Conservation Genet Resour (2017) 9:451–458
DOI 10.1007/s12686-016-0680-2

METHODS AND RESOURCES ARTICLE

Using cytochrome *b* to identify nests and museum specimens of cryptic songbirds

A. Elizabeth Arnold[1] · Erik M. Andersen[2] · Michael J. Taylor[2] · Robert J. Steidl[2]

Received: 21 June 2016 / Accepted: 28 December 2016 / Published online: 15 February 2017

Abstract Understanding demography of vertebrate populations requires quantifying reproductive success. For taxa that are difficult to distinguish, estimates of reproductive success can be biased if species are misidentified or if breeding concludes before identification is confirmed. We surveyed desert grasslands where three species of cryptic sparrows breed: *Peucaea botterii, P. cassinii*, and *Aimophila ruficeps* (Emberizidae). Nests, eggs, and nestlings of these species are similar, making it difficult to differentiate nests without observing adults, which can be challenging. We collected seven types of material from nests for DNA analysis: maternal cells from exterior surfaces of unhatched eggs, epithelial cells from the oropharyngeal cavity of nestlings, eggshells, feathers, feather sheaths, feces, and fecal sacs. From these materials, we amplified and sequenced a fragment of the diagnostic locus, mitochondrial cytochrome *b* (cyt *b*), and analyzed the data in a phylogenetic framework to classify nests to species. We validated our classification by sequencing the same locus from feathers of museum specimens. Overall, 72% of samples from nests yielded high-quality sequences. We identified to species 44 of 51 nests and identified museum specimens with archival ages of up to 47 years. Our study extends previous research by demonstrating the efficacy of standard kits, inexpensive reagents, low DNA concentrations, and diverse materials in classifying nests of grassland sparrows. Compared to more invasive methods of acquiring DNA, the approaches we describe are less likely to affect nesting behavior and bias estimates of nesting success of birds, issues especially important for species of conservation concern.

Keywords *Aimophila* · Cyt *b* · Emberizidae · Mitochondrial · Nest success · *Peucaea* · Polymerase chain reaction

Introduction

Understanding mechanisms that govern demography and dynamics of vertebrate populations requires quantifying reproductive success. A key component of reproduction for birds is nest success, but classifying nests by species can be challenging for groups of species where nesting females are secretive, have similar morphology and cryptic coloration, and nest sympatrically. Identification is especially difficult when eggs, nestlings, and nest structure cannot be used to distinguish among species reliably. In grassland bird communities, for example, researchers often locate nests by flushing females that are incubating eggs or brooding young in well-concealed nests (Winter et al. 2003). Frequently, adults will flush and escape to dense vegetation before they can be identified. Researchers will attempt to classify nests during subsequent surveys, but nests often fail before identification has been confirmed, especially when nesting periods are relatively short and intervals between surveys are long. Because the probability of classifying the nesting species is a function of the duration of time that a nest is

Electronic supplementary material The online version of this article (doi:10.1007/s12686-016-0680-2) contains supplementary material, which is available to authorized users.

✉ A. Elizabeth Arnold
arnold@ag.arizona.edu

[1] School of Plant Sciences and Department of Ecology and Evolutionary Biology, The University of Arizona, 1140 E. South Campus Drive, Forbes 303, Tucson, AZ 85721, USA

[2] School of Natural Resources and the Environment, The University of Arizona, Tucson, USA

active, excluding unidentified nests from analyses is likely to bias estimates of success (see Runge et al. 2007).

In southeastern Arizona, we surveyed nests of grassland birds to estimate success, including three similar species of sparrows in the family Emberizidae: *Peucaea botterii, P. cassinii*, and *Aimophila ruficeps*. Nests, eggs, and nestlings of these species are similar in size and appearance, making it difficult to differentiate their nests without observing cryptic adults during the 10–13 day incubation stage or 8–10 day nestling stage.

To increase our ability to classify nests by species, we collected materials from nests for DNA analysis, including maternal cells from exterior surfaces of unhatched eggs, epithelial cells from the oropharyngeal cavity of nestlings, eggshells, feathers, disintegrated feather sheaths, feces, and fecal sacs. From these materials we amplified and sequenced a fragment of the diagnostic locus, mitochondrial cytochrome *b* (cyt *b*), and analyzed the data in a phylogenetic framework to classify nests to species. We validated our classification by amplifying and sequencing the same locus from feathers of museum specimens. Our study extends previous research by demonstrating the efficacy of standard kits, inexpensive reagents, low DNA concentrations, and diverse materials in classifying nests of grassland sparrows.

Methods

From 2013 to 2015, we surveyed nests of birds in three areas of desert grassland and *Prosopis* savanna in southeastern Arizona, USA: Appleton-Whittell Research Ranch (Santa Cruz County), Fort Huachuca Military Reservation (Cochise County), and Las Cienegas National Conservation Area (Pima and Santa Cruz Counties). Elevations of the survey area ranged from 1350 to 1550 m. Approximately 60% of the ca. 400 mm annual precipitation falls between July and September (McClaran 1995), and dominant plants include perennial C_4 bunchgrasses (*Eragrostis, Bouteloua, Hilaria, Bothriochloa*, and *Aristida*) interspersed with many species of forbs, succulents, and low shrubs (Steidl et al. 2013). Velvet mesquite (*Prosopis velutina*) has encroached into many upland areas, increasing the amount of woody structure in these grasslands in recent decades (Van Auken 2000).

Field collections

Between 28 July and 6 September in 2015, we collected materials from 51 nests (Table 1, Supplementary Table 1), 50 of which we anticipated belonging to *P. botterii, P. cassinii*, or *Aim. ruficeps*. Overall, we collected 68 samples of material from nests, including maternal cells from the exterior surfaces of unhatched eggs, epithelial cells from the oropharynx of nestlings, eggshells with attached membranes, feathers, disintegrated feather sheaths, feces, and fecal sacs (Table 1, Supplementary Table 1). We included feathers from one nest of *Ammodramus savannarum* to test our approach across genera and to confirm that nest classifications were not constrained by using materials from only *Peucaea* or *Aimophila*.

We collected maternal cells from surfaces of unhatched eggs by swabbing the exterior of intact eggs with a sterile cotton swab dipped in phosphate-buffered saline (PBS; $n=28$ of 68 samples). We swabbed all exposed surfaces of eggs without removing them from nests. We collected epithelial cells from nestlings between 1 and 6 days old by swabbing inside the oropharyngeal cavity of 1 individual per brood with a dry, sterile cotton swab ($n=12$). Nestlings responded to proximity of the swab by begging with open gapes, which facilitated collection. We did not sample nestlings older than 6 days to avoid causing them to fledge prematurely. Although available less commonly than intact eggs or nestlings, we collected eggshells ($n=10$), feathers ($n=7$), disintegrated feather sheaths ($n=6$), feces ($n=1$), and fecal sacs ($n=4$) with forceps that were sterilized with 95% ethanol between samples.

Table 1 Types of sample material collected from nests of Emberizidae in grasslands in southeastern Arizona; number and percentage of samples sequenced successfully for cyt *b*; DNA concentrations prior to PCR; and sequence length for edited sequences

Sample material	Samples sequenced for cyt *b*, n (%)	DNA concentration, ng/μL [95% CI]	cyt *b* sequence length, basepairs [95% CI]
Epithelial cells, nestling oropharynx	12 (100)	0.81 [0.04–0.12]	890.4 [721.3–1059.0]
Sheath	6 (100)	1.24 [−0.73–3.22]	976.0 [731.6–1220.4]
Feces	1 (100)	0.29 [NA]	672.0 [NA]
Eggshell	9 (90.0)	4.09 [−0.15–8.31]	901.8 [682.2–1121.4]
Feather	6 (85.7)	0.72 [0.52–1.95]	880.5 [633.4–1127.6]
Fecal sac	3 (75.0)	0.27 [−0.29–0.83]	680.3 [626.1–734.5]
Maternal cells, egg surfaces	12 (42.9)	0.67 [0.04–0.10]	683.5 [670.9–696.1]
Total (%) or mean [95% CI]	49 (72.1)	1.04 [0.26–1.82]	837.9 [769.8–906.0]

At the time of collection, we placed feathers in empty, sterile 1.5 mL sample tubes and other sample materials in sterile 1.5 mL sample tubes with 1 mL of sterile PBS. Samples were stored in a cooler for 1–3 days in the field before being transferred to a commercial freezer where they were stored for up to 3 weeks at −20 °C. We then transferred all samples to a −80 °C freezer where they were stored until DNA was extracted.

Museum collections

We used forceps sterilized with 95% ethanol to collect single breast feathers from 20 preserved specimens classified as *P. botterii* and *P. cassinii* in the University of Arizona Natural History Museum. Specimens were preserved as skins from individuals collected in southern Arizona (Santa Cruz, Cochise, and Pima Co.) from 1969 to 1996 (*P. botterii*) and 1982–1996 (*P. cassinii*) (Table 2, Supplementary Table 1). We obtained collection data from the museum's online database and verified information by examining labels on preserved skins (http://www.eebweb.arizona.edu/Collections/Birds/BirdPage.htm). We transferred feathers to collection tubes as described above and stored them at −80 °C until DNA was extracted.

Molecular analyses

We used flame-sterilized forceps to transfer each feather, each sheath, half of each swab, or ca. 2mm^3 subsamples of feces, eggshell, or fecal sac into a 1.5 mL cryotube with ~10 sterile beads (zirconium oxide, 1.0 mm, NextAdvance Inc., Averill Park, NY, USA). We then pulverized samples with a Bullet Blender bead beater (NextAdvance, model BBY24M) following the manufacturer's instructions.

We used the MoBio PowerSoil DNA Isolation Kit (MoBio Laboratories, Carlsbad, CA, USA) to extract total genomic DNA from all materials except feathers. We altered the manufacturer's protocol by incubating samples for 10 min at 65 °C after adding the first buffer (buffer C1, MoBio Laboratories). To extract DNA from feathers, we used the DNEasy Feather Protocol, a user-developed application of the DNEasy Tissue and Blood Kit (Qiagen, Hilden, Germany; see http://www.qiagen.com). We modified the protocol for feathers by extending the pre-extraction incubation time to 24 h following Bush et al. (2005). We used a Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA) to quantify DNA concentration for each sample (Tables 1, 2).

We used the polymerase chain reaction (PCR) to amplify a diagnostic portion of cyt *b* (Harshman 1996; Klicka et al. 1999, 2014; DaCosta et al. 2009). Initially, we used the flanking primers L14764 (Sorenson et al. 1999) and H4A (Harshman 1996) in 20 µL reactions. Reactions consisted of 10 µL of Extract-N-Amp Amp solution (Sigma–Aldrich, St. Louis, MO, USA), 1.3 µL of bovine serum albumin (15 mg/mL), 0.8 µL of each primer (10 µM), 4 µL of template, and PCR quality water. We modified the PCR cycling parameters from Harshman (1996), Klicka and Spellman (2007), and Klicka et al. (2014) to match optima for the PCR chemistry. Cycling reactions consisted of an initial denaturation for 3 min at 94 °C; 35 cycles of denaturation at 94 °C for 30 s, annealing at 51 °C for 45 s, and an extension stage at 72 °C for 2 min; and a final extension at 72 °C for 10 min.

If amplification with these primers failed, we used 0.5 µL of PCR product in a semi-nested approach with the internal primer LCBA (Klicka et al. 1999) and H4A. The PCR reaction mix and cycling parameters remained the same. If that amplification failed, we repeated the PCR with the flanking primers (above) and 4 µL of genomic DNA, but replaced the Amp solution with 10 µL of Phusion Flash high fidelity master mix (Thermo Scientific). PCR cycling parameters were 10 s at 98 °C; 35 cycles of 98 °C for 1 s, 51 °C for 5 s, and 72 °C for 10 s; and 72 °C for 1 min. If amplification was unsuccessful or weak, we then repeated the semi-nested PCR using primers LCBA and H4A, but implementing the Phusion chemistry and cycling parameters.

We prepared all PCR mixtures in a dedicated area that was separated physically from amplified products. We used pipettes designated only for PCR setup, which we cleaned frequently with DNA Away (Molecular BioProducts, Inc., San Diego, CA, USA). We used sterilized barrier tips for all steps in DNA extraction and PCR and worked in a space

Table 2 Number and percentage of museum samples (i.e., single breast feathers from archived skins) from the University of Arizona Natural History Museum that were sequenced successfully for cyt *b*, DNA concentrations prior to PCR, and sequence length for edited sequences

Sample material	Samples sequenced for cyt *b*, n (%)	DNA concentration, ng/µL [95% CI]	Sequence length, basepairs [95% CI]
Feather, *P. botterii*	2 (20.0 %)	0.03 [0.01–0.07]	692.3 [603.4–781.3]
Feather, *P. cassinii*	5 (50.0 %)	3.30 [−1.82–8.42]	660.0 [634.2–685.8]
Total (%) or mean [95% CI]	7 (35.0 %)	2.07 [−0.89–5.04]	672.1 [647.1–697.1]

Species names reflect putative identification based on collectors' notes

in which there had been no previous research on birds. We used double-sterilized water in place of template as a negative control in each reaction. Negative controls were always clear, indicating a lack of contamination during PCR preparation.

We verified amplification by gel electrophoresis, visualizing products on a 2% agarose gel in tris–acetate ethylenediaminetetraacetic acid with staining by SYBR Green (Thermo Scientific). All products that amplified successfully generated single bands of the appropriate length given the primer sets. We used ExoSAP-IT (Affymetrix; Santa Clara, CA, USA) to clean samples that amplified successfully. Cleaned samples were normalized prior to bidirectional Sanger sequencing with the PCR primers (5 μM) at the University of Arizona Genetics Core.

We assembled DNA sequences, called bases, and assigned quality scores using *phred* and *phrap* (Ewing and Green 1998; Ewing et al. 1998) with orchestration by Mesquite v. 1.06 (http://www.mesquiteproject.org). We edited consensus sequences manually in Sequencher 5.1 (Gene Codes Corporation, Ann Arbor, MI, USA) and compared sequences manually against the NCBI GenBank database to estimate taxonomic placement based on BLASTn (Altschul et al. 1990). We used phylogenetic analyses to confirm taxonomic identification (below). We deposited sequence data to GenBank under accession nos. KY608288-KY608340 (Supplementary Table 1). We compared DNA concentrations and sequence length among sample materials using Kruskal–Wallis tests.

Phylogenetic analyses

We used phylogenetic analysis to verify taxonomic placement of nest and museum material. We downloaded published cyt *b* sequences from GenBank for vouchered specimens to represent the clade containing *P. botterii* and *P. cassinii* (sensu DaCosta et al. 2009), including *P. botterii, P. cassinii, P. aestivalis, P. ruficauda*, and *P. humeralis* (Supplementary Table 2). Sequences of these species, representing 7 GenBank accessions, were aligned with representative sequences obtained in this study, including the 6 highest-quality sequences from museum specimens and 47 sequences representing 42 nests (we did not include data from nests identified by BLASTn as belonging to *Aim. ruficeps* or *Amm. savannarum* because these fall outside the clade containing *P. botterii* and *P. cassinii*). We aligned these data in MUSCLE in the EMBL portal (http://www.ebi.ac.uk) and verified the resulting alignment by eye in Mesquite v. 3.04 (Maddison and Maddison 2015). We trimmed the alignment to consistent start and end points, reducing the alignment from 1537 to 629 characters. Original and trimmed alignments have been archived at TreeBASE (study ID S20370).

We analyzed the aligned data with maximum parsimony methods via a heuristic search in PAUP* v. 4.0a147 (Swofford 2002) and by maximum likelihood in GARLI v 0.95 (Zwickl 2006). The former analysis identified 42 variable characters that were parsimony-uninformative and 87 parsimony-informative characters. The GTR + I + gamma model was implemented for the maximum likelihood analysis following evaluation of the best model fit in jModelTest (Posada and Crandall 1998). We used 1000 bootstrap replicates in the likelihood framework to evaluate branch support. Similar analyses confirmed classification of nest material belonging to *Aim. ruficeps* and *Amm. savannarum* (data not shown).

Results

By sequencing cyt *b* from diverse materials collected from nests, we classified to species 44 of 51 nests of grassland sparrows (86.3%). The majority of classified nests (93.1%) belonged to *P. botterii*, followed by single nests of *P. cassinii* and *Aim. ruficeps* (Fig. 1, Supplementary Table 1). The nest that we anticipated belonging to *Amm. savannarum* was confirmed (Supplementary Table 1).

Analysis of field-collected material

DNA extraction, PCR, and sequencing were successful for 49 of 68 samples from nests (72.1%) (Table 1). We obtained high-quality sequence data from all oropharyngeal swabs, sheath material, and the single sample of fecal material (Table 1). The majority of eggshells (90.0%), feathers (85.7%), and fecal sacs (75.0%) also were sequenced successfully (Table 1). We were less successful obtaining sequence data from swabs of egg surfaces (42.9%) (Table 1).

DNA concentrations ranged from 0.27 to 4.09 ng/μL and varied markedly among material types ($\chi^2_6 = 20.84$, $P = 0.0020$) (Table 1). Amplification and sequencing often were successful even when DNA concentrations were low. Sequence length after editing ranged from 672 to 976 base-pairs and did not differ appreciably among material types ($\chi^2_6 = 7.41$, $P = 0.28$) (Table 1).

Samples that did not yield high-quality sequence data typically failed at the DNA extraction step (feathers, swabs of egg surfaces) or sequencing step (eggshells, swabs of egg surfaces) (Supplementary Table 3). PCR was unsuccessful for one fecal sac and one egg swab from which measurable DNA had been obtained (Supplementary Table 3).

We used the flanking primers L14764 and H4A to amplify cyt *b* from 20 of 49 samples that ultimately yielded high-quality sequences; the remainder were amplified successfully by semi-nested PCR (primers LCBA and H4A).

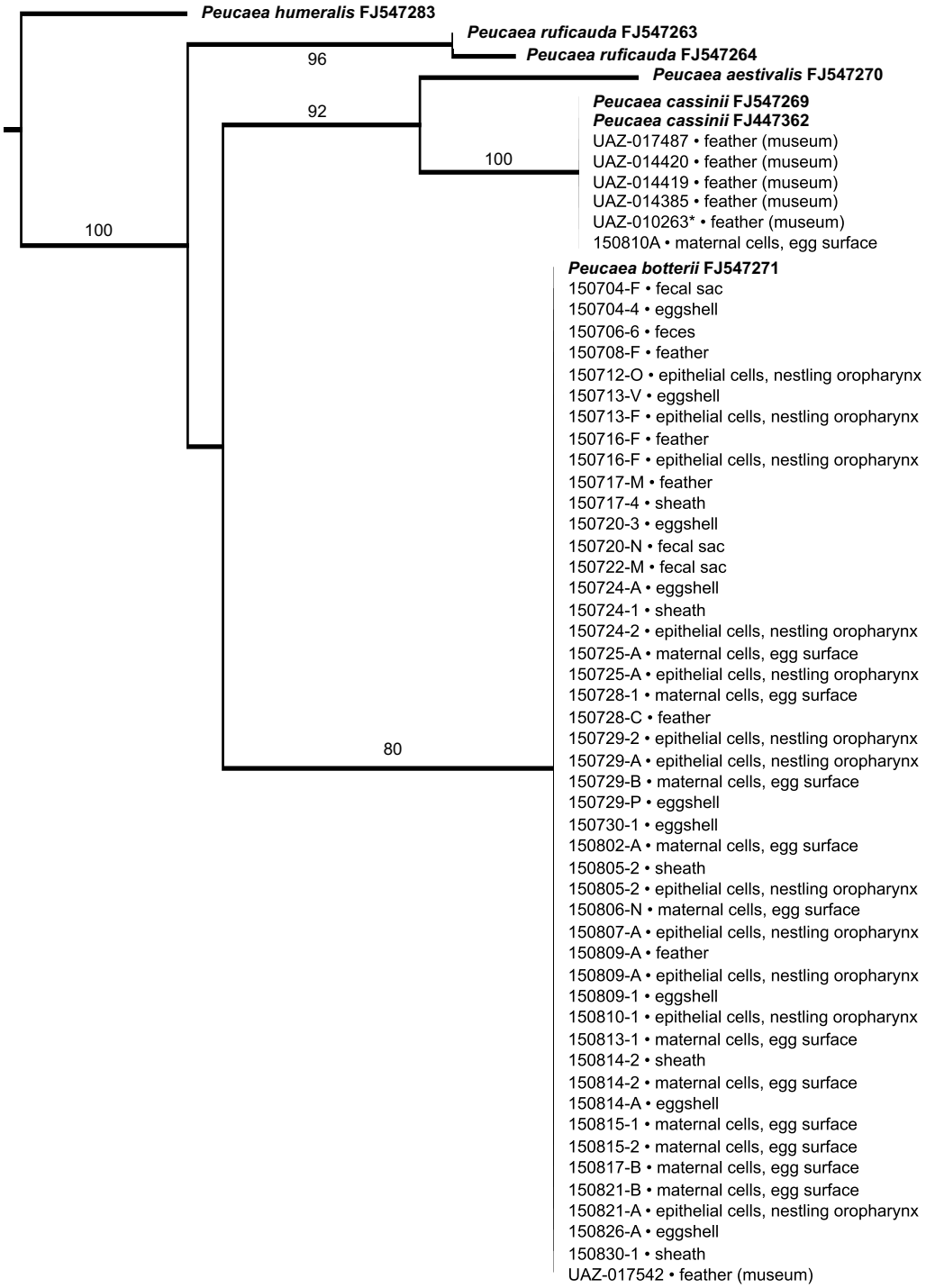

Fig. 1 Phylogenetic classification of nest and museum samples of Emberizid sparrows from southeastern Arizona, including 47 samples from 42 nests, 6 representative museum specimens, and 7 reference sequences representing *Peucaea* species that are related closely to *P. botterii* and *P. cassinii*. The *tree* represents the results of a maximum likelihood analysis, depicted here with branch lengths proportional to change. Parsimony analysis yielded the same topology. *Numbers* near branches indicate bootstrap support (values ≥ 70% are shown). Nest codes (150XXX-X) and museum accession numbers (UAZ-01XXXX) are followed by sample materials. Some nests are represented by more than one sample type; in each case, different samples from the same nests yielded the same classification. The *asterisk* indicates a juvenile specimen (UAZ-010263) identified by the collector as '*Aimophila botterii?*'. It is here classified as *P. cassinii*

Among samples sequenced successfully, 66.7% of oropharyngeal and feather samples, 55.6% of eggshell samples, and 50% of sheath samples were amplified using the flanking primer set. The semi-nested approach was needed to amplify cyt *b* for all samples of feces, fecal sacs, and the exterior surfaces of eggs.

Analysis of museum specimens

We obtained high-quality sequence data from single breast feathers of 7 of 20 museum specimens that had been archived for up to 47 years (Table 2). Success (35%) was less than half that for feathers collected in the field (85.7%) (Tables 1, 2).

DNA concentrations were more variable for feathers on preserved skins than for feathers collected in the field (Tables 1, 2). Semi-nested PCRs were needed consistently to amplify cyt *b* from museum material, and DNA sequences were shorter than those obtained from field-collected feathers (Tables 1, 2).

Failure to obtain sequence data from museum material typically reflected a failure at the PCR step (Supplementary Table 4). We were more successful obtaining sequence data from specimens collected within the last 35 years vs. older specimens (Table 3), but we did obtain high-quality sequence data from one specimen collected 47 years ago (Supplementary Table 1).

Phylogenetic analyses

Phylogenetic analyses confirmed taxonomic placement for nest and museum samples that corresponded to *P. botterii* and *P. cassinii* (Fig. 1). We found that one museum specimen that had been classified as *P. botterii* by the collector in 1969 (UA-010263) is *P. cassinii* (Fig. 1). Different types of material from the same nests always yielded identical species identification (Fig. 1). Within each species, cyt *b* sequences from nest material, museum specimens, and previously published data were ≥99% identical over their shared length.

Table 3 Archival age of museum specimens (i.e., years since collection), number of samples from preserved individuals of *P. botterii* and *P. cassinii* that were sequenced successfully for cyt *b*, and number of samples for which sequence data were not obtained

Archival age	Sequencing for cyt *b*	
	Successful	Failed
≤35 years	6	9
≥40 years	1	4
Total	7	13

Discussion

Quantifying reproductive success for birds depends on identifying species reliably, because misidentifying nests or excluding unidentified nests from analyses can bias estimates of demographic rates (see Runge et al. 2007) and misinform subsequent inferences. We increased our ability to classify nests to species by analyzing a diagnostic portion of the mitochondrial cytochrome *b* gene, which we amplified and sequenced from diverse materials collected from nests. Overall, we classified 44 of 51 nests of focal Emberizid species.

Our methods identified nests of *P. botterii, P. cassinii, Aim. ruficeps*, and, as expected based on nest traits, *Amm. savannarum*. Our inferences benefited from previous publication of high-quality sequence data (e.g., DaCosta et al. 2009; Yuri and Mindell 2002), which permitted BLASTn comparisons with GenBank records and subsequent phylogenetic analyses. Overall, phylogenetic relationships inferred here corresponded to current phylogenetic hypotheses for the group (DaCosta et al. 2009), providing strong evidence for accurate classification of nest material and museum specimens.

The majority of nests we sampled were classified as *P. botterii*. Of the three focal species, *P. botterii* was most common in the study area: during three field seasons, we found and identified 223 nests of *P. botterii*, 11 nests of *A. ruficeps*, and 2 nests of *P. cassinii* based on adults at nests (Andersen and Steidl, unpublished).

Although the approaches we describe were less successful for identifying museum specimens, we were able to identify 7 of 20 skins sampled from sparrows collected in southeastern Arizona with archival ages ranging from 20 to 47 years. High-quality data were obtained from specimens that had been preserved for 20, 33, 34, and 47 years. We anticipate that focusing on shorter DNA fragments, ensuring that feather sheaths are obtained when feathers are plucked, or using two feathers per individual rather than one could increase successful identification without damaging specimens.

We found that one specimen in the collection labeled as *P. botterii* likely was a misidentified *P. cassinii*. This specimen was a recently fledged juvenile, a stage with limited diagnostic characters to distinguish congeners. It was collected soon after *P. botterii* had been documented as breeding in Arizona for the first time following earlier extirpation (Ohmart 1968), and was labeled as the first *P. botterii* fledgling collected in Arizona.

Perspective on molecular methods

We used kits that are available commercially to extract DNA from most materials from nests. One protocol was

optimized for feathers; the other was standard for DNA extraction from soils and has been used effectively to extract high-quality DNA from diverse and often recalcitrant materials (e.g., Wood et al. 2012; Eichmiller et al. 2016; Oakley and Kogut 2016).

We recovered DNA concentrations that generally were comparable to those of recent studies (see Handel et al. 2006). For example, single retrix feathers from black-capped chickadees (*Poecile atricapillus*) yielded an average of 1.16 ng/mL of DNA (Harvey et al. 2006); we obtained an average of 0.72 ng/mL from field samples to 2.07 ng/mL from museum samples (Tables 1, 2). Buccal swabs from black-capped chickadees yielded DNA concentrations of 2.7–4.3 ng/mL (Handel et al. 2006), which were four- to eightfold greater than we obtained in the field; however, we still obtained high-quality sequence data from all 12 oropharyngeal samples (Table 1). Swabs of eggshell surfaces across 23 species yielded average DNA concentrations of 8.38 ng/mL (Martín-Gálvez et al. 2011), 10-fold greater than we obtained; here, we sampled without removing eggs from nests and swabbed only small areas of exposed surfaces (Table 1). Although DNA concentrations were low, current methods and reagents for PCR allowed us to amplify and sequence cyt *b* from these samples frequently, suggesting that DNA quality and purity may be more important than concentration alone. Overall, only 2 of 68 samples from nests failed at the PCR step (2.9%) and 7 failed at the sequencing step (10.3%), most of which (77.8%) were obtained by swabbing the exterior surfaces of eggs (see below; see also Supplementary Table 3). We do not believe that inhibitors were problematic in the PCR step given that amplification failed so rarely, but if evidence for such inhibition is observed, reducing template concentrations may prove useful.

We used Amp solution (Sigma) in our initial PCRs to manage costs: this is a relatively inexpensive, pre-prepared solution for amplification that has been used in diverse applications in molecular ecology (e.g., Tanigawa et al. 2012). The flanking- and semi-nested approaches we used with this chemistry amplified 50% of the samples that ultimately yielded high-quality cyt *b* data. Phusion master mix, which contains a high-fidelity DNA polymerase, often amplified samples that failed with Amp chemistry. For future work, we suggest using a high-fidelity polymerase to accelerate data collection. If costs outweigh urgency, then the combined process of using a less-expensive reagent set first followed by a high-fidelity polymerase may be most cost effective. In turn, the semi-nested approach to PCR increased greatly our success overall, and resulting sequences—although short relative to the reads produced by PCR and sequencing with the flanking primers—were sufficient to differentiate clearly among the sparrow species considered here. We recommend the use of such an approach, provided that sufficient care is taken to prevent cross-contamination.

Recommendations for field surveys

At least 75% of oropharyngeal swabs, sheath material, eggshells, feathers, feces, and fecal sac samples yielded high-quality sequence data, suggesting that these materials could work well to classify to species nests of grassland sparrows and potentially other species in future efforts. Limited success with swabs from unhatched eggs reflects frequent failure at the DNA extraction step, indicating that DNA quantities on swabs were low (Table 1) or that DNA extraction was inhibited by foreign material on egg surfaces. Some of these samples also failed at the sequencing step despite amplifying successfully, consistent with contaminants from other environmental sources in the field. Molecular cloning could be used to generate high quality data from such samples. We suggest using egg swabs as a last-priority choice of field material for the kind of analyses described here, although they may be the only material available from nests that fail before chicks hatch and ca. 43% of these samples did yield useful sequence data. Instead of swabbing only the exposed surfaces of eggs, we recommend removing eggs and swabbing the entire surface to increase the quantity of DNA (see Martín-Gálvez et al. 2011).

Overall, this study complements a growing body of literature describing an array of non-invasive methods for collecting DNA from birds (e.g., Taberlet and Bouvet 1991; Segelbacher 2002; Handel et al. 2006; Schmaltz et al. 2006; Gebhardt et al. 2009; Guerrini and Barbanera 2009; Martín-Gálvez et al. 2011). The methods we describe extend the scope of previous research by demonstrating the efficacy of commercial kits, relatively inexpensive and industry-standard reagents, low concentrations of DNA, and diverse materials for classifying nests of grassland sparrows. Compared to more invasive methods of acquiring DNA, the approaches we describe are less likely to affect nesting behavior and bias estimates of nesting success of birds, issues especially important for species of conservation concern.

Acknowledgements We thank Scott Freeman, Marion Cambrelin, and Joann Wang for assisting in the field; Ming-min Lee, Nicholas P. Garber, Justin P. Shaffer, and Sarah Araldi-Brondolo for assisting with laboratory analyses; and Peter N. Reinthal and Melanie Bucci for providing access to specimens at the University of Arizona Natural History Museum. Funding for molecular analyses was provided by the College of Agriculture and Life Sciences at The University of Arizona. Funding for the field effort was provided by the Bureau of Land Management and Arizona Game and Fish Department. In-kind support for the field effort was provided by Audubon's Appleton-Whittell Research Ranch and U.S. Department of Defense, Fort Huachuca; Linda Kennedy, Sheridan Stone, and Karen Simms provided administrative and logistical support. The University of Arizona

Undergraduate Biology Research Program provided a research opportunity for MJT.

Author contributions AEA and EMA conceived of the study; EMA collected field samples; EMA, AEA, and MJT sampled museum specimens; MJT and AEA completed molecular analyses; AEA, EMA, MJT, and RJS wrote the manuscript.

References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410

Bush KL, Vinsky MD, Aldridge CL, Paszkowski CA (2005) A comparison of sample types varying in invasiveness for use in DNA sex determination in an endangered population of greater sage-grouse (*Centrocercus urophasianus*). Conserv Genet 6:867–870

DaCosta JM, Spellman GM, Escalante P, Klicka J (2009) A molecular systematic revision of two historically problematic songbird clades: *Aimophila* and *Pipilo*. J Avian Biol 40:206–216

Eichmiller JJ, Miller LM, Sorensen PW (2016) Optimizing techniques to capture and extract environmental DNA for detection and quantification of fish. Mol Ecol Resour 16:56–68.

Ewing B, Green P (1998) Basecalling of automated sequencer traces using phred: error probabilities. Genome Res 8:186–194

Ewing B, Hiller L, Wendl M, Green P (1998) Basecalling of automated sequencer traces using phred: Accuracy assessment. Genome Res 8:175–185

Gebhardt KJ, Brightsmith D, Powell G, Waits LP (2009) Molted feathers from clay licks in Peru provide DNA for three large macaws (*Ara ararauna, A. chloropterus*, and *A. macao*). J Field Ornithol 80:183–192

Guerrini M, Barbanera F (2009) Noninvasive genotyping of the red-legged partridge (*Alectoris rufa*, Phasianidae): semi-nested PCR of mitochrondrial DNA from feces. Biochem Genet 47:873–883

Handel CM, Pajot LM, Talbot SL, Sage GK (2006) Use of buccal swabs for sampling DNA from nestling and adult birds. Wildlife Soc B 34:1094–1100

Harshman J (1996) Phylogeny, evolutionary rates, and ducks. Dissertation, University of Chicago, Chicago

Harvey MG, Bonter DN, Stenzler LM, Lovette IJ (2006) A comparison of plucked feathers versus blood samples as DNA sources for molecular sexing. J Field Ornithol 77:136–140

Klicka J, Spellman GM (2007) A molecular evaluation of the North American "grassland" sparrow clade. Auk 124:537–551

Klicka J, Barlow JC, McGillivray WB, Doyle TJ, Zink RM (1999) Evidence supporting the recent origin and species status of the timberline sparrow. Condor 101:577–588

Klicka J, Barker FK, Burns KJ, Lanyon SM, Lovette IJ, Chaves JA, Bryson RW (2014) A comprehensive multilocus assessment of sparrow (Aves: Passerellidae) relationships. Mol Phylogenet Evol 77:177–182

Maddison WP, Maddison DR (2015) Mesquite: a modular system for evolutionary analysis. Version 3.04: http://mesquiteproject.org.

Martín-Gálvez D, Peralta-Sánchez JM, Dawson DA, Martín-Platero AM, Martínez-Bueno M, Burke T, Soler JJ (2011) DNA sampling from eggshell swabbing is widely applicable in wild bird populations as demonstrated in 23 species. Mol Ecol Resour 11:481–493.

McClaran MP (1995) The desert grassland. University of Arizona Press, Tucson

Oakley BB, Kogut MH (2016) Spatial and temporal changes in the broiler chicken cecal and fecal microbiomes and correlations of bacterial taxa with cytokine gene expression. Frontiers Vet Sci 3:11.

Ohmart RD (1968) Breeding of Botteri's sparrow (*Aimophila botterii*) in Arizona. Condor 70:277

Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14:817–818

Runge JP, Hines JE, Nichols JD (2007) Estimating species-specific survival and movement when species identification is uncertain. Ecology 88:282–288

Schmaltz G, Somers CM, Sharma P, Quinn JS (2006) Non-destructive sampling of maternal DNA from the external shell of bird eggs. Conserv Genet 7:543–549

Segelbacher G (2002) Noninvasive genetic analysis in birds: testing reliability of feather samples. Mol Ecol Notes 2:367–369

Sorenson MD, Ast JC, Dimcheff DE, Yuri T, Mindell DP (1999) Primers for a PCR-based approach to mitochondrial genome sequencing in birds and other vertebrates. Mol Phylogenet Evol 12:105–114

Steidl RJ, Litt AR, Matter WJ (2013) Effects of plant invasions on wildlife in desert grasslands. Wildlife Soc B 37:527–536

Swofford DL (2002) PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland

Taberlet P, Bouvet J (1991) A single plucked feather as a source of DNA for bird genetic studies. Auk 108:959–960

Tanigawa M, Sato Y, Eijiri H, Imura T, Chiba R, Kawaguchi M, Tsuda Y, Murata K, Yukawa M (2012) Molecular identification of avian haemosporidia in wild birds and mosquitoes on Tshusima Island, Japan. J Vet Med Sci 75:319–326

Van Auken OW (2000) Shrub invasions of North American semiarid grasslands. Annu Rev Ecol Syst 31:197–215

Winter M, Hawks S, Shaffer JA, Johnson DH (2003) Guidelines for finding nests of passerine birds in tallgrass prairie. Prairie Naturalist 35:197–211

Wood JR, Wilmshurst JM, Wagstaff SJ, Worthy TH, Rawlence NJ, Cooper A (2012) High-resolution coproecology: using coprolites to reconstruct the habit and habitats of New Zealand's extinct upland moa (*Megalapteryx didinus*). PLoS ONE 7:e40025

Yuri T, Mindell DP (2002) Molecular analysis of Frgingillidae, 'New World nine-primaried oscines' (Aves: Passeriformes). Mol Phylogenet Evol 23:229–243

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Dissertation, The University of Texas, Austin