

## Extensive introgressive hybridization within the northern oriole group (Genus *Icterus*) revealed by three-species isolation with migration analysis

Frode Jacobsen & Kevin E. Omland

Department of Biological Sciences, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD, 21250, USA

### Keywords

Allele sharing, coalescent, *Icterus*, IMa2, incomplete lineage sorting, introgression, multipopulation, northern orioles.

### Correspondence

Frode Jacobsen, Department of Biological Sciences, University of Maryland Baltimore County, 1000 Hilltop Circle, Baltimore, MD 21250, USA. Tel: +1 410 455 1704; Fax: +1 410 455 3875; E-mail: frode1@umbc.edu

### Funding Information

This work was funded by National Science Foundation grants to K.E.O. (DEB-0347083 and DEB-1119506), and Maryland Ornithological Society Avian Research Grants and an American Ornithologist's Union Research Grant to F.J.

Received: 20 June 2012; Revised: 30 July 2012; Accepted: 31 July 2012

*Ecology and Evolution* 2012; 2(10): 2413–2429

doi: 10.1002/ece3.365

### Abstract

Until recently, studies of divergence and gene flow among closely-related taxa were generally limited to pairs of sister taxa. However, organisms frequently exchange genes with other non-sister taxa. The “northern oriole” group within genus *Icterus* exemplifies this problem. This group involves the extensively studied hybrid zone between Baltimore oriole (*Icterus galbula*) and Bullock's oriole (*I. bullockii*), an alleged hybrid zone between *I. bullockii* and black-backed oriole (*I. abeillei*), and likely mtDNA introgression between *I. galbula* and *I. abeillei*. Here, we examine the divergence population genetics of the entire northern oriole group using a multipopulation Isolation-with-Migration (IM) model. In accordance with Haldane's rule, nuclear loci introgress extensively beyond the *I. galbula*–*I. bullockii* hybrid zone, while mtDNA does not. We found no evidence of introgression between *I. bullockii* and *I. abeillei* or between *I. galbula* and *I. abeillei* when all three species were analyzed together in a three-population model. However, traditional pairwise analysis suggested some nuclear introgression from *I. abeillei* into *I. galbula*, probably reflecting genetic contributions from *I. bullockii* unaccounted for in a two-population model. Thus, only by including all members of this group in the analysis was it possible to rigorously estimate the level of gene flow among these three closely related species.

## Introduction

The study of phylogeography and population divergence is challenged by the difficulty of distinguishing between shared retained ancestral polymorphism and introgressed alleles (Nielsen and Wakeley 2001). The much larger effective population size of nuclear DNA relative to mtDNA slows the rate of lineage sorting and the speed at which reciprocal monophyly is achieved at the nuclear level (Avice 2004). Slow and stochastic lineage sorting thus causes a pattern of allele sharing among species that can seem similar to patterns caused by gene flow through introgressive hybridization, even between long-divergent lineages (Hudson and Turelli 2003). Multiple unlinked loci and coalescent-based “divergence population genet-

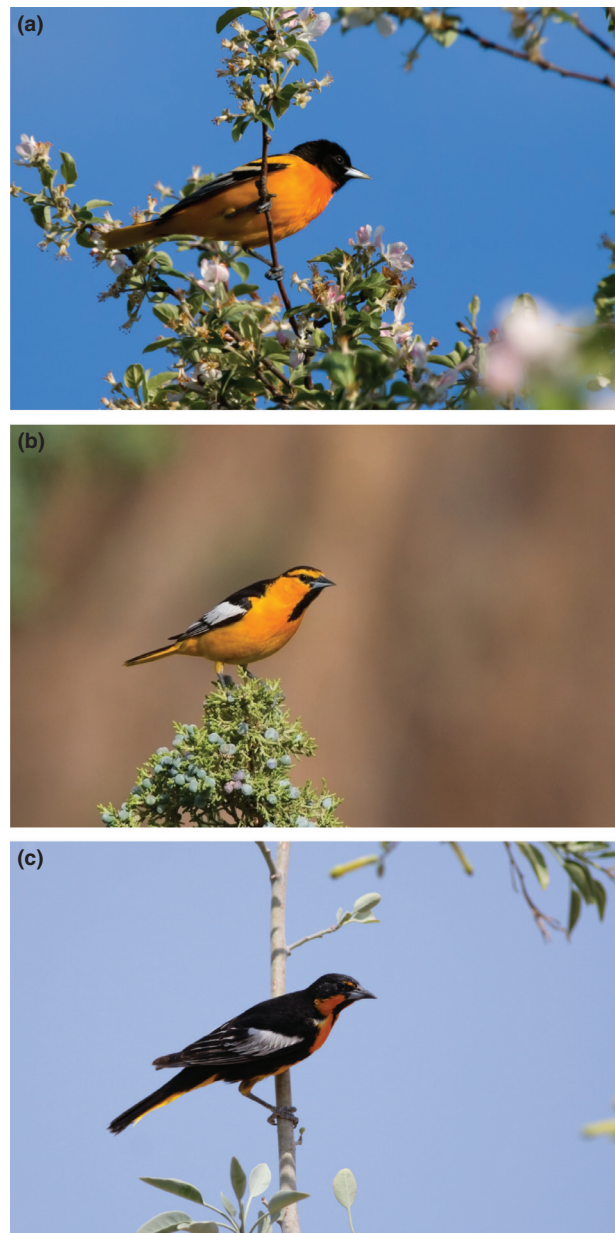
ics” methods are needed to distinguish between these two processes (e.g., Kliman et al. 2000; Machado et al. 2002).

The Isolation-with-Migration (IM) model (Hey and Nielsen 2004, 2007) provides a powerful tool for the study of population divergence and for testing different speciation models and demographic scenarios (Hey 2006; Peters et al. 2007; Kondo et al. 2008; Pinho and Hey 2010). Shared alleles allow multilocus coalescent approaches such as IM to reconstruct the divergence history of closely related species (Nielsen and Wakeley 2001; Hey and Nielsen 2004, 2007). The original IM model was designed for two-population studies between sister taxa that have not exchanged genes with other taxa since they split (Hey and Nielsen 2004, 2007). However, organisms frequently exchange genes with other non-sister taxa (Arnold 2006;

Nosil 2008; Petit and Excoffier 2009). Genetic contributions from other source populations not included in the analysis (i.e., “ghost” populations) could thus severely bias or invalidate inference about the history of closely related taxa (Beerli 2004; Slatkin 2005; Strasburg and Rieseberg 2010). Thus, until recently, IM analyses were limited to pairs of taxa that share more recent common ancestry and gene flow with each other. However, this limitation was removed with the release of IMA2, allowing the simultaneous consideration of up to ten populations granted a population tree (i.e., phylogeny) is available for the taxa involved (IMA2: Hey 2010b).

In this study, we used IMA2 to analyze the three species in the “northern oriole” group within the New World orioles (Jacobsen and Omland 2011; Fig. 1). This group includes the well-known hybrid zone between the eastern Baltimore oriole (*Icterus galbula*) and Bullock’s oriole (*I. bullockii*) (Sutton 1938, 1968; Sibley and Short 1964; Rising 1969, 1970, 1983, 1996; Corbin et al. 1979). *I. galbula* breeds across the eastern United States south to Louisiana, and across most of Canada from Alberta in the west to Nova Scotia in the east (Fig. 2). The breeding range of *I. bullockii* extends across western North America from British Columbia, Canada in the north to Durango, Mexico in the south. Both *I. galbula* and *I. bullockii* are long-distance migrants and mainly overwinter in southern Mexico and Central America (Rising et al. 1998, 1999). Their winter ranges overlap with the range of the third member of the group, *I. abeillei*, a Mexican endemic limited to the central Mexican plateau (Jaramillo and Burke 1999) and traditionally considered a close relative and subspecies within *I. bullockii* due to their overall phenotypic similarity (Jaramillo and Burke 1999).

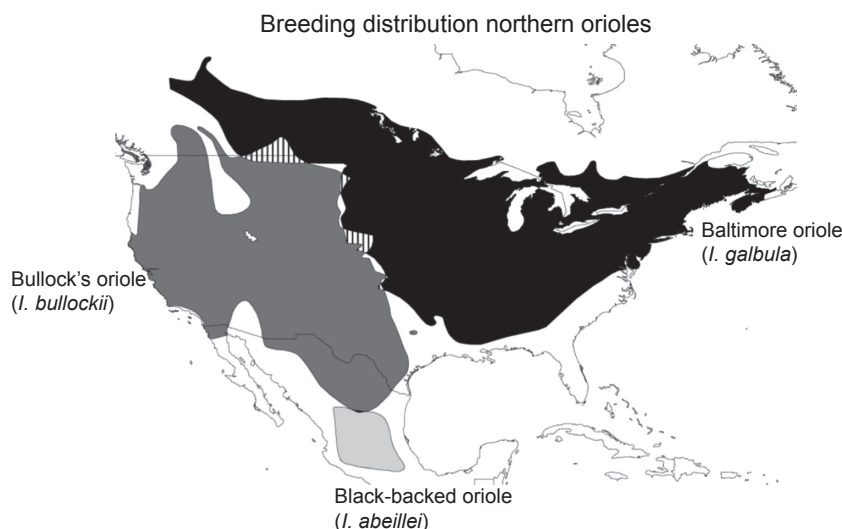
The parapatric breeding ranges of *I. galbula* and *I. bullockii* overlap along a long and narrow transect (150 km in Kansas, Allen 2002) that stretches over 1800 km from southern Alberta to central Texas (Rising 1983). The two orioles interbreed extensively in this contact zone, which falls within the “Rocky Mountain–Great Plains” suture zone (Remington et al. 1968), a hotspot of avian hybrid zones between eastern and western species pairs (e.g., Carling and Brumfield 2008; Flockhart and Wiebe 2009; Mettler and Spellman 2009). Documentation of viable hybrid offspring within this contact zone (Sutton 1938; Sibley and Short 1964; Sutton 1968; Rising 1970) led to the lumping of *I. galbula* and *I. bullockii* (including *I. abeillei*) into a single species, the northern oriole (AOU 1973). Later studies found that the *I. galbula*–*I. bullockii* hybrid zone was stable (Rohwer and Manning 1990; Rohwer and Johnson 1992; Rising 1996; Allen 2002), and full species status was restored for all three members of the group (AOU 1995). Hybridization has never been firmly documented between *I. abeillei* and the other two mem-



**Figure 1.** Adult males of the three species in the northern oriole group; (a) Baltimore oriole *Icterus galbula* (photo: Frode Jacobsen), (b) Bullock’s oriole *I. bullockii* (photo: Frode Jacobsen), and (c) black-backed oriole *I. abeillei* (photo: Jonathan Hiley)

bers of the group, but there is some evidence suggesting hybridization with *I. bullockii* where their ranges abut in northern Durango, Mexico (Miller 1906; but see Rising 1973). We are unaware of any documented instances of hybridization between members of the northern oriole group and other oriole species.

Despite the extensive research on the northern oriole group, we have limited knowledge about the extent of nuclear gene flow within this species complex. Allen



**Figure 2.** Breeding distributions of the three northern oriole species: *Icterus abeillei* (light gray), *I. bullockii* (dark gray), and *I. galbula* (black). Hatched areas indicate range overlap between *I. bullockii* and *I. galbula*. Digital distribution maps for each species were downloaded from NatureServe (Ridgely et al. 2007).

(2002) examined clinal variation of phenotypic and molecular markers across the *I. galbula*–*I. bullockii* hybrid zone in Kansas, and found that neither male plumage traits nor mtDNA haplotypes introgressed beyond the zone of contact. These findings were recently corroborated by Carling et al. (2011), who estimated the mitochondrial cline to be 328 km wide. In contrast, of 152 neutral AFLP markers examined by Allen (2002), there were no fixed differences between the two species, and only ten AFLP bands differed more than 40% in frequency across the hybrid zone. Allen (2002) interpreted this apparent lack of differentiation in AFLP markers between *I. galbula* and *I. bullockii* as evidence of unhampered nuclear introgression since the two species came into secondary contact. However, such extensive allele sharing among closely related species could also result from the slow sorting of retained ancestral polymorphism (Moore 1995; Hudson and Turelli 2003).

The goals of this study were to (1) estimate levels of introgression among all three species in the northern oriole group while accounting for their recent shared ancestry using a three-population IMA2 model and (2) examine the effects of genetic exchange with unsampled taxa in two-population analyses.

## Material and Methods

### Taxonomic sampling

We sought a broad representation from throughout the ranges of all three species whilst avoiding known areas of

sympatry (see Appendix 1 for sampling locations). Our sampling scheme provided a total of 21 to 24 individuals of each of the three species, *I. abeillei*, *I. bullockii* and *I. galbula*. In addition, a single individual of western meadowlark (*Sturnella neglecta*) was included as an out-group taxon.

### Multilocus nuclear data

This study included sequence data from eight intron-spanning loci that map to six autosomes and the Z sex chromosome in the zebra finch (*Taeniopygia guttata*) genome (see Table 1 for details). The two Z-linked loci are well separated on the zebra finch Z chromosome (>7 Mb apart) and were therefore treated as independent (Backström et al. 2006, 2010). All eight loci were amplified and sequenced following procedures described in previous phylogenetic studies on *Icterus* (Jacobsen et al. 2010; Jacobsen and Omland 2011). Resulting DNA chromatograms were aligned and edited in SEQUENCHER 4.7 (Gene Codes Corp., Ann Arbor, MI). Edited sequences for all loci were imported into MEGA5 (Tamura et al. 2011) and aligned using MUSCLE (Edgar 2004) with final manual adjustments. The haplotypes of individuals heterozygous at more than one position at a given locus were determined using the software PHASE v.2.1.1 (Stephens et al. 2001; Stephens and Scheet 2005), with the following parameter settings: burn-in = 1000, number of iterations = 10,000, and thinning interval = 1. Each analysis was repeated 10 times using different random starting seed and the run that received the highest

**Table 1.** Locus information. Chromosome: intron location determined based on BLAST searches of zebra finch reference genome assembly, with start coordinates for Z-linked loci provided in parentheses; Length: size (in base pairs) of largest non-recombinant alignment segment; *N*: number of non-recombinant alleles included in the study.

Locus <sup>1</sup>	Chromosome	Length	<i>N</i>			Primer source
			<i>I. abeillei</i>	<i>I. bullockii</i>	<i>I. galbula</i>	
<i>β-ACT2</i>	22	558	24	26	24	(Waltari and Edwards 2002)
<i>α-ENO8</i>	21	252	46	42	46	(Kondo et al. 2008)
<i>FGB4</i>	4	575	24	30	28	(Barker 2004)
<i>GAPDH11</i>	1	327	36	30	30	(Primmer et al. 2002)
<i>RDP2</i>	12	298	36	28	26	(Waltari and Edwards 2002)
<i>TGFβ5</i>	3	575	30	32	28	(Bureš et al. 2002)
<i>ALDOB5</i>	Z (400,071)	253	36	35	24	(Kondo et al. 2008)
<i>BRM15</i>	Z (7,528,458)	362	40	35	36	(Borge et al. 2005)

<sup>1</sup>Full locus information is as follows: Beta-actin gene, intron 2 (*β-ACT2*); alpha-enolase gene, intron 8 (*α-ENO8*); fibrinogen B beta polypeptide gene, intron 4 (*FGB4*); glyceraldehyde-3-phosphate dehydrogenase gene, intron 11 (*GAPDH11*); rhodopsin gene, intron 2 (*RDP2*); and transforming growth factor beta-2, intron 5 (*TGFβ5*); aldolase-B fructose-biphosphate intron 5 (*ALDOB5*); Brahma protein gene, intron 15 (*BRM15*).

log-likelihood was used to infer the haplotypes. To avoid systematic bias in downstream coalescent analyses, all haplotypes resolved at a probability of 0.5 or higher were included in the study (Garrick et al. 2010). An ambiguously aligned single-nucleotide repeat (poly-A) region within locus *FGB4* was excluded from the alignment prior to analyses. Linkage analysis of the two Z-linked loci (*ALDOB5* and *BRM15*) using HAPLOVIEW v.4.2 (Barrett et al. 2005) revealed no evidence of LD among polymorphic sites (results not shown). We tested for intralocus recombination using the difference in sums of squares (DSS) test in TOPALi v.2.5 (Milne et al. 2004), with 500 bootstrap replicates to determine statistical significance. A significant DSS peak was detected at locus *β-ACT2*, and only the largest non-recombinant sequence block (558 bp) was retained for downstream analysis. Accession numbers for sequences obtained from previous studies on this group are listed in Appendix 2. All new sequences collected in this study have been deposited in GenBank under accession numbers JX403068–JX403342 and JX412950–JX412955.

Three common measures of nucleotide polymorphism were calculated using ARLEQUIN v.3.11 (Excoffier and Lischer 2010): number of haplotypes (*h*), haplotype diversity (*Hd*), and the average number of nucleotide differences per site between two sequences (*π*). Haplotype networks illustrating the relationships among inferred alleles were constructed for each locus using the median-joining algorithm in NETWORK v.4.5.1.0 (Bandelt et al. 1999, <http://www.fluxus-engineering.com>). The networks were edited using NETWORK PUBLISHER v.1.1.0.7 (<http://www.fluxus-engineering.com>).

Tests for selection were conducted using a multilocus HKA test (Hudson et al. 1987) implemented in the HKA program available at <http://lifesci.rutgers.edu/~hey/hey/HeylabSoftware.htm>. We chose western meadowlark

(*Sturnella neglecta*) as outgroup for this test, the most distant relative within the family Icteridae available for comparison across the eight loci examined in this study. Kimura 2P distances between the three northern oriole species and *S. neglecta* were calculated using DNASP v.5.10.0.1 (Librado and Rozas 2009). Statistical significance of the HKA test was determined through 10,000 coalescent simulations based on the number of sampled alleles and observed levels of polymorphism and divergence across loci. These simulations also allowed us to conduct a multilocus test of Tajima's *D* statistic (Tajima 1989). Fu's (1997) *F<sub>s</sub>* and Ramos-Onsins and Rozas (2002) *R<sub>2</sub>* neutrality statistics were also calculated using DnaSP. Confidence intervals and statistical significance of test values were calculated using 10,000 random permutations and coalescent simulations.

## Multilocus coalescent analyses

We examined patterns of gene flow among all three members of the northern oriole group using the Bayesian IM model implemented in the program IMA2 (Hey 2010b), which assumes random mating, constant population sizes, selective neutrality, free recombination among loci and no recombination within loci (Hey and Nielsen 2004, 2007; Hey 2010b).

By default, all model parameters estimated by IMA2 are scaled to the population mutation rate. The parameter estimates were converted into demographic units (i.e., effective population sizes in number of individuals and divergence times in years) using a neutral mutation rate of  $1.35 \times 10^{-9}$  substitutions per site per year for autosomal loci (Ellegren 2007) and  $1.45 \times 10^{-9}$  substitutions per site per year for Z-linked loci (Axelsson et al. 2004; Ellegren 2007). To reflect the different modes of inheritance, auto-



somal (1.0) and Z-linked (0.75) loci were assigned respective inheritance scalars in the IMA2 input file. Population sizes were converted into effective number of individuals using a generation time of 1.7 years, the average age at reproduction documented for multiple passerine lineages (Sæther et al. 2005). The actual generation time in orioles may well be slightly longer, as most oriole species exhibit delayed plumage maturation, which could delay the onset of reproduction (for males in particular) (e.g., Flood 1984; Rohwer and Manning 1990; Butcher 1991; Richardson and Burke 1999). Effective population sizes reported in this study are therefore likely conservative estimates of the actual effective population sizes. The HKY mutation model was specified for all loci.

### Three-population IMA2 analysis

First, we estimated levels of gene flow among all three species within the northern oriole group simultaneously in a three-population IMA2 analysis. We conducted separate analyses with two alternative population trees: (1) the nuclear species tree inferred by our recent species tree analysis of *Icterus* (Jacobsen and Omland 2011): (*I. galbula*, (*I. abeillei*, *I. bullockii*)) and (2) the mtDNA gene tree inferred by Omland et al. (1999): (*I. bullockii*, (*I. abeillei*, *I. galbula*)).

Short preliminary runs were conducted to determine appropriate upper bounds of the demographic parameters. Subsequently, two identical long runs of  $3\text{--}5 \times 10^6$  MCMC steps with a burnin period of  $1 \times 10^6$  steps were started with different starting seeds and run until minimum effective sample sizes (ESS) of 200 were achieved. We assessed convergence by monitoring the mixing properties of the MCMC during each run and by ensuring that similar parameter estimates were obtained from the two independent runs. Acceptable chain mixing (i.e., absence of long-term trends in plots of  $L[P]$  and  $t$  over the course of each run) and low autocorrelations ( $<0.1$ ) were achieved by running 100 chains under a

geometric heating scheme (with heating parameters set to  $a = 0.975$  and  $b = 0.75$ ). Population-specific parameter priors were provided in a separate file. We used exponential priors on migration rates (option-j7), and determined appropriate prior distribution means based on the peak posterior parameter estimates from preliminary runs.

### Two-population IMA2 analyses

To assess the effects of violating the crucial assumption of no allelic contributions from other source (i.e., “ghost”) populations, we also conducted three separate pairwise IMA2 analyses of the three northern oriole species. Run settings were kept unchanged, except for slight adjustment of parameter priors where posterior density distributions exceeded the upper bounds determined by preliminary runs. Each MCMC (M) mode analysis was replicated using different random seeds to ensure convergence onto similar posterior distributions, using the same criteria for assessing convergence as described above. Finally, we conducted nested model testing by running IMA2 in “Load genealogies” (L) mode to specifically compare the fit of reduced models that do not allow for post divergence gene flow to the full IMA2 model (i.e., including all six parameters  $\Theta_1$ ,  $\Theta_2$ ,  $\Theta_a$ ,  $m_1$ ,  $m_2$ ,  $t$ ). This was performed by sampling  $2.5 \times 10^5$  genealogies evenly from the two MCMC runs performed on each population pair and compared the log-likelihood scores of all possible nested models provided in a separate file in the IMA2 distribution package.

## Results

### Polymorphism and divergence

The three species within the northern oriole group were polymorphic at all loci used in this study, except for *I. abeillei* that was monomorphic at *ALDOB5*. Sample sizes and data for individual loci are summarized in

**Table 2.** Locus-specific estimates of number of haplotypes ( $h$ ), haplotype diversity ( $Hd$ ), and Jukes & Cantor nucleotide diversity ( $\pi$ ).

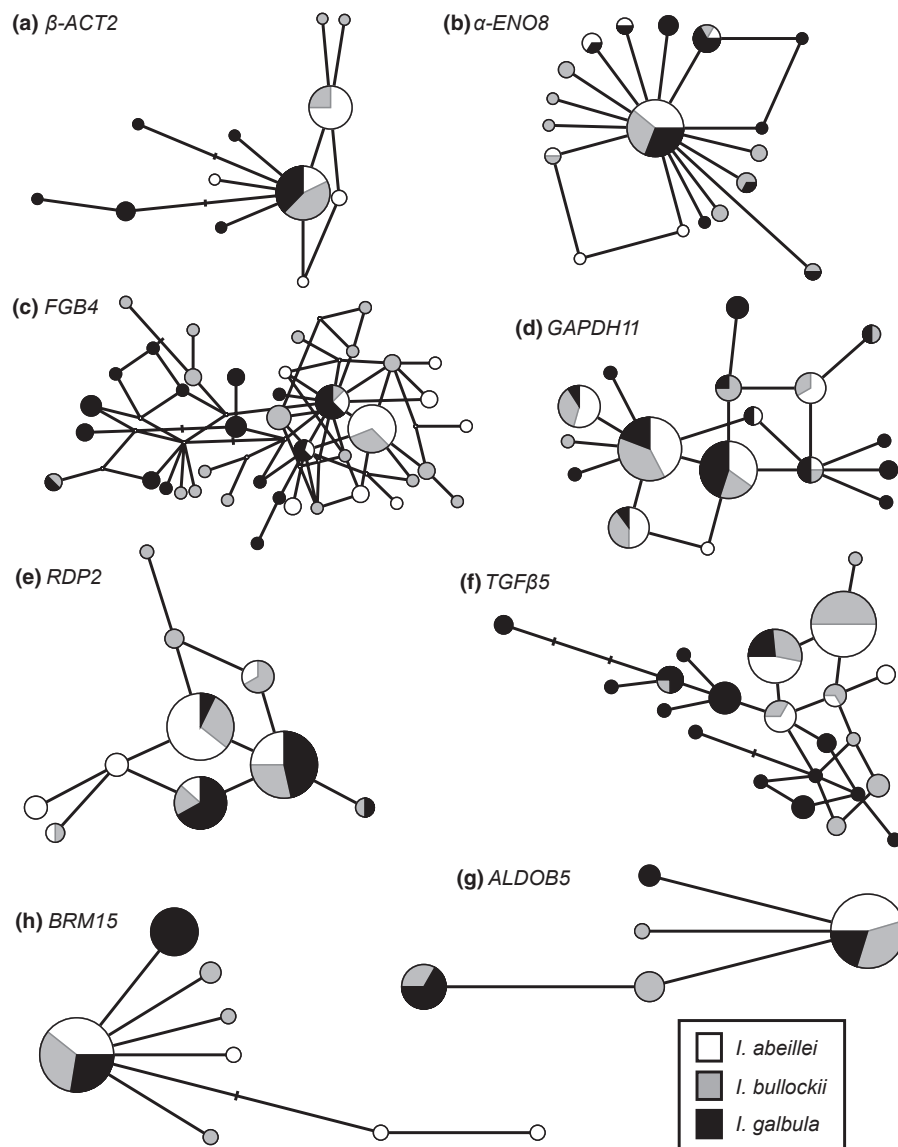
Locus	<i>I. abeillei</i>			<i>I. bullockii</i>			<i>I. galbula</i>		
	$h$	$Hd$	$\pi$	$h$	$Hd$	$\pi$	$h$	$Hd$	$\pi$
<i><math>\beta</math>-ACT2</i>	5	0.656	0.00147	4	0.397	0.00094	4	0.239	0.00060
<i><math>\alpha</math>-ENO8</i>	7	0.283	0.00136	9	0.490	0.00224	9	0.481	0.00234
<i>FGB4</i>	10	0.786	0.00300	20	0.956	0.00741	15	0.942	0.00603
<i>GAPDH11</i>	8	0.830	0.00472	8	0.823	0.00492	14	0.883	0.00590
<i>RDP2</i>	7	0.711	0.00385	7	0.804	0.00438	3	0.578	0.00216
<i>TGF<math>\beta</math>5</i>	3	0.393	0.00087	6	0.585	0.00213	14	0.923	0.00430
<i>ALDOB5</i>	1	0	0	4	0.395	0.00217	3	0.507	0.00372
<i>BRM15</i>	3	0.145	0.00068	4	0.217	0.00062	2	0.413	0.00114

Table 1. The three measures of sequence polymorphism were all lower in *I. abeillei* than in *I. bullockii* and *I. galbula* (Table 2), but not significantly so (single factor ANOVA: h:  $F_{2,21} = 0.656$ ,  $P = 0.529$ ; Hd:  $F_{2,21} = 0.585$ ,  $P = 0.566$ ;  $\pi$ :  $F_{2,21} = 0.943$ ,  $P = 0.405$ ; respectively). The differences in mean genetic diversity between *I. abeillei* and the two other members of the group are mainly driven by the lack of sequence variation at the sex-linked locus *ALDOB5* (Table 2).

The haplotype networks revealed extensive allele sharing and few fixed differences among the three northern

oriole species (Fig. 3). Some loci were characterized by a distinct star-shaped topology (e.g.,  $\alpha$ -*ENO8*, *ALDOB5* and *BRM15*), with a common central (and presumed ancestral) haplotype surrounded by several rare, derived haplotypes one or two mutational steps away from the ancestral haplotype (Fig. 3b,g,h).

The three species showed varying degrees of genetic differentiation across loci, but generally exhibited extensive haplotype sharing across loci (Table 3, also see Fig. 3). Strongest differentiation was observed between the highly disjunct *I. abeillei* and *I. galbula*, with highly significant



**Figure 3.** Haplotype networks of eight nuclear intron loci (6 autosomal + 2 Z-linked) revealing extensive haplotype sharing among the three northern orioles: *Icterus abeillei* (white), *I. bullockii* (gray), and *I. galbula* (black). Each circle corresponds to a unique haplotype and the size of each circle reflect the relative total sample size of each haplotype. Small open circles in (c) indicate unsampled/extinct haplotypes, and perpendicular bars along lines connecting haplotypes indicate additional mutations distinguishing neighboring haplotypes.

**Table 3.** Population differentiation among three northern oriole populations, *Icterus abeillei*, *I. bullockii*, and *I. galbula*, estimated based on pairwise genetic distances between populations. Significant  $F_{ST}$  values are indicated at the  $0.05 > P > 0.01$  (\*) and  $P < 0.01$  (\*\*) level. Bolded values are significant at the Bonferroni-adjusted  $\alpha$ -level of 0.00625.

Locus	$F_{ST}$		
	<i>I. abeillei</i> – <i>I. bullockii</i>	<i>I. abeillei</i> – <i>I. galbula</i>	<i>I. bullockii</i> – <i>I. galbula</i>
$\beta$ -ACT2	0.151**	<b>0.372**</b>	0.096*
$\alpha$ -ENO8	0.010	0.019*	0.019*
FGB4	<b>0.102**</b>	<b>0.383**</b>	<b>0.173**</b>
GAPDH11	0	0.064*	0.052*
RDP2	0.069*	<b>0.289**</b>	<b>0.177**</b>
TGF $\beta$ 5	0.065*	<b>0.400**</b>	<b>0.289**</b>
ALDO5	<b>0.131**</b>	<b>0.241**</b>	0.017
BRM15	0.018	<b>0.186**</b>	<b>0.186**</b>

$F_{ST}$  values at six of the eight loci examined. The two widely hybridizing *I. bullockii* and *I. galbula* were strongly differentiated at only half (4/8) of the loci examined. The two sister species, *I. abeillei* and *I. bullockii*, were only weakly differentiated and only two loci (FGB4 and ALDO5) remained significant after Bonferroni corrections (Table 3).

The multilocus HKA test indicated no significant departures from neutrality in our dataset. The multilocus Tajima's  $D$  test was significant only for *I. bullockii*, a result that is mainly driven by significantly more negative than expected  $D$  values at loci  $\alpha$ -ENO8 and BRM15 (Table 4).  $F_s$  values were negative at the majority of loci, indicative of recent population growth, but this statistic deviated only significantly from neutral expectations at less than half the loci following Bonferroni corrections (Table 4). The  $R_2$  test was only significant at a single locus  $\alpha$ -ENO8 (Table 4).

### Three-population IMA2 analysis

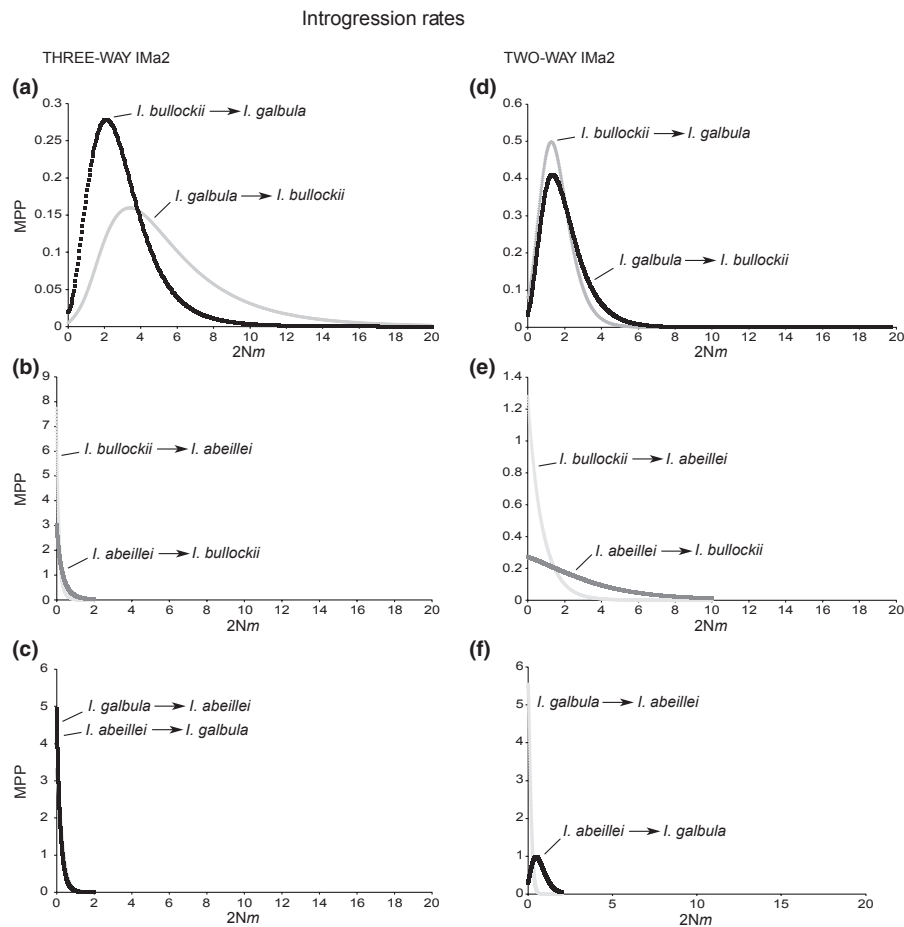
The inclusion of all three members of this species complex in a single IMA2 analysis revealed some interesting patterns of gene introgression (Fig. 4). The two competing population trees (nDNA vs. mtDNA) yielded similar parameter estimates (with one important exception discussed below). We therefore only report the results based on the nuclear population tree (*I. galbula*, (*I. abeillei*, *I. bullockii*)).

First, we found clear evidence of substantial introgression across the hybrid zone between *I. galbula* and *I. bullockii*, ( $2Nm_{I. galbula > I. bullockii} = 3.388$ , 95% HPD = 0.41–12.30, vs.  $2Nm_{I. bullockii > I. galbula} = 2.125$ , 95% HPD = 0.193–6.504, Fig. 4a). Interestingly, the gene flow appears to be slightly skewed in the direction of *I. bullockii*. In contrast, we did not detect any gene flow between the sister species *I. bullockii* (Fig. 4b). The gene flow estimates virtually peaked at zero ( $2Nm = 0.001$  in both directions) and the 95% HPD intervals ranged from 0 to 1.148 and 0 to 0.495 in the direction of *I. bullockii* and *I. abeillei*, respectively. Similarly, the three-population analysis did not detect any gene flow between the highly allopatric *I. galbula* and *I. abeillei* (Fig. 4c), with gene flow estimates peaking near zero and 95% HPD intervals ranging from 0 to 0.631 and 0 to 0.719 in the direction of *I. galbula* and *I. abeillei*, respectively. There was no evidence of historical gene flow between the *I. galbula* lineage and the ancestral *I. abeillei*/*I. bullockii* lineage.

IMA2 provided very concise estimates of the timing of divergence events within the group. The analysis based on the nuclear species tree suggested that the initial divergence between the *I. galbula* lineage and the ancestor of *I. abeillei* and *I. bullockii* appears to have occurred approximately 350,000 years ago (95% HPD = 215,838–568,125), whereas *I. abeillei* and *I. bullockii* split much more recently approximately 95,000 years ago (41,865–159,087, Fig. 5a). In contrast, the analysis based on the

**Table 4.** Locus-specific tests for deviations from selective neutrality and constant population sizes in *Icterus abeillei* (a), *I. bullockii* (b), and *I. galbula* (g). Significance of observed values is indicated at the  $0.05 > P > 0.01$  (\*) and  $P < 0.01$  (\*\*) level. Tests were inapplicable for *I. abeillei* at ALDO5 due to lack of sequence variation. Bolded values remained significant after Bonferroni corrections for multiple testing (adjusted  $\alpha$ -level of 0.00625).

Locus	Tajima's $D$			Fu's $F_s$			Ramos-Onsins & Rozas $R_2$		
	a	b	g	a	b	g	a	b	g
$\beta$ -ACT2	0.048	−0.821	− <b>1.884**</b>	−1.406	−1.227	−1.398	0.14	0.108	0.118
$\alpha$ -ENO8	−1.725*	− <b>1.97**</b>	−1.708*	− <b>6.561**</b>	− <b>7.748**</b>	− <b>7.315**</b>	0.048	<b>0.043**</b>	0.05*
FGB4	−1.412	−0.24	0.711	− <b>4.64**</b>	− <b>11.431**</b>	− <b>6.177**</b>	0.085	0.118	0.152
GAPDH11	0.712	0.152	−0.993	−1.889	−2.089	− <b>8.774**</b>	0.154	0.131	0.078
RDP2	−0.136	0.019	0.474	−1.978	−1.962	0.464	0.114	0.126	0.16
TGF $\beta$ 5	0.024	−0.047	−0.144	0.058	−0.973	− <b>7.268**</b>	0.125	0.13	0.128
ALDO5	—	−0.559	0.434	—	−0.88	1.258	—	0.099	0.157
BRM15	−1.423	−1.559*	1.047	−1.112	− <b>2.963**</b>	1.369	0.061	0.08	0.206



**Figure 4.** Posterior probability plots of bi-directional migration rates (given in effective number of gene copies -  $2Nm$ ) estimated between *Icterus bullockii* and *I. galbula* (a,d), between *I. abeillei* and *I. bullockii* (b,e), and between *I. abeillei* and *I. galbula* (c,f) based on a combined three-population analysis and three separate pairwise analyses in IMA2.

mtDNA population tree (*I. bullockii*, (*I. abeillei*, *I. galbula*)) resulted in perfectly overlapping estimates of the two splitting events around 215–250,000 years ago, with nearly perfectly overlapping HPD intervals (Fig. 5b). Effective population sizes ( $N_e$ ) were approximately 398,000 black-backed orioles, 557,000 Baltimore orioles, and 921,000 Bullock's orioles (Fig. 6).

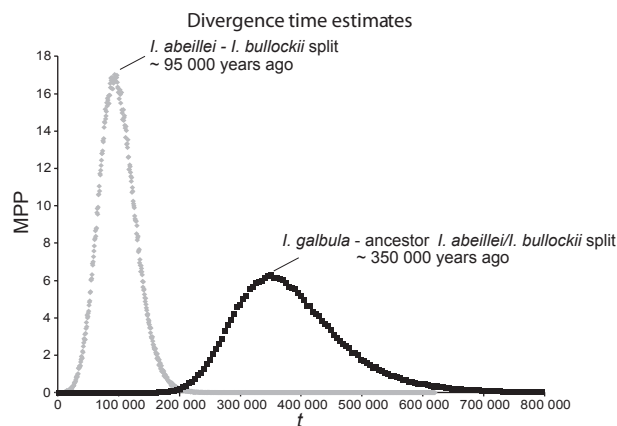
### Two-population IMA2 analyses

A summary of the three separate two-population analyses is presented in Table 3. Estimates of gene flow were mostly consistent between the pairwise two-population analyses and the three-population analysis. However, the two-population analysis of *I. abeillei* and *I. galbula* suggested a small but significant amount of introgression from *I. abeillei* into *I. galbula* ( $2Nm_{I. abeillei > I. galbula} = 0.477$  (Fig. 4f). Although the 95% HPD interval included zero (see Table 5), nested model testing conducted in

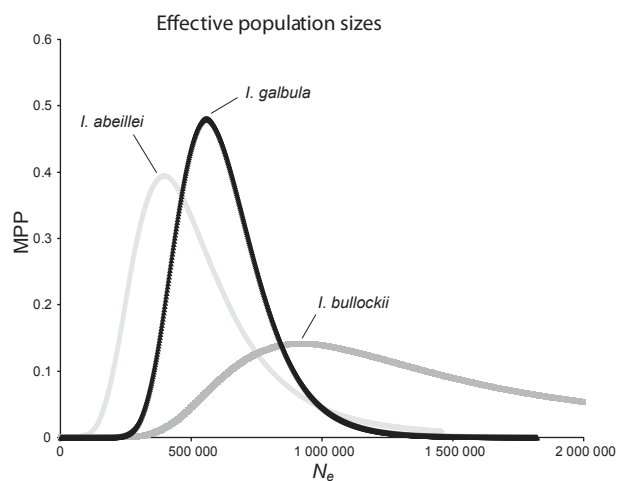
IMA2 indicated that a reduced model with zero migration was a significantly worse fit to the data than the full model allowing for migration (results not shown). Furthermore, the two-population analysis of *I. galbula* and *I. bullockii* indicated lower but equal rates of introgression in both directions ( $2Nm_{I. bullockii > I. galbula} = 1.331$  and  $2Nm_{I. galbula > I. bullockii} = 1.275$ , see Fig. 4d).

The divergence time estimates from the pairwise two-population analyses were comparable to those produced by the three-population IMA2 analysis, except that both divergence events (*I. galbula*–*I. abeillei*/*I. bullockii* split, and *I. abeillei*–*I. bullockii* split) were estimated to have occurred roughly 70,000 years earlier than that indicated by the three-population analysis (see previous section, Fig. 5). In agreement with the nuclear species tree of this group (Jacobsen and Omland 2011), the most recent divergence in this group between *I. abeillei* and *I. bullockii* was estimated at approximately 163,000 years ago, well within the 95% HPD interval of the three-population





**Figure 5.** Posterior probability plot of divergence (splitting) times within the northern oriole group. According to IMA2, the initial split between *Icterus galbula* and the ancestor of *I. abeillei* and *I. bullockii* occurred roughly 350,000 years ago, whereas the more recent split between *I. abeillei* and *I. bullockii* occurred approximately 95,000 years ago.



**Figure 6.** Posterior probability plot of effective population sizes ( $N_e$ ) of extant populations of *I. abeillei*, *I. bullockii*, and *I. galbula*. Peak estimates of  $N_e$  are very similar for all three populations, with highly overlapping confidence intervals (95% HPDs).

estimate (see previous section). Similarly, the older split between the *I. galbula* lineage and the sister taxa *I. abeillei* and *I. bullockii* were estimated at 411,000 and 424,000 years ago, respectively.

Effective population size ( $N_e$ ) estimates from the pairwise two-population analyses were also roughly similar to those estimated by the three-population analysis. The peak estimates were slightly higher in the two-population analyses, but the 95% HPD intervals overlap greatly with the estimates from the three-population analysis. The most striking deviations in population size estimates arose

**Table 5.** Summary of pairwise two-population IMA2 analyses of black-backed oriole (*Icterus abeillei*), Bullock's oriole (*I. bullockii*), and Baltimore oriole (*I. galbula*). Given are highest posterior parameter estimate (HiPt) and lower (HPD95Lo) and upper (HPD95Hi) bounds of 95%HPD intervals for the effective population size of population 1 ( $N_1$ ), population 2 ( $N_2$ ), and ancestral population ( $N_a$ ), splitting time ( $t$ ), and population migration rate from population 2 into population 1 ( $2Nm1 > 2$ ) and from population 1 into population 2 ( $2Nm2 > 1$ ) forward in time. Population sizes are given in 1000 individuals and time since divergence in 1000 years.

	$N_1$	$N_2$	$N_a$	$t$	$2Nm1 > 2$	$2Nm2 > 1$
<i>I. abeillei</i> (1) vs. <i>I. galbula</i> (2)						
HiPt	328	951	169	411	0.001	0.477
HPD95Lo	200	532	66	216	0	0
HPD95Hi	527	1,937	332	750	0.351	1.486
<i>I. abeillei</i> (1) vs. <i>I. bullockii</i> (2)						
HiPt	517	1,198	225	163	0.005	0.025
HPD95Lo	252	528	118	79	0	0
HPD95Hi	1,184	4,732	373	299	2.404	7.461
<i>I. bullockii</i> (1) vs. <i>I. galbula</i> (2)						
HiPt	595	688	163	424	1.275	1.331
HPD95Lo	350	397	62	245	0	0.069
HPD95Hi	1,005	1,215	313	749	3.206	4.25

when the non-sister species, *I. galbula* and *I. abeillei*, were analyzed separately, which resulted in overestimation of  $N_{I. galbula}$  and a slight underestimation of  $N_{I. abeillei}$  (Table 5). Similarly,  $N_{I. bullockii}$  was substantially smaller when paired with the non-sister *I. galbula* than when correctly paired with its sister species *I. abeillei* (Table 5, Fig. 6).

## Discussion

The recent development of a multispecies coalescent framework for the study of divergence and gene flow among three or more taxa simultaneously in the program IMA2 (Hey 2010b) has proven useful for understanding the population dynamics of complex taxonomic groups consisting of several recently diverged species (e.g., Hey 2010a; Illera et al. 2011; Schoville et al. 2011; So et al. 2011; Li et al. 2012). In this study, we were able to estimate levels of nuclear introgression between allopatric populations of the three species in the northern oriole group while also accounting for their recent shared ancestry.

The hybrid zone between the non-sister species, *I. galbula* and *I. bullockii*, represents perhaps one of the most blatant case examples for the need of a multipopulation framework to disentangle the demographic history of larger groups of taxa. The reproductive barriers that so effectively prevent mtDNA and phenotypes of both *I. galbula* and *I. bullockii* from introgressing beyond the

stable hybrid zone appear to be “leaky”, allowing near-neutral nuclear alleles to introgress and successfully spread in a heterospecific genetic background without erasing the diagnostic species differences. Such a substantial rate of introgression evident in populations several hundred miles away from the nearest known areas of sympatry has major implications for our understanding of the population dynamics of all three species within the group. Limited mtDNA introgression (Allen 2002; Carling et al. 2011) but extensive nDNA introgression (Allen 2002; this study) follows the predictions of Haldane’s rule (Haldane 1922), wherein reduced female hybrid fitness prevents maternally inherited mtDNA, but not bi-parentally inherited nDNA from introgressing (Brookfield 1993; Turelli and Orr 1995; Orr 1997). Sex-biased (i.e., male driven) gene flow due to female heterogamety and the effects of Haldane’s rule have increasingly been documented in a wide range of avian groups (e.g., Borge et al. 2005; Carling and Brumfield 2008; Carling et al. 2010; Storchová et al. 2010; Backström and Väli 2011).

A question of interest to researchers studying divergence and gene flow is to determine the timing of migration events between diverging populations. If possible, knowing when gene flow occurred in the speciation process would be of great importance to distinguish between an allopatric speciation model with recent secondary contact following range expansion (gene flow late) and a parapatric speciation model where divergence occurred in the presence of gene flow with subsequent decrease, then cessation of gene flow once reproductive barriers evolved. However, it is now clear that IMA2 and similar methods are unable to distinguish between these different modes of speciation (Sousa et al. 2011; Strasburg and Rieseberg 2011). We therefore cannot infer with confidence whether the currently hybridizing *I. galbula* and *I. bullockii* diverged in allopatry and only recently came back into secondary contact following postglacial range expansion, or if they diverged in parapatry while still exchanging genes throughout the early stages of divergence. Nonetheless, the fact that these two orioles are nearly 5% divergent in mtDNA (Omland et al. 1999) indicates that they have been evolving independently for an extended period of time, consistent with allopatric speciation followed by one or more episodes of gene flow during the last glacial cycles of the Pleistocene epoch. Pliocene or Pleistocene divergence followed by postglacial expansion has been proposed to explain the current distribution of many parapatric species found in North America and Eurasia (Klicka and Zink 1999; Hewitt 2004; Johnson and Cicero 2004). A recent exchange of nuclear variation through introgressive hybridization during secondary contact in the late Pleistocene would explain the wide gap between the mtDNA and nDNA divergence time estimates for these two species,

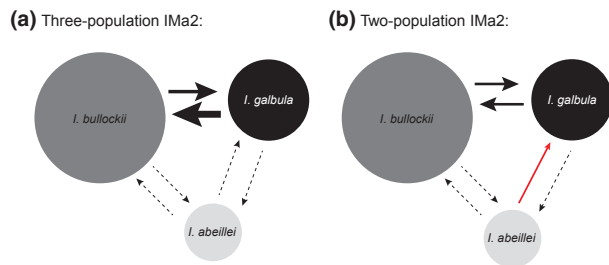
even though the TMRCA for a single locus (e.g., mtDNA) will necessarily always predate the divergence between two species (Edwards and Beerli 2000).

In contrast, the IMA2 analyses revealed no evidence of introgression between *I. abeillei* and *I. bullockii*. The region of parapatry in Durango, Mexico where hybridization between these two species is alleged to have taken place was unfortunately not represented in this study. It is therefore possible that some hybridization could occur within this region, but restricted enough to prevent widespread introgression into the core ranges of either species. Thus, the ubiquitous nuclear haplotype sharing observed between these sister species most likely represents ancestral polymorphism retained in both descendent lineages since their recent split in the late Pleistocene (Hudson and Turelli 2003).

The low rate of unidirectional introgression from *I. abeillei* into *I. galbula* that was indicated by the two-population IMA2 analysis is somewhat puzzling, given that no evidence of gene flow between these two species was revealed in the three-population analysis. Previous studies on these two species using IM based on mtDNA and two nuclear introns also found no evidence of gene flow between these two mitochondrial sister taxa (Kondo et al. 2004, 2008). The fact that they are reciprocally monophyletic at the mtDNA level separated by a single fixed substitution also strongly suggests that there is no recent or ongoing mitochondrial gene flow between these highly allopatric species (Kondo et al. 2008).

Rather, we suspect that this signature of gene flow is a spurious outcome arising when violating the assumption of the IM model that no genetic exchange has occurred between the two focal populations and ghost populations (Beerli 2004; Slatkin 2005). As *I. abeillei* shares a more recent evolutionary history with *I. bullockii* than with *I. galbula* (Jacobsen et al. 2010; Jacobsen and Omland 2011), retained ancestral polymorphisms shared between *I. abeillei* and *I. bullockii* could introgress into *I. galbula* via *I. bullockii* over the extensive hybrid zone. This study revealed just how easily neutral nuclear alleles can cross this species boundary and freely introgress across the entire range of the other species. By excluding *I. bullockii* from the analysis, such shared ancestral alleles could mistakenly be inferred by IMA2 to have introgressed directly from *I. abeillei* into *I. galbula* (Fig. 7). Furthermore, the many alleles introgressing from *I. bullockii* into *I. galbula* have probably contributed to the high nucleotide diversity in *I. galbula*, causing IMA2 to overestimate the current effective population size of *I. galbula* in analyses that do not account for *I. bullockii*.

Thus, exclusion from the analysis of other taxa that have exchanged genes with either of the two focal taxa through shared ancestry or gene flow can bias estimation



**Figure 7.** Patterns of gene flow among the three species in the northern oriole group estimated using IMA2. Circle sizes reflect relative  $N_e$  estimates. Relative rates of gene introgression are indicated by the width of solid arrows. Dashed arrows indicate little or no introgression. (a) The full three-population analysis inferred gene flow only between the hybridizing *Icterus galbula* and *I. bullockii*, whereas (b) two-population analyses also indicated a low level of introgression (highlighted in red) from *I. abeillei* into *I. galbula*. The introgression between these highly allopatric species is best ascribed to ancestral polymorphisms shared between *I. abeillei* and *I. bullockii* introduced to *I. galbula* across the *I. galbula*–*I. bullockii* hybrid zone. This study clearly illustrates how introgression from other taxa not included in the analysis (in this case *I. bullockii*) can bias gene flow estimation if not accounted for.

of population sizes and gene flow in IM, and the severity of the bias depends on how large the genetic contribution from the third taxon has been (Strasburg and Rieseberg 2010).

## Conclusions

Our study reveals the complexity of evolutionary processes (recent divergence, slow sorting of ancestral variation, and introgressive hybridization) that have contributed to produce the observed pattern of rampant allele sharing among the three species within the northern oriole group. Furthermore, our findings highlight the importance of accounting for all potentially interbreeding taxa when interpreting the divergence history of a group of taxa, and the extent to which gene flow may retard the rate of divergence at neutral loci. Thus, this study provides an interesting perspective on the broad-scale impacts of introgressive hybridization, which has important implications for everything from understanding adaptive evolution to inferring species trees. The ease with which genome-scale datasets can now be obtained will soon enable researchers to study even more complex clades with four or more taxa that may interbreed.

## Acknowledgments

We thank the curatorial staff at ANSP, LSUMNS, MZFC, BMNH, FMNH, UWBM, and H. L. Gibbs for providing

the tissue samples used in this study. We especially thank A. G. Navarro-Sigüenza, B. E. Hernández-Baños, and M. Honey for help in collecting in Mexico. We further thank D. Kenny for laboratory assistance, and J. L. Peters for assistance with coalescent analyses. This work was funded by National Science Foundation grants to K. E. O. (DEB-0347083 and DEB-1119506), and Maryland Ornithological Society Avian Research Grants and an American Ornithologist's Union Research Grant to F. J.

## Conflict of Interest

None declared.

## References

- Allen, E. S. 2002. Long-term hybridization and the maintenance of species identity in orioles (*Icterus*). Ph.D. diss. Indiana University, Bloomington.
- AOU. 1973. Thirty-second supplement to the American Ornithologist's Union check-list of North American birds. *Auk* 90:411–419.
- AOU. 1995. Fortieth supplement to the American Ornithologist's Union check-list of North American birds. *Auk* 112:819–830.
- Arnold, M. 2006. *Evolution through genetic change*. Oxford Univ. Press, Oxford, U.K.
- Avise, J. C. 2004. *Molecular markers, natural history, and evolution*. Sinauer Associates, Sunderland, MA.
- Axelsson, E., N. G. C. Smith, H. Sundström, S. Berlin, and H. Ellegren. 2004. Male-biased mutation rate and divergence in autosomal, Z-linked and W-linked introns of chicken and turkey. *Mol. Biol. Evol.* 21:1538–1547.
- Backström, N., and Ü. Väli. 2011. Sex- and species-biased gene flow in a spotted eagle hybrid zone. *BMC Evol. Biol.* 11:100.
- Backström, N., M. Brandström, L. Gustafsson, A. Qvarnström, 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.
- Backström, N., W. Forstmeier, H. Schielzeth, H. Mellenius, K. Nam, E. Bolund, et al. 2010. The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Res.* 20:485–495.
- Bandelt, H. J., P. Forster, and A. Rohl. 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16:37–48.
- Barker, F. K. 2004. Monophyly and relationships of wrens (Aves: Troglodytidae): a congruence analysis of heterogeneous mitochondrial and nuclear DNA sequence data. *Mol. Phyl. Evol.* 31:486–504.
- Barrett, J. C., B. Fry, J. Maller, and M. J. Daly. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* 21:263–265.

- Beerli, P. 2004. Effect of unsampled populations on the estimation of population sizes and migration rates between sampled populations. *Mol. Ecol.* 13:827–836.
- Borge, T., M. T. Webster, G. Andersson, and G.-P. Sætre. 2005. Contrasting patterns of polymorphism and divergence on the Z chromosome and autosomes on two *Ficedula* flycatcher species. *Genetics* 171:1861–1873.
- Brookfield, J. 1993. Haldane's rule is significant. *Evolution* 47:1885–1888.
- Bureš, S., P. Nádvorník, and G.-P. Sætre. 2002. Hybridization and apparent hybridization between meadow pipit (*Anthus pratensis*) and water pipit (*A. spinoletta*). *Hereditas* 136: 254–256.
- Butcher, G. S. 1991. Mate choice in female northern orioles with a consideration for the role of the black male coloration in female choice. *Condor* 93:82–88.
- Carling, M. D., and R. T. Brumfield. 2008. 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., I. J. Lovette, and R. T. Brumfield. 2010. Historical divergence and gene flow: coalescent analyses of mitochondrial, autosomal and sex-linked loci in *Passerina* buntings. *Evolution* 64:1762–1772.
- Carling, M. D., L. G. Serene, and I. J. Lovette. 2011. Using historical DNA to characterize hybridization between Baltimore (Icterus galbula) and Bullock's Orioles. *Auk* 128:61–68.
- Corbin, K. W., C. G. Sibley, and A. Ferguson. 1979. Genic changes associated with the establishment of sympatry in orioles of the genus *Icterus*. *Evolution* 33:624–633.
- Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32:1792–1797.
- Edwards, S. V., and P. Beerli. 2000. Gene divergence, population divergence, and the variance in coalescence time in phylogenetic studies. *Evolution* 54:1839–1854.
- Ellegren, H. 2007. Molecular evolutionary genomics of birds. *Cytogenet. Genome Res.* 117:120–130.
- Excoffier, L., and H. E. L. Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* 10:564–567.
- Flockhart, D. T. T., and K. L. Wiebe. 2009. Absence of reproductive consequences of hybridization in the Northern Flicker (*Colaptes auratus*) hybrid zone. *Auk* 126:351–358.
- Flood, N. J. 1984. Adaptive significance of delayed plumage maturation in male Northern Orioles. *Evolution* 38:267–279.
- Fu, Y.-X. 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. *Genetics* 147:915–925.
- Garrick, R., P. Sunnucks, and R. Dyer. 2010. Nuclear gene phylogeography using PHASE: dealing with unresolved genotypes, lost alleles, and systematic bias in parameter estimation. *BMC Evol. Biol.* 10:118.
- Haldane, J. B. S. 1922. Sex ration and unisexual sterility in animal hybrids. *Genetics* 12:101.
- Hewitt, G. M. 2004. Genetic consequences of climatic oscillations in the Quaternary. *Philos. Trans. R Soc. London Ser. B Biol. Sci.* 359:183–195.
- Hey, J. 2006. Recent advances in assessing gene flow between diverging populations and species. *Curr. Opin. Genet. Dev.* 16:592–596.
- Hey, J. 2010a. The divergence of chimpanzee species and subspecies as revealed in multipopulation isolation-with-migration analyses. *Mol. Biol. Evol.* 27:921–933.
- Hey, J. 2010b. Isolation with migration models for more than two populations. *Mol. Biol. Evol.* 27:905–920.
- Hey, J., and R. Nielsen. 2004. Multilocus methods for estimating population sizes, migration rates and divergence time, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. *Genetics* 167:747–760.
- Hey, J., and R. Nielsen. 2007. Integrating within the Felsenstein equation for improved Markov chain Monte Carlo methods in population genetics. *Proc. Natl. Acad. Sci. USA* 104:2785–2790.
- Hudson, R. R., and M. Turelli. 2003. Stochasticity overrules the “three-times rule”: genetic drift, genetic draft, and coalescence times for nuclear loci versus mitochondrial DNA. *Evolution* 57:182–190.
- Hudson, R. R., M. Kreitman, and M. Aguadé. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153–159.
- Illera, J. C., K. Koivula, J. Broggi, and M. PÄCKert, J. Martens, and L. Kvist. . 2011. A multi-gene approach reveals a complex evolutionary history in the *Cyanistes* species group. *Mol. Ecol.* 20:4123–4139.
- Jacobsen, F., and K. E. Omland. 2011. Species tree inference in a recent radiation of orioles (genus *Icterus*): multiple markers and methods reveal cytonuclear discordance in the northern oriole group. *Mol. Phyl. Evol.* 61:460–469.
- Jacobsen, F., N. R. Friedman, and K. E. Omland. 2010. Congruence between nuclear and mitochondrial DNA: Combination of multiple nuclear introns resolves a well-supported phylogeny of New World orioles (*Icterus*). *Mol Phyl Evol* 56:419–427.
- Jaramillo, A., and P. Burke. 1999. New World blackbirds: the Icterids. Princeton University Press, Princeton, NJ.
- Johnson, N. K., and C. Cicero. 2004. New mitochondrial DNA data affirm the importance of Pleistocene speciation in North American birds. *Evolution* 58:1122–1130.
- Klicka, J., and R. M. Zink. 1999. Pleistocene effects on North American songbird evolution. *Proc. R Soc. London Ser. B Biol. Sci.* 266:695–700.
- Kliman, R. M., P. Andolfatto, J. A. Coyne, F. Depaulis, M. Kreitman, A. J. Berry, et al. 2000. The population genetics



- of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* 156:1913–1931.
- Kondo, B., J. M. Baker, and K. E. Omland. 2004. Recent speciation between the Baltimore Oriole and the Black-backed Oriole. *Condor* 106:674–680.
- Kondo, B., J. L. Peters, B. B. Rosensteel, and K. E. Omland. 2008. Coalescent analyses of multiple loci support a new route to speciation in birds. *Evolution* 62:1182–1191.
- Li, Z., J. Zou, K. Mao, K. Lin, H. Li, J. Liu, et al. 2012. Population genetic evidence for complex evolutionary histories of four high altitude juniper species in the Qinghai-Tibetan Plateau. *Evolution* 66:831–845.
- Librado, P., and J. Rozas. 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452.
- Machado, C. A., R. M. Kliman, J. A. Markert, and J. Hey. 2002. Inferring the history of speciation from multilocus dna sequence data: the case of *Drosophila pseudoobscura* and close relatives. *Mol. Biol. Evol.* 19:472–488.
- Mettler, R. D., and G. M. Spellman. 2009. A hybrid zone revisited: molecular and morphological analysis of the maintenance, movement, and evolution of a Great Plains avian (Cardinalidae: *Pheucticus*) hybrid zone. *Mol. Ecol.* 18:3256–3267.
- Miller, W. D. 1906. Lists of birds collected in northwestern Durango, Mexico, by F. H. Batty, during 1903. *Bull. Am. Mus. Nat. Hist.* 22:161–183.
- Milne, I., F. Wright, G. Rowe, D. F. Marshall, D. Husmeier, and G. McGuire. 2004. TOPALi: software for automatic identification of recombinant sequences within DNA multiple alignments. *Bioinformatics* 20:1806–1807.
- Moore, W. S. 1995. Inferring phylogenies from mtDNA variation: mitochondrial-gene trees versus nuclear-gene trees. *Evolution* 49:718–726.
- Nielsen, R., and J. Wakeley. 2001. Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* 158:885–896.
- Nosil, P. 2008. Speciation with gene flow could be common. *Mol. Ecol.* 17:2103–2106.
- Omland, K. E., S. M. Lanyon, and S. J. Fritz. 1999. A molecular phylogeny of the New World orioles (*Icterus*): the importance of dense taxon sampling. *Mol. Phylog. Evol.* 12:224–239.
- Orr, H. A. 1997. Haldane's rule. *Annu. Rev. Ecol. Syst.* 28:195–218.
- Peters, J. L., Y. Zhuravlev, I. Fefelov, A. Logie, and K. E. Omland. 2007. Nuclear loci and coalescent methods support ancient hybridization as cause of mitochondrial paralogy between gadwall and falcated duck (*Anas* spp.). *Evolution* 61:1992–2006.
- Petit, R. J., and L. Excoffier. 2009. Gene flow and species delimitation. *Trends Ecol. Evol.* 24:386–393.
- Pinho, C., and J. Hey. 2010. Divergence with gene flow: models and data. *Annu. Rev. Ecol. Syst.* 41:215–230.
- Primmer, C. R., T. Borge, J. Lindell, and G.-P. Sætre. 2002. Single-nucleotide polymorphism characterization in species with limited available sequence information: high nucleotide diversity revealed in the avian genome. *Mol. Ecol.* 11:603–612.
- Ramos-Onsins, S. E., and J. Rozas. 2002. Statistical properties of new neutrality tests against population growth. *Mol. Biol. Evol.* 19:2092–2100.
- Remington, C. L. 1968. Suture-zones of hybrid interaction between recently joined biotas. Pp. 321–428 in T. Dobzhansky, M. K. Hecht and W. C. Steere, eds. *Evolutionary biology*. Plenum Press, New York.
- Richardson, D. S., and T. Burke. 1999. Extra-pair paternity in relation to male age in Bullock's oriole. *Mol. Ecol.* 8:2115–2126.
- Ridgely, R. S., T. F. Allnutt, T. Brooks, D. K. McNicol, D. W. Mehlman, B. E. Young, et al. 2007. Digital distribution maps of the birds of the Western Hemisphere, version 3.0. NatureServe, Arlington, VA.
- Rising, J. D. 1969. A comparison of metabolism and evaporative water loss of Baltimore and bullock orioles. *Comp. Biochem. Physiol.* 31:915–925.
- Rising, J. D. 1970. Morphological variation and evolution in some North American orioles. *Syst. Zool.* 19:315–351.
- Rising, J. D. 1973. Morphological variation and status of the orioles, *Icterus galbula*, *I. bullockii*, and *I. abeillei*, in the northern Great Plains and in Durango, Mexico. *Can. J. Zool.* 51:1267–1273.
- Rising, J. D. 1983. The progress of Oriole hybridization in Kansas. *Auk* 100:885–897.
- Rising, J. D. 1996. The stability of the oriole hybrid zone in Western Kansas. *Condor* 98:658–663.
- Rising, J. D., and N. J. Flood. 1998. Baltimore Oriole (*Icterus galbula*). in A. Poole, ed. *The Birds of North America*. Cornell Lab of Ornithology, Ithaca, NY.
- Rising, J. D., and P. L. Williams. 1999. Bullock's Oriole (*Icterus bullockii*) in A. Poole, ed. *The Birds of North America*. Cornell Lab of Ornithology, Ithaca, NY.
- Rohwer, S., and M. S. Johnson. 1992. Scheduling differences of molt and migration for Baltimore and Bullock's Orioles persist in a common environment. *Condor* 94:992–994.
- Rohwer, S., and J. Manning. 1990. Differences in timing and number of molts for Baltimore and Bullock's orioles: implications to hybrid fitness and theories of delayed plumage maturation. *Condor* 92:125–140.
- Sæther, B.-E., R. Lande, S. Engen, H. Weimerskirch, M. Lillegard, R. Altwegg, et al. 2005. Generation time and temporal scaling of bird population dynamics. *Nature* 436:99–102.
- Schoville, S. D., T. S. Tustall, V. T. Vredenburg, A. R. Backlin, E. Gallegos, D. A. Wood, et al. 2011. Conservation genetics of evolutionary lineages of the endangered mountain yellow-legged frog, *Rana muscosa* (Amphibia: Ranidae), in southern California. *Biol. Conserv.* 144:2031–2040.



- Sibley, C. G., and L. L. J. Short. 1964. Hybridization in the Orioles of the Great Plains. *Condor* 66:130–150.
- Slatkin, M. 2005. Seeing ghosts: the effect of unsampled populations on migration rates estimated for sampled populations. *Mol. Ecol.* 14:67–73.
- So, J., S. Uthicke, J.-F. Hamel, and A. Mercier. 2011. Genetic population structure in a commercial marine invertebrate with long-lived lecithotrophic larvae: *Cucumaria frondosa* (Echinodermata: Holothuroidea). *Mar. Biol.* 158:859–870.
- Sousa, V. C., A. Grelaud, and J. Hey. 2011. On the nonidentifiability of migration time estimates in isolation with migration models. *Mol. Ecol.* 20: 3956–3962.
- Stephens, M., and P. Scheet. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am. J. Hum. Genet.* 76:449–462.
- Stephens, M., N. J. Smith, and P. Donnelly. 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68:978–989.
- Storchová, R., J. Reif, and M. W. Nachman. 2010. Female heterogamety and speciation: reduced introgression of the Z chromosome between two species of nightingales. *Evolution* 64:456–471.
- Strasburg, J. L., and L. H. Rieseberg. 2010. How robust are “isolation with migration” analyses to violations of the IM model? a simulation study. *Mol. Biol. Evol.* 27:297–310.
- Strasburg, J. L., and L. H. Rieseberg. 2011. Interpreting the estimated timing of migration events between hybridizing species. *Mol. Ecol.* 20:2353–2366.
- Sutton, G. M. 1938. Oddly plumaged orioles from western Oklahoma. *Auk* 55:1–6.
- Sutton, G. M. 1968. Oriole hybridization in Oklahoma. *Bull. Ok Ornith Soc* 1:1–7.
- Tajima, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* 28:2731–2739.
- Turelli, M., and H. A. Orr. 1995. The dominance theory of Haldane’s rule. *Genetics* 140:389–402.
- Waltari, E., and S. V. Edwards. 2002. Evolutionary dynamics of intron size, genome size, and physiological correlates in Archosaurs. *Am. Nat.* 160:539–552.

**Appendix 1.** Taxonomic sampling

Species	Museum <sup>a</sup>	ID	Sex	Collection Locality <sup>b</sup>
<i>Icterus abeillei</i>	MZFC	2339	M	Mexico, Morelos, Carretera
<i>Icterus abeillei</i>	MZFC	5381	M	Mexico, Guerrero, Apaxtla
<i>Icterus abeillei</i>	MZFC	3468	M	Mexico, Guanajuato, San Pedro
<i>Icterus abeillei</i>	MZFC	KEO27	F	Mexico: Guanajuato, Salvatierra
<i>Icterus abeillei</i>	MZFC	KEO28	M	Mexico: Guanajuato, Salvatierra
<i>Icterus abeillei</i>	MZFC	KEO29	M	Mexico: Guanajuato, Salvatierra
<i>Icterus abeillei</i>	MZFC	1087	M	Mexico, Guanajuato, Yuriria
<i>Icterus abeillei</i>	MZFC	5441	M	Mexico, Guanajuato, Yuriria
<i>Icterus abeillei</i>	MZFC	KEO30	M	Mexico: Guanajuato, Yuriria
<i>Icterus abeillei</i>	MZFC	KEO31	M	Mexico: Guanajuato, Yuriria
<i>Icterus abeillei</i>	MZFC	3467	M	Mexico, Jalisco, Laguna de Chapala
<i>Icterus abeillei</i>	MZFC	7966	M	Mexico: Jalisco
<i>Icterus abeillei</i>	MZFC	10733	M	Mexico: Jalisco
<i>Icterus abeillei</i>	MZFC	11483	M	Mexico: Jalisco
<i>Icterus abeillei</i>	MZFC	KEO32	M	Mexico: Jalisco, Chapala
<i>Icterus abeillei</i>	MZFC	KEO33	M	Mexico: Jalisco, Chapala
<i>Icterus abeillei</i>	MZFC	KEO34	M	Mexico: Jalisco, Chapala
<i>Icterus abeillei</i>	MZFC	KEO35	F	Mexico: Jalisco, Chapala
<i>Icterus abeillei</i>	MZFC	KEO37	F	Mexico: Jalisco, Chapala
<i>Icterus abeillei</i>	MZFC	KEO45	M	Mexico: Michoacán, Briseñas
<i>Icterus abeillei</i>	MZFC	KEO46	M	Mexico: Michoacán, Briseñas
<i>Icterus abeillei</i>	MZFC	KEO48	F	Mexico: Michoacán, Briseñas
<i>Icterus abeillei</i>	MZFC	KEO49	M	Mexico: Michoacán, Briseñas
<i>Icterus bullockii</i>	MVZ	176515	M	USA: NM, Mora Co.
<i>Icterus bullockii</i>	LSUMNS	3874	F	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3875	F	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3876	M	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3877	F	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3878	M	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3879	M	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	LSUMNS	3880	M	USA: CA, Inyo Co.
<i>Icterus bullockii</i>	FMNH	330029	M	USA: CA, San Bernardino Co.
<i>Icterus bullockii</i>	FMNH	341933	M	USA: CA, Imperial Co.
<i>Icterus bullockii</i>	FMNH	341934	M	USA: CA, Monterrey Co.
<i>Icterus bullockii</i>	FMNH	341937	M	USA: CA, Monterrey Co.
<i>Icterus bullockii</i>	FMNH	341938	M	USA: CA, Monterrey Co.
<i>Icterus bullockii</i>	UWBM	55942	M	USA: WA, Grant Co.
<i>Icterus bullockii</i>	UWBM	55951	F	USA: WA, Douglas Co.
<i>Icterus bullockii</i>	UWBM	55975	M	USA: WA, Douglas Co.
<i>Icterus bullockii</i>	UWBM	55976	F	USA: WA, Douglas Co.
<i>Icterus bullockii</i>	UWBM	55978	F	USA: WA, Douglas Co.
<i>Icterus bullockii</i>	UWBM	59055	F	USA: WA, Asotin Co.
<i>Icterus bullockii</i>	UWBM	59056	M	USA: WA, Asotin Co.
<i>Icterus bullockii</i>	UWBM	59058	M	USA: WA, Asotin Co.
<i>Icterus galbula</i>	LSUMNS	131146	M	USA: LA, Cameron Co.
<i>Icterus galbula</i>	ANSP	10126	M	USA: PA, Bucks Co.
<i>Icterus galbula</i>	ANSP	10134	M	USA: PA, Bucks Co.
<i>Icterus galbula</i>	ANSP	10148	F	USA: PA, Bucks Co.
<i>Icterus galbula</i>	USFWS	17860	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17861	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17862	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17864	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17866	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17867	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	USFWS	17869	M	USA: MD, Baltimore Co.

(Continued)

**Appendix 1.** (Continued).

Species	Museum <sup>a</sup>	ID	Sex	Collection Locality <sup>b</sup>
<i>Icterus galbula</i>	USFWS	17870	M	USA: MD, Baltimore Co.
<i>Icterus galbula</i>	UWBM	67201	M	USA: MD, Prince Georges Co.
<i>Icterus galbula</i>	BMNH	42547	M	USA: MN, Becker Co.
<i>Icterus galbula</i>	BMNH	X7630	M	USA: MN, Hennepin Co.
<i>Icterus galbula</i>	BMNH	X7763	F	USA: MN, Chisago Co.
<i>Icterus galbula</i>	BMNH	X7799	M	USA: MN, Chisago Co.
<i>Icterus galbula</i>	FMNH	350604	M	USA: IL, Cook Co.
<i>Icterus galbula</i>	FMNH	395866	F	USA: IL, Cook Co.
<i>Icterus galbula</i>	HLG	LG03	M	Canada: Manitoba
<i>Icterus galbula</i>	HLG	LG04	M	Canada: Manitoba
<i>Icterus galbula</i>	HLG	LG05	M	Canada: Manitoba
<i>Icterus galbula</i>	HLG	LG06	F	Canada: Manitoba
<i>Sturnella neglecta</i>	NMNH	586115		

<sup>a</sup>Museums are abbreviated as follows: ANSP = Academy of Natural Sciences, Philadelphia; BMNH = J. F. Bell Museum of Natural History, University of Minnesota; FMNH = Field Museum of Natural History; HLG = H. Lisle Gibbs, blood samples; LSUMNS = Louisiana State University, Museum of Natural Science; MVZ = Museum of Vertebrate Zoology, University of California, Berkeley; MZFC = Museo de Zoología, Facultad de Ciencias, Universidad Nacional Autónoma de México; NMNH = National Museum of Natural History; UWBM = University of Washington, Burke Museum and USFWS = United States Fish and Wildlife Service band numbers.

<sup>b</sup>San Pedro is the short form of San Pedro de Los Naranjos, and Briseñas is the short form of Briseñas de Matamoros.

**Appendix 2.** GenBank accession numbers for DNA sequences obtained from previous studies on *Icterus* (Kondo et al. 2008; Jacobsen et al. 2010; Jacobsen and Omland 2011).

Species	GenBank Accession Nos.
<i>I. abeillei</i>	DQ898412-DQ898433, DQ898358-DQ898385, GU972843, GU972875, JN169143-JN169144, JN169230-JN169239, JN169317-JN169326, JN169404-JN169413, JN169491-JN169500, JN169578-JN169587, JN169665-JN169674
<i>I. bullockii</i>	JN169159-JN169170, JN169246-JN169257, JN169333-JN169344, JN169422-JN169431, JN169507-JN169518, JN169594-JN169605, JN169681-JN169692
<i>I. galbula</i>	DQ898316-DQ898331, DQ898344-DQ898347, DQ898356-DQ898357, DQ898386-DQ898392, DQ898407, GU972855, GU972887, JN169191- JN169196, JN169274-JN169283, JN169361-JN169370, JN169448-JN169457, JN169535-JN169544, JN169622-JN169631, JN169709- JN169718
<i>S. neglecta</i>	GU972873, GU972905