

EXPERIMENTAL EVIDENCE FOR MULTIVARIATE STABILIZING SEXUAL SELECTION

ROBERT BROOKS,^{1,2} JOHN HUNT,^{1,3} MARK W. BLOWS,⁴ MICHAEL J. SMITH,^{1,5} LUC F. BUSSIÈRE,^{1,6} AND MICHAEL D. JENNIONS⁷

¹*School of Biological, Earth and Environmental Sciences, University of New South Wales, Sydney, NSW 2052, Australia*

²*E-mail: rob.brooks@unsw.edu.au*

⁴*School of Integrative Biology, University of Queensland, St. Lucia 4072, Queensland, Australia*

E-mail: mblows@zoology.uq.edu.au

⁷*School of Botany and Zoology, Australian National University, Canberra, Australian Capital Territory 0200, Australia*

E-mail: michael.jennions@anu.edu.au

Abstract.—Stabilizing selection is a fundamental concept in evolutionary biology. In the presence of a single intermediate optimum phenotype (fitness peak) on the fitness surface, stabilizing selection should cause the population to evolve toward such a peak. This prediction has seldom been tested, particularly for suites of correlated traits. The lack of tests for an evolutionary match between population means and adaptive peaks may be due, at least in part, to problems associated with empirically detecting multivariate stabilizing selection and with testing whether population means are at the peak of multivariate fitness surfaces. Here we show how canonical analysis of the fitness surface, combined with the estimation of confidence regions for stationary points on quadratic response surfaces, may be used to define multivariate stabilizing selection on a suite of traits and to establish whether natural populations reside on the multivariate peak. We manufactured artificial advertisement calls of the male cricket *Teleogryllus commodus* and played them back to females in laboratory phonotaxis trials to estimate the linear and nonlinear sexual selection that female phonotactic choice imposes on male call structure. Significant nonlinear selection on the major axes of the fitness surface was convex in nature and displayed an intermediate optimum, indicating multivariate stabilizing selection. The mean phenotypes of four independent samples of males, from the same population as the females used in phonotaxis trials, were within the 95% confidence region for the fitness peak. These experiments indicate that stabilizing sexual selection may play an important role in the evolution of male call properties in natural populations of *T. commodus*.

Key words.—Adaptive landscape, call structure, nonlinear selection, selection analysis, selection gradient, stabilizing selection.

Received October 28, 2004. Accepted January 14, 2005.

By extending Wright's (1931) concept of the fitness landscape in genotypic space to the phenotype, Simpson (1944, 1953) provided what remains the major conceptual bridge between microevolutionary processes (e.g., selection, genetic drift) and patterns of macroevolutionary change (e.g., speciation, adaptive radiation) (Fear and Price 1998; Schluter 2000; Arnold et al. 2001; Gavrilets 2004). Although Simpson's conceptualization of the adaptive landscape has enjoyed considerable theoretical development (e.g., Lande 1976, 1979), its importance has largely gone unappreciated by empiricists (Arnold et al. 2001). A fundamental principle of evolution on adaptive landscapes is that in the absence of frequency-dependent selection the population mean will tend to evolve uphill toward an adaptive peak (Simpson 1953; Lande 1976; Fear and Price 1998; Arnold 2003). Analytical (Lande 1976) and simulation (Bürger and Lande 1994) studies have provided formal support for Simpson's (1953) prediction that on an adaptive landscape with a single peak of elevated fitness (i.e., under stabilizing selection), the population mean should evolve to match that peak. This prediction is central to much of evolutionary biology, yet surprisingly few studies have formally demonstrated that populations

evolve to match adaptive peaks (Travis 1989; but see Armbruster 1990; Benkman 1993, 2003; Schluter 2000).

Although adaptive landscapes (which describe the relationship between the phenotype and population mean fitness) are rarely estimated (Arnold et al. 2001), the estimation of individual fitness surfaces (which describe the relationship between the phenotype and individual relative fitness within a population) have been the target of numerous empirical studies (Kingsolver et al. 2001). Considerable attention has been paid to estimates of linear (directional) selection, which appears to be widespread (Kingsolver et al. 2001), although it remains unclear how strong it may be in natural populations (Hersch and Phillips 2004). Less attention has been paid to estimating nonlinear selection (Kingsolver et al. 2001; Blows and Brooks 2003). Stabilizing selection is convex nonlinear selection in which the optimum phenotype is at an intermediate point in the range of phenotypes in the population (Lande and Arnold 1983; Mitchell-Olds and Shaw 1987; Phillips and Arnold 1989; Travis 1989). It is through the characterization of quadratic fitness surfaces, and in particular, the presence of stabilizing selection on those surfaces, that the heuristic fitness peaks envisaged by Simpson (1944, 1953) may be empirically investigated and visualized.

There is surprisingly little direct evidence for the importance of stabilizing selection (Travis 1989; Schluter 2000), and therefore fully characterized fitness peaks, in natural populations. Estimates of convex quadratic selection gradients are generally smaller than linear selection gradients (Kingsolver et al. 2001), and of a similar magnitude to concave

³ Present address: Faculty of Life Sciences, University of Manchester, 3.614 Stopford Building, Oxford Road, Manchester, M13 9PT, United Kingdom; E-mail: john.hunt@manchester.ac.uk.

⁵ Present address: Arthur Rylah Institute for Environmental Research, 123 Brown Street, Heidelberg, Victoria, Australia, 3084.

⁶ Present address: Zoologisches Museum der Universität Zürich, Winterthurerstrasse 190, CH 8057 Zürich, Switzerland; E-mail: luc.bussiere@access.unizh.ch.

quadratic selection gradients, leading to the claim that nonlinear selection in general (Kingsolver et al. 2001) and stabilizing selection in particular (Barton and Keightley 2002; Zhang et al. 2002; Zhang and Hill 2003) may be relatively weak. However, most analyses of stabilizing selection have been inherently univariate in nature (e.g., Gerhardt 1991; Benkman 1993; Cresswell 2000), even though multiple regression approaches (Lande and Arnold 1983) may have been used to estimate quadratic selection gradients. Blows and Brooks (2003) pointed out that the widespread practice of ignoring correlational selection (Kingsolver et al. 2001) was responsible for underestimating the strength of nonlinear selection. Canonical rotation (Lande and Arnold 1983; Phillips and Arnold 1989) identifies major axes of nonlinear selection on fitness surfaces that comprise both quadratic and correlational selection gradients. Canonical rotation facilitates the detection of multivariate stabilizing (or disruptive) selection because the quadratic terms describing nonlinear selection on each axis can be interpreted independently of one another. Consequently, there remains considerable scope for empirical estimation of the strength of nonlinear selection, and novel tests to assess the evolutionary consequences of its presence.

Stabilizing Sexual Selection on Male Acoustic Signals

Mate choice may often exert directional sexual selection on male signal traits (Andersson 1994). However, it is also possible that intermediate male phenotypes are most preferred under some circumstances (Andersson 1994; Mead and Arnold 2004). Preference for intermediate trait values is commonly the case for the structural and spectral properties of acoustic calls in frogs and crickets because the most effective call will be the one whose structure and frequency best match the peak sensitivity of the receiver's psychosensory apparatus (Ryan and Wilczynski 1988; Murphy and Gerhardt 2000; Stumpner and von Helversen 2001). For example, the central auditory system of female cricket frogs, *Acris crepitans*, is most sensitive to an intermediate frequency within the range of dominant call frequencies made by males (Ryan and Wilczynski 1988).

There is evidence that preferences for intermediate frequencies exert stabilizing sexual selection on male call frequency in cricket frogs (Ryan and Wilczynski 1988). Several other univariate studies (e.g., Ryan and Keddy-Hector 1992; Polakow et al. 1995; Ritchie 1996; Murphy and Gerhardt 2000) indicate that preferences for intermediate values of call frequency and various temporal structure measures may likewise result in stabilizing selection on these properties. However, analyses need to account for the simultaneous effects of selection on all relevant aspects of the phenotype (Jang and Greenfield 1998; Ryan and Rand 2003), particularly when there are genetic or phenotypic correlations among traits. The components of call structure are often phenotypically correlated in grasshoppers (Klappert and Reinhold 2003), field crickets (Scheuber et al. 2003; J. Hunt, M. Jennions, and R. Brooks, unpubl. data) and frogs (Ryan and Wilczynski 1991; Castellano et al. 2002). The use of selection analysis techniques that are designed specifically to deal with correlated traits (Lande and Arnold 1983; Phillips and Arnold

1989; Ryan and Rand 2003) is therefore especially pertinent to studies of selection on call structure.

In this paper we investigated the form and strength of sexual selection on a suite of five acoustic traits of male *Teleogryllus commodus*, a native Australian field cricket. Our primary objective was to describe the pattern of selection that female phonotaxis to male calls imposes on male call structure in a way that is free of the problems associated with unmeasured variables and correlations among traits (Lande and Arnold 1983; Mitchell-Olds and Shaw 1987). We overcame these traditional limitations of selection analyses by constructing artificial male calls, based on the distributions of the acoustic traits of males from a single natural population. We played the artificial calls back to females in binary phonotaxis trials as a bioassay of their attractiveness, and used multiple regression methods to estimate the response surface describing male attractiveness (which we refer to as the fitness surface). We demonstrated that this surface was a single multivariate peak, and used the Box-Hunter (1954) technique to estimate the confidence region for the peak. Finally, the mean trait values from the calls of four independent samples of males from the same population were mapped onto the fitness surface to test the prediction (Simpson 1944, 1953; Lande 1976) that multivariate stabilizing selection should lead populations to evolve to match the fitness peak of such a surface.

MATERIALS AND METHODS

*The Advertisement Call of *Teleogryllus commodus**

The advertisement call of *Teleogryllus commodus* contains two discrete pulse types beginning with a single chirp sequence followed by a variable number of trill sequences that are repeated at a faster rate (Hill et al. 1972; Bentley and Hoy 1972; Fig. 1). To obtain a sample of calls, we collected adult males in the field at Smith's Lakes, New South Wales, Australia (32°22'S, 152°30'E), and brought them back to the University of New South Wales campus in Sydney. Over the following week, we opportunistically recorded the calls of 15 males in a greenhouse at 20–22°C using a Sony Professional Walkman (WM-D6C) equipped with a Sennheiser (Wedemark, Germany; e845) microphone. We subsequently digitized the calls from tape and used Canary software (version 1.2.4; Bioacoustics Research Program, Cornell University, Ithaca, NY) to measure the following properties of 10 randomly selected calls per male: the number of pulses in the chirp, chirp duration, chirp pulse duration, chirp pulse interval, the interval between the chirp and the first trill, the number of trills, the number of pulses in and duration of each trill, the interval between each trill, trill pulse duration and interpulse duration, the interval between the last trill pulse of one call and the first chirp pulse of the next call, the duration of the entire call, and the dominant frequency of the call. We present a schematic illustration of the temporal attributes of the male advertisement call in Figure 1.

Many of the temporal measures were positively skewed; before further analysis we transformed these measures using natural logarithms to generate normal distributions. We calculated the mean for the appropriate normally distributed form of each measure for each male and then estimated the

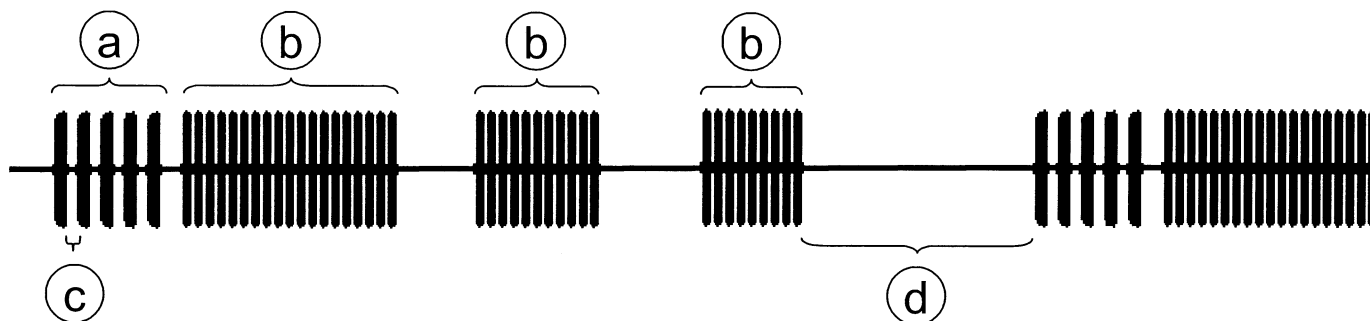


FIG. 1. Schematic representation of the advertisement call of *Teleogryllus commodus*. Each call comprises a chirp (a) followed by one or more trills (b). We manipulated the number of pulses in the chirp (CPN), the duration of the interval between pulses in the chirp (CIPD, c), the number of trills (TN), the intercall interval duration (ICD, d) and the dominant frequency of the call (DF, not shown). All other call parameters were kept constant. See Table 1 for values used.

grand mean and standard deviation about that mean for the sample of 15 males.

Design of Artificial Calls

Lande and Arnold's (1983) multiple-regression-based methods for the analysis of linear and nonlinear selection were developed explicitly to account for the effects of correlations among traits. However, strong correlations among traits still generate potential problems. Strong phenotypic correlations restrict the type of trait combinations that occur in real individuals, and thus estimates of selection gradients tend to be based on a restricted subset of the possible trait values and combinations and may be biased as a result (Mitchell-Olds and Shaw 1987). It is also conceivable that some combinations of traits that are particularly attractive to females may be biophysically impossible for males to generate (Ryan and Rand 2003), a situation that would constrain the population from reaching some adaptive peaks.

A further limitation of selection analyses based on multiple regression is that they provide correlative (e.g., Blows et al. 2003) rather than manipulative (e.g., Cresswell 2000) evidence of selection. When variables that are not measured (and thus not included in the selection analysis) have strong effects on fitness, estimates of selection on the measured traits may be biased (Lande and Arnold 1983; Mitchell-Olds and Shaw 1987). Therefore selection analysis should ideally be accompanied by manipulative experimentation or path analyses to determine the causal basis of selection (Mitchell-Olds and Shaw 1987; Sinervo and Basolo 1996; Sinervo 1998).

We therefore chose to manipulate the following five call parameters that do not have strong functional relationships: the number of pulses in the chirp (chirp pulse number: CPN), the mean duration of the interval between pulses in the chirp (chirp interpulse duration: CIPD), the number of trills (trill number: TN), the interval between the last trill pulse of one call and the first chirp pulse of the next call (intercall duration: ICD), and the dominant frequency of the call (dominant frequency: DF). We constructed 300 artificial calls that varied only in the above five manipulated traits. Further, we ensured that the distribution of each trait followed the same univariate distribution as the original 15 measured males, but that the five measures were uncorrelated. We first drew five independent random numbers between 0 and 1. We then used

the NORMSINV function in MS Excel (Microsoft Corp., Redmond, WA) to obtain the inverse of the standard normal cumulative distribution corresponding to each of the random numbers (i.e., to turn a probability into a z -score). The resulting number, z , is the number of standard deviations above or below the mean that the observation would take. We then multiplied the value by the measured standard deviation for the relevant trait and added the measured mean (i.e., to turn a z -score into a measure, x , in the original units). For the two variables that had been transformed to natural logarithms before we estimated means and standard deviations, we transformed this value, x , to the original biological scale by raising e^x .

We generated 300 random calls in this way, resulting in a distribution of uncorrelated call properties with the same univariate means and standard deviations as the original sample of 15 males. Our method allowed us to explore selection on each of the five traits without the complications that arise due to strong correlations among traits. It provides a practical alternative to factorial designs in which we could only have examined three values per trait if we had used a similar number (i.e., 243) of calls. We held the remaining call properties constant, using the mean values from the 15 measured males. The unmanipulated values used in our calls are reported in Table 1. Each focal call was played against a control call (whose attributes are provided in Table 1) in our two-speaker phonotaxis trials.

We synthesized calls using Cool Edit Pro software (version 1.2; now available as Adobe Audition, Adobe Systems Inc., San Jose, CA). We created the focal calls on one channel of a stereo computer file and a control call on the other channel. The calls were constructed so that at the start of the file, the control call began exactly in the middle and the focal call at the beginning. During playback phonotaxis trials, the file was started at an arbitrary place and run as a continuous loop to ensure that there were no systematic "leader-follower" effects (Dyson and Passmore 1988; Greenfield and Roizen 1993; Dyson et al. 1994).

Phonotaxis Experiment

We conducted 1200 two-speaker phonotaxis trials on 300 females. Each female was used in four phonotaxis trials, each trial with a different focal call, run over consecutive days.

TABLE 1. Values of call properties that were manipulated and those kept constant in our construction of artificial calls. The values used in the control call are given in the right hand column. ms, milliseconds.

Call property	Abbreviation	Mean	SD	Control
Properties manipulated				
Chirp pulse number	CPN	6.447	0.803	6
Chirp inter-pulse duration (\log_e) (ms)	CIPD	3.325	0.128	3.333
Trill number	TN	2.70	1.14	3
Intercall duration (\log_e) (ms)	ICD	5.897	0.737	5.897
Dominant frequency (kHz)	DF	3.83	0.14	3.83
Properties kept constant				Value used
Chirp pulse duration (ms)				29
Chirp-trill 1 interval (ms)				31
Trill 1–2 interval (ms)				117
Trill 2–3 interval (ms) ¹				141
Trill pulse duration (ms)				20
Trill interpulse duration (ms)				8
Trill 1 pulse number				19
Trill 2 pulse number				11
Trill 3 pulse number ²				6

¹ Value also used for all subsequent intertrill intervals.

² Value also used for all subsequent trills.

Each focal-control-call pair was used in four phonotaxis trials: once on a female's first day, once on another female's second day, and so on. The order in which calls were used in the experiment was otherwise random.

We presented the calls using SoundEdit 16 (version 2.0; Macromedia, San Francisco, CA) software running on an iMac G3 (Apple Computers, Cupertino, CA) computer through a BOSS (Oxnard, CA) BA-250 audio amplifier to two Koss (Hazelwood, MO) HD50 speakers. The speakers were positioned centrally on opposite sides of a 106×106 cm arena with their surface flush against the arena walls. The arena was fitted with sound insulation acoustic foam to reduce reverberation. The speaker through which the focal call was played was alternated. Before every trial, we used a Radio Shack (Fort Worth, TX) SPL meter to ensure both calls had amplitudes of 75 dB (root mean square) at the center of the arena.

Trials were conducted at 20–25°C under red light to minimize observer disturbance. The female was first placed in the center of the arena under a piece of egg carton that was, in turn, under an inverted hemispheric wire tea strainer (diameter 6 cm). The calls were then played for 1 min before we removed the strainer. An observer then watched the female through a small hole in the side of the arena until she entered the 15-cm radius semicircular choice zone in front of a speaker, in which case she was said to have chosen that call. Female response time was moderately repeatable (repeatability = 0.428 ± 0.03 , $F_{297,925} = 4.079$, $P < 0.0001$), and we obtained similar results when we used the absolute preference for a call (whether it was preferred over the control call) or a time-weighted preference score. Due to the more straightforward biological interpretation (Brodie and Janzen 1996), we used the absolute preference score.

Fitness Surface Estimation

Each trial was assigned a score of 0 if the female preferred the control call and 1 if the female preferred the focal call. We transformed this score to a mean of one and all traits to zero means and unit variances as suggested by Lande and

Arnold (1983). We then fitted a linear regression including all five manipulated call property values in order to estimate the vector of linear selection gradients, β (Lande and Arnold 1983). We then used a quadratic regression model incorporating all linear, quadratic, and cross-product terms to estimate the matrix of nonlinear selection gradients, γ , that describe the response surface, which we interpret as the multivariate individual fitness surface.

It is possible to underestimate the strength of nonlinear selection if one attempts to interpret the size and significance of the γ -coefficients individually (Phillips and Arnold 1989; Blows and Brooks 2003). We therefore explored the extent of nonlinear selection by performing a canonical analysis to find the major axes of the response surface (Phillips and Arnold 1989; see method in appendix 1 of Blows and Brooks 2003). The strength of nonlinear selection along each of the eigenvectors (\mathbf{m}_i) of the response surface is given by their eigenvalues, λ_i .

To estimate selection gradients in a way that is relevant to equations of evolutionary change, no transformations of the fitness measure apart from conversion to relative fitness should be made (Lande and Arnold 1983). However, violation of distributional assumptions does create problems for testing significance for estimated gradients (Draper and Smith 1981; Mitchell-Olds and Shaw 1987). There were thus two problems with applying traditional significance tests to our data. First, assumptions of normality were violated for our binary response variable and for at least two of the five original variables, chirp interpulse duration and intercall duration, which were right skewed. Second, the use of 300 females to obtain 1200 observations meant that some observations (and their residuals) were potentially nonindependent. We therefore used bootstrap resampling methods (as suggested by Mitchell-Olds and Shaw 1987) to test for bias in our estimates of selection gradients based on all 1200 observations, and randomization tests to assess the significance of estimated gradients. For bootstrap tests for bias we drew one observation at random from each female. We then used these 300 observations in a new multiple regression model to estimate

TABLE 2. The vector of standardized linear selection gradients (β) and the matrix of standardized quadratic and correlational selection gradients (γ). Randomization test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

	β	γ				
		CPN	CIPD	TN	ICD	DF
Chirp pulse number (CPN)	0.007	0.006				
Chirp interpulse duration (CIPD)	-0.003	0.017	-0.006			
Trill number (TN)	0.015	0.019	0.039	-0.040		
Intercall duration (ICD)	-0.214***	-0.022	-0.036	0.086*	-0.080**	
Dominant frequency (DF)	0.059	0.024	-0.031	0.041	-0.013	-0.047*

each selection gradient. We then replaced all observations and repeated the bootstrap 10,000 times to obtain confidence intervals for each gradient. This procedure established that our overall estimates were not biased by assuming independence of choice among trials by the same female ($P = 0.185-0.659$). To assess the significance of the overall estimates, we then conducted a randomization test in which response scores were randomized across the 1200 trials to obtain an expected distribution of each gradient under random choice. Probabilities are the number of randomizations (of 9999) in which the gradient pseudo-estimate was equal to or farther away from zero than the original estimate. We conducted separate bootstrap and randomization analyses for the multiple regression models for directional selection (i.e., model containing only the linear terms) and for the full quadratic model. We used the same resampling procedures to assess the significance of linear and nonlinear selection on each of the eigenvectors extracted from canonical rotation of the response surface (Bisgaard and Ankenman 1996; Blows and Brooks 2003).

Fitness Surface Visualization

We used thin-plate splines (Green and Silverman 1994) to visualize the major axes of the fitness surface extracted from the canonical rotation. Thin-plate splines are a nonparametric approach that provides a less constrained view of the surface than the best quadratic approximation (Blows et al. 2003). We used the Tsp function in the *fields* package in R (available via <http://www.r-project.org>) to fit a spline surface using the value of the smoothing parameter, λ , that minimized the generalized cross-validation (GCV) score. We then plotted the surface using R.

Estimating the Stationary Point and Its 95% Confidence Region

One can estimate the position of the stationary point on each eigenvector by differentiating the equation for the rotated response surface with respect to each eigenvector in turn (Phillips and Arnold 1989). Placing confidence on the location of this stationary point is somewhat more complicated. A method for computing a confidence region for the location of a stationary point on a response surface was developed by Box and Hunter (1954). This method has not been applied to evolutionary fitness surface analysis and has, until recently, seldom been used in response-surface analysis because it is algebraically complicated, and appropriate software has generally been unavailable. Recently, tools have been developed to plot Box-Hunter confidence regions (del

Castillo and Cahya 2001; Peterson et al. 2002). We used a program written for the MAPLE computer algebra software package (del Castillo and Cahya 2001) to plot the conditional 95% confidence regions for our rotated fitness surface.

Test of Whether the Population Mean is at the Peak

Having established that there was evidence of stabilizing selection (see Results), we obtained four independent samples of males to test whether the calls of Smith's Lake males are, on average, at the fitness peak. We individually reared males that were second-generation laboratory descendants of animals collected with the original males (on which the call designs were based). These descendants were reared on one of three diets: high protein (pellets comprising 100% high protein fish food), medium protein (75% fish food, 25% oatmeal by weight), or low protein (50% fish food, 50% oatmeal; for more details on feeding and rearing conditions see Hunt et al. 2004). We recorded males between 10 and 25 days posteclosion (age did not significantly affect call structure; Hunt et al. 2004) in a temperature controlled room set to 28°C. We also collected 30 males from the same field as the original males one year later (in March 2003), and recorded their calls under the same greenhouse conditions as the original sample of males. We then measured each of the recordings in the same way as for the original males, obtained scores for each of the eigenvectors identified above as major axes of nonlinear selection for each male, and obtained means of those measurements for each of the four samples.

We tested for differences among the samples in each of the dimensions of the fitness surface using multivariate analysis of variance (MANOVA). Our approach to assessing whether the four independent samples of males is at or near to the estimated fitness peak was to compare the mean eigenscore of each sample visually with the 95% confidence region for the peak.

RESULTS

Standardized linear, quadratic, and correlational selection gradients and information on significance from randomization tests are presented in Table 2. There was significant directional selection favoring shorter intercall durations, but not on any other trait. There was also significant nonlinear selection involving intercall duration, including convex (negative γ) selection on intercall duration and positive correlational selection on trill number and intercall duration. There was significant convex selection on dominant frequency.

We conducted a canonical rotation of the γ matrix of quadratic selection gradients presented in Table 2 to find the

TABLE 3. The \mathbf{M} matrix of eigenvectors from the canonical analysis of γ . The linear (θ_i) and quadratic (λ_i) gradients of selection along each eigenvector are given in the last two columns. The quadratic selection gradient (λ_i) of each eigenvector (\mathbf{m}_i) is equivalent to the eigenvalue. Randomization test: * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

	M					Selection	
	CPN	CIPD	TN	ICD	DF	θ_i	λ_i
\mathbf{m}_1	0.800	0.497	0.305	-0.057	0.130	0.028	0.018
\mathbf{m}_2	-0.446	0.806	-0.001	-0.091	-0.377	-0.017	-0.001
\mathbf{m}_3	-0.302	0.003	0.776	0.500	0.240	-0.082*	-0.010
\mathbf{m}_4	-0.257	0.208	-0.102	-0.405	0.846	0.144***	-0.054**
\mathbf{m}_5	0.068	0.244	-0.543	0.758	0.258	-0.160***	-0.119***

major axes of nonlinear selection. We present the resulting \mathbf{M} matrix of eigenvectors and their associated eigenvalues in Table 3. This table also contains the gradients of linear (θ_i) and nonlinear (λ_i) selection on each eigenvector and information on significance derived from randomization tests. Four of the five eigenvectors, including the only three along which there was any significant selection (\mathbf{m}_3 , \mathbf{m}_4 , and \mathbf{m}_5) had negative eigenvalues indicating convex selection. The remaining eigenvector (\mathbf{m}_1) had a positive eigenvalue, but there was no evidence that selection along this eigenvector was significantly different from zero. Therefore, although this single eigenvector of positive quadratic selection formally indicates that the response surface contains a saddle, the lack of significant selection on this axis suggests that the surface is not significantly different from a multivariate peak in shape.

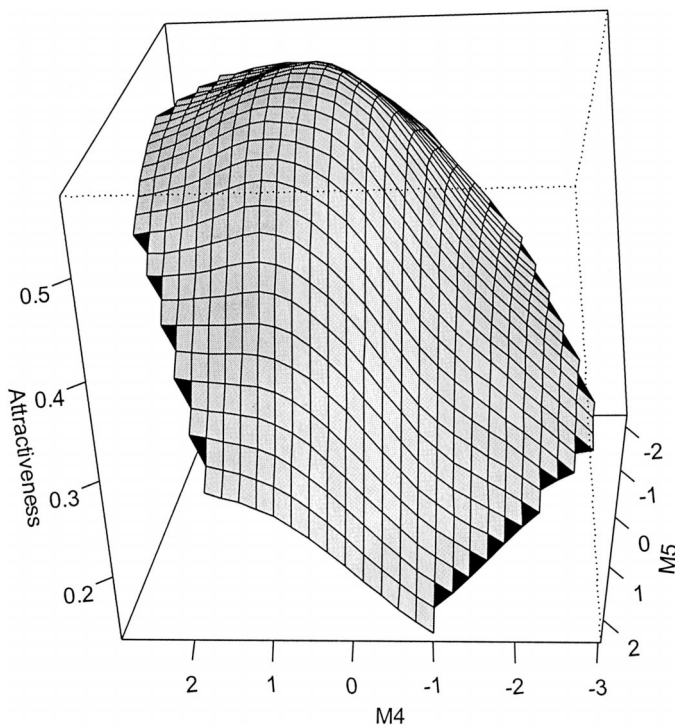


FIG. 2. Thin-plate spline perspective-view visualization of the fitness surface on the two major axes of nonlinear selection, \mathbf{m}_4 and \mathbf{m}_5 . Attractiveness is the probability of a female responding positively to the focal call rather than the control.

The Fitness Surface and the Adaptive Peak

Examining the eigenvectors along which there was significant linear or nonlinear selection revealed that all three had negative eigenvalues, suggesting multivariate stabilizing selection. However, along \mathbf{m}_3 , there was only significant linear selection rather than nonlinear selection. In contrast, there was significant linear and nonlinear selection on \mathbf{m}_4 and \mathbf{m}_5 . The response surface on these two major axes of nonlinear selection is shown in Figure 2. The surface and the positions in this space occupied by the 300 artificial calls on which the surface is based are shown in Figure 3a. It has been suggested that the term *stabilizing selection* only be used to describe situations in which there is convex selection and an intermediate maximum within the space sampled (Mitchell-Olds and Shaw 1987). The local maxima for \mathbf{m}_4 and \mathbf{m}_5 are within the sampled space ($\mathbf{m}_4 = 1.44$, $\mathbf{m}_5 = -0.67$; Fig. 2).

The Box-Hunter (1954) conditional 95% confidence region for the location of the stationary point in \mathbf{m}_4 - \mathbf{m}_5 space is shown in Figure 3b. The stationary point, the calls of individual males, and the means of the four independent samples from Smith's Lakes are also plotted on Figure 3. All four sample means are within half a standard deviation of the stationary point and well within the 95% confidence region. Even though the sample means are aligned with the peak of the fitness surface, many individual males from each of the four samples fall outside the 95% confidence region, suggesting that males are not physiologically or mechanically constrained to this area of the fitness surface. MANOVA revealed that the only eigenvector on which the four samples differed significantly from one another was \mathbf{m}_2 (Table 4), the eigenvector of weakest (and not significant) linear and nonlinear selection (Table 3).

DISCUSSION

We have shown that female *Teleogryllus commodus* from Smith's Lakes orient to male calls in the laboratory in a manner that strongly suggests the action of multivariate stabilizing sexual selection on male advertisement call structure. Moreover, the peak of the multivariate response surface successfully predicted the mean phenotype of the natural population reared under a variety of environmental conditions. Mean call parameters from four independent samples of males from the Smith's Lakes population were positioned within the 95% confidence region of the stationary point on the multivariate response surface. This finding is consistent with the theoretical prediction (Simpson 1944, 1953; Lande

1979) that populations should evolve to match adaptive peaks where single peaks occur in an adaptive landscape.

Stabilizing Sexual Selection on Male Acoustic Traits

Although tests for nonlinear sexual selection in a range of taxa have applied multivariate methods (Jang and Greenfield 1998; Blows et al. 2003, 2004; Ryan and Rand 2003), none of these studies have found clear evidence of multivariate stabilizing selection. Ryan and Rand (2003) constructed artificial calls to explore the relationship between call properties and fitness in the Túngara frog (*Physalemus pustulosus*) unconstrained by phenotypic correlations among call properties. They used a multidimensional scaling approach to understand variation in calls, and showed that although the most attractive part of the acoustic landscape includes the mean call, it was not centered on the mean. This finding suggested that the population had not evolved to match an acoustic fitness peak, although a formal demonstration of the shape of the acoustic landscape would help to clarify whether this was the case. Here we have demonstrated not only that there is an intermediate peak in the acoustic fitness surface in *T. commodus*, but also that the population mean appears to be within half a standard deviation of the peak attractiveness. Although this evidence is consistent with call convergence on the peak, sexual selection theory suggests that the match between female preference and male call structure is likely to have also involved female adaptation and preference-call coevolution (Houde 1993; Andersson 1994; Mead and Arnold 2004).

Historically, traits important in mate recognition have been thought to be under strong stabilizing natural selection (Paterson 1982, 1985; Butlin et al. 1985), and consequently, to be relatively invariant. However, univariate studies of mate choice in several orthopteran (Butlin et al. 1985; Ritchie 1996) and anuran (Gerhardt 1991; Ryan et al. 1992) species suggest that mate choice alone may be sufficient to exert stabilizing sexual selection on the temporal and spectral components of the male's advertisement call. Gerhardt (1991) showed that in three species of frogs, mate choice tends to exert stabilizing selection on relatively "static" components of male signals (i.e., signals that display relatively little variation, i.e., $CV < 5\%$, among males within populations). By contrast, preferences for more variable or "dynamic" components of male advertisement often appear to be strongly directional (Gerhardt 1991). Although one explanation for these patterns is that stabilizing mating preferences tend to evolve when they focus on less variable components of mating signals, an alternative explanation is that selection shapes the phenotypic variance in these traits. Quantitative genetic theory predicts that stabilizing selection may result in the genetic variance-covariance matrix of the suite of traits conforming to the orientation of the fitness surface (Lande 1980; Cheverud 1984). Therefore, it may be that multivariate stabilizing selection imposed by the limitations of the female sensory system may result in certain components of the signaling system becoming more stereotyped.

At a more proximate level, stabilizing selection on male call structure may result directly from the properties of the female neurosensory system (Greenfield 2002). Selection will

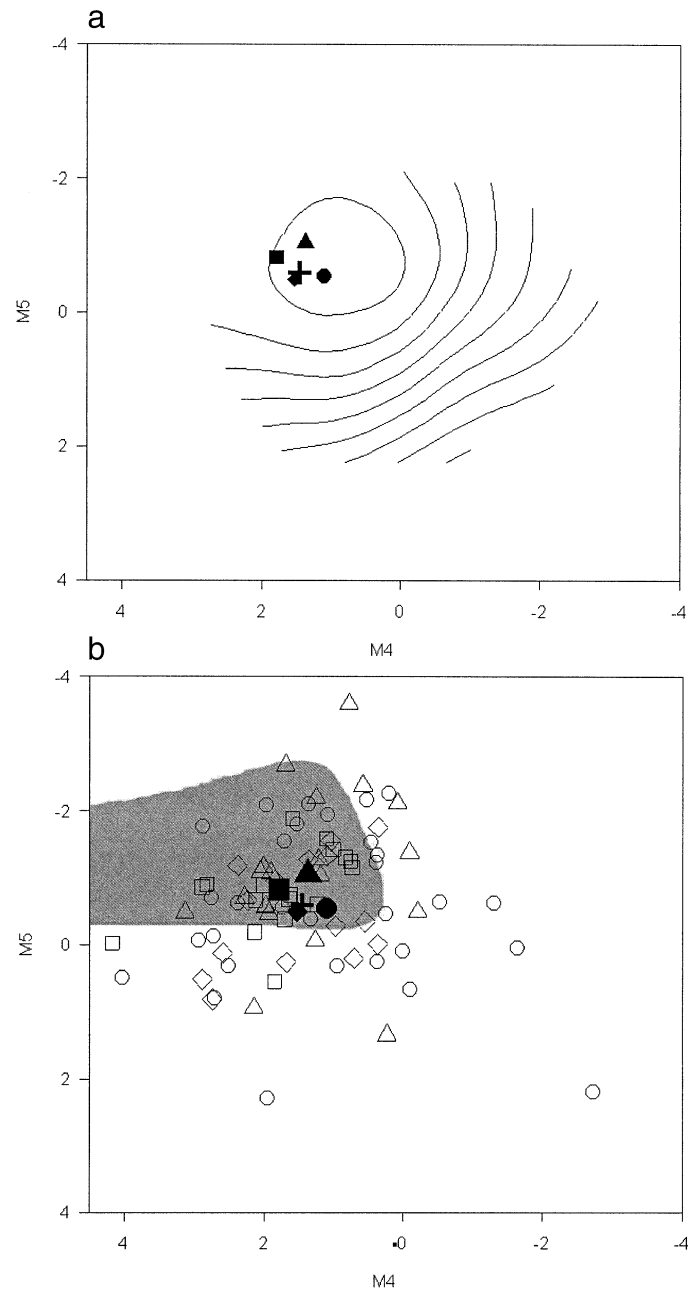


FIG. 3. (a) Thin-plate spline contour map visualization of the fitness surface on the two major axes of nonlinear selection, m_4 and m_5 , showing the peak of the surface (+) and the mean call values of the four independent samples of males (circles, wild caught males; squares, high protein diet; diamonds, medium protein diet; triangles, low protein diet). (b) The Box-Hunter conditional 95% confidence region (shaded area) for the position of the stationary point on the surface, the mean call values of the four independent samples of males (closed symbols, as above), and the actual call values of the males from these four samples (open symbols).

favor males that make calls with acoustic and temporal properties that best match the greatest sensitivity of the receiver (Ryan and Wilczynski 1988; Ryan et al. 1992; Stumpner and von Helversen 2001). The most attractive male calls are often those with a dominant carrier frequency equal or close to the

TABLE 4. Multivariate analysis of variance (MANOVA) comparing means of the four independent samples of males on each of the five eigenvectors resulting from canonical rotation of the quadratic fitness surface.

MANOVA (Pillai's trace = 0.422, $F_{15,228} = 2.49$, $P = 0.002$)						
	Sample means \pm standard error				ANOVA	
	Wild	High protein	Medium protein	Low protein	$F_{3,78}$	P
\mathbf{m}_1	-0.501 \pm 0.244	-1.240 \pm 0.443	-1.566 \pm 0.412	-1.641 \pm 0.419	2.52	0.064
\mathbf{m}_2	-0.866 \pm 0.218	-2.467 \pm 0.350	-2.330 \pm 0.348	-2.786 \pm 0.332	10.23	0.000
\mathbf{m}_3	0.566 \pm 0.157	-0.017 \pm 0.231	0.115 \pm 0.186	0.371 \pm 0.288	1.54	0.212
\mathbf{m}_4	1.089 \pm 0.277	1.777 \pm 0.218	1.518 \pm 0.229	1.367 \pm 0.202	1.35	0.264
\mathbf{m}_5	-0.547 \pm 0.219	-0.820 \pm 0.147	-0.496 \pm 0.209	-1.049 \pm 0.260	1.26	0.295
N	30	17	15	20		

frequencies to which female auditory organs are most sensitive (Ryan and Wilczynski 1988; Hennig and Weber 1997; Stumpner and von Helversen 2001) with more extreme frequencies discriminated against. Likewise, female response to the temporal parameters of male calls are subject to filtering by the nervous system (Hennig and Weber 1997; Stumpner and von Helversen 2001), with calls that have extreme properties not eliciting a phonotactic response. For example, female *T. commodus* have different temporal filters tuned to pulse duration: one is specific to the duration of pulses in the chirp and the other is specific to trill pulse duration (Hennig and Weber 1997). Within this proximate framework, it is possible to envisage how call properties such as dominant frequency, chirp pulse duration and chirp interpulse duration might be subject to stabilizing selection due solely to the properties of the female receptor and processing systems. However, more work on the neurophysiology of female *T. commodus* is needed before the proximate mechanisms that underlie the pattern of multivariate stabilizing selection are fully understood.

Linear Sexual Selection on Male Acoustic Traits

Although the predominant form of sexual selection operating on the call parameters that we examined was stabilizing, it is important to note that the presence of stabilizing selection has not precluded the detection of linear sexual selection. There was significant directional selection on intercall duration and on three of the five eigenvectors (\mathbf{m}_3 – \mathbf{m}_5), including the two eigenvectors on which there was significant nonlinear selection. Intercall duration was an important component of each of these eigenvectors, loading in the direction opposite to the direction of linear selection in each case. Therefore, the significant linear selection gradients associated with each of these eigenvectors may be due largely to the significant directional selection toward shorter intercall intervals.

Male advertisement traits are often under strong directional sexual selection (Andersson 1994). Directional selection favoring shorter intercall intervals may be adaptive in *T. commodus* because ICD is not only an important component of call structure but also contributes to a male's calling effort. The effort a male invests in calling is a composite measure and includes call intensity (i.e., sound pressure level), bout length, the number of calling bouts in a night, and the duty cycle of each call. That these traits are consistently preferred by females in a variety of acoustic species (Ryan and Keddy-Hector 1992; Welch et al. 1998; Castellano et al. 2000; Sim-

mons et al. 2001; Klappert and Reinhold 2003; Holzer et al. 2003) is not surprising, because such call attributes are likely to be the basis for condition-dependent signals of male quality (Welch et al. 1998; Holzer et al. 2003; Hunt et al. 2004). Indeed, playback trials conducted in the field at Smith's Lake, in which we manipulated the number of calls performed in a night in addition to the same five call parameters used in this study, suggest that a male's nightly calling effort is under strong directional selection (Hunt et al. 2004).

However, at least some of the directional selection on intercall duration reported here is likely to be an artifact of the method we used to construct our artificial calls. The 15 field-collected males whose calls we used to estimate the range of phenotypes used in this study were a small sample. Four of them had a bimodal distribution of intercall duration estimates, in which several of the sampled calls were separated by intervals of more than one second. Such distributions are typical of males calling at the start or end of a bout and atypical of a male calling in the middle of a bout. In our call recordings of the four independent samples of males, ICD rarely exceeded 300 milliseconds in duration. Therefore, it is conceivable that the mean of the artificial calls that we used in this selection analysis may have been biased away from the actual population mean for males in midbout. If there is indeed such a bias, then some of the estimated linear selection in our analyses may be an artifact of how we designed the artificial calls rather than a representation of directional selection occurring in the field.

Phenotypic Manipulation and the Fitness Surface

Although the prediction that populations should evolve toward fitness peaks is a powerful heuristic image, demonstrating that the population mean is at a peak in the fitness surface is complicated by the presence of strong correlations between traits that may constrain the range of possible phenotypes in multivariate space. Such strong correlations not only bias selection analyses, but also may prevent populations from reaching adaptive peaks that comprise untenable trait combinations as a consequence of genetic correlations (Mitchell-Olds and Shaw 1987) or biophysical constraints (Ryan and Rand 2003). For example, in the guppy, *Poecilia reticulata*, genetic correlations among different components of male ornamentation may prevent populations converging on one of the three peaks identified in a multivariate fitness surface describing attractiveness (Blows et al. 2003).

The ease with which acoustic signals can be manipulated using computer call construction has allowed us and others

(Ryan and Rand 2003) to further demonstrate the powerful inferences about selection that can be gained from manipulative experiments. By artificially constructing calls with random combinations of the five manipulated call parameters, our selection analysis was freed from the usual constraints imposed by correlations among trait values (see Mitchell-Olds and Shaw 1987; Arnold et al. 2001). The fitness peak was within the range of natural, unmanipulated phenotypes, indicating that in *T. commodus*, as in Túngara frogs (Ryan and Rand 2003), biophysical constraints on the types of call a male can make do not prevent the population from converging on the most attractive phenotypic combination. This is only likely to be the case when traits are under stabilizing selection: on surfaces where there is strong directional or disruptive selection, evolutionary change may be constrained by biophysical and energetic constraints.

In conclusion, stabilizing selection and the associated heuristic imagery of a fitness peak on an adaptive landscape are fundamental to how evolutionary biologists attempt to understand evolution in natural populations. It is important to note that the common approach of analyzing and interpreting univariate and bivariate selection gradients will often be inadequate to establish the form and strength of selection acting on a set of functionally related traits. We have shown how it is possible to statistically define a multivariate fitness peak that successfully predicts the position of the mean of the natural population within the 95% confidence region of the fitness peak. Although a multivariate stabilizing peak was found in this system, canonical analysis of the few multivariate fitness surfaces available in the database of Kingsolver et al. (2001) that report correlational selection gradients suggests that such multivariate stabilizing peaks may be uncommon (Blows and Brooks 2003). The presence of multivariate stabilizing selection therefore remains an important but relatively unexplored aspect of selection in natural populations.

ACKNOWLEDGMENTS

This research was supported by the Australian Research Council. LFB was supported by NSERC (Canada). Many thanks to F. Zajitschek, A. Lindholm, S. Arnold, B. Walsh, and T. Tregenza for helpful comments on the manuscript and A. Gaskett for assistance with playback trials.

LITERATURE CITED

- Andersson, M. 1994. Sexual selection. Princeton Univ. Press, Princeton, NJ.
- Armbruster, W. S. 1990. Estimating and testing the shapes of adaptive surfaces: the morphology and pollination of *Dalechampia* blossoms. *Am. Nat.* 135:14–31.
- Arnold, S. J. 2003. Performance surfaces and adaptive landscapes. *Integr. Comp. Biol.* 43:367–375.
- Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macro-evolution. *Genetica* 112–113:9–32.
- Barton, N. and P. D. Keightley. 2002. Understanding quantitative genetic variation. *Nat. Rev. Genet.* 3:11–20.
- Benkman, C. W. 1993. Adaptation to single resources and the evolution of crossbill (*Loxia*) diversity. *Ecol. Monogr.* 63:305–325.
- . 2003. Divergent selection drives the adaptive radiation of crossbills. *Evolution* 57:1176–1181.
- Bentley, D. R., and R. R. Hoy. 1972. Genetic control of the neural network generating cricket (*Teleogryllus*: *Gryllus*) song patterns. *Anim. Behav.* 20:478–492.
- Bisgaard, S., and B. Ankenman. 1996. Standard errors for the eigenvalues in second-order response surface models. *Technometrics* 38:238–246.
- Blows, M. W., and R. Brooks. 2003. Measuring non-linear selection. *Am. Nat.* 162:815–820.
- Blows, M. W., R. Brooks, and P. G. Kraft. 2003. Exploring complex fitness surfaces: multiple ornamentation and polymorphism in male guppies. *Evolution* 57:1622–1630.
- Blows, M. W., S. F. Chenoweth, and E. Hine. 2004. Orientation of the genetic variance-covariance matrix and the fitness surface for multiple male sexually selected traits. *Am. Nat.* 163:329–340.
- Box, G. E. P., and J. S. Hunter. 1954. A confidence region for the solution of a set of simultaneous equations with an application to experimental design. *Biometrika* 41:190–199.
- Brodie, E. D., III, and F. J. Janzen. 1996. On the assignment of fitness values in statistical analyses of selection. *Evolution* 50:437–442.
- Bürger, R., and R. Lande. 1994. On the distribution of the mean and the variance of a quantitative trait under mutation-selection-drift balance. *Genetics* 138:901–912.
- Butlin, R. K., G. M. Hewitt, and S. F. Webb. 1985. Sexual selection for intermediate optimum in *Chorthippus brunneus* (Orthoptera: Acrididae). *Anim. Behav.* 1985:1281–1292.
- Castellano, S., B. Cuatros, R. Rinella, A. Rosso, and C. Giacomini. 2002. The advertisement call of the European treefrogs (*Hyla arborea*): a multilevel study of variation. *Ethology* 108:75–89.
- Castellano, S., A. Ross, F. Laoretti, S. Doglio, and C. Giacomini. 2000. Call intensity and female preferences in the European green toad. *Ethology* 106:1129–1141.
- Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* 110:155–171.
- Cresswell, J. E. 2000. Manipulation of female architecture in flowers reveals a narrow optimum for pollen deposition. *Ecology* 81:3244–3249.
- del Castillo, E., and S. Cahya. 2001. A tool for computing confidence regions on the stationary point of a response surface. *Am. Stat.* 55:358–365.
- Draper, N., and H. Smith. 1981. Applied regression analysis. 2nd ed. John Wiley and Sons, New York.
- Dyson, M. L., S. P. Henzi, and N. I. Passmore. 1994. The effect of changes in the relative timing of signals during female phonotaxis in the reed frog *Hyperolius marmoratus*. *Anim. Behav.* 48:679–685.
- Dyson, M. L., and N. I. Passmore. 1988. The combined effect of intensity and the temporal relationship of stimuli on phonotaxis in female painted reed frogs. *Anim. Behav.* 36:1555–1556.
- Fear, K. K., and T. Price. 1998. The adaptive surface in ecology. *Oikos* 82:440–448.
- Gavrilets, S. 2004. Fitness landscapes and the origin of species. Princeton Univ. Press, Princeton, NJ.
- Gerhardt, H. C. 1991. Female mate choice in treefrogs: static and dynamic acoustic criteria. *Anim. Behav.* 42:615–635.
- Green, P. J., and B. W. Silverman. 1994. Nonparametric regression and generalised linear models. Chapman and Hall, London.
- Greenfield, M. D. 2002. Signalers and receivers: mechanisms and evolution of arthropod communication. Oxford Univ. Press, Oxford, U.K.
- Greenfield, M. D., and I. Roizen. 1993. Katydid synchronous chorusing is an evolutionary stable outcome of female choice. *Nature* 364:618–620.
- Hennig, R. M., and T. Weber. 1997. Filtering of temporal parameters of the calling song by cricket females of two closely related species: a behavioural analysis. *J. Comp. Physiol. A* 180:621–630.
- Hersch, E. I., and P. C. Phillips. 2004. Power and potential bias in field studies of natural selection. *Evolution* 58:479–485.
- Hill, K. G., J. J. Loftus-Hills, and D. F. Gartside. 1972. Pre-mating isolation between the Australian field crickets *Teleogryllus com-*

- modus* and *T. oceanicus* (Orthoptera: Gyllidae). *Aust. J. Zool.* 20:153–163.
- Holzer, B., A. Jacot, and M. W. G. Brinkhof. 2003. Condition-dependent signaling affects male sexual attractiveness in field crickets, *Gryllus campestris*. *Behav. Ecol.* 353:359.
- Houde, A. E. 1993. Evolution by sexual selection: What can population comparisons tell us? *Am. Nat.* 141:796–803.
- Hunt, J., R. Brooks, M. J. Smith, M. D. Jennions, C. L. Bensten, and L. F. Bussière. 2004. High quality male field crickets invest heavily in sexual display but die young. *Nature* 432:1024–1027.
- Jang, Y., and M. D. Greenfield. 1998. Absolute versus relative measurements of sexual selection: assessing the contributions of ultrasonic signal characters to male attraction in lesser wax moths, *Achroia grisella* (Lepidoptera: Pyralidae). *Evolution* 52:1383–1393.
- Kingsolver, J. G., H. E. Hoekstra, J. M. Hoekstra, D. Berrigan, S. N. Vignieri, C. E. Hill, A. Hoang, P. Gibert, and P. Beerli. 2001. The strength of phenotypic selection in natural populations. *Am. Nat.* 157:245–261.
- Klappert, K., and K. Reinhold. 2003. Acoustic preference functions and sexual selection on the male calling song in the grasshopper *Chorthippus biguttulus*. *Anim. Behav.* 65:225–233.
- Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314–334.
- . 1979. Quantitative genetical analysis of multivariate evolution, applied to brain:body size allometry. *Evolution* 33:402–416.
- . 1980. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* 94:203–215.
- Lande, R., and S. J. Arnold. 1983. The measurement of selection on correlated characters. *Evolution* 37:1210–1226.
- Mead, L. S., and S. J. Arnold. 2004. Quantitative genetic models of sexual selection. *Trends Ecol. Evol.* 19:264–271.
- Mitchell-Olds, T., and R. G. Shaw. 1987. Regression analysis of natural selection: statistical inference and biological interpretation. *Evolution* 41:1149–1161.
- 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.
- Paterson, H. E. H. 1982. Perspective on speciation by reinforcement. *S. Afr. J. Sci.* 78:53–57.
- . 1985. The recognition concept of species. Pp. 21–29. in E. Vrba, ed. *Species and speciation*. Transvaal Museum Monograph No. 4. Transvaal Museum, Pretoria.
- Peterson, J. J., S. Cahya, and E. del Castillo. 2002. A general approach to confidence regions for optimal factor levels of response surfaces. *Biometrics* 58:422–431.
- Phillips, P. C., and S. J. Arnold. 1989. Visualizing multivariate selection. *Evolution* 43:1209–1222.
- Polakow, D. A., P. R. Y. Backwell, N. Caithness, and M. D. Jennions. 1995. Stabilizing or directional selection in signalling systems: investigations in a population of painted reed frogs, *Hyperolius marmoratus*. *S. Afr. J. Sci.* 91:270–273.
- Ritchie, M. G. 1996. The shape of female mating preferences. *Proc. Natl. Acad. Sci. USA* 93:14628–14631.
- Ryan, M. J., and A. Keddy-Hector. 1992. Directional patterns of female mate choice and the role of sensory biases. *Am. Nat.* 139:S4–S35.
- Ryan, M. J., S. A. Perrill, and W. Wilczynski. 1992. Auditory tuning and call frequency predict population-based mating preferences in the cricket frog, *Acris crepitans*. *Am. Nat.* 139:1370–1383.
- Ryan, M. J., and A. S. Rand. 2003. Sexual selection in female perceptual space: how female Túngara frogs perceive and respond to complex population variation in acoustic mating signals. *Evolution* 57:2608–2618.
- Ryan, M. J., and W. Wilczynski. 1988. Coevolution of sender and receiver: effect on local mate preferences in cricket frogs. *Science* 240:1786–1788.
- . 1991. Evolution of intraspecific variation in the advertisement call of a cricket frog (*Acris crepitans*, Hylidae). *Biol. J. Linn. Soc.* 44:249–271.
- Scheuber, H., A. Jacot, and M. W. G. Brinkhof. 2003. Condition dependence of a multicomponent sexual signal in the field cricket *Gryllus campestris*. *Anim. Behav.* 65:721–727.
- Schluter, D. 2000. *The ecology of adaptive radiation*. Oxford Univ. Press, Oxford, U.K.
- Simmons, L. W., M. Zuk, and J. T. Rotenberry. 2001. Geographic variation in preference functions and male songs of the field cricket *Teleogryllus oceanicus*. *Evolution* 55:1386–1394.
- Simpson, G. G. 1944. *Tempo and mode in evolution*. Columbia Univ. Press, New York.
- . 1953. *The major features of evolution*. Columbia Univ. Press, Garden City, NY.
- Sinervo, B. 1998. Adaptation of maternal effects in the wild: path analysis of natural variation and experimental tests of causation. Pp. 288–306 in T. A. Mousseau and C. W. Fox, eds. *Maternal effects as adaptations*. Oxford Univ. Press, New York.
- Sinervo, B., and A. Basolo. 1996. Testing adaptation using phenotypic manipulation. Pp. 149–185 in M. R. Rose and G. V. Lauder, eds. *Adaptation*. Academic Press, London.
- Stumpner, A., and D. von Helversen. 2001. Evolution and function of auditory systems in insects. *Naturwissenschaften* 88:159–170.
- Travis, J. 1989. The role of optimizing selection in natural populations. *Annu. Rev. Ecol. Syst.* 20:279–296.
- 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* 280:1928–1930.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:159.
- Zhang, X.-S., and W. G. Hill. 2003. Multivariate stabilizing selection and pleiotropy in the maintenance of quantitative genetic variation. *Evolution* 57:1761–1775.
- Zhang, X.-S., J. Wang, and W. G. Hill. 2002. Pleiotropic model of maintenance of quantitative genetic variation at mutation-selection balance. *Genetics* 161:419–433.

Corresponding Editor: T. Tregenza