Song and Male Quality in Prairie Warblers

Bruce E. Byers*, Michael E. Akresh† & David I. King‡

* Department of Biology, University of Massachusetts Amherst, Amherst, MA, USA
† Department of Environmental Conservation, University of Massachusetts Amherst, Amherst, MA, USA
‡ U.S. Forest Service Northern Research Station, University of Massachusetts Amherst, Amherst, MA, USA

Abstract

To determine if the songs of male prairie warblers could potentially reveal to female listeners information about the quality of singers, we compared various aspects of prairie warbler song structure and performance to attributes that might reflect a male singer's potential to enhance the fitness of his mate. We found that all the tested male attributes—arrival date, age, body size, annual survival, and fledging success—were associated with singing, most with multiple aspects of singing. Several of the song traits that were associated with potential indicators of male quality had also been found previously to be good predictors of female social mate choice. In particular, longer songs with longer elements, performed at lower frequency and with greater consistency, were associated with both female mate choice and potential indicators of quality. Thus, female prairie warblers may assess potential mates with the help of a set of song characteristics that collectively reveal an array of attributes that together indicate overall male quality.

Introduction

Although an arbitrary signal that conveys no information about the signaler can in principle evolve (Lande 1981; Kirkpatrick 1982), it is likely that at least some signals evolve because they reveal differences among signalers in traits that affect the receiver (Zahavi 1975; Grafen 1990). Selection should favor the persistence of such signals if receivers gain material or genetic benefits by responding to them (Maynard Smith & Harper 2003). Receivers might gain such benefits if, for example, the signal encodes useful information about a potential rival or mate.

In the context of female mate choice, the degree to which a male possesses attributes that would increase his mate's fitness is often characterized as his "quality" (Wilson & Nussey 2010; Bergeron et al. 2011). For example, a male bird might be of high quality if he has an unusually strong predilection to feed offspring and thereby increases his mate's reproductive success by providing an above-average level of parental care (Siefferman & Hill 2003; Stodola et al. 2010). Or a high-quality male might be one with an unusually

robust immune system, who increases his mate's fitness by contributing alleles that improve the immunocompetence of offspring that inherit them (Johnsen et al. 2000; Garvin et al. 2006; Raveh et al. 2014).

Other male attributes may not themselves be inherently beneficial to mates' fitness, but may nonetheless indicate a generally higher capacity to prove material or genetic benefits. For example, a male bird's body size might reflect his condition or viability, and therefore his ability to provide resources (Candolin & Voigt 2001), parental care (Petrie 1983) or, "good genes" (Briggs 2013). Similarly, a male bird's age might reflect his ability to provide resources (Part 2001), beneficial alleles (Hegyi et al. 2006) or, viable sperm (Møller et al. 2009).

If a variable signal had been shaped by selection to reliably reveal male quality, we would expect that the signal and male quality would be correlated. Such a correlation is not by itself sufficient evidence to demonstrate that the signal evolved via selection for reliable indication of male quality (Kokko et al. 2003; Prum 2010), but any signal that evolved in this

fashion would yield such correlations. Thus, testing for correlations between signal features and male attributes that might reflect quality can test a condition that is necessary, even if not sufficient, to support an inference that a signal serves as a reliable indicator of male quality and has evolved for that function.

We searched for evidence of such correlations in a population of prairie warblers (*Setophaga discolor*). In particular, we compared measurements of various aspects of song performance with various male attributes that might indicate quality. The song performance features tested were the same ones we considered in an earlier comparison (Byers et al. 2015) of songs and first-egg date, a proxy for female social mate choice. The male attributes we considered included body size and age, as well as three attributes —arrival date, fledging success and, survival—that are not necessarily quality measures *per se*, but that might be expected to vary as a direct consequence of variation in male quality. Males that arrive at their breeding grounds earlier may be those whose higher quality allowed them to be in better condition and leave earlier for migration (Kokko 1999; Newton 2008). Males that successfully fledge more young may be those whose higher quality enabled them to acquire territories with good nest sites and ample food (Siikamaki 1995; Sergio et al. 2009). Males that live longer may do so at least in part due to advantages attributable to their high quality (Bize et al. 2008; Saino et al. 2012).

Prairie warbler songs, like those of many other wood-warbler species, fall into two categories, dubbed first category and second category (Spector 1992). In prairie warblers, songs of the two categories differ in acoustic structure and in the contexts in which they are used (Nolan 1978; Houlihan 2000). Second category songs are of longer duration and lower pitch than first category songs, and are composed of longer elements (Fig. 1). In addition, second category songs (but not first category songs) are used in a distinctive mode of singing that takes place at dawn and consists of steadily delivered songs interspersed with characteristic chips. Because songs and singing in the two categories are so different, we considered each category separately in our analyses.

Methods

Arrival Date, Age, and Size

Our study is based on data collected in the Montague Plains Wildlife Management Area (N42°34′, W72°31′), a 600 ha pitch pine-scrub oak barren in
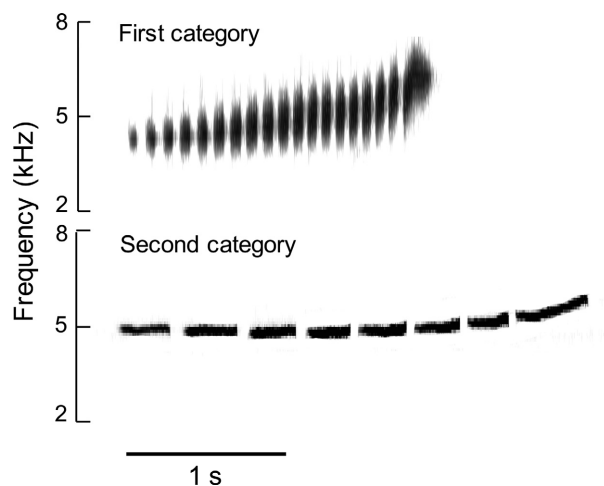

Fig. 1: Spectrograms (512 point FFT) of prairie warbler songs, showing an example of a first category and an example of a second category song.

Franklin County, Massachusetts that contains a mosaic of thickets, thinned pine forests, and powerline corridors. In 2009 and 2010, we monitored the prairie warblers in the study area, beginning when the first males arrived on the breeding grounds in late April. Intensive surveys during the first few weeks of the breeding season enabled us to determine the arrival dates of almost all breeding males.

Most males and many females were captured in mist nets and fitted with a USGS aluminum band and a unique combination of plastic color bands. At the time of capture, birds were aged on the basis of plumage characteristics (Pyle 1997) and classified as second year (SY) or after second year (ASY). Captured birds were also weighed and measured (tarsus length, wing length, and tail length). (Some birds in the study area were captured, measured, and banded in 2008, prior to the current study; size measurements from these birds were excluded from our analysis).

In subsequent analysis of body size, we opted not to consider weight, because time between arrival and capture varied among males, so males were not all weighed at the same stage of the breeding cycle, and weight measurements were therefore not readily comparable among males. The remaining three size variables were intercorrelated, so we saw little value in increasing our number of comparisons by considering each one separately. We instead constructed a single size variable by performing a principal components analysis on the three variables for the 102 males for which we measured all three variables 2008–2010. In the results of this analysis, the first principal component was the only component with an eigenvalue

above one and it accounted for 68% of variance in the variables (component loadings: wing length 0.65, tail length 0.63, tarsus length 0.44). We used the equation for the first principal component to calculate principal component scores for the 30 males for which we had both song samples and size measurements made in the year in which the songs were recorded. We used these scores to represent body size.

Survival and Fledging Success

Although the current study focuses on 2009–2010, we monitored the study area for the presence of banded prairie warblers from 2008 to 2015, and measured the fledging success of the birds in the study area from 2009 to 2013. We were therefore able to base our assessments of survival and fledging success on multi-year evaluations. We reasoned that if variation in these attributes is a consequence of variation in male quality, then the intrinsic quality of a male would be best assessed by estimating the two attributes over as much of his life as possible.

From our records of the banded males detected in each year, we measured survival as probability of survival from one year to the next, estimated with a Cormack-Jolly-Seber model (Lebreton et al. 1992). Because this estimation method assumes that surviving birds remain in the study population, it would underestimate the survival probability of birds that survived but nested far outside of our study area (our surveys for banded birds included the study area and suitable habitat within 5 km of it). The impact of such underestimation would be small, however, if surviving males generally returned to the vicinity of their breeding site. Evidence that such site fidelity may be prevalent in our study population includes a relatively high observed rate of return among breeding males, e.g., 72% annual return rate of banded males 2008–2012 (Akresh et al. 2015).

To measure fledging success, we searched for and monitored nests to determine how many, if any, young were fledged. To find nests, we observed parent behavior and conducted systematic searches (Martin & Geupel 1993). After a nest was discovered, we checked it every 2–4 d until the nestlings fledged or the contents of the nest had disappeared. Nestlings were color-banded when they were approximately 8 d old. To determine whether and how many young fledged, we searched systematically for fledglings in territories where nests were found empty and young could have fledged based on their stage of development in the previous nest check. Because our visits to each territory were relatively frequent and fledglings

were generally loud, conspicuous, and attended by color-banded adults, we are confident that we detected most of the fledglings produced by the birds in our sample. In some years, territories in certain portions of the study area were visited less frequently than our norm; all nesting data from these territories were excluded from our analysis in order to avoid potential bias resulting from reduced likelihood of detecting nesting activity.

Songs

During the breeding seasons of 2009 and 2010, we recorded approximately 9500 songs from 73 male prairie warblers, eight of which were recorded in both years. The sampling period extended from 4 May to 27 June in 2009, and from 6 May to 28 June in 2010. Our sampling protocol ensured that each bird contributed songs recorded at various times of day (although not all birds were recorded during the dawn bout) and yielded a sample in which the proportion of songs recorded in each stage of the breeding cycle was similar across birds. From our song sample, we selected for analysis a set of song bouts that met all of the following criteria: (1) The bout included least five consecutive songs; (2) the recording quality for all songs in the bout was good (minimal reverberation and background noise); (3) the recorded individual had been recorded on at least four different days in the same year; (4) the recorded individual was one for which we had collected data on arrival date, size, age, fledging success, and/or survival. The resulting sample included 5391 songs in 563 bouts (mean songs/bout 9.6, standard deviation 8.1, range 5–43) from 45 birds (mean bouts/bird 12.5, standard deviation 7.2, range 4–34).

For each song in our focal sample, we measured or calculated the same features that we had previously considered as potential predictors of a proxy for female mate choice (Table 1). The song features were chosen to include a range of traits that have been linked to aspects of mate choice in other songbird species (see Byers et al. 2015 for details). The directly measured features were intersong interval (a measure of singing rate), lowest frequency, frequency of peak power, frequency bandwidth, song duration, and the number of elements in a song. We also calculated the element rate for each song (number of elements divided by song duration). For subsequent analysis, we used the per-bout means of these variables. For dawn samples only, we calculated the chip rate of each bout (the number of chips in a bout divided by the sum of the bout's intersong intervals). In addition,

Table 1: Means and standard deviations of traits measured in a sample of prairie warbler songs

	1st category (mean ± SD)	2nd category (mean ± SD)
Intersong interval (s)	13.8 ± 3.4	13.9 ± 6.3
Lowest frequency (Hz)	4189.4 ± 155.6	4268.1 ± 172.0
Peak frequency (Hz)	5510.1 ± 280.2	5100.5 ± 239.5
Bandwidth (Hz)	2946.1 ± 269.3	1788.3 ± 466.3
Duration (s)	1.9 ± 0.2	2.7 ± 0.4
Number of elements	19.3 ± 3.9	8.9 ± 3.6
Element rate (elements/s)	10.2 ± 1.8	3.4 ± 1.6
Chip rate (chips/s)	NA	1.6 ± 0.5
Consistency (PC1 score)	0.2 ± 0.4	0.3 ± 0.5

Values are shown separately for first category songs (n = 50 birds) and second category songs (n = 28 birds).

we calculated song performance consistency for each bout. To derive this measure, we first calculated the within-bout coefficient of variation (CV) for each song variable (except chip rate) in a larger dataset that included all good quality song bouts recorded in 2009 and 2010. We then performed a principal components analysis on the CVs, which were highly intercorrelated. The first principal component in this analysis loaded positively on all eight contributing variables and explained about 90% of the variance in the CVs. We therefore used each bout's PC1 score to represent its performance consistency, reversing its sign, so that higher values indicated lower coefficients of variation (i.e., greater consistency) across the spectral and temporal features of the songs in a bout.

All song measurements were made in RAVEN 1.4 (Bioacoustics Research Program 2011), using onscreen cursors. Time measurements were made on spectrograms (128 point FFT) and frequency measurements were made on power spectra (1024 point FFT). Highest frequency (measured to obtain bandwidth) and lowest frequency were measured as the upper and lower frequencies with amplitudes −30 dB relative to that of the peak frequency, respectively.

Statistical Analysis

To determine if song features were associated with arrival date, size, or age, we fit a series of linear mixed models with these male attributes as predictors and song features as response variables. Because our data included multiple non-independent song samples (bouts) from each male, bird ID was included in each model as a random factor. We chose to analyze each male attribute separately because measurements for each attribute were available for a different subset of

the males for which we had suitable song samples (Table 2), and conducting separate analyses allowed us to maximize sample size for each trait. In addition, each trait was analyzed separately for first category songs and second category songs.

To model the association between song features and annual survival, we included per-bird means of song variables as covariates in a Cormack-Jolly-Seber model that estimated annual survival from encounter history data. Before adding song covariates to the model, we fit models in which estimates of survival probability and detection probability were either constrained to be constant over time or allowed to vary from year to year. We fit four models based on the four possible combinations of these conditions, and found that of these models the one that best fit the data (lowest AIC value) was the one in which survival probability varied between years but probability of detecting a bird that was present remained constant. We therefore used this model as the baseline to which the song covariates were added.

To determine if song features were associated with fledging success, we fit mixed models with per-male means of song variables as predictors and number of fledglings as the response variable. Because the fledging success of many males was measured in 2 or more years, we included bird ID in the models as a random factor. The distribution of fledgling counts was not a good fit to a Poisson distribution, due to a large number of territories with zero counts (36% of the male territories in our sample failed to fledge young). We therefore fit zero-altered Poisson models (hurdle models) in which fledging success was modeled as a two-part process in which success (some offspring) vs. failure (zero offspring) was modeled as a binary outcome with a binomial distribution, and successful

Table 2: Numbers of males in the subsets of our overall sample for which measurements of various male attributes were available, and the number of song bouts measured from those males

Attribute	n birds	n song bouts
First category songs		
Arrival date	44	318
Age	36	250
Size	30	206
Survival	41	284
Fledging success	37	241
Second category songs		
Arrival date	26	195
Size	18	141
Survival	22	159
Fledging success	20	155

outcomes entered a sub-model in which fledgling counts were assumed to be drawn from a truncated Poisson distribution (Zuur et al. 2009). This approach assumed that the process leading to variation in the binary outcome was independent of the process leading to variation in the conditional count outcome, as would be the case if, for example, complete nesting failure is due largely to predation (as we believe to be the case in our study population), while the number of offspring in nests that are not depredated is determined mainly by other factors.

Except for the binary variable age (SY or ASY) and the binary and count variables that characterized fledging success, all analyzed variables were centered and standardized prior to analysis, to have a mean of 0 and a standard deviation of 1. Such standardization ensures that the magnitudes of estimated coefficients in the models are comparable across variables, and can be viewed as standardized effect sizes (Schielzeth 2010). For the binary predictors, standardized estimated coefficients represent the predicted difference in the mean value of the response variable, in standard deviations, between the two levels of the predictor. We coded the binary variables so that the reference level (coded as 0) was SY for age, and failure to fledge young for fledging success.

For each male (non-song) attribute, we screened our set of fitted models to identify informative associations between male attributes and song traits. For our models of arrival date, size, and age (each of which fit a single main effect), we defined informative associations as those with estimated coefficients whose 95% confidence intervals did not span zero (roughly equivalent to $p < 0.05$) or that came very close to this criterion. For our models of annual survival and fledging success (which fit multiple main effects), we first used a likelihood ratio test to compare a full model containing all of the song variables to a baseline model that excluded the song variables. If the full model provided a better fit ($p < 0.05$) to the data than the baseline model, we used likelihood ratio tests to compare the full model to models from which a single song variable had been removed, and deemed an association informative if this test yielded a p-value <0.05. Although it is likely that some associations detected by our criteria arose as a result of cumulative type 1 error (our statistical analysis included 86 different comparisons), we felt that it would be most useful to use a comparatively relaxed criterion, given that studies comparable to ours tend to have relatively low statistical power (Jennions & Møller 2003) even as they attempt to

detect effects that tend to be of relatively small magnitude (Møller & Jennions 2002).

Two-part (hurdle) models for fledging success were fit using the R (R Core Team 2014) package glmmADMB (Fournier et al. 2012; Skaug et al. 2015). Models for annual survival were fit with MARK (White & Burnham 1999), using the interface provided by the R package RMARK (Laake 2013). All other statistical analyses were performed using STATA 13 (StataCorp 2013).

Results

In this section, we first describe our findings for each of the male attributes we analyzed. We conclude the section by summarizing our findings, with the summary broken down by song category.

Arrival Date

Five first category song traits yielded associations that met our screening criterion with respect to arrival date (Fig. 2). Lowest frequency and element rate were positively correlated with arrival date (i.e., males that arrived earlier sang at lower frequency and with lower element rate and therefore longer elements). Performance consistency, duration, and bandwidth were negatively correlated with arrival date (i.e., males that arrived earlier sang songs with higher consistency, longer duration, and larger bandwidth).

Associations with three-second category song traits met our criterion with respect to arrival date (Fig. 2). Bandwidth was positively correlated with arrival date, whereas lowest frequency and intersong interval were negatively correlated with arrival date (i.e., males that arrived earlier sang with smaller bandwidth, higher frequency, and slower rate).

Age

Only one association between first category song traits and age met our criterion (Fig. 3). ASY (older) birds tended to sing at lower frequency than did SY birds.

We were unable to fit models relating second category song traits to age because our sample of second category songs contained songs from only one SY bird.

Size

Two associations between first category song traits and body size met our criterion (Fig. 4). Element rate and number of elements per song were positively

Fig. 2: Standardized coefficients and their 95% confidence intervals from linear mixed models in which the dependent variable was the date of male prairie warbler arrival on the breeding grounds, and the predictor variables were aspects of male song performance. Vertical reference lines indicate zero (solid line) and ± 0.1 (dotted lines).
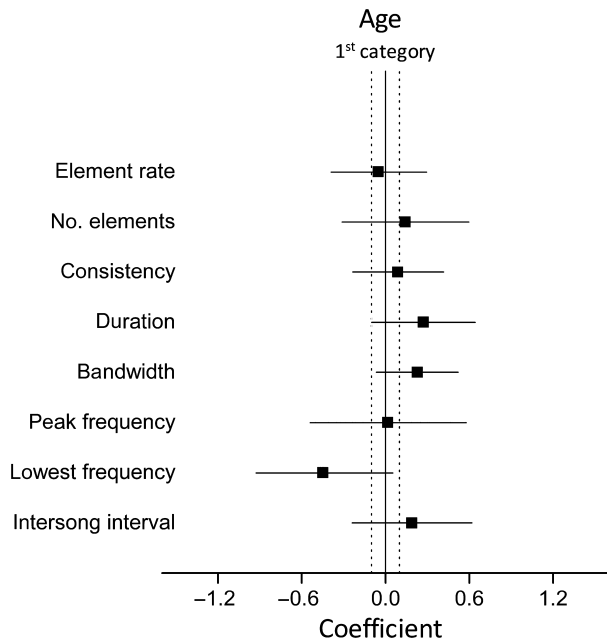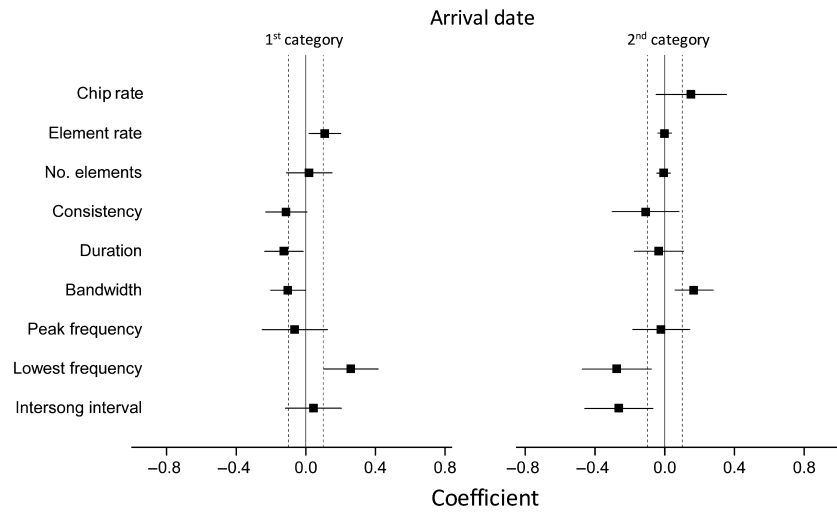

Fig. 3: Standardized coefficients and their 95% confidence intervals from linear mixed models in which the dependent variable was the age of male prairie warblers [second year (SY) vs. after second year (ASY)], and the predictor variables were aspects of male song performance.

correlated with size (i.e., smaller birds tended to have fewer elements and lower element rates, and therefore longer elements).

Three associations between second category song traits and age met our criterion (Fig. 4). Bandwidth and chip rate were negatively correlated with bird size, and intersong interval was positively correlated with bird size (i.e., smaller birds tended to have songs with broader bandwidths and to chip and sing more rapidly).

Survival

One association between first category song traits and survival met our criterion. Performance consistency was positively correlated with annual survival probability (Table 3).

For second category songs, three associations with survival met our criterion (Table 3). Song duration was positively correlated with annual survival probability (i.e., birds with higher probability of annual survival sang longer songs). Element rate and number of elements were negatively correlated with annual survival probability (i.e., birds with higher probability of annual survival sang songs that contained fewer, longer elements).

Fledging Success

In our hurdle models for fledging success, the model component that modeled fledging success vs. failure yielded four associations with first category song variables (Table 4). Peak frequency, song duration, element rate, and number of elements were negatively correlated with success (i.e., males that fledged young sang shorter, lower frequency songs with fewer, longer elements).

Among the potential associations with second category song traits, three met our criterion with respect to fledging success vs. failure (Table 4). Song duration was positively correlated with success, whereas element rate and number of elements were negatively correlated with success (males that fledged young had longer songs with fewer, longer elements).

In the hurdle model component that modeled fledgling number conditional on fledging success,
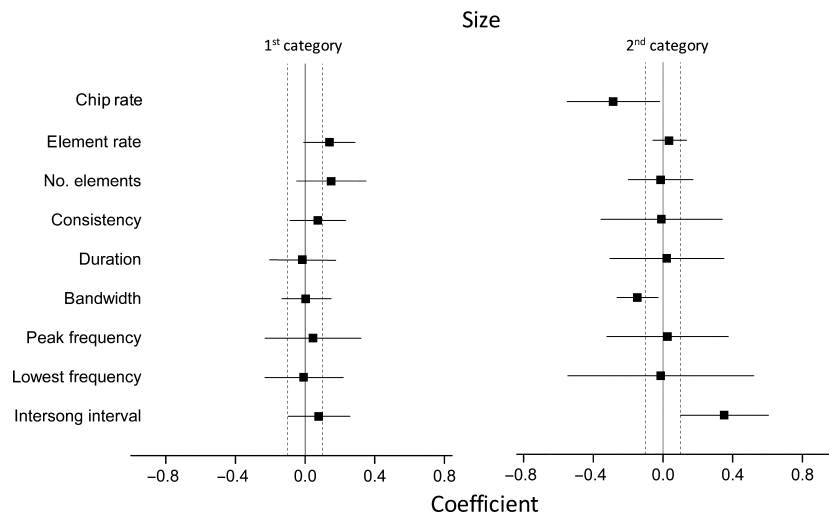
Fig. 4: Standardized coefficients and their 95% confidence intervals from linear mixed models in which the dependent variable was a principal component score representing the size of male prairie warbler arrival on the breeding grounds, and the predictor variables were aspects of male song performance.

Table 3: Estimated standardized coefficients of covariates from Cormack-Jolly-Seber models that estimated annual survival probability

Song variable	Coefficient	SE	95% CI	χ^2	df	p
First category						
Element rate	0.25	1.43	−2.5, 3.1	0.03	1	0.9
No. elements	0.19	1.45	−2.7, 3.0	0.02	1	0.9
Consistency	0.37	0.15	0.1, 0.8	5.19	1	0.02
Duration	0.57	0.73	−0.9, 2.0	0.62	1	0.4
Bandwidth	−0.38	0.54	−1.4, 0.7	0.49	1	0.5
Peak frequency	0.47	0.33	−0.2, 1.1	2.11	1	0.1
Lowest frequency	−0.12	0.40	−0.9, 0.7	0.09	1	0.8
Intersong interval	−0.12	0.24	−0.6, 0.3	0.25	1	0.6
Second category						
Chip rate	1.08	0.61	−0.1, 2.3	3.64	1	0.05
Element rate	−6.20	2.81	−11.7, −0.7	4.96	1	0.03
No. elements	−6.74	3.01	−12.6, −0.8	5.14	1	0.02
Consistency	−0.23	0.60	−1.4, 0.9	0.14	1	0.7
Duration	8.12	3.78	0.7, 15.5	5.26	1	0.02
Bandwidth	−0.53	0.56	−1.6, 0.6	0.83	1	0.4
Peak frequency	0.60	0.61	−0.6, 1.8	0.98	1	0.3
Lowest frequency	−0.61	0.59	−1.8, 0.5	1.10	1	0.3
Intersong interval	−0.03	0.40	−0.8, 0.7	0.05	1	0.9

Coefficients, standard errors, and confidence intervals are from a full model containing all song variables. Chi-square and p values are from likelihood ratio tests comparing the full model to a model with a song variable removed. The full model fit the data significantly better than a baseline model with no song covariates (first category songs: $\chi^2 = 18.59$, df = 8, p = 0.017; second category songs: $\chi^2 = 21.10$, df = 9, p = 0.012).

the full model containing all song variables yielded no significant improvement over a baseline (intercept-only) model, for both first category and second category songs. Thus, we detected no associations between song variables and number of fledglings.

Summary, First Category Songs

We found 13 correlations between male attributes and first category song traits. The correlations involved six of the eight song traits we considered. Males with lower frequency first category songs tended to arrive earlier, be older, and fledge young. Males singing first category songs with greater consistency tended to arrive earlier and have higher annual survival probability. Males whose first category songs had lower element rate (longer elements) tended to arrive earlier, be smaller, and fledge young. Males whose first category songs contained fewer elements tended to be smaller and to fledge young. Males that sang longer first category songs tended to arrive earlier but fail to fledge young. Males whose first category songs had broader bandwidths tended to arrive earlier.

Summary, Second Category Songs

We found 14 correlations between male attributes and second category song traits. The correlations involved eight of the nine song traits we considered. Males that sang second category songs with greater consistency tended to arrive earlier. Males whose second category songs had a lower element rate (longer elements) tended to have higher annual survival probability and fledge young. Males that sang longer second category songs tended to have higher annual survival probability and fledge young. Males that sang second category songs at low frequency tended to arrive later. Males that used second category songs with broader bandwidth tended to arrive later and be smaller. Males with faster rates of second category

Table 4: Estimated standardized coefficients from the binary response component of zero-altered Poisson (hurdle) mixed models that modeled predictors of fledging success

Song variable	Coefficient	SE	95% CI	χ^2	df	p
First category						
Element rate	−5.66	2.75	−11.1, −0.3	5.02	1	0.03
No. elements	−7.67	3.74	−15.0, −0.3	4.65	1	0.03
Consistency	0.26	0.43	−0.6, 1.1	0.39	1	0.5
Duration	−6.45	3.40	−12.1, 0.2	5.57	1	0.02
Bandwidth	0.81	0.68	−0.5, 2.2	1.48	1	0.2
Peak frequency	−1.27	0.58	−2.1, −0.1	5.76	1	0.02
Lowest frequency	−0.09	0.53	−1.1, 0.9	0.03	1	0.8
Intersong interval	0.53	0.41	−0.3, 1.4	1.89	1	0.2
Second category						
Chip rate	1.3	2.01	−2.7, 5.3	0.57	1	0.4
Element rate	−4.85	1.88	−8.5, −1.0	9.33	1	0.002
No. elements	−3.79	1.47	−6.7, −0.9	9.36	1	0.002
Consistency	−2.12	1.84	−5.7, 1.5	2.93	1	0.09
Duration	5.41	2.19	1.1, 9.7	7.37	1	0.007
Bandwidth	−3.43	3.12	−9.5, 2.7	3.42	1	0.07
Peak frequency	0.77	1.36	−1.9, 3.4	0.36	1	0.5
Lowest frequency	0.33	1.55	−2.7, 3.4	0.04	1	0.8
Intersong interval	−0.35	1.00	−2.3, 1.6	0.12	1	0.7

Coefficients, standard errors, and confidence intervals are from a full model containing all song variables. Chi-square and p values are from likelihood ratio tests comparing the full model to a model with a song variable removed. The full model fit the data significantly better than an intercept-only model with no song covariates (first category songs: $\chi^2 = 16.74$, df = 8, p = 0.033; second category songs: $\chi^2 = 18.88$, df = 9, p = 0.026).

singing tended to arrive later and be smaller. Males whose second category songs contained fewer elements tended to have higher annual survival rates and fledge young. Males with faster chipping rates at dawn tended to be smaller.

Discussion

Song Traits Associated with Social Mate Choice

In an earlier study (Byers et al. 2015), we identified six song traits—faster singing rate, lower song frequency, and greater performance consistency for first category songs; slower element rate (longer elements), longer song duration, and greater performance consistency for second category songs—that were good predictors of a proxy for social mate choice in our study population. Our current results suggest that these song traits also reveal information about the quality of the singer. With one exception (singing

rate of first category songs), each of the first and second category song traits that best predicted social mate choice was also correlated with at least one male attribute that might indicate higher phenotypic quality, and most of these song traits were associated with multiple male attributes.

Thus, our data reveal a pattern of broad statistical association between a set of song traits that collectively predict social mate choice and an array of male attributes that may collectively indicate male quality. This pattern suggests that listening female prairie warblers could in principle use a combination of different song performance cues to assess the quality of singers. The pattern is also consistent with the prediction from theory that when multiple signals function in mate choice, it is more likely that the different signals collectively signal overall quality than that each signal reveals a different component of male quality (Møller & Pomiankowski 1993; Johnstone 1996; Kokko et al. 2003).

Song Traits Not Associated with Social Mate Choice

Some song features that were excluded from our best models for predicting social mate choice (Byers et al. 2015) were nonetheless correlated with male attributes in the current study. Given the association of these "extra" song traits with attributes that potentially indicate male quality, why were they overlooked in the models that best predicted social mate choice? One possibility is that they do predict social mate choice, but with an effect size too small to have been detected with the sample (n = 51 birds) that we used to estimate our models.

The pattern of associations with the "extra" song traits is somewhat ambiguous and difficult to interpret. In most cases, these song traits—first category songs with fewer, longer elements and broader bandwidths, and second category songs with longer fewer elements and faster chipping rates—were correlated with male traits that might reflect higher quality, if we accept the possibility that smaller rather than larger males are of higher quality. Smaller males might be superior if their body condition is better, as in common moorhens (Petrie 1983; Alvarez et al. 2005), or because they are more agile, as in chironomid midges (Crompton et al. 2003). Some songs traits, however, were involved in seemingly contradictory associations. In particular, males that sang longer first category songs tended to arrive earlier but fail to fledge young, and males that sang second category songs more rapidly and with broader bandwidth tended to be smaller but arrive later. It appears that for some of

the song traits that were not identified as good predictors of social mate choice, the relationship between song traits and male quality is more complex than it is for song traits that predict social mate choice. Nonetheless, listeners could in principle extract additional information about male quality from the "extra" song traits by focusing on the ones that are most consistently correlated with quality.

Can Song Structure Signal Male Quality?

Some of the song traits that we identified as possibly used by females to assess male quality seem well suited for that role. In particular, song traits that are not linked to the structure of individual songs seem especially suitable. A male's degree of song performance consistency, for example, does not depend on the particular song types he uses, so it is intuitively plausible that a high-quality male might be able to perform whatever song types he possesses more consistently than a low-quality male would perform his song types.

Other song traits that we identified as potential signals of male quality are a bit more puzzling, because they seem to depend on the form of individual songs. Because songbirds typically acquire their songs by imitation (Marler 1997; Hultsch & Todt 2004), the set of song types that a male acquires seems to depend in large measure on the types being sung within earshot at the time of acquisition (Nordby et al. 1999; Liu & Kroodsma 2006), a determinant with no obvious connection to the bird's quality, given that the imitated model is generally not the father (Kroodsma 1974; Wheelwright et al. 2008; Soha et al. 2009). So how might traits such as song frequency, song duration, and element rate come to be associated with male quality? It seems possible only if, during or after the song learning process, birds routinely alter the frequency, duration, or element rate of the songs they sing, relative to the values of these traits in the song types that were imitated. Unfortunately, for prairie warblers the degree of fidelity in song imitation is currently unknown, so a solution to this puzzle must await further study.

Acknowledgements

Literature Cited

Akresh, M. E., King, D. I. & Brooks, R. T. 2015: Demographic response of a shrubland bird to habitat creation, succession, and disturbance in a dynamic landscape. For. Ecol. Manag. **336**, 72—80.

Alvarez, F., Sanchez, C. & Angulo, S. 2005: Small, fat male moorhens *Gallinula chloropus* are healthier. Ardeola **52**, 279—285.

Bergeron, P., Baeta, R., Pelletier, F., Reale, D. & Garant, D. 2011: Individual quality: tautology or biological reality? J. Anim. Ecol. **80**, 361—364.

Bioacoustics Research Program. 2011: Raven Pro: Interactive Sound Analysis Software (Version 1.4). Cornell Lab of Ornithology, Ithaca, NY.

Bize, P., Devevey, G., Monaghan, P., Doligez, B. & Christe, P. 2008: Fecundity and survival in relation to resistance to oxidative stress in a free-living bird. Ecology **89**, 2584—2593.

Briggs, V. S. 2013: Do big dads make big babies? Paternal effects on larval performance in red-eyed treefrogs of Belize (*Agalychnis callidryas, A. moreletti*). Herpetol. J. **23**, 131—138.

Byers, B. E., Akresh, M. E. & King, D. I. 2015: A proxy of social mate choice in prairie warblers is correlated with consistent, rapid, low-pitched singing. Behav. Ecol. Sociobiol. **69**, 1275—1286.

Candolin, U. & Voigt, H. R. 2001: Correlation between male size and territory quality: consequence of male competition or predation susceptibility? Oikos **95**, 225—230.

Crompton, B., Thomason, J. C. & McLachlan, A. 2003: Mating in a viscous universe: the race is to the agile, not to the swift. Proc. R. Soc. B **270**, 1991—1995.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., Nielsen, A. & Sibert, J. 2012: AD Model Builder: using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optim. Methods Softw. **27**, 233—249.

Garvin, J. C., Abroe, B., Pedersen, M. C., Dunn, P. O. & Whittingham, L. A. 2006: Immune response of nestling warblers varies with extra-pair paternity and temperature. Mol. Ecol. **15**, 3833—3840.

Grafen, A. 1990: Sexual selection unhandicapped by the Fisher process. J. Theor. Biol. **144**, 473—516.

Hegyi, G., Rosivall, B. & Torok, J. 2006: Paternal age and offspring growth: separating the intrinsic quality of young from rearing effects. Behav. Ecol. Sociobiol. **60**, 672—682.

Houlihan, P. W. 2000: The singing behavior of prairie warblers (*Dendroica discolor*). Ph.D. Thesis, Univ. of Massachusetts, Amherst, MA.

Hultsch, H. & Todt, D. 2004: Learning to sing. In: Nature's Music: The Science of Birdsong. (Marler, P. & Slabbekoorn, H., eds). Academic Press, New York. pp. 80—107.

Jennions, M. D. & Møller, A. P. 2003: A survey of the statistical power of research in behavioral ecology and animal behavior. Behav. Ecol. **14**, 438—445.

Johnsen, A., Andersen, V., Sunding, C. & Lifjeld, J. T. 2000: Female bluethroats enhance offspring immunocompetence through extra-pair copulations. Nature **406**, 296—299.

Johnstone, R. A. 1996: Multiple displays in animal communication: 'Backup signals' and 'multiple messages'. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **351**, 329—338.

Kirkpatrick, M. 1982: Sexual selection and the evolution of female choice. Evolution **36**, 1—12.

Kokko, H. 1999: Competition for early arrival in migratory birds. J. Anim. Ecol. **68**, 940—950.

Kokko, H., Brooks, R., Jennions, M. D. & Morley, J. 2003: The evolution of mate choice and mating biases. Proc. R. Soc. Lond. B **270**, 653—664.

Kroodsma, D. E. 1974: Song learning, dialects, and dispersal in the Bewick's wren. Zeitschrift für Tierpsychologie **35**, 352—380.

Laake, J. L. 2013: RMark: An R Interface for Analysis of Capture-Recapture Data With MARK. National Marine Fisheries Service, Seattle, WA.

Lande, R. 1981: Models of speciation by sexual selection on polygenic traits. Proc. Natl. Acad. Sci. U. S. A. **78**, 3721—3725.

Lebreton, J. D., Burnham, K. P., Clobert, J. & Anderson, D. R. 1992: Modeling survival and testing biological hypotheses using marked animals: a unified approach with case-studies. Ecol. Monogr. **62**, 67—118.

Liu, W. C. & Kroodsma, D. E. 2006: Song learning by chipping sparrows: when, where, and from whom. Condor **108**, 509—517.

Marler, P. 1997: Three models of song learning: evidence from behavior. J. Neurobiol. **33**, 501—516.

Martin, T. E. & Geupel, G. R. 1993: Nest-monitoring plots: methods for locating nests and monitoring success. J. Field Ornithol. **64**, 507—519.

Maynard Smith, J. & Harper, D. 2003: Animal Signals. Oxford Univ. Press, Oxford.

Møller, A. P. & Jennions, M. D. 2002: How much variance can be explained by ecologists and evolutionary biologists? Oecologia **132**, 492—500.

Møller, A. P. & Pomiankowski, A. 1993: Why have birds got multiple sexual ornaments? Behav. Ecol. Sociobiol. **32**, 167—176.

Møller, A. P., Mousseau, T. A., Rudolfsen, G., Balbontin, J., Marzal, A., Hermosell, I. & De Lope, F. 2009: Senescent sperm performance in old male birds. J. Evol. Biol. **22**, 334—344.

Newton, I. 2008: Migration Ecology of Birds. Academic Press, London.

Nolan, V. 1978: The Ecology and Behavior of the Prairie Warbler, *Dendroica discolor*. American Ornithologists' Union, Tampa.

Nordby, J. C., Campbell, S. E. & Beecher, M. D. 1999: Ecological correlates of song learning in song sparrows. Behav. Ecol. **10**, 287—297.

Part, T. 2001: Experimental evidence of environmental effects on age-specific reproductive success: the importance of resource quality. Proc. R. Soc. B **268**, 2267—2271.

Petrie, M. 1983: Female moorhens compete for small fat males. Science **220**, 413—415.

Prum, R. O. 2010: The Lande-Kirkpatrick mechanism is the null model of evolution by intersexual selection: implications for meaning, honesty, and design in intersexual signals. Evolution **64**, 3085—3100.

Pyle, P. 1997: Identification Guide to North American Passerines. Slate Creek Press, Bolinas, CA.

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

Raveh, S., Sutalo, S., Thonhauser, K. E., Thoss, M., Hettyey, A., Winkelser, F. & Penn, D. J. 2014: Female partner preferences enhance offspring ability to survive an infection. BMC Evol. Biol. **14**, 14.

Saino, N., Romano, M., Ambrosini, R., Rubolini, D., Boncoraglio, G., Caprioli, M. & Romano, A. 2012: Longevity and lifetime reproductive success of barn swallow offspring are predicted by their hatching date and phenotypic quality. J. Anim. Ecol. **81**, 1004—1012.

Schielzeth, H. 2010: Simple means to improve the interpretability of regression coefficients. Methods Ecol. Evol. **1**, 103—113.

Sergio, F., Blas, J., Baos, R., Forero, M. G., Antonio Donazar, J. & Hiraldo, F. 2009: Short- and long-term consequences of individual and territory quality in a long-lived bird. Oecologia **160**, 507—514.

Siefferman, L. & Hill, G. E. 2003: Structural and melanin coloration indicate parental effort and reproductive success in male eastern bluebirds. Behav. Ecol. **14**, 855—861.

Siikamaki, P. 1995: Habitat quality and reproductive traits in the pied flycatcher - an experiment. Ecology **76**, 308—312.

Skaug, H., Fournier, D., Bolker, B., Magnusson, A. & Nielsen, A. 2015: Generalized linear mixed models using AD Model Builder. R package version 0.8.3.2.

Soha, J. A., Lohr, B. & Gill, D. E. 2009: Song development in the grasshopper sparrow, *Ammodramus savannarum*. Anim. Behav. **77**, 1479—1489.

Spector, D. A. 1992: Wood-warbler song systems: a review of paruline singing behaviors. Curr. Ornithol. **9**, 199—238.

StataCorp. 2013: Stata Statistical Software: Release 13. StataCorp LP, College Station, TX.

Stodola, K. W., Linder, E. T., Buehler, D. A., Franzreb, K. E., Kim, D. H. & Cooper, R. J. 2010: Relative influence of male and female care in determining nestling mass in a migratory songbird. J. Avian Biol. **41**, 515—522.

Wheelwright, N. T., Swett, M. B., Levin, I. I., Kroodsma, D. E., Freeman-Gallant, C. R. & Williams, H. 2008: The influence of different tutor types on song learning in a natural bird population. Anim. Behav. **75**, 1479—1493.

White, G. C. & Burnham, K. P. 1999: Program MARK: survival estimation from populations of marked animals. Bird Study **46**, 120—139.

Wilson, A. J. & Nussey, D. H. 2010: What is individual quality? An evolutionary perspective. Trends Ecol. Evol. **25**, 207—214.

Zahavi, A. 1975: Mate selection — selection for a handicap. J. Theor. Biol. **53**, 205—214.

Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A. & Smith, G. M. 2009: Mixed Effects Models and Extensions in Ecology with R. Springer-Verlag, New York.