

## SEXUAL SELECTION IN FEMALE PERCEPTUAL SPACE: HOW FEMALE TÚNGARA FROGS PERCEIVE AND RESPOND TO COMPLEX POPULATION VARIATION IN ACOUSTIC MATING SIGNALS

MICHAEL J. RYAN<sup>1,2,3</sup> AND A. STANLEY RAND<sup>2,4</sup>

<sup>1</sup>Section of Integrative Biology C0930, University of Texas, Austin, Texas 78712

<sup>2</sup>Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Panama

<sup>3</sup>E-mail: mryan@mail.utexas.edu

<sup>4</sup>E-mail: rands@ic.si.edu

**Abstract.**—Female preferences for male mating signals are often evaluated on single parameters in isolation or small suites of characters. Most signals, however, are composites of many individual parameters. In this study we quantified multivariate traits in the advertisement call of the túngara frog, *Physalaemus pustulosus*. We represented the calls in multidimensional scaling space and chose nine test calls to represent the range of population variation. We then tested females for phonotactic preference between calls in each pair of the nine test calls. We used statistics developed for paired comparisons in such “round robin” competitions to evaluate the null hypothesis of equal attractiveness, and to examine the degree to which females responded to calls as being different from or similar to one another in attractiveness. We then examined the attractiveness of each test call relative to all other test calls as a function of their location in multivariate acoustic space (the acoustic landscape) to visualize sexual selection on calls. Finally, we used methods from cognitive psychology to illustrate the females’ perception of call attractiveness in multivariate space, and compared this perceptual landscape to the acoustic landscape of quantitative call variation.

We show that correlations between individual call characters are not strong and thus there are few biomechanical constraints on their independent evolution. Most call variables differed among males, and there was high repeatability of call characters within males. Females often discriminated between pairs of calls from the population, and there were significant differences among calls in their attractiveness. Female preferences for calls were not stabilizing. The region of the acoustic landscape that was most attractive to females included the mean call but was not centered around it. The females’ perceptual or preference landscape did not correlate with the call’s acoustic landscape, and female perception of calls decreased rather than enhanced call differences.

**Key words.**—Animal communication, mate recognition, perception, *Physalaemus pustulosus*, sexual selection.

Received December 23, 2002. Accepted June 3, 2003.

Darwin’s conjecture (Darwin 1859, 1871) that female choice of mates can influence male reproductive success and generate sexual selection has now been demonstrated exhaustively (Andersson 1994). Female mate choice is often based on male courtship signals. Thus, to understand how sexual selection is generated it becomes important to understand how females perceive and react to signal variation.

Many studies have shown how female preferences vary as a function of an individual signal parameter (Gerhardt 1978; Ryan 1980; Basolo 1990; Gerhardt and Schul 1999; Gerhardt and Huber 2002 (Appendix 4)), small suites of characters (Murphy and Gerhardt 2000; Gerhardt and Huber 2002 (Appendix 4); Rosenthal 2002; Schul and Bush 2002), or entire signal complexes (Ryan and Wagner 1987; Zuk et al. 1990; Jang and Greenfield 1998). Some of these studies have measured preference functions for single traits or pairs of traits and show how the strength of preference, which is sometimes translated to a selection gradient, covaries with the stimulus (Ritchie 1996; Wagner 1998; Murphy and Gerhardt 2000; Schul and Bush 2002).

The approaches used above have been invaluable in elucidating how female mating preferences (and sometimes male mating preferences) can generate selection on individual or small suites of signal traits. There are several limitations to these approaches, however. One limitation is that the female preferences for individual traits might not predict how females respond to comparisons of entire signals. The interaction effect of individual traits on female preference might

not be additive, and it could even reverse signs. As a specific example, studies of jungle fowl show that females attend to a suite of characters but the relative importance of each character can vary with time (Zuk et al. 1992). Another limitation is that examining small numbers of traits in isolation does not take into account how those traits are correlated among individuals. Interpreting the selection consequences of a study showing that females prefer signals of higher frequency and longer duration, for example, depends on how those two characters, duration and frequency, co-occur in nature. These limitations have been pointed out by other authors as well (e.g., Zuk et al. 1992; Jang and Greenfield 1998; Rosenthal 2000).

The purpose of our study is to document how females perceive and respond to population variation in multivariate signals that are used in mate attraction. We analyze mating signals of a population and represent these signals in multivariate space, which, since they are mating calls, we refer to as the *acoustic landscape* or *acoustic space*. We then choose calls, the *test calls*, that are representative of the population variation. We determine female preference for all possible pairs of test calls to determine the *relative attractiveness* of each call, thus giving a view of how attractiveness varies within the known distribution of population variation. We then use the same phonotaxis data to calculate how females perceive relative call attractiveness, and we represent these calls in *perceptual space* or in a *perceptual landscape*. Finally, we compare differences in how the same signals are distributed in *acoustic space* and *perceptual space*.

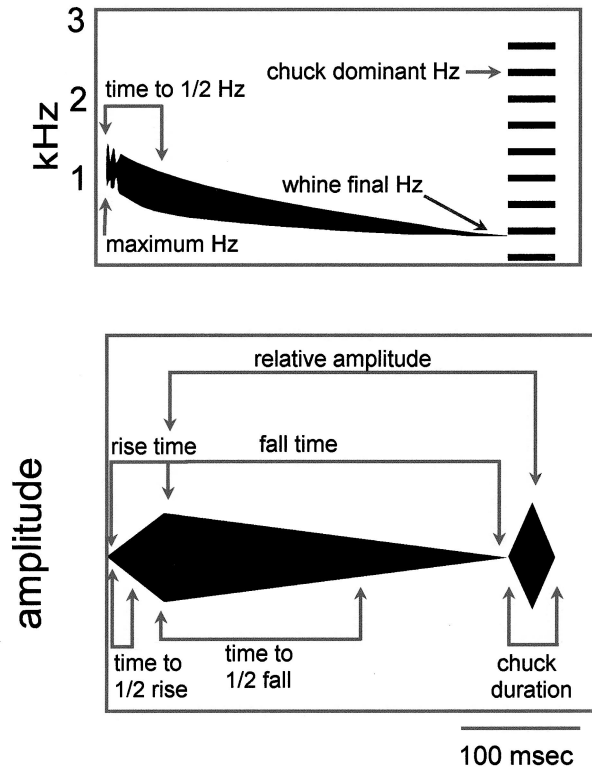


FIG. 1. Schematic representation of a typical mating call of the túngara frog *Physalameus pustulosus* and the call variables measured for this study. The top illustration is a sonogram of the whines fundamental frequency and one chuck from which we measure spectral characteristics of the call. The bottom illustration is an oscillogram from which we measure amplitude characteristics of the call.

## MATERIALS AND METHODS

### Signal Recording and Analysis

The túngara frog's advertisement call consists of a frequency sweep of about 300 msec, the whine, that can be produced by itself or followed by a shorter, about 40 msec, burst of sound, the chuck (Ryan 1985). We recorded calls of male túngara frogs, *Physalameus pustulosus*, from a population near one of the laboratories of the Smithsonian Tropical Research Institute in Gamboa, Panama (09°07'0"N, 79°41'53"W). Calls were recorded in July 1996 with a Marantz PMD 420 recorder and a Sennheiser ME 80 microphone with K3U power module on magnetic cassette tape. Air temperatures at the calling males were always within a few degrees of 25°C thus call variables were not corrected for temperature. We recorded at least five complex calls (Figs. 1, 2), that is, a call with one whine and at least one chuck, in each recording session. We recorded calls from 50 males. Nine of those males were recorded on a second night during the month, and one male was recorded on a third night during the month. The total number of calls recorded and analyzed was 300. We used all of these calls in the analysis because we are interested in the variation in signals that a female encounters during a reproductive bout and how females react to and perceive such call variation. That is, we are interested in the relative attractiveness and perception of signals rather than that of males.

Calls were digitized at a rate of 20 kHz and analyzed with Signal. All calls analyzed consisted of a whine plus one chuck. Some calls had more than one chuck when recorded but these call were truncated after the first chuck for analysis. We measured a suite of call characters that have proven useful for quantifying signal variation in these frogs over the last two decades (e.g., Ryan 1980; Ryan and Rand 2001). Measures of spectral aspects of the whine are restricted to the fundamental frequency-sweep. In previous experiments we

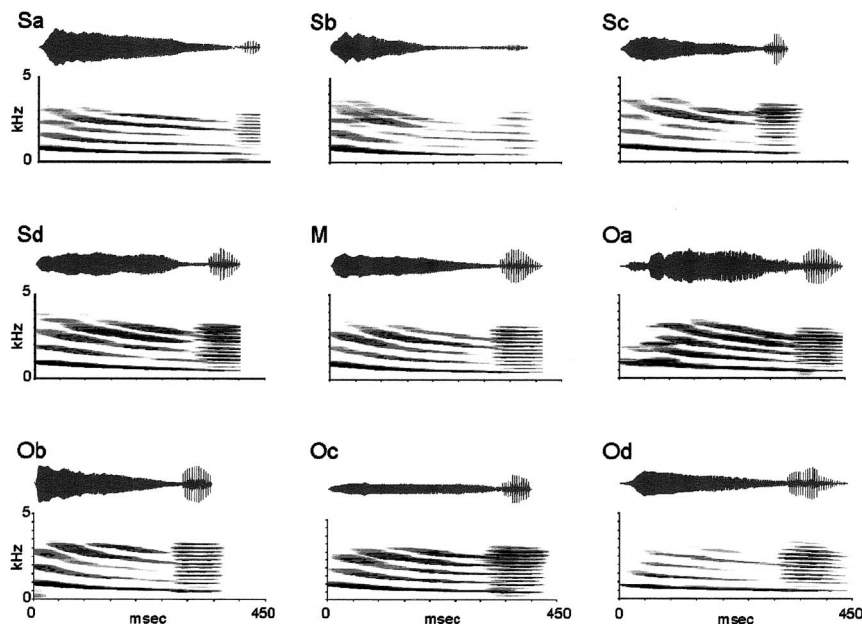


FIG. 2. Oscillograms and sonograms of the nine test calls.

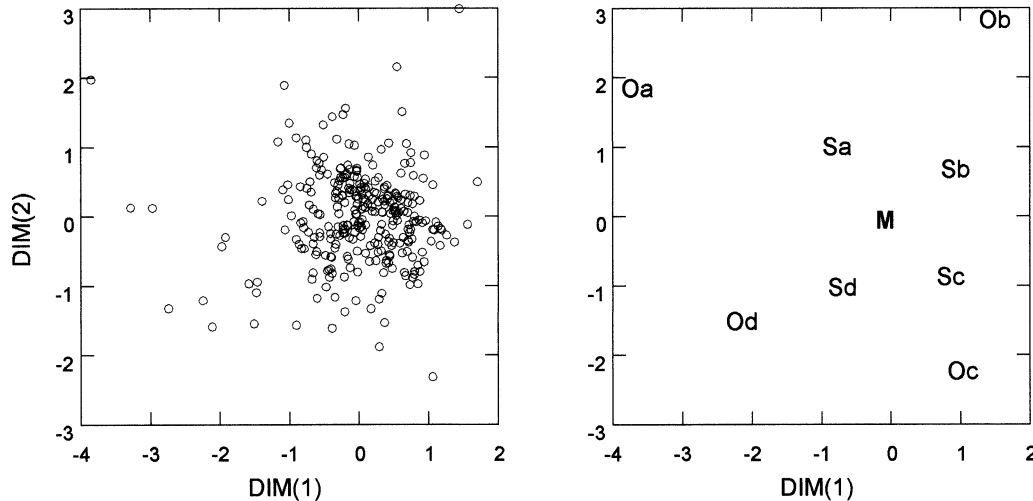


FIG. 3. (Left) A plot of the first two dimensions of a multidimensional scaling analysis of the 300 túngara frog calls sampled from the populations. (Right) The same plot but only the nine test calls are represented.

used synthetic calls that varied the presence of whine harmonics and found that the upper harmonics of the whine do not influence female phonotaxis (Rand et al. 1992; Wilczynski et al. 1995). This does not prove that upper whine harmonics could not have subtle interactions with other variables to influence preferences, but we have not tested this possibility.

Calls were analyzed in batches with programs written in Signal. Batch processing enforces a degree of standardization that is sometimes lost when calls are analyzed individually. Call variables were extracted from spectrograms, oscillograms and fast fourier transforms.

We measured fifteen call variables (Fig. 1). Previous studies have shown that calls synthesized using these variables are as attractive as natural calls (M. J. Ryan and A. S. Rand, unpubl. data). These variables, which seem to capture the full salience of the signal, are as follows: *Call duration*: the duration of the entire call. *Call dominant frequency*: the dominant frequency of the entire call. *Initial frequency*: the initial frequency of the whine's fundamental frequency. *Maximum frequency*: the maximum frequency of the whine's fundamental. Often this is the same as the initial frequency, but in some cases the whine exhibits a rapid and slight increase in frequency. *Time to maximum frequency*: the time interval from the beginning of the whine to the whine's maximum frequency. If the initial and maximum frequency are the same, the time is 0 msec. *Final frequency*: the whine's frequency at the end of the call, which is always the lowest frequency. *Whine duration*: the duration of the whine. The whine gradually ends before initiation of the chuck, and the demarcation between whine and chuck is somewhat arbitrary. We defined the end of the whine as the minimum amplitude between the maximum amplitude of the whine and the maximum amplitude of the chuck. *Rise time*: the time from the whine's onset to its maximum amplitude. *Fall time*: the time from the whine's maximum amplitude to the end of the call. *Half frequency*: the whine's duration from the onset to its mid-frequency; this is a measure of the shape of the whine's frequency sweep. *Half rise*: the duration from the onset of

the call to one-half the maximum amplitude during the rise. *Half fall*: the whine's duration from the whine's maximum amplitude to one-half the maximum amplitude during the fall. *Chuck duration*: the duration of the chuck. As noted above, the end of the whine and the beginning of the chuck can be subjective. The amplitude envelope of the chuck is approximately symmetrical around the chuck's maximum amplitude. We measured the time from the maximum amplitude of the chuck to the end of the chuck, which was always easy to discern. We then determined the onset of the chuck as this time interval before the chuck's maximum amplitude. These points defined the beginning and end of the chuck and were used to derive the chuck's duration. Thus the end of the whine (see *Whine duration*, above) and the beginning of the chuck are not necessarily the same point. *Chuck dominant frequency*: the frequency of the chuck with the most energy. *Relative chuck-whine amplitude*: the maximum amplitude of the chuck divided by the maximum amplitude of the whine.

As described below, we chose nine natural calls that represented the population variation (Figs. 2, 3). These calls were manipulated to remove as much background noise as possible. This was accomplished using a procedure in Signal that defines the spectral and amplitude characteristics of each harmonic of the call and then resynthesizes the call. Thus, any background noise, unless it overlaps the call, is eliminated from the signal.

#### Phonotaxis Experiments

We tested the female preference for each pair of the nine test calls. There were 36 call comparisons and 20 females were tested in each experiment, for a total of 720 female choices. Tests were conducted during three years, the number of females tested and the total number of tests in each year were as follows: 1997: 31 females, 46 tests; 1998: 113 females, 400 tests; 2001: 85 females, 379 tests. The total number of tests conducted, 825, exceeds the number of tests analyzed because some experiments had sample sizes greater than 20. In such cases we analyzed the results from only the

first 20 responses in order to have equal sample sizes among experiments to facilitate statistical analyses. Most females were tested in a number of experiments, but the window of reproductive receptivity relative to the time involved in experimentation did not make it possible to test all females in all experiments. In other studies we have shown that there does not appear to be a polymorphism in female response characteristics (Kime et al. 1998).

We tested female túngara frogs from Gamboa, Panama and most females were from the population of males whose calls we analyzed. Typically, females were collected at choruses between 1900 and 2200 h and tested between 2300 and 0700 h. Females are usually only found at a chorus when ready to mate and most females were in amplexus, that is, clasped by a male, when collected. After testing, females were released within 12 hours of capture allowing them the opportunity to reproduce in the wild. They were toe-clipped so as not to be tested again in the same experiment if recaptured.

We conducted the phonotaxis experiments in an Acoustic Systems (Austin, TX) sound attenuation chamber that measured 1.8 m × 2.7 m. The female's behavior was observed on a video monitor equipped with an infrared light source connected to a wide-lens video camera. The camera was located on the chamber's ceiling. We placed a female under a small cone in the center of the chamber which could be raised remotely to initiate testing. We broadcast the test stimuli antiphonally from speakers in the center of walls opposite one another such that the peak amplitude of the whine of each test call at the center of the arena was 82 dB SPL (re. 20 µPa). Calls were broadcast at a rate of one call per two seconds from each speaker. A positive phonotactic response was noted if a female approached within 10 cm of one of the speakers as long as this response did not result from the female following the chamber's walls. A female did not exhibit phonotaxis to the test stimulus if she approached the speaker broadcasting noise, if she remained motionless for the first five minutes or any subsequent two minutes of the trial, or did not exhibit phonotaxis after 15 minutes. Each experiment was continued until 20 females were tested.

*Statistical Analysis*

Standard statistical calculations were made for all of the 15 call variables and are reported in Table 1. We also transformed the call variables to z scores for various analyses. We computed the Pearson product moment correlation between all pairs of call variables (Table 2). We compared call variables among males using an analysis of variance (ANOVA). We also computed the repeatability for each call measure by calculating the coefficient of the intraclass correlation for each call variable. This is a measure of the proportion of variation among males compared to that within males. A value of 1.0 signifies high repeatability of measures within males (relative to among males), 0 means that the repeatability within males is no greater than among males (Sokal and Rohlf 1995).

We calculated the Euclidean distance for each pair of calls using the transformed measures of call variables (z scores). This acoustic similarity/dissimilarity matrix was used to generate a multidimensional scaling (MDS) plot of call variation

TABLE 1. Results of the statistical analyses (n, sample size; SE, standard error; SD, standard deviation; CV, coefficient of variation; ICC, intraclass correlation coefficient) for the call variables measured (metric for each is indicated in parentheses): ALLCLDUR, duration of the entire call (ms); ALLCLDZH, dominant frequency of the entire call (Hz); INITHZ, initial frequency of fundamental frequency of the whine (Hz); ENDDHZ, final frequency of the whine (Hz); MAXHZ, maximum frequency of fundamental frequency of the whine (Hz); TIMMXHZ, time to maximum frequency of the whine (ms); WHDUR, duration of the whine (ms); RISE, rise time of the whine (ms); HFRISE, time to half amplitude of the whine (ms); FALL, fall time of the whine (ms); HFFALL, time to half amplitude of the fall from the call's end (ms); HFFHZ, time to mid-frequency of the fundamental frequency of the whine (Hz); CKDUR, duration of the chuck (ms), dominant frequency of the chuck (Hz); and RELAMP, peak amplitude of chuck divided by peak amplitude of the whine.

	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300
	ALLCLDUR	ALLCLDZH	INITHZ	ENDDHZ	MAXHZ	TIMMXHZ	WHDUR	RISE	HFRISE	FALL	HFFALL	HFFHZ	CKDUR	CKDOMHZ	RELAMP		
n	300	300	300	300	300	300	300	300	300	300	300	300	300	300	300		
Mean	409.20	581.85	1000.66	466.52	1040.17	6.19	325.88	42.49	15.32	283.38	102.95	59.83	43.83	2591.87	1.11		
SE	2.38	2.17	5.93	2.15	5.32	0.46	1.98	1.71	0.75	2.18	2.92	1.18	0.63	23.10	0.28		
SD	41.30	37.73	102.83	37.31	92.18	8.09	34.44	29.62	13.07	37.81	50.72	20.49	10.97	400.18	0.49		
CV	0.10	0.06	0.10	0.07	0.08	0.13	0.10	0.69	0.85	0.13	0.49	0.34	0.25	0.15	0.44		
ICC	0.78	0.68	0.33	0.62	0.78	0.62	0.73	0.76	0.57	0.66	0.76	0.64	0.73	0.41	0.63		



TABLE 2. Call character correlation matrix. Definitions of abbreviations are as in Table 1.

	ALLCLDUR	ALLCLDHZ	INITHZ	ENDHZ	MAXHZ	TIMMXHZ	WHDUR	RISE	HFRISE	FALL	HFFALL	HFHZ	CKDUR	CKDOMHZ	RELAMP
ALLCLDUR	1.000														
ALLCLDHZ	-0.413	1.000													
INITHZ	-0.081	0.088	1.000												
ENDHZ	-0.214	0.252	0.145	1.000											
MAXHZ	-0.015	0.097	0.597	0.184	1.000										
TIMMXHZ	0.242	-0.033	0.128	0.085	0.602	1.000									
WHDUR	0.829	-0.389	-0.066	-0.305	-0.040	0.123	1.000								
RISE	0.361	0.014	0.091	-0.066	0.195	0.242	0.311	1.000							
HFRISE	0.403	0.047	-0.004	-0.051	0.173	0.309	0.346	0.635	1.000						
FALL	0.472	-0.366	-0.131	-0.226	-0.190	-0.077	0.667	-0.500	-0.183	1.000					
HFFALL	0.109	0.072	-0.036	-0.045	-0.010	0.144	0.050	0.003	0.202	0.043	1.000				
HFHZ	0.221	0.167	-0.380	-0.393	-0.510	-0.286	0.282	0.327	0.311	0.001	0.088	1.000			
CKDUR	0.379	-0.051	-0.097	-0.032	-0.046	0.110	0.249	0.185	0.350	0.081	0.197	0.195	1.000		
CKDOMHZ	-0.082	0.198	0.085	0.084	0.143	0.065	-0.192	0.176	0.053	-0.313	0.093	0.020	-0.042	1.00	
RELAMP	0.134	-0.079	-0.011	0.094	0.134	0.251	0.065	0.098	0.131	-0.017	0.217	0.020	0.201	0.204	1.00

in two dimensions (Wilkinson 1996). Multidimensional scaling is a technique that represents any proximity matrix in multidimensional space such that the spatial distance between any pair of points reflects their similarity/dissimilarity in the proximity matrix (Kruskal and Wish 1978). We will refer to this MDS plot of the acoustic similarity of calls as the *acoustic landscape* and areas within this landscape as *acoustic space*.

The MDS axes of call variation are in units of standard deviations. We used the plot of the acoustic landscape to choose nine test calls that were representative of the population variation (Figs. 2, 3). We chose one call near the population mean ( $M$ ), four calls that were  $\pm 1$  SD from the mean ( $S_{a-d}$ ), and four calls from near the extremes of the population variation ( $O_{a-d}$ ; Figs. 2, 3; Table 3).

Female preferences between pairs of calls were used to calculate the average attractiveness of a call; the more females responded to one of the calls in a pair the more attractive that call was in that particular comparison. The total number of responses to each of the nine calls for all the experiments combined is a measure of the attractiveness of all calls relative to one another. We compared the relative attractiveness of calls to one another with a nonparametric paired comparison test (David 1988). Post hoc tests were conducted with a least significant difference test (David 1988). These statistics are specifically designed to analyze results of experiments in which judges (in this case females) are presented pairs of objects (in this case calls) and indicate a choice between the two objects. These statistics are commonly used in ‘round-robin’ experiments in which all objects are competed against one another in pairs. To test null hypothesis of equal preference among objects, we compared  $D_n$  to  $X^2_{df,0.05}$ , in which

$$D_n = 4 \left[ \sum a_i^2 - \frac{1}{4} t n^2 (t - 1)^2 \right] / (n t),$$

where  $a$  is the score of each call versus all other calls (1 = chosen, 0 = not chosen),  $t$  is the number of calls compared, and  $n$  is the number of females tested in each comparison. The least significant difference ( $P = 0.05$ ) between scores,  $m_c$ , was calculated as

$$m_c = 1.96 \left( \frac{1}{2} n t \right)^{1/2} + \frac{1}{2}.$$

We also used the female preference data to construct a similarity/dissimilarity matrix. This matrix is similar to the acoustic similarity matrix, but instead of being based on physical properties of the calls it is based on the female’s perception of call differences. For example, if 20 females respond to one call in an experiment and 0 females to the alternative call then the females perceive these calls as quite different from one another. If, on the other hand, 10 females exhibit phonotaxis to one call and 10 to the alternative, it suggests the females perceive these two calls as equivalent. There is an important matter of definition here. If females show a preference for one call in a pair then they are capable of discriminating the acoustic properties of the two signals and these differences are meaningful to the females. In the parlance of psychophysics, there are both just noticeable dif-

TABLE 3. Call variables for test calls. Listed in order of attractiveness (see Fig. 4) from most attractive (top) to least attractive (bottom). Definitions of abbreviations are as in Table 1.

	ALLCLDUR	ALLCLDZH	INITHZ	ENDHZ	MAXHZ	TIMMXHZ	WHDUR	RISE	HFRISE	FALL	HFFALL	HFHZ	CKDUR	CKDOMHZ	RELAMP
411.7	631.10	1013.69	471.11	1013.69	0.00	339.20	118.05	30.35	221.15	141.80	97.21	52.10	2822.27	1.16	
485.5	617.68	1088.53	474.19	1176.06	29.66	370.05	91.90	74.40	278.15	167.15	66.00	84.05	2519.53	1.44	
409.6	581.05	1017.81	464.55	1017.81	0.00	329.60	55.70	8.05	273.90	131.20	74.64	52.05	2871.09	1.50	
516.5	531.01	947.61	434.70	949.21	2.35	391.55	33.65	17.85	357.90	113.55	55.60	34.45	1933.59	0.43	
396.6	625.00	1016.34	704.28	1016.34	0.00	341.00	55.50	11.35	285.50	135.80	43.74	48.20	2773.44	2.46	
344.8	598.14	961.99	474.33	945.69	5.23	290.15	13.35	6.05	276.80	136.30	81.99	38.05	2949.22	1.14	
447.9	620.12	965.42	463.20	1090.16	9.57	367.35	79.30	60.15	288.05	88.15	112.63	55.45	2294.92	0.95	
371.1	639.65	1092.48	539.40	1092.48	0.00	279.45	85.55	14.40	193.90	69.10	47.04	41.45	3085.94	1.04	
382.8	549.32	943.92	461.50	943.92	0.00	281.75	31.50	9.95	250.25	78.40	54.77	33.75	1250.00	0.25	

ferences (JND) and just meaningful differences (JMD; Nelson and Marler 1990). If females do not discriminate between calls in a pair, we conclude there are no JMDs. In such a case, however, we can make no conclusion about JNDs; females may or may not hear the calls as acoustically different. Thus the perceptual equivalence of calls refers to the call's meaning not its sounds.

We calculated the similarity between calls in a pair as  $Similarity = ABS(10 - x) \times 0.05$ , in which  $ABS$  is the absolute value and  $x$  is the number of responses to a stimulus given a total sample size of 20. Thus, if 10 females respond to each call, the calls are most similar (similarity = 0), if 20 females respond to one call and 0 to the other, the two calls are most different (similarity = 1). This similarity matrix is used in a multidimensional scaling (MDS) analysis, as described above, to represent graphically the perceived similarity among calls, and we refer to this graph as the *perceptual space*. This is similar to using the Euclidean distances of acoustic variables to construct an MDS map of acoustic similarity. But in this analysis the females, rather than a quantitative analysis of acoustic features, classifies the degree to which calls are a similar or not similar.

To explore what individual call variables might explain variation in call attractiveness, we conducted a principal components analysis to reduce the call variation to five axes. We then used these factors as the independent variables in a stepwise multiple regression analysis with female responses as the dependent variable.

We used a Mantel test to determine if the pairwise distances between calls in acoustic space and perceptual space are correlated (Mantel nonparametric calculator, Vers. 2.0, Liedloff 1999). The probability of rejecting the null hypothesis was based on 1000 randomization simulations.

All analyses were conducted in Systat 10 (Wilkinson 1996) unless otherwise noted.

## RESULTS

### Call Variation

Three hundred calls of 50 males were analyzed (Table 1). We examined the correlations between call characters using the transformed data ( $z$  scores), and those results are presented in Table 2. About one-third (38 of 91) of the pairwise correlations were statistically significant (Bonferonni adjusted  $P = 0.0005$ , critical  $r = 0.189$ ). There are few correlations, however, that are sufficiently strong to suggest that there are major biomechanical constraints that influence the evolution of individual call characters. In only seven of 91 correlations does the variation in one variable of the pair account for more than 25% of the variation in the other variable (i.e.,  $r > 0.50$ , coefficient of determination  $>25\%$ ). Most of these correlations appear to be a trivial result of how we define and measure calls rather than indicating an unexpected constraint. For example, whine duration and the duration of the entire call are strongly correlated ( $r = 0.829$ ), but this is expected since the whine makes up almost the entire call. Whine duration and fall time are correlated ( $r = 0.667$ ), as expected since the fall time is most of the whine. Similarly, we would expect a strong correlation between rise time and the time to half rise time ( $r = 0.635$ ), and the call's

TABLE 4. Female preferences between calls in a pair. The abbreviations in the row and column axes refer to calls tested (see Fig. 2 and text). Total number of choices in all contrasts is 20. The cells show the number of responses to the call in the column. The "score" is the number of phonotactic responses to the call in the column versus all other calls tested.

	Oc	Sd	Ob	Sc	Sb	Sa	M	Oa	Od
Oc		10	10	6	8	5	18	10	12
Sd	10		8	4	1	6	12	4	15
Ob	10	12		8	7	14	18	12	12
Sc	14	16	12		10	13	17	13	15
Sb	12	19	13	10		16	16	8	15
Sa	15	14	6	7	4		11	10	9
M	2	8	2	3	4	9		9	13
Oa	10	16	8	7	12	10	11		16
Od	8	5	8	5	5	11	7	4	
Score	81	100	67	50	51	84	110	70	107

initial frequency and its maximum frequency ( $r = 0.597$ ), and this is what the results show.

Two correlations suggest some interesting biological constraints. The strong negative correlation between rise time and fall time of the whine ( $r = -0.500$ ) indicates a stronger constraint on the call's duration than on the precise shape of the amplitude envelope. The negative correlation ( $r = -0.510$ ) between the call's maximum frequency and the shape of the frequency sweep suggests that the frequency range and time constant of the whine's sweep might be functionally related to the frequencies with which the call begins.

There were statistically significant differences in all of the call variables among males (ANOVA all  $P < 0.001$ ). The intraclass correlation coefficients for each call variable revealed that call variables were moderately repeatable, more variation among than within males—all but two variables had intraclass correlation coefficients above 0.50. One call variable that had low repeatability was the dominant frequency of the chuck ( $r = 0.41$ ). The chuck has more than a dozen harmonics with substantial energy. A small change in the difference in total energy among harmonics could yield a deceptively large difference in the chuck's dominant frequency. For example, if a call had most energy in harmonic 14 (e.g., 3500 Hz) and only slightly less in harmonic 10 (2500 Hz), then a relatively small change in the energy between only these two harmonics would yield calls that differed by 1000 Hz in their dominant frequency. The other call variable with an intraclass correlation coefficient less than 0.50 was initial frequency. This is a variable that is probably under active control of the male. We assume that the tension of the vocal cords is increased from resting condition at the onset of the whine, and the precise tension, which appears not to be very repeatable, should influence the call's initial frequency (Drewry et al. 1982; Dudley and Rand 1991).

#### Female Call Preferences

The matrix of responses for the phonotaxis experiments comparing all pairs of the nine test calls is shown in Table 4. Thirty-nine percent (14 of 36) of the phonotaxis experiments showed a statistically significant preference for one of the calls (if number of responses  $\geq 15$ ,  $P < 0.05$ , exact binomial probability).

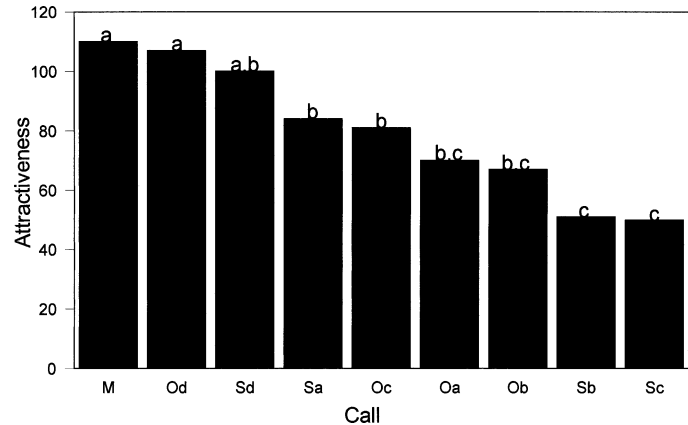


FIG. 4. The attractiveness of each test call relative to the other test calls estimated as the total responses to each call. Letters represent calls that are not significantly different from one another in attractiveness.

The attractiveness of calls was estimated as the number of phonotactic responses to that call versus all other calls against which it was tested. There are significant differences in attractiveness among calls ( $D_n = 90.03$ ,  $P < 0.005$ ). The least significant difference ( $P = 0.05$ ),  $m_c$ , in call attractiveness was 19.09. Figure 4 illustrates the calls that were considered significantly different in attractiveness from one another by this analysis.

We coded the attractiveness of the test calls and plotted them on the acoustic landscape of call variation (Fig. 5). Figure 5 shows that there is not a pattern of strong stabilizing selection centered around the mean call. There is also no evidence for enhanced attraction of calls that depart substantially from the mean. Instead, we see a contiguous region of the acoustic landscape, which includes the mean, in which call attractiveness is highest. But call attractiveness does not vary from highest to lowest predictably across the acoustic landscape. The two least attractive calls ( $Sb$ ,  $Sc$ ) are closer to the area of highest attractiveness in the acoustic landscape than are the moderately attractive calls that are distant outliers ( $Oa$ ,  $Ob$ ).

We performed a principal components analysis to reduce variation in the fifteen call variables to five principal components. The components explained 91% of the total variation. The first five components 25, 20, 19, 14 and 13% of the variation, respectively. The loadings of the call variables on each of the principal components are shown in Table 5.

We performed a stepwise multiple regression analysis of call attractiveness on the first five principal component factors. Three components explained a significant amount of the variation in call attractiveness: principal component five explained 44% of the variation by itself, when principal component two was added 68% of the variation was explained, and the further addition of principal component three explained 82% of the variation in call attractiveness. The call variables that loaded most heavily on principal components five and two were related to various measures of call duration. Call duration loaded most heavily on principal component five, and fall time of the whine loaded most heavily on principal component two. Principal component three was distin-

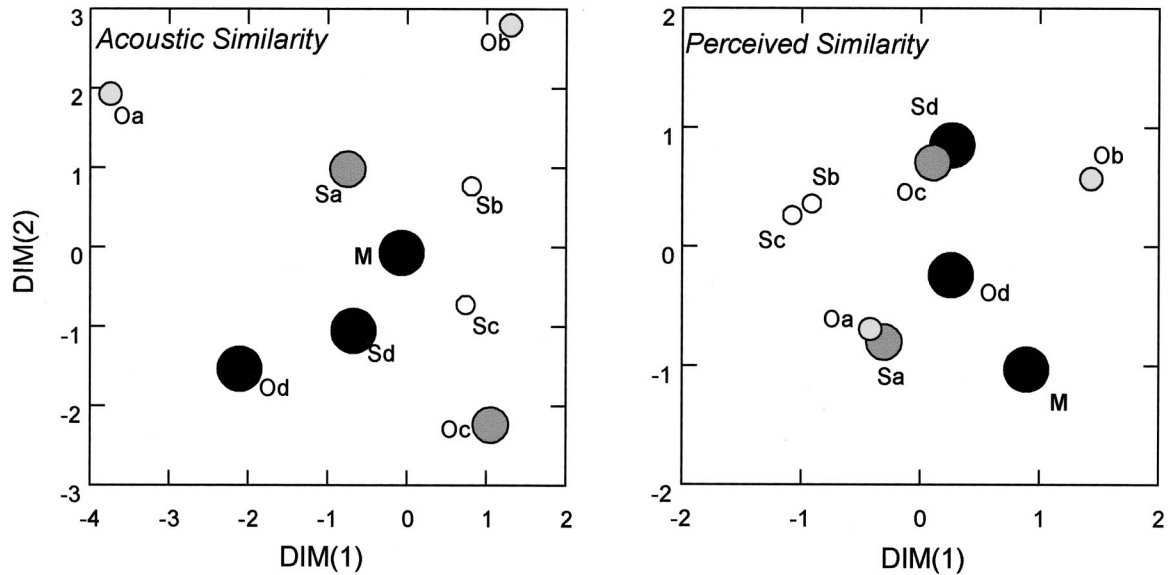


FIG. 5. (Left) The nine test calls are represented in MDS acoustic space (as in Fig. 3) and the calls' relative attractiveness is coded with larger symbols and darker shading indicating most attractive and smaller symbols and lighter shading indicating least attractive. (Right) The same test calls with the same coding of attractiveness as above are plotted in MDS perceptual space based on the female's assessment of how similar or different the calls are from one another.

guished by the relative amplitude of the chuck to the whine, the dominant frequency of the chuck, and the ending frequency of the whine.

We used the phonotaxis data in Table 4 to construct a matrix of perceived call similarity. These data were used in a MDS analysis in which the similarity of calls to females, rather than their physical acoustic similarity, is plotted in perceptual space. The first two dimensions of MDS distances between calls is shown in Figure 5, with the calls coded for mean attractiveness as is also done for the calls plotted in acoustic space. The distributions of the test calls in physical acoustic space and in female perceptual space were not significantly correlated with one another (Mantel test, 1000 iterations,  $g = 0.031$ ,  $P = 0.44$ , Fig. 5). Thus, the females' perception of signal variation is quite different from what one would surmise by examining the quantitative variation of the signals. These distributions differ significantly in two

details, the mean distances of the test calls to the center of the distribution (acoustic space:  $\bar{x} = 3.13$ ; perceptual space:  $\bar{x} = 0.940$ ,  $U$  test,  $U = 63$ ,  $P = 0.047$ ,  $n = 9,9$ ), and the variance (and coefficient of variation,  $CV$ ) of these distances (acoustic space:  $s^2 = 8.82$ ,  $CV = 0.95$ ; perceptual space,  $s^2 = 0.129$ ,  $CV = 0.38$ ; Levene's test for inequality of variances  $F_{1,16} = 16.50$ ,  $P = 0.0009$ ). Both the mean and variance are greater in physical than in perceptual space.

## DISCUSSION

There are four main results of this study. The first is that variation in natural calls, the complex suite of characters that constitute the advertisement call of túngara frogs, influences female call preferences. Extrapolating to mate choice in the wild, our results suggest that calls are under sexual selection by female choice. The second main result of this study is that there is weak directional selection on calls. The mean call is one of the most attractive calls but other calls that deviate somewhat from the mean call are similarly as attractive. The third result is that call characters are not highly correlated with one another and thus their independent evolution does not appear to be seriously constrained by biomechanics. The final result is that female perception of call attractiveness is not correlated with the unbiased quantitative measures of acoustic similarity, and female perception attenuates rather than enhances acoustic variation.

### *Multivariate Traits and Female Preference*

Many studies have shown that female mating preferences are influenced by variation in male advertisement displays (Andersson 1994). Some of these experimental studies have utilized the entire signal phenotype. Some examples are studies of mate choice in *Drosophila* (Kaneshiro 1988), call preferences in insects (Jang and Greenfield 1998; Table 7.2 in

TABLE 5. Component loadings for call variables of test calls. See Table 1 for definitions.

Call variables	Factor(1)	Factor(2)	Factor(3)	Factor(4)	Factor(5)
ALLCLDUR	0.345	0.084	-0.285	-0.127	0.864
ALLCLDHZ	0.338	0.560	0.596	0.165	-0.375
INITHZ	-0.013	0.309	0.043	-0.893	0.110
ENDHZ	-0.113	-0.031	0.744	-0.405	-0.175
MAXHZ	0.820	0.449	0.182	-0.073	-0.024
TIMMXHZ	0.961	-0.112	-0.001	0.070	0.202
WHDUR	0.267	0.104	0.034	0.082	0.936
RISE	0.371	0.908	0.135	-0.027	0.059
HFRISE	0.851	0.287	-0.106	0.223	0.302
FALL	-0.033	-0.570	-0.068	0.093	0.789
HFFALL	0.351	-0.117	0.568	0.052	0.449
HFHZ	0.123	0.319	-0.082	0.901	0.139
CKDUR	0.867	0.262	0.281	0.017	0.230
CKDOMHZ	0.019	0.338	0.764	0.183	-0.191
RELAMP	0.133	0.053	0.958	-0.130	0.040



Gwynne 2001), anurans (Marques and Bosch 2001), and birds (Catchpole et al. 1986), and olfactory preferences in moths (Conner et al. 1981) and fish (McLennan and Ryan 1999). The advantage of such studies is that they demonstrate a critical point—natural signal variation can influence female mating preferences. This point has been so well documented that one sometimes forgets the great skepticism that greeted this notion and ensued for an entire century after it was first introduced by Darwin (Cronin 1991).

The disadvantage of studies that utilize the entire signal is that the precise signal parameters that influence female preferences can not be uncovered. As we know from Tinbergen's concept of sign stimuli (Tinbergen 1969), many features of signals are not salient to their intended receivers. Furthermore, when female preferences for entire signals are tested the statistical distribution of these signals are often not specified. Thus it is difficult to ascertain how selection might influence the population distribution of traits.

An alternative approach to investigating preferences for sexual signals has been the parameterization of signals and the attempt to identify their salient properties. The goal of such studies are usually motivated by one of three concerns: to understand how the receiver decodes signals (Gerhardt and Schul 1999); to construct a preference function or selection gradient to predict how selection might influence the evolution of a individual signal parameters (Ritchie 1996; Wagner 1998); to search for correlations between aspects of the signal and male quality to hypothesize why preferences for signal parameters have evolved (Welch et al. 1998). All of these approaches have been successful and contributed substantially to our understanding of sexual communication (reviewed in Andersson 1994; Hauser 1996; Bradbury and Vehrencamp 1998; Gerhardt and Huber 2002; Greenfield 2002). Many of these studies, like the present study, also show that the female's perceptual space does not coincide with the population distribution of signal traits (Ryan 1980; Schul 1998; Schul and Bush 2002).

However, there can also be problems with testing only one or a few signal variables. We usually have little idea of how various signal parameters interact, and how robust these interactions are (e.g., Zuk et al. 1992). The interactions could be additive—if females prefer long tails and bright tails, then a long bright tail would be even more attractive. The interactions could be hierarchical—if a female detects one trait that is attractive, a long tail or a bright tail, it might ignore the other trait. Or the interaction effects could be unanticipated. On the other hand, in some systems it seems or it has been shown that females are attendant to only a pair of critical variables. In such cases manipulation of these variables is in essence manipulating the entire signal, or at least those parts that are salient to the female. We feel the interaction of individual signal parameters is an empirical question; it simply needs to be tested for each system.

There has been considerable theoretical discussion as to how females should evolve preferences for multivariate signals as well as multiple traits. Iwasa and Pomiankowski (1994) assumed that when females select for male traits indicative of "good genes," preferences should cycle through their dependence on traits in making mating decision, whereas multiple female preferences are more likely to occur si-

multaneously under a Fisherian process. Holland and Rice (1988), as part of their hypothesis of chase-away sexual selection, suggested that females evolve preferences with higher thresholds for traits until eventually those traits play no role in attraction. The result is a "graveyard of ineffectual traits," as opposed to Rosenthal's (2000) characterization of multiple traits as a "buffet of marginally effective ones." Although these models and discussions raise some interesting issues of how multiple preferences for multiple traits could evolve, it seems crucial to obtain more data on how females actually respond to multivariate traits.

In this study we have shown that there are female preferences among multivariate signals that represent the extant variation in a túngara frog population. This is consistent with the general result of previous studies of this species demonstrating that signal variation within the species is subject to sexual selection by female mate choice (Ryan 1985). Furthermore, these preference experiments suggest that this selection is not stabilizing. The region of the acoustic landscape with the most attractive calls includes the mean call, but attraction is just as high for some calls that are approximately one and two standard deviations away. There are also regions of call space that are clearly not very attractive.

Even though this study suggests that there will be selection on calls, we can hardly guess how the call, as a suite of multivariate characters, would evolve in response to selection. We have no information on heritability of call characters, and given the weak but significant correlations among call characters, it is not clear to what extent traits might evolve together or independently. But the lack of any very strong correlations among call characters suggests a lack of biomechanical constraints that would impede the independent evolution of individual call traits.

Our study was not designed to investigate the role of individual call characters in call attractiveness; we have addressed these issues in other studies (Ryan 1980, 1985; Wilczynski et al. 1995, 1999; Ryan and Rand 1999, 2001). In this study we used a statistical analysis to elucidate the call characters correlated with attractiveness. Some of the call characters identified in the stepwise multiple regression analysis are consistent with previous studies, chuck dominant frequency (Ryan 1980, 1985; Wilczynski et al. 1995; Ryan and Rand 2001) and the relative amplitude of the chuck to the whine (M. J. Ryan and A. S. Rand, unpubl. data), although others, such as call duration, have not been tested in great detail (Ryan et al. 1990). We are somewhat skeptical of strong conclusions drawn from this statistical approach. Instead, we think these results should be used to generate experimentally testable hypotheses rather than reach final conclusions about what signal parameters are salient to receivers. This might not be easily done in many cases since, as we have argued above, the attractiveness of one trait might depend on the state of other call traits.

#### *Call Variation and Female Preference*

We used multidimensional scaling to plot female perception of call similarity. As indicated earlier, the perception we measure is not one of acoustic phenotype but one of meaning. If females perceive two calls as very attractive it does not

necessarily indicate that they sound similar to her, but that their meaning is similar. We think there are two interesting results from this exercise.

First, there is not a significant correlation between the location of the nine test calls in acoustic and perceptual MDS space (Fig. 5). This might be expected if there are sign stimuli. Only if females were attending to all signal parameters and weighting them equally would we expect strong concordance between acoustic and perceptual landscapes.

Second, although some calls are outliers in acoustic space, in fact the *Oa-d* calls were chosen because they were outliers, there are no such outliers in perceptual space. More generally, female perception tends to collapse acoustic variation in this system. This was not necessarily expected. It was also possible that some signals might have been found especially more attractive or unattractive given their acoustic similarity to other calls. This result reinforces the notion that interpretations of how sexual selection by female choice influences male trait evolution need to be based not only on the quantitative variation in male traits but how females perceive such variation (see also Cohen 1984).

#### ACKNOWLEDGMENTS

For discussion of some of the issues addressed here or comments on the manuscript we thank R. Cocroft, C. Gerhardt, M. Greenfield, M. Hauser, M. Kirkpatrick, J. Podos, B. Searcy, and anonymous reviewers. This research was supported by grants from the National Science Foundation (IBN-93-16185, 98-16564, 99-81631), and from the Smithsonian Institution's Scholarly Studies program. We benefited greatly from the logistical support of the Smithsonian Tropical Research Institution, and we are especially grateful to the research assistants who conducted these phonotaxis experiments: A. Angulo, X. Bernal, M. Dillon, L. Ferrari, J. Fiano, and F. Vera-Condioti.

#### LITERATURE CITED

- Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.
- Basolo, A. L. 1990. Female preference predates the evolution of the swordtail fish. *Science* 250:808–810.
- Bradbury, J. W., and S. L. Vehrencamp. 1998. Principles of animal communication. Sinauer Associates, Sunderland, MA.
- Catchpole, C. K., B. Leisler, and J. Dittani. 1986. Sexual differences in the responses of captive great reed warblers (*Acrocephalus arundinaceus*) to variation in song structure and size. *Ethology* 73:69–77.
- Cohen, J. 1984. Sexual selection and the psychophysics of female choice. *Z. Tierpsychol.* 64:1–8.
- Conner, W. E., T. Eisner, R. K. Vander Meer, A. Guerrero, and J. Meinwald. 1981. Precopulatory sexual interaction in an arctiid moth (*Utetheisa ornatrix*). Role of a pheromone derived from dietary alkaloids. *Behav. Ecol. Sociobiol.* 9:227–235.
- Cronin, H. 1991. The ant and the peacock: altruism and sexual selection from Darwin to today. Cambridge Univ. Press, Cambridge, U.K.
- David, H. A. 1988. The method of paired comparisons. Oxford Univ. Press, Oxford, U.K.
- Darwin, C. 1859. On the origin of species. Murray, London.
- . 1871. The descent of man and selection in relation to sex. Murray, London.
- Drewry, G., H. W. R. Heyer, and A. S. Rand. 1982. A functional analysis of the complex call of the frog *Physalaemus pustulosus*. *Copeia* 1982:636–645.
- Dudley, R., and A. S. Rand. 1991. Sound production and vocal sac inflation in the túngara frog, *Physalaemus pustulosus* (Leptodactylidae). *Copeia* 1991:460–470.
- Gerhardt, H. C. 1978. Discrimination of intermediate sounds in a synthetic call continuum by female green tree frogs. *Science* 199:1089–1091.
- Gerhardt, H. C., and F. Huber. 2002. Acoustic communication in insects and anurans. Univ. of Chicago Press, Chicago, IL.
- Gerhardt, H. C., and J. Schul. 1999. A quantitative analysis of behavioral selectivity for pulse rise-time in the gray treefrog, *Hyla versicolor*. *J. Comp. Physiol.* 185:33–40.
- Greenfield, M. D. 2002. Signalers and receivers: mechanisms and evolution of arthropod communication. Oxford Univ. Press, Oxford, U.K.
- Gwynne, D. T. 2001. Katydid and bush-crickets: reproductive behavior and evolution of the Tettigoniidae. Cornell Univ. Press, Ithaca, NY.
- Hauser, M. D. 1996. The evolution of communication. The MIT Press, Cambridge, MA.
- Holland, B., and W. R. Rice. 1988. Perspective: chase-away sexual selection: antagonistic seduction versus resistance. *Evolution* 52:1–7.
- Iwasa, Y., and A. Pomiankowski. 1994. The evolution of mate preferences for multiple sexual ornaments. *Evolution* 48:853–867.
- Jang, Y., and M. D. Greenfield. 1998. Absolute versus relative measures of sexual selection: assessing the contributions of ultrasonic signal characters to mate attraction in lesser wax moths, *Achroia grisella* (Lepidoptera: Pyralidae). *Evolution* 52:1383–1393.
- Kaneshiro, K. Y. 1988. Speciation in the Hawaiian *Drosophila*. *Bioscience* 38:258–263.
- Kime, N. M., A. S. Rand, M. Kapfer, and M. J. Ryan. 1998. Consistency of female choice in the túngara frog: a permissive preference for complex characters. *Anim. Behav.* 55:641–649.
- Kruskal, J. B., and M. Wish. 1978. Multidimensional scaling. Sage Publications, Newbury Park, CA.
- Liedloff, A. 1999. Mantel nonparametric calculator. Vers. 2.0. Available from the author via e-mail at: adam.liedloff@terc.csiro.au.
- Marques, R., and J. Bosch. 2001. Communication and mating in the midwife toads (*Alytes obstetricans* and *Alytes cisternassi*). Pp. 220–231 in M. J. Ryan, ed. Anuran communication. Smithsonian Institution Press, Washington, DC.
- McLennan, D. A., and M. J. Ryan. 1999. Interspecific recognition and discrimination based upon olfactory cues in swordtails. *Evolution* 53:880–888.
- Murphy, C. G., and H. C. Gerhardt. 2000. Mating preference functions of individual female barking treefrogs, *Hyla gratiosa*, for two properties of male advertisement calls. *Evolution* 54:660–669.
- Nelson, D. A., and P. Marler. 1990. The perception of birdsong and an ecological concept of signal space. Pp. 443–478 in W. C. Stebbins and M. A. Berkeley, eds. Comparative perception: complex signals. John Wiley, New York.
- Rand, A. S., M. J. Ryan, and W. Wilczynski. 1992. Signal redundancy and receiver permissiveness in acoustic mate recognition by the túngara frog *Physalaemus pustulosus*. *Am. Zool.* 32:81–90.
- Ritchie, M. G. 1996. The shape of female mating preferences. *Proc. Natl. Acad. Sci. USA* 93:14628–14631.
- Rosenthal, G. G. 2000. The behavioral ecology of visual signaling in swordtails. Ph.D. diss. Department of Zoology, University of Texas, Austin, TX.
- Ryan, M. J. 1980. Female mate choice in a Neotropical frog. *Science* 209:523–525.
- . 1985. The túngara frog: a study in sexual selection and communication. Univ. of Chicago Press, Chicago, IL.
- Ryan, M. J., and A. S. Rand. 1999. Phylogenetic inference and the evolution of communication in túngara frogs. Pp. 535–557 in M. K. Hauser, ed. The design of animal communication. The MIT Press, Cambridge, MA.
- . 2001. Feature weighting in signal recognition and discrim-

- ination by the túngara frog. Pp. 86–101 in M. J. Ryan, ed. Anuran communication. Smithsonian Institution Press, Washington DC.
- Ryan, M. J., and W. E. Wagner, Jr. 1987. Asymmetries in mating preferences between species: female swordtails prefer hetero-specific males. *Science* 236:595–597.
- Ryan, M. J., J. H. Fox, W. Wilczynski, and A. S. Rand. 1990. Sexual selection for sensory exploitation in the frog *Physalaemus pustulosus*. *Nature* 343:66–67.
- Schul, J. 1998. Song recognition by temporal cues in a group of closely related bush cricket species (genus *Tettigonia*). *J. Comp. Physiol.* 184:401–410.
- Schul, J., and S. L. Bush. 2002. Non-parallel coevolution of sender and receiver in the acoustic communication system of treefrogs. *Proc. R. Soc. Lond. B. Biol. Sci.* 269:1847–1852.
- Sokal, R. R., and F. J. Rohlf. 1995. *Biometry: the principles and practice of statistics in biological research*. W.H. Freeman, New York.
- Tinbergen, N. 1969. *The study of instinct*. Clarendon Press, Oxford, U.K.
- Wagner, W. E., Jr. 1998. Measuring female mating preferences. *Anim. Behav.* 55:1029–1042.
- Welch, A. M., R. D. Semlitsch, and H. C. Gerhardt. 1998. Call duration as an indicator of genetic quality in male gray tree frogs. *Science*. 280:1928–30.
- Wilczynski, W., A. S. Rand, and M. J. Ryan. 1995. The processing of spectral cues by the call analysis system of the túngara frog, *Physalaemus pustulosus*. *Anim. Behav.* 49:911–929.
- Wilczynski, W., A. S. Rand, and M. J. Ryan. 1999. Female preferences for temporal order of call components in the túngara frog: a Bayesian analysis. *Anim. Behav.* 58:841–851.
- Wilkinson, L. 1996. *Systat 10*. SPSS Inc., Chicago, IL.
- Zuk, M., D. Ligon, and R. Thornhill. 1992. Effects of experimental manipulation of male secondary sex characters on female mate preference in red jungle fowl. *Anim. Behav.* 44:999–1006.
- Zuk, M., R. Thornhill, J. D. Ligon, K. Johnson, S. Austad, S. H. Ligon, N. W. Thornhill, and C. Costin. 1990. The role of male ornaments and courtship behavior in female mate choice of red jungle fowl. *Am. Nat.* 136:459–473.

Corresponding Editor: T. Tregenza