BioOne COMPLETE

Geographic variation in call structure, likelihood, and call-song associations across subspecies boundaries, migratory patterns, and habitat types in the Marsh Wren (Cistothorus palustris)

Authors: Luttrell, Sarah A. M., and Lohr, Bernard

RESEARCH ARTICLE

Geographic variation in call structure, likelihood, and call-song associations across subspecies boundaries, migratory patterns, and habitat types in the Marsh Wren (*Cistothorus palustris*)

Sarah A. M. Luttrell* and Bernard Lohr

Department of Biological Sciences, University of Maryland Baltimore County, Baltimore, Maryland, USA
* Corresponding author: luttrell.sa@gmail.com

ABSTRACT
Geographic variation in acoustic signals can be important in species divergence, especially the maintenance of prezygotic barriers to gene flow. Furthermore, selective pressures on acoustic signals likely vary both across geographic distances and among vocalizations used in different behavioral contexts. We described the call repertoire of 5 subspecies of Marsh Wren (*Cistothorus palustris*) in eastern North America and tested for variation in both the acoustic structure and likelihood of production of each call type at 3 functional–ecological levels: subspecies identity, migratory pattern, and habitat type. Three of the 7 described call types exhibited acoustic variation best explained by either migratory pattern or habitat type. These calls were used principally in courtship–territorial patrol contexts, whereas 4 calls that did not exhibit geographic variation were used in agonistic interactions. How often a call is used in a population may be indicative of the behavior or breeding phenology associated with that vocalization. We found that 4 calls varied in how commonly they were produced among the subspecies and/or habitat types. We also described and quantified the degree to which these 5 subspecies produce calls in association with song—a little reported, but possibly more widespread, behavior in birds. Marsh Wrens commonly embedded 3 call types into song in a nonrandom pattern. This behavior was more common in freshwater-marsh populations than in saltmarsh populations, and we discuss several possible functions for call-song associations. Overall, when geographic variation in call structure occurred, it was most commonly explained by differences in habitat type and, therefore, may be indicative of local adaptations that could limit gene flow across environments.

Keywords: call, *Cistothorus palustris*, geographic variation, habitat, Marsh Wren, migration, saltmarsh, subspecies

Variación geográfica en estructura de la llamada, en probabilidad y en asociaciones entre llamadas y cantos a través de los límites de las subespecies, de los patrones migratorios y de los tipos de hábitat en *Cistothorus palustris*

RESUMEN
La variación geográfica en las señales acústicas puede ser importante en la divergencia de las especies, especialmente en el mantenimiento de las barreras precigóticas al flujo génico. Más aún, las presiones selectivas sobre las señales acústicas probablemente varían tanto a través de distancias geográficas como entre las vocalizaciones usadas en diferentes contextos de comportamiento. Describimos el repertorio de llamadas de cinco subespecies de *Cistothorus palustris* en el este de América del Norte y evaluamos la variación tanto en la estructura acústica como en la probabilidad de producción de cada tipo de llamada a tres niveles funcionales/ecológicos: identidad de la subespecie, patrón migratorio y tipo de hábitat. Tres de los siete tipos de llamadas que describimos mostraron una forma de variación acústica que pudo ser explicada de mejor manera ya sea por el patrón migratorio o por el tipo de hábitat. Estas llamadas fueron usadas principalmente en contextos de cortejo / patrulla territorial, mientras que cuatro llamadas que no mostraron variación geográfica fueron usadas en interacciones agonísticas. La frecuencia con que se usa una llamada en una población puede ser un indicador del comportamiento o de la fenología reproductiva asociada con esa vocalización. Encontramos que cuatro llamadas variaron según la frecuencia con que fueron producidas entre las subespecies y/o los tipos de hábitats. Por último, describimos y cuantificamos el grado al cual todas las subespecies de *C. palustris* que estudiamos producen llamadas asociadas al canto; un comportamiento posiblemente más extendido en las aves, pero poco estudiado. *C. palustris* comúnmente incluyó tres tipos de llamadas en el canto siguiendo un patrón no azaroso. Este comportamiento fue más común en las poblaciones de marismas de agua dulce que en las poblaciones de marismas saladas, y discutimos varias posibles funciones de las asociaciones entre llamadas y cantos. En general, cuando hubo variación geográfica en la estructura de la llamada, fue generalmente explicada por diferencias en el tipo de hábitat, y por ende puede ser un indicador de las adaptaciones locales que podrían limitar el flujo génico entre ambientes.

INTRODUCTION

Acoustic signals can play an important role in the divergence of populations and the reinforcement of past divergence events. Geographic variation in acoustic signals has been demonstrated in a wide variety of taxa, including insects, anurans, fishes, birds, and mammals (e.g., Baptista 1977, Ryan and Wilczynski 1991, Wielgart and Whitehead 1997, Irwin 2000, Parmentier et al. 2005, Stephan and Zuberbühler 2008, Marshall et al. 2011, Jiang et al. 2015). When acoustic signals play a role in mate selection, they have been shown to act as strong prezygotic barriers to gene flow (e.g., Ryan and Wilczynski 1991, Wells and Henry 1992, Mendelson and Shaw 2002, Danner et al. 2011) and can have demonstrated fitness costs to hybridization among individuals with heterotypical vocalizations (e.g., Grant and Grant 1996, Snowberg and Benkman 2007).

A great deal of what we know about geographic variation in acoustic signals has come from decades of study on birdsong and song dialects (reviewed in Podos and Warren 2007). Birdsong plays a well-known role in mate attraction and territory defense (Catchpole and Slater 2008). However, song is only a small portion of the repertoire of avian vocal signals (Krebs and Kroodsma 1980). Signaling behaviors in general are under myriad selective pressures that may vary in their strength and direction across a geographic range, including differences in the acoustic environment (Morton 1975, Nicholls et al. 2006), signaler physiology or morphology (Bertelli and Tubaro 2002, Ballentine 2006), genetic drift (Laiolo et al. 2001a, Nicholls et al. 2006, Roach and Phillmore 2017), cultural drift (Soha et al. 2004, Roach and Phillmore 2017), or sexual selection (Irwin et al. 2001). Similarly, within an individual, different vocalizations in the repertoire may be under different types or intensities of selection (Laiolo et al. 2001a, Sturge et al. 2016). In contrast to most songs, calls are used in broader functional contexts (not largely mate attraction or territory defense), and they are typically shorter in duration and less acoustically complex than song (Marler 2004a). Furthermore, unlike song, which involves learning and therefore includes a cultural component in its vertical and horizontal transmission within oscine passerines ("songbirds"), most calls are thought to be innate, though there are now numerous exceptions to this generality (see Groth 1993, Ficken and Popp 1995, Greenlaw et al. 1998, Hughes et al. 1998, Riebel and Slater 1998). Studying calls alongside song therefore affords the opportunity to understand patterns of divergence in functionally and acoustically distinct signals that may be subject to different types of selection. By exploring multiple signals within a species' vocal repertoire across a wide geographic area, we can build on the powerful framework of song dialects to create a more complete view of the evolution of vocal signal repertoires.

The Marsh Wren is a well-developed model for song learning and has an impressive and well-described song repertoire (Verner 1975, Kroodsma and Canady 1985, Kroodsma and Verner 1987, Luttrell et al. 2016). Although several early naturalists described Marsh Wren calling behavior anecdotally, there has been no comprehensive description of the call repertoire in this species (Townsend 1905, Allen 1923, Welter 1935, Verner 1975). Marsh Wrens comprise a species complex with 14 current subspecies (Kroodsma and Verner 2013) that vary in morphology, migratory pattern, and habitat type. Further, the Marsh Wren may represent 2 cryptic species divided into eastern and western groups in North America (Kroodsma and Canady 1985, Hebert et al. 2004). Here, we explore in detail the call repertoire of only the eastern group. We recorded 5 subspecies of Marsh Wren in eastern North America, including nonmigratory, partially migratory, and fully migratory populations that are endemic to 2 different habitat types, tidal saltmarsh and freshwater inland marsh (Kroodsma and Verner 2013). Partially migratory subspecies are short-distance, non-obligatory migrants that likely migrate along suitable habitat corridors; the fully migratory subspecies is an obligatory migrant that likely migrates over unsuitable habitat.

This study has 3 main goals. First, we describe the acoustic structure and behavioral context of the Marsh Wren call repertoire in qualitative and quantitative detail. We focus on breeding-season calls, because several Marsh Wren subspecies are sympatric during the nonbreeding season and discrimination among subspecies may be difficult during this time. Second, we test for variation in acoustic structure of each described call type at 3 functional–ecological levels: subspecies identity, migratory pattern, and habitat type. Third, we provide preliminary evidence of differences in the likelihood of call types being produced at each of the same 3 functional–ecological levels, and discuss the potential importance of those differences for inferring underlying behavioral and ecological causes.

Each of the functional–ecological categories we explore is likely to support a different regime of selective pressures. At the level of subspecies identity, we expect few vocal differences, especially if calls are innate, given that subspecies of Marsh Wren are likely quite young geologically. The current geographic range of suitable Marsh Wren habitat stabilized only in the past 5,000–21,000 yr (Malamud-Roam et al. 2006). Despite this recent range
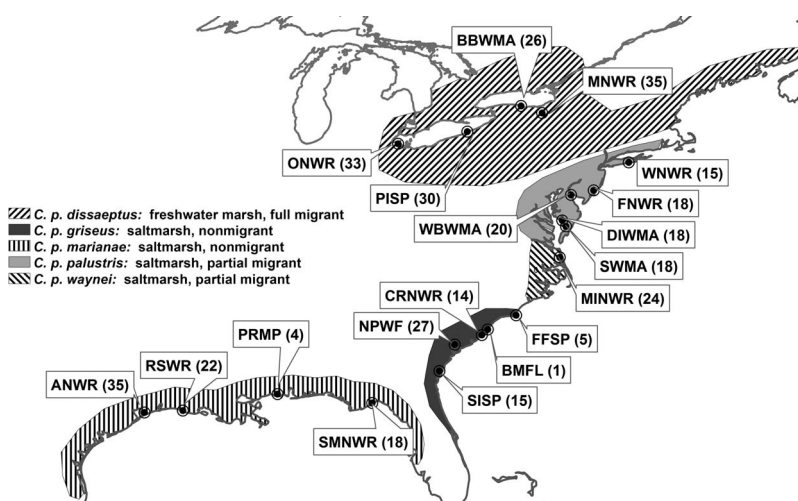
FIGURE 1. Distributions of 5 Marsh Wren (*Cistothorus palustris*) subspecies in eastern North America, based on descriptions in Kroodsma and Verner (2013). Sample sites are abbreviated as follows, clockwise from top left (with total sample size at each site in parentheses): ONWR = Ottawa National Wildlife Refuge, Ohio; PISP = Presque Isle State Park, Pennsylvania; BBWMA = Braddock Bay Wildlife Management Area, New York; MNWR = Montezuma National Wildlife Refuge, New York; WNWR = Wertheim National Wildlife Refuge, New York; FNWR = Edwin B. Forsythe National Wildlife Refuge, New Jersey; WBWMA = Woodland Beach Wildlife Management Area, Delaware; DIWMA = Deal Island Wildlife Management Area, Maryland; SWMA = Saxis Wildlife Management Area, Virginia; MINWR = Mackay Island National Wildlife Refuge, North Carolina; FFSP = Fort Fisher State Recreation Area, North Carolina; BMFL = Baruch Marine Field Laboratory, University of South Carolina; CRNWR = Cape Romaine National Wildlife Refuge, South Carolina; NWPF = Nemours Plantation, Nemours Wildlife Foundation, South Carolina; SISP = Sapelo Island National Estuarine Research Reserve, Georgia; SMNWR = St. Marks National Wildlife Refuge, Florida; PRMP = Pascagoula River Marsh Preserve, Mississippi; RSWR = Rockefeller State Wildlife Refuge, Louisiana; ANWR = Anahuac National Wildlife Refuge, Texas.

stabilization, Marsh Wrens demonstrate multiple migratory patterns among different populations. Mechanistically, differences in migratory patterns result from alterations in hormone expression that may also affect breeding phenology (Ramenofsky and Wingfield 2007). Thus, populations with different migratory patterns have already accumulated key physiological differences and are more likely than subspecies within a migratory pattern to demonstrate differences in vocal characteristics. Lastly, Marsh Wrens are endemic to 2 distinct habitat types, freshwater marsh and saltmarsh. Other subspecies pairs of birds that exhibit this habitat dichotomy have marked differences in physiology and morphology related to heat dispersal (Greenberg and Danner 2012) and osmoregulation (Goldstein 2006). Differences in bill morphology related to heat dispersal directly affect vocal signal production and receiver preference in other saltmarsh endemics (Ballentine 2006, Liu et al. 2008). We expect the largest degree of vocal differences to occur at the habitat level, whether that difference is due to divergence in morphology or to selection acting on reduced fitness of ecotype-hybrids (Maley 2012).

METHODS

Sample Sites

We collected recordings from 5 of the 14 subspecies of Marsh Wren (see Figure 1). *Cistothorus palustris dissaep-*

tus (hereafter *dissaeptus*) is the only fully migratory subspecies endemic to freshwater marshes in our sample. The remaining 4 subspecies are either partial migrants or nonmigrants and are found in coastal or estuarine marshes that vary in salinity and tidal action. *Cistothorus p. palustris* (hereafter *palustris*) and *C. p. waynei* (hereafter *waynei*) are considered partially migratory. *Cistothorus p. griseus* (hereafter *griseus*) and *C. p. marianae* (hereafter *marianae*) are considered nonmigratory (Kroodsma and Verner 2013).

We recorded Marsh Wrens throughout their distribution in eastern North America at 19 sites along the Gulf of Mexico, the Atlantic Coast, and the Great Lakes region during the breeding season (April 30–July 21, 2012–2014; see Figure 1). We focus on the breeding season because although Marsh Wren subspecies are geographically separate during the breeding season, the migratory and nonmigratory populations are sympatric during the nonbreeding season. While variation in plumage may allow a wintering individual to be correctly identified to its breeding population, focusing on breeding birds eliminates error in population assignment. For Atlantic coast populations, we worked from south to north to minimize differences in breeding phenology among populations. For the Great Lakes and Gulf Coast birds, we collected recordings beginning on June 1 in each year to ensure that we recorded local birds on established territories. The

number of individuals recorded at a site varied from 1 to 35, averaging 18.6 individuals site^{-1}.

Field Recordings

We recorded individual Marsh Wrens on territory from sunrise (earliest time: 0525 hours EST) until midday (latest time: 1245 hours), using directional microphones (Sennheiser ME67 with K6 power module) and solid-state digital recorders (Marantz PMD 660) set to 24-bit sample depth and 48 kHz sample rate. Birds were not color banded for individual identification, but each territory was identified by GPS location and adjacent territories were recorded on the same day, and simultaneously when possible, by different observers to ensure recording of distinct individuals one time each. Recording sessions ranged from 20 to 60 min. We dictated the behavior of the bird during each observation onto the recording and included information about the presence of any conspecifics and whether we could locate an active nest on the territory. Sex was either assessed via behavior (female Marsh Wrens are not known to sing; Kroodsma and Verner 2013) or was already known for previously captured birds, which we had fit with a federal band as part of a related study, at which time we ascertained sex via visual inspection of cloacal protuberance or brood patch.

Acoustic Description of Song and Call Types

We annotated entire recordings—including songs, calls, and behavioral observations—for each individual, using Raven Pro 1.4 (Cornell Lab of Ornithology, Ithaca, New York, USA). From the recordings, we identified 8 qualitatively distinct vocalizations: songs (Figure 2) and 7 call types (hereafter "buzz," "chuck," "churr," "rattle," "scream," "trill," and "twitter"; Figure 3). We defined a call bout as a series of vocalizations separated by >0.5 s of silence. A preliminary analysis in 3 subspecies examining the spacing (time) between all vocal elements (i.e. traces on a spectrogram; excluding notes within songs) showed that most traces were within 100 ms of one another, representing notes within syllables. There was a second major drop-off between 300 and 500 ms with a long tail that continued out past 5 s, which we identified as intervals between call bouts. This suggested that 0.5 s was a conservative estimate for identifying a break in vocal bouts that would be reasonable for an observer to quickly visualize on a spectrogram. Each call bout was extracted from the original file, named individually, and high-pass filtered to reduce background noise (buzz = 500 Hz; chuck, churr, rattle, scream = 800 Hz; trill, twitter = 1,000 Hz). We standardized the amplitude of each recording and removed background noise between notes before making detailed measurements of each call.

We measured spectral and temporal parameters for each call type using SIGNAL/RTSD (Engineering Design 2015).
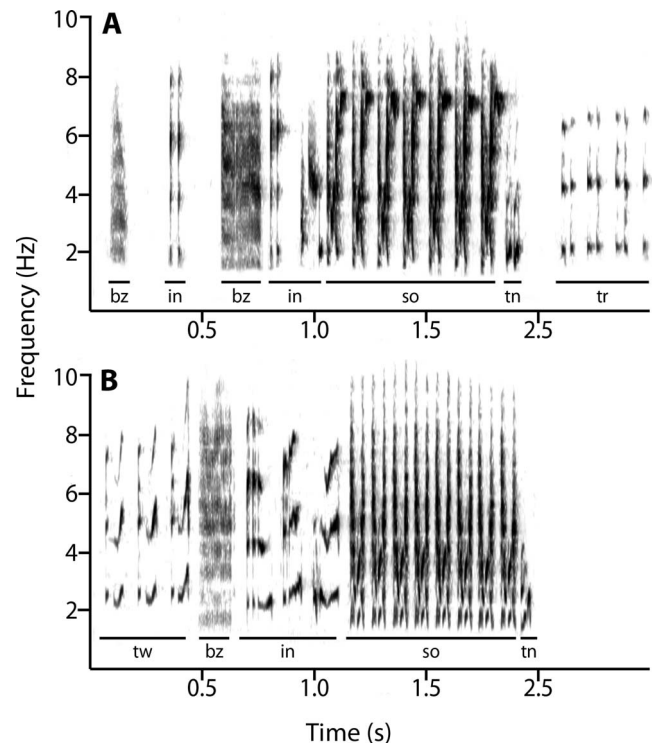

FIGURE 2. Representative spectrograms of Marsh Wren song for a freshwater-marsh bird (**A**) and a saltmarsh bird (**B**), showing typical relationships between calls and song elements. The number of call elements embedded in a song can vary between different renditions of the same song type. Elements of the vocalizations are abbreviated as follows: bz = buzz, in = introductory note, so = song, tn = terminal note, tr = trill, tw = twitter.

In keeping with common practice, we defined a note as a single trace on a spectrogram, and a syllable as a series of consistently repeating notes that occur clustered together. For each call, we took the following acoustic measurements (using a 32,768-point fast Fourier transform [FFT] for frequency measures; spectral resolution = 0.73 Hz; time measurements taken from a time waveform aligned with a 256-point FFT; temporal resolution = 5.3 ms): total duration of the call, number of notes in the call (or, in the case of the churr call, number of syllables), duration of an individual note in the calling bout (for the churr call, the duty cycle of a single note), temporal spacing between notes in the bout (hereafter "inter-note interval"), peak frequency (frequency of greatest power), bandwidth at 15 dB down from peak, and lower and upper frequency limits 15 dB down from peak. For the buzz call, which consisted of a sequence of amplitude modulations (hereafter "pulses") separated by silence, we also measured the duty cycle of a single pulse, as well as the average depth of amplitude modulation in the pulses. To measure depth of amplitude modulation, we tracked the amplitude envelope of the signal with a 2 ms exponential decay. We then
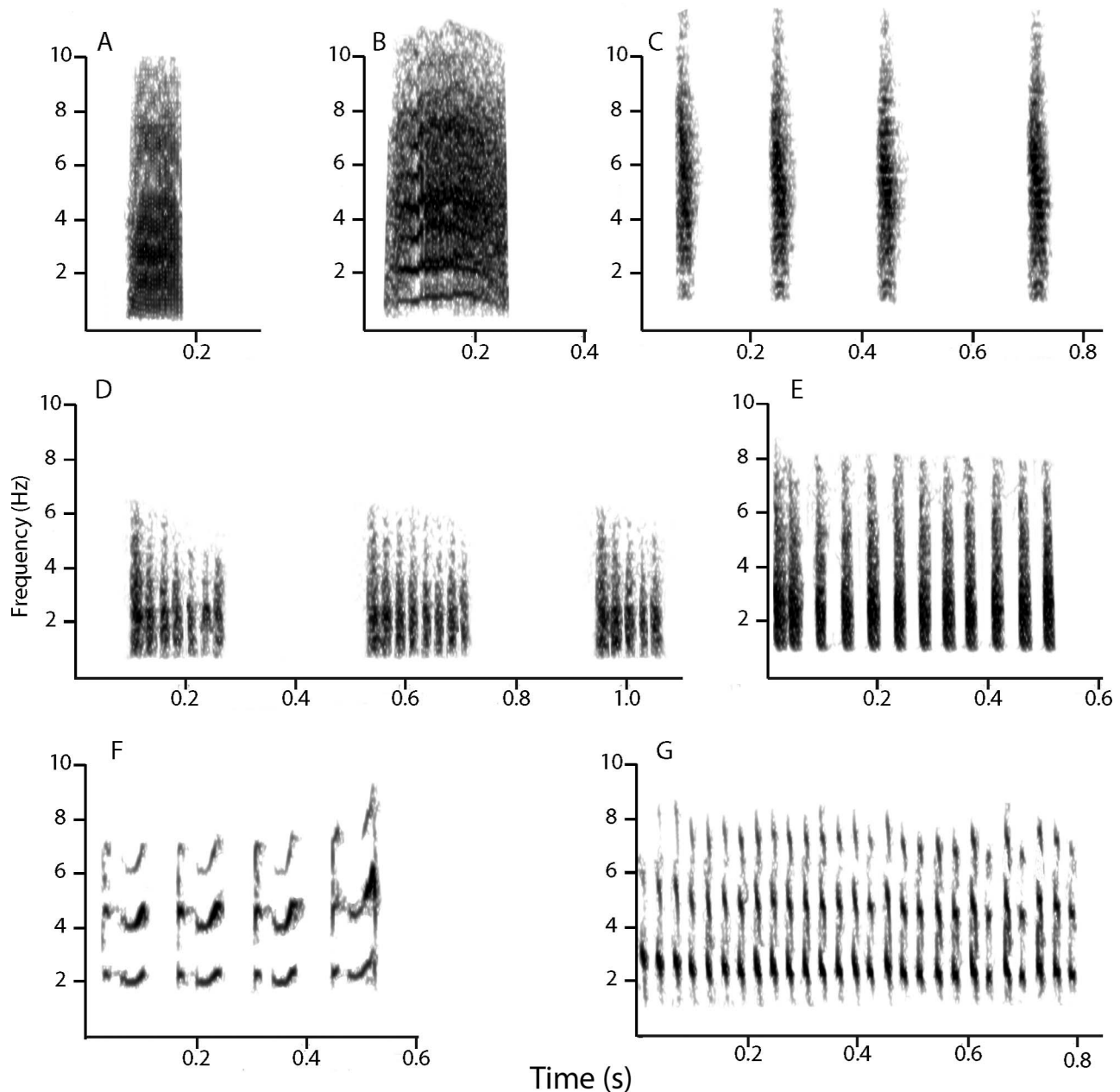
FIGURE 3. Spectrograms of call types identified in eastern Marsh Wrens (512 pt FFT, 135 Hz frequency resolution, 5.3 ms time resolution). (**A**) Buzz call. (**B**) Scream call. (**C**) Chuck call. (**D**) Churr call. (**E**) Rattle call. (**F**) Twitter call. (**G**) Trill call. Individual renditions of a call may vary, most notably in inter-note intervals of the chuck, churr, rattle, and trill calls.

calculated a long-term contour (running average width in milliseconds = duration of call divided by 10) of the envelope and subtracted this contour from the original amplitude envelope, removing any remaining DC offset, to create a modulated signal with an average value of zero. We then calculated the standard error of the resulting envelope over the duration of the call.

We measured 1–20 high-quality call exemplars for each individual (when we recorded >20 calls of a given type, we measured the 20 highest-quality calls). Within an individual, we averaged all measures of a given call type and retained for subsequent analyses only individuals with ≥3 examples of a call type. Subsequent analyses were then performed on the individual average value of a call variable for those individuals with ≥3 examples of the call type.

In early inspections of spectrograms of 2 call types (the trill and twitter), the frequency and bandwidth of notes appeared to change throughout the duration of the call. To

quantify this pattern, we measured peak frequency, bandwidth, lower and upper frequency limits, note duration, and inter-note intervals for 3 groups of 4 consecutive notes representing the beginning, middle, and end of each call. Then we used repeated-measures analysis of variance (ANOVA) to test for variation in each measure across call duration. We corrected for multiple testing using a false-discovery-rate adjustment. The false discovery rate controls the proportion of type 1 errors accepted and is more powerful but less stringent than the Bonferroni correction. We ranked P values from smallest to largest and then accepted as significant only P values that exceeded the value of family-wise alpha*rank/total number of tests ($\alpha = 0.05$; Benjamini-Hochberg 1995). For each variable that was not significantly different among positions in the call, we chose only a single section of the call for further analyses.

Quantitative Discrimination of Call Types

We based our initial identification of call types on visual inspection of spectrograms and observations of Marsh Wren behavior. While most individual calls were visually and acoustically distinct and easily classified, several represented intermediate forms. We used multivariate tests and t-tests to analyze the chuck, churr, and rattle calls (broadband, short duration) and the trill and twitter calls (narrowband, longer duration) in greater detail to determine whether these call classes were statistically distinct. For these analyses, we used individual average measurements for each call type (any individual with ≥ 3 instances of a call type) regardless of subspecies designation.

Chuck, churr, rattle comparison. The chuck, churr, and rattle are all broadband signals with rapid onset and offset and are produced as a sequence of notes (chuck, churr, rattle) or syllables (churr) (Figure 3C–3E). To establish the distinctiveness of these calls, we employed a linear-model-based categorization method. At this level of analysis, many of our acoustic measurements did not conform to multivariate normality, so we used a multinomial logistic regression with a logit link function (a nonparametric equivalent to linear discriminant analysis) to test for differences among call types (SPSS 24.0; IBM, Armonk, New York, USA). We used means of original variables, rather than a composite measure such as a principal component score, in the multinomial regression in order to interpret directly how the call types compared with one another. Several of our acoustic measures were tightly correlated, such as bandwidth, peak frequency, and upper frequency limit. In these cases, we retained only one measure in our model to avoid issues of collinearity. Because call duration and number of notes in a call can vary widely within each of the call types, we did not include them in this model, focusing instead on the properties of individual notes. Our final model included the 3 most

normally distributed, non-collinear measures of individual notes: bandwidth, lower frequency limit, and note duration (or, for churr calls, pulse duration).

Trill and twitter comparison. The trill and twitter calls are narrowband calls consisting of a series of repeated notes (Figure 3F, 3G). Although the trill and twitter have visible differences in call structure based on the spectrogram, the 2 calls are used in similar contexts and are often observed with trill-like notes interspersed with twitter-like notes in calls that we categorized as "mixed." Multivariate analyses using a binomial model and logit link function to differentiate the trill and twitter failed to converge. As a consequence, we tested trill and twitter calls using paired t-tests or Wilcoxon signed-rank tests (depending on whether the data met the assumptions of parametric tests) to determine whether any of the acoustic measurements from the 2 subjective categories were quantitatively different. We corrected significance using a false-discovery-rate error adjustment for a family-wise alpha ($\alpha = 0.05$; Benjamini-Hochberg 1995).

Behavioral Association of Call Types

For each call type, we observed 47–171 individuals with sufficient quality to describe behavior associated with that call type. In each case, we noted the sex of the caller, whether there was another Marsh Wren detected on the territory at the time of the vocalization, and whether that territory contained an active nest that we could locate. When Marsh Wrens were visible through the dense vegetation, we also noted behaviors related to the following categories: courtship (including male courtship display, copulation solicitation, or rebuffing courtship; Welter 1935); territory patrol (moving frequently around the territory without engaging in any other identifiable behavior); agonistic interactions (pursuing another Marsh Wren or being pursued by another Marsh Wren); post-agonistic interaction (within 30 s of agonistic interaction); indeterminate pursuit (an adult bird following and closely associating with another Marsh Wren, but without engaging in any courtship or agonistic behaviors); nonspecific interaction (interaction with no clearly delineated function); interaction with young (including provisioning nestlings); preening; foraging; nest building; flight; a heterospecific interaction; and reaction related to researcher presence.

Geographic Call Variation

For each functional–ecological level of analysis (subspecies, migratory pattern, or habitat type), we performed a linear discriminant analysis (LDA) on the acoustic properties of each of the 7 call types (lda function, MASS package; Venables and Ripley 2002). LDA is a multivariate technique that fits orthogonal, linear functions from a series of predictor variables to divide individuals into

assigned categorical groups with the least amount of error. It is a useful technique for determining whether proposed groups can be reliably identified on the basis of the simultaneous assessment of a small number of measureable traits. Various thresholds have been proposed regarding what constitutes a reliable, correct classification of groups, all based on some variation of a "75% rule" (75% of population A lies outside the range of population B; reviewed in Patten and Unitt 2002). Here, rather than designate a number as a critical threshold for diagnosability, we discuss instead the relative ability of each of our functional–ecological levels of analysis to explain variation in our data. All LDAs were performed using jackknife, leave-one-out cross-validation and group sizes proportional to original group membership. In each test, only variables that met the following assumptions of an LDA were included in the model: normality, equal variance, not covarying with any other variables.

Testing for variation among subspecies. We ran the LDA model with subspecies as the response variable and only the acoustic measures that conformed to the assumptions of the LDA as predictor variables. For all acoustic analyses examining subspecies' differences, we excluded *waynei* because we were only able to sample this population at a single site (24 individuals).

Testing for variation among migratory patterns. We categorized populations to migratory pattern on the basis of previous natural-history descriptions of the populations (Kroodsma and Verner 2013). We considered *dissaeptus* to be full migrants, *palustris* and *waynei* to be partial migrants, and *griseus* and *marianae* to be nonmigrants. We ran the LDA model with migratory pattern as the response variable and as many acoustic measures as conformed to the assumptions of the LDA as predictor variables.

Testing for variation among habitat types. For habitat-type analysis, all of the tidal marsh subspecies were treated together as a saltmarsh group (*griseus*, *marianae*, *palustris*, *waynei*) and compared with the freshwater-marsh subspecies (*dissaeptus*). We treated habitat type as the response variable and as many acoustic measures as conformed to the assumptions of the LDA as predictor variables.

To address the differences in sample sizes between the 2 habitat categories, we established a 95% confidence interval (CI) for the saltmarsh population's linear discriminant (LD) values and determined the number of freshwater-marsh individuals that fell outside the 95% CI. In order to establish an LD value for each individual bird, we reran the LDA for each call type without jackknifing, because the jackknifing procedure calculates probabilities of group membership, not explicit LD values. Then we bootstrapped the saltmarsh sample by sampling with replacement to simulate 1,000 new samples of saltmarsh

LD values that were equal in size to the smaller freshwater-marsh sample. We calculated the mean ± SD for each of our 1,000 bootstrapped samples, and then used the mean-of-the-mean ± 1.96 times the mean-of-the-standard deviation to create a 95% CI for the saltmarsh LD values.

Call Presence–Absence

In addition to variation in acoustic qualities of calls, we were interested in the commonness of each call type within each of the 3 functional–ecological categories. We examined the commonness of call types by coding observations of individual birds as binary responses; a call was considered "present" for an individual (that individual produced the call type at least one time) or "absent" (that individual was never observed to produce the call type). Because all birds within a sample site share the same predictor categories (subspecies, migratory pattern, or habitat type), we used sample site as a unit of comparison by calculating the proportion of individuals with a call type "present" at each site (Crawley 2007).

We included data from 17 sites representing 4 subspecies. We excluded *waynei* because we sampled it only at a single site, and one sample site for *griseus* because we were able to record only a single bird at that site. Sample sizes for the remaining sites varied from 4 to 35 individuals, with an average of 20 individuals observed per location. Excluding the sites with <14 individuals did not change the outcome of the analyses. For these analyses, we were able to include all occurrences of all call types regardless of quality, including the twitter from *dissaeptus*, because we were evaluating presence or absence and not taking detailed acoustic measurements.

We evaluated whether the log likelihood of a call type being present at a site could be predicted by a fixed effect of grouping strategy (subspecies, migratory pattern, or habitat type) and a random effect of ordinal date to account for seasonality using a generalized linear mixed model with a binomial distribution and logit link function (glmer, lme4 package; Bates et al. 2015). All models were ranked against the null model of intercept and the random effect of date. We used a combination of ΔAIC_c scores and P values from a likelihood ratio test to determine the model that best explained differences in probability of call observation ($AIC_c()$, qpcR package, Spiess 2014; anova(), stats package, R Core Team 2014). The best-fit model was considered a model that was significantly different from the null model under a likelihood ratio test and either ≥2 ΔAIC_c values better than the next best model (in the case that the top 2 models were ≤2 ΔAIC_c values apart, the model with the fewest number of parameters was selected). Call occurrence was coded as either "present" or "absent" for each bird, and individual responses were combined to find an overall probability of a call being present at each sampling location. For each call type, we ranked 4 binomial

models using the probability of observing the call (response variable) at a given sampling location and predictor variables of either a null random intercept, subspecies designation, migratory pattern designation, or habitat-type designation while controlling for the random effect of sample date. When the highest-ranking model was significantly different from the null model, we performed a post hoc analysis using least-squares-means Tukey's pairwise independent contrasts (lsmeans package; Lenth 2016) to determine which comparisons contributed to the difference while accounting for unbalanced samples and the inclusion of a random effect in the model. From the post hoc analysis, we found predictions of the group averages with 95% CIs and odds ratios (magnitude of difference) for the pairwise contrasts.

Calls Embedded in Songs

Often, Marsh Wren songs have at least one embedded call element, including the buzz, trill, and/or twitter. This phenomenon was quite common in all of our observations, but since call elements may play an important role in courtship in the Marsh Wren, we wanted to know whether the inclusion of embedded call elements in songs varied at any of our 3 levels of functional–ecological analysis: subspecies, migratory pattern, or habitat type. We considered any vocalization with one or more calls embedded in a song as a subcategory of song, "call-song." The commonness of a call-song can vary in 2 ways. First, some functional–ecological groups might be more or less likely to produce call-songs. We analyzed whether the probability of a call-song's presence at a sample site varied at each functional–ecological level using the "call presence–absence" model-ranking method described above. Second, individual birds within a site might be more or less likely to produce call-songs. To address whether the proportion of songs vs. call-songs that an individual produced varied within vs. among functional–ecological groups, we employed an ANOVA-like framework. We limited our analysis to include only individuals for which we recorded ≥30 songs. For each individual, we calculated the proportion of total songs recorded that contained embedded call elements. The distribution of proportions was non-normal, could not be transformed to approximate normality, and was bounded from 0 to 1. As a result, we used a nonparametric Kruskal-Wallis test to compare groups according to rank rather than raw proportion values. Lastly, in order to balance the sample design at the subspecies level, we randomly sampled 50 *dissaeptus* individuals to include in the analysis. Sample sizes for the subspecies analysis were as follows: *dissaeptus* = 50, *griseus* = 22, *marianae* = 50, and *palustris* = 44. All individuals with ≥30 songs were retained for the migratory-pattern and habitat-type analyses. Sample sizes for the migratory-pattern analysis

were as follows: full migrants = 98, partial migrants = 54, and nonmigrants = 72. Samples sizes for the habitat-type analysis were as follows: freshwater-marsh = 98 and saltmarsh = 124.

RESULTS

Acoustic Description of Song and Call Types

Song. Marsh Wren songs begin with a series of introductory notes followed by a rapid trill of 4–16 syllables (groups of repeated notes) and end with 2–6 short pure-tone notes. *Cistothorus p. dissaeptus* typically produced 2 or 3 very short, broadband introductory notes (Figure 2A). The rest of the eastern subspecies produced introductory notes that consisted of a series of harmonic sweeps that increased in frequency and bandwidth (Figure 2B). Songs could be produced independently, but 85.3% of all songs observed were associated with one or more calls. If produced in the context of a song, the buzz call was added before the song or embedded in the introductory notes. Twitter calls were commonly added before the introductory notes in the saltmarsh populations, were rarely added in freshwater-marsh populations, and could transition into introductory notes of the song in some cases. The twitter call was rarely produced after the song but in that case was often followed immediately by another song. The trill call was typically added after the song was completed and, in some cases, bridged 2 songs together. The subspecies studied here have song repertoires of 30–60 song types (Kroodsma and Verner 1987, Luttrell et al. 2016).

Buzz. The buzz call is a low-amplitude, noisy sound composed of 7–59 broadband pulses separated by short periods of silence (Figure 3A and Table 1). Rarely, the buzz included some tonal harmonic elements produced at the end of the vocalization. Marsh Wrens may produce a single buzz call or a series of buzz calls that can vary in duration or intensity. We treated each individual buzz as a separate unit.

Chuck. The chuck call is a short, loud, broadband vocalization with rapid onset and offset (Figure 3C and Table 1). Chucks were produced in bouts of individual notes, though the temporal spacing of notes within a bout varied. We considered each note as an individual unit within a bout.

Churr. The churr call is a sequence of syllables that can vary in amplitude from quiet to loud, each consisting of several short, broadband notes with rapid onset and offset (Figure 3D and Table 1). Bouts varied in length from 1 to 217 syllables. The duration of the syllable could be increased or decreased by altering the number of notes within the syllable (the duty cycle of the notes was relatively invariant to the number of notes in the syllable; Table 1).

TABLE 1. Acoustic measurements of the 7 Marsh Wren call types. Grand mean ± mean SD is presented for all measures to show the overall mean for eastern Marsh Wrens. Coefficient of variation for within-individual variation is shown in parentheses. In the case of the buzz call, "note duration" represents the duty cycle of the pulse duration.

Call type	Basic unit	Minimum–maximum number of basic units	Note rate (notes s^{-1})	Note duration or *duty cycle (ms)	Call duration (s)	Bandwidth (Hz) ± SD (CV)	Peak frequency (Hz)	Low frequency (Hz)	High frequency (Hz)
Buzz	Entire call	1–1	–	5.8 ± 0.7* (7.8)	0.12 ± 0.05 (28.9)	6,362 ± 1,052 (18.8)	2,810 ± 716 (25.8)	842 ± 310 (24.9)	7,204 ± 1008 (15.1)
Chuck	Note	1–165	5.9 ± 1.3 (29.4)	26.1 ± 2.8 (9.3)	1.29 ± 1.41 (77.4)	5,772 ± 696 (10.7)	4,576 ± 635 (14.9)	1,299 ± 158 (18.9)	7,071 ± 735 (7.3)
Churr	Syllable	1–217	–	20.0 ± 2.1 (6.1)*	1.28 ± 0.80 (78.8)	2,759 ± 401 (15.5)	2,346 ± 290 (12.9)	1,086 ± 142 (19.1)	3,844 ± 417 (9.3)
Rattle	Note	2–81	2.0 ± 0.6 (39.6)	15.5 ± 2.5 (19.1)	0.98 ± 0.66 (85.1)	3,464 ± 586 (18.6)	2,917 ± 517 (20.0)	1,232 ± 214 (19.5)	4,696 ± 615 (13.8)
Scream	Entire call	1–1	–	–	0.108 ± 0.04 (41.4)	6,305 ± 803 (16.1)	4,230 ± 755 (28.2)	1,349 ± 386 (30.3)	7,654 ± 704 (11.5)
Trill	Note	3–503	21.2 ± 3.8 (18.9)	12.6 ± 2.4 (21.5)	2.32 ± 1.36 (60.5)	565 ± 143 (29.4)	2,606 ± 444 (24.4)	1,916 ± 199 (10.4)	2,481 ± 247 (10.9)
Twitter	Note	3–195	14.3 ± 2.7 (18.4)	21.9 ± 15.5 (46.8) to 36.9 ± 10.5 (50.5)	1.60 ± 0.92 (41.5)	828 ± 316 (33.4)	2,612 ± 347 (12.5)	2,089 ± 239 (9.5)	2,917 ± 467 (10.4)

Call type	Basic unit	Minimum number of syllables	Minimum number of pulses	Maximum number of syllables	Maximum number of pulses	Syllable rate (syllables s^{-1})	Syllable duration (ms)	SE of pulse depth*1,000
Buzz	Entire call	–	7	–	59	–	–	15.6 ± 3.7 (21.2)
Chuck	Note	–	–	–	–	–	–	–
Churr	Syllable	1	–	217	–	0.6 ± 0.1 (28.1)	85.0 ± 24.0	–
Rattle	Note	–	–	–	–	–	–	–
Scream	Entire call	–	–	–	–	–	–	–
Trill	Note	–	–	–	–	–	–	–
Twitter	Note	–	–	–	–	–	–	–

Rattle. The rattle call is a loud sequence of short-duration, broadband notes with rapid onset and offset that have a descending frequency sweep (Figure 3E and Table 1). It is produced as a series of individual notes.

Scream. The scream is a broadband call that contains both noisy and harmonic tonal elements (Figure 3B and Table 1). It is amplitude modulated, like the buzz, but the modulations are less regular in time and have a smaller dynamic range. The relative contribution of noise vs. tonal elements within a call varied within an individual and across bouts. It appears to be functionally and structurally analogous to the distress calls of many species of passerines (Jurisevic and Sanderson 1994).

Trill. The trill call consists of a series of short, narrowband notes with harmonics (and sometimes non-harmonic elements), produced as a sequence (Figure 3G and Table 1). Notes could be spaced evenly (single), could be pairs of notes spaced closely together with longer gaps between pairs (pairs), could be 3 quickly produced notes with longer gaps between these groups (triplets), or could be a mix of these rhythmic pattern types within a single bout. Note structure and duration were independent of rhythmic pattern. Within an individual bout, the bandwidth and peak frequency of notes was consistent over the duration of the call. Apparent small shifts in frequency during a call were not significant after correcting for multiple comparisons (Appendix Table 4). Although the trill was produced independently, it was also commonly associated with the buzz or the song. Independently produced trill calls were considered for all further behavioral and acoustic analyses.

Twitter. Like the trill, the twitter call is a series of narrowband notes with harmonics. However, the twitter most commonly consisted of a 2-note syllable (a short-duration note followed by a longer-duration note with an upward frequency sweep), although there was some variation in this pattern (Figure 3F and Table 1). Multiple measures of note frequency increased from the beginning to the middle portion of the call, then remained invariant from the middle to the end of the call (Appendix Table 4). The lower frequency limit increased by a mean of 112 Hz, the upper frequency limit increased by a mean of 157 Hz, and the peak frequency increased by a mean of 190 Hz. Subtle changes in bandwidth were not significant after corrections for multiple testing (Appendix Table 4). Like the trill, the twitter was often produced as an independent vocalization but could be produced immediately following a buzz call or immediately prior to a song. Subsequent analyses included only twitter calls produced independently of other vocalizations.

Trill–twitter mix. Twitter notes are often interspersed with trill notes to produce a "mixed" call type, which we observed 1,311 times from 129 individuals (32% of all trill- or twitter-like calls were mixed call types, and 32% of all individuals that produced trill- or twitter-like calls produced mixed call types). In this mixed call type, birds transitioned between a series of trill-like and twitter-like notes, and the number of times the note types switched back and forth varied. All subsequent measures for trill and twitter calls came only from unmixed versions of each call.

Behavioral Association of Call Types

We observed both sexes producing all call types, but for every call type one sex was more likely to produce the call than the other. Behavioral observation data are summarized in Table 2. We observed chucks, churrs, rattles, and screams more often when a conspecific was detected on the territory. The remaining calls (buzz, trill, and twitter) and song were observed in similar proportions of recordings when conspecifics were either detected or not detected on the territory.

All call types fell most commonly into 2 broad behavioral categories: territory patrol or agonistic interactions. The buzz, churr, trill, twitter, and song were all most commonly associated with territory-patrol behaviors, nest building, and courtship, and males were more likely to produce all of these calls than females (ranging from 58% to 100% of observations from known males). The chuck, rattle, and scream were most commonly observed in agonistic encounters, and the sex of the caller varied depending on the behavioral context of the calling event but was most often female or of unknown sex.

Quantitative Discrimination of Call Types

Chuck, churr, rattle comparison. The chuck, churr, and rattle were readily distinguishable on the basis of bandwidth and note duration. A multinomial logistic regression of call type using bandwidth, lower frequency limit, and note duration as predictor variables was significantly different from a null model of intercept only ($\chi^2 = 331.35$, df = 6, $P < 0.001$; Nagelkerke's $R^2 = 0.956$). The lower frequency limit did not contribute significantly to the model in a likelihood ratio test ($\chi^2 = 4.74$, df = 2, $P = 0.093$) but was left in, because it may be biologically relevant. The full model separated the chuck, churr, and rattle with a high degree of accuracy (Figure 4A). The model identified 95.8% of chucks correctly, 93.0% of churrs correctly, and 96.4% of rattles correctly. Bandwidth differed significantly between the churr and chuck calls (churrs having narrower bandwidth; Wald statistic = 11.54, df = 1, $P = 0.001$), and between the churr and rattle calls (churrs having narrower bandwidth; Wald statistic = 17.58, df = 1, $P < 0.001$), but did not differentiate between the chuck and rattle calls (Wald statistic = 3.10, df = 1, $P = 0.078$). Note duration differed between the rattle and chuck calls (rattles having shorter duration; Wald statistic = 9.63, df = 1, $P = 0.002$) and between the

TABLE 2. Behavioral observations for each Marsh Wren call type reported as a proportion of total observations for each call type. First, second, and third most common behaviors associated with each call type are in bold and are indicated with 1, 2, and 3 asterisks, respectively.

Call type	Total number of individuals observed	Sex of caller			Conspecifics detected	Active nest found on territory
		Male	Female	Unknown		
Buzz	170	0.75	0.05	0.20	0.47	0.02
Chuck	141	0.28	0.50	0.21	0.85	0.22
Churr	80	0.58	0.09	0.34	0.61	0.13
Rattle	93	0.05	0.49	0.45	0.89	0.16
Scream	47	0.04	0.23	0.72	0.96	0.13
Trill	75	0.71	0.03	0.27	0.54	0.12
Twitter	66	0.82	0.03	0.15	0.53	0.12
Song	171	1.00	0.00	0.00	0.35	0.05

	Behavior						
Call type	During courtship (male)	During courtship (female)	Territory patrol	Agonistic	After agonistic encounter	Pursuit: non-agonistic, nonsexual	Conspecific: non-agonistic, nonsexual
Buzz	***0.14	0.00	*0.35	0.04	0.00	0.04	0.06
Chuck	0.02	**0.25	0.00	*0.32	0.00	0.00	0.05
Churr	***0.15	0.00	*0.31	0.06	0.08	0.00	0.08
Rattle	0.00	0.03	0.03	*0.42	0.00	0.00	0.00
Scream	0.00	0.00	0.03	*0.79	0.03	0.00	0.00
Trill	0.07	0.00	*0.38	0.02	0.03	0.03	0.06
Twitter	***0.10	0.00	*0.44	0.06	0.00	0.03	0.04
Song	0.07	0.00	*0.61	0.05	0.01	0.03	0.00

	Behavior						
Call type	Interacting with young	Preening	Foraging	Nest building	Flight	Heterospecific interaction	Researcher presence
Buzz	0.01	0.03	0.05	**0.17	0.09	0.00	0.00
Chuck	0.16	0.04	0.06	0.06	0.00	0.04	***0.20
Churr	0.06	0.00	0.00	**0.21	0.00	0.04	0.13
Rattle	***0.11	0.00	0.00	0.05	**0.29	0.08	0.05
Scream	0.00	0.00	0.00	0.00	**0.15	0.00	0.03
Trill	0.03	0.02	0.03	***0.12	**0.18	0.01	0.00
Twitter	0.04	0.03	0.04	0.06	**0.16	0.00	0.00
Song	0.00	0.02	0.01	**0.17	0.04	0.00	0.00

rattle and churr calls (rattle having shorter note duration; Wald statistic = 22.41, df = 1, $P < 0.001$) but did not differentiate between the chuck and the churr calls (Wald statistic = 1.40, df = 1, $P = 0.236$).

Trill and twitter comparison. We found many significant differences between the trill and the twitter, mostly related to the up-sweeping pure-tone note found in the twitter that was typically the second note in a 2-note repeated pattern (absent in the trill) (Figure 4B). Using pairwise comparisons of all calls designated by eye from spectrograms as pure trill or pure twitter, we found that 13 of 16 variables were significantly different after correcting for multiple testing via the Benjamini-Hochberg false discovery rate (Bejamini-Hochberg 1995; Appendix Table 5). Only the first inter-note interval, peak frequency of the beginning of the call, and peak frequency of the middle of

the call did not differ significantly between the trill call and the twitter call.

Geographic Call Variation

Variation among subspecies. At the subspecific level as a whole, LDA did not reliably discriminate any call type. In all cases, we excluded *waynei* from our analyses because its sample size was too small in relation to the number of variables used in the LDA. Therefore, for each call type, 3 LD functions were fit to differentiate 4 subspecies based on a suite of normally distributed, noncorrelated acoustic measurements. Overall, the percentage of individuals assigned to the correct subspecies by each LDA ranged from 28% for the rattle to 65% for the buzz. Two call types, the buzz and twitter, had overall correct classifications >50%, but within-call-type correct classification among
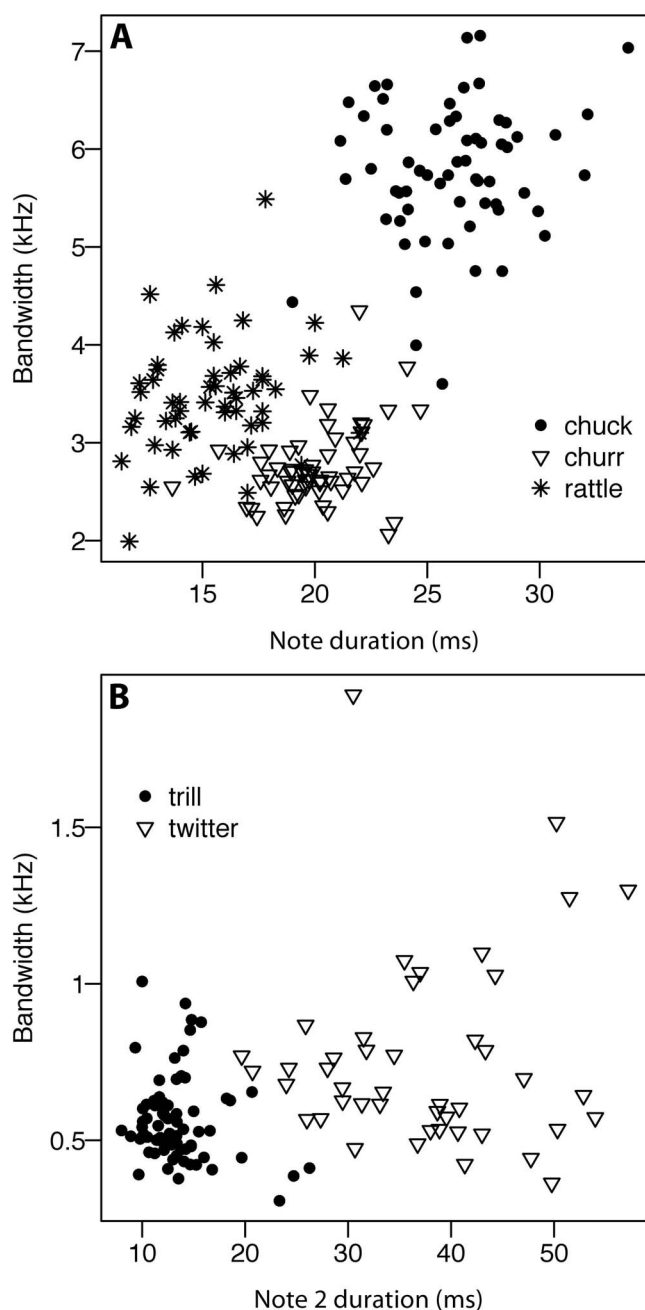
FIGURE 4. Chuck, churr, and rattle calls were reliably categorized using multinomial regression. (**A**) Bandwidth and note duration were the most important variables for discriminating among the chuck, churr, and rattle in a multinomial regression. (**B**) Bandwidth and note duration (note 2 in a repeated, stereotypical 2-note sequence) separate trill and twitter calls as well (for the trill and twitter calls, 13 spectral and temporal measures were significantly different between the 2 call types).

subspecies was highly variable (10–87% correct; Appendix Table 6). A 25% correct assignment is expected by chance alone. There was no consistent pattern across call types in which subspecies were best or worst classified.

Variation among migratory patterns. For the buzz, rattle, trill, and twitter, we found that LDA differentiated migratory pattern with greater accuracy than subspecies identity. The chuck, churr, and scream were slightly less well categorized by migratory pattern than by subspecies identity (Appendix Table 7). For each call type except the twitter, 2 LD functions were fit to differentiate 3 migratory patterns based on a suite of normally distributed, noncorrelated acoustic measurements. We had only 3 examples of poor-quality twitters from nonmigratory birds (*dissaeptus*), and so nonmigrants were excluded from the migratory-pattern model for the twitter analysis, resulting in one LD function fit to differentiate between partial and nonmigrants. The percent correct classifications for the chuck, churr, rattle, and scream were poor, ranging from 21% for the scream to 37% for the chuck and churr. A 33% correct assignment was expected by chance. The LDA models performed better for the buzz, trill, and twitter (69–71% overall correct classification; Appendix Table 7). For both the buzz and the trill, only the nonmigratory and fully migratory birds were well classified, whereas the partially migratory populations were only correctly classified 40% and 31% of the time, respectively (Appendix Table 7).

The final model for the twitter included peak frequency, lower frequency limit, bandwidth, note rate, length of the fourth note, duration of the first inter-note interval, and duration of the second inter-note interval. The most important variable of the LD function was note rate (LD coefficient = 0.50). Nonmigratory birds had a lower note rate (13.7 notes s^{-1}) than partially migratory birds (14.8 notes s^{-1}). Nonmigratory birds also had a shorter first-note duration in the 2-note series (39.5 vs. 43.0 ms) than partially migratory birds (LD coefficient = 0.07).

Variation among habitat types. Because freshwater-marsh birds (*dissaeptus*) rarely produced twitter-like calls (see above), we were unable to analyze the twitter at the habitat level. For each call type, one LD function was fit to differentiate 2 groups: freshwater-marsh birds and salt-marsh birds. For each call type, the habitat-type models performed best compared to the subspecies and migratory categorizations, classifying between 56% and 90% of individuals correctly. The chuck, rattle, and scream were still poorly categorized (56–61% correct; 50% correct classification expected by chance). However, the accuracies of all habitat-level LDAs were skewed as a result of sample-size differences (a much larger sample of saltmarsh birds as a whole compared with freshwater-marsh birds). Because our LDAs included prior probabilities of group membership based on original group sizes, they would, by design, assign more individuals to the saltmarsh category. In all calls except the buzz, we had a high percentage of correct classification for the saltmarsh birds and a low percentage of correct classification for the freshwater-marsh birds (Appendix Table 8; the buzz LDA had nearly

TABLE 3. Results of best-fit logistic models predicting the likelihood of observing each call type based on fixed effects of either a random intercept (null) model, subspecies model, migratory-pattern model, or habitat-type model each with random effect of ordinal date. Least-squares-mean post hoc analysis with Tukey contrasts demonstrates the magnitude of the likelihood difference as an odds-ratio with the first pairwise object as the reference. P value of least-squares mean indicates the significance of the pairwise comparison. Odds ratio of 1.0 is equivalent. Bold indicates significant differences.

Call type	Best-fit model	ΔAIC_c	P value for likelihood ratio test against null model (df = 1)	Subspecies pairwise comparison	Odds-ratio least-squares-mean post hoc	P value least-squares-mean post hoc
Buzz	Habitat	1.6	**0.004**	**Fresh–salt**	**3.87**	**0.001**
Chuck	Subspecies	3.7	0.017	*dissaeptus–marianae*	1.29	0.817
				griseus–dissaeptus	1.89	0.21
				griseus–marianae	2.45	0.061
				palustris–dissaeptus	**2.43**	**0.012**
				palustris–griseus	1.28	0.897
				palustris–marianae	**3.13**	**0.002**
Churr	Null	3.1	–	–	–	–
Rattle	Null	2.7	–	–	–	–
Scream	Null	1.8	–	–	–	–
Twitter	Habitat	1.8	**<<0.001**	**Salt–fresh**	**20.00**	**<0.001**
Trill	Null	2.0	–	–	–	–
Trill–twitter mix	Habitat	1.2	**<<0.001**	**Salt–fresh**	**8.20**	**<0.001**
Song-call	Null	2.6	–	–	–	–

equal sample sizes and assigned 89% of freshwater-marsh birds correctly and 82% of saltmarsh birds correctly).

To address the disparity in sample sizes, we created a 95% CI for the LD values for the saltmarsh subspecies using 1,000 bootstrap replicate samples equal in size to the freshwater sample (see Figure 5). We found that for the chuck, churr, rattle, and scream, fewer than a third of the freshwater-marsh individuals had LD values that fell outside of the saltmarsh 95% CI (15%, 32%, 29%, and 0% respectively). Therefore, although there may be acoustic differences in the chuck, churr, and rattle between habitat types, the magnitude of those differences was small.

For the buzz, although the LDA discriminated readily between the 2 habitat types, only 39% of freshwater-marsh birds fell outside the range of 95% CIs of the saltmarsh birds. The model included peak frequency, pulse duration, and SE of pulse depth. The highest weighted coefficients in the LD function for the buzz were mean pulse duration (LD coefficient = 0.338) and SE of pulse depth (LD coefficient = –0.264). Freshwater-marsh birds had shorter pulse duration (5.5 vs. 6.1 ms) and a greater degree of amplitude modulation (pulse depth) (SE*1,000 = 17 vs. SE*1,000 = 14) than saltmarsh birds.

For the trill, we found that 72% of the freshwater-marsh birds fell outside of the 95% CI for saltmarsh birds, demonstrating that habitat-level differences in the trill call type are sufficiently large that we were able to detect them even with a skewed sample size. The final model included peak frequency, upper frequency limit, bandwidth, note rate, first and second inter-note duration, and total call duration. The highest weighted coefficients in the LDA were mean note rate (LD coefficient = 0.25) and total call

duration (LD coefficient = 0.37). Freshwater-marsh birds had a lower note rate (17 vs. 22 notes s^{-1}) and shorter total call duration (1.5 vs. 2.2 s) than saltmarsh birds.

Call Presence–Absence

Using a model selection approach, we asked whether the functional–ecological designation of a population predicted the likelihood of observing a call type. The results are summarized in Table 3. For the churr, rattle, scream, and trill, there were no significant differences in the likelihood of observing a call at any functional–ecological level. In these cases, we found that the null model was the best fit.

For the buzz call, the habitat model was the best-fit model and was 10.8 ΔAIC_c values better than the null model ($\chi^2 = 12.77$, df = 1, $P = 0.001$). We were more likely to observe a buzz from freshwater-marsh birds than from saltmarsh birds ($P = 0.005$; Figure 6A). For the chuck call, the subspecies model was the best-fit model and was 5.2 ΔAIC_c values better than the null model ($\chi^2 = 5.66$, df = 1, $P = 0.017$). In general, we were more likely to observe a chuck from an Atlantic coast subspecies (*griseus* and *palustris*) than from any other subspecies. Post hoc contrasts indicated several significant pairwise subspecies comparisons. We were significantly more likely to observe a chuck from *palustris* vs. *dissaeptus* ($P = 0.018$) and from *palustris* vs. *marianae* ($P = 0.002$). Additionally, we were more likely to observe a chuck from *griseus* vs. *marianae*, but this difference only approached significance ($P = 0.061$; Figure 6D).

For the twitter call, we found that habitat was the best-fit model and was 44.5 ΔAIC_c values lower than the null model ($\chi^2 = 46.43$, df = 1, $P << 0.001$). Post hoc analysis
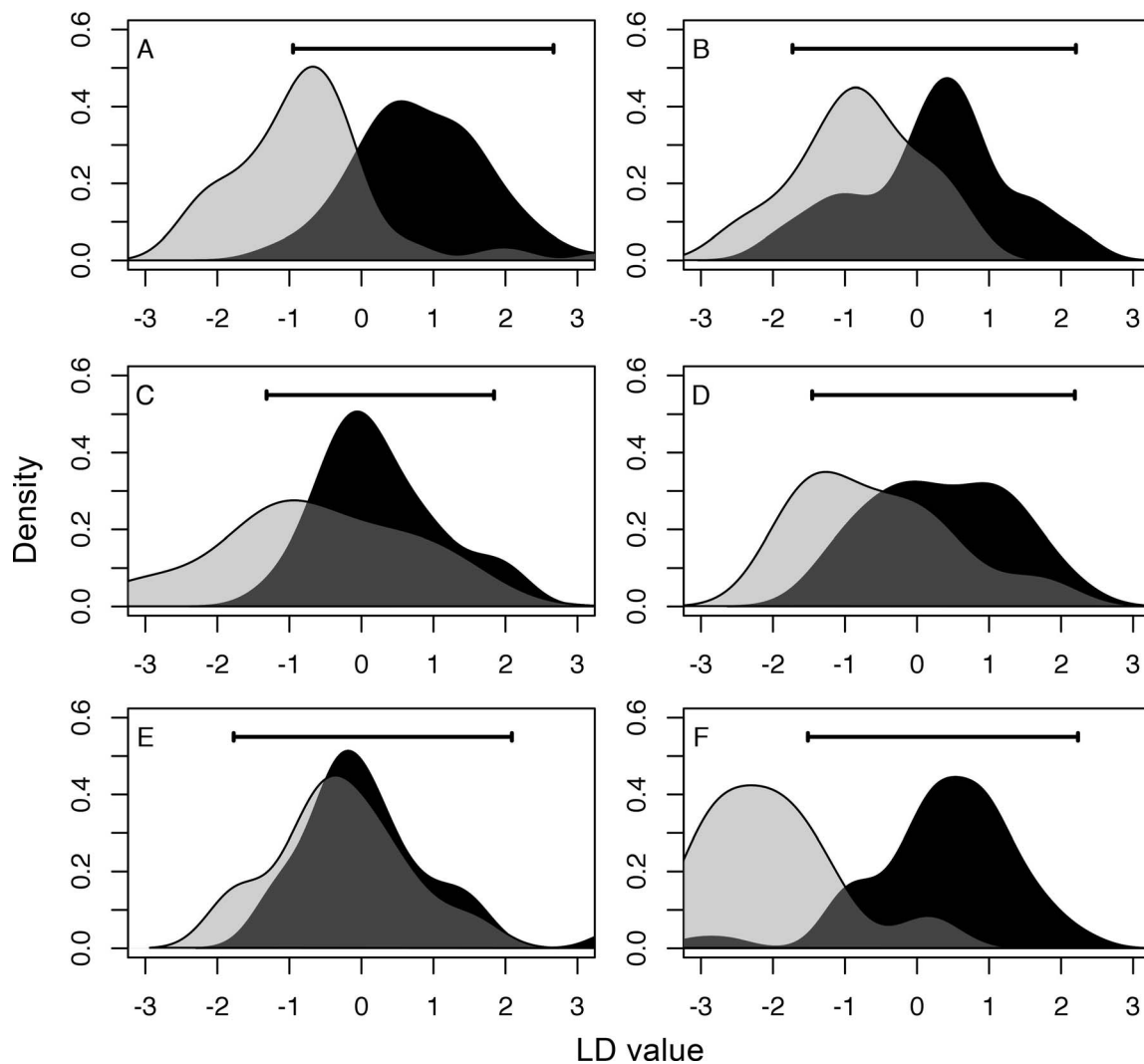
FIGURE 5. The distribution of linear discriminant (LD) values for each call of freshwater-marsh (light gray) and saltmarsh (black) birds, shown as density curves standardized to an area under the curve of 1.0. Horizontal bars represent 95% confidence intervals of bootstrapped saltmarsh LD values. (**A**) Buzz call. (**B**) Chuck call. (**C**) Churr call. (**D**) Rattle call. (**E**) Scream call. (**F**) Trill call. Twitter call is not shown because sample-size differences prevented analysis of acoustic parameters between freshwater-marsh and saltmarsh birds. Linear discriminant analysis of 2 groups by design generates 2 distributions. The power of the analysis is shown by the degree of overlap between the 2 distributions (distributions with a high degree of overlap may be able to be correctly classified if sample sizes are skewed but do not represent easily diagnosable groups).

of the model indicated that we were much more likely to observe a twitter in saltmarsh birds than in freshwater-marsh birds ($P < 0.001$; Figure 6B). For the trill–twitter mix, the habitat model was the best-fit model, at 26.9 ΔAIC_c values lower than the null model ($\chi^2 = 23.71$, df $= 1$, $P < 0.001$). Similar to the twitter, we found that the trill–twitter mix was rare in freshwater-marsh birds compared to saltmarsh birds ($P < 0.001$; Figure 6C).

Calls Embedded in Songs
Within an individual, the proportion of songs produced that included one or more call elements was high in all

subspecies, ranging from 72.5% to 96.7% of all songs produced. At each functional–ecological level of testing, we found significant differences in the proportion of songs produced that had embedded call elements (Kruskal-Wallis test; habitat level: $\chi^2 = 29.95$, df $= 1$, $P << 0.001$; migratory-pattern level: $\chi^2 = 30.04$, df $= 2$, $P << 0.001$; subspecies level: $\chi^2 = 33.56$, df $= 3$, $P << 0.001$; Figure 7). Post hoc Dunn's tests with Bonferroni adjustment revealed significant pairwise differences at the subspecies level between *dissaeptus* and *marianae* ($P << 0.001$; *dissaeptus* mean [$\pm$ SD] $= 0.93 \pm 0.10$, *marianae* mean $= 0.70 \pm 0.29$), between *dissaeptus* and *palustris* ($P < 0.001$;

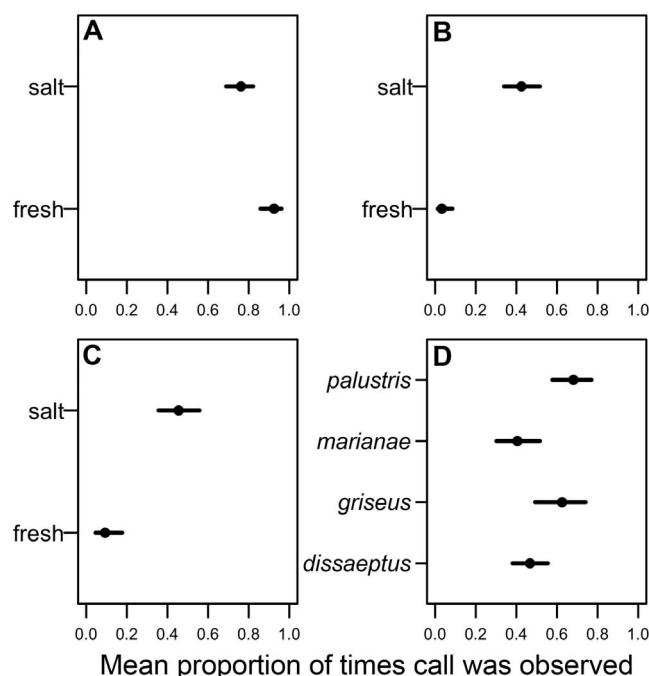The Auk: Ornithological Advances 135:127–151, © 2018 American Ornithological Society

FIGURE 6. Mean (± 95% confidence interval) probability of observing a call type at the functional–ecological level (subspecies, migratory pattern, habitat type) best fit by the data. The likelihood of observing 3 call types was best explained by habitat type: (**A**) buzz call, (**B**) twitter call, and (**C**) trill–twitter mix. The likelihood of observing one call type was best explained by subspecies: (**D**) chuck call.
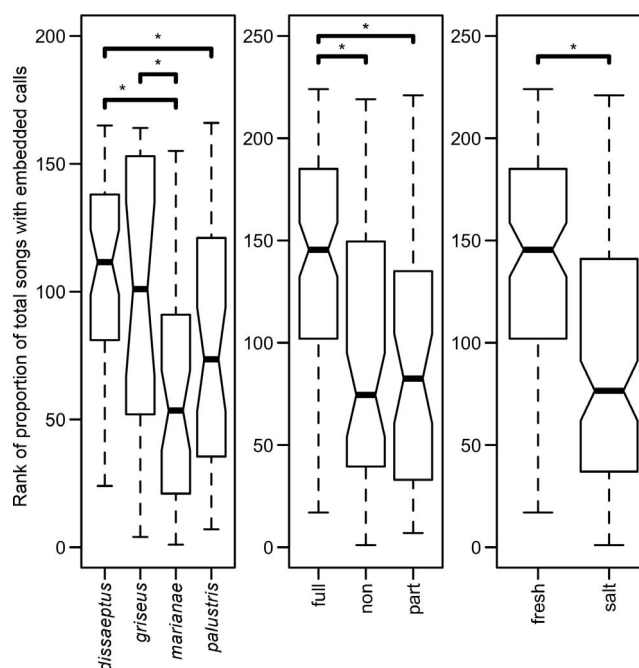
FIGURE 7. Box plots showing ranks of the proportion of songs with embedded calls at each functional–ecological level (subspecies, migratory pattern, habitat type). Tied ranks were broken by random assignment. Median = thick line, second and third quartiles = boxes, first and fourth quartiles = whiskers. Notches indicate a confidence interval around the median. Asterisks indicate significant differences in Kruskal-Wallis non-parametric tests on ranks.

palustris mean = 0.80 ± 0.22), and between *griseus* and *marianae* ($P = 0.020$; *griseus* mean = 0.85 ± 0.23). Post hoc tests also revealed differences at the migratory-pattern level between full migration and nonmigration ($P \ll 0.001$; full-migrant mean = 0.92 ± 0.12, nonmigrant mean = 0.75 ± 0.28), and between full migration and partial migration ($P \ll 0.001$; partial-migrant mean = 0.76 ± 0.24). Freshwater-marsh birds had a significantly higher proportion of call-songs in their repertoire (freshwater-marsh birds: mean = 0.92 ± 0.11; saltmarsh birds: mean = 0.75 ± 0.26). Although we were unable to determine which level of analysis was the best for explaining the variation in proportion of songs with embedded calls because the 3 Kruskal-Wallis tests are not directly comparable in all cases, the single freshwater, fully migratory subspecies, *dissaeptus*, had a higher proportion of call-songs in its repertoire than any other combination of habitat, subspecies, or migratory pattern, and the magnitude of this difference is most striking at the habitat level.

DISCUSSION

Differences in acoustic signals that correspond to taxonomy, geography, or habitat type have been noted across a wide variety of taxa and may play an important role in understanding the mode and tempo of species diversification (e.g., Ryan and Wilczynski 1991, Wielgart and Whitehead 1997, Irwin 2000, Mendelson and Shaw 2002, Parmentier et al. 2005, Stephan and Zuberbühler 2008, Jiang et al. 2015). One way in which acoustic signals may facilitate the divergence of populations is by acting as barriers to hybridization, since acoustic signals can diverge faster than postzygotic barriers to reproduction (Grant and Grant 1996, Wilkins et al. 2013). As mate attraction signals, bird songs have a clear potential to act as prezygotic barriers to hybridization. Calls might also act as barriers to hybridization, particularly if they function in contexts related to mate attraction, though structural and functional differences in calls used in other contexts may still reflect the degree of underlying diversification (while not acting as barriers themselves). Barriers to hybridization could be especially important in recently diverged taxa that inhabit different physiologically challenging environments that favor local adaptation, such as saltmarsh endemics (Basham and Mewaldt 1987, Greenberg and Droege 1990, Greenberg and Danner 2012, Maley 2012, Danner et al. 2017). Signal divergence can occur through multiple mechanisms, including indirect selection on the morphol-

ogy of the signal-production apparatus (e.g., Ballentine 2006), direct selection of the habitat on the signal (e.g., acoustic transmission properties; Morton 1975), cultural selection on the signal form (e.g., Price 1998, Soha et al. 2004), social or sexual selection on the signal form (e.g., Colbeck et al. 2010, Dingle et al. 2010), and the cultural or genetic drift of the signal (e.g., Irwin et al. 2008). One of the challenges in comparing acoustic signals among populations is that animals with complex acoustic repertoires may demonstrate different patterns of selection and differentiation on different calls within the repertoire, depending on their form and function (Laiolo et al. 2001a, Laiolo and Rolando 2002, Irwin et al. 2008, Sturge et al. 2016).

Acoustic Descriptions of Call Types

Marsh Wrens, like many birds, have complicated call repertoires that facilitate interactions among individuals. We described the acoustic properties and behavioral context of calling behavior in Marsh Wrens, identified 7 discrete call types, and quantified the acoustic properties of those calls across all eastern subspecies. Despite the fact that Marsh Wrens are models of vocal learning with well-studied song repertoires (Verner 1975, Kroodsma and Verner 1987, Luttrell et al. 2016), Marsh Wren calls have not been formally described prior to this work.

Quantitative Discrimination of Call Types

Multivariate analysis among similar calls confirmed our initial qualitative categorization of 7 call types. Intermediate call types may be produced but are uncommon. Variation within individuals ranged from stereotyped (most notably in note duration of alarm calls) to highly variable (e.g., total call duration), depending on the call type and measurement. Flexibility in delivery in each call type may allow for graded signal information that parallels underlying signaler motivation but falls within a defined acoustic space for each of the 7 call types (Morton 1977, Davis 1988, Ficken 1990; Figure 4). Graded signals are common in the alarm calls of some mammals (Fichtel and Kappeler 2002) and birds, including Phasianidae, Corvidae, and Paridae (Suzuki 2016). Within Troglodytidae, all species with at least partially described call repertoires have at least one call that could be considered graded (Hejl et al. 2002, Hamilton et al. 2011, Toews and Irwin 2012, Haggerty and Morton 2014, Johnson 2014).

Behavioral Association of Call Types

The behavioral contexts of Marsh Wren vocalizations fall into 2 broad categories: calls associated with mate attraction or territory patrol, and calls associated with alarm or distress. The buzz, churr, trill, and twitter calls, as well as Marsh Wren song, were all most commonly associated with territory-patrol behaviors, nest building,

and courtship (Table 2). Males were more likely to produce all of these vocalizations than females. Our observations of Marsh Wren calling behavior were made during the breeding season. Nonmigratory populations of Marsh Wren do not defend territories in the nonbreeding season, when social interactions may differ (Kale 1965, Verner 1965). Future work during the nonbreeding season would help identify any behavioral differences in the described calls or additional call types not used during the breeding season (Marler 2004b).

The buzz call is an integral part of male Marsh Wren courtship. Welter (1935) described the buzz call as part of the song and reported that it was rarely heard after the beginning of June. By contrast, we observed the buzz call more often than any other call type, and it was produced throughout the entirety of our field seasons (May–July), both independently and in conjunction with song. Marsh Wrens have a distinctive courtship display during which the male raises the tail over the head and rocks forward and backward while primarily producing the buzz call (Welter 1935, Kale 1965). The buzz, churr, trill, and twitter calls were also used during pursuit of females around the territory (song is less commonly used than calls during these courtship pursuits). The buzz and associated display have also been reported during territory intrusions (Welter 1935, Kale 1965), but during simulated territory intrusions we found that the male typically produced the buzz and display when a female also approached the simulated intruder.

In addition to territory patrol, male Marsh Wrens also frequently produced the buzz call during nest building; and, to a lesser degree, they also produced the churr, trill, and twitter calls, as well as song, in this context. Eastern male Marsh Wrens build, on average, 5–12 nests on a single territory that become courtship display centers, although these nests likely serve multiple functions (Bent 1948, Kale 1965, Picman 1977). The trill and twitter calls were also often produced during flight while entering or leaving the territory. The trill, in particular, was associated with song during pop-up flight displays (as noted by Welter 1935, Kale 1965).

The remaining call types—chuck, rattle, and scream—were most commonly observed in agonistic encounters and could be produced by either the instigator or the recipient of conspecific agonistic encounters. The chuck is an alarm call used by both sexes, usually when a conspecific was present, and most commonly by females, confirming Welter's (1935) observations. Marsh Wrens may produce chuck calls in response to other Marsh Wren intrusions, during an unwanted courtship display, or in response to researcher presence. The rate of delivery and number of notes are variable within and among individuals, which suggests a graded structure to the signal (Davis 1988, Fichtel and Kappeler 2002). The rattle call was

typically produced by females or by birds of unknown sex and was nearly always observed when conspecifics were detected. It is likely analogous with Welter's (1935) "kek," which he describes as a female-only alarm call. In keeping with Welter's observations, we observed the rattle most often immediately before or after an agonistic encounter, but also often during flight away from the same location on repeated observations, which suggests that this call may be associated with nest departure. Nest-departure calls are common in marsh-breeding species and are found frequently in species that breed in high densities (McDonald and Greenberg 1991). The broad-bandwidth, steeply modulated, repetitive structure of the rattle call is consistent with nest-departure vocalizations that are easily localizable and can be used to track flight trajectory (McDonald and Greenberg 1991, Haff et al. 2015). Marsh Wrens readily destroy conspecific nests, and females guard nests closely (Picman 1977). If females use rattle calls to alert males when the nest is left unattended, it may function to increase vigilance against potentially agonistic interactions (Yasukawa 1989, Grunst et al. 2014). The scream was most commonly observed from birds of unknown sex or females that were the recipients of aggressive encounters. It was also produced when birds were restrained for banding and measurements. The broadband, noisy, frequency- and amplitude-modulated nature of the scream was typical of distress calls used by a variety of passerines (Conover 1994, Jurisevic and Sanderson 1994).

Geographic Call Variation

We examined call variation in the context of 3 levels of functional–ecological categorization—subspecies-level grouping, migratory behavior, and habitat type—and found that 4 call types were not diagnosable by any level of functional–ecological categorization. That is, these calls were similar in acoustic structure across the range of Marsh Wren populations we examined. By contrast, another call, the twitter, was most accurately categorized according to migratory pattern, and 2 calls (buzz and trill) were most accurately categorized according to habitat type. A variety of evolutionary mechanisms could lead to variation in call structure among taxonomic groups, migratory patterns, or habitat types, including drift (Benedict and Krakaur 2013), indirect selection related to variation in body size or physiology (e.g., Laiolo et al. 2001a), selection via acoustic transmission properties of the habitat (Morton 1975), or social or sexual selection (West-Eberhard 1983). Although many studies exploring geographic variation in vocalizations focus on a single vocal signal, those examining multiple signals have found that different signals may differ in the amount of variation exhibited across geographic ranges (Laiolo et al. 2001a, 2001b, Irwin et al. 2008, Sturge et al. 2016). Varying

patterns and degrees of divergence among vocalization types may be the rule, rather than the exception. Because different call types can have different anatomical or production constraints and different behavioral functions, they may also experience different selective pressures.

The 4 call types that did not exhibit any acoustic differences across the geographic range that we examined—the chuck, churr, rattle, and scream calls—share some similar acoustic features. All 4 calls are broadband vocalizations with rapid onset and offset. Alarm and distress calls of passerines such as the chuck, rattle, and scream, especially those that are highly localizable, often draw the attention of both conspecific and heterospecific individuals to investigate the source of the call and, in some cases, mob potential predators or other threats (Stefanski and Falls 1972a, 1972b, Hurd 1996, Chu 2001, Lee et al. 2015). As a result, these call types may be under convergent selection to be broadly recognizable and easy to localize for both conspecifics and heterospecifics. In contrast to these 3 calls, it is less clear from a functional perspective why the churr call did not show evidence of geographic variation. Like the other calls, the churr is a broadband call, though it is lower in amplitude and has a narrower bandwidth than any of the alarm or distress calls. The churr was used in territory-patrol and courtship contexts but was more stereotyped within an individual than other territory-patrol or courtship calls.

The twitter call was most divergent among migratory patterns. Nonmigratory populations produced a slower-paced twitter with a shorter first note in the 2-note delivery than partially migratory populations. Other call features were more similar; for example, there was little variation in frequency of the twitter call between the 2 groups. While slower note delivery may be advantageous in closed habitat types because of increased reverberation, there are no consistent differences in habitat acoustics among the nonmigratory or partially migratory populations (Wiley and Richards 1978). In addition, there are no reported body-size differences among the nonmigratory or partially migratory populations (Pyle 1997, S. A. M. Luttrell personal observation), which suggests that pleiotropic effects of body size on signal structure are also unlikely. The nonmigratory and partially migratory populations of Marsh Wren are parapatric, distributed in a wide latitudinal gradient, and isolated during the breeding season, meeting several key criteria for drift. Despite differences between nonmigratory and partially migratory twitter calls, we found no evidence of differences in the twitter between the 2 nonmigratory subspecies that also meet key criteria for drift, *griseus* and *marianae*, which are isolated during both the breeding and nonbreeding seasons, though divergence times between these 2 subspecies are unknown in comparison with the divergence between *dissaeptus* and the saltmarsh subspecies as

a whole. Because freshwater-marsh populations and salt-marsh populations occur over disjunct rather than continuous geographic ranges, it is difficult to apply an isolation-by-distance approach in this instance. More detailed sampling between the freshwater-marsh subspecies *dissaeptus* and the parapatric saltmarsh subspecies *palustris* could be used to further test an isolation-by-distance hypothesis in this call type. The lack of divergence in the twitter call among allopatric nonmigratory populations suggests a minor role for drift in this signal but doesn't rule out drift as an important force in this system.

The twitter call is produced primarily by males and is used in territory patrol and courtship, behaviors related to either intrasexual or intersexual communication. Partially migratory populations winter within the breeding range of nonmigratory populations, and it is possible that nonmigrants begin breeding before all partial migrants have left the wintering range. Subspecies of birds are capable of interbreeding, though hybrid migratory patterns have been shown to have direct fitness costs in other species (Helbig 1991, Delmore and Irwin 2014). Given that variation in the twitter is best explained by migratory pattern, it is possible that sexual selection on the twitter call could reinforce selective breeding within individuals that have a similar migratory pattern. Although it is not known whether coastal Marsh Wren populations with different migratory patterns hybridize, the genetically and vocally divergent subspecies of the Great Plains, *C. p. illiacus* and *C. p. laingi*, do occasionally hybridize, which suggests that such hybridization across migratory patterns is possible (Kroodsma and Verner 2013). Female-preference tests with twitter calls from birds of different migratory populations are needed to confirm that birds can perceive these differences, and that these differences are meaningful in an analogue for a mate-selection context. Although these tests would not determine whether drift preceded preferences for local signals or whether selection against hybrid migratory patterns drove divergence, they would provide support for the idea that the observed differences are reinforced through sexual selection between birds with different migratory patterns.

Two call types, the buzz and the trill, were most accurately grouped according to habitat type, rather than subspecies or migratory pattern. Freshwater-marsh birds produced buzzes with higher peak frequency, faster pulse rates, and greater amplitude variation between pulses than saltmarsh birds. Freshwater-marsh birds produced trills that were shorter in duration, with higher frequency and slower pace of delivery. Marsh Wrens from freshwater marshes are larger than saltmarsh birds (S. A. M. Luttrell personal observation). However, increased body size is more often correlated with decreased fundamental frequencies rather than increased frequencies (Wallschläger 1980, Ryan and Brenowitz 1985, Martin et al. 2011). Both

freshwater marshes and saltmarshes are open habitats over water. Given that the structural components of these habitats are quite similar, direct selection on the frequency elements of calls as a result of habitat differences seems unlikely in these 2 habitats (Wiley and Richards 1978, Cosens and Falls 1984).

Given that neither direct selection of the habitat nor pleiotropic effects of body size are likely to explain differences in the buzz and trill calls between freshwater-marsh and saltmarsh populations, we are left with neutral isolation by distance (drift) and/or direct effects of sexual selection to explain the observed differences. The effect of drift may be enhanced if calls or call components are learned, given that cultural changes may proceed more rapidly than genetic variation over time (Lynch 1996). Examples of drift or isolation by distance are common among species in which geographic variation in calling behavior has been documented, and many of these cases involve or are suspected to involve vocalizations that are learned (Miyasato and Baker 1999, Laiolo et al. 2001a, Wright et al. 2005, Mulard et al. 2009). We do not know whether any acoustic components of the buzz or trill calls involve learning. In the absence of learning, the observed variation in these calls may be indicative of underlying genetic drift (Benedict and Krakauer 2013). As in the prior comparison with the twitter call, however, we found no discernible evidence of drift or variation in the buzz or trill calls between the 2 nonmigratory saltmarsh subspecies, *griseus* and *marianae*.

Lastly, whether divergence in the buzz and trill calls was initiated via neutral or selective processes, both calls are used in the context of intrasexual and intersexual communication, and, as a consequence, they are candidates for ongoing sexual selection. Although freshwater-marsh and saltmarsh populations are allopatric during the bulk of the breeding season, saltmarsh populations may begin breeding before freshwater-marsh populations have completely left the wintering grounds (Kroodsma and Verner 2013). If hybrids exist, as they do at the boundary between eastern and western subspecies-complexes (Kroodsma and Verner 2013), sexual selection could facilitate or reinforce divergence in these call types. Playback tests in same-ecotype vs. different-ecotype songs have demonstrated discrimination in freshwater-marsh vs. saltmarsh subspecies of the Swamp Sparrow (*Melospiza georgiana*; Liu et al. 2008). Similar tests with Marsh Wrens could identify whether the differences we measured are perceptible and meaningful to birds in a territorial context, and confirm whether vocal variation in saltmarsh endemic populations is a more widespread phenomenon in birds.

Call Presence–Absence

In addition to geographic variation in acoustic structure in the calls of Marsh Wrens, we found that some call types

varied in how commonly they occurred across subspecies and habitat types. The freshwater-marsh, fully migratory subspecies, *dissaeptus*, produced the twitter and trill–twitter mixed calls rarely compared with all other subspecies and habitat types but produced the buzz call more frequently than saltmarsh populations. Although both the buzz and twitter are used during territory patrol and courtship, the buzz seems to play an integral role in courtship behavior over short distances, whereas the twitter seems to occur primarily in a long-distance context as a broadcast signal similar to song. This preliminary analysis suggests that there could be behavioral differences between freshwater-marsh and saltmarsh populations with respect to courtship or advertising patterns. Although we did not evaluate the breeding density or degree of polygyny among our populations, both density and polygyny vary among populations of Marsh Wrens and could contribute to differences in mate-advertisement and territory-defense patterns (Welter 1935, Kale 1965, Verner 1965, Leonard 1986). For example, if territories are densely distributed in suitable habitat, we might expect higher proportions of the low-amplitude, short-distance buzz call in the repertoire than of the trill or twitter calls, which are broadcast signals. Furthermore, since the buzz call is predominantly used in courtship or during territory intrusions when a female is present, populations with higher levels of polygyny might be expected to have higher proportions of the buzz call. Lastly, because the buzz and twitter seem to play different roles in courtship and territory patrol, differences in production of the buzz and twitter calls could result from differences in breeding stage. We controlled for the effect of date in our analyses, but we did not systematically quantify the breeding stage of individuals. Differences in the occurrence of the buzz and twitter could be related to differences in breeding phenology between saltmarsh and freshwater-marsh Marsh Wrens, or some combination of the effects of breeding density, breeding phenology, and levels of polygyny.

We also found differences in the likelihood of observing a chuck call at the subspecies level, with individuals of 2 Atlantic coast subspecies (*griseus* and *palustris*) producing the chuck call more often than individuals of other subspecies. The chuck call is primarily an alarm call. If the number of agonistic encounters or risk of predation is greater in Atlantic coast populations than in other populations, then the propensity to use this call type may be increased. Increased breeding density could also result in increased alarm calling, given that Marsh Wrens commonly destroy conspecific nests and females may aggressively defend nest areas from conspecifics (Picman 1977, Leonard and Picman 1987). The propensity to produce alarm calls, and the composition of alarm calls in response to threats, has been shown to vary among individuals and sex classes in another North American

passerine, the Tufted Titmouse (*Baeolophus bicolor*; Freeberg and Branch 2013), further suggesting that the age or sex composition, or the dominance dynamics of a population as a whole, could influence how common a vocalization is within a population. Freeberg (2012) has demonstrated variation in both the relative proportion and context of components of the *chick-a-dee* call in Carolina Chickadees (*Poecile carolinensis*) between 2 populations in central North America, which suggests a functional divergence in these calls despite a lack of variation in the acoustic structure of those components. Future studies on geographic variation in complex calling behavior, coupled with detailed demographic information, could clarify whether geographic differences in how commonly a call is produced, or the context of its production, are more widespread in passerines and perhaps act as additional mechanisms by which vocal divergence can occur.

Calls Embedded in Songs

Eastern Marsh Wrens commonly embed 3 call types—the buzz, trill, and twitter calls—into the song. The incidence of such "call-songs" may be less common in western populations (Verner 1975). Calls can be included or excluded in different renditions of the same song type, and they can be inserted between multiple songs within a single, continuous song bout. When calls are embedded in songs, they follow a nonrandom organizational pattern. Twitter calls typically precede introductory notes, the buzz call is often embedded among introductory notes, and the trill call usually follows the terminal notes of the song. Each of these call types is regularly produced independent of song as well, and all 3 calls appear to function primarily in territory patrol and courtship. Although the proportion of total songs produced that included one or more call elements was high in all populations, the freshwater-marsh, fully migratory subspecies, *dissaeptus*, consistently had the highest proportion of songs with embedded call elements. Systematic differences in breeding density or breeding phenology between freshwater-marsh and saltmarsh populations of Marsh Wren could contribute to our observed differences. For example, if call elements are directed at different individuals than song elements, or function as local vs. broadcast signals, increased breeding density may result in an increase in the use of call-songs. Welter (1935) suggested that males reduce the incidence of buzz calls produced in conjunction with the song as the breeding season progresses. Although we did not observe an effect of date on the occurrence of songs with embedded calls, Welter's observation could be a consequence of a difference in the likelihood of call-songs at different stages in the breeding cycle. Thus, a greater proportion of embedded calls in the freshwater populations could be indicative of differences in breeding phenology. We did not quantify breeding density or breeding stage in our analysis. Alternatively, embedded calls may function to increase the

diversity of the already complex song repertoires of eastern Marsh Wrens. Prior work on Marsh Wren repertoires has demonstrated that all eastern subspecies have similar repertoire sizes regardless of habitat type or migratory pattern, in particular when only the main portion of the song is considered (Kroodsma and Verner 1987, Luttrell et al. 2016). If embedded calls serve to increase song repertoire sizes in fully migratory, freshwater-marsh populations, then this would support the hypothesis that migratory populations of birds should have larger song repertoire sizes due to increased intensity of sexual selection pressures during shorter breeding seasons (Read and Weary 1992).

The use of call notes embedded in songs may be common in other species, but if so it is underreported. In most cases where this behavior has been reported, the calls of heterospecifics are included in a small proportion of overall songs produced and are hypothesized to increase the repertoire size of the singer (Howard 1974, Hindmarsh 1984, Greenlaw et al. 1998). Although it is not a perfectly analogous case, we could find only one other reported bird species, the Zebra Finch (*Taeniopygia guttata*), that consistently embeds conspecific call types between song syllables (e.g., Zann 1993). The function of call syllables within Zebra Finch songs is unknown. As in Marsh Wrens, the calls appear in a nonrandom order within the song; but unlike in Marsh Wrens, call notes in Zebra Finch song represent a crystallized song motif rather than a syllable that may be flexibly added or removed. More detailed observations of when and how calls are embedded in bird songs generally, and in Marsh Wren songs specifically, or the response of Marsh Wrens to songs with and without calls, are needed to determine their function in this context.

Conclusions

We describe the call repertoire of Marsh Wrens and report evidence of geographic variation in the acoustic structure of 3 of 7 call types, as well as geographic and taxonomic variation in how commonly 4 call types are produced. Marsh Wrens are semicolonial, gregarious birds that have complex vocal and social behavior. While it may not be surprising that calls show varying levels of divergence in this species, it is nonetheless interesting that the signals demonstrating the greatest degree of divergence are those most likely to be under sexual selection pressures based on the behavioral context of the calls. The effects of drift cannot be explicitly ruled out in these circumstances, but if drift is playing a role in the divergence of calls involved in intrasexual and intersexual communication, we expect that sexual selection may reinforce or strengthen the effects of neutral divergence when birds from different populations are sympatric. The variation that we observed is not well explained by selective effects of the habitat for sound transmission or by pleiotropic effects on acoustics due to body size. Because within-species divergence is likely

recent and ongoing, focusing on within-species rather than between-species variation in behavior is useful for inferring the mode and tempo of divergence in populations. In particular, such a focus avoids the confounding effects of the accumulation of additional differences over long time spans (Foster 1999). Extensive work on song dialects within species of birds has provided us with ample evidence of geographic divergence in vocal signals (reviewed in Podos and Warren 2007), in particular as a consequence of cultural evolution. However, the vocal repertoires of birds are typically complex, and song is only one component of the overall repertoire (Marler 2004a). Geographic variation in calls has been demonstrated in many species, but in such studies often only a single call type is examined (e.g., Rothstein and Fleischer 1987, Miyasato and Baker 1999, Wright et al. 2005, Nicholls et al. 2006, Snowberg and Benkman 2007; but see Laiolo et al. 2001a, 2001b). By comparing multiple vocal signals across a wide geographic range, we can help build a more complete picture of the overall processes of selection and divergence involved in speciation.

LITERATURE CITED

Allen, A. A. (1923). April in the marshes. Bird Lore 25:145–146.

Ballentine, B. (2006). Morphological adaptation influences the evolution of a mating signal. Evolution 60:1936–1944.

Baptista, L. F. (1977). Geographic variation in song and dialects of the Puget Sound White-crowned Sparrow. The Condor 79: 356–370.

Basham, M. P., and L. R. Mewaldt (1987). Salt water tolerance and the distribution of south San Francisco Bay Song Sparrows. The Condor 89:697–709.

Bates, D., M. Mächler, B. Bolker, and S. Walker (2015). Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Benedict, L., and A. H. Krakauer (2013). Kiwis to pewees: The value of studying bird calls. Ibis 155:225–228.

Benjamini, Y., and Y. Hochberg (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B 57: 289–300.

Bent, A. C. (1948). Life histories of familiar North American birds. U.S. National Museum Bulletin 195:248–259.

Bertelli, S., and P. L. Tubaro (2002). Body mass and habitat correlates of song structure in a primitive group of birds. Biological Journal of the Linnean Society 77:423–430.

Catchpole, C. K., and P. J. B. Slater (2008). Bird Song: Biological Themes and Variations, second edition. Cambridge University Press, New York, NY, USA.

Chu, M. (2001). Heterospecific responses to scream calls and vocal mimicry by phainopeplas (Phainopepla nitens) in distress. Behaviour 138:775–787.

Colbeck, G. J., T. S. Sillett, and M. S. Webster (2010). Asymmetric discrimination of geographical variation in song in a migratory passerine. Animal Behaviour 80:311–318.

Conover, M. R. (1994). Stimuli eliciting distress calls in adult passerines and response of predators and birds to their broadcast. Behaviour 131:19–37.

Cosens, S. E., and J. B. Falls (1984). A comparison of sound propagation and song frequency in temperate marsh and grassland habitats. Behavioral Ecology and Sociobiology 15: 161–170.

Crawley, M. J. (2007). The R Book. Wiley, Chichester, UK.

Danner, J. E., R. M. Danner, F. Bonier, P. R. Martin, T. W. Small, and I. T. Moore (2011). Female, but not male, tropical sparrows respond more strongly to the local song dialect: Implications for population divergence. The American Naturalist 178:53–63.

Danner, R. M., E. R. Gulson-Castillo, H. F. James, S. A. Dzielski, D. C. Frank III, E. T. Sibbald, and D. W. Winkler (2017). Habitat-specific divergence of air conditioning structures in bird bills. The Auk 134:65–75.

Davis, W. J. (1988). Acoustic communication in the Belted Kingfisher: An example of temporal coding. Behaviour 106:1–23.

Delmore, K. E., and D. E. Irwin (2014). Hybrid songbirds employ intermediate routes in a migratory divide. Ecology Letters 17: 1211–1218.

Dingle, C., J. W. Poelstra, W. Halfwerk, D. M. Brinkhuizen, and H. Slabbekoorn (2010). Asymmetric response patterns to subspecies-specific song differences in allopatry and para-patry in the Gray-breasted Wood-wren. Evolution 64:3537–3548.

Engineering Design (2015). SIGNAL/RTSD. Berkeley, CA, USA. http://engdes.com

Fichtel, C., and P. M. Kappeler (2002). Anti-predator behavior of group-living Malagasy primates: Mixed evidence for a referential alarm call system. Behavioral Ecology and Sociobiology 51:262–275.

Ficken, M. S. (1990). Acoustic characteristics of alarm calls associated with predation risk in chickadees. Animal Behaviour 39:400–401.

Ficken, M. S., and J. W. Popp (1995). Long-term persistence of a culturally transmitted vocalization of the Black-capped Chickadee. Animal Behaviour 50:683–693.

Foster, S. A. (1999). The geography of behaviour: An evolutionary perspective. Trends in Ecology & Evolution 14:190–195.

Freeberg, T. M. (2012). Geographic variation in note composition and use of chick-a-dee calls of Carolina Chickadees (Poecile carolinensis). Ethology 118:555–565.

Freeberg, T. M., and C. L. Branch (2013). Call variation in Tufted Titmice (Baeolophus bicolor) in a distress context: Potential sex differences. Ethology 119:680–691.

Goldstein, D. L. (2006). Osmoregulatory biology of saltmarsh passerines. Studies in Avian Biology 32:110–118.

Grant, B. R., and P. R. Grant (1996). Cultural inheritance of song and its role in the evolution of Darwin's finches. Evolution 50: 2471–2487.

Greenberg, R., and R. M. Danner (2012). The influence of the California marine layer on bill size in a generalist songbird. Evolution 66:3825–3835.

Greenberg, R., and S. Droege (1990). Adaptations to tidal marshes in breeding populations of the Swamp Sparrow. The Condor 92:393–404.

Greenlaw, J. S., C. E. Shackelford, and R. E. Brown (1998). Call mimicry by Eastern Towhees and its significance in relation to auditory learning. The Wilson Bulletin 110:431–434.

Groth, J. G. (1993). Call matching and positive assortative mating in Red Crossbills. The Auk 110:398–401.

Grunst, M. L., A. S. Grunst, and J. T. Rotenberry (2014). Phenotypic plasticity in nest departure calls: Weighing costs and benefits. Animal Behaviour 90:31–39.

Haff, T. M., A. G. Horn, M. L. Leonard, and R. D. Magrath (2015). Conspicuous calling near cryptic nests: A review of hypotheses and a field study on White-browed Scrubwrens. Journal of Avian Biology 46:289–302.

Haggerty, T. M., and E. S. Morton (2014). Carolina Wren (Thryothorus ludovicianus). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/carwre

Hamilton, R. A., G. A. Proudfoot, D. A. Sherry, and S. L. Johnson (2011). Cactus Wren (Campylorhynchus brunneicapillus). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/cacwre

Hebert, P. D. N., M. Y. Stoeckle, T. S. Zemlak, and C. M. Francis (2004). Identification of birds through DNA barcodes. PLOS Biology 2:e312.

Hejl, S. J., J. A. Holmes, and D. E. Kroodsma (2002). Winter Wren (Troglodytes hiemalis). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/winwre3

Helbig, A. J. (1991). Inheritance of migratory direction in a bird species: A cross-breeding experiment with SE- and SW-migrating Blackcaps (Sylvia atricapilla). Behavioral Ecology and Sociobiology 28:9–12.

Hindmarsh, A. M. (1984). Vocal mimicry in starlings. Behaviour 90:302–324.

Howard, R. D. (1974). The influence of sexual selection and interspecific competition on Mockingbird song. Evolution 28:428–438.

Hughes, M., S. Nowicki, and B. Lohr (1998). Call learning in Black-capped Chickadees (*Parus atricapillus*): The role of experience in the development of "chick-a-dee" calls. Ethology 104:232–249.

Hurd, C. R. (1996). Interspecific attraction to the mobbing calls of Black-capped Chickadees (*Parus atricapillus*). Behavioral Ecology and Sociobiology 38:287–292.

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

Irwin, D. E., S. Bensch, and T. D. Price (2001). Speciation in a ring. Nature 409:333–337.

Irwin, D. E., M. P. Thimgan, and J. H. Irwin (2008). Call divergence is correlated with geographic and genetic distance in Greenish Warblers (*Phylloscopus trochiloides*): A strong role for stochasticity in signal evolution? Journal of Evolutionary Biology 21:435–448.

Jiang, T., H. Wu, and J. Feng (2015). Patterns and causes of geographic variation in bat echolocation pulses. Integrative Zoology 10:241–256.

Johnson, L. S. (2014). House Wren (*Troglodytes aedon*). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/houwre

Jurisevic, M. A., and K. J. Sanderson (1994). Alarm vocalisations in Australian birds: Convergent characteristics and phylogenetic differences. Emu–Austral Ornithology 94:69–77.

Kale, H. W., II (1965). Ecology and bioenergetics of the Long-billed Marsh Wren *Telmatodytes palustris griseus* (Brewster) in Georgia salt marshes. Publications of the Nuttall Ornithological Club 5.

Krebs, J. R., and D. E. Kroodsma (1980). Repertoires and geographical variation in bird song. Advances in the Study of Behavior 11:143–177.

Kroodsma, D. E., and R. A. Canady (1985). Differences in repertoire size, singing behavior, and associated neuroanatomy among Marsh Wren populations have a genetic basis. The Auk 102:439–446.

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

Kroodsma, D. E., and J. Verner (2013). Marsh Wren (*Cistothorus palustris*). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/marwre

Laiolo, P., and A. Rolando (2002). The begging call of the Red-billed Chough *Pyrrhocorax pyrrhocorax*: An analysis of its spectrotemporal variability. Ardea 90:139–147.

Laiolo, P., A. Rolando, A. Delestrade, and A. de Sanctis (2001a). Geographical variation in the calls of the choughs. The Condor 103:287–297.

Laiolo, P., A. Rolando, A. Delestrade, and A. De Sanctis (2001b). Geographic diversification in the call repertoire of the genus *Pyrrhocorax* (Aves, Corvidae). Canadian Journal of Zoology 79:1568–1576.

Lee, J.-W., J.-Y. Lee, H.-J. Noh, M.-S. Kim, K.-H. Kim, and J.-C. Yoo (2015). Habitats matter: The incidence of and response to fear screams in a habitat generalist, the Vinous-throated Parrotbill *Paradoxornis webbianus*. Behavioral Ecology and Sociobiology 69:1575–1584.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. Journal of Statistical Software 69(1).

Leonard, M. L. (1986). Female choice and the evolution of polygyny in the Marsh Wren (*Cistothorus palustris*). M.Sc. thesis, University of Ottawa, Ottawa, Canada.

Leonard, M. L., and J. Picman (1987). Female settlement in Marsh Wrens: Is it affected by other females? Behavioral Ecology and Sociobiology 21:135–140.

Liu, I. A., B. Lohr, B. Olsen, and R. Greenberg (2008). Macrogeographic vocal variation in subspecies of Swamp Sparrow. The Condor 110:102–109.

Luttrell, S. A. M., M. E. Gallagher, and B. Lohr (2016). Multiple estimation methods suggest similar repertoire sizes for Gulf Coast and eastern Marsh Wrens with no correlation between repertoire size and migratory distance. Behaviour 153:287–311.

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. pp. 181–197.

Malamud-Roam, K. P., F. P. Malamud-Roam, E. B. Watson, J. N. Colins, and B. L. Ingram (2006). The quaternary geography and biogeography of tidal salt marshes. Studies in Avian Biology 32:11–31.

Maley, J. M. (2012). Ecological speciation of King Rails (*Rallus elegans*) and Clapper Rails (*Rallus longirostris*). Ph.D. dissertation, Louisiana State University, Baton Rouge, LA, USA.

Marler, P. (2004a). Bird calls: A cornucopia for communication. In Nature's Music: The Science of Birdsong (P. Marler and H. Slabbekorn, Editors). Elsevier Academic Press, San Diego, CA, USA. pp. 132–177.

Marler, P. (2004b). Bird calls: Their potential for behavioral neurobiology. Annals of the New York Academy of Science 1016:31–44.

Marshall, D. C., K. B. R. Hill, J. R. Cooley, and C. Simon (2011). Hybridization, mitochondrial DNA phylogeography, and prediction of the early stages of reproductive isolation: Lessons from New Zealand cicadas (genus *Kikihia*). Systematic Biology 60:482–502.

Martin, J. P., S. M. Doucet, R. C. Knox, and D. J. Mennill (2011). Body size correlates negatively with the frequency of distress calls and songs of Neotropical birds. Journal of Field Ornithology 82:259–268.

McDonald, M. V., and R. Greenberg (1991). Nest departure calls in female songbirds. The Condor 93:365–373.

Mendelson, T. C., and K. L. Shaw (2002). Genetic and behavioral components of the cryptic species boundary between *Laupala cerasina* and *L. kohalensis* (Orthoptera: Gryllidae). Genetica 116:301–310.

Miyasato, L. E., and M. C. Baker (1999). Black-capped Chickadee call dialects along a continuous habitat corridor. Animal Behaviour 57:1311–1318.

Morton, E. S. (1975). Ecological sources of selection on avian sounds. The American Naturalist 109:17–34.

Morton, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. The American Naturalist 111:855–869.

Mulard, H., T. Aubin, J. F. White, R. H. Wagner, and É. Danchin (2009). Voice variance may signify ongoing divergence among Black-legged Kittiwake populations. Biological Journal of the Linnean Society 97:289–297.

Nicholls, J. A., J. J. Austin, C. Moritz, and A. W. Goldizen (2006). Genetic population structure and call variation in a passerine bird, the Satin Bowerbird, *Ptilonorhynchus violaceus*. Evolution 60:1279–1290.

Parmentier, E., J. P. Lagardère, P. Vandewalle, and M. L. Fine (2005). Geographical variation in sound production in the anemonefish *Amphiprion akallopisos*. Proceedings of the Royal Society B 272:1697–1703.

Patten, M. A., and P. Unitt (2002). Diagnosability versus mean differences of Sage Sparrow subspecies. The Auk 119:26–35.

Picman, J. (1977). Intraspecific nest destruction in the Long-billed Marsh Wren, *Telmatodytes palustris palustris*. Canadian Journal of Zoology 55:1997–2003.

Podos, J., and P. S. Warren (2007). The evolution of geographic variation in birdsong. Advances in the Study of Behavior 37:403–458.

Price, J. J. (1998). Family- and sex-specific vocal traditions in a cooperatively breeding songbird. Proceedings of the Royal Society of London, Series B 265:497–502.

Pyle, P. (1997). Identification Guide to North American Birds, part 1 Columbidae to Ploceidae, second edition. Slate Creek Press, Bolinas, CA, USA.

Ramenofsky, M., and J. C. Wingfield (2007). Regulation of migration. Bioscience 57:135–143.

R Core Team (2014). R: A Language and Evironment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

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.

Riebel, K., and P. J. B. Slater (1998). Male Chaffinches (*Fringilla coelebs*) can copy calls from a tape tutor. Journal für Ornithologie 139:353–355.

Roach, S. P., and L. S. Phillmore (2017). Geographic variation in song structure in the Hermit Thrush (*Catharus guttatus*). The Auk: Ornithological Advances 134:612–626.

Rothstein, S. I., and R. C. Fleischer (1987). Vocal dialects and their possible relation to honest status signalling in the Brown-headed Cowbird. The Condor 89:1–23.

Ryan, M. J., and E. A. Brenowitz (1985). The role of body size, phylogeny, and ambient noise in the evolution of bird song. The American Naturalist 126:87–100.

Ryan, M. J., and W. Wilczynski (1991). Evolution of intraspecific variation in the advertisement call of a cricket frog (*Acris crepitans*, Hylidae). Biological Journal of the Linnean Society 44:249–271.

Snowberg, L. K., and C. W. Benkman (2007). The role of marker traits in the assortative mating within Red Crossbills, *Loxia curvirostra* complex. Journal of Evolutionary Biology 20:1924–1932.

Soha, J. A., D. A. Nelson, and P. G. Parker (2004). Genetic analysis of song dialect populations in Puget Sound White-crowned Sparrows. Behavioral Ecology 15:636–646.

Spiess, A.-N. (2014). qpcR: Modelling and anlaysis of real-time PCR data. https://cran.r-project.org/web/packages/qpcR/index.html

Stefanski, R. A., and J. B. Falls (1972a). A study of distress calls of Song, Swamp, and White-throated sparrows (Aves: Fringillidae). I. Intraspecific responses and functions. Canadian Journal of Zoology 50:1501–1512.

Stefanski, R. A., and J. B. Falls (1972b). A study of distress calls of Song, Swamp, and White-throated sparrows (Aves: Fringillidae). II. Interspecific responses and properties used in recognition. Canadian Journal of Zoology 50:1513–1525.

Stephan, C., and K. Zuberbühler (2008). Predation increases acoustic complexity in primate alarm calls. Biology Letters 4:641–644.

Sturge, R. J., K. E. Omland, J. J. Price, and B. Lohr (2016). Divergence in calls but not songs in the Orchard Oriole complex: *Icterus spurius* and *I. fuertesi*. Journal of Avian Biology 47:109–120.

Suzuki, T. N. (2016). Semantic communication in birds: Evidence from field research over the past two decades. Ecological Research 31:307–319.

Toews, D. P. L., and D. E. Irwin (2012). Pacific Wren (*Troglodytes pacificus*). In Birds of North America Online (P. G. Rodewald, Editor). Cornell Lab of Ornithology, Ithaca, NY, USA. https://birdsna.org/Species-Account/bna/species/pacwre1

Townsend, C. W. (1905). The birds of Essex County, Massachusetts. Memoirs of the Nuttall Ornithological Club 3.

Venables, W. N., and B. D. Ripley (2002). Modern Applied Statistics with S, fourth edition. Springer, New York, NY, USA.

Verner, J. (1965). Breeding biology of the Long-billed Marsh Wren. The Condor 67:6–30.

Verner, J. (1975). Complex song repertoire of male Long-billed Marsh Wrens in eastern Washington. The Living Bird 14:263–300.

Wallschläger, D. (1980). Correlation of song frequency and body weight in passerine birds. Cellular and Molecular Life Sciences 36:412.

Weilgart, L., and H. Whitehead (1997). Group-specific dialects and geographical variation in coda repertoire in South Pacific sperm whales. Behavioral Ecology and Sociobiology 40:277–285.

Wells, M. M., and C. S. Henry (1992). The role of courtship songs in reproductive isolation among populations of Green Lacewings of the genus *Chrysoperla* (Neuroptera: Chrysopidae). Evolution 46:31–42.

Welter, W. A. (1935). The natural history of the long-billed marsh wren. The Wilson Bulletin 47:3–34.

West-Eberhard, M. J. (1983). Sexual selection, social competition, and speciation. The Quarterly Review of Biology 58:155–183.

Wiley, R. H., and D. G. Richards (1978). Physical constraints on acoustic communication in the atmosphere: Implications for the evolution of animal vocalizations. Behavioral Ecology and Sociobiology 3:69–94.

Wilkins, M. R., N. Seddon, and R. J. Safran (2013). Evolutionary divergence in acoustic signals: Causes and consequences. Trends in Ecology & Evolution 28:156–166.

Wright, T. F., A. M. Rodriguez, and R. C. Fleischer (2005). Vocal dialects, sex-biased dispersal, and microsatellite population structure in the parrot *Amazona auropalliata*. Molecular Ecology 14:1197–1205.

Yasukawa, K. (1989). The costs and benefits of a vocal signal: The nest-associated "chit" of the female Red-winged Blackbird, *Agelaius phoenicus*. Animal Behaviour 38:866–874.

Zann, R. (1993). Structure, sequence and evolution of song elements in wild Australian Zebra Finches. The Auk 110:702–715.

APPENDIX

APPENDIX TABLE 4. Repeated-measures ANOVA (twitter: df = 2 and 72; trill: df = 2 and 130) demonstrated that several measures of the Marsh Wren twitter call vary significantly throughout the duration of the call after correcting for multiple testing (Benjamini-Hochberg false discovery rate, α = 0.05). Variables are ranked from lowest to highest P value; values in bold are significant.

Rank	Variable	Call type	F	P
1	Lower frequency limit	Twitter	23.56	<**0.001**
2	Upper frequency limit	Twitter	22.46	<**0.001**
3	Note 2 duration	Twitter	10.91	<**0.001**
4	Peak frequency	Twitter	9.04	<**0.001**
5	Lower frequency limit	Trill	4.32	**0.015**
6	Bandwidth	Twitter	3.85	**0.026**
7	Inter-note space 1 duration	Twitter	3.64	**0.031**
8	Inter-note space 1 duration	Trill	2.51	0.086
9	Upper frequency limit	Trill	1.85	0.161
10	Peak frequency	Trill	1.28	0.281
11	Note 4 duration	Trill	1.21	0.303
12	Inter-note space 2 duration	Trill	1.09	0.339
13	Inter-note space 3 duration	Trill	1.02	0.37
14	Note 1 duration	Twitter	0.86	0.426
15	Note 1 duration	Trill	0.76	0.468
16	Note 3 duration	Twitter	0.66	0.519
17	Note 2 duration	Trill	0.66	0.52
18	Note 4 duration	Twitter	0.59	0.562
19	Bandwidth	Trill	0.37	0.691
20	Note 3 duration	Trill	0.34	0.711
21	Inter-note space 2 duration	Twitter	0.05	0.947

APPENDIX TABLE 6. Percent correct classification at the subspecies level using linear discriminant analysis (LDA).

Call type	Sample size	Percent correct classification from original LDA	Overall percent correct classification
Buzz	*dissaeptus*: 61	87	65
	griseus: 15	33	
	mariainae: 10	10	
	palustris: 26	54	
Chuck	*dissaeptus*: 14	36	42
	griseus: 10	30	
	marianae: 14	43	
	palustris: 24	50	
Churr	*dissaeptus*: 17	50	38
	griseus: 9	0	
	marianae: 7	14	
	palustris: 22	52	
Rattle	*dissaeptus*: 21	62	28
	griseus: 6	0	
	marianae: 16	6	
	palustris: 10	10	
Scream	*dissaeptus*: 14	64	33
	griseus: 10	0	
	marianae: 7	0	
	palustris: 8	50	
Trill	*dissaeptus*: 11	73	48
	griseus: 28	61	
	marianae: 17	24	
	palustris: 11	27	
Twitter	*dissaeptus*: 0	NA	53
	griseus: 11	45	
	marianae: 13	62	
	palustris: 16	50	

APPENDIX TABLE 5. Comparison of acoustic measures for trill vs. twitter calls. Thirteen of 16 pairwise tests (Welch's t-test or Wilcoxon signed-rank test) showed significant differences after correcting for multiple testing (Benjamini-Hochberg false discovery rate, α = 0.05). Variables ranked from lowest to highest P value; values in bold are significant.

Rank	Variable	Test	df	Test statistic	P
1	Note rate	Welch	112.684	11.869	<**0.001**
2	Duration of note 2, beginning	Welch	51.396	−16.438	<**0.001**
3	Duration of note 2, middle	Welch	47.007	−16.954	<**0.001**
4	Duration of note 1, beginning	Wilcoxon	–	438	<**0.001**
5	Number of notes	Welch	109.3	7.016	<**0.001**
6	Inter-note space 2 duration, beginning	Welch	106.355	−6.091	<**0.001**
7	Call duration	Wilcoxon	–	2305	<**0.001**
8	Upper frequency limit, middle	Welch	62.576	−5.657	<**0.001**
9	Upper frequency limit, beginning	Wilcoxon	–	938	<**0.001**
10	Bandwidth	Welch	55.469	−3.948	<**0.001**
11	Lower frequency limit, middle	Wilcoxon	–	991	<**0.001**
12	Inter-note space 3 duration	Welch	103.787	3.5256	<**0.001**
13	Lower frequency limit, beginning	Welch	99.466	−2.528	**0.013**
14	Inter-note space 1 duration	Wilcoxon	–	1376	0.173
15	Peak frequency, beginning	Welch	112.033	1.243	0.217
16	Peak frequency, middle	Wilcoxon	–	1635	0.935

The Auk: Ornithological Advances 135:127–151, © 2018 American Ornithological Society

APPENDIX TABLE 7. Percent correct classification at the migratory-pattern level (full migrants, partial migrants, and nonmigrants) using linear discriminant analysis (LDA).

Call type	Sample size	Percent correct classification from original LDA	Overall percent correct classification
Buzz	Full: 61	90	71
	Part: 30	40	
	Non: 25	57	
Chuck	Full: 14	28	37
	Part: 25	21	
	Non: 24	58	
Churr	Full: 17	50	37
	Part: 26	33	
	Non: 16	31	
Rattle	Full: 21	48	36
	Part: 12	0	
	Non: 22	41	
Scream	Full: 14	0	21
	Part: 9	13	
	Non: 17	41	
Trill	Full: 11	64	71
	Part: 16	31	
	Non: 45	87	
Twitter	Full: 0	NA	69
	Part: 21	62	
	Non: 24	75	

APPENDIX TABLE 8. Percent correct classification at the habitat-type level using linear discriminant analysis (LDA). Sample-size differences make the percent correct classification data difficult to interpret. In place of overall correct classification, we show the percentage of freshwater-marsh birds that fall outside a 95% confidence interval (CI) for bootstrapped saltmarsh-bird linear discriminant values. Twitter could not be analyzed at this level (see text).

Call type	Sample size	Percent correct classification from original LDA	Percentage of freshwater-marsh birds outside the 95% CI for saltmarsh birds
Buzz	Fresh: 61	89	39
	Salt: 55	82	
Chuck	Fresh: 14	14	15
	Salt: 48	92	
Churr	Fresh: 16	31	32
	Salt: 37	89	
Rattle	Fresh: 21	52	29
	Salt: 32	75	
Scream	Fresh: 14	0	0
	Salt: 26	88	
Trill	Fresh: 11	64	72
	Salt: 61	95	
Twitter	Fresh: 0	NA	NA
	Salt: 45	NA	