Incomplete reproductive isolation following host shift in brood parasitic indigobirds

Christopher N. Balakrishnan^{1,*}, Kristina M. Sefc² and Michael D. Sorenson¹

¹Department of Biology, Boston University, 5 Cummington Street, Boston, MA 02215, USA

²Department of Zoology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria

Behavioural and molecular studies suggest that brood parasitic indigobirds (*Vidua* spp.) rapidly diversified through a process of speciation by host shift. However, behavioural imprinting on host song, the key mechanism promoting speciation in this system, may also lead to hybridization and gene flow among established indigobird species when and if female indigobirds parasitize hosts already associated with other indigobird species. It is therefore not clear to what extent the low level of genetic differentiation among indigobird species is due to recent common ancestry versus ongoing gene flow. We tested for reproductive isolation among three indigobird species in Cameroon, one of which comprises two morphologically indistinguishable host races. Mimicry of host songs corresponded with plumage colour in 184 male indigobirds, suggesting that females rarely parasitize the host of another indigobird species. Paternity analyses, however, suggest that imperfect specificity in host and/or mate choice allows for continuing gene flow between recently formed host races of the Cameroon Indigobird *Vidua camerunensis*; while 63 pairs of close relatives were associated with the same host, two strongly supported father–son pairs included males mimicking the songs of the two different hosts of *V. camerunensis*. Thus, complete reproductive isolation is not necessarily an automatic consequence of host shifts, a result that suggests an important role for natural and/or sexual selection in indigobird speciation.

Keywords: brood parasitism; divergence with gene flow; parentage analysis; sympatric speciation; *Vidua*

1. INTRODUCTION

Populations in the initial stages of speciation are of particular interest because they provide an opportunity to observe the processes contributing to adaptive diversification and the evolution of reproductive isolation (reviewed in Edwards *et al.* 2005; Price 2007). In sympatric speciation by host shift, the colonization of a novel host can produce immediate reproductive isolation if cues involved in mate choice are acquired from the host species (e.g. Payne *et al.* 1998; Linn *et al.* 2003). Following a host shift, we may therefore expect strong premating reproductive isolation even in the absence of genetic differentiation (Via 2001). An observed lack of genetic differentiation is difficult to interpret, however, because it could be the result of recent common ancestry, ongoing gene flow or both (e.g. Nielsen & Wakeley 2001; Bulgin *et al.* 2003; Palsbøll *et al.* 2004). Evaluating whether and to what extent gene flow continues between diverging populations is important because the level of gene flow determines the strength of selection required to fix alternative adaptive traits in diverging populations (e.g. Wright 1931; Slatkin 1987; Morjan & Rieseberg 2004).

Indigobirds (genus *Vidua*) are obligate, host-specific brood parasites of finches (genus *Lagonosticta*) and other estrildid finches (family Estrildidae). Phylogenetic

analyses suggest that 10 indigobird species rapidly diversified via the colonization of new hosts (Klein & Payne 1998; Sorenson *et al.* 2003, 2004). Following host colonization, song learning and the mimicry of host song by male indigobirds leads to assortative mating between individuals reared by the novel host (Payne *et al.* 1998, 2000). Premating reproductive isolation is therefore possible in the first generation following host colonization. This same mechanism, however, may also lead to hybridization if females occasionally lay in the nests of host species that are already parasitized by other indigobird species (figure 1a; see also Payne & Sorenson 2004). Male progeny produced from such ‘errors’ in host choice will learn the songs of another indigobird species (Payne *et al.* 1993), and thus should attract heterospecific females as mates. Likewise, because female indigobirds also imprint on host song (Payne *et al.* 2000), female progeny resulting from mislaid eggs should be attracted to males of another indigobird species. Incomplete genetic differentiation among indigobird species may therefore be a result of imperfect fidelity in either host selection or mate choice (figure 1b), both of which will promote interspecific gene flow, or it may simply reflect recent common ancestry (Sorenson *et al.* 2003; Sefc *et al.* 2005).

Despite their genetic similarity and a plausible behavioural mechanism promoting hybridization, sympatric indigobird species are generally distinct in male plumage colour and have different, culturally inherited non-mimicry songs (Payne 1973, 1982, 1996; Payne *et al.* 1993; Balakrishnan & Sorenson 2006, in preparation). In addition, many indigobird species have evolved nestling mouth markings that match those of their respective hosts

* Author and address for correspondence: Institute for Genomic Biology 2500N, University of Illinois, Urbana-Champaign, 1206 West Gregory Drive MC-195, Urbana, IL 61801, USA (cbala@igb.uiuc.edu).

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspb.2008.0733> or via <http://journals.royalsociety.org>.

Figure 1. Behavioural mechanisms potentially contributing to gene flow between indigobird species or host races. (a) Gene flow via imperfect host fidelity or 'error' in host choice. In this scenario, a female indigobird parasitizes two different host species. A male offspring raised by the alternate host will acquire songs from the alternate host but will have the plumage colour of his parents' species, as hybridization has not yet occurred. Thus, the male will differ from his father in mimicry song but not in plumage colour. Likewise, full siblings will share the same plumage colour, but differ in song mimicry. Assuming that the mis-imprinted male attracts heterospecific females with his song, hybridization will probably occur in the next generation. (b) Gene flow via 'error' in mate choice. In this scenario, a female mates with a male reared by a different host species but then parasitizes her own host species. Fathers and sons will again sing different mimicry songs. Depending on the genetics of plumage colour, the hybrid son may be morphologically intermediate.

(Payne 2005), thereby enabling indigobird young to better compete for parental care in host nests (Payne *et al.* 2001; Schuetz 2005). If hybridization among species occurs at an appreciable rate, this would suggest that diversifying selection on morphology has been strong enough to counteract the homogenizing influence of gene flow. Given the difficulty of distinguishing between gene flow and recent common ancestry as explanations for the genetic similarity of populations, we used genetic parentage analyses to directly test for potential gene flow among indigobird species and host races.

Three indigobird species occur syntopically near Tibati, Cameroon: the Cameroon indigobird *Vidua camerunensis*; the jambandu indigobird *Vidua raricola*; and Wilson's indigobird *Vidua wilsoni*. *Vidua camerunensis* is of particular interest because this species is associated with two different hosts at Tibati and elsewhere in Cameroon: African firefinch *Lagonosticta rara* and black-bellied firefinch *Lagonosticta rubricata*. Across West Africa, *V. camerunensis* is also known to parasitize Dybowski's twinspot *Euschistospiza dybowskii* and brown twinspot *Clytospiza monteiri* (Payne *et al.* 2005). Populations associated with each of these hosts fit the definition of host races (Bush 1969), presumably formed as a result of recent host colonization (see Payne *et al.* 2002, 2005). These host races appear to be behaviourally isolated—mimicking the songs of their respective hosts (Payne *et al.* 2005), singing distinct repertoires of non-mimicry songs (Balakrishnan & Sorenson *in preparation*) and responding differentially to playback of each other's songs (Balakrishnan & Sorenson 2006)—but they are morphologically indistinguishable (Payne *et al.* 2005) and show no

significant differentiation in neutral genetic markers (Sefc *et al.* 2005; Balakrishnan & Sorenson *in preparation*). These host races therefore allow a test for reproductive isolation at the earliest stages of speciation. At the same time, we test for potential gene flow between *V. camerunensis* and two other morphologically distinct indigobird species, *V. raricola* and *V. wilsoni*, which parasitize the zebra waxbill *Amandava subflava* and bar-breasted firefinch *Lagonosticta rufopicta*, respectively.

2. MATERIAL AND METHODS

(a) Field methods

Fieldwork was conducted near Tibati, Cameroon (6°28' N, 13°34' E) in October and November 2000 and from October to December in 2001–2003. The end of the rainy season in October coincides with the start of the breeding season for indigobirds and their hosts in Cameroon.

The indigobird mating system has been described as a highly polygynous, dispersed lek (Payne & Payne 1977). Male indigobirds defend exclusive territories during the breeding season and devote most of their time to singing from a 'call site' at or near the top of a tree within the territory (Payne 1973). Territorial males were recorded using a Sony TCDM-5 cassette recorder and a Sennheiser ME-66 shotgun microphone. Recording continued until approximately 15–20 min of song had been acquired, after which males were captured using playback of their own song from behind a mist net. Females and/or young males in brown plumage were occasionally captured while interacting with a singing male. A few indigobirds were also captured opportunistically by placing nets in wet areas where finches were observed to drink.

Indigobird song repertoires are complex, including both non-mimetic indigobird songs and mimicry of host songs and calls (Payne 1973). Mimetic songs are easily identified in the field and provide direct evidence of the host species that reared each male indigobird. The identity of mimicry songs was confirmed by comparing sonograms generated in AVISOFT SAS LAB PRO (AviSoft Bioacoustics) with published sonograms of indigobirds and their hosts (Payne 1973; Payne & Payne 1994; Payne *et al.* 2005). Female indigobirds do not sing and females of the three species at Tibati are morphologically similar; thus, host association and species identity of females can only be inferred through observations of social interactions with males.

A total of 265 birds were captured, of which 140 were sacrificed via thoracic compression and prepared as voucher specimens. Animal protocols were approved by Boston University's Institutional Animal Care and Use Committee and followed the guidelines of the Ornithological Council (Gaunt *et al.* 1997). Vouchers allow for rigorous morphometric and plumage colour measurements, important for quantifying the subtle differences among indigobird species and testing for the possibility of hybrid individuals with intermediate plumage colour. All other birds were banded and released with a unique combination of two coloured plastic leg bands and one numbered metal band. Recapture and resighting data from banded birds allowed a qualitative assessment of survival and philopatry between years. We collected two feathers and a blood sample from each captured bird. Blood samples and muscle tissue (from birds prepared as voucher specimens) were stored in lysis buffer and a dimethyl sulfoxide storage buffer, respectively (Seutin *et al.* 1991). Voucher specimens have been deposited in the University of Michigan Museum of Zoology.

Males removed from call sites were generally replaced by a new male of the same species (or the same host race), often within minutes. While removal of males probably reduced the reproductive success of individual males and therefore the likelihood of capturing their offspring in subsequent years, the opportunity to record and capture replacement males greatly increased our ability to sample the adult male population. As removal of territorial males did not change the number or species identity of active male call sites, we reasoned that this perturbation to the system would not have significant effects on patterns of mate choice by females.

(b) Plumage colour measurements

We measured plumage colour of the 140 voucher specimens using an Ocean Optics USB2000 spectrometer. As breeding male indigobirds are uniform in plumage colour, average values of hue, brightness, contrast and UV brightness were calculated for each bird from 15 spot readings following the approach of Doucet & Montgomerie (2003; details in Balakrishnan & Sorenson *in preparation*). The colour of birds that were banded and released was categorized as 'blue', 'green' or 'purple' (corresponding to the three morphologically distinct species at Tibati) and photo vouchers were taken of all captured birds.

(c) Genetic analyses

DNA was extracted from 25 mg of muscle tissue or the calami of two feathers using a QIAamp tissue kit (Qiagen). Three milligrams of dithiothreitol were added to the extraction buffer to aid in the lysis of feather samples. Each of 265 indigobirds was genotyped for 22 nuclear loci. The nuclear dataset comprised 11 microsatellite loci, six single nucleotide

polymorphism (SNP) loci and five DNA sequence loci (see the electronic supplementary material for primer and probe sequences). Amplification and scoring of microsatellite loci were as described by Sefc *et al.* (2001). SNP loci were identified by sequencing nuclear introns for a panel of three individuals representing the three species in this study. SNP genotyping used TaqMan probes (Applied Biosystems) designed to interrogate one polymorphic site in each intron and an ABI 7900HT real-time PCR machine following the manufacturer's instructions. The five nuclear sequence loci included four anonymous loci obtained from an indigobird genomic library and intron 3 of the lecithin cholesterol acyltransferase (LCAT) gene (table 1). Direct sequences of PCR products from heterozygous individuals were computationally resolved into haplotypes using PHASE v. 2.1.1 (Stephens *et al.* 2001). Haplotypes for nuclear sequence loci were generally resolved with high confidence (more than 95%), although the phase of polymorphisms unique to a single individual ($n=4$ sites in three loci) and a few other sites showing evidence of recombination or homoplasy ($n=3$ sites in three loci) were uncertain. We excluded these problematic sites, thereby collapsing rare alleles into more common ones. This approach, which affected only a small number of individuals, prevented the exclusion of potential parents on the basis of incorrectly inferred haplotypes.

We collected data from three additional loci for selected individuals (see §3 and the electronic supplementary material, table 3) using similar methods; these included one additional anonymous sequence locus and the flanking regions of two microsatellite loci, the latter scored as HapSTR loci (Hey *et al.* 2004), as described by Sorenson & DaCosta (*in preparation*).

Finally, for each individual sampled, we sequenced a 1100 bp region of the maternally inherited mitochondrial DNA comprising most of the ND6 gene, tRNA glutamine and the 5' half of the control region, using primers and protocols described in Sorenson & Payne (2001).

(d) Parentage analyses

Parentage analyses of brood parasites are challenging because putative parents and offspring cannot be sampled at the nest. Given no *a priori* knowledge about relationships, all 265 individuals were included as both potential parents and potential offspring in a single analysis. We used CERVUS v. 3.0 (Marshall *et al.* 1998; Kalinowski *et al.* 2007), which uses a likelihood approach to identify the most likely parents for each individual. Given that nearly all individuals in our dataset with known host association were males, we used CERVUS to find the single most likely father of each individual, treating the maternal genotype as unknown. Confidence in these assignments is determined through analyses of simulated genotypes in which the true parent is known. We systematically varied two parameters in the simulations to explore the extent to which our results were dependent on the settings used. First, we assumed that there were 500 candidate fathers in the population and that our sample included either 10, 20 or 30 per cent of these individuals. Second, we varied the genotyping error rate, using values of 1, 0.1 or 0 per cent. Our estimates for population size and the proportion of males sampled were based on observations of approximately 70 male territories each year plus numerous subordinate males, including those replacing territorial males that were removed, and the fact that parents of birds caught in the first year of the study and/or around the perimeter of the study area were less likely to be sampled.

Table 1. Descriptive statistics for the nuclear loci used in this study. (N , number of individuals genotyped out of 265; K , number of alleles; H , heterozygosity. F_{ST} estimates and statistical significance ($*p < 0.05$; $**p < 0.01$) for each locus and across all loci are given for a comparison across the three morphologically distinct species ($n = 3$) and for *V. camerunensis* host races only ($n = 2$). The non-exclusion probability (1-exclusion probability) is the probability that an unrelated individual will be included as a possible parent or offspring (Marshall *et al.* 1998).

	N	K	H	F_{ST} all species	F_{ST} host races	non-exclusion probability
<i>microsatellite loci</i>						
Indigo7	264	12	0.742	0.010	0.007	0.675
Indigo8	263	11	0.821	0.007	0.004	0.496
Indigo15	265	8	0.758	0.011*	−0.002	0.659
Indigo27	258	14	0.733	0.016*	0.002	0.519
Indigo28	265	12	0.819	0.014**	−0.003	0.557
Indigo29	262	12	0.893	0.015**	−0.007	0.428
Indigo30	265	11	0.845	0.009*	−0.002	0.525
Indigo37	264	22	0.837	0.029**	0.004	0.391
Indigo38	265	13	0.811	0.037**	−0.002	0.432
Indigo40	263	8	0.726	0.018*	−0.004	0.615
Indigo41	265	8	0.774	0.011	−0.006	0.615
<i>SNP loci</i>						
GAPDH	255	2	0.478	0.022*	−0.006	0.875
LDH-B	258	2	0.031	0.053**	−0.009	1.000
PEPCK9	258	2	0.453	0.016	−0.007	0.896
RP40	251	2	0.231	0.053*	−0.009	0.964
TGF2-B	262	2	0.046	0.047*	−0.009	0.999
Trop	258	2	0.527	0.006	0.019	0.875
<i>sequence loci</i>						
B103	262	13	0.569	0.015	−0.004	0.789
P30FD2	265	10	0.672	0.074**	−0.006	0.718
C125	265	9	0.683	0.002	0.006	0.760
P31RD5	264	15	0.803	0.020**	0.003	0.557
LCAT	258	8	0.725	0.011*	0.007	0.694
				overall F_{ST} : 0.019**	overall F_{ST} : 0.000	combined non-exclusion probability: 0.000101

CERVUS was also used to calculate non-exclusion probabilities and to test for deviations from the Hardy–Weinberg equilibrium. To assess genetic substructure among species and host races, we estimated F_{ST} for each nuclear locus and across all loci using ARLEQUIN v. 3.01 (Excoffier *et al.* 2005).

The inclusion of both parents and offspring as candidate parents violates an assumption of the model implemented in CERVUS. The primary consequence, however, is that siblings may be identified as parent–offspring (PO) pairs. Thus, significance values returned by CERVUS should not be interpreted literally with respect to paternity (T. C. Marshall 2006, personal communication). This is not a serious problem for our analysis because both sibling pairs and PO pairs are useful for testing the general prediction that first-degree male relatives are always associated with the same host (see below). To test whether PO pairs found by CERVUS might actually be full-sibs and to identify additional pairs of close relatives, we also used ML-RELATE (Kalinowski *et al.* 2006). ML-RELATE determines the most likely relationship (i.e. PO; full siblings, FS; half siblings, HS; or unrelated) for every pair of individuals in the dataset. For each pair of individuals for which the most likely relationship was PO, FS or HS, we conducted a likelihood ratio test comparing the identified relationship against the null hypothesis of ‘unrelated’. This approach allowed us to apply a more stringent threshold for rejection of the null hypothesis and thereby avoid false positives (see §3 for additional details). Significant FS and HS pairs were also tested against mtDNA data, with the expectation that full-sibs or maternal half-sibs should share the same haplotype.

To further assess the risk of type I and II errors, we calculated relatedness coefficients based on 22 nuclear loci for all pairs of birds using KINSHIP v. 1.2 (Queller & Goodnight 1989). The distribution of observed values was then compared with expected distributions for unrelated individuals, PO pairs and FS pairs, the latter values simulated in KINSHIP using empirical allele frequencies.

(e) Rationale for tests of host specificity and assortative mating

By combining data on parentage, songs and plumage colour, we test for possible deviations from perfect host specificity and assortative mating, either of which could lead to hybridization and genetic introgression between indigobird species and host races. If female indigobirds always mate with conspecifics and always parasitize nests of the host species that raised them, then male song mimicry should always be concordant with species-specific differences in plumage colour. In addition, first-order male relatives should always mimic songs of the same host species and should have the same plumage colour.

If female indigobirds show imperfect host fidelity in egg-laying, occasionally parasitizing the host of another indigobird species, we expect to find fathers and sons or male siblings that match in plumage colour but mimic the songs of different hosts. In addition, the male reared by the alternate host should have a song that is unexpected given its morphology (figure 1a). This prediction of a mismatch in plumage colour is not applicable, however, if a *V. camerunensis*

female parasitizes both hosts associated with *V. camerunensis*. Nonetheless, any pair of full-sibs that mimic the songs of different hosts provides evidence that individual females sometimes parasitize more than one host species.

If a female indigobird occasionally mates with a male of a different indigobird species but then parasitizes her usual host, father and son will again mimic the songs of different hosts, but they might also differ in plumage colour (figure 1b). Depending on the genetics of colour, the hybrid son might be intermediate or have the colour of either his mother's or father's species. Again, plumage colour will be potentially informative only in cases of hybridization between morphologically distinct species; a father and son sharing blue plumage but mimicking the two different hosts of *V. camerunensis* could result from either imperfect host fidelity in egg-laying or non-assortative mating by females.

3. RESULTS

(a) Captured birds and banding

We captured 265 indigobirds, including 23, 96, 69 and 77 new individuals in each of the 4 years of the study. Thirteen individuals were captured in multiple years, including one male that was caught in three consecutive years. Direct evidence of host association based on recorded songs was available for 188 males (184 of which were in breeding plumage). Of these, 80 were *V. camerunensis* mimicking *L. rara*, 47 were *V. camerunensis* mimicking *L. rubricata*, 48 were *V. raricola* and 13 were *V. wilsoni*. Host association was inferred with less certainty for another 37 individuals (adding 10, 11, 11 and 5 individuals to the above totals, respectively) based on behavioural observations (association with known males, songs that were identified in the field but not recorded, response to playback) and/or plumage colour (*V. wilsoni* and *V. raricola* only). We had no host information for the remaining 40 birds, which were either juveniles, females or males in non-breeding plumage, but included them in the parentage analyses to improve estimates of allele frequencies.

During the first three years of the study, 50 captured birds bore the streaky brown plumage characteristic of juveniles, males in non-breeding plumage and adult females. Of these, nine were recaptured or resighted in male breeding plumage in the year subsequent to their initial capture. Given an expectation of moderate to high mortality for juveniles, and an unknown proportion of females among the 50 captured brown birds, a return of nine males suggests strong natal philopatry. Adult males were also philopatric, as 12 out of 32 adult males recorded at a call site during the first three years of the study were captured or resighted again the next year (see also Payne & Payne 1977; Payne 1985). These data suggest that fathers, sons and siblings are likely to be found in the same general area and that close relatives were likely to be captured during our 4-year study.

(b) Comparison of morphology and behaviour

A plot of plumage hue and brightness measured from voucher specimens yields three non-overlapping distributions corresponding to the three species at Tibati (Balakrishnan *et al.* in preparation). Although there was some variation in hue and brightness within each species, all photographed individuals could be assigned to green, blue

or purple, and none were obviously intermediate in plumage colour as might be expected of interspecific hybrids. All 184 males that were recorded and captured in breeding plumage bore plumage of the colour expected given their mimicry songs. Forty-seven birds mimicking goldbreast were green (*V. raricola*), 124 African firefinch or black-bellied firefinch mimics were blue (*V. camerunensis*) and 13 bar-breasted firefinch mimics were purple (*V. wilsoni*).

(c) Genetic markers and parentage analysis

Nuclear sequence loci approached microsatellite loci in numbers of alleles and heterozygosity, whereas the bi-allelic SNP loci contributed relatively little to the power of the parentage analysis (table 1). No significant deviations from Hardy–Weinberg equilibrium were found and F_{ST} values between species and host races were low (table 1), reflecting the minimal differentiation of indigobird species at neutral genetic markers (Seftc *et al.* 2005). These results suggest that genetic differentiation among species and host races should have relatively little effect on the inference of individual-level relationships and that background allele frequencies calculated for the three species combined are suitable for the simulations used to determine probabilities in the paternity analyses.

Simulation of relatedness coefficients for unrelated pairs, HS pairs and FS pairs in KINSHIP (Queller & Goodnight 1989) suggests that our data should perform well in distinguishing full-sibs from unrelated pairs (type II error rate, failing to reject the null hypothesis of unrelated when 'full-sibs' is true = 7.2% at the $p < 0.001$ significance level), but will often fail to distinguish half-sibs from unrelated pairs (type II error rate = 70%). Non-exclusion probabilities combined across loci indicate a reasonably low likelihood of type I error or false positives (table 1), but the large number of pairs tested in our analysis (25 200 unique pairs among 225 individuals with known or inferred host association) suggests that a few unrelated pairs could by chance share at least one allele at all of the sampled loci. Distributions of relatedness values also suggest the potential for false positives in our analysis; r -values between 0.325 and 0.45 are just as likely to be encountered among the large sample of unrelated individuals in our analysis as in a small sample of close relatives (figure 2). Given these considerations, we sought to limit the likelihood of false positives (i.e. incorrectly inferring that unrelated individuals are related) by (i) attempting to exclude a subset of significant PO pairs using additional loci, and (ii) reporting FS and HS pairs only if the hypothesis of unrelated could be rejected at the $p < 0.0001$ significance level using the likelihood ratio test implemented in ML-RELATE.

Results from CERVUS and ML-RELATE were broadly consistent, generally identifying the same pairs of individuals as close relatives (details provided in the electronic supplementary material). Forty-five putative PO pairs involving birds of known or inferred host association were identified in one or more of the nine analyses completed in CERVUS (table 2). Increasing the assumed proportion of candidate males sampled from 10 to 20 to 30 per cent and/or reducing the assumed genotyping error rate from 1 to 0.1 to 0 per cent generally increased statistical significance (i.e. reduced the p values), although a genotyping error rate of 0 per cent excluded those pairs that did not share at least one allele at every

Figure 2. Comparison of expected and observed distributions of relatedness values. (a) Observed distributions (solid line with diamond) of relatedness values compared with simulated distributions of relatedness values for 34 980 pairs of unrelated (solid line) individuals (the number of pairwise comparisons among 265 birds), 100 FS pairs (dashed line) and 100 PO pairs (dotted line), reflecting an approximate number for each category in the real data. Note that r -values above 0.4 are more frequent in the observed data than expected if all individuals were unrelated, indicating the presence of close relatives in our dataset. (b) Distributions of relatedness values for pairs of relatives identified by CERVUS (PO pairs) and ML-RELATE (FS pairs). The dashed vertical line is at $r=0.45$, below which the number of unrelated pairs (false positives) is expected to be equal or greater than the number of truly related pairs. Each box represents a pair of birds coded according to the hosts with which they were associated (grey box, same host; white box, two hosts of *V. camerunensis*; black box, hosts of two different indigobird species). In the FS plots, pairs significant at $p < 0.0001$ are shown above the line; pairs not meeting this stringent test are shown below. Note that FS pairs with matching mtDNA have higher r -values on average and are more likely to meet the $p < 0.0001$ threshold. These observations support the conclusion that some pairs with relatively low r -values are false positives.

locus ($n=18$). ML-RELATE, which does not allow for genotyping error, identified these pairs as full-sibs ($n=14$) or half-sibs ($n=4$). Of these, 10 had matching mtDNA haplotypes, as expected for full-sibs or maternal half-sibs. The remaining eight pairs had mismatching alleles at a single microsatellite locus, and in seven of these cases an increase or decrease of a single repeat unit in one individual would be sufficient to achieve a match. Given the high mutation rate of microsatellites (reviewed in Ellegren 2000), the data for these seven pairs are perhaps most consistent with PO relationships, as also suggested by CERVUS analyses that allowed for a low level of genotyping error.

ML-RELATE identified 63 pairs of close relatives significant at $p < 0.0001$ (table 2). These included 39 out of the 45 pairs discussed above (the remaining six were

significant only at $p < 0.001$ in the ML-RELATE analysis) plus 24 additional pairs. Twenty-three of these additional pairs also appeared in the output from CERVUS, but were of marginal significance in the CERVUS analysis ($n=8$) or their significance was not tested because one or both individuals was involved in another PO pair with higher likelihood ($n=15$).

Of the 69 putative pairs of close relatives identified in the above analyses, 63 pairs comprise two individuals associated with the same host species. Most of these involve two males in breeding plumage, both mimicking songs of the same host species and matching in plumage colour. Three pairs (including two PO pairs and one HS pair) comprise two males with blue plumage singing the songs of the two *V. camerunensis* host races, one male mimicking *L. rara* and the other mimicking *L. rubricata*.

Table 2. Statistically significant pairs of close relatives identified by CERVUS and ML-RELATE. (Results for CERVUS include all pairs significant at $p < 0.05$ in at least one of the nine analyses. For ML-RELATE, which examined all possible pairs of individuals, only pairs significant at $p < 0.0001$ are shown; this includes 27 out of 30 pairs, 19 out of 62 pairs and 14 out of 2603 pairs for which the most likely relationship was PO, FS or HS, respectively. Expected values for *V. camerunensis* are based on the assumption of no specificity in female mate choice and/or host choice. FS and HS pairs with matching mtDNA haplotypes are indicated in parentheses. Full details on each pair are provided in the electronic supplementary material.)

species or host race male 1–male 2	CERVUS	ML-RELATE			total	expected
		PO	FS	HS		
<i>camerunensis</i> – <i>camerunensis</i>						
<i>rara</i> mimic– <i>rara</i> mimic	26	14	13(11)	7(4)	38	17.16
<i>rubricata</i> mimic– <i>rubricata</i> mimic	4	3			4	6.97
<i>rubricata</i> mimic– <i>rara</i> mimic	3	3		1(0)	4	21.87
<i>raricola</i> – <i>raricola</i>	10	5	8(4)	5(1)	19	
<i>wilsoni</i> – <i>wilsoni</i>	0		1(0)	1(0)	2	
<i>raricola</i> – <i>rara</i> mimic	1 ^a	1 ^a			1 ^a	
<i>wilsoni</i> – <i>rubricata</i> mimic	1 ^a	1 ^a			1 ^a	

^a Pair excluded based on data from additional loci.

A fourth PO pair comprised a blue male mimicking *L. rara* and a female indigobird captured the next year at the call site of a male mimicking *L. rubricata*. Finally, the last two pairs (both PO pairs) include males that were assigned to two different species based on plumage colour and that were also mismatched in song, each male singing the song of their expected host. None of the latter six pairs had matching mtDNA haplotypes.

Further examination of our results indicates that the two pairs in which males both mismatched in plumage colour and mimicked the hosts of two different indigobird species also (i) include individuals with relatively high non-exclusion probabilities, (ii) are relatively weakly supported as measured by delta values from ML-RELATE (see the electronic supplementary material), and (iii) have relatively low r -values (figure 2), reflecting the fact that these pairs share common alleles at many loci. To test whether these pairs might represent false positives, we collected DNA sequence data for three additional loci for all four mismatched PO pairs that involved recorded males in breeding plumage plus six additional PO pairs comprising males that matched in both plumage colour and songs. Both pairs that mismatched in both song and plumage were excluded by one or two of the three additional loci. Mismatches at these loci involve SNPs which should have much lower mutation rates and be less susceptible to genotyping errors than microsatellites. By contrast, none of the eight intraspecific pairs (including the two comprising *V. camerunensis* males associated with different hosts) were excluded by this additional testing.

Pairs of related *V. camerunensis* individuals associated with different hosts were far less frequent than would be expected if female *V. camerunensis* mated with males irrespective of song mimicry and/or parasitized the two hosts of *V. camerunensis* indiscriminately (table 2; $G_{adj} = 49.8$, d.f. = 2, $p < 0.0001$).

4. DISCUSSION

Phylogenetic and population genetic analyses strongly support a model of recent sympatric speciation by host shift in the brood parasitic indigobirds (Sorenson *et al.* 2003, 2004). Sympatric indigobirds associated with different hosts are in most cases morphologically distinct

and show limited but often statistically significant differentiation at both nuclear and mitochondrial loci, suggesting limited gene flow between species (Sorenson *et al.* 2003; Sefc *et al.* 2005; table 1). Here, we use parentage analyses to directly assess the possibility of ongoing gene flow among sympatric indigobird species and host races, testing for imperfect fidelity of female indigobirds in mate choice and/or host choice, both of which appear to be dependent on the host species by which a particular female was raised (Payne *et al.* 2000).

Consistent with strong host specificity by female indigobirds and assortative mating between males and females reared by the same host species (figure 1), pairs of close relatives identified in our analyses generally comprised two individuals associated with the same host species. In addition, all 184 males that were recorded and captured in breeding plumage were of the colour expected given their songs; thus, we detected no obvious cases of female indigobirds parasitizing the host of another indigobird species. Similarly, Payne *et al.* (1993) found that just 4 out of 494 male indigobirds (0.8%) recorded in southern Africa mimicked the host of another indigobird species.

An alternative scenario yielding male relatives with different songs posits a female indigobird mating with a heterospecific male before laying an egg in a nest of her usual host (i.e. the species that raised her; figure 1b). In this case, the male offspring would be a hybrid genetically and might be expected to have intermediate plumage colour, unless differences in colour are controlled by a simple genetic polymorphism, with, for example, blue dominant to green or vice versa. Unfortunately, nothing is known of the genetics of plumage colour in indigobirds, though it does not appear to be a developmentally plastic trait (R. B. Payne 2005, personal communication). Although our analyses identified two PO pairs involving males of different species, the large number of pairwise comparisons completed in our search for close relatives resulted in an appreciable risk of false positives. Relatively low relatedness values for these pairs (figure 2) and exclusion by additional sequence loci support the conclusion that both pairs were in fact false positives. In addition, if these were legitimate PO pairs, the results would require that each son inherited a dominant colour gene from his mother, as all four males involved had songs

consistent with their morphology. Thus, our dataset does not include any strongly supported examples of interbreeding between indigobird species.

An inference of host specificity and assortative mating also applies to the two host races of *V. camerunensis*, in which first-degree male relatives generally mimicked songs of the same hosts, suggesting that two sympatric and morphologically indistinguishable populations are, for the most part, reproductively isolated due to their association with two different hosts. Our analyses, however, revealed four exceptions to this overall pattern. Two PO pairs and one significant HS pair comprised two blue-plumaged males mimicking the two different hosts of *V. camerunensis*. The two PO pairs were strongly supported in all of our analyses and were further supported by data from three additional, highly variable loci. Due to the absence of plumage colour differences between the *V. camerunensis* host races, there is no basis for distinguishing between the alternative scenarios described above. Either laying in an alternate host nest or mating with a male of the other host race could produce a father–son pair in which the two males mimic the songs of the two different hosts associated with this species. Likewise, either mechanism could produce a father–daughter pair in which the female prefers males of the other host race; the fourth significant PO pair involved a female whose host association was indirectly inferred based on capture near a male call site.

While the frequency of related individuals associated with different hosts was much lower than expected if female *V. camerunensis* mated indiscriminately or were not generally faithful to their hosts, two to four such pairs in our relatively small sample of close relatives are potentially significant from a population genetic perspective. Imperfect fidelity in mating and/or egg-laying of the order of approximately 5 per cent could lead to a substantial level of ongoing gene flow, sufficient to counteract genetic drift and perhaps also impede adaptive differentiation of the two host races (Wright 1931; Slatkin 1987).

Song playback experiments suggest that *V. camerunensis* host races at Tibati are behaviourally isolated (Balakrishnan & Sorenson 2006); territorial males respond strongly to playbacks of homotypic song, whereas similarly weak responses are elicited by playback of the other *V. camerunensis* host race or other indigobird species. Given that random mating by females should select for generalized territoriality by male indigobirds, this result is consistent with premating reproductive isolation among both species and host races. Whether the lack of plumage differences between host races makes mixed matings more likely than between indigobird species is not known, but the central importance of song in indigobird mate choice is clear (Payne *et al.* 2000). Even if mating is perfectly assortative with respect to song, however, gene flow might continue via imperfect fidelity in host choice (figure 1a).

Reproductive isolation arising as a consequence of imprinting on host song is the key element of the indigobird speciation model (Payne *et al.* 1998, 2000). Our evidence of ongoing gene flow between *V. camerunensis* host races, however, suggests that complete reproductive isolation may not be an automatic or immediate consequence of host colonization. Rather, following colonization, divergent natural and/or sexual selection on morphology, ecology and/or behaviour may be necessary to complete the speciation process. This is

analogous to cases of allopatric speciation in which geographical separation alone is insufficient to produce reproductive isolation (e.g. Rice & Hostert 1993). An important factor probably influencing the extent of reproductive isolation following a host shift by indigobirds is host nesting ecology. Among the four host species at Tibati, the two associated with *V. camerunensis* are the most closely related phylogenetically (Sorenson *et al.* 2003, 2004), show the greatest similarity in nestling morphology (Payne 2005) and show the greatest overlap in nesting habitat at Tibati (C.N. Balakrishnan 2001–2003, personal observations). Together, these factors may facilitate both egg-laying errors by female indigobirds and subsequent survival of offspring in the nest of a closely related alternate host. Cophylogenetic analyses of brood parasitic finches and their hosts indicate a pattern of ‘clade-limited’ colonization wherein host shifts have most often involved closely related host species (Sorenson *et al.* 2004). Thus, the initial colonization of hosts with similar nesting ecology and nestling morphology may be more frequent and/or successful, but the same factors may have the countervailing effect of facilitating ongoing gene flow following colonization.

Animal protocols were approved by Boston University’s Institutional Animal Care and Use Committee and followed the guidelines of the Ornithological Council.

We thank T. Kunz, I. Lovette, C. Schneider and F. Wasserman for their comments on this manuscript. R. Payne provided advice on fieldwork and help with identification of host song mimicry. K. Njabo, J. Dongmo and J. Cooper provided assistance in the field and Jeff Tetrault and Jeff DaCosta assisted with data collection in the laboratory. We thank R. Fotso for helping us initiate our research in Cameroon. T. Marshall and S. Kalinowski provided helpful guidance in the use of parentage analysis software. Permits for work in Cameroon were kindly provided by the Institute of Agricultural Research for Development and the Ministry of the Environment and Forests. This paper is based on work supported by the National Science Foundation under grant nos. 0089757 and 0309249.

REFERENCES

- Balakrishnan, C. N. & Sorenson, M. D. 2006 Song discrimination suggests premating reproductive isolation among sympatric indigobird species and host races. *Behav. Ecol.* **17**, 473–478. (doi:10.1093/beheco/arj052)
- Balakrishnan, C. N. & Sorenson, M. D. In preparation. Genetic and morphological similarity of Cameroon indigobird host races suggests recent divergence and historical gene flow.
- Bulgin, N. L., Gibbs, H. L., Vickery, P. & Baker, A. J. 2003 Ancestral polymorphisms in genetic markers obscure detection of evolutionarily distinct populations in the endangered Florida grasshopper sparrow (*Ammodramus saviarum floridanus*). *Mol. Ecol.* **12**, 831–844. (doi:10.1046/j.1365-294X.2003.01774.x)
- Bush, G. L. 1969 Sympatric host race formation and speciation in frugivorous flies of the genus *Rhagoletis* (Diptera: Tephritidae). *Evolution* **23**, 237–251. (doi:10.2307/2406788)
- Doucet, S. M. & Montgomerie, R. 2003 Structural plumage colour and parasites in satin bowerbirds *Ptilonorhynchus violaceus*: implications for sexual selection. *J. Avian Biol.* **34**, 237–242. (doi:10.1034/j.1600-048X.2003.03113.x)

- Edwards, S. V., Kingan, S. B., Calkins, J. D., Balakrishnan, C. N., Jennings, W. B., Swanson, W. J. & Sorenson, M. D. 2005 Speciation in birds: genes, geography, and sexual selection. *Proc. Acad. Natl Sci. USA* **102**, 6550–6557. (doi:10.1073/pnas.0501846102)
- Ellegren, H. 2000 Microsatellite mutations in the germline: implications for evolutionary inference. *Trends Genet.* **16**, 551–558. (doi:10.1016/S0168-9525(00)02139-9)
- Excoffier, L. G., Laval, G. & Schneider, S. 2005 ARLEQUIN v. 3.0: an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* **1**, 47–50.
- Gaunt, A. S., Oring, L. W., Anderson, D. W., Baptista, L. F., Barlow, J. C. & Wingfield, J. C. 1997 *Guidelines to the use of wild birds in research*. Washington, DC: The Ornithological Council.
- Hey, J., Won, Y. J., Sivasundar, A., Nielsen, R. & Markert, J. A. 2004 Using nuclear haplotypes with microsatellites to study gene flow between recently separated cichlid species. *Mol. Ecol.* **13**, 909–919. (doi:10.1046/j.1365-294X.2003.02031.x)
- Kalinowski, S. T., Wagner, A. P. & Taper, M. L. 2006 ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol. Ecol. Notes* **6**, 576–579. (doi:10.1111/j.1471-8286.2006.01256.x)
- Kalinowski, S. T., Taper, M. L. & Marshall, T. C. 2007 Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.* **16**, 1099–1106. (doi:10.1111/j.1365-294X.2007.03089.x)
- Klein, N. K. & Payne, R. B. 1998 Evolutionary associations of brood parasitic finches (*Vidua*) and their host species: analyses of mitochondrial DNA restriction sites. *Evolution* **52**, 566–582. (doi:10.2307/2411091)
- Linn, C., Feder, J. L., Nojima, S., Dambroski, H. R., Berlocher, S. H. & Roelofs, W. 2003 Fruit odor discrimination and sympatric host race formation in *Rhagoletis*. *Proc. Natl Acad. Sci. USA* **100**, 11 490–11 493. (doi:10.1073/pnas.1635049100)
- Marshall, T. C., Slate, J., Kruuk, L. E. B. & Pemberton, J. M. 1998 Statistical confidence for likelihood-based paternity inference in natural populations. *Mol. Ecol.* **7**, 639–655. (doi:10.1046/j.1365-294x.1998.00374.x)
- Morjan, C. L. & Rieseberg, L. H. 2004 How species evolve collectively: implications of gene flow and selection for the spread of advantageous alleles. *Mol. Ecol.* **13**, 1341–1356. (doi:10.1111/j.1365-294X.2004.02164.x)
- Nielsen, R. & Wakeley, J. 2001 Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* **158**, 885–896.
- Palsbøll, P. J., Bérubé, M., Aguilar, A., Notarbartolo-Di-Sciara, G. & Nielsen, R. 2004 Discerning between recurrent gene flow and recent divergence under a finite-site mutation model applied to North Atlantic and Mediterranean Sea fin whale (*Balaenoptera physalus*) populations. *Evolution* **58**, 670–675. (doi:10.1554/02-529)
- Payne, R. B. 1973 Behavior, mimetic songs and song dialects, and relationships of the parasitic indigobirds (*Vidua*) of Africa. *Ornithol. Monogr.* **11**, 1–333.
- Payne, R. B. 1982 Species limits in the indigobirds (Ploceidae, *Vidua*) of West Africa: mouth mimicry, song mimicry, and description of new species. *Misc. Pub. Mus. Zool. Univ. Mich.* **162**, 1–96.
- Payne, R. B. 1985 Behavioral continuity and change in local song populations of village indigobirds *Vidua chalybeata*. *Z. Tierpsychol.* **70**, 1–44.
- Payne, R. B. 1996 Field identification of the indigobirds. *Bull. Afr. Bird Club* **3**, 14–25.
- Payne, R. B. 2005 Nestling mouth markings and colors of Old World finches Estrildidae: mimicry and coevolution of nestling finches and their brood parasite. *Misc. Pub. Mus. of Zool. Univ. Mich.* **194**, 1–45.
- Payne, R. B. & Payne, K. 1977 Social organization and mating success in local populations of village indigobirds, *Vidua chalybeata*. *Z. Tierpsychol.* **45**, 113–173.
- Payne, R. B. & Payne, L. L. 1994 Song mimicry and species associations of West-African indigobirds *Vidua* with quail-finch *Ortygospiza atricollis*, goldbreast *Amandava subflava* and brown twain-spot *Clytospiza monteiri*. *Ibis* **136**, 291–304. (doi:10.1111/j.1474-919X.1994.tb01098.x)
- Payne, R. B. & Sorenson, M. D. 2004 Behavioral and genetic identification of a hybrid *Vidua*: maternal origin and mate choice in a brood-parasitic finch. *Auk* **121**, 156–161. (doi:10.1642/0004-8038(2004)121[0156:BAGIOA]2.0.CO;2)
- Payne, R. B., Payne, L. L., Nhlane, M. E. D. & Hustler, K. 1993 Species status and distribution of the parasitic indigo-birds *Vidua* in east and southern Africa. In *Proc. VIII Pan-African Ornithological Congress 1992*, pp. 40–52.
- Payne, R. B., Payne, L. L. & Woods, J. L. 1998 Song learning in brood-parasitic indigobirds *Vidua chalybeata*: song mimicry of the host species. *Anim. Behav.* **55**, 1537–1553. (doi:10.1006/anbe.1997.0701)
- Payne, R. B., Payne, L. L., Woods, J. L. & Sorenson, M. D. 2000 Imprinting and the origin of parasite–host species associations in brood-parasitic indigobirds, *Vidua chalybeata*. *Anim. Behav.* **59**, 69–81. (doi:10.1006/anbe.1999.1283)
- Payne, R. B., Woods, J. L. & Payne, L. L. 2001 Parental care in estrildid finches: experimental tests of a model of *Vidua* brood parasitism. *Anim. Behav.* **62**, 473–483. (doi:10.1006/anbe.2001.1773)
- Payne, R. B., Hustler, K., Sternstedt, R., Sefc, K. M. & Sorenson, M. D. 2002 Behavioural and genetic evidence of a recent population switch to a novel host species in brood parasitic indigobirds *Vidua chalybeata*. *Ibis* **144**, 373–383. (doi:10.1046/j.1474-919X.2002.00065.x)
- Payne, R. B., Barlow, C. R., Balakrishnan, C. N. & Sorenson, M. D. 2005 Song mimicry of black-bellied firefinch *Lagonosticta rara* and other finches by brood-parasitic indigobirds *Vidua camerunensis* in West Africa. *Ibis* **147**, 130–143. (doi:10.1111/j.1474-919x.2004.00378.x)
- Price, T. 2007 *Speciation in birds*. Greenwood Village, CO: Roberts and Company.
- Queller, D. C. & Goodnight, K. F. 1989 Estimating relatedness using genetic markers. *Evolution* **43**, 258–275. (doi:10.2307/2409206)
- Rice, W. R. & Hostert, E. E. 1993 Perspective: laboratory experiments on speciation: what have we learned in 40 years? *Evolution* **47**, 1637–1653. (doi:10.2307/2410209)
- Schuetz, J. G. 2005 Reduced growth but not survival of chicks with altered gape patterns: implications for the evolution of nestling similarity in a parasitic finch. *Anim. Behav.* **70**, 839–848. (doi:10.1016/j.anbehav.2005.01.007)
- Sefc, K. M., Payne, R. B. & Sorenson, M. D. 2001 Characterization of microsatellite loci in village indigobirds *Vidua chalybeata* and cross-species amplification in estrildid and ploceid finches. *Mol. Ecol. Notes* **1**, 252–254. (doi:10.1046/j.1471-8278%20.2001.00095.x)
- Sefc, K. M., Payne, R. B. & Sorenson, M. D. 2005 Genetic continuity of brood-parasitic indigobird species. *Mol. Ecol.* **14**, 1407–1419. (doi:10.1111/j.1365-294X.2005.02492.x)
- Seutin, G., White, B. N. & Boag, P. T. 1991 Preservation of avian blood and tissue samples for DNA analyses. *Can. J. Zool.* **69**, 82–90. (doi:10.1139/z91-013)

- Slatkin, M. 1987 Gene flow and the geographic structure of natural populations. *Science* **236**, 787–792. (doi:10.1126/science.3576198)
- Sorenson, M. D. & DaCosta, J. M. In preparation. Genotyping Hap STR loci: phase determination from direct sequencing of PCR products.
- Sorenson, M. D. & Payne, R. B. 2001 A single ancient origin of brood parasitism in African finches: implications for host–parasite coevolution. *Evolution* **55**, 2550–2567. (doi:10.1554/0014-3820(2001)055[2550:ASAOOB]2.0.CO;2)
- Sorenson, M. D., Sefc, K. M. & Payne, R. B. 2003 Speciation by host switch in brood parasitic indigobirds. *Nature* **424**, 928–931. (doi:10.1038/nature01863)
- Sorenson, M. D., Balakrishnan, C. N. & Payne, R. B. 2004 Clade-limited colonization in brood parasitic finches (*Vidua* spp.). *Syst. Biol.* **53**, 140–153. (doi:10.1080/10635150490265021)
- Stephens, M., Smith, N. J. & Donnelly, P. 2001 A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* **68**, 978–989. (doi:10.1086/319501)
- Via, S. 2001 Sympatric speciation in animals: the ugly duckling grows up. *Trends Ecol. Evol.* **16**, 381–390. (doi:10.1016/S0169-5347(01)02188-7)
- Wright, S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97–159.