2003). Nodes receiving a posterior probability (PP)  $\geq 0.95$  by Bayesian inference were considered statistically robust.

### 2.4. Tests of congruence

To assess the level of phylogenetic concordance between the nuclear (Z-intron DNA) and mitochondrial (mtDNA) data sets we employed a suite of incongruence tests. First, a partition homogeneity test (ILD tests: Farris et al., 1994) was conducted in PAUP\* 4.0b10 (Swofford, 2002) to test for phylogenetic conflict between the mtDNA and Z-intron DNA data sets. Only ingroup taxa and parsimony-informative characters were considered for this test. Second, comparisons of ML trees as inferred by GARLI were performed in PAML (Yang, 2007) using the Kishino–Hasegawa

(KH) test (Kishino and Hasegawa, 1989). Both the ILD test and the KH test have limitations and problems attached to them (Planet, 2006). Yet, they provide reasonable metrics of the degree of incongruence in the underlying data. Since the objective was to test the congruence of *Icterus* relationships as inferred by independent mitochondrial and nuclear data the outgroup taxa were excluded from these tests. For completeness, we reconstructed a “total evidence” tree (sensu de Queiroz et al., 1995) based on combined analysis of mtDNA and Z-intron DNA data using the same methods of inference as described above. This not only allowed assessment of the effects potentially contradicting mtDNA and combined Z-intron DNA signals might have a on total evidence analysis, but also the relative fit of the total evidence tree to the two underlying independent data sets. Finally, we estimated branch lengths on this tree using PAUP\* based on both mtDNA and Z-intron DNA data sets. A Spearman's ranks test was used to determine whether the branch lengths calculated based on mtDNA correlated with branch lengths calculated based on independent Z-intron DNA.

### 3. Results

#### 3.1. Intron characters

Amplification of six Z-linked intron loci for the 32 taxa included in the study yielded a total of 2540 nucleotide characters, of which 80 were parsimony-informative (PI) overall (66 PI for ingroup taxa; Table 1). A total of 24 alignment gaps (indels) were present in the data set. Of these, 22 indel characters were coded and added to the data matrix (five PI for ingroup; Table 1). For comparison, the existing mtDNA data set (2005 bp) on the same 29 ingroup taxa contained 475 PI characters (see Table 1). Only three of 192 total amplifications failed; *Icterus wagleri* at ADAM-6, *Icterus auratus* at CHD-18, and *Icterus prothemelas* at SLC-9. While *I. wagleri* and *I. auratus* were represented as missing data in these instances, a second individual collected at the same location was chosen to represent *I. prothemelas* at locus SLC-9 (see Appendix A). All females included in study were homozygous across all loci sequenced, supporting the assumption that these six loci indeed are Z-linked. See Table 1 for details regarding mtDNA data included in study.

#### 3.2. Nuclear gene tree inference

Individual nuclear gene trees were generally poorly resolved with only one or a few strongly supported nodes (receiving  $\geq 70\%$  bootstrap proportion (BP) and  $\geq 0.95$  Bayesian posterior probability (PP); see Table 2). The intron loci differed greatly in the degree of support for *Icterus* relationships, as a given gene tree

typically only supported some of the deeper relationships within the genus (i.e., clade A, B, or C with little or no resolution near the tips). Whereas no node in the mtDNA tree was supported by all nuclear gene trees, a majority of loci strongly supported the monophyly of *Icterus* (node 14, Table 2) and the sister relationship between *Icterus graduacauda* and *Icterus chrysater* (node 13, Table 2). Importantly, no strongly supported ingroup nodes in any nuclear gene tree were incongruent with other nuclear gene trees or in discord with the mtDNA tree. Conflicting nodes were supported by only a single intron locus and with weak or no bootstrap support or posterior probabilities ( $< 50\%$  BP and  $< 0.70$  PP; except SLC-9 supporting *Cacicus* as nested within clade C with 61% MP BP, 79% ML BP, and 1.0 Bayesian PP).

#### 3.3. Combined analysis of multiple introns

Combined analysis of all six Z-linked introns yielded a well-resolved nuclear phylogeny of *Icterus* (Figs. 1 and 2). The three major clades (A, B, and C) were all recovered with moderate to strong support (Figs. 1 and 2, Table 2). Furthermore, most sub-clades strongly supported by the mtDNA data (hereafter referred to as clades A<sub>1-2</sub>, B<sub>1-2</sub>, and C<sub>1-2</sub>, see Fig. 2) were also strongly supported by the combined nuclear data (Figs. 1 and 2 and Table 2). Overall, the three methods of inference used (MP, ML, and BA) provided similar levels of support for nodes in the combined Z-intron tree.

#### 3.4. Topological concordance between Z-intron DNA and mtDNA

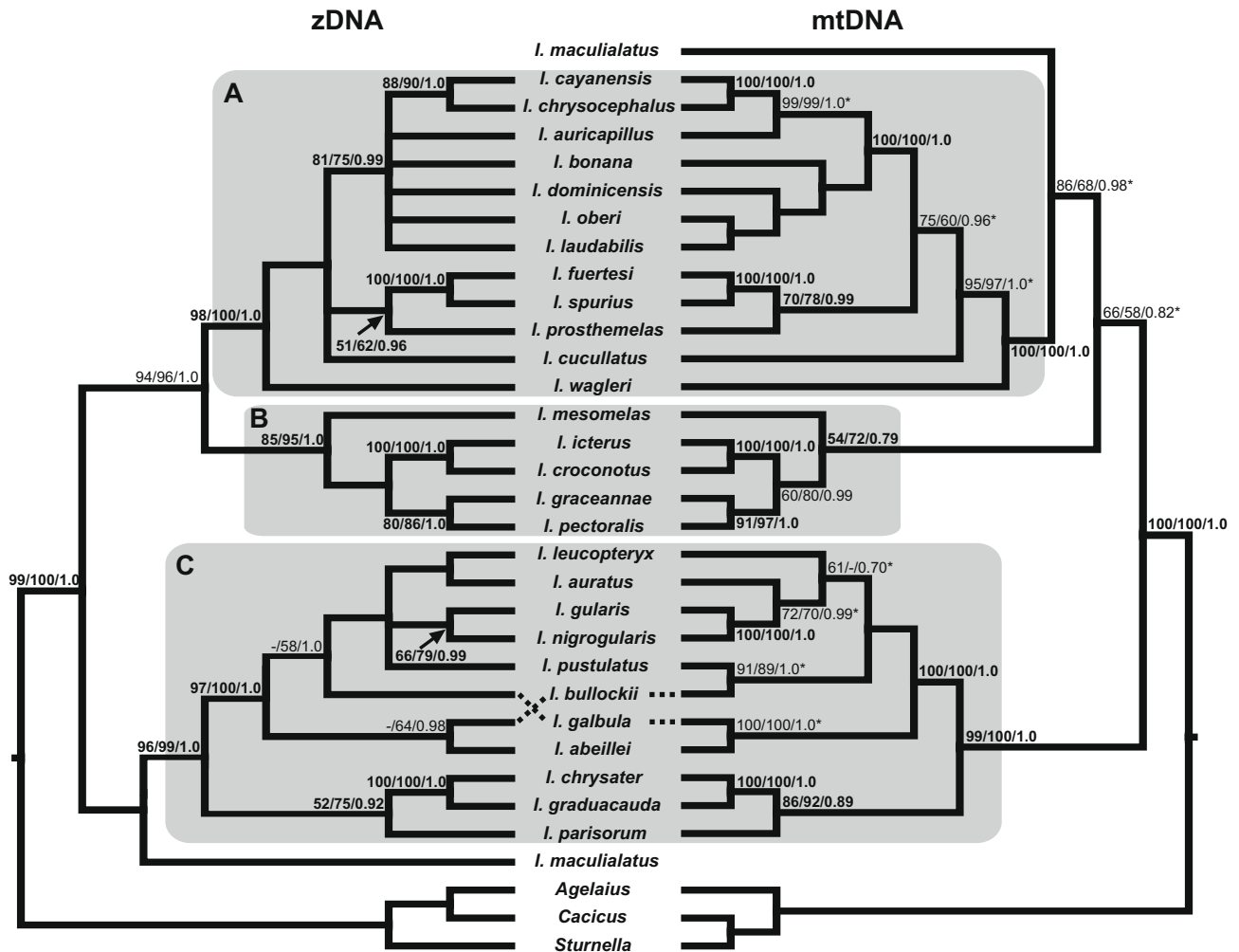
The well-supported nuclear DNA tree resolved by combined analysis of multiple sex-linked introns is remarkably congruent with the published mitochondrial tree of *Icterus* (Omland et al., 1999, also see Figs. 1 and 2 this paper). Of 14 nodes that are strongly supported by at least two of three methods used to analyze the concatenated Z-intron data, 12 (86%) are also strongly supported by mtDNA data (see Figs. 1, 2 and Table 2). Furthermore, branch lengths reconstructed onto the total evidence tree ( $n = 59$ ) based on nuclear and mitochondrial DNA were also highly correlated (Spearman's rank test;  $r_s = 0.69$ ,  $t_{57} = 7.18$ ,  $p_{two-tailed} < 0.000001$ , also see Fig. 2).

There are some interesting points of disagreement between the two data sets as well. One concerns the position of the deep and species-poor lineage leading to *I. maculialatus*, reconstructed as sister to clade A in the mtDNA tree and sister to clade C in the Z-intron tree. This conflict will be dealt with in greater detail below. Another source of conflict involves two sister relationships within sub-clade C<sub>1</sub> (see Fig. 2). The mtDNA tree strongly supports *Icterus abeillei* as sister to *Icterus galbula*, whereas the Z-intron tree places

**Table 2**  
Confidence indices based on MP and ML bootstrap proportions (BP), and Bayesian posterior probabilities (PP), respectively.

Locus/Node	zDNA	zDNA*	mtDNA	ADAM-5	ALDO-5	BRM-15	CHD-18	MUSK-3	SLC-9
1 (A)	98/100/1.0	99/NA/NA	95/97/1.0	–/54/0.74			90/90/1.0	87/94/1.0	
2 (A <sub>1</sub> )	81/75/0.99	82/NA/NA	100/100/1.0	95/95/1.0			61/60/0.56		
3	88/90/1.0	90/NA/NA	100/100/1.0					65/65/0.99	67/62/0.88
4 (A <sub>2</sub> )	51/62/0.96	61/NA/NA	70/78/0.99			63/63/0.92			
5	100/100/1.0	100/NA/NA	100/100/1.0		61/62/0.92		67/68/0.99	97/96/1.0	60/67/0.96
6 (B)	85/95/1.0	85/NA/NA	54/72/0.79			65/62/0.93	–/–/0.52		
7 (B <sub>1</sub> )	100/100/1.0	100/NA/NA	100/100/1.0	62/78/0.97		64/66/0.97	78/88/1.0		
8 (B <sub>2</sub> )	80/86/1.0	76/NA/NA	91/97/1.0		84/88/1.0	61/69/0.98			
9 (C)	96/99/1.0	94/NA/NA	99/100/1.0			–/52/0.92		–/60/0.85	
10 (C <sub>1</sub> )	97/100/1.0	99/NA/NA	100/100/1.0	64/64/0.70	–/62/0.88		–/88/0.72		61/65/0.82
11	66/79/0.99	67/NA/NA	100/100/1.0				–/50/–		59/65/0.72
12 (C <sub>2</sub> )	52/75/0.92		86/92/0.89				54/68/0.84		
13	100/100/1.0	100/NA/NA	100/100/1.0	62/71/0.97	54/65/0.92	97/98/1.0	91/92/1.0	84/92/1.0	–/0.52/0.77
14 ( <i>Icterus</i> )	99/100/1.0	100/NA/NA	100/100/1.0	80/77/0.96	63/80/0.90	96/98/1.0	56/83/0.99		

\* MP bootstrap proportions for Z-intron tree with the inclusion of 22 indel characters.



**Fig. 1.** Parsimony Z-intron DNA (left) and mtDNA (right) trees of 29 species of *Icterus* orioles (we randomly chose one of the 54 and one of the three most parsimonious trees, respectively). Trees are rooted by three outgroups from family Icteridae. Shaded areas outline three major clades of orioles (A, B, and C) within genus *Icterus*. Nodal confidence is indicated by MP and ML bootstrap proportions and Bayesian posterior probabilities. Support values in bold refer to nodes supported by both nuclear and mitochondrial data. Support values highlighted by asterisks indicate nodes that are only resolved by mtDNA. Note the conflicting arrangement of *I. bullockii* and *I. galbula* as indicated by the dotted lines. While mtDNA strongly supports a close sister relationship between *I. galbula* and *I. abeillei*, nuclear DNA instead suggests a close sister relationship between *I. bullockii* and *I. abeillei*.

*I. abeillei* sister to *Icterus bullockii* with moderate to strong support (see Figs. 1, 2 and Table 2).

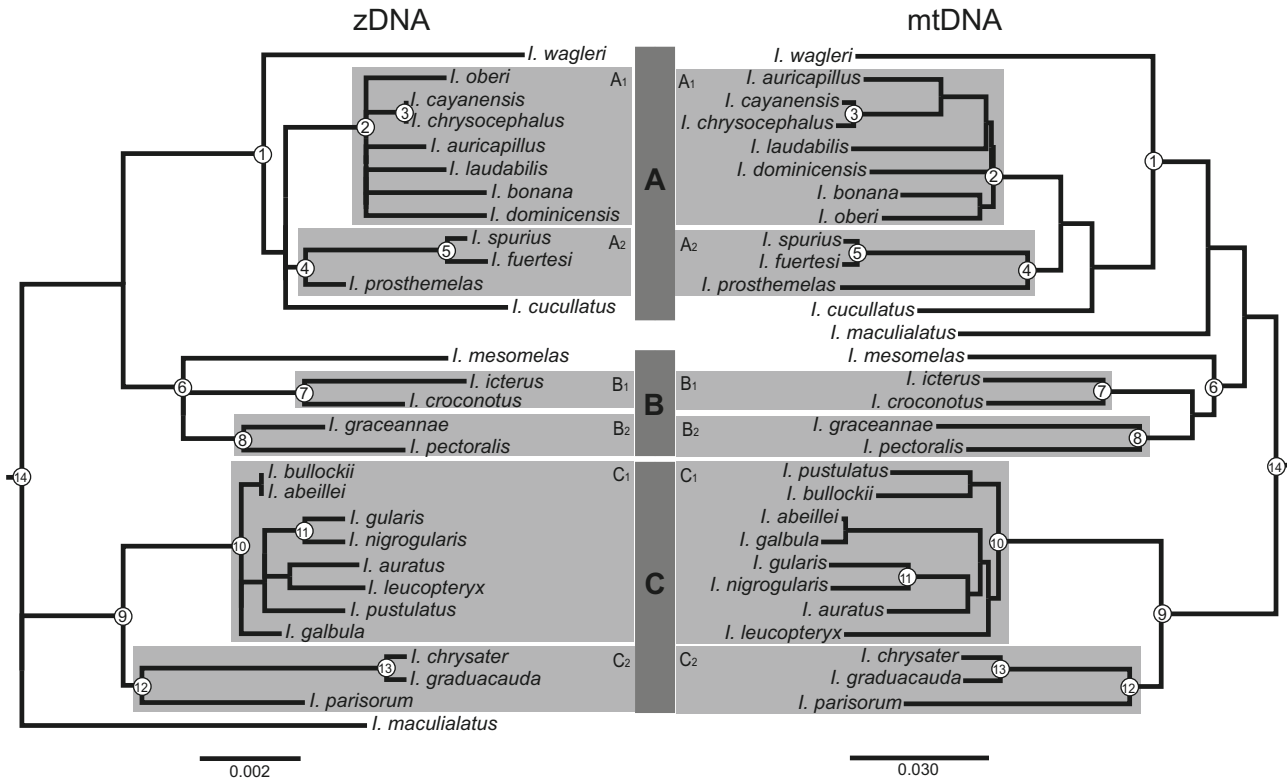
The ILD test revealed no strong conflict between these two independent estimates of the underlying species tree ( $p = 0.655$ ). However, maximum likelihood comparisons of mtDNA and Z-intron trees (topology plus branch lengths) indicated significant incongruence between nuclear and mitochondrial data sets (Table 3). The most obvious conflict between the mtDNA and Z-intron trees concerns the placement of *Icterus maculialatus* (Figs. 1 and 2). A Z-intron tree constrained to fit this particular aspect of the mtDNA tree topology (i.e., with *I. maculialatus* sister to clade A) was a significantly worse fit to the concatenated Z-intron data set than the unconstrained tree (KH test: diff  $L = -17.44$ ;  $p = 0.033$ ; see Table 3). On the other hand, a mtDNA tree constrained to fit the Z-intron tree topology (i.e., with *I. maculialatus* as sister to clade C) was not a significantly worse fit than the unconstrained mtDNA tree (KH test: diff  $L < -0.001$ ;  $p = 0.495$ ).

#### 4. Discussion

Combined analysis of multiple Z-linked introns resolved a well-supported nuclear tree of the New World orioles (*Icterus*).

Obtaining such high resolution within a recent radiation like *Icterus* using slowly evolving nuclear introns is quite surprising, especially given the few informative characters present in the data set (71 ingroup characters including indels). Although the nuclear Z-intron characters are much fewer in number, they display little homoplasy compared to mtDNA as indicated by the Consistency Index and the Rescaled Consistency index (CI = 0.89 and 0.42; RC = 0.81 and 0.23 for Z-intron DNA and mtDNA, respectively). Clearly, the Z-intron characters are highly informative once analyzed in combination. While individual gene trees generally provided poor resolution for oriole relationships within the three major clades, combining data from multiple Z-linked introns resolved nearly all but the most recent divergences (2–3 million years ago or about 5% mtDNA divergence: see Omland et al., 1999). Strikingly, almost all of the deeper nodes in this Z-intron tree agree with mtDNA.

The New World orioles present perhaps a unique opportunity to evaluate the utility of the concatenation approach at the species-level, because rarely is a robust mtDNA tree of such a densely sampled genus of birds available for direct comparison. Specifically, the presence of short (in coalescent units) internal branches in the species tree greatly increases the probability of incongruence among



**Fig. 2.** ML trees of 29 *Icterus* species based on concatenation of six Z-linked introns (left) and two mitochondrial genes (right). Fourteen labeled nodes are strongly supported by both data sets (see Table 2, for details). Shaded areas outline six sub-clades within the three major lineages of orioles ( $A_1$ ,  $A_2$ ,  $B_1$ ,  $B_2$ ,  $C_1$ ,  $C_2$ ) that agree between Z-intron DNA and mtDNA. Support indices for individual intron partitions are listed in Table 2. Branch lengths were estimated using GARLI – note the similarity in internode length between the two data sets.

**Table 3**

Tests of congruence between nuclear, mitochondrial, and total evidence data sets as evaluated using maximum likelihood based Kishino–Hasegawa (KH) tests. In addition to the overall fit of inferred tree topologies to the three data sets, the specific fit of conflicting placements of *I. maculialatus* to the mtDNA and Z-intron (zDNA) data sets was evaluated. Only ingroup taxa were considered for congruence testing. *P*-values were corrected for multiple testing.

Tree Topology Data Set	mtDNA	DNA	Total evidence (mt + zDNA)	<i>I. maculialatus</i> sister to A	<i>I. maculialatus</i> sister to C
mtDNA	Best	<0.001	0.215	Best	0.495
zDNA	0.008	Best	0.039	0.033	Best
mt + zDNA	0.173	<0.001	Best		

gene trees (i.e., the “anomaly zone”: Degnan and Rosenberg, 2006) and decreases the chances that concatenation infers the true species tree. Hence, when strong gene tree-species tree discord is suspected other non-concatenation approaches are required to achieve a robust estimate of the phylogeny (reviewed in Degnan and Rosenberg (2009)). Here, combined analysis of multiple Z-linked introns resolved a tree that largely corroborates a robust mtDNA tree, both with respect to branching order and branch lengths.

Whether the *Icterus* phylogeny is simply a relatively “easy” problem, or if our apparent success results from the use of faster sorting and less recombining loci linked to the Z chromosome that are perhaps more likely to reflect concordant histories is difficult to assess at this point. But recent work suggests that sex-linked loci (e.g., Z-introns) may indeed be more useful than autosomal loci near the species boundary (see Carling and Brumfield, 2008a; Petit and Excoffier, 2009; Storchová et al., 2010). Our ongoing population genetic studies on orioles provide even more compelling

evidence. A comparison of multiple Z-linked and autosomal intron loci sequenced across four closely related taxa revealed that Z-linked introns seem to have a significantly higher proportion of fixed differences between species relative to autosomal introns (F. Jacobsen, unpublished data). While mutations likely arise at a similar rate at Z-linked loci as expected at autosomal loci, lower  $N_e$  as well as hitchhiking and background selection could increase the rate of fixation at Z-linked loci (Betancourt et al., 2004; Singh and Petrov, 2007).

The most obvious disagreement between the combined Z-intron tree and mtDNA tree involving the position of *I. maculialatus* actually only represents weak conflict in the underlying data. Different weighting of mtDNA data result in different placements of this species (see Omland et al., 1999, for details), the concatenated Z-intron DNA only weakly supports a sister relationship with clade C, whereas the total evidence tree places it as sister to clades A and B with moderate support. The close sister relationship inferred between *I. abeillei* and *I. bullockii* by Z-intron DNA, to the exclusion of their respective mitochondrial sister taxa, was more unexpected given previous studies on these taxa (e.g., Kondo et al., 2004, 2008). We are now investigating this species complex using rigorous population-level studies; hybridization and mtDNA replacement is one of possible hypotheses we are testing to explain these conflicting relationships.

This study represents one of the first attempts to reconstruct a multilocus nuclear phylogeny of a species-rich genus of closely related vertebrate species. Our nuclear Z-intron DNA tree largely corroborates the published mitochondrial tree of the New World orioles (*Icterus*) and resolves many important clades within the genus. The few unresolved nodes involve very recent divergences (<2 million years ago: see Omland et al., 1999), possible gene introgression, and short branches associated with rapid radiations (e.g., coloniza-

tion of the Caribbean archipelago). We are currently testing a range of species-tree methods (e.g., Maddison and Knowles, 2006; Ané et al., 2007; Liu and Pearl, 2007; Heled and Drummond, 2010) and additional loci to resolve these challenging relationships.

## Acknowledgments

The authors are very grateful to S. Olson, G. Graves, T. Chesser, and J. Dean at USNM; R. Brumfield and D. Dittmann at LSUWMS; J. Klicka at MBM; J. Bates, S.J. Hackett, and D. Willard at FMNH; N.

Rice at ANSP; S.M. Lanyon and F.K. Barker at BMNH; and E. Bermingham and O. Sanjur at STRI for kindly providing tissue samples to this work. The authors also thank former and current members of the Omland Lab group for lab support and valuable discussions on this manuscript. Part of this work was carried out by using the resources of the Computational Biology Service Unit from Cornell University, which is partially funded by the Microsoft Corporation. This work was funded by a National Science Foundation CAREER Grant (DEB-0347083) to K.E.O., Maryland Ornithological Society Avian Research Grants and American Ornithologist Union Research Grant to F.J.

## Appendix A. Voucher specimen information

Genus/species	Sub-species	Museum	Catalog/field No.	Sampling locality
<i>Icterus abeillei</i>	Monotypic	MZFC	KEO-028	Mexico, Guanajuato
<i>Icterus auratus</i>	Monotypic	UAM	7222	Mexico, Yucatan, El Coyo
<i>Icterus auricapillus</i>	Monotypic	MBM	GMS-1823	Panama, Lago Bayano
<i>Icterus bonana</i>	Monotypic	STRI	MA-IBO1	Martinique, Fond Baron
<i>Icterus bullockii</i>	<i>bullockii</i>	FMNH	341938	USA, CA, Monterey Co.
<i>Icterus cayanensis</i>	<i>cayanensis</i>	USNM	610174	Guyana, Berbice
<i>Icterus chrysater</i>	<i>chrysater</i>	UWBM	DAB-1573	Nicaragua, Casitta
<i>Icterus chrysocephalus</i>	Monotypic	FMNH	339734	Venezuela, Sucre, Guanoco
<i>Icterus croconotus</i>	<i>croconotus</i>	USNM	632483	Guyana, Upper Takutu
<i>Icterus cucullatus</i>	<i>igneus</i>	MZFC	KEO-011	Mexico, Campeche, Xpujil
<i>Icterus dominicensis</i>	<i>dominicensis</i>	AMNH	NKK-1112	Dominican Republic, Peravia
<i>Icterus fuertesi</i>	Monotypic	MZFC	KEO-024	Mexico, Veracruz
<i>Icterus galbula</i>	Monotypic	BMNH	42547	USA, MN
<i>Icterus graduacauda</i>	<i>graduacauda</i>	BMNH	BMM-212	Mexico, Oaxaca
<i>Icterus graceannae</i>	Monotypic	ANSP	181810	Equador, Loja, Celica
<i>Icterus gularis</i>	<i>tamaulipensis</i>	MZFC	KEO-003	Mexico, Veracruz, Tlacoalpan
<i>Icterus icterus</i>	<i>ridgewayi</i>	LSUMZ	11328	Puerto Rico, Cabo Jojo
<i>Icterus laudabilis</i>	Monotypic	STRI	SL-ILA4	St. Lucia, Anse la Sorciere
<i>Icterus leucopteryx</i>	Monotypic	FMNH	331142	Jamaica, Cornwall Co.
<i>Icterus maculialatus</i>	Monotypic	INIREB	SRF-387	Mexico, Chiapas, Tuxtla Gut.
<i>Icterus mesomelas</i>	<i>mesomelas</i>	UWBM	52153	Mexico, Chiapas, Estacion Juarez
<i>Icterus nigrogularis</i>	<i>nigrogularis</i>	USNM	610091	Guyana, West Berbice
<i>Icterus oberi</i>	Monotypic	STRI	MO-IOB4	Montserrat, Soufriere
<i>Icterus parisorum</i>	Monotypic	FMNH	334367	USA, CA, San Bernardino Co.
<i>Icterus pectoralis</i>	<i>pectoralis</i>	BMNH	42544	USA, FL, Dade Co.
<i>Icterus prothemelas</i>	<i>prothemelas</i>	BMNH/MZFC	42543/KEO-018	Mexico, Campeche, Xpujil
<i>Icterus pustulatus</i>	<i>formosus</i>	UWBM	52129	Mexico, Chiapas, Ocozucualta
<i>Icterus spurius</i>	Monotypic	FMNH	389579	USA, IL, Cook Co.
<i>Icterus wagleri</i>	<i>wagleri</i>	MZFC	04 Sierra de Huatla	Mexico, Morelos
<i>Agelaius phoeniceus</i>	–	USNM	626365	USA, FL, Hillsborough Co.
<i>Cacicus solitarius</i>	–	USNM	609510	Argentina, Corrientes
<i>Sturnella neglecta</i>	–	USNM	586115	USA, WA, Grant Co.

## References

- Allen, E.S., Omland, K.E., 2003. Novel intron phylogeny supports plumage convergence in orioles (*Icterus*). *Auk* 120, 961–969.
- Ané, C., Larget, B., Baum, D.A., Smith, S.D., Rokas, A., 2007. Bayesian estimation of concordance among gene trees. *Mol. Biol. Evol.* 24, 412–426.
- Backström, N., Brandström, M., Gustafsson, L., Qvarnström, A., Cheng, H., Ellegren, H., 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.
- Backström, N., Karaiskou, N., Leder, E.H., Gustafsson, L., Primmer, C.R., Qvarnström, A., Ellegren, H., 2008. A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene order conservation during 100 million years of avian evolution. *Genetics* 179, 1479–1495.
- Ballard, J.W.O., Rand, D.M., 2005. The population biology of mitochondrial DNA and its phylogenetic implication. *Annu. Rev. Ecol. Evol. Syst.* 36, 621–642.
- Barker, F.K., Vandergon, A.J., Lanyon, S.M., 2008. Assessment of species limits among yellow-breasted meadowlarks (*Sturnella* spp.) using mitochondrial and sex-linked markers. *Auk* 125, 869–879.
- Benavides, E., Baum, R., McClellan, D., Jack, W., Sites, J., 2007. Molecular phylogenetics of the lizard genus *Microlophus* (Squamata: Tropicoduridae): aligning and retrieving indel signal from nuclear introns. *Syst. Biol.* 56, 776–797.
- Berlin, S., Smith, N.G.C., Ellegren, H., 2004. Do avian mitochondria recombine? *J. Mol. Evol.* 58, 163–167.
- Betancourt, A.J., Kim, Y., Orr, H.A., 2004. A pseudohitchhiking model of X vs. autosomal diversity. *Genetics* 168, 2261–2269.
- Borge, T., Webster, M.T., Andersson, G., Sætre, G.-P., 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes on two *Ficedula* flycatcher species. *Genetics* 171, 1861–1873.
- Carling, M.D., Brumfield, R.T., 2008a. Haldane's rule in an avian system: using cline theory and divergence population genetics to test for differential introgression of mitochondrial, autosomal, and sex-linked loci across the *Passerina* bunting hybrid zone. *Evolution* 62, 2600–2615.
- Carling, M.D., Brumfield, R.T., 2008b. Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. *Genetics* 178, 363–377.
- Carstens, B.C., Knowles, L.L., 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Syst. Biol.* 56, 400–411.

- Clark, W.S., Witt, C.C., 2006. First known specimen of a hybrid *Buteo*: Swainson's Hawk (*Buteo swainsoni*) × Rough-legged Hawk (*B. Lagopus*) from Louisiana. *Wilson J. Ornithol.* 118, 42–52.
- Clements, J.F., 2007. The Clements Checklist of Birds of the World. Comstock Publishing Associates, Ithaca, NY.
- de Queiroz, A., Donoghue, M.J., Kim, J., 1995. Separate versus combined analysis of phylogenetic evidence. *Annu. Rev. Ecol. Syst.* 26, 657–681.
- Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. *PLoS Genet.* 2, 762–768.
- Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. *Trends Ecol. Evol.* 24, 332–340.
- Delsuc, F., Stanhope, M.J., Douzery, E.J.P., 2003. Molecular systematics of armadillos (Xenarthra, Dasypodidae): contribution of maximum likelihood and Bayesian analyses of mitochondrial and nuclear genes. *Mol. Phyl. Evol.* 28, 261–275.
- Ellegren, H., Fridolfsson, A.-K., 1997. Male-driven evolution of DNA sequences in birds. *Nat. Genet.* 17, 182–184.
- Ezaz, T., Stiglec, R., Veyrunes, F., Graves, J.A.M., 2006. Relationships between vertebrate ZW and XY sex chromosome systems. *Curr. Biol.* 16, R736–R743.
- Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1994. Testing significance of incongruence. *Cladistics* 10, 315–319.
- Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39, 783–791.
- Fridolfsson, A.-K., Cheng, H., Copeland, N.G., Jenkins, N.A., Liu, H.-C., Raudsepp, T., Woodage, T., Chowdhary, B., Halvorson, J., Ellegren, H., 1998. Evolution of the avian sex chromosome from an ancestral pair of autosomes. *Proc. Natl. Acad. Sci. USA* 95, 8147–8152.
- Friedman, N.R., Hofmann, C.M., Kondo, B., Omland, K.E., 2009. Correlated evolution of migration and sexual dichromatism in New World orioles (*Icterus*). *Evolution* 63, 3269–3274.
- Friesen, V.L., Congdon, B.C., Kidd, M.G., Birt, T.P., 1999. Polymerase chain reaction (PCR) primers for the amplification of five nuclear introns in vertebrates. *Mol. Ecol.* 8, 2147–2149.
- Funk, D.J., Omland, K.E., 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. *Annu. Rev. Ecol. Syst.* 34, 397–423.
- Hackett, S.J., Kimball, R.T., Reddy, S., Bowie, R.C.K., Braun, E.L., Braun, M.J., Chojnowski, J.L., Cox, W.A., Han, K.-L., Harshman, J., Huddleston, C.J., Marks, B.D., Miglia, K.J., Moore, W.S., Sheldon, F.H., Steadman, D.W., Witt, C.C., Yuri, T., 2008. A phylogenomic study of birds reveals their evolutionary history. *Science* 320, 1763–1768.
- Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. *Mol. Biol. Evol.* 27, 570–580.
- Hofmann, C.M., Cronin, T.W., Omland, K.E., 2006. Using spectral data to reconstruct evolutionary changes in coloration: carotenoid color evolution in New World Orioles. *Evolution* 60, 1680–1691.
- Hofmann, C.M., Cronin, T.W., Omland, K.E., 2007a. Melanin coloration in New World orioles II: ancestral state reconstruction reveals lability in the use of carotenoids and pheomelanins. *J. Avian Biol.* 38, 172–181.
- Hofmann, C.M., Cronin, T.W., Omland, K.E., 2008a. Evolution of sexual dichromatism. 1. Convergent losses of elaborate female coloration in New World orioles (*Icterus* spp.). *Auk* 125, 778–789.
- Hofmann, C.M., Cronin, T.W., Omland, K.E., 2008b. Evolution of sexual dichromatism. 2. Carotenoids and melanins contribute to sexual dichromatism in New World orioles (*Icterus* spp.). *Auk* 125, 790–795.
- Hofmann, C.M., McGraw, K.J., Cronin, T.W., Omland, K.E., 2007b. Melanin coloration in New World Orioles I: carotenoid masking and pigment dichromatism in the orchard oriole complex. *J. Avian Biol.* 38, 163–171.
- Hudson, R.R., Turelli, M., 2003. Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57, 182–190.
- Huelsenbeck, J.P., Ronquist, F., 2001. MRBAYES: Bayesian inference of phylogenetic trees. *BMC Bioinformatics* 17, 754–755.
- Irimia, M., Roy, S.W., 2008. Spliceosomal introns as tools for genomic and evolutionary analysis. *Nucleic Acids Res.* 36, 1703–1712.
- Jarne, P., Lagoda, P.J.L., 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* 11, 424–429.
- Kishino, H., Hasegawa, M., 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea. *J. Mol. Evol.* 29, 170–179.
- Knowles, L.L., 2009. Estimating species trees: methods of phylogenetic analysis when there is incongruence across genes. *Syst. Biol.* doi:10.1093/sysbio/syp1061.
- Kondo, B., 2006. Speciation and the Evolution of Migration: A Phylogenetic Examination using New World Orioles. Ph.D. Diss UMBC Baltimore, MD, 92.
- Kondo, B., Baker, J.M., Omland, K.E., 2004. Recent speciation between the Baltimore Oriole and the Black-backed Oriole. *Condor* 106, 674–680.
- Kondo, B., Peters, J.L., Rosensteel, B.B., Omland, K.E., 2008. Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution* 62, 1182–1191.
- Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. *Syst. Biol.* 56, 17–24.
- Lanyon, S.M., Omland, K.E., 1999. A molecular phylogeny of the blackbirds (Icteridae): five lineages revealed by cytochrome-*b* sequence data. *Auk* 116, 629–639.
- Liu, L., Edwards, S.V., 2009. Phylogenetic analysis in the anomaly zone. *Syst. Biol.* 58, 452–460.
- Liu, L., Pearl, D.K., 2007. Species trees from gene trees: reconstructing Bayesian posterior distributions of a species phylogeny using estimated gene trees distributions. *Syst. Biol.* 56, 504–514.
- Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55, 21–30.
- Mank, J.E., Promislow, D.E.L., Avise, J.C., 2006. Evolution of alternative sex-determining mechanisms in teleost fishes. *Biol. J. Linn. Soc.* 87, 83–93.
- Mendelson, T.C., Shaw, K.L., 2005. Rapid speciation in an arthropod. *Nature* 433, 375–376.
- Mendelson, T.C., Simons, J.N., 2006. AFLPs resolve cytonuclear discordance and increase resolution among barcheek darters (Percidae: *Etheostoma*: *Catonotus*). *Mol. Phyl. Evol.* 41, 445–453.
- Meudt, H.M., Clarke, A.C., 2007. Almost forgotten or latest practice? AFLP applications, analyses and advances. *Trends Plant Sci.* 12, 106–117.
- Moore, W.S., 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49, 718–726.
- Müller, K., 2006. Incorporating information from length-mutational events into phylogenetic analysis. *Mol. Phyl. Evol.* 38, 667–676.
- Müller, K.F., 2005. SeqState-primer design and sequence statistics for phylogenetic DNA data sets. *Appl. Bioinformatics* 4, 65–69.
- Nakamura, M., 2009. Sex determination in amphibians. *Semin. Cell Dev. Biol.* 20, 271–282.
- Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P., Nieves-Aldrey, J.L., 2004. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* 53, 47–67.
- Omland, K.E., Lanyon, S.M., Fritz, S.J., 1999. A molecular phylogeny of the New World orioles (*Icterus*): the importance of dense taxon sampling. *Mol. Phyl. Evol.* 12, 224–239.
- Pagès, M., Calvignac, S., Klein, C., Paris, M., Hughes, S., Hanni, C., 2008. Combined analysis of fourteen nuclear genes refines the Ursidae phylogeny. *Mol. Phyl. Evol.* 47, 73–83.
- Pamilo, P., Nei, M., 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* 5, 568–583.
- Parra, J.L., Remsen Jr., J.V., Alvarez-Rebolledo, M., McGuire, J.A., 2009. Molecular phylogenetics of the hummingbird genus *Coeligena*. *Mol. Phyl. Evol.* 53, 425–434.
- Peters, J.L., McCracken, K.G., Zhuravlev, Y.N., Lu, Y., Wilson, R.E., Johnson, K.P., Omland, K.E., 2005. Phylogenetics of wigeons and allies (Anatidae: *Anas*): the importance of sampling multiple loci and multiple individuals. *Mol. Phyl. Evol.* 35, 209–224.
- Petit, R.J., Excoffier, L., 2009. Gene flow and species delimitation. *Trends Ecol. Evol.* 24, 386–393.
- Petren, K., Grant, B.R., Grant, P.R., 1999. A phylogeny of Darwin's finches based on microsatellite DNA length variation. *Proc. R. Soc. B* 266, 321–329.
- Planet, P.J., 2006. Tree disagreement: measuring and testing incongruence in phylogenies. *J. Biomed. Inform.* 39, 86–102.
- Price, J.J., Friedman, N.R., Omland, K.E., 2007. Song and plumage evolution in the New World orioles (*Icterus*) show similar lability and convergence in patterns. *Evolution* 61, 850–863.
- Rambaut, A., Drummond, A.J., 2004. Tracer v1.3. Available from: <http://evolve.zoo.ox.ac.uk/software.html>.
- Rokas, A., Ladoukakis, E., Zouros, E., 2003a. Animal mitochondrial DNA recombination revisited. *Trends Ecol. Evol.* 18, 411–417.
- Rokas, A., Williams, B.L., King, N., Carroll, S.B., 2003b. Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature* 425, 798–804.
- Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19, 1572–1574.
- Sætre, G.-P., Borge, T., Lindroos, K., Haavie, J., Sheldon, B.C., Primmer, C., Syvänen, A.-C., 2002. Sex chromosome evolution and speciation in *Ficedula* flycatchers. *Proc. R. Soc. B* 270, 53–59.
- Singh, N.D., Petrov, D.A., 2007. Evolution of gene function on the X chromosome versus the autosomes. *Genome Dyn.* 3, 101–118.
- Sota, T., Sasabe, M., 2006. Utility of nuclear allele networks for the analysis of closely related species in the genus *Carabus*, subgenus *Ohomopterus*. *Syst. Biol.* 55, 329–344.
- Steppan, S.J., Adkins, R.M., Spinks, P.Q., Hale, C., 2005. Multigene phylogeny of the Old World mice, Murinae, reveals distinct geographic lineages and the declining utility of mitochondrial genes compared to nuclear genes. *Mol. Phyl. Evol.* 37, 370–388.
- Storchová, R., Reif, J., Nachman, M.W., 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution* 64, 456–471.
- Sturge, R.J., Jacobsen, F., Rosensteel, B.B., Neale, R.J., Omland, K.E., 2009. Colonization of South America from Caribbean Islands confirmed by molecular phylogeny with increased taxon sampling. *Condor* 111, 575–579.
- Sullivan, J.P., Lavoué, S., Arnegard, M.E., Hopkins, C.D., 2004. AFLPs resolve phylogeny and reveal mitochondrial introgression within a species flock of African electric fish (Mormyroidea: Teleostei). *Evolution* 58, 825–841.
- Swofford, D.L., 2002. PAUP: Phylogenetic Analysis using Parsimony (and other methods) Version 4.0. Sinauer Associates, Sunderland, MA.
- Takahata, N., 1989. Gene genealogy in three related populations: consistency probability between gene and population trees. *Genetics* 122, 957–966.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL X windows interface: flexible strategies for multiple sequence

- alignment aided by quality analysis tools. *Nucleic Acids Res.* 25, 4876–4882.
- Traut, W., Sahara, K., Marec, F., 2007. Sex chromosomes and sex determination in Lepidoptera. *Sex Dev.* 1, 332–346.
- Tsaousis, A.D., Martin, D.P., Ladoukakis, E.D., Posada, D., Zouros, E., 2005. Widespread recombination in published animal mtDNA sequences. *Mol. Biol. Evol.* 22, 925–933.
- Wiens, J.J., Kuczynski, C.A., Smith, S.A., Mulcahy, D.G., Jack, W., Sites, J., Townsend, T.M., Reeder, T.W., 2008. Branch lengths, support, and congruence: testing the phylogenomics approach with 20 nuclear loci in snakes. *Syst. Biol.* 57, 420–431.
- Yang, Z., 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591.
- Zink, R.M., Barrowclough, G.F., 2008. Mitochondrial DNA under siege in avian phylogeography. *Mol. Ecol.* 17, 2107–2121.
- Zwickl, D.J., 2006. GARLI Version 0.951. Available from: [www.bio.utexas.edu/faculty/antisense/garli/Garli.html](http://www.bio.utexas.edu/faculty/antisense/garli/Garli.html), Austin, Texas, p. Genetic Algorithm for Rapid Likelihood Inference.