Migration and song elaboration in wood-warblers (Geothlypis)

Author: Bruce E. Byers

RESEARCH ARTICLE

Migration and song elaboration in wood-warblers (*Geothlypis*)

Bruce E. Byers

Department of Biology, University of Massachusetts, Amherst, Massachusetts, USA
bbyers@bio.umass.edu

ABSTRACT

Although some comparative studies of oscine songbirds have found that long-distance migration is positively correlated with elaborate songs, an analysis of singing by species in the genus *Geothlypis* (Parulidae) found no evidence of such a correlation. The migratory species in the genus sing relatively simple songs, whereas the singing of nonmigratory species varies; some species have simple songs, and others have more elaborate songs. Elaborate songs are found in the nonmigratory species *G. semiflava*, *G. aequinoctialis*, and *G. poliocephala*. For example, in Costa Rican populations of these species, songs are longer, contain more notes, have greater note-type diversity, and (in *G. semiflava* and *G. aequinoctialis*) have more phrase types than the songs of the migratory *Geothlypis* species. However, in other nonmigratory species (*G. nelsoni*, *G. flavovelata*, *G. speciosa*, and *G. rostrata*), the duration, note count, and note-type diversity of songs are similar to those of the migratory species. Thus, there seems to be no consistent relationship between migration and song elaboration in *Geothlypis*. In accordance with this inconsistency, ancestral-character-state reconstruction showed that evolutionary loss of migratory behavior was associated with increased song elaboration in some clades within the genus, but not in others. Overall, song variation in wood-warblers (*Geothlypis*) provides no support for the hypothesis that long-distance migration favors the evolution of elaborate songs.

Keywords: birdsong, *Geothlypis*, migration, vocalization, yellowthroat

Migration et élaboration du chant chez les parulines du genre *Geothlypis*

RÉSUMÉ

Bien que certaines études comparatives sur les oiseaux chanteurs aient trouvé que la migration sur de longues distances soit positivement corrélée avec des chants élaborés, une analyse du chant par espèce chez les parulines du genre *Geothlypis* n'a trouvé aucune preuve d'une telle corrélation. Les espèces migratrices de ce genre émettent des chants relativement simples, alors que les chants des espèces non migratrices varient; certaines espèces ont des chants simples et d'autres ont des chants plus élaborés. Des chants élaborés se rencontrent chez les espèces non migratrices *G. semiflava*, *G. aequinoctialis* et *G. poliocephala*. Par exemple, dans les populations costaricaines de ces espèces, les chants sont plus longs, contiennent plus de notes, ont une plus grande diversité de types de notes et (chez *semiflava* et *aequinoctialis*) ont plus de types de phrases que les chants des espèces migratrices de *Geothlypis*. Toutefois, chez d'autres espèces non migratrices (*G. nelsoni*, *G. flavovelata*, *G. speciosa* et *G. rostrata)*, la durée, le nombre de notes et la diversité des types de notes dans les chants sont similaires à ceux des espèces migratrices. Ainsi, il ne semble pas y avoir de relation cohérente entre la migration et l'élaboration du chant chez *Geothlypis*. Conformément à cette incohérence, la reconstruction des caractères ancestraux a montré que la perte évolutive du comportement migratoire était associée à une augmentation de la complexité du chant chez certains clades du genre, mais pas chez d'autres. En général, la variation du chant chez les parulines du genre *Geothlypis* ne supporte pas l'hypothèse que la migration sur de longues distances favorise l'évolution des chants élaborés.

Mots-clés: chant d'oiseau, vocalisation, *Geothlypis*, migration

INTRODUCTION

In the years since Catchpole (1982) hypothesized that migratory songbird species might be expected to sing more elaborate songs than nonmigratory species, several broad comparative studies have found a positive association between song elaboration and migratory distance (Read and Weary 1992, Mountjoy and Leger 2001) or breeding latitude (Botero et al. 2009, Mahler and Gil 2009, Weir and Wheatcroft 2011, Cardoso et al. 2012, Greig et al. 2013). These empirical findings are intriguing, but they are also puzzling, because it is not clear why migration and song elaboration would be linked. The studies cited above generally invoke the explanation advanced by Catchpole (1982): If sexual selection by female mate choice is more intense in migratory species, this increased intensity

TABLE 1. Sizes of song samples (*n*) analyzed to characterize the singing of 9 wood-warbler species. For the first 3 species listed, samples were recorded in Costa Rica in 1998, 2001, and 2011; for the other 6 species, samples were drawn from song archives. The archived songs were recorded between 1981 and 2012 in Mexico (*G. flavovelata*, *G. nelsoni*, and *G. speciosa*), the Bahamas (*G. rostrata*), and North America (*G. formosa* and *Oporornis agilis*).

	Birds (*n*)	Songs (*n*)	Songs bird^{-1} (range)
Costa Rican recordings			
Olive-crowned Yellowthroat (*G. semiflava*) [a]	11	676	51–86
Masked Yellowthroat (*G. aequinoctialis*) [a]	14	916	44–93
Gray-crowned Yellowthroat (*G. poliocephala*) [a]	28	1223	10–116
Archive recordings			
Kentucky Warbler (*G. formosa*)	20	226	3–27
Altamira Yellowthroat (*G. flavovelata*)	6	21	1–8
Hooded Yellowthroat (*G. nelsoni*)	11	114	2–31
Bahama Yellowthroat (*G. rostrata*)	10	101	2–45
Black-polled Yellowthroat (*G. speciosa*)	16	117	1–19
Connecicut Warbler (*O. agilis*)	21	340	2–40

[a] Additional archive recordings were also analyzed for the 3 species recorded in Costa Rica; see text for details.

should result in more elaborate songs. This reasoning, however, assumes that song elaboration in songbirds has resulted, in general, from sexual selection—a generalization that is not well supported by evidence (Wiley 2000, Byers and Kroodsma 2009, Soma and Garamszegi 2011).

An alternative explanation arises from Kroodsma's (1999) hypothesis that singing should be more elaborate when breeding density is high, because high density leads to frequent and intense vocal interactions between males, and this vocal competition favors the evolution of diverse signal arrays to help mediate the competition. A potential corollary of this hypothesized relationship between density and song elaboration is that, because population density tends to be higher in species that live at high latitudes than in tropical species (Terborgh et al. 1990, Gaston and Blackburn 1996, Johnson 1998, Symonds et al. 2006), songs will be most elaborate in long-distance migrant species that tend to breed at high latitudes. This explanation, however, remains tenuous, because neither the assumed association between breeding density and song elaboration nor the assumed association between population density and latitude has been systematically investigated in songbirds.

Overall, it seems to me that the correlations between singing and migration that have emerged from the broad comparative studies cannot currently be placed in the context of a biologically compelling theoretical explanation. In addition, broad comparative studies sometimes fail to find a relationship between migration and song elaboration (e.g., Medina and Francis 2012). It thus seems prudent to await additional evidence before concluding that the putative association is robust. In the meantime, a workable null hypothesis might be that among-species variation in song elaboration arose through divergent cultural evolution driven by stochastic, nonadaptive changes that can accompany cultural transmission (Lynch

1996, Bentley et al. 2004, Byers et al. 2010). Under this null hypothesis, variation in birdsong elaboration is a function of historical contingency and random cultural drift, rather than a predictable outcome of variation in ecological or social conditions.

To help determine whether rejecting the null hypothesis is justified with respect to the effects of migration, it is useful, in my view, to supplement the broad but shallow comparative studies with more detailed comparisons involving larger per species song samples from a smaller set of species. Larger song samples per species are more likely to accurately characterize the often highly variable singing of each sampled species than are the very small song samples used in most broad comparative studies.

Here, I present the results of an investigation that incorporates some relatively large song samples from yellowthroats in the genus *Geothlypis* (Parulidae). I first describe the songs of 3 nonmigratory Central American *Geothlypis* species. The songs are described in some detail, in order to identify distinctive features that might be compared across the genus. To make this comparison, I gathered information on the songs of other *Geothlypis* species from the literature and, for species whose songs have not been formally described, from archived song recordings. Then, making use of the well-resolved phylogeny constructed by Escalante et al. (2009), I mapped key song features onto the *Geothlypis* phylogeny, in order to determine whether evolutionary shifts in migratory behavior were associated with changes in aspects of song elaboration.

METHODS

My analysis of singing by Central American yellowthroats is based mainly on recordings from Costa Rica. In May and June 1998, I recorded birds at 3 locations: (1) Coris,
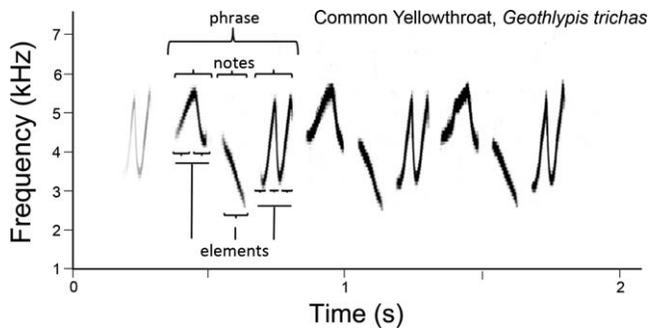
FIGURE 1. Spectrogram of a song of *Geothlypis trichas*, illustrating the nomenclature used to describe yellowthroat song components. The song is composed of a repeated "phrase" that is itself composed of multiple "notes" that incorporate one or more "elements." The song shown was recorded in Massachusetts in 2012.

Cartago (9°51′N, 83°58′W; *poliocephala*), (2) Tapantí National Park (9°47′N, 83°55′W; *poliocephala* and *semiflava*), and (3) in the vicinity of Las Alturas field station (8°56′N, 82°50′W; *aequinoctialis*). I also analyzed recordings made in July 2001 near Cañitas, Puntarenas (10°19′N, 84°49′W; *poliocephala*), and in July 2011 at Rincon de la Vieja National Park (10°47′N, 85°17′W; *poliocephala*).

All the Costa Rican recordings were made opportunistically, beginning before dawn and continuing until singing became sporadic. Each day's recording session took place at a different site within a locale, with the recordist rotating among birds, such that each individual's singing was included in multiple samples of 2–15 min. Birds were not marked, so individuals were distinguished by their singing locations. Because the recordist could usually hear multiple birds singing simultaneously and did not attempt to revisit individuals on subsequent days, I am confident that assignments of songs to individuals were largely accurate. Overall, the Costa Rican recordings include 2,815 songs from 53 birds (Table 1).

In 1998, songs were recorded with a Sony TC-D5ProII cassette recorder. In 2001 and 2011, the recorders were an HHB Portadat PDR-1000 digital audio tape recorder and a Sound Devices 722 solid state recorder, respectively. Microphones included a Sennheiser ME-62 (in 1998) or MKD-2H (in 2011) mounted on a Telinga Universal parabolic reflector, and a Telinga Pro5W integrated parabolic microphone (in 2001).

I supplemented the Costa Rican recordings with recordings of the 3 species from the sound archives at the Cornell Lab of Ornithology Macaulay Library, the Ohio State University Borror Laboratory of Bioacoustics, the Florida Museum of Natural History, and xeno-canto (http://xeno-canto.org). I reviewed 142 archived recordings, each of which contained 2–45 songs. The recordings spanned a wide range of dates and locations: 1981–2009 in

Ecuador and Costa Rica (for *semiflava*, 19 recordings); 1956–2009 in Brazil, Bolivia, Venezuela, Paraguay, Peru, Argentina, and Costa Rica (for *aequinoctialis*, 84 recordings); and 1955–2010 in Guatemala, Nicaragua, Mexico, Texas, El Salvador, Belize, and Costa Rica (for *poliocephala*, 39 recordings).

In addition to the Costa Rican and archived recordings of *semiflava*, *aequinoctialis*, and *poliocephala*, I also analyzed archived recordings (again from the Cornell, Borror, Florida, and xeno-canto archives) of other *Geothlypis* species, and of the clade's sister species, *Oporornis agilis*. I aimed to analyze 20 recordings of each species, but <20 recordings were available for several species (see Table 1). As with the archive recordings of the Central and South American species, these songs were recorded over a rather wide span of years, and the details of song form may well have changed over that time. However, I do not think it likely that the fundamental structure of songs underwent significant modification over time, a supposition consistent with Bolus's (2014) finding that *trichas* song types persisted over periods of ≥50 yr.

The combined Costa Rican and archived recordings accounted for all the species in *Geothlypis*, except for *trichas*, *philadelphia*, *tolmiei*, and *beldingi*. For the songs of *trichas*, *philadelphia*, and *tolmiei*, I was able to glean the information I required from the very thorough published descriptions of Borror (1967; *trichas* songs from 411 birds throughout the breeding range), Wunderle (1979; *trichas*, 40 birds, New York), Ritchison (1995; *trichas*, 7 birds, Kentucky), and Pitocchelli (1990, 2011; *philadelphia* and *tolmiei*, ~765 and ~115 birds, throughout breeding range). For *beldingi*, neither archived song recordings nor published song measurements were available. My analysis excludes "flight songs," which are uncommonly uttered vocalizations (often performed in flight) that have been described for some *Geothlypis* species but that are rarely recorded and have been documented in only a minority of species in the genus (reviewed in Spector 1992).

To describe the songs of *Geothlypis* species, I have adopted the nomenclature devised by Borror (1967) to describe *trichas* songs (Figure 1). Under this nomenclature, a song component that produces a continuous trace on a spectrogram is designated a "note," and each note is composed of one or more upsweeping or downsweeping "elements." A recurring, stereotyped series of different notes is a "phrase."

For all songs in the Costa Rican and archive samples, spectrograms (512 point FFT) and oscillograms were generated with Raven Pro version 1.3 (Bioacoustics Research Program 2008). Representative examples of spectrograms, chosen to capture the between-rendition variation in each recording, were printed out. On the printed spectrograms, I examined the notes in each song, classified the notes into distinctive note types, and labeled

each note with its type. These note-type classifications were strictly per individual; I did not attempt to compare notes across the songs of different birds. Distinctive sequences of note types that recurred within an individual's songs were designated as phrase types.

For each song in the sample (except for the archived songs of the 3 Central American species), on-screen cursors were used to measure the duration of each song, and the number of notes in each song was counted. Several other variables were also measured: the number of notes in each phrase, the number of elements in each note (counted once for each note type in a song), the number of different note types in each song, and the number of different phrase types in each song. For each measurement, I calculated a mean per bird for each sampled individual and then used those values to calculate species means.

To supplement and aid interpretation of the raw measurements, I performed a principal component analysis to generate an integrated measure of song elaborateness, and a k-means cluster analysis to help determine whether songs fell into discrete levels of elaborateness. The input of the principal component and cluster analyses consisted of the species means of song duration, notes per song, elements per note, and note types per song (the 4 variables that could be measured for all species). Statistical analyses were performed with R (R Development Core Team 2014).

The measurements and descriptions I compiled revealed song characteristics that varied among species (see below). I mapped 2 of these traits (phrase structure and song elaborateness) onto Escalante et al.'s (2009) *Geothlypis* phylogeny and performed maximum-likelihood ancestral state reconstructions to recover the most likely evolutionary history of the song features and of migratory behavior. Ancestral character states were reconstructed using the StochChar module in Mesquite version 2.75 (Maddison and Maddison 2011), assuming a Markov k-state 1-parameter (mk1) model of character evolution, in which a single rate of change between states is estimated from the data for each character, and all state changes of a character are assumed to be equally probable. In the reconstruction of song elaborateness, each species was coded as having elaborate or simple song, because maximum-likelihood reconstruction in Mesquite requires categorical characters, and statistical analysis suggested that song elaboration in *Geothlypis* is a 2-state character (see below). As a check, I also performed a maximum-parsimony reconstruction of song elaboration with elaborateness coded as a continuous variable (i.e., as a principal component score).

RESULTS

To facilitate comparisons among species, I use the singing of *G. trichas* as the frame of reference for my descriptions of singing by *Geothlypis* species. The *trichas* song consists of 2–5 repetitions of a 2- to 6-note phrase, yielding a song with a duration of ~2 s (Table 2).

Olive-crowned Yellowthroat

Like *trichas*, the individual *semiflava* in the Costa Rican sample each sang a single song type composed of repeating phrases (Figure 2A). The *semiflava* songs, however, were more elaborate. Instead of a single repeated phrase type, a *semiflava* song contained 3 phrase types. A song consisted of an initial phrase that was repeated 1–3 times, followed by 2–5 repetitions of a second phrase and 1–3 utterances of a third, concluding phrase. This structure yielded greater note-type diversity than is present in *trichas* songs (an average of ~15 different note types per song; see Table 2). In most songs, the phrases became more elaborate as the song progressed; the first phrase typically contained 3 notes, the second phrase typically contained 4 notes, and the final phrase was a more elaborate "coda" that contained ~15 notes. The notes in the coda phrase were generally more complex (i.e. contained more elements) and had larger bandwidth than the notes in the first 2 phrases; some coda phrases included noisy broadband notes or buzzy notes with rapid frequency modulation.

Although each individual's repertoire included only a single song type, different renditions of a bird's song varied. In each of the 3 sections of a song, the number of times the section's phrase was repeated differed among renditions, and the final iteration of a phrase was sometimes truncated (lacked the final note or notes). The terminal coda phrase was occasionally omitted altogether. In addition, in some songs the initial 2-section sequence was produced twice in succession (yielding the impression that the song restarted midway through).

One consequence of the structural variation among renditions is that *semiflava* songs varied in duration. The duration of the shortest songs was about the same as the duration of a *trichas* song, but the average song was longer and contained more notes (see Table 2). Some songs were quite long. The longest song that I recorded was >13 s long and consisted of 3 full, 3-phrase-type sequences strung together in a continuous utterance.

Most of the *semiflava* songs in the archived recordings that I reviewed conformed to the song structure observed in the Costa Rican sample. A few songs from Ecuador, however, diverged from the Costa Rican norm. One song had 4 (instead of 3) different repeated phrase types, and 1 song had a repeating introductory phrase containing 9 notes. Several songs contained nonterminal segments consisting of multiple repetitions of a single note (rather than a multinote phrase).

Masked Yellowthroat

Songs in the Costa Rican sample of *aequinoctialis* had a general structure similar to that of *semiflava* songs, but

TABLE 2. Song characteristics of 11 *Geothlypis* and 1 *Oporornis* species. Measurements for *G. semiflava, G. aequinoctialis,* and *G. poliocephala* are from samples recorded in Costa Rica. Measurements for *G. trichas* are from Wunderle (1979); duration and notes per song) and Borror (1967; all other measurements); those for *G. philadelphia* and *tolmiei* are from Pitocchelli (1990), except for elements per note, which was measured from the song library presented in Pitocchelli (1988). Measurements for all other species are from samples drawn from song archives. Values shown are means, with 95% confidence intervals in brackets and ranges in parentheses (except for PC1 score, which is a dimensionless indicator of elaborateness; see text).

	Duration (s)	Notes song⁻¹ (n)	Notes phrase⁻¹ (n)	Elements note⁻¹ (n)	Note types song⁻¹ (n)	Phrase types song⁻¹ (n)	PC1 score
G. semiflava	3.8 [3.4, 4.2] (1.4–13.7)	18.8 [17.5, 20.0] (4–48)	4.1 [3.7, 4.4] (2–18)	1.7 [1.4, 2.0] (1–5) [b]	14.9 [14.2, 15.6] (11–19)	3 [a]	2.84
G. aequinoctialis	4.7 [4.4, 5.0] (1.2–22.5)	15.8 [14.9, 16.6] (5–63)	3.6 [3.2, 4.0] (2–7)	1.8 [1.3, 2.0] (1–3)	9.3 [8.7, 9.8] (5–12)	3	2.27
G. poliocephala	3.4 [3.1, 3.7] (1.1–7.5)	18.6 [17.8, 19.3] (5–47)	NA [b]	3.1 [2.8, 3.3] (1–7) [c]	18.3 [17.5, 19.0] (5–27)	NA [b]	2.47
G. trichas	1.9 [1.83, 1.95] (1.4–2.3)	12.4 [11.6, 13.1] (10–18)	3.5 [3.4, 3.6] (2–6)	1.9 [1.89, 1.96] (1–5)	3.5 [3.4, 3.6] (2–6)	1 [d]	−0.27
G. formosa	1.8 [1.7, 1.9] (1.3–2.2)	10.8 [8.8, 12.8] (6–21)	2 [e]	3.4 [2.8, 4.0] (2–6)	1.7 [1.4, 2.0] (1–2)	1 [d]	−0.25
G. flavovelata	1.8 [1.7, 2.0] (1.6–2.1)	11.5 [9.4, 13.6] (8–16)	3.7 [3.1, 4.2] (3–4)	1.8 [1.5, 2.2] (1–4)	3.7 [3.1, 4.2] (3–4)	1	−0.40
G. nelsoni	1.9 [1.7, 2.2] (1.3–2.7)	16.2 [13.0, 19.3] (11–23)	2.6 [1.4, 3.7] (2–3)	1.7 [1.3, 2.0] (1–3)	2.9 [1.8, 4.5] (2–5)	1 [d]	0.29
G. rostrata	1.9 [1.7, 2.1] (1.3–2.4)	11.9 [10.2, 13.5] (5–17)	3.1 [2.1, 4.2] (2–4)	1.6 [1.4, 1.9] (1–4)	4.5 [2.9, 6.0] (2–5)	1 [d]	−0.14
G. speciosa	1.8 [1.7, 2.0] (1.1–2.4)	12.7 [10.7, 14.8] (6–24)	NA [b]	2.6 [2.3, 2.8] (1–4)	2.6 [2.0, 3.1] (2–4)	NA [b]	−0.61
G. philadelphia	1.1 [1.08, 1.14] [f]	5.7 [5.5, 5.8] [f]	NA [b]	3.6 [2.8, 4.4] (2–6) [g]	1.7 [1.6, 1.8] (1–3) [g]	NA [b]	−2.42
G. tolmiei	1.3 [1.25, 1.35] [f]	7.0 [6.7, 7.3] [f]	NA [b]	4.3 [3.9, 4.7] (2–6) [g]	2.0 [1.9, 2.1] (1–3) [g]	NA [b]	−2.34
Oporornis agilis	1.5 [1.4, 1.6] (0.8–2.3)	11.3 [10.2, 12.3] (6–16)	3.9 [3.7, 4.1] (3–4)	1.4 [1.3, 1.5] (1–2)	3.8 [3.6, 4.0] (3–4)	1	−0.43

[a] Complete *semiflava* songs include 3 phrase types; shortened renditions may contain fewer phrase types.
[b] Songs of *poliocephala, speciosa, philadelphia,* and *tolmiei* do not contain repeating multinote phrases.
[c] Broadband noisy and buzzy notes excluded.
[d] A few *trichas, formosa, nelsoni,* and *rostrata* individuals have 2 or 3 phrase types song⁻¹.
[e] Phrases contain 2 notes in all *formosa* songs with a repeating phrase, but some songs consist of a single repeating note only.
[f] Ranges of duration and notes per song for *philadelphia* and *tolmiei* were not presented in Pitocchelli (1990) and so are not shown.
[g] Components of *philadelphia* and *tolmiei* songs are not directly comparable to notes as defined here; values shown are for the analogous components designated "syllables" in Pitocchelli (1990).
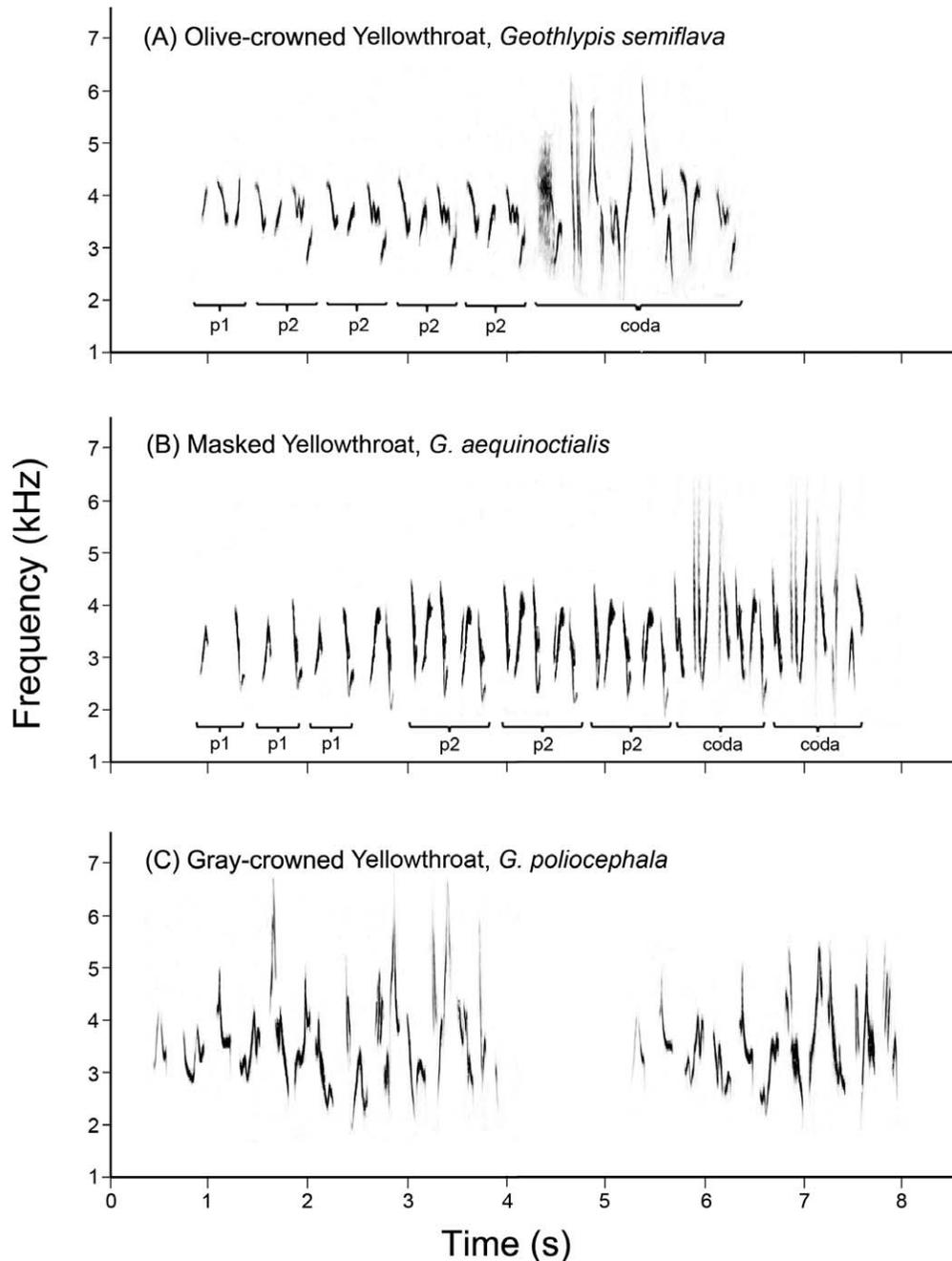
FIGURE 2. Spectrograms of songs of 3 species in the genus *Geothlypis*, recorded in Costa Rica. The spectrograms for (**A**) *G. semilflava* and (**B**) *G. aequinoctialis* are annotated to delineate the 3 phrase types (labeled p1, p2, and coda) that make up the songs. In the *aequinoctialis* song, the first full instance of p2 is preceded by a partial version. The spectrograms for (**C**) *G. poliocephala* show 2 consecutive songs from the same individual, to illustrate 2 different song types from a bird's repertoire. The interval between the 2 songs has been artificially shortened.

with a longer middle section and a shorter coda section (Figure 2B). Each *aequinoctialis* individual had a single song type that had a repeating-phrase structure with multiple phrase types. A typical song consisted of 2–5 repetitions of a 2-note phrase followed by 4–13 repetitions of a 5-note phrase and 1–3 repetitions of a concluding 7-

note coda phrase. In almost all songs, the second section of the song was introduced by a transitional motif consisting of the last 2 notes of the second section's phrase.

The *aequinoctialis* phrases, like the *semiflava* phrases, became increasingly elaborate with each succeeding song section, but elaboration did not increase as sharply in the

coda section of *aequinoctialis* songs as it did in *semiflava* songs. Thus, note-type diversity (on average, ∼9 note types song^{-1}; see Table 2) was somewhat lower than in *semiflava* songs. Also, most notes in the *aequinoctialis* coda were no more complex (in terms of element number) than the notes in the preceding song sections (though many of the coda notes had a broader bandwidth), and coda phrases did not include the kind of noisy or buzzy notes that were present in *semiflava* coda phrases.

As in *semiflava* singing, different renditions of an *aequinoctialis* bird's song varied in duration as a result of variation in the number of repetitions of the phrase in each song section. Songs lacking the coda were more common than in *semiflava* singing, though even songs without the coda could be long. The songs of *aequinoctialis* had the highest average duration among the study species (see Table 2). The longest song recorded was an extraordinary 22.5 s long.

Among the *aequinoctialis* songs from archive recordings, those recorded in Costa Rica had structural characteristics identical to those of the songs in my main sample. Songs recorded at various sites in South America, however, were quite different. None of these songs had the 3-part, repeating-phrase structure characteristic of *aequinoctialis* in Costa Rica. Instead, they exhibited a variety of different structures, most of which included multiple short series of similar or identical notes. Many songs also included sequences in which each note was different.

Gray-crowned Yellowthroat

The *poliocephala* songs in the Costa Rican sample were quite unlike those of the other *Geothlypis* species considered here. Unlike the single-song repertories of *trichas*, *semiflava*, and *aequinoctialis*, *poliocephala* repertoires contained multiple song types. At least 2 song types were recorded from all but 2 of 28 birds sampled; the maximum detected repertoire size was 5 song types (found in the samples of 6 birds). The sampled birds, however, did not sing with immediate variety, so it is possible that, for many individuals, recorded samples did not capture a bird's full song repertoire.

The singing of *poliocephala* also differed from that of the other Central and South American species in that *poliocephala* songs did not exhibit a repeating-phrase structure (Figure 2C). Instead, the songs consisted of a series of notes, each different from the one preceding it; with few exceptions, a note type did not recur anywhere in a song. Thus, note-type diversity per song was higher than in the other species (an average of 18 note types song^{-1}; see Table 2). Some of a bird's note types were shared between its different song types, so that the repertoires of the best-sampled birds contained 50–60 different note types. Notes were generally more complex (contained more elements) than those in the songs of the other yellowthroat species. These structural features yielded songs that sounded more musical and less rhythmic than the songs of the other Costa Rican yellowthroats.

As in the other Costa Rican yellowthroat species, *poliocephala* songs varied in duration. For a given song type, some renditions were complete and other renditions were truncated at different points to yield songs of varying duration. Average song duration was longer than that of *trichas* songs and similar to that of *semiflava* songs (see Table 2), but the longest *poliocephala* songs were not as long as the longest songs of the other 2 Costa Rican species.

The *poliocephala* songs in the archived recordings all had the same basic structure as the songs in the Costa Rican sample. The only exceptions constituted minor deviations: a few songs in which the same note type appeared twice in succession and some instances in which different song types in a bout had no note types in common.

Mexican and Bahamas Endemics

Four *Geothlypis* species (*beldingi*, *flavovelata*, *nelsoni*, and *speciosa*) are endemic to regions in Mexico, and 1 species is endemic to the Bahamas (*rostrata*). These species have limited geographic ranges, and all are nonmigratory. Setting aside *beldingi* for the moment (because its songs were not represented in the song archives), the sampled songs of all the endemic species were similar to those of *trichas* in terms of duration, note-type diversity, and note complexity (see Table 2).

Except for *speciosa*, all the nonmigratory endemic species also shared, with *trichas*, songs that generally consisted of a single repeating phrase type, with the number of notes per phrase likewise similar to *trichas* (Figure 3A–3D). However, the archive recordings contained hints that the occurrence of outlier songs with >1 phrase type may be more frequent in *nelsoni* and *rostrata* than in *trichas*. Among the archived recordings, 1 of 11 *nelsoni* recordings and 2 of 10 *rostrata* recordings contained 2-phrase songs, which suggests frequencies of occurrence that may be higher than the 3% of birds in Borror's (1967) sample of *trichas* that sang such songs. In contrast to the songs of the other endemics, *speciosa* songs did not have a repeating multinote phrase. Instead, a *speciosa* song contained 2 or 3 segments, each of which consisted of repetitions of a single note.

None of the archive recordings of the endemics contained >1 song type, which suggests that these species do not have multitype song repertoires. Song samples were relatively small for these species, and no descriptions of their singing have been published, so there is some uncertainty in the conclusion that individuals of all these species use only a single song type. Nonetheless, the overall similarity of the endemic species' songs to those of *trichas* suggests that the conclusion is plausible.

Although I was not able to measure *beldingi* songs, published anecdotal accounts (e.g., Griscom and Sprunt
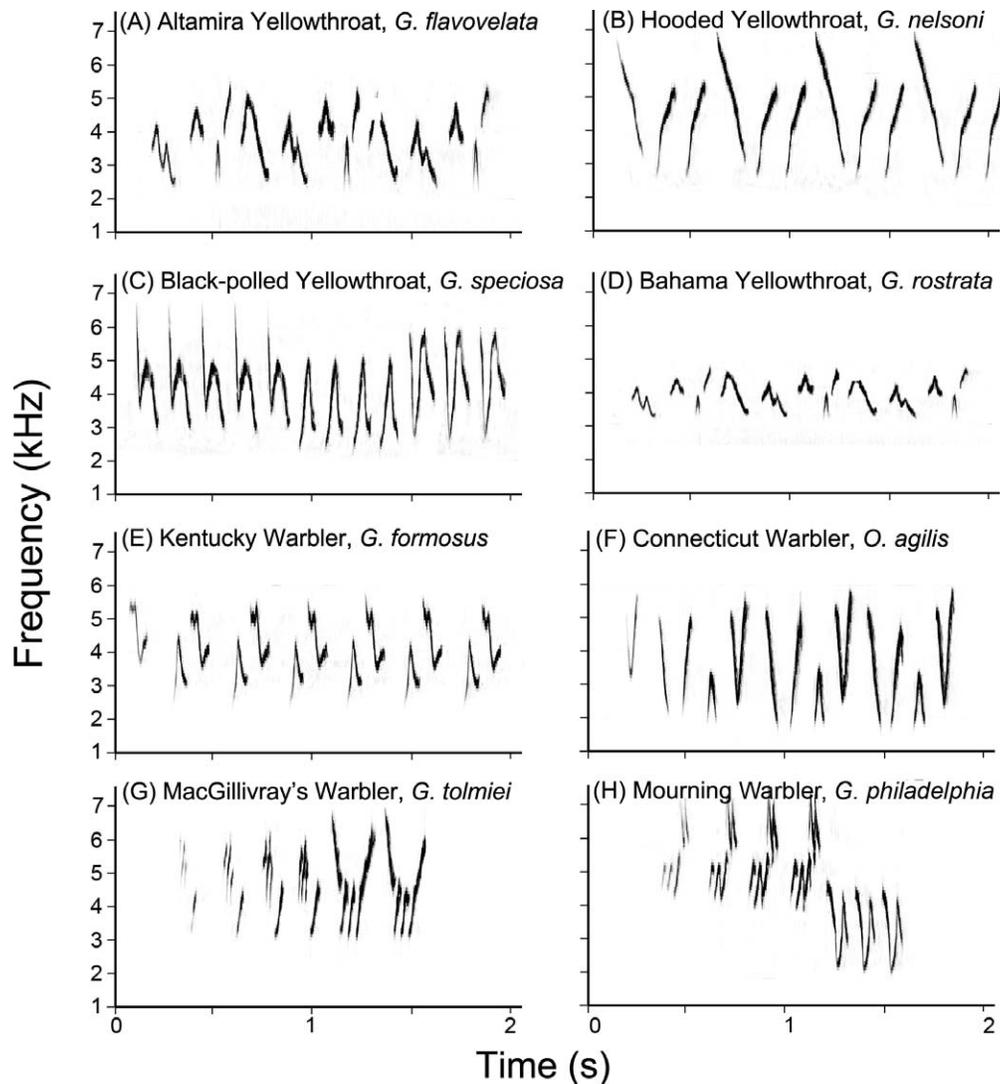
FIGURE 3. Spectrograms of songs of 7 *Geothlypis* and 1 *Oporornis* species, from archived song recordings. The songs were recorded from (**A**) *G. flavovelata* in Tamaulipas, Mexico, 1956 (Macaulay 10076); (**B**) *G. nelsoni* in Mexico City, Mexico, 2006 (xeno-canto XC5863); (**C**) *G. speciosa* in Capulhuac, Mexico, 2012 (Macaulay 172425); (**D**) *G. rostrata* in New Providence, Bahamas, 1971 (Macaulay 10139); (**E**) *G. formosa* in Indiana, USA, 1992 (Macaulay 73967); (**F**) *O. agilis* in Quebec, Canada, 2012 (Macaulay 173597); (**G**) *G. tolmiei* in Oregon, USA, 1988 (Macaulay 42249); and (**H**) *G. philadelphia* in Vermont, USA, 1962 (Macaulay 10104).

1957, Curson et al. 1994, Dunn and Garrett 1997) state that they are similar to *trichas* songs, though perhaps with a richer tonal quality. Therefore, for purposes of comparison, I tentatively assume that *beldingi* individuals have single-song repertoires of short-duration songs consisting of a single repeated phrase-type.

Long-distance Migrants

In addition to *trichas*, 3 other *Geothlypis* species (*formosa*, *philadelphia*, and *tolmiei*) are long-distance migrants, as is *Geothlypis*'s sister species *O. agilis*. The songs of all these migratory species were simple, with mean durations as short as, or shorter than, those of *trichas* songs, and with

note numbers and note-type diversity similar to or lower than those of *trichas* songs (see Table 2).

Although all the long-distance migrants had short, simple songs, the details of song structure varied somewhat among the migrant species (Figure 3E–3H). The structure of *O. agilis* songs was similar to that of *trichas* songs, with repetitions of a single multinote phrase; the number of notes per phrase and number of elements per note were also similar to values from *trichas*. Most *formosa* songs in the archive song sample (those in 14 of 20 samples) also consisted of a single repeated phrase, but in all such songs, the phrase contained only 2 notes. McDonald (2013) reported that this song structure is

TABLE 3. Comparison of *k*-means cluster analyses that specify different numbers of clusters (*k*) in the first principal component of song measurements from 12 wood-warbler species. The comparison shows that an analysis with $k = 2$ yields the highest gap statistic and that $k = 2$ is the smallest *k* for which gap(*k*) ≥ gap(*k*+1) − s_{k+1} (the criterion for estimating the number of clusters in a data set).

Number of clusters (*k*)	Gap statistic	Standard error (*s*)	Gap(*k*+1) − s_{k+1}
1	−0.081	0.026	0.231
2	0.259	0.028	0.083
3	0.116	0.032	−0.015
4	0.028	0.043	NA

common in the species. However, the remaining songs in the *formosa* archive sample did not exhibit a repeating-phrase structure, but instead consisted of repetitions of a single note type (the population studied by Tsipoura and Morton [1988] also contained songs with this structure). A few *formosa* individuals in the archive sample used songs in which a second phrase or note was repeated in a second segment of the song.

Songs of *tolmiei* and *philadelphia* did not contain repeating multinote phrases, but instead contained 1, 2, or 3 segments (most commonly 2), each of which consisted of repetitions of a single note (Pitocchelli 1990, 2011). The "notes," however, differed somewhat from those present in the songs of the other members of the genus. Instead of a linked series of elements that form a continuous trace in a spectrograph, many notes in *tolmiei* and *philadelphia* songs consist of tightly bunched clusters of separate elements. Some of the *formosa* notes illustrated in Tsipoura and Morton (1988) also have this kind of structure. The average complexity of these notes (in terms of mean number of elements per note) is higher than in the notes of most other *Geothlypis* (though comparable to the notes of *poliocephala*).

Song Summary

(1) The songs of the 3 nonmigratory Central and South American yellowthroat species are longer and more elaborate than those of migratory and other nonmigratory *Geothlypis* species. (2) The songs of *poliocephala* and *semiflava* from Costa Rica are generally similar to these species' songs from other locations, but the songs of Costa Rican *aequinoctialis* are quite different from those of South American *aequinoctialis*. (3) Individuals of all *Geothlypis* species have single-song-type repertoires, except for *poliocephala*, in which individuals have repertoires of multiple song types. (4) Most *Geothlypis* species have songs with a repeating-phrase structure, but the songs of *poliocephala*, *speciosa*, *tolmiei*, and *philadelphia* lack the repeating-phrase structure.

Statistical Evaluation of Song Elaboration

Statistical analysis confirmed the conclusion, drawn from inspection of spectrograms and song measurements, that elaborateness is distinctly greater in *poliocephala*, *aequinoctialis*, and *semiflava* than in the other *Geothlypis* species. In a principal component analysis, the first component (PC1) accounted for 69% of the variation in the contributing variables and loaded heavily on song duration, notes per song, and note types per song; it thus serves as a good proxy for song elaboration. Examination of PC1 scores (see Table 2) reveals that 3 species (*poliocephala*, *aequinoctialis*, and *semiflava*) have scores with substantial positive values (>2); that 2 species (*philadelphia* and *tolmiei*) have substantial negative values (<−2); and that the remaining species have small, mostly negative values close to zero. Thus, 3 species stand clumped together on the high-elaborateness end of this scale, 2 species are on the low end of the scale, and the remaining species are clumped together with small absolute values that are somewhat closer to the low end of the scale than to the high end. A *k*-means cluster analysis specifying 3 groups recovered the same grouping. However, a *k*-means analysis specifying only 2 groups lumped *philadelphia* and *tolmiei* together with the larger group from the 3-group analysis. Further analysis suggested that this 2-cluster solution best accounts for the pattern of variability in the elaborateness data. In particular, a gap statistic analysis (Tibshirani et al. 2001) comparing *k*-means clustering with 1, 2, 3, or 4 groups found that the optimal number of clusters was 2 (Table 3).

Ancestral State Reconstruction

Maximum-likelihood ancestral state reconstruction indicates that the common ancestor of today's *Geothlypis* species (Figure 4, node 1) was most likely a long-distance migrant (proportional likelihood = 0.86; Figure 5A), and that migratory behavior was subsequently lost in the species that are now sedentary. In this scenario, migratory behavior re-evolved 2 or 3 times, in the eastern and western lineages of *trichas* and perhaps in *formosa* (depending on whether the lineage leading to node 3 was migratory or nonmigratory). This reconstruction of migratory behavior agrees with the maximum-parsimony reconstruction presented in Escalante et al. (2009) and is similar to the pattern presented in Winger et al. (2012), though that analysis suggests that the migratory ancestor of *Geothlypis* predated the origin of the genus.

Reconstruction of the evolution of song elaboration suggests that the migratory ancestor of *Geothlypis* had simple songs (proportional likelihood = 0.99; Figure 5B). Elaborate songs arose twice: in the lineage leading from node 3 to *poliocephala* and South American *aequinoctialis*, and in the lineage leading from node 5 to *semiflava* and Costa Rican *aequinoctialis*. (A maximum-parsimony
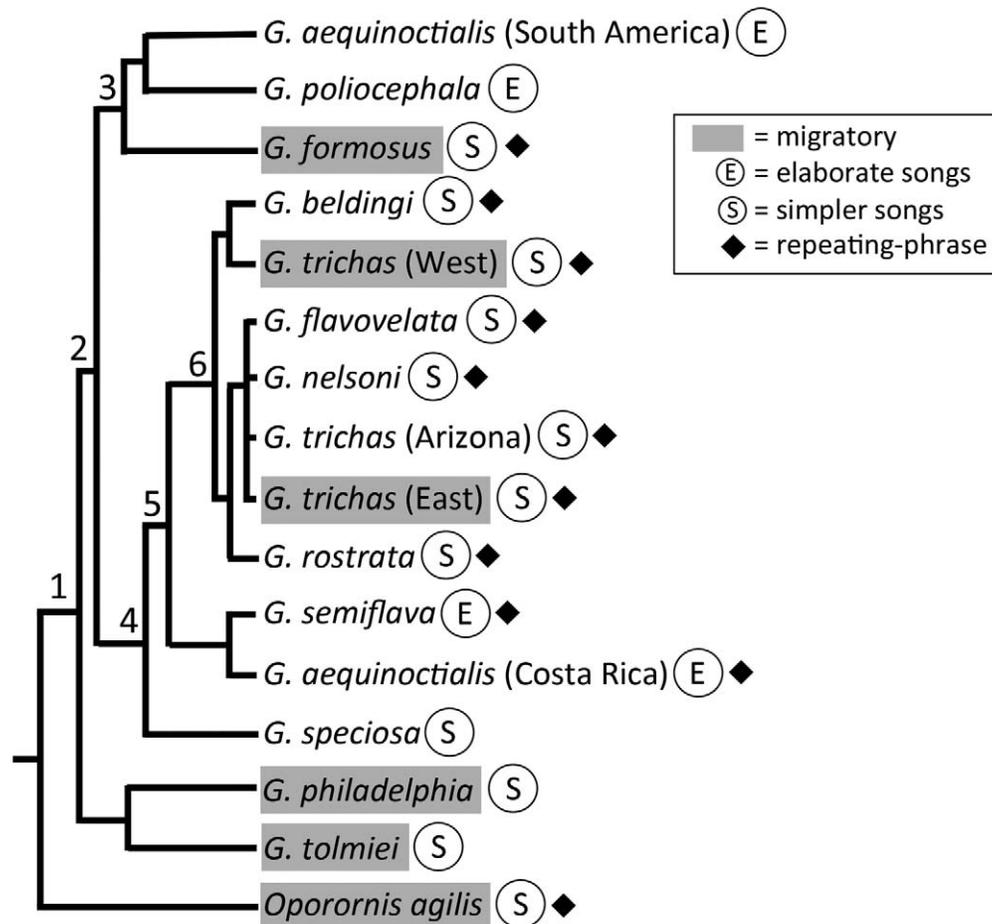
FIGURE 4. Phylogeny of *Geothlypis* (redrawn from Escalante et al. 2009), showing how migratory behavior, elaborate and simpler songs, and repeating-phrase song structure are distributed across the genus. The eastern and western lineages of *G. trichas*, marked as migratory, also contain some sedentary populations. *Geothlypis formosa*, marked as having repeating-phrase songs, includes some individuals that lack them.

reconstruction based on the continuous values of PC1 scores yielded an essentially identical result.)

Although a repeating-phrase song structure is present in both the outgroup *O. agilis* and many *Geothlypis* species, reconstruction analysis suggests that it is unlikely to have been present in the common ancestor of *Geothlypis* (proportional likelihood = 0.36; Figure 5C). The analysis further suggests that the most likely scenario is that the repeating-phrase structure arose twice, in the lineage leading to *formosa* and at node 5, after which it was retained in all descendant species.

DISCUSSION

Phylogenetic Signal and Evolutionary Lability
Many *Geothlypis* species have songs whose structure is characterized by a pattern of repeated multinote phrases, and the distribution of this pattern suggests that song structure carries a phylogenetic signal within the genus.

For example, the repeating-phrase structure is present in all lineages in the clade that contains all descendants of the ancestor (Figure 5C, node 5).

Similarly, song structure within the nominal species *aequinoctialis* reflects the phylogenetic status of the taxon. The phylogenetic analysis of Escalante et al. (2009) revealed that the Costa Rican lineage of nominal *aequinoctialis* is actually embedded within *semiflava* and is not closely related to South American *aequinoctialis*. In accordance with this finding, the songs of both *semiflava* and Costa Rican *aequinoctialis* have the repeating-phrase structure, but the songs of South American *aequinoctialis* do not.

Nonetheless, song characteristics are not always phylogenetically informative in *Geothlypis*. Consider, for example, the 3 species in the small clade that includes *poliocephala*, South American *aequinoctialis*, and *formosa*. Despite the small evolutionary distance separating these 3 taxa, 1 species (*poliocephala*) has
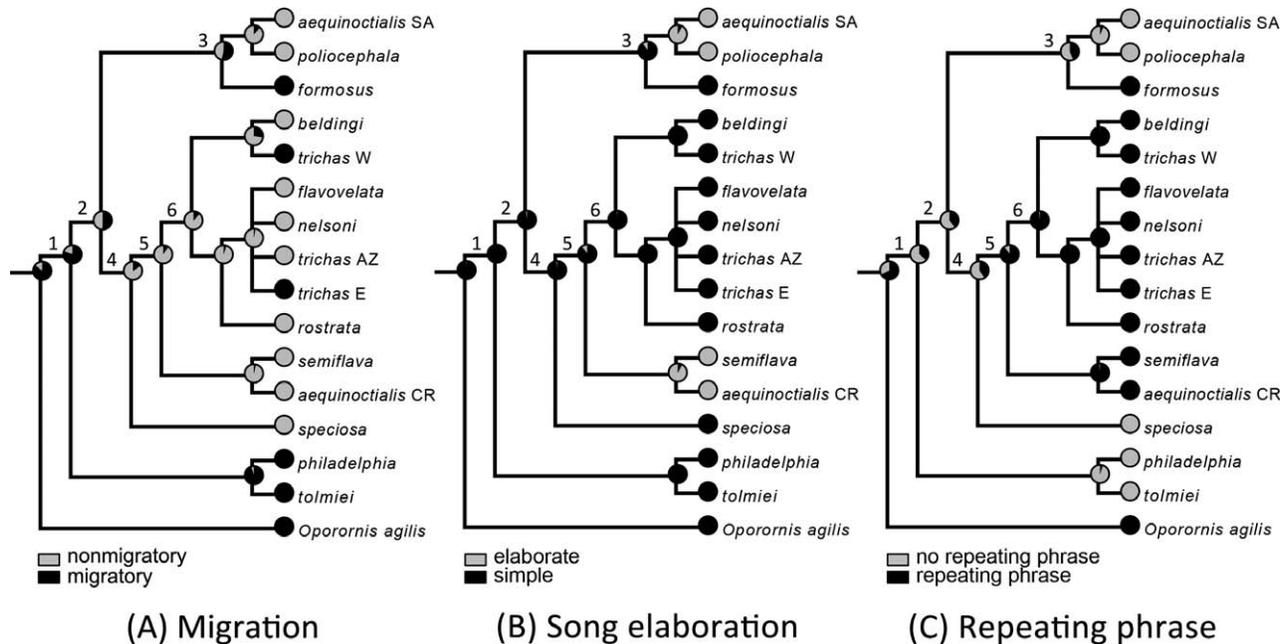
FIGURE 5. Ancestral state reconstructions of 3 characters. Reconstructions are shown for presence versus absence of (**A**) migratory behavior, (**B**) elaborate songs, and (**C**) songs consisting of one or more repeating phrases. Pie charts at nodes represent relative maximum-likelihood support for each character state.

multisong repertoires of elaborate songs that lack repeating phrases, another (South American *aequinoctialis*) uses elaborate non-repeating-phrase songs in which each bird sings only a single type, and the third (*formosa*) sings simple repeating-phrase songs, 1 type per bird. Thus, even song features that seem generally conserved within *Geothlypis*, such as the repeating-phrase structure and single-type repertoires, can differ sharply between closely related taxa. These differences demonstrate how remarkably evolutionarily labile oscine birdsong can be.

Migration and Song Elaboration

All the *Geothlypis* species that retained the clade's ancestral migratory behavior also retained simple songs, as did the lineages that secondarily re-evolved migratory behavior. Conversely, the elaborate songs of the nonmigratory *semiflava*, *aequinoctialis*, and *poliocephala* suggest that song elaboration in these species arose (on 2 independent occasions) in association with their evolutionary loss of migratory behavior. Although the contrast between the migratory species and the nonmigratory Central–South American ones suggests that sedentary behavior is linked to elaborate singing in *Geothlypis*, this relationship does not hold throughout the genus. In the clade that Escalante et al. (2009) designated "core *Geothlypis*" (i.e. all species descended from the ancestor at node 4 in Figure 4), all the species that lost migratory behavior nonetheless retained simple songs.

Geothlypis contains no examples of species that retained or regained migratory behavior and also gained elaborate songs, so song variation in the genus is not consistent with the hypothesis that migratory behavior fosters song elaboration. However, the group of *Geothlypis* species that lost migratory behavior includes both species that have gained elaborate songs and species that have not, so the data are also inconsistent with the hypothesis that sedentary behavior generally leads to more elaborate singing.

Although no study of a single genus can reveal whether a broad pattern exists across all songbird species, the lack of a pattern in *Geothlypis* is reflected by the mixed results of comparable studies. Some of these studies have found that migration is associated with elaboration. For example, Irwin (2000) discovered that songs of the Greenish Warbler (*Phylloscopus trochiloides*) become increasingly elaborate with increased latitude, and Kroodsma et al. (2001) found that the sedentary Mérida Wren (*Cistothorus meridae*) has song repertoires much smaller than those used in migratory populations of congeneric Marsh Wrens (*C. palustris*) or Sedge Wrens (*C. platensis*). Other investigations, however, have found that migratory birds do not necessarily sing more elaborate songs. For example, song elaboration does not differ between migratory and nonmigratory populations of Marsh Wrens (Kroodsma and Verner 1987) or Sedge Wrens (Kroodsma et al. 1999). In Eastern Towhees (*Pipilo erythrophthalmus*) and Song Sparrows (*Melospiza melodia*), the singing of migratory

populations is less elaborate than the singing of nonmigratory populations (Ewert and Kroodsma 1994, Peters et al. 2000).

Overall, the available data do not seem to permit any confident generalization about the relationship between song elaboration and migration. Instead, the data suggest that it is not yet time to reject the null hypothesis that there is no general and predictable casual relationship between these variables, nor to reject the more general null hypothesis that among-species variation in song elaboration arises through stochastic processes.

ACKNOWLEDGMENTS

LITERATURE CITED

Bentley, R. A., M. W. Hahn, and S. J. Shennan (2004). Random drift and culture change. Proceedings of the Royal Society of London, Series B 271:1443–1450.

Bioacoustics Research Program (2008). Raven Pro: Interactive Sound Analysis Software, version 1.3. Cornell Lab of Ornithology. http://www.birds.cornell.edu/raven

Bolus, R. T. (2014). Geographic variation in songs of the Common Yellowthroat. The Auk: Ornithological Advances 131:175–185.

Borror, D. J. (1967). Songs of the yellowthroat. Living Bird 6:141–161.

Botero, C. A., N. J. Boogert, S. L. Vehrencamp, and I. J. Lovette (2009). Climatic patterns predict the elaboration of song displays in mockingbirds. Current Biology 19:1151–1155.

Byers, B. E., K. L. Belinsky, and R. A. Bentley (2010). Independent cultural evolution of two song traditions in the Chestnut-sided Warbler. American Naturalist 176:476–489.

Byers, B. E., and D. E. Kroodsma (2009). Female mate choice and songbird song repertoires. Animal Behaviour 77:13–22.

Cardoso, G. C., Y. Hu, and P. G. Mota (2012). Birdsong, sexual selection, and the flawed taxonomy of canaries, goldfinches and allies. Animal Behaviour 84:111–119.

Catchpole, C. K. (1982). The evolution of bird sounds in relation to mating and spacing behavior. In Acoustic Communication in Birds (D. E. Kroodsma and E. H. Miller, Editors). Academic Press, New York, NY, USA.

Curson, J., D. Quinn, and D. Beadle (1994). Warblers of the Americas. Houghton Mifflin, New York, NY, USA.

Dunn, J., and K. Garrett (1997). A Field Guide to Warblers of North America. Houghton Mifflin, New York, NY, USA.

Escalante, P., L. Márquez-Valdelamar, P. de la Torre, J. P. Laclette, and J. Klicka (2009). Evolutionary history of a prominent North American warbler clade: The Oporornis–Geothlypis complex. Molecular Phylogenetics and Evolution 53:668–678.

Ewert, D. N., and D. E. Kroodsma (1994). Song sharing and repertoires among migratory and resident Rufous-sided Towhees. The Condor 96:190–196.

Gaston, K. J., and T. M. Blackburn (1996). Global scale macroecology: Interactions between population size, geographic range size and body size in the Anseriformes. Journal of Animal Ecology 65:701–714.

Greig, E. I., J. J. Price, and S. Pruett-Jones (2013). Song evolution in Maluridae: Influences of natural and sexual selection on acoustic structure. Emu 113:270–281.

Griscom, L., and A. Sprunt, Jr. (1957). The Warblers of America. Devin-Adair, New York, NY, USA.

Johnson, C. N. (1998). Rarity in the tropics: Latitudinal gradients in distribution and abundance in Australian mammals. Journal of Animal Ecology 67:689–698.

Irwin, D. E. (2000). Song variation in an avian ring species. Evolution 54:998–1010.

Kroodsma, D. E. (1999). Making ecological sense of song development in songbirds. In The Design of Animal Communication (M. D. Hauser and M. Konishi, Editors). MIT Press, Cambridge, MA, USA.

Kroodsma, D. E., J. Sanchez, D. W. Stemple, E. Goodwin, M. L. da Silva, and J. M. E. Vielliard (1999). Sedentary life style of Neotropical sedge wrens promotes song imitation. Animal Behaviour 57:855–863.

Kroodsma, D. E., and J. Verner (1987). Use of song repertoires among Marsh Wren populations. The Auk 104:63–72.

Kroodsma, D. E., K. Wilda, V. Salas, and R. Muradian (2001). Song variation among Cistothorus wrens, with a focus on the Mérida Wren. The Condor 103:855–860.

Lynch, A. (1996). The population memetics of birdsong. In Ecology and Evolution of Acoustic Communication in Birds (D. E. Kroodsma and E. H. Miller, Editors). Cornell University Press, Ithaca, NY, USA.

Maddison, W. P., and D. R. Maddison (2011). Mesquite: A Modular System for Evolutionary Analysis, version 2.75. http://mesquiteproject.org

Mahler, B., and D. Gil (2009). The evolution of song in the Phylloscopus leaf warblers (Aves: Sylviidae): A tale of sexual selection, habitat adaptation, and morphological constraints. Advances in the Study of Behavior 40:35–66.

McDonald, M. V. (2013). Kentucky Warbler. In Birds of North America Online (A. Poole, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA.

Medina, I., and C. D. Francis (2012). Environmental variability and acoustic signals: A multi-level approach in songbirds. Biology Letters 8:928–931.

Mountjoy, J., and D. W. Leger (2001). Vireo song repertoires and migratory distance: Three sexual selection hypotheses fail to explain the correlation. Behavioral Ecology 12:98–102.

Peters, S., W. A. Searcy, M. D. Beecher, and S. Nowicki (2000). Geographic variation in the organization of Song Sparrow repertoires. The Auk 117:936–942.

Pitocchelli, J. (1988). Character variation in the Oporornis philadelphia–tolmiei complex. Ph.D. dissertation, City University of New York, New York City, NY, USA.

Pitocchelli, J. (1990). Plumage, morphometric, and song variation in Mourning (Oporornis philadelphia) and MacGillivray's (O. tolmiei) warblers. The Auk 107:161–171.

Pitocchelli, J. (2011). Macrogeographic variation in the song of the Mourning Warbler (Oporornis philadelphia). Canadian Journal of Zoology 89:1027–1040.

R Development Core Team (2014). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

Read, A. F., and D. M. Weary (1992). The evolution of bird song: Comparative analyses. Philosophical Transactions of the Royal Society of London, Series B 338:165–187.

Ritchison, G. (1995). Characteristics, use and possible functions of the perch songs and chatter calls of male Common Yellowthroats. The Condor 97:27–38.

Soma, M., and L. Z. Garamszegi (2011). Rethinking birdsong evolution: Meta-analysis of the relationship between song complexity and reproductive success. Behavioral Ecology 22: 363–371.

Spector, D. A. (1992). Wood-warbler song systems: A review of paruline singing behaviors. Current Ornithology 9:199–238.

Symonds, M. R. E., L. Christidis, and C. N. Johnson (2006). Latitudinal gradients in abundance, and the causes of rarity in the tropics: A test using Australian honeyeaters (Aves: Meliphagidae). Oecologia 149:406–417.

Terborgh, J., S. K. Robinson, T. A. Parker III, C. A. Munn, and N. Pierpont (1990). Structure and organization of an Amazonian forest bird community. Ecological Monographs 60:213–238.

Tibshirani, R., G. Walther, and T. Hastie (2001). Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society, Series B 63:411–423.

Tsipoura, N., and E. S. Morton (1988). Song-type distribution in a population of Kentucky Warblers. Wilson Bulletin 100:9–16.

Weir, J. T., and D. Wheatcroft (2011). A latitudinal gradient in rates of evolution of avian syllable diversity and song length. Proceedings of the Royal Society of London, Series B 278: 1713–1720.

Wiley, R. H. (2000). A new sense of the complexities of bird song. The Auk 117:861–868.

Winger, B. M., I. J. Lovette, and D. W. Winkler (2012). Ancestry and evolution of seasonal migration in the Parulidae. Proceedings of the Royal Society of London, Series B 279: 610–618.

Wunderle, J. M., Jr. (1979). Components of song used for species recognition in the Common Yellowthroat. Animal Behaviour 27:982–996.