

Limited plasticity in the phenotypic variance-covariance matrix for male advertisement calls in the black field cricket, *Teleogryllus commodus*

W. R. PITCHERS*§, R. BROOKS†, M. D. JENNIONS‡, T. TREGENZA§, I. DWORKIN* & J. HUNT§

*Department of Zoology, Program in Ecology Evolutionary Biology and Behavior, Michigan State University, East Lansing, MI, USA

†School of Biological, Earth and Environmental Sciences, The University of New South Wales, Kensington, Sydney, NSW, Australia

‡Evolution, Ecology & Genetics, Research School of Biology, Australian National University, Canberra, ACT, Australia

§Centre for Ecology & Conservation, University of Exeter, Cornwall Campus, Penryn, UK

Keywords:

advertisement call;
common garden;
diet;
field cricket;
matrix comparison;
P matrix;
phenotypic integration;
phenotypic plasticity;
sexual selection.

Abstract

Phenotypic integration and plasticity are central to our understanding of how complex phenotypic traits evolve. Evolutionary change in complex quantitative traits can be predicted using the multivariate breeders' equation, but such predictions are only accurate if the matrices involved are stable over evolutionary time. Recent study, however, suggests that these matrices are temporally plastic, spatially variable and themselves evolvable. The data available on phenotypic variance-covariance matrix (**P**) stability are sparse, and largely focused on morphological traits. Here, we compared **P** for the structure of the complex sexual advertisement call of six divergent allopatric populations of the Australian black field cricket, *Teleogryllus commodus*. We measured a subset of calls from wild-caught crickets from each of the populations and then a second subset after rearing crickets under common-garden conditions for three generations. In a second experiment, crickets from each population were reared in the laboratory on high- and low-nutrient diets and their calls recorded. In both experiments, we estimated **P** for call traits and used multiple methods to compare them statistically (Flury hierarchy, geometric subspace comparisons and random skewers). Despite considerable variation in means and variances of individual call traits, the structure of **P** was largely conserved among populations, across generations and between our rearing diets. Our finding that **P** remains largely stable, among populations and between environmental conditions, suggests that selection has preserved the structure of call traits in order that they can function as an integrated unit.

Introduction

Phenotypic traits seldom evolve independently and responses to selection are often correlated due to shared genetic effects such as pleiotropy and linkage (Lande & Arnold, 1983). The phenotypes expressed in a population may further covary because traits arise from shared developmental processes or are influenced by the same

environmental factors (West-Eberhard, 2003). The resulting nonindependence of traits is generally referred to as 'phenotypic integration' (e.g. Zelditch *et al.*, 1992; Cheverud, 1995; Olson & Miller, 1999; Merilä & Björklund, 1999; Pigliucci & Kolodnynska, 2002b) and can be quantified as the pattern of phenotypic variance in and covariances between traits in a population.

Much of quantitative genetic theory on how complex traits evolve is based on the assumption that the phenotypic (**P**) and genetic (**G**) variance-covariance matrices will remain stable over evolutionary time (Lande, 1980; Turelli, 1988), or vary only proportionally (Roff *et al.*, 2000; Stirling & Roff, 2000). However, there is

Correspondence: William R. Pitches, Department of Zoology, Program in Ecology Evolutionary Biology and Behavior, Michigan State University, East Lansing MI 48824, USA. Tel.: +1 (517) 432-6730; fax: +1 (517) 432-2789; e-mail: pitches@msu.edu

now a substantial body of theoretical and empirical work suggesting that these matrices may vary (Roff & Mousseau, 1999; Agrawal *et al.*, 2001; Eroukhanoff, 2009; de Oliveira *et al.*, 2009), and that they can themselves evolve, both in response to selection (Blows & Higgie, 2003), as a result of mutation (Camara *et al.*, 2000) and through random genetic drift (Phillips *et al.*, 2001). Thus, it is important to study empirically the extent to which **P** varies, and to quantify the scale over which this variation occurs (Arnold & Phillips, 1999; Ackermann & Cheverud, 2000; Bégin *et al.*, 2004; Roff *et al.*, 2004; Game & Caley, 2006; de Oliveira *et al.*, 2009).

The issue of the stability of **P** is of importance to evolutionary biology because the response of complex traits to selection is modelled by the multivariate extension of the breeders' equation: $\Delta\bar{z} = \mathbf{GP}^{-1}\mathbf{s}$, where $\Delta\bar{z}$ is the vector of change in trait means, and \mathbf{s} is the vector of selection differentials (Lande, 1979). Following the breeders' equation, **P** is expected to play an important role in the trajectory of phenotypic evolution, since selection (\mathbf{s}) may only act on trait combinations that are expressed (**P**). However, most organisms encounter heterogeneity in their environment and phenotypes are often plastic in their expression. As **P** has an environmental component, an understanding of the plasticity of **P** is required for prediction of the capacity for complex phenotypes to evolve in response to selection (Lande & Arnold, 1983; Schluter, 2001).

Sexually selected signals are expected to show a high degree of plasticity if their expression is to accurately reflect an individuals' condition and therefore communicate information about mate quality to the choosing sex (Zahavi, 1975; Andersson, 1982). However, there are conflicting predictions about the level of integration between individual traits that comprise a complex sexual signal, as well as between the sexual signal and other traits expressed by the organism (Badyaev, 2004). On one hand, if sexual signals function as an honest indicator of male quality or general vigour, we might predict strong integration between signals and the traits that are advertised by those signals (Wedekind, 1992; Johnstone, 1995; Bischoff *et al.*, 2009). On the other, when selection favours the expression of sexual signals to become more exaggerated, we might predict that this should minimize the integration between sexual signals and other elements of the phenotype, as this would allow the sexual signal to evolve independently of traits that are not subject to sexual selection (Emlen & Nijhout, 2000).

From the state of current research it is not clear at what taxonomic level we ought to predict divergence in **P**. Empirical work has shown stability of **P** among populations of garter snakes (Phillips & Arnold, 1999) and among populations of a number of different species of coral reef fishes (Game & Caley, 2006) and New World monkeys (de Oliveira *et al.*, 2009). Conversely,

studies have shown **P** to vary between populations of house finches (Badyaev & Hill, 2000), and there is some evidence for stability of **P** at the species level in tamarins (Ackermann & Cheverud, 2000). The relationship between integration and plasticity has also been investigated, most prominently in the morphology of *Arabidopsis* (Pigliucci & Kolodynska, 2002a,b; Kolodynska & Pigliucci, 2003; Bossdorf & Pigliucci, 2009). In addition to differences in **P** that have been documented across populations of *Arabidopsis* (Pigliucci & Kolodynska, 2002a; Bossdorf & Pigliucci, 2009), a number of studies have also examined patterns of phenotypic integration under experimentally induced stress. Despite finding plastic responses (in terms of trait means and variances) associated with changes in soil moisture content (Pigliucci & Kolodynska, 2002a), light intensity (Pigliucci & Kolodynska, 2002b) and wind speed (Bossdorf & Pigliucci, 2009), these studies found that patterns of phenotypic covariance tended to remain stable. This study has been focused on morphology, and we know of no study examining the plasticity of integration in sexual signals. Such data are needed, given that the stability of (co)variance structures is of particular relevance to how complex sexual signals evolve (Badyaev, 2004).

The advertisement call produced by male black field crickets (*Teleogryllus commodus*) is a complex sexual signal. It commences with a single chirp sequence, which is followed by a variable number of trill sequences (Bentley & Hoy, 1972; Hill *et al.*, 1972). Female preference for this call structure has been studied, and females have been shown to be sensitive to both temporal (Pollack & Hoy, 1979) and spectral (Hennig & Weber, 1997) call properties. The selection imposed by these female preferences has been measured both in the laboratory (Brooks *et al.*, 2005) and in the field (Bentsen *et al.*, 2006), and the dominant form of selection on call structure was found to be multivariate stabilizing. If persistent, this is a selection regime that would be predicted to facilitate trait integration (Lande, 1980; Cheverud, 1984; McGlothlin *et al.*, 2005).

Here, we present estimates of **P** for the structure of the advertisement call of *T. commodus* from 6 geographically isolated populations across the southern distribution of this species in Australia. We demonstrate that the advertisement calls vary among populations, and compare **P** estimated from call recordings of wild-caught males and from males reared under common-garden conditions. Using a suite of statistical analyses (Flury hierarchy, geometric subspace comparisons and random skewers), we evaluate the stability of **P** among populations and between rearing environments. The sexual advertisement calls of some crickets are also known to show phenotypic plasticity in response to environmental conditions, most famously relating to temperature (e.g. Elliott & Koch, 1985; Olvido & Mousseau, 1995). In addition, call plasticity has been

measured in a number of cricket species in response to diet, both in terms of calling effort (Wagner & Hoback, 1999; Hunt *et al.*, 2004; Hedrick, 2005; Judge *et al.*, 2008; Zajitschek *et al.*, 2009) and in the structure of the call (Scheuber *et al.*, 2003a,b), though this result is not ubiquitous (Gray & Eckhardt, 2001; Hartbauer *et al.*, 2006). In a second experiment, we reared crickets from the same six populations on high- and low-nutrient diets. We then employ the same suite of approaches to test for plasticity of **P** for the advertisement call of this species across diets. In this study, we therefore compare the stability of **P** between allopatric populations, across generations and in response to different rearing environments.

Materials and methods

Experiment 1. Collection and common-garden rearing of crickets from six populations

Approximately, 400 adult *T. commodus* were collected from each of six populations spanning the southern distribution of this species in Australia: Western Australia (WA), South Australia (SA), Tasmania (TAS), Kioloa (KL), Australian Capital Territory (ACT) and Smith's Lake (SL) (Fig. 1). These populations are widely distributed across both latitude and longitude in southern Australia, and the abiotic (e.g. rainfall, temperature) and biotic (e.g. species composition and amount of vegetation) environmental conditions they experience (Pitchers, personal observation). However, these populations do not vary according to a simple environmental cline. Crickets were collected between January and April in 2004 and air-freighted back to the University of New South Wales. In the laboratory, each population was established in a separate 80L culture container and provided with Go-Cat Senior cat pellets as food (Purina, Croydon, UK) and water *ad libitum*, and an abundance of egg cartons for shelter. Cultures were maintained in a constant temperature room at $28 \pm 1^\circ\text{C}$ with a 14L : 10D light regime. A random sample of males from each population was recorded within three nights of being established in the laboratory. As this sample of males had completed their development and reached sexual maturity in the field, we refer to these as *field* males.

In each generation, populations were maintained by randomly pairing 100 males each with a virgin female. Each pair was placed in an individual clear plastic container ($5 \times 5 \times 5$ cm) for 3 days to mate. Males were then removed and females provided with food, water and a 4-cm Petri-dish of moist cotton wool for a week to lay eggs. At hatching, approximately 25 nymphs were randomly selected from each mating pair and reared in a small clear plastic container ($15 \times 10 \times 15$ cm) until they reached 4th instar. Nymphs were then sexed and five crickets of each sex

per container were randomly selected to contribute to the next generation. These crickets were reared to eclosion in a single, large plastic container (80L) for each population. Containers were checked daily and adults removed and maintained in sex-specific containers (80L) until sexually mature. When sexually mature, these crickets were used to propagate the next generation following the protocol outlined above.

These populations were maintained in this fashion under common environmental conditions for three generations prior to measuring the calls of a random sample from each population for a second time. As this sample of males has been maintained in the laboratory for three generations, we refer to them as laboratory males. Even in the absence of genetic differences among populations, maternal effects are known to induce adaptive plastic responses that can resemble local adaptation (Agrawal *et al.*, 1999), but after three generations of rearing under lab-standard conditions any such effects are predicted to be reduced to a negligible level (Roach & Wulff, 1987).

Experiment 2. The plasticity of **P** under dietary manipulation

Between February and March, 2007, we collected approximately 400 adult crickets from each of the same six populations outlined above. Crickets were air-freighted to the Australian National University in Canberra, where they were maintained in 80L culture containers and provided with moist cotton wool to oviposit. These egg-pads were then air-freighted to the University of Exeter in Cornwall (UK) to establish the new cultures for each population. The laboratory populations derived from these egg-pads were maintained in common-garden conditions following the same protocol outlined above for five generations before we began our diet manipulation study.

In each population, a total of 300 nymphs on their day of hatching were established in individual plastic containers ($5 \times 5 \times 5$ cm) and provided with a small water bottle (small plastic vial, 4 cm long \times 1 cm diameter plugged with cotton wool) and a piece of cardboard egg carton for shelter. Half of the nymphs per population were provided with a high-nutrient diet and the remaining half with a low-nutrient diet ($n = 900$ nymphs per diet, total $n = 1800$ nymphs). The high-nutrient diet consisted of Go-Cat Senior cat pellets (Purina) that were used to maintain our stock cultures, whereas the low-nutrient diet consisted of a 50 : 50 mixture (by dry weight) of ground cat pellets and ground oats. Cat food and oats were ground to a fine powder using a food processor (BL10450H model, Black and Decker, London, UK) and sieved to remove larger particles. Each diet was then mixed with a small quantity of water and stirred to make a paste, which was then spread across a 1 cm thick rigid polymer sheet

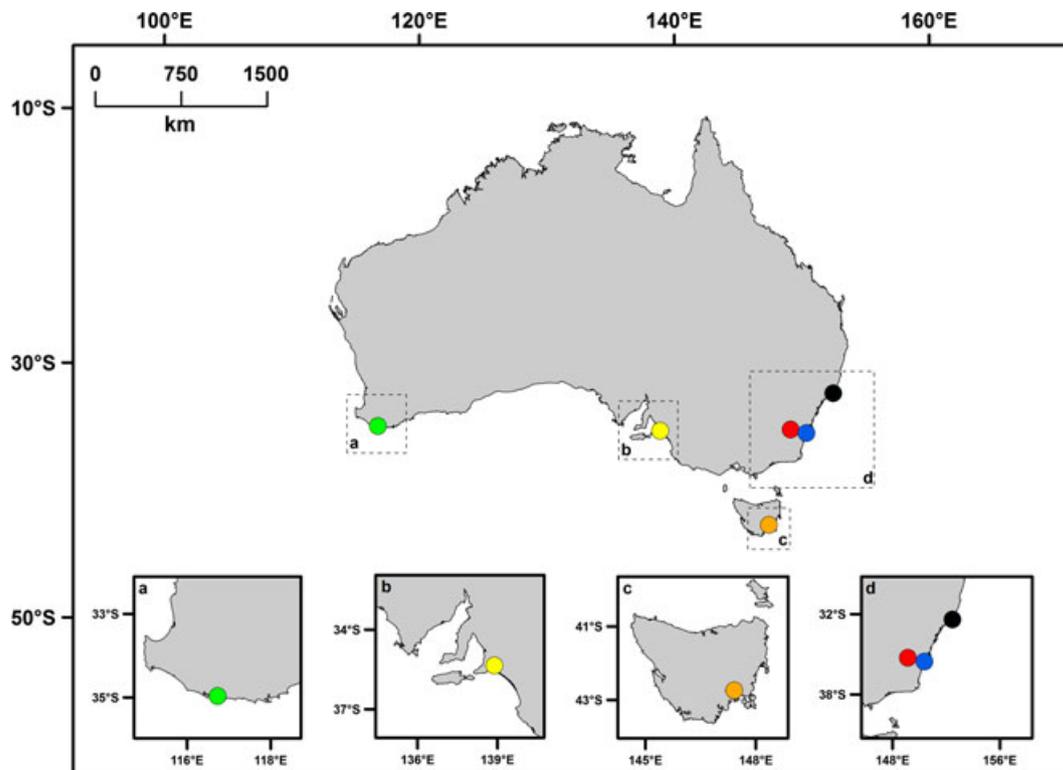


Fig. 1 The location of the six populations that crickets were collected. Populations are colour coded: Western Australia (WA, green), South Australia (SA, yellow), Tasmania (TAS, orange), Australian Capital Territory (ACT, red), Kioloa (KL, blue) and Smith's Lakes (SL, black).

perforated with 4 cm holes. After drying for 24 h at 30 °C, the dried diet was pushed out from each hole in pellet form. Pellets were stored in sealed containers at room temperature. Both diets were presented as powder (in the lid of a 1.5-mL centrifuge tube) for the first 4 weeks of feeding because small nymphs can have difficulty feeding on the larger food pellets. Food and water were replaced and containers cleaned weekly. Upon reaching final instar, crickets were checked daily for eclosion to adulthood to determine development time. Since some aspects of some cricket calls have been found to covary with body size (e.g. Bennet-Clark, 1999), we measured the body weight of each cricket at eclosion to adulthood on an analytic balance (Mettler-Toledo, 345G). We also measured the width of the pronotum, another commonly used measure of size in crickets, using an eyepiece graticule on a binocular dissecting microscope (Leica, MZ5). All crickets were maintained in a constant temperature room set to 28 ± 1 °C with a 14L : 10D light regime.

Call recording and analysis

At 10 days of age post-eclosion, all males were moved to a separate constant temperature room set to 28 ± 1 °C to record their advertisement call. Each male remained housed in their original plastic container that

they were reared in through development, but the lid of the container was replaced with one that had a small condenser microphone (C1163, Dick Smith, Chullora, Australia) mounted onto its underside. Prior to recording, each container housing a male was placed inside a foam box (15 × 15 × 15 cm) to minimize interference from other calling males when recording. The calls of field and laboratory males in Experiment 1 were recorded at the University of New South Wales using a Sony Professional Walkman (WM-D6C) on high-quality chrome tapes (BASF CEII, Germany). Calls were recorded opportunistically between 21:00 and 03:00 each night by scanning recording chambers hourly. The recordings were then digitized for measurement and analysis. In total, we measured the calls of 142 males (WA = 20, SA = 25, TAS = 17, ACT = 25, KL = 25, SL = 30) collected from the field and 173 males (WA = 25, SA = 35, TAS = 32, ACT = 26, KL = 28, SL = 27) after three generations of common-garden rearing in the laboratory.

The calls of the high- and low-nutrient males reared under common-garden conditions at the University of Exeter were recorded using an automated multi-channel digital recorder using an NI-DAQ digital acquisition interface running at 48 K samples-s (NI USB-6009, National Instruments, www.uk.ni.com) controlled by

software written in LabView (National Instruments, www.uk.ni.com) using the full sampling speed of the interface to record at 48 KHz. This system automatically monitored up to 256 channels and made 1-minute digital recordings only when the sound amplitude on a channel exceeds a user-defined threshold. In total, we measured the calls of 174 males (WA = 34, SA = 27, TAS = 28, ACT = 28, KL = 25, SL = 31) reared on the high-nutrient diet and 183 males (WA = 31, SA = 26, TAS = 30, ACT = 31, KL = 27, SL = 37) reared on the low-nutrient diet.

All calls were measured using 'Raven' software version 1.1 (Bioacoustics Research Group: Cornell Lab of Ornithology, Ithaca, NY, USA). We measured the following properties of six randomly selected calls per male: the number of pulses in the chirp (CPN, chirp pulse number), the interval between pulses in the chirp (CIPD, chirp inter-pulse duration (ms)), the number of trills in the call (TN, trill number), the interval between the last trill pulse of one call and the first chirp pulse of the next call (ICD, inter-call duration (ms)) and the dominant frequency of the call (DF, dominant frequency (mHz)) (Fig. 2). We used the mean of these call properties for each male in all subsequent analyses.

Statistical analysis

Unless otherwise stated, analyses were conducted in **R** (version 2.13.0). To determine if the structure of the advertisement call had genetically diverged across our populations, we analysed our call data using a multivariate analysis of covariance (MANCOVA) using the 'Manova' function from the 'car' package. We used a model including population, generation (i.e. field and common-garden males) and their interaction as fixed effects and the 5 call traits as the response variables.

Male size (measured as the maximum width of the pronotum) differed significantly across populations (ANOVA: $F_{5,303} = 4.61$, $P < 0.001$), between rearing environments ($F_{1,303} = 221.16$, $P < 0.001$), and there was a significant interaction between population and rearing environment ($F_{5,303} = 15.19$, $P < 0.001$). As body size is known to influence the structure of the advertisement call in other field cricket species (e.g. Simmons, 1995; Simmons & Zuk, 1992), we included male size as a covariate in our MANCOVA of call traits. Our model was therefore: $\mathbf{Y} = \mu + \beta_p + \beta_r + \beta_{p \times r} + \beta_s + \epsilon$, where \mathbf{Y} is the matrix of call variables, β_p is the vector of coefficients for population, β_r is that for rearing environment (field or lab), $\beta_{p \times r}$ is that for the interaction of population and rearing and β_s is that vector of coefficients for size. The model was the same for the second experiment, with diet treatment in place of rearing environment. All call structure parameters, as well as male size, were log transformed prior to analysis. Some of the call data from Experiment 2 were not normally distributed and could not be transformed to normal, so we used a permutation test to determine the significance of terms in our model. Briefly, for each of 10 000 iterations we resampled \mathbf{Y} (without replacement), ran a MANCOVA and retained the P -values. These P -values then give us an expected distribution under the null hypothesis, and the permuted P -value is obtained as the proportion of iterations wherein the null P is smaller than the P -value obtained from a MANCOVA of the original data.

We visualized the results of our MANCOVA's by using linear discriminant analysis ('lda' function from the 'MASS' package), to find the rotation that best separated the call structure data by both population and generation or diet. We then calculated means and standard errors for each population/rearing environment

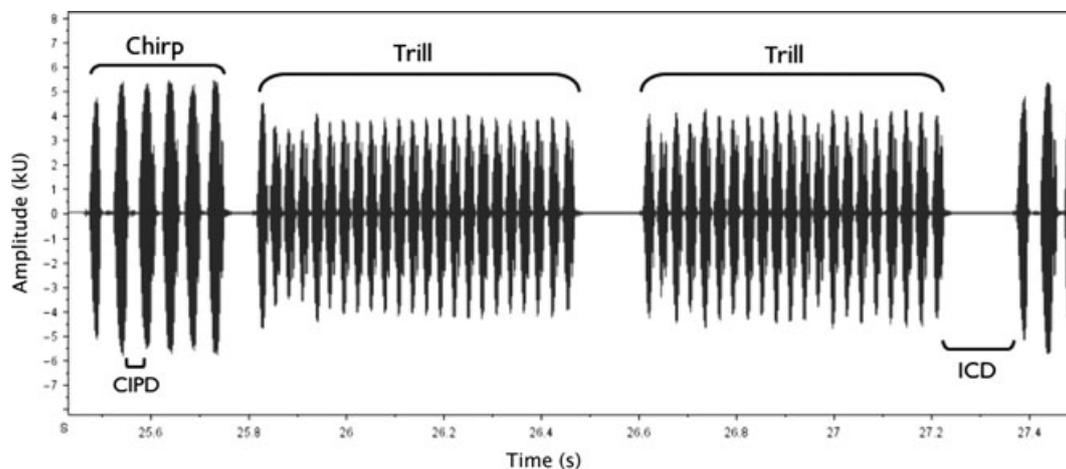


Fig. 2 Screen capture from 'Raven' analysis software (v1.1), with annotations to show the main features of the call. Measures recorded from each call were the number of pulses in the chirp (CPN), the interval between pulses in the chirp (CIPD), the number of trills in the call (TN), the interval between the last trill and the next call (ICD) and the dominant frequency of the call (DF – not shown).

combination. To determine which call traits were driving the patterns revealed by these multivariate analyses, we also ran *post hoc* ANOVA's for each trait. The right-hand side of each *post hoc* model was the same as the corresponding multivariate model. For the data from Experiment 2, we used the same permutation approach as above to assign *P*-values, and *post hoc* tests were also run for 10 000 permutations. For Experiment 2, we also ran ANOVA's to test for effects of population, diet and population \times diet interaction on development time (number of days from hatching to eclosion), weight and pronotum width (measured at eclosion).

We calculated **P** matrices for the five call traits of each population and rearing environment in Experiment 1 and for each population and diet in Experiment 2, using the 'cov' function from the 'stats' package. With six populations and two treatments in each experiment, we had 12 call **P** matrices for each experiment. Although our focus was to compare the **P** matrices for call traits, we were also able to calculate a second set of **P** matrices for the life-history traits (development time, weight and pronotum width) that we measured in Experiment 2. There is currently no single best method of comparing matrices, with each available technique testing slightly different aspects of the matrices being compared (Roff *et al.*, 2012). We therefore used a suite of complementary methods to test whether **P** differed across populations and our treatments (generation and diet).

First, we used Common Principal Component (CPC) analysis (Phillips & Arnold, 1999). CPC analysis sequentially tests for differences between matrices in a hierarchical manner (Phillips & Arnold, 1999). We used the 'CPC' program (Phillips, 1998) to perform analyses and evaluated the results using the 'jump-up' approach; evaluating each model against the null model of unrelated matrices. CPC analysis is quite well-used in the field (e.g. Bégin & Roff, 2001; Game & Caley, 2006; Sherrard *et al.*, 2009), but is not without problems. In particular, CPC analysis compares all principal components between matrices, but does not take account of their relative eigenvalues, meaning that similarity may be inferred between principal components that have a small eigenvalue in one matrix but a large eigenvalue in another (Blows *et al.*, 2004).

We also compared the eigenstructure of the **P** matrices using geometric subspace analysis (Krzanowski, 1979). Krzanowski's geometric method compares *k*-dimensional subspaces, where *k* is the number of eigenvectors being compared in each sample, by calculating the angles between the most similar pairs of orthogonal components. To be meaningful, no more than half the eigenvectors of an *n*-dimensional matrix can be included in any comparison because to do so would constrain the analysis to recover common dimensions (Blows *et al.*, 2004). In the extreme case, including all eigenvectors will, by definition, recover

identical subspaces. In our analyses, we are therefore only able to include the first 2 eigenvectors (of our 5-dimensional matrices) in the subsets. However, these account for greater than 99% of the variation in the structure of the advertisement call. For a pair of matrices to be compared (in our case; two 5×5 (co)variance matrices) the subset of eigenvectors to be compared are designated as matrices **A** & **B** (in our case; two 5×2 matrices), and used to calculate a new matrix (**S**) as: $\mathbf{S} = \mathbf{A}^T \mathbf{B} \mathbf{B}^T \mathbf{A}$. The angle between most closely aligned vectors can then be found as $\cos^{-1} \sqrt{\lambda_1}$, where λ_1 is the largest eigenvalue of **S**. In addition, the sum of **S** matrix eigenvalues, which is equal to the sum of squares of the cosines of the angles between the two sets of orthogonal axes, can be used as a metric of similarity. This metric will have a possible range from 0 (orthogonal subspaces) to *k* (identical subspaces). Since our estimates of **P** contain may be influenced by sampling error we bootstrapped our Krzanowski tests (10 000 iterations) to place confidence intervals on the resulting statistics. Krzanowski comparisons were made using the 'krzanowski.test' function in the 'MCMCglmm' package of **R** (Hadfield, 2010).

As discussed above, one of the main reasons that **P** is of interest is their potential to influence a population's evolutionary response to selection. The 'random skewers' method developed by Cheverud (1996) is intuitive, in that it compares matrices by comparing the direction of 'response' to set of randomly drawn 'skewers'. These are sometimes described as hypothetical vectors of selection and so are represented as $\beta_{\mathbf{R}}$. The elements of the random skewers are randomly drawn from a uniform distribution from -1 to $+1$, and each vector is then standardized so that the sum of the squares of its elements is equal to 1 (Cheverud & Marroig, 2007). For each pair of matrices (**A** and **B**) and each random skewer, 'response' vectors are computed as $\mathbf{R}_{\mathbf{A}} = \mathbf{A} \beta_{\mathbf{R}}$ and $\mathbf{R}_{\mathbf{B}} = \mathbf{B} \beta_{\mathbf{R}}$, and the vector correlation between $\mathbf{R}_{\mathbf{A}}$ and $\mathbf{R}_{\mathbf{B}}$ are calculated following Calsbeek & Goodnight (2009). As we are testing **P** matrices, the skewers are not directly analogous to the vector of standardized selection gradients (β , *sensu* Lande & Arnold, 1983). We report the mean vector correlation from the application of 1000 random skewers as a measure of matrix similarity. We used a randomization approach to assess statistical differences between matrices (Roff *et al.*, 2000). After estimating a mean vector correlation for each pair of matrices, we randomized our data among populations to generate two new matrices and computed their vector correlation. This process was also iterated 1000 times. Following Calsbeek & Goodnight (2009) and Roff *et al.* (2000), we took the null hypothesis to be that matrices **A** and **B** are identical. *P*-values are therefore assigned as $P = n/R$, where *n* is the number of $r_{\text{rand}} \geq r_{\text{obs}}$ and *R* is the number of randomizations. Random skewers analyses were performed using **R** code adapted from Roff *et al.* (2000).

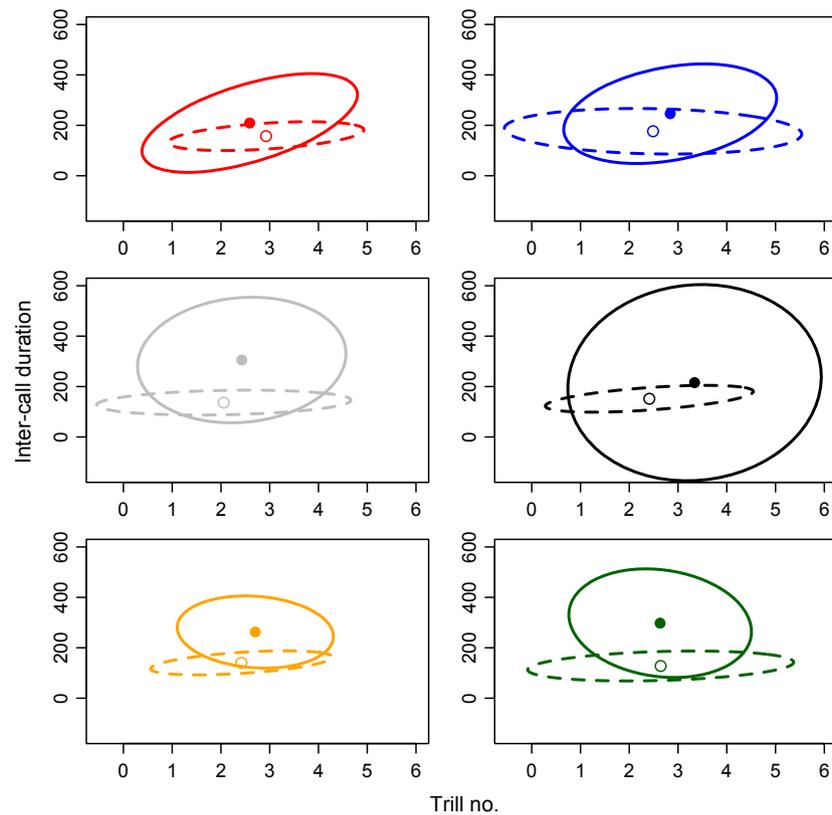


Fig. 3 Ellipses representing the first two dimensions of the call-trait \mathbf{P} matrices from experiment 1. The ellipses show the first two principal components projected onto the ICD and TN traits. These traits were chosen for illustration since they had the highest loadings for PC1 and PC2 respectively. The solid point and ellipse represent the call structure of the wild-caught males; the open point and the dashed ellipse represent the call structure of common-garden-reared males. Population colours and abbreviations are as shown for Fig. 1.

Lastly, we used the Bayesian ‘ d method’ of matrix comparison as described by Ovaskainen *et al.* (2008). Briefly, \mathbf{P} matrices were estimated from centred data using an MCMCglmm and compared pairwise. We calculated the divergence metric ‘ d ’ using the ‘Ddivergence’ function (‘MCMCglmm’ package for \mathbf{R}) at each position in the posterior distributions. The posterior mean and standard deviation from the distribution of d values provide an estimate of the magnitude of the difference between matrices in each pairwise comparison. We also calculated posterior distributions for Ovaskainen’s Ψ_d measure by resampling from the posterior distributions. Drawing two random samples from the posterior distribution for each matrix (\mathbf{P}^A and \mathbf{P}^B), we calculated the difference over 10 000 resampling iterations as:

$$\Psi_d(\mathbf{A}, \mathbf{B}) = [d(\mathbf{P}_1^A, \mathbf{P}_2^A) + d(\mathbf{P}_1^B, \mathbf{P}_2^B)] - [d(\mathbf{P}_1^A, \mathbf{P}_2^B) + d(\mathbf{P}_1^B, \mathbf{P}_2^A)]$$

The proportion of the posterior Ψ_d distribution where $\Psi_d(\mathbf{A}, \mathbf{B}) < 0$ can then be used to infer statistical support for difference between the two matrices.

To illustrate the covariance matrices in two dimensions, we extracted principal components from the data for the five call traits (i.e. populations, generation and diet treatments were pooled to give 1 set of PCs for each experiment) using the ‘prcomp’ function from the ‘stats’ package in \mathbf{R} . As the call traits are measured in a number of different units and therefore are quite different in magnitude, extractions were performed on correlation matrices. Inter-call duration and trill number were the two traits with the highest loadings on PC1 and PC2, respectively, and so PC1 and PC2 were projected onto these traits to draw 95% confidence ellipses for Figs 3 and 4.

Results

Experiment 1: Common-garden rearing of crickets from six populations

A MANCOVA (Table 1) revealed significant differences in the structure of the advertisement call in *T. commodus* both among populations and between rearing environments, and there was also a significant interaction

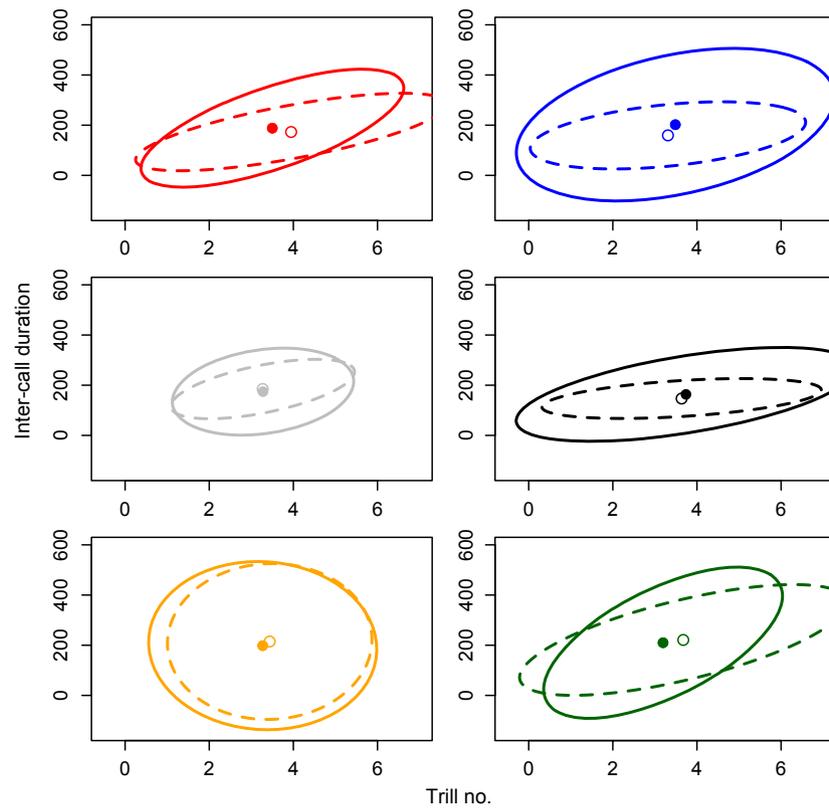


Fig. 4 Ellipses representing the first two dimensions of the call-trait **P** matrices from Experiment 2. The ellipses show the first two principal components projected onto the ICD and TN traits. These traits were chosen for illustration since they had the highest loadings for PC1 and PC2 respectively. The solid point and ellipse represent the call structure of the males reared on the high-nutrient diet; the open point and the dashed ellipse represent the call structure of males reared on the low-nutrient diet. Population colours and abbreviations are as shown for Fig. 1.

Table 1 MANCOVA statistics from experiment 1. Significant p-values are italicized.

Effect	Pillai	F approx.	hypoth. df	resid. df	P-value
Population (A)	0.38	5.01	25	1510	<0.0001
Rearing Environment (B)	0.66	117.86	5	298	<0.0001
A × B	0.58	7.90	25	1510	<0.0001
Size (PW)	0.05	3.14	5	298	<0.01
Trait Means (±SE) by Population					
Wild-Caught	CPN (#)	TN (#)	DF (MHz)	ICD (ms)	CIPD (ms)
ACT	6.43 (0.30)	2.59 (0.21)	3.87 (0.03)	209.29 (18.25)	26.67 (0.94)
KL	6.73 (0.25)	2.84 (0.20)	3.78 (0.03)	246.09 (18.40)	30.33 (1.19)
SA	6.73 (0.26)	2.43 (0.20)	3.66 (0.05)	305.33 (23.21)	29.28 (0.71)
SL	6.02 (0.20)	3.34 (0.22)	4.00 (0.04)	215.4 (33.11)	25.32 (0.83)
TAS	6.95 (0.21)	2.7 (0.18)	3.49 (0.04)	262.54 (16.19)	32.75 (1.43)
WA	8.15 (0.48)	2.63 (0.20)	3.89 (0.08)	297.56 (22.45)	37.4 (0.87)
Common-Garden	CPN (#)	TN (#)	DF (MHz)	ICD (ms)	CIPD (ms)
ACT	5.59 (0.14)	2.93 (0.18)	4.06 (0.02)	157.15 (5.13)	18.97 (0.75)
KL	6.08 (0.17)	2.49 (0.28)	4.04 (0.03)	176.44 (8.42)	21.5 (0.76)
SA	5.79 (0.15)	2.06 (0.23)	4.15 (0.03)	136.42 (4.34)	18.08 (0.61)
SL	5.77 (0.12)	2.41 (0.17)	3.97 (0.03)	151.52 (4.19)	18.1 (0.38)
TAS	5.79 (0.12)	2.42 (0.16)	4.08 (0.03)	140.17 (4.00)	18.13 (0.49)
WA	5.56 (0.14)	2.64 (0.24)	4.15 (0.03)	127.78 (5.32)	18.63 (0.57)
Post hoc P < 0.05	A, B, AxB	B, AxB	A, B, AxB	B, AxB	A, B, AxB, PW

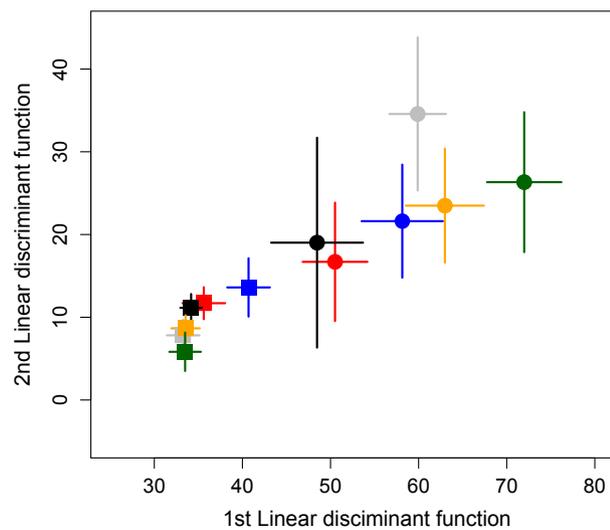


Fig. 5 Call structure variation between populations and between rearing environments in experiment 1. Population mean linear discriminant scores for the first and second linear discriminant functions of call phenotype calculated with population and rearing environment as grouping variables. Populations are colour coded as per the legend of Fig. 1. The round points are for calls measured from wild-collected males, whereas the square points represent common-garden males. The bars represent the standard error around the mean.

effect of population \times rearing environment. Male body size (pronotum width) was also a significant predictor of call traits and the inclusion of this term in our model ensures that the differences in the main effects are not driven solely by allometric relationships with body size. Among the crickets collected in the wild, males from the WA and TAS populations were most distinct from the other populations in call structure (Fig. 5). Differences between populations were dramatically reduced when the environmental component of these differences were minimized by laboratory rearing, while the effect of lab rearing differed between populations. The KL population was the most distinct under laboratory conditions. The maintenance of detectable differences across populations under common-garden rearing conditions is indicative that the divergence in advertisement calls has a heritable genetic basis as opposed to being the result of maternal effects. *Post hoc* ANCOVA's revealed a significant effect of population on CPN ($F_{5,302} = 3.51$, $P = 0.004$), DF ($F_{5,302} = 5.41$, $P = 0.0001$) and CIPD ($F_{5,302} = 13.47$, $P = 0.001$) and a significant effect of generation on all 5 call traits (CPN: $F_{1,302} = 66.63$, $P = 0.0001$; TN: $F_{1,302} = 5.87$, $P = 0.016$; DF: $F_{1,302} = 172.78$, $P = 0.0001$; ICD: $F_{1,302} = 126.03$, $P = 0.0001$; CIPD: $F_{1,302} = 602.48$, $P < 0.0001$). Body size, included in these analyses as pronotum width, only significantly influenced CIPD ($F_{1,302} = 13.30$, $P < 0.001$), but the interaction effect of popula-

tion \times generation was found to be significant for all five call traits (CPN: $F_{5,302} = 6.74$, $P = 0.0001$; TN: $F_{5,302} = 2.23$, $P = 0.05$; DF: $F_{5,302} = 20.19$, $P = 0.0001$; ICD: $F_{5,302} = 6.24$, $P = 0.0001$; CIPD: $F_{5,302} = 14.68$, $P = 0.0001$).

Comparing the covariance matrices for the wild-collected crickets from all six populations using CPC analysis found the most parsimonious model among the three approaches (step-up, model building and jump-up) was that for two common principal components ('CPC(2)'). This model implies that 2 of the 5 eigenvectors of each matrix are closely aligned among all six matrices (Phillips & Arnold, 1999). When we performed a similar comparison of matrices from the common-garden reared crickets, we found that the best fitting model was for three common principal components ('CPC(3)'). We also performed pairwise tests between rearing environments within each population, and again found the proportional model to be most parsimonious in the majority of cases. The only exception to this was the comparison of calls from field and lab-reared crickets from the ACT population, where the most parsimonious model was found to be that of 'common principal components'; the next model down Flury's hierarchy, also indicating differences confined to eigenvalues, but in this case the eigenvalues do not differ proportionately. These comparisons suggest that while these matrices have a conserved structure, they differ in the overall magnitude of variation they contain (Fig. 3). The comparative reduction in the amount of variation contained in \mathbf{P} after common-garden rearing suggests an important role for environmental variance in \mathbf{P} for wild-caught crickets.

We also compared \mathbf{P} from Experiment 1 using Krzanowski's geometric subspace comparison method (Krzanowski, 1979). This approach can only compare 2 matrices at a time and therefore all comparisons are pairwise (Table 2). Population comparisons of \mathbf{P} in the wild-collected males found all sums of $\mathbf{S}\lambda$ values to ≥ 0.56 (where the maximum possible value of two would indicate identical subspaces). Closest eigenvector angles were all $\leq 44.4^\circ$, indicating strong similarity of matrices. The results of \mathbf{P} comparisons among populations in the lab-reared males were somewhat more similar, with the sum of $\mathbf{S}\lambda$ values all being ≥ 0.89 and the closest eigenvector angles all being $\leq 18.8^\circ$. We found that \mathbf{P} matrices compared between field and lab-reared males within populations were of an intermediate level of similarity; with the sum of $\mathbf{S}\lambda$ values all being ≥ 0.80 and the closest eigenvector angles all being $\leq 27.9^\circ$. The consistency of these results and those of the common principal analyses indicate that, although not identical, \mathbf{P} is very similar in structure across populations and generations.

Matrix comparison by random skewers is also a pairwise procedure. The random skewers vector

correlations among populations in the field cohort were all ≥ 0.57 (mean \pm SE = 0.81 ± 0.06), and 9 of the 15 comparisons were found to be significantly different

from identical in randomization tests (Table 3). Among populations in the common-garden cohort random skewers vector correlations were all ≥ 0.44

Table 2 The results of geometric matrix comparisons among **P** matrices from the experiment 1. The upper matrix contains the sum of **S** matrix eigenvalues (a value between 0 & 2 in this case – 2 indicating that matrices span an identical subspace, with their 95% confidence interval below them). The lower *italicized* matrix contains the closest angles between principal vectors describing the shared subspace between matrices (in degrees), with 95% confidence intervals. Within each matrix, cells on the diagonal (**bold**) are within-population comparisons between wild-caught and common-garden calls. Cells below the diagonal are between-population comparisons of wild-caught calls – between-population comparisons of common-garden are calls above the diagonal. Population abbreviations are as for Fig. 1.

	ACT	KL	SA	SL	TA	WA	
ACT	0.92 0.41–1.36	0.74 0.29–1.29	1.06 0.55–1.85	1.66 0.90–1.91	1.31 0.62–1.74	0.72 0.18–1.34	Common-Garden Males
KL	1.00 0.69–1.59	0.78 0.20–1.33	1.02 0.34–1.63	0.57 0.13–1.14	1.16 0.42–1.65	1.22 0.27–1.68	
SA	1.24 0.60–1.71	1.01 0.66–1.76	1.62 0.62–1.90	0.97 0.46–1.91	1.05 0.37–1.77	0.99 0.12–1.40	
SL	1.47 0.87–1.81	1.15 0.76–1.61	0.98 0.67–1.63	0.87 0.35–1.36	1.55 0.67–1.83	0.99 0.22–1.55	
TA	0.79 0.26–1.56	0.63 0.14–1.30	1.16 0.40–1.80	0.65 0.21–1.51	0.69 0.26–1.38	1.53 0.48–1.84	
WA	0.65 0.26–1.30	0.93 0.61–1.49	0.55 0.19–1.29	1.01 0.31–1.54	1.26 0.29–1.60	0.91 0.02–1.42	
ACT	16.95 1.70–46.6	34.96 3.94–56.9	8.88 0.89–43.5	12.83 0.96–27.7	13.38 1.21–39.2	36.54 4.85–60.2	
KL	3.63 0.37–29.5	29.57 2.06–57.8	12.02 1.01–50.7	49.20 14.3–69.5	19.32 1.48–48.5	6.46 2.17–55.5	
SA	14.71 0.91–39.4	7.15 0.29–34.1	11.88 0.93–42.0	11.69 0.69–48.1	3.59 0.60–50.7	6.82 2.14–63.0	
SL	8.78 0.96–28.9	6.13 0.73–30.7	8.83 0.34–36.0	24.76 3.94–49.8	11.41 1.62–40.7	39.98 3.19–59.5	
TA	30.71 2.66–55.6	40.51 2.45–58.6	11.58 1.13–49.8	40.87 3.62–58.1	35.54 4.35–57.0	8.20 1.23–47.6	
WA	40.34 4.21–55.7	22.30 4.49–39.2	47.81 8.02–61.0	22.86 2.36–50.9	18.88 3.31–52.4	27.94 4.15–67.2	
	Wild-Caught Males						

Table 3 Random skewers correlations between and within populations in experiment 1. Correlations between populations in the wild-caught cohort are below the diagonal, and between the common-garden cohort are above the diagonal. Correlations along the diagonal are within population and between cohorts. Below each correlation is its associated p-value in italics, where a significant value indicates difference from equality.

	ACT	KL	SA	SL	TA	WA		
ACT	0.96 0.25	0.54 0.00	0.97 0.35	0.88 0.05	0.91 0.06	0.70 0.00	Common-Garden Males	
KL	0.77 0.01	0.89 0.04	0.58 0.00	0.69 0.00	0.47 0.00	0.63 0.00		
SA	0.94 0.19	0.74 0.00	0.96 0.28	0.90 0.10	0.94 0.17	0.70 0.00		
SL	0.82 0.01	0.96 0.38	0.71 0.00	0.90 0.12	0.71 0.00	0.92 0.26		
TA	0.93 0.13	0.76 0.00	0.98 0.64	0.72 0.00	0.98 0.51	0.44 0.00		
WA	0.66 0.00	0.94 0.23	0.57 0.00	0.94 0.22	0.64 0.00	0.83 0.03		
	Wild-Caught Males							

(mean \pm SE = 0.73 \pm 0.06), with 9 of 15 comparisons again being different from identical (Table 3). Within-populations between rearing environments, random skewer vector correlations were all \geq 0.83 (mean \pm SE = 0.92 \pm 0.03), but **P** for the KL and WA populations were significantly different from identical between cohorts.

Finally, we tested for differences between matrices using the Bayesian approach of Ovaskainen *et al.* (2008). The resulting 'd' divergence statistics were small for all comparisons. The mean 'd' value among within-population tests between rearing treatments was 0.08 (SD = 0.01), with a similar mean value (mean = 0.05, SD = 0.01) for between-population comparisons among wild-caught males. The mean 'd' value for between-population comparisons among common-garden reared males was slightly higher, at 0.16 (SD = 0.03). However, the ψ_d statistics for all comparisons were high (all $>$ 0.95), which implies a lack of statistical support for these matrices being different. A table of results for individual tests is available in Online Supplement 1.

Experiment 2. The plasticity of P under dietary manipulation

We used ANOVA'S to test for the influence of diet treatment on the development time and size of crickets from all six populations. Males reared on the low-nutrient diet took 2.6 (\pm 1.2) days longer to reach eclosion than those on the high-nutrient diet ($F_{1,719} = 21.97$, $P = 0.0001$). Development time varied significantly among populations ($F_{5,719} = 61.17$, $P = 0.0001$), but there was no significant interaction effect of population

x diet. In addition to developing more slowly, males reared on the low-nutrient diet also weighed 0.03 g (\pm 0.01) more and were 0.22 mm (\pm 0.01) wider in the pronotum (weight: $F_{1,719} = 49.87$, $P = 0.0001$; PW: $F_{1,719} = 43.52$, $P = 0.0001$). Weight and pronotum width also varied significantly among populations (weight: $F_{5,719} = 30.03$, $P = 0.001$; PW: $F_{5,719} = 32.56$, $P = 0.0001$). However, unlike development time, weight and pronotum width did show a significant interaction between population and diet (weight: $F_{5,719} = 5.08$, $P = 0.0001$; PW: $F_{5,719} = 3.36$, $P = 0.005$). Given that these effects were measured in common-garden conditions, this suggests genotype x environment interactions for these phenotypic traits.

We tested the effects of population, diet and their interaction on male call structure using the same model as specified for Experiment 1. However, due to the distribution of the call data we obtained *P*-values using a permutation test rather than from the parametric MANCOVA (Table 4). As in Experiment 1, we found that the call traits showed significant differences among populations. These among-population differences (Fig. 4) appear to be similar to the pattern found in the common-garden reared cohort from experiment 1 (dashed ellipses in Fig. 3), with the distributions being of similar size and shape among populations. There is no evidence for a diet effect on call traits. Although the between-diet differences in population means appear to be in different directions (Fig. 6), the population \times diet effect is also nonsignificant. As in Experiment 1, pronotum width was also a significant predictor of call traits. *Post hoc* ANOVA'S, with *P*-values again assigned by permutation, showed that the among-population differences

Table 4 Statistics from the randomisation tests on the data from the Experiment 2. Significant *P*-values are italicized.

Effect	#permutations	#permutations $P_p \leq P$	<i>P</i> -value		
Population (A)	10 000	0	<0.0001		
Diet (B)	10 000	4952	0.50		
A \times B	10 000	1696	0.17		
Size (PW)	10 000	276	0.03		
Trait Means (\pm SE) by Population					
High-Nutrient Diet	CPN	TN	DF	ICD	CIPD
ACT	5.64 (0.38)	3.5 (0.25)	4.81 (0.30)	188.11 (19.36)	17.37 (1.22)
KL	6.19 (0.50)	3.48 (0.32)	5.25 (0.35)	202.05 (25.84)	20.70 (1.40)
SA	6.00 (0.39)	3.73 (0.32)	4.49 (0.22)	163.37 (15.17)	17.32 (1.21)
SL	5.22 (0.28)	3.27 (0.18)	4.36 (0.18)	174.35 (15.25)	18.40 (1.87)
TAS	6.27 (0.30)	3.27 (0.22)	4.19 (0.17)	198.31 (28.04)	18.57 (1.50)
WA	5.69 (0.23)	3.19 (0.22)	4.10 (0.11)	209.88 (23.38)	20.26 (0.87)
Low-Nutrient Diet	CPN	TN	DF	ICD	CIPD
ACT	5.7 (0.35)	3.94 (0.29)	4.74 (0.27)	172.87 (12.3)	19.74 (1.43)
KL	5.26 (0.27)	3.31 (0.28)	4.32 (0.19)	159.41 (11.5)	16.95 (0.89)
SA	5.8 (0.33)	3.27 (0.19)	4.52 (0.25)	184.79 (10.4)	20.57 (2.04)
SL	5.48 (0.19)	3.63 (0.24)	4.42 (0.18)	146.94 (5.92)	16.59 (0.96)
TAS	5.72 (0.34)	3.43 (0.20)	4.31 (0.21)	214.72 (25.5)	18.31 (1.34)
WA	5.90 (0.31)	3.67 (0.30)	4.41 (0.22)	220.96 (17.3)	22.30 (1.33)
<i>Post hoc P</i> < 0.05	<i>PW</i>	-	<i>A, PW</i>	<i>A</i>	<i>A, PW</i>

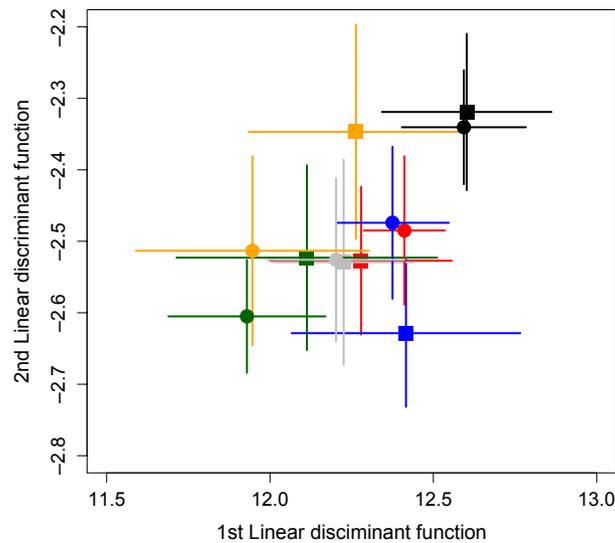


Fig. 6 Variation in call structure among populations and between diet treatments in Experiment 2. Population mean linear discriminant scores for the first and second linear discriminant functions were calculated for call structure using population and rearing environment as grouping variables. Populations are colour coded and abbreviated according to the legend of Fig. 1. The round points represent calls measured from males reared on the high-nutrient diet, whereas the square points represent males reared on the low-nutrient diet, and the bars represent the standard error around the mean.

were driven by significant differences in ICD, CIPD and DF (10 000 permutations; ICD $P = 0.001$; CIPD $P = 0.002$; DF $P = 0.048$) and the effect of pronotum width was driven by CPN, CIPD and DF (10 000 permutations; CPN $P = 0.003$; CIPD $P = 0.008$; DF $P = 0.001$).

Our CPC analysis of \mathbf{P} from Experiment 2 found that for males reared on both diets, the most parsimonious model comparing \mathbf{P} among populations was the proportional model. This result is in agreement with the finding from Experiment 1 that these \mathbf{P} matrices differ in their eigenvalues, but not in the direction of their eigenvectors. When we compared \mathbf{P} within populations across our diet treatments, the most parsimonious model was that of proportional matrices for all populations, with the exception of the KL population whose best model was for the common principal components. These results indicate that the structure of \mathbf{P} for call structure was largely unaffected by the difference between the diet treatments, with only the KL population showing evidence of change but this was for eigenvalues only and the eigenvectors remained the same across diets.

Results from matrix comparison by Krzanowski's geometric method (Table 5) were similar to those for Experiment 1. Among-population comparisons between males reared on the high-nutrient diet revealed differences of a similar scale to those found in the common-garden cohort in Experiment 1: sums of $\mathbf{S}\lambda$ values were ≥ 0.99 (identical matrices would return a value of 2), and the angles between most closely aligned vectors

were $\leq 13.8^\circ$. The among-population differences between \mathbf{P} matrices were of similar magnitude when reared on the low-nutrient diet, with sum of $\mathbf{S}\lambda$ values ≥ 1.14 , and all maximum angles between closest vectors $\leq 17.1^\circ$. The between-diet differences within populations were somewhat smaller than the inter-population differences within either diet treatment, with sum of $\mathbf{S}\lambda$ values ≥ 1.26 and angles between most closely aligned vectors $\leq 7.79^\circ$.

We compared matrices in a pairwise fashion by the application of random skewers. The random skewers vector correlations among populations in the cohort reared on the high-nutrient diet were all ≥ 0.83 (mean \pm SE = 0.93 ± 0.06), and only 4 of 15 comparisons tested as significantly different from identical (Table 6). Among populations in the cohort reared on the low-nutrient diet random skewer vector correlations were all ≥ 0.78 (mean \pm SE = 0.92 ± 0.06), and 5 of 15 comparisons were significantly different from identical (Table 6). Within-populations between diet treatments, random skewer vector correlations were all ≥ 0.85 (mean \pm SE = 0.95 ± 0.03), and only populations KL and SA tested as different from identical between treatments.

Lastly, we tested for differences between \mathbf{P} matrices using the Bayesian ' d ' method (see supplementary material for full details). Divergence (' d ') values were somewhat higher than for Experiment 1. The mean ' d ' value for within-population comparisons between diet treatments was 1.22 (SD = 0.21), with a mean of 1.32 (SD = 0.25) for between population tests on the

Table 5 The results of geometric matrix comparisons among **P** matrices from Experiment 2. The upper matrix contains the sum of **S** matrix eigenvalues (a value between 0 & 2 in this case – 2 indicating that matrices span an identical subspace, with their 95% confidence interval below them). The lower *italicized* matrix contains the closest angles between principal vectors describing the shared subspace between matrices (in degrees), with 95% confidence intervals. Within each matrix, cells on the diagonal (**bold**) are within-population comparisons between males reared on the high-nutrient diet and those reared on the low-nutrient diet. Cells below the diagonal are between-population comparisons of calls from the high-nutrient diet treatment – between-population comparisons of calls from the low-nutrient treatment are above the diagonal. Population abbreviations are as for Fig. 1.

	ACT	KL	SA	SL	TA	WA		
ACT	1.63 1.28–1.91	1.53 1.04–1.98	1.52 0.36–1.90	1.26 0.99–1.82	0.99 0.31–1.94	1.87 0.99–1.99	Low-Nutrient Diet	
KL	1.73 0.84–1.99	1.03 0.75–1.98	1.01 0.35–1.96	1.70 0.98–1.96	1.23 0.32–1.94	1.20 0.74–1.99		
SA	1.80 1.05–1.99	1.60 0.80–1.99	1.63 0.51–1.99	1.02 0.27–1.85	1.43 0.49–1.88	1.68 0.38–1.98		
SL	1.03 0.95–1.37	1.00 0.78–1.43	1.05 0.93–1.77	1.16 0.96–1.93	1.02 0.30–1.79	1.04 0.61–1.97		
TA	1.32 0.94–1.96	1.26 0.52–1.84	1.03 0.73–1.96	1.71 0.60–1.97	1.94 0.59–1.96	1.04 0.38–1.95		
WA	1.08 0.97–1.72	1.04 0.73–1.59	1.45 0.97–1.90	1.74 0.99–1.96	1.29 0.56–1.98	1.24 0.70–1.80		
ACT	0.72 0.13–4.83	3.62 0.15–10.2	1.41 0.04–53.1	3.68 0.13–8.06	6.61 0.15–58.8	4.73 0.01–16.9		
KL	1.92 0.12–30.6	5.21 0.10–32.0	3.62 0.02–55.2	1.58 0.10–9.58	3.24 0.07–58.1	7.76 0.08–36.7		
SA	3.29 0.06–7.73	4.21 0.02–31.1	3.19 0.02–47.2	3.54 0.12–54.7	7.55 0.14–45.1	1.20 0.06–53.8		
SL	4.14 0.23–11.2	4.59 0.17–25.3	6.79 0.40–13.5	0.40 0.01–11.6	4.07 0.03–52.6	7.50 0.08–41.2		
TA	5.45 0.23–22.4	5.19 0.15–43.2	1.47 0.04–36.0	1.66 0.08–42.5	3.66 0.20–41.9	6.85 0.12–54.6		
WA	1.51 0.09–6.93	3.15 0.07–27.8	3.77 0.10–11.6	1.23 0.05–8.64	5.94 0.09–46.3	5.44 0.15–34.0		
	High-Nutrient Diet							

Table 6 Random skewers correlations between and within populations in Experiment 2. Correlations between populations reared on the high-nutrient diet are below the diagonal, and between those on the low-nutrient diet are above the diagonal. Correlations along the diagonal are within population and between diets. Below each correlation is its associated *P*-value in *italics*, where a significant value indicates difference from equality.

	ACT	KL	SA	SL	TA	WA		
ACT	0.97 0.33	0.90 0.05	0.86 0.01	0.99 0.63	0.99 0.86	0.90 0.05	Low-Nutrient Diet	
KL	0.91 0.02	0.85 0.00	0.94 0.12	0.93 0.11	0.93 0.11	0.93 0.12		
SA	0.99 0.71	0.87 0.01	0.90 0.04	0.86 0.01	0.90 0.04	0.78 0.00		
SL	0.97 0.31	0.97 0.27	0.94 0.11	0.99 0.74	0.99 0.74	0.96 0.25		
TA	1.00 0.89	0.92 0.05	0.98 0.49	0.96 0.27	0.99 0.73	0.92 0.10		
WA	0.91 0.06	0.83 0.00	0.95 0.24	0.89 0.03	0.88 0.05	0.97 0.44		
	High-Nutrient Diet							

high-nutrient diet and 0.86 (SD = 0.15) for between population tests on the low-nutrient diet. However, as with Experiment 1 these differences lacked statistical support, with ψ_d statistics for all tests being > 0.98.

We also applied the same suite of matrix comparisons to **P** matrices estimated for the three life-history traits we measured in Experiment 2. These analyses provided little evidence to suggest that the life-history **P** matrices differed among populations or between diet treatments (see Online Supplement 2).

Discussion

Quantifying the scale of variation in the phenotypic variance-covariance matrix (**P**) is vital to our understanding of the evolution of complex traits (Roff *et al.*, 2000). Over short timescales, **G** is predicted to influence the direction of response to selection (Lande, 1979; Schluter, 1996), but over longer timescales selection is expected to influence the structure of **G** (Stephan *et al.*, 2002). Since **P** comprises **G**, this is also true of **P**, but with the added complication that we must also consider changes due to the environment (**E**).

Comparisons of **P** have sometimes found substantial changes between divergent populations (Badyaev & Hill, 2000), although more frequently **P** has been found to be stable at this level of comparison (Arnold & Phillips, 1999; Game & Caley, 2006; de Oliveira *et al.*, 2009). Notably, we are not aware of any previous comparisons of **P** for a sexual signal. In this study, we compared **P** for the advertisement call of six allopatric populations of the field cricket *T. commodus*, as well as estimating **P** from the recorded calls of both wild-caught and common-garden-reared males in each population to assess population divergence in **P**. We then re-measured advertisement calls from an independent collection of these populations after rearing them on high- and low-nutrient diets to assess the plasticity of **P**. We found that call structure varied among populations, with evidence for only small changes in **P** that were of a similar degree in wild-caught and common-garden-reared males. Although our manipulation of diet was sufficient to measurably affect the life-history of crickets (i.e. development time and body size), it did not result in a plastic response in the mean structure of advertisement calls nor did it alter the structure of **P**. Together, the results of our two experiments suggest that although minor differences in the structure of **P** were detected on various scales (i.e. populations, generations or diets), the overarching pattern was that **P** remained relatively stable in each of these comparisons. **P** matrix stability has potential consequences for the evolution of advertisement call, depending on the nature of selection. If the structure of **P** is stable, then it has the potential to influence the rate of response to selection; allowing a stronger response if **P** and **G** are aligned, and slowing the response if **P** and **G** are poorly aligned.

We identified differences in the five call traits across populations and between generations, and these differences remained after common-garden rearing, indicating that the divergence in call structure observed across of our six populations has a genetic basis. This pattern is not altogether surprising as geographical variation in the call structure of Gryllids has been demonstrated before (Ferreira & Ferguson, 2002; Jang & Gerhardt, 2006), including in the genus *Teleogryllus* (Simmons *et al.*, 2001), and these structural components of *T. commodus* call are known to have a strong genetic basis (Hunt *et al.*, 2007). The effect of generation (wild-caught vs. common-garden reared) indicates a plastic developmental response in addition to the inferred genetic differences. The calls of the six populations were less variable and more similar to each other after rearing in common-garden, but as the interaction between generation and population demonstrates, the six populations responded differently to the common-garden conditions of the laboratory. Although not conclusive, this pattern suggests a genotype x environment interaction for advertisement call

traits in *T. commodus* (Olvido & Mousseau, 1995; Tolle & Wagner, 2011). Male body size was also a significant predictor of calling traits, but differing body size between rearing environments or among populations does not appear to be driving all the differences in call traits, since we find that body size is predictive only of CIPD (Table 1). This finding does not fit the general hypothesis that dominant call frequency signals body size in Orthopterans (e.g. Bennet-Clark, 1999). Thus, to summarize, we find convincing evidence that the call traits we measured differ among populations, and are responsive to common-garden rearing environment in the laboratory.

As all the call traits were affected by the interaction of population and generation, it is perhaps not surprising that we detect differences among **P** matrices for advertisement call structure. However, the evidence suggests that these differences are small and therefore unlikely to lead to biologically significant differences in the evolution of call structure under a given regime of selection. Among the wild-caught cohort, the general outcome from CPC analysis was for two common principal components, which fits well with the results from the Krzanowski test where the mean closest eigenvector angle was 25.2°, indicating that the subspaces are closer to alignment than to being orthogonal. These results indicate a comparable level of among-population divergence to that found between populations of house finches by (Badyaev & Hill, 2000), with partial common principal components and statistical support for some differences (Tables 3 and 5). Somewhat smaller differences have been reported from other inter-population matrix comparison, for example, in garter snake morphology (Arnold & Phillips, 1999) and in the body shape of six species of reef fish (Game & Caley, 2006) where all principal components were found to be common. However, despite this evidence of differences among populations, the mean random skewers correlation was high (0.81) and there was statistical support for matrices being different in only 60% of pairwise comparisons. Among the common-garden reared cohort, there is some evidence that **P** matrices were slightly more similar, with the consensus CPC verdict being a step higher on Flury's hierarchy (i.e. CPC(3)), Kzanowski tests giving a lower mean closest eigenvector angle (10.2°) and a slightly lower mean correlation of random skewers (0.73), but with only 60% of pairwise comparison showing statistical support for divergence on **P**. The within-population differences in **P** across generations were of a similar magnitude to the between population differences, with a consensus verdict of 3 common PCs, a mean closest eigenvector angle of 22.5° and a high mean random skewers correlation (0.92). This indicates that the **P** matrices for advertisement calls have a very similar structure in both the wild and the laboratory. Together, these results show small but detectable differences among populations in **P**

matrices and illustrate the value of using multiple tests to assess matrix similarity (Roff *et al.*, 2012).

When we re-collected from these same populations three years later (Experiment 2), we once again detected the significant inter-population differences in advertisement call. However, the differences were confined to DF, TN and CIPD, as opposed to Experiment 1 where we found evidence for inter-population differences in all five call traits. Notably, in addition to its association with CIPD in Experiment 1, body size is a significant predictor of both CPN and DF. The relationship between size and call traits appears to have changed and this suggests that the link between call frequency and body size (Simmons, 1995; Brown *et al.*, 1996; Simmons & Ritchie, 1996; Scheuber *et al.*, 2003a,b) may be mediated, in part, by some other additional factors. We know from Experiment 1 that the call traits we measured are responsive to the number of generations of laboratory common-garden rearing, and a number of other studies have found call trait plasticity in response to diet (e.g. Wagner & Hoback, 1999; Scheuber *et al.*, 2003b; Hedrick, 2005). Our manipulation of diet had effects on body size and development time, in line with what has been observed in other field crickets (Scheuber *et al.*, 2003a,b), but we found no evidence of differences in **P** for these life-history traits, nor any associated changes in call structure. This is not unprecedented, since diet manipulations have previously been shown to result in body size effects without affecting advertisement calls in bushcrickets (Hartbauer *et al.*, 2006), and dietary effects on call structure without differences in body size have been found in a field cricket (Wagner & Hoback, 1999). Therefore, it seems that the relative magnitudes of change in diet required to affect calling behaviour and body size may be somewhat species-specific.

Overall, the results of Experiment 2 found that **P** matrices for advertisement call structure were relatively more stable than those estimated from Experiment 1. At all three levels of comparison: among-populations within the low- and high-nutrient diets and within-populations between diets, the consensus verdict from our CPC analysis was that all principal components were common. The mean random skewers correlations were also very high, although some of the pairwise comparisons on both diets were significantly different from identical and the Krzanowski test also indicated a greater degree of similarity for comparisons than in Experiment 2. Taken together these results strongly suggest that **P** for the structure of the advertisement call, like the individual trait means, were robust in response to the diet manipulation. This was in strong contrast to the plasticity in development time and body size shown in this experiment.

Plasticity in the face of a heterogeneous environment is an extremely common adaptation in a variety of species (West-Eberhard, 2003; DeWitt & Scheiner, 2004), and the advertisement calls of numerous cricket species

are known to exhibit this kind of plasticity (Elliott & Koch, 1985; Wagner & Hoback, 1999; Scheuber *et al.*, 2003b; Hunt *et al.*, 2004; Grace & Shaw, 2004; Hedrick, 2005; Jacot *et al.*, 2005; Judge *et al.*, 2008; Zajitschek *et al.*, 2009). If sexual signals are too readily plastic, however, their reliability as signals is predicted to break down (Greenfield & Rodriguez, 2004; Ingleby *et al.*, 2010). When considering complex traits that function as sexual signals, this leads to somewhat paradoxical predictions (Badyaev, 2004): they should be integrated with the rest of the phenotype in order that any variability may function as an honest indicator of physiological quality, yet response to sexual selection is also expected to uncouple sexual signals from the rest of the phenotype to enable them to become plastic in their expression and evolve independently from the rest of the phenotype (Emlen & Nijhout, 2000). Natural selection can be expected to favour integration as it contributes to the development of robust and functionally resilient body-plans (Pigliucci, 2003), but this does not mean that the processes of plasticity and integration need be thought of as mutually exclusive (Phillips & McGuigan, 2006).

The plasticity we measured in *T. commodus* advertisement calls between field conditions and common-garden rearing involved both shifts in the position of the population means and changes in some of the eigenvectors of **P** (Figs 5 and 3). The eigenvector angles show small to intermediate changes between wild and common-garden matrices (confidence intervals in Table 2), but the strong random skewers correlations (Table 3) suggest that these effects are relatively unimportant. The plastic response was similar in magnitude to the differences found among populations. In comparison, in most other cases where the stability of integration has been assessed by comparing covariance matrices between populations they have often found intermediate levels of difference (Cowley & Atchley, 1990; Shaw & Billington, 1991; Platenkamp & Shaw, 1992; Brodie, 1993; Fenster, 1994; Jernigan *et al.*, 1994) or more recently, detected conserved eigenstructure (Roff *et al.*, 2004; Game & Caley, 2006). In general, we might predict that inter-population differences would be predominantly described by changes in means and variances, with changes in covariances being less important.

A number of researchers studying multivariate evolution have recently used the concept of modularity (Wagner & Altenberg, 1996; Hansen *et al.*, 2003; Mezey & Houle, 2003; Badyaev, 2004; Kraft *et al.*, 2006; Bossdorf & Pigliucci, 2009; de Oliveira *et al.*, 2009). A phenotypic module is a suite of traits that display enhanced integration with each other and reduced integration with the rest of the organism (Wagner & Altenberg, 1996). In light of the limited plasticity we observe in **P** and the apparent insensitivity of call traits to the plasticity of the animals' life-history, it seems reasonable to suggest that the structure of the advertisement call in

T. commodus meets this criterion. Badyaev (2004) specifically addressed the issue of modularity and integration in sexual signals and made three predictions: that sexual traits should exhibit (1) weakened developmental integration with other traits, (2) strengthened functional integration and modularity, and (3) weaker genetic integration with the rest of the organism. We found evidence supporting prediction (1). Although pronotum width was a significant predictor of call traits in both experiments (Table 1 and 4), the correlations between call traits and pronotum width were consistently low (Experiment 1, $r = 0.20$; Experiment 2, $r = 0.17$). This contrasts the tight correlations observed between body size and body weight (Experiment 1, $r = 0.79$; Experiment 2, $r = 0.87$) and between body size and development time (Experiment 1, $r = 0.49$; Experiment 2, $r = 0.54$). Our finding of limited plasticity in the structure of **P** for call traits also provides support for prediction (2). Although the structure of **P** varied across populations and time in Experiment 1, the magnitude of these differences was small. This pattern was even more pronounced in Experiment 2 where body size, weight and development time showed a plastic response to diet, but neither mean call structure nor the structure of **P** for call traits differed across diets. This finding suggests that call structure is unlikely to signal a male's underlying condition (i.e. condition-dependence) but instead represents a tightly integrated functional module. Taken collectively, this suggests that call structure is tightly integrated in *T. commodus* and largely independent of other elements of the phenotype which should, in theory, allow call structure to evolve independently of traits that are not subject to sexual selection (Badyaev, 2004). Unfortunately, as our work consists of phenotypic measurements only, we are unable to directly test prediction (3), although our current work is addressing this question.

Sexual selection imposed by female mate choice on the structure of advertisement call has been measured in our study populations and there are significant differences in linear, quadratic and correlational selection gradients across populations (Hunt *et al.* unpublished data). Given these differences in sexual selection, it is therefore not surprising that populations also differ in the call traits that are the target of female choice. However, we found little evidence to suggest that the differences between populations in the structure of **P** was any greater in magnitude than the response of **P** to common-garden rearing in the laboratory. This is surprising as a regime of both quadratic and correlational selection is likely to influence the evolution of phenotypic integration and thus the structure of **P** (Cheverud, 1984). Although not conclusive, this finding is further evidence of the constraining influence that the integration of **P** has on phenotypic evolution, enabling some divergence and plasticity in call traits but ultimately conserving the underlying structure of **P**.

Acknowledgments

This study was supported by Natural and Environment Research Council studentship to WP. Collections of animals from the field would have been impossible without the assistance from members of the Brooks and Jennions labs. Advice on an earlier version of this MS was appreciated from Fiona Ingleby and Chris Chandler. JH and TT were funded by Natural Environment Research Council and JH by a Royal Society Fellowship. RB and MDJ were funded by the Australian Research Council. ID was funded by the National Institutes of Health and Michigan State University.

References

- Ackermann, R. & Cheverud, J. 2000. Phenotypic covariance structure in tamarins (genus *Saguinus*): A comparison of variation patterns using matrix correlation and common principal component analysis. *Am. J. Phys. Anthropol.* **111**: 489–501.
- Agrawal, A.A., Laforsch, C. & Tollrian, R. 1999. Transgenerational induction of defences in animals and plants. *Nature* **401**: 60–63.
- Agrawal, A.F., Brodie, E. & Rieseberg, L. 2001. Possible consequences of genes of major effect: transient changes in the **G**-matrix. *Genetica* **112**: 33–43.
- Andersson, M. 1982. Sexual selection, natural-selection and quality advertisement. *Biol. J. Linn. Soc.* **17**: 375–393.
- Arnold, S. & Phillips, P. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake. *Thamnophis elegans*. *Evol.* **53**: 1516–1527.
- Badyaev, A.V. 2004. Integration and Modularity in the Evolution of Sexual Ornaments: An Overlooked Perspective. In: *Phenotypic Integration* (M. Pigliucci & K. Preston, eds), pp. 50–79. Oxford University Press, Oxford.
- Badyaev, A. & Hill, G. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. *Evolution* **54**: 1784–1794.
- Bégin, M. & Roff, D.A. 2001. An analysis of **G** matrix variation in two closely related cricket species, *Gryllus firmus* and *G. pennsylvanicus*. *J. Evol. Biol.* **14**: 1–13.
- Bégin, M., Roff, D.A. & Debat, V. 2004. The effect of temperature and wing morphology on quantitative genetic variation in the cricket *Gryllus firmus*, with an appendix examining the statistical properties of the Jackknife-MANOVA method of matrix comparison. *J. Evol. Biol.* **17**: 1255–1267.
- Bennet-Clark, H. 1999. Resonators in insect sound production: How insects produce loud pure-tone songs. *J. Exp. Biol.* **202**: 3347–3357.
- Bentley, D.R. & Hoy, R.R. 1972. Genetic control of the neuronal network generating cricket (*Teleogryllus Gryllus*) song patterns. *Anim. Behav.* **20**: 478–492.
- Bentsen, C., Hunt, J., Jennions, M. & Brooks, R. 2006. Complex multivariate sexual selection on male acoustic signaling in a wild population of *Teleogryllus commodus*. *Am. Nat.* **167**: E102–E116.
- Bischoff, L.L., Tschirren, B. & Richner, H. 2009. Long-term effects of early parasite exposure on song duration and singing strategy in great tits. *Behav. Ecol.* **20**: 265–270.

- Blows, M. & Higgie, M. 2003. Genetic constraints on the evolution of mate recognition under natural selection. *Am. Nat.* **161**: 240–253.
- Blows, M.W., Chenoweth, S.F. & Hine, E. 2004. Orientation of the Genetic Variance-Covariance Matrix and the Fitness Surface for Multiple Male Sexually Selected Traits. *Am. Nat.* **163**: 329–340.
- Bosssdorf, O. & Pigliucci, M. 2009. Plasticity to wind is modular and genetically variable in *Arabidopsis thaliana*. *Evol. Ecol.* **23**: 669–685.
- Brodie, E. 1993. Homogeneity of the Genetic Variance-Covariance Matrix for Antipredator Traits in 2 Natural Populations of the Garter Snake *Thamnophis ordinoides*. *Evolution* **47**: 844–854.
- Brooks, R., Hunt, J., Blows, M., Smith, M., Bussiere, L. & Jennions, M. 2005. Experimental evidence for multivariate stabilizing sexual selection. *Evolution* **59**: 871–880.
- Brown, W., Wideman, J., Andrade, M., Mason, A. & Gwynne, D. 1996. Female choice for an indicator of male size in the song of the black-horned tree cricket *Oecanthus nigricornis* (Orthoptera: Gryllidae: Oecanthinae). *Evolution* **50**: 2400–2411.
- Calsbeek, B. & Goodnight, C.J. 2009. Empirical Comparison of G Matrix Test Statistics: Finding Biologically Relevant Change. *Evolution* **63**: 2627–2635.
- Camara, M., Ancell, C. & Pigliucci, M. 2000. Induced mutations: A novel tool to study phenotypic integration and evolutionary constraints in *Arabidopsis thaliana*. *Evol. Ecol. Res.* **2**: 1009–1029.
- Cheverud, J.M. 1984. Quantitative genetics and developmental constraints on evolution by selection. *J. Theor. Biol.* **110**: 155–171.
- Cheverud, J.M. 1995. Morphological Integration in the Saddle-Back Tamarin (*Saguinus fuscicollis*) Cranium. *Am. Nat.* **145**: 63–89.
- Cheverud, J. 1996. Developmental integration and the evolution of pleiotropy. *Am. Zool.* **36**: 44–50.
- Cheverud, J.M. & Marroig, G. 2007. Comparing covariance matrices: Random skewers method compared to the common principal components model. *Genet. Mol. Biol.* **30**: 461–469.
- Cowley, D. & Atchley, W. 1990. Development and Quantitative Genetics of Correlation Structure Among Body Parts of *Drosophila melanogaster*. *Am. Nat.* **135**: 242–268.
- DeWitt, T. & Scheiner, S. 2004. *Phenotypic Plasticity: Functional and Conceptual Approaches*. Oxford University Press, Oxford.
- Elliott, C. & Koch, U. 1985. The Clockwork Cricket. *Naturwissenschaften* **72**: 150–153.
- Emlen, D. & Nijhout, H. 2000. The development and evolution of exaggerated morphologies in insects. *Annu. Rev. Entomol.* **45**: 661–708.
- Eroukhmanoff, F. 2009. Just how much is the G-matrix actually constraining adaptation? *Evol. Biol.* **36**: 323–326.
- Fenster, C. 1994. Levels of genetic variation and covariation for *Mimulus* (Scrophulariaceae) floral traits. *Heredity*.
- Ferreira, M. & Ferguson, J. 2002. Geographic variation in the calling song of the field cricket *Gryllus bimaculatus* (Orthoptera: Gryllidae) and its relevance to mate recognition and mate choice. *J. Zool.* **257**: 163–170.
- Game, E.T. & Caley, M.J. 2006. The stability of P in coral reef fishes. *Evolution* **60**: 814–823.
- Grace, J. & Shaw, K. 2004. Effects of developmental environment on signal-preference coupling in a Hawaiian cricket. *Evolution* **58**: 1627–1633.
- Gray, D.A. & Eckhardt, G. 2001. Is cricket courtship song condition dependent? *Anim. Behav.* **62**: 871–877.
- Greenfield, M.D. & Rodriguez, R.L. 2004. Genotype-environment interaction and the reliability of mating signals. *Anim. Behav.* **68**: 1461–1468.
- Hadfield, J.D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmm R Package. *J. Stat. Softw.* **33**: 1–22.
- Hansen, T.F., Armbruster, W.S., Carlson, M.L. & Pélabon, C. 2003. Evolvability and genetic constraint in Dalechampia blossoms: genetic correlations and conditional evolvability. *J. Exp. Zool. B Mol. Dev. Evol.* **296**: 23–39.
- Hartbauer, M., Kratzer, S. & Romer, H. 2006. Chirp rate is independent of male condition in a synchronising bushcricket. *J. Insect Physiol.* **52**: 221–230.
- Hedrick, A. 2005. Environmental condition-dependent effects on a heritable, preferred male trait. *Anim. Behav.* **70**: 1121–1124.
- Hennig, R.M. & Weber, T. 1997. Filtering of temporal parameters of the calling song by cricket females of two closely related species: a behavioral analysis. *J. Comp. Physiol., A* **180**: 621–630.
- Hill, K., Loftus-Hills, J. & Gartside, D. 1972. Pre-mating isolation between the Australian field crickets *Teleogryllus commodus* and *T. oceanicus* (Orthoptera: Gryllidae). *Aust. J. Zool.* **20**: 153–163.
- Hunt, J., Brooks, R., Jennions, M.D., Smith, M.J., Bentsen, C.L. & Bussiere, L.F. 2004. High-quality male field crickets invest heavily in sexual display but die young. *Nature* **432**: 1024–1027.
- Hunt, J., Blows, M., Zajitschek, F., Jennions, M. & Brooks, R. 2007. Reconciling strong stabilizing selection with the maintenance of genetic variation in a natural population of black field crickets (*Teleogryllus commodus*). *Genetics*.
- Ingleby, F.C., Hunt, J. & Hosken, D.J. 2010. The role of genotype-by-environment interactions in sexual selection. *J. Evol. Biol.* **23**: 2031–2045.
- Jacot, A., Scheuber, H., Kurtz, J. & Brinkhof, M. 2005. Juvenile immune status affects the expression of a sexually selected trait in field crickets. *J. Evol. Biol.* **18**: 1060–1068.
- Jang, Y. & Gerhardt, H. 2006. Divergence in the calling songs between sympatric and allopatric populations of the southern wood cricket *Gryllus fultoni* (Orthoptera: Gryllidae). *J. Evol. Biol.* **19**: 459–472.
- Jernigan, R., Culver, D. & Fong, D. 1994. The Dual Role of Selection and Evolutionary History as Reflected in Genetic Correlations. *Evolution* **48**: 587–596.
- Johnstone, R. 1995. Honest Advertisement of Multiple Qualities Using Multiple Signals. *J. Theor. Biol.* **177**: 87–94.
- Judge, K.A., Ting, J.J. & Gwynne, D.T. 2008. Condition Dependence of Male Life Span and Calling Effort in a Field Cricket. *Evolution* **62**: 868–878.
- Kolodnynska, A. & Pigliucci, M. 2003. Multivariate responses to flooding in *Arabidopsis*: an experimental evolutionary investigation. *Funct. Ecol.* **17**: 131–140.
- Kraft, P., Wilson, R., Franklin, C. & Blows, M. 2006. Substantial changes in the genetic basis of tadpole morphology of *Rana lessonae* in the presence of predators. *J. Evol. Biol.* **19**: 1813–1818.
- Krzanowski, W. 1979. Between-Groups Comparison of Principal Components. *J. Am. Stat. Assoc.* **74**: 703–707.

- Lande, R. 1979. Quantitative Genetic-Analysis of Multivariate Evolution, Applied to Brain - Body Size Allometry. *Evolution* **33**: 402–416.
- Lande, R. 1980. The Genetic Covariance between Characters Maintained by Pleiotropic Mutations. *Genetics* **94**: 203. Genetics Society of America.
- Lande, R. & Arnold, S. 1983. The Measurement of Selection on Correlated Characters. *Evolution* **37**: 1210–1226.
- McGlothlin, J.W., Parker, P.G., Nolan, V. & Ketterson, E.D. 2005. Correlational Selection Leads to Genetic Integration of Body Size and an Attractive Plumage Trait in Dark-Eyed Juncos. *Evolution* **59**: 658–671.
- Merilä, J. & Björklund, M. 1999. Population Divergence and Morphometric Integration in the Greenfinch (*Carduelis chloris*) - Evolution Against The Trajectory Of Least Resistance? *J. Evol. Biol.* **12**: 103–112.
- Mezey, J.G. & Houle, D. 2003. Comparing **G** matrices: Are common principal components informative? *Genetics* **165**: 411–425.
- de Oliveira, F.B., Porto, A. & Marroig, G. 2009. Covariance structure in the skull of Catarrhini: a case of pattern stasis and magnitude evolution. *J. Hum. Evol.* **56**: 417–430.
- Olson, E.C. & Miller, R.L. 1999 *Morphological Integration*. University Of Chicago Press.
- Olvido, A. & Mousseau, T. 1995. Effect of rearing environment on calling-song plasticity in the striped ground cricket. *Evolution* **49**: 1271–1277.
- Ovaskainen, O., Cano, J.M. & Merilä, J. 2008. A Bayesian framework for comparative quantitative genetics. *Proc. Biol. Sci.* **275**: 669–678.
- Phillips, P.C. 1998. *CPC: Common Principal Components Analysis*. University of Oregon, Eugene, OR. URL: pages.uoregon.edu/pphill/programs/cpc/cpc.htm
- Phillips, P.C. & Arnold, S.J. 1999. Hierarchical comparison of genetic variance-covariance matrices. I. Using the Flury hierarchy. *Evolution* **53**: 1506–1515.
- Phillips, P.C. & McGuigan, K. 2006 Evolution of Genetic Variance-Covariance Structure. In: *Evolutionary Genetics: concepts and case studies* (C.W. Fox, J.B. Wolf, eds), pp. 310–325. Oxford University Press, New York.
- Phillips, P.C., Whitlock, M.C. & Fowler, K. 2001. Inbreeding changes the shape of the genetic covariance matrix in *Drosophila melanogaster*. *Genetics* **158**: 1137–1145.
- Pigliucci, M. 2003. Phenotypic integration: studying the ecology and evolution of complex phenotypes. *Ecol. Lett.* **6**: 265–272.
- Pigliucci, M. & Kolodynska, A. 2002a. Phenotypic plasticity and integration in response to flooded conditions in natural accessions of *Arabidopsis thaliana* (L.) Heynh (Brassicaceae). *Ann. Bot.* **90**: 199–207.
- Pigliucci, M. & Kolodynska, A. 2002b. Phenotypic plasticity to light intensity in *Arabidopsis thaliana*: invariance of reaction norms and phenotypic integration. *Evol. Ecol.* **16**: 27–47.
- Platenkamp, G. & Shaw, R. 1992. Environmental and Genetic Constraints on Adaptive Population Differentiation in *Anthoxanthum odoratum*. *Evolution* **46**: 341–352.
- Pollack, G. & Hoy, R. 1979. Temporal Pattern as a Cue for Species-Specific Calling Song Recognition in Crickets. *Science* **204**: 429–432.
- R Core Team. 2012. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Roach, D. & Wulff, R. 1987. Maternal Effects in Plants. *Annu. Rev. Ecol. Syst.* **18**: 209–235.
- Roff, D.A. & Mousseau, T.A. 1999. Does natural selection alter genetic architecture? An evaluation of quantitative genetic variation among populations of *Allonemobius socius* and *A. fasciatus*. *J. Evol. Biol.* **12**: 361–369.
- Roff, D.A., Prokkola, J.M., Krams, I., Rantala, M.J., Dochtermann, N.A., et al. 2000. The evolution of the **G** matrix: selection or drift? *Heredity* **84**: 135–142.
- Roff, D.A., Mousseau, T., Møller, A.P., de Lope, F. & Saino, N. 2004. Geographic variation in the **G** matrices of wild populations of the barn swallow. *Heredity* **93**: 8–14.
- Roff, D.A., Prokkola, J.M., Krams, I. & Rantala, M.J. 2012. There is more than one way to skin a **G** matrix. *J. Evol. Biol.* **25**: 1113–1126.
- Scheuber, H., Jacot, A. & Brinkhof, M. 2003a. Condition Dependence of a Multicomponent Sexual Signal in the Field Cricket *Gryllus campestris*. *Anim. Behav.* **65**: 721–727.
- Scheuber, H., Jacot, A. & Brinkhof, M.W.G. 2003b. The Effect of Past Condition on a Multicomponent Sexual Signal. *Proc. Biol. Sci.* **270**: 1779–1784.
- Schluter, D. 1996. Adaptive Radiation Along Genetic Lines of Least Resistance. *Evolution* **50**: 1766–1774.
- Schluter, D. 2001. Ecology and the Origin of Species. *Trends Ecol. Evol.* **16**: 372–380.
- Shaw, R. & Billington, H. 1991. Comparison of Variance-Components Between 2 Populations of *Holcus lanatus* - a Reanalysis. *Evolution* **45**: 1287–1289.
- Sherrard, M.E., Maherali, H. & Latta, R.G. 2009. Water Stress Alters the Genetic Architecture of Functional Traits Associated with Drought Adaptation in *Avena barbata*. *Evolution* **63**: 702–715.
- Simmons, L. 1995. Correlates of male quality in the field cricket, *Gryllus campestris* L: Age, size, and symmetry determine pairing success in field populations. *Behav. Ecol.* **6**: 376–381.
- Simmons, L. & Ritchie, M. 1996. Symmetry in the songs of crickets. *Proc. R. Soc. Lond., Ser. B: Biol. Sci.* **263**: 1305–1311.
- Simmons, L. & Zuk, M. 1992. Variability in call structure and pairing success of male field crickets, *Gryllus bimaculatus* - the effects of age, size and parasite load. *Anim. Behav.* **44**: 1145–1152.
- Simmons, L., Zuk, M. & Rotenberry, J. 2001. Geographic variation in female preference functions and male songs of the field cricket *Teleogryllus oceanicus*. *Evolution* **55**: 1386–1394.
- Steppan, S., Phillips, P. & Houle, D. 2002. Comparative quantitative genetics: evolution of the **G** matrix. *Trends Ecol. Evol.* **17**: 320–327.
- Stirling, G. & Roff, D. 2000. Behaviour plasticity without learning: phenotypic and genetic variation of naive *Daphnia* in an ecological trade-off. *Anim. Behav.* **59**: 929–941.
- Tolle, A.E. & Wagner, W.E. 2011. Costly signals in a field cricket can indicate high- or low-quality direct benefits depending upon the environment. *Evolution* **65**: 283–294.
- Turelli, M. 1988. Phenotypic Evolution, Constant Covariances, and the Maintenance of Additive Variance. *Evolution* **42**: 1342–1347.
- Venables, W.N. & Ripley, B.D. 2002. *Modern Applied Statistics with S*. 4th edn. Springer, New York.
- Wagner, G. & Altenberg, L. 1996. Perspective: Complex adaptations and the evolution of evolvability. *Evolution* **50**: 967–976.
- Wagner, W. & Hoback, W. 1999. Nutritional effects on male calling behaviour in the variable field cricket. *Anim. Behav.* **57**: 89–95.

- Wedekind, C. 1992. Detailed Information About Parasites Revealed by Sexual Ornamentation. *Proc. R. Soc. Lond., Ser. B: Biol. Sci.* **247**: 169–174.
- West-Eberhard, M.J. 2003 *Developmental plasticity and evolution*. Oxford University Press, USA.
- Zahavi, A. 1975. Mate selection – a selection for a handicap. *J. Theor. Biol.* **53**: 205–214.
- Zajitschek, F., Hunt, J., Jennions, M.D., Hall, M.D. & Brooks, R.C. 2009. Effects of juvenile and adult diet on ageing and reproductive effort of male and female black field crickets, *Teleogryllus commodus*. *Funct. Ecol.* **23**: 602–611.
- Zelditch, M.L., Bookstein, F.L. & Lundrigan, B.L. 1992. Ontogeny of Integrated Skull Growth in the Cotton Rat *Sigmodon fulviventer*. *Evolution* **46**: 1164–1180.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Appendix S1 Additional Statistics for comparing P for advertisement call structure in *T. commodus* when taking error in P into account.

Appendix S2 Comparison of P for Life-History Traits in *T. commodus* from Experiment 2.

Data deposited at Dryad doi:10.5061/dryad.h1v65

Received 25 September 2012; revised 3 January 2013; accepted 4 January 2013