

Correlated evolution of multivariate traits: detecting co-divergence across multiple dimensions

E. J. MÁRQUEZ & L. L. KNOWLES

Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, USA

Keywords:

constraint;
correlated evolution;
covariation;
functionally coupled traits;
genitalia;
multidimensional traits;
multivariate evolution;
sexual selection.

Abstract

Tests of correlated evolution typically treat phenotypic characters as univariate variables, even though different trait attributes may contribute to their association with other traits. In this study, patterns of character covariation among species are analysed in a multivariate framework to test for both correlated rates and directions of evolutionary change in traits forming the genitalic complex of male grasshoppers. Although the covariation structure differs among traits, and among the constituent species of two grasshopper clades, significant co-divergence was detected among the most closely interacting genitalic traits (i.e. intromittent characters) in both clades. Co-divergence across shape space is not accompanied by similar rates of evolution among species, although the intromittent characters tend to show accelerated evolution (relative to nonintromittent characters). Differences in the evolutionary trajectories among traits may relate to their varied roles during mating. The study emphasizes the importance of a multivariate framework for detecting macroevolutionary patterns of correlated change.

Introduction

Correlated evolutionary divergence among species is expected for functionally coupled traits, whose individual performances depend on their mutual interactions (e.g. vertebrate jaws; Olson & Miller, 1958; Berg, 1960; Cheverud, 1982; Burger, 1986; Wagner, 1988; Armbruster & Schwaegerle, 1996; Leamy *et al.*, 1999). Concerted evolution is typically measured as the extent to which traits diverge jointly, for which bivariate statistics, such as correlations, partial correlations and covariances, are commonly used (e.g. Olson & Miller, 1958; Cheverud, 1982; Magwene, 2001; Armbruster *et al.*, 2004). However, phenotypic traits are rarely, if ever, truly unidimensional features (see Bookstein *et al.*, 2003; Klingenberg *et al.*, 2003; Bastir & Rosas, 2005; Marguán-Lobón & Buscaglioni, 2006), and treating them as such carries the hidden assumption that the covariation among co-diverging traits also tracks along a single dimension across all sampled species. Yet, the interactions

between two multivariate traits can actually span multiple dimensions (as discussed below). To avoid potentially misleading predictions about the directions of phenotypic change (e.g. Schluter, 1996), and interpretations of constraint on evolutionary trajectories (e.g. Hansen *et al.*, 2003), detecting co-divergence in multivariate traits among species requires considering the multiple dimensions in which correlated evolution might occur.

To illustrate the potential consequences of not taking into account the multiple dimensions relevant to correlated evolutionary divergence among phenotypic traits, consider the following hypothetical scenarios. With the relatively simple shapes of two co-diverging traits (Fig. 1a), correlations between univariate measures obtained from different traits (i.e. between y , x_1 and x_2 ; Fig. 1b) may be relatively high. However, when variables measured on each trait are analysed in a multivariate framework (i.e. combining x_1 and x_2 in a single factor), a substantial increase in the correlation between the two traits is obtained (Fig. 1c) because the simplest trait (i.e. the triangle) has not actually co-diverged with any single aspect of the more complex trait, but instead with a combination of its features. As is evident in Fig. 2, as the shape of a trait becomes more complex – and thus increasing the available directions of co-divergence – the

Correspondence: L. Lacey Knowles, Department of Ecology and Evolutionary Biology, University of Michigan, MI, Ann Arbor, USA.
Tel.: 1 734 763 5603; fax: 1 734 763 4080;
e-mail: knowlesl@umich.edu

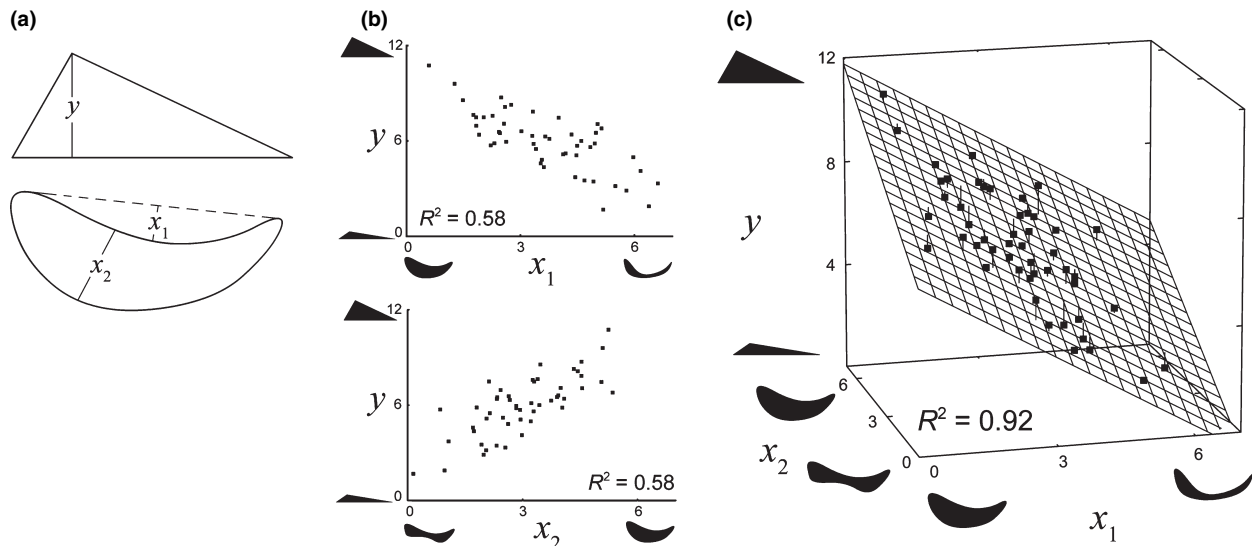


Fig. 1 Contrast between bivariate and multivariate tests of association between two hypothetical traits. (a) Univariate measurements $-y$, x_1 and x_2 , and (b) the association between the two traits was measured (Pearson correlation coefficients are $r_{x_1,y} = -0.76$ and $r_{x_2,y} = 0.76$), where each point represents a species. Illustrations next to the axes represent how the traits differ, as implied by variation of each of the measurements. (c) A trivariate plot shows how the association between the two traits is actually better fit by a plane (correlation from a 2B-PLS multivariate analysis is $r = 0.96$), showing that y -values can co-evolve independently or simultaneously with both x_1 and x_2 -values.

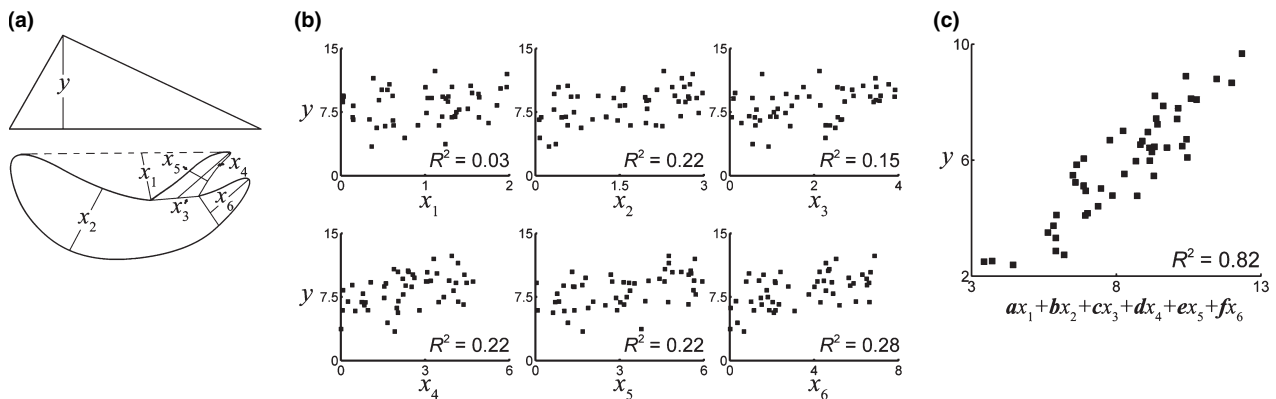


Fig. 2 Contrast between bivariate and multivariate tests of association when one of the two hypothetical traits is quite complex. (A) Univariate measurements were considered $-y$ and x_1 to x_6 , and (B) separate tests of the association between the two traits were conducted for each measurement (Pearson correlation coefficients are $r_{x_1,y} = 0.18$, $r_{x_2,y} = 0.47$, $r_{x_3,y} = 0.39$, $r_{x_4,y} = 0.47$, $r_{x_5,y} = 0.47$, and $r_{x_6,y} = 0.53$). (C) Multivariate combination of variables x_1, \dots, x_6 plotted against variable y ; this particular combination accounts for the highest portion of the covariance between the two traits, as computed using 2B-PLS (correlation coefficient $r = 0.90$), and represents a six-dimensional space (i.e. all six variables are relevant for the association between these traits).

results from tests of trait association from any single pair of univariate variables may not indicate any strong evidence of correlated change between the traits across species (Fig. 2b), when in fact a strong association may exist (Fig. 2c).

These examples illustrate how increased dimensionality of individual traits will generally tend to lower the correlation expected between any single pair of univariate measurements between the traits, even if the traits do indeed exhibit a pattern of co-divergence among

species (e.g. Fig. 1 compared with Fig. 2). This is because when conducting tests of correlated divergence between traits, it is the combination of features, not individual aspects of a trait, which must be considered as there are many different ways (i.e. different shapes) in which a trait might diverge in response to divergence in another trait to maintain an association. As the dimensionality of the individual traits increases, so too does the diversity of responses across different species lineages (formally referred to as dimensionality of the co-divergence),

making it more likely that tests of correlated evolution will be misleading if the multiple dimensions of character covariation are not taken into account.

The present contribution deals with the methodological issue of testing for correlated evolution among multivariate traits. In addition to: (i) estimating shared directionality of variation among multivariate traits, we also (ii) explicitly consider the extent of shared magnitude of correlated change, to examine whether characters that diverge along a common trajectory also diverge at similar rates of evolution. This approach to comparative study of multivariate traits complements other multivariate tests for correlated evolution in several significant ways. Traditional multivariate methods (e.g. two-block partial least squares and canonical correlation analysis, CCA; Bookstein, 1986; Rohlf & Corti, 2000) identify sets of dimensions that capture the greatest covariation among multivariate traits, and as such focus on detecting the directions of trait co-divergence among species. However, shared magnitude of divergence (as illustrated in this study) is also a valuable indicator of the presence of correlating factors in species diversification. For example, different phenotypic characters may evolve at similar rates, irrespective of whether the traits diverge in similar directions. Likewise, the evolutionary consequences of whether characters evolve at similar rates among species might be of interest (e.g. whether or not correlated rates of divergence are associated with accelerated amounts of characters change among species).

To assess shared directionality of divergence across traits, two related approaches are used in the present study: a multidimensional extension of covariance and a test of congruence among intertaxon distance matrices (Monteiro *et al.*, 2005). Both of these approaches ask whether interspecific differences accumulate in a similar manner in two multivariate traits; however, by standardizing traits by their disparity values in the latter case, the impact of trait disparity on estimates of trait correlation can be explored. On the other hand, shared magnitude of correlated evolution among multivariate traits is estimated as the extent to which two traits depart similarly from a reference value (e.g. the centroid of the trait space or a trait's ancestral value), herein termed 'co-disparity'. These approaches are used to test for correlated evolution among five individual components of a functional trait complex, specifically the male genitalia in species of montane grasshoppers, genus *Melanoplus*. In addition to being well suited for the study of correlated evolution because of their functional relationships, these traits are particularly interesting because genitalia are also the primary site of interspecific divergence in these taxa (Deyrup, 1996; Squitier *et al.*, 1998; Knowles & Otte, 2000; Knowles, 2000) – a pattern that suggests a major role for sexual selection in causing species divergence (Eberhard, 1985, 2004; Arnqvist, 1998; Arnqvist & Rowe, 2002; Hosken & Stockley, 2004). Furthermore, among the sampled structures,

two are particularly close in functional terms, the dorsal and ventral aedeagal valves of the intromittent organ (Whitman & Loher, 1984), and therefore are expected to exhibit stronger patterns of concerted evolution compared with the other two genitalic, nonintromittent traits (the cercus and furcula) (Scudder, 1898).

Results from separate analyses of two closely related *Melanoplus* clades (Knowles & Otte, 2000) illustrate the usefulness of this approach. Namely, the methods provide estimates of different aspects of interspecific associations among traits that can be compiled for a large number of traits. The morphological changes associated with trait co-divergence can be visualized with multi-dimensional plots (e.g. as obtained from 2B-PLS). As predicted, the tight functional relationship between traits of the intromittent organ results in a significant pattern of correlated directionality in divergence. This pattern, however, becomes evident only after separate consideration of congruence and multivariate covariance. Intriguingly, the co-disparity analysis reveals that even though most of the interspecific variation is concentrated in the intromittent organ, most sampled structures show no evidence of correlated magnitudes of their deviations from their means, implying that each structure seems to be diverging at its own rate across the sampled taxa. Differences between the species groups in the patterns of covariation in these high-dimensional traits also suggest that the amount of constraint structuring divergence in these genitalic traits differs among grasshopper clades.

Materials and methods

Species

The focus of this study is on montane *Melanoplus* species that diversified across the sky islands of the northern Rocky Mountains (Table 1). These flightless species are morphologically very similar, except for pronounced differences in the shape of the male genitalia (Knowles, 2000). The females also exhibit very little morphological differences, and some species are virtually indistinguishable; lack of sclerotization of the female genitalic tract makes quantification of potential shape differences extremely difficult. The taxa studied here belong to two species groups, Montanus and Indigens. These species groups overlap in distribution throughout the sky islands of Montana, Idaho, Wyoming and Colorado, and actually co-occur in the same montane meadows. However, the two species groups differ in the number of constituent members, with 28 species in Montanus compared with just nine species in the Indigens species group. Similarity in the degree of genetic distinctiveness of the species from the Montanus and Indigens species groups suggests that the differences in recognized diversity do not reflect an inherent bias of species delimitation based on the phenotypic differences (i.e. genitalic differences) (Knowles, 2001).

Table 1 List of *Melanoplus* species sampled in the present study indicating the clade (species group) to which they belong.

Species	Acronym	Species group
<i>M. dodgei</i>	<i>do</i>	Montanus
<i>M. glymma</i>	<i>gl</i>	Montanus
<i>M. huroni</i>	<i>hu</i>	Montanus
<i>M. hatu</i>	<i>ha</i>	Montanus
<i>M. gothicus</i>	<i>go</i>	Montanus
<i>M. ourayensis</i>	<i>ou</i>	Montanus
<i>M. cumbres lacertus</i>	<i>cl</i>	Montanus
<i>M. cumbres cumbres</i>	<i>cc</i>	Montanus
<i>M. truchensis</i>	<i>tr</i>	Montanus
<i>M. tusharensis</i>	<i>tu</i>	Montanus
<i>M. adapi</i>	<i>ad</i>	Montanus
<i>M. cumbres hister</i>	<i>ch</i>	Montanus
<i>M. gunnisoni</i>	<i>gu</i>	Montanus
<i>M. pahgrense</i>	<i>pa</i>	Montanus
<i>M. washingtonius</i>	<i>wa</i>	Montanus
<i>M. molothrus</i>	<i>ml</i>	Montanus
<i>M. repetinus</i>	<i>re</i>	Montanus
<i>M. oreophilus</i>	<i>or</i>	Montanus
<i>M. montanus</i>	<i>my</i>	Montanus
<i>M. moyense</i>	<i>mn</i>	Montanus
<i>M. crux</i>	<i>cr</i>	Indigens
<i>M. payette</i>	<i>py</i>	Indigens
<i>M. indigens</i>	<i>in</i>	Indigens
<i>M. triangularis</i>	<i>tg</i>	Indigens
<i>M. oregonensis</i>	<i>og</i>	Indigens
<i>M. marshalli</i>	<i>ma</i>	Indigens

These species are members of a radiation of grasshoppers that took place during the dynamic Pleistocene, which has made it difficult to attain reliable estimates of species relationships (Knowles & Otte, 2000; Knowles, 2000; Carstens & Knowles, 2007a,b). Although the phylogenetic relationships among the constituent taxa of the two species groups analysed were not considered here, the relatively large morphological difference in these species is restricted to traits presumably under strong sexual selection, thereby making a large contribution of phylogeny to the observed patterns of divergence less likely (e.g. Rheindt *et al.*, 2004). Moreover, constraints associated with the maintenance of functionality among the individual traits comprising the genitalic complex would be expected to provide a stronger factor in producing correlated evolution than mere phylogenetic 'inertia'.

Data acquisition

Morphometric data were obtained from digital images of dissected male genitalic structures of *Melanoplus* species, acquired using a Leica 3500 digital imaging system (Leica Microsystems, Inc., Bannockburn, IL, USA). A total of 28 species were sampled for the study, including 21 (75%) of the 28 species described in the Montanus species

group, and seven (78%) of the nine species described in the Indigens species group (see Table 1). Because present analyses focused only on the interspecific genitalic variation, one specimen was sampled per species. Five traits were photographed in each species (Fig. 3): the left cercus, right furcula, right dorsal valve of the aedeagus (DAV), left ventral valve of the aedeagus (VAV) and lateral flexion of the aedeagus (LFA). Images of intermittent structures (i.e. the aedeagal valves) were acquired from standardized orientations of the intact endophalli dissected from the genital complex; cerci and furculae were photographed from dissected epiprocts (Fig. 3).

Morphometric analyses were based on an elliptical Fourier analysis of shape outlines (Kuhl & Giardina, 1982; Lestrel, 1997). Digital images were transformed into monochrome silhouettes to facilitate chain coding of outlines (Fig. 3). Shape outlines were manually rotated as commonly used rotation criteria (i.e. first harmonic and longest radius) often produced incorrect superimpositions. The most distal point of each organ with respect to the insect abdomen was consistently used as the starting point for outline computation in all structures. Four coefficients and 20 harmonics were then extracted from shape outlines and treated as shape variables. Chain coding, rotation and computation of harmonics were carried out using the software SHAPE (Iwata & Ukai, 2002).

All analyses were programmed and carried out using MATLAB® (The Mathworks, Inc., 2006). Most of the methods described in the following sections have been compiled in the software CORIANDIS, which is freely available for download at <http://www-personal.umich.edu/~emarquez/morph>.

Measurement error

To minimize sampling error associated with acquiring the digital images, five replicate images were acquired for each specimen and structure; images were taken by the same person, letting a period of at least 24 h to elapse between replicates. The proportion of the total variation due to error was quantified by dividing the trace of the pooled within-specimen covariance matrix, by the trace of the total covariance matrix. In addition, MANOVA tests were performed to assess whether interspecific variation was significantly higher than the measurement error.

MANOVA tests demonstrated significant differences in the measurement error among species (Wilk's Lambda $P \approx 0$). The proportion of error also varied among genitalic structures, ranging from 6% (cerci) to 26% (DAV), which is consistent with the differences in the complexity, and presumably dimensionality, of the genitalic traits. The covariance structure of measurement error was not significantly correlated with the covariance structure of interspecific differences in the Montanus or Indigens species groups. Measurement error was similarly patterned in the Indigens and Montanus groups (matrix correlation $r = 0.94$), despite much less

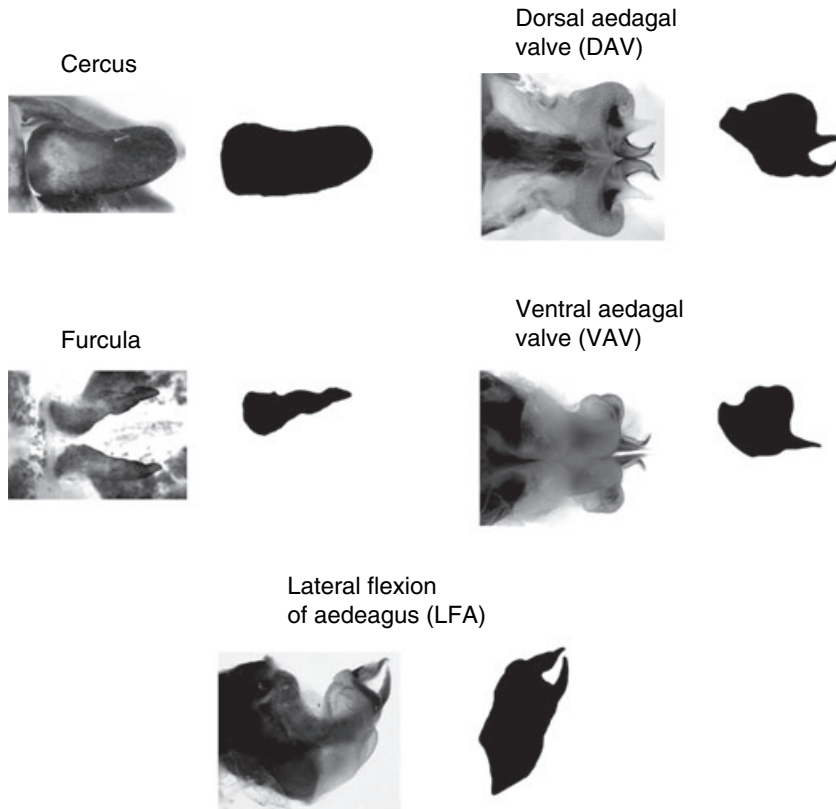


Fig. 3 Genital traits analysed in this study, illustrating the extraction of silhouettes used to identify traits' outlines.

similarity in patterns of interspecific variation between the groups ($r = 0.41$), indicating that observed patterns of divergence are not a result of measurement error. All tests of correlated evolution are based on mean shape values from the five replicates per structure.

Testing for shared directionality of divergence among traits

Congruence analysis

Congruence refers to the similarity among interspecific distance matrices computed from each trait, and it estimates the extent to which pairs of traits have diverged in the same relative directions, irrespective of the magnitude of such divergence. In the absence of trait data for ancestral taxa that allow computing patristic distances, nonpatristic pairwise distance metrics (e.g. Euclidean and Procrustes) have to be used. For the present analyses, matrices of interspecific Euclidean distances were calculated for each genitalic character.

To highlight the relationship between congruence and multivariate covariance, matrix correlations are calculated from the distance matrices for each of the five traits following a methodology described by Abdi *et al.*, (2005), derived from a family of techniques known as ACT (acronym for the French term 'Analyse Conjointe de Tableaux', Lavit *et al.*, 1994) or STATIS (Structuration des

Tableaux A Trois Indices de la Statistique, Escoufier, 1973), that are analogous to Procrustes methods for matrix comparisons (Meyners *et al.*, 2000). In this approach, distance matrices are standardized to remove differences in trait disparity, and then combined in a congruence matrix (**R**) whose entries, also known as R_V coefficients (Robert & Escoufier, 1976), are equivalent to Pearson correlation coefficients between the elements of the distance matrices, and thus measure similarity (proportionality) between the divergence patterns of each pair of traits. Congruence values in **R** were further standardized to remove the effect of the sample size, which is necessary to compare across groups with different numbers of taxa (see Appendix 1 for mathematical details).

Significance of congruence values among traits was assessed using permutation tests (Heo & Gabriel, 1998). To test the null hypothesis that observed congruence values among trait patterns are undistinguishable from random patterns, Mantel permutation tests were used (Dietz, 1983). In these tests, matching rows and columns of one (randomly chosen) of the matrices from each pairwise comparison were randomly permuted. Congruence was computed between the permuted and nonpermuted matrices, and the procedure was repeated 9999 times. These values thus comprise distributions of congruence under the null hypothesis of random

congruence, against which the observed congruence for each pair was compared and rejected when the observed value exceeded 95% of permuted values. *P*-values for this hypothesis were computed as the number of instances in which permuted congruence exceeded observed congruence plus one, divided by the total number of permutations plus one (Dietz, 1983).

Estimation of multivariate covariance

As neither individual shape variables nor equivalently ranked PCs are in general directly comparable across multivariate traits, comparisons across multivariate traits require values of all traits and taxa to be projected onto a space or basis that is common to all traits. In the present approach, such a common basis is obtained from eigen decomposition of the congruence matrix **R**. Because **R** contains information about similarity among traits, its eigenvectors contain information about similarity patterns among those traits (Lavitz *et al.*, 1994). The first eigenvector of matrix **R** can be used to compute a consensus or 'compromise' space in which atypical traits weigh less than more 'normal' ones, providing an appropriate common basis for projecting the distance matrices from individual traits (see Appendix 2). The multivariate covariance (δ_{ij}) between traits *i* and *j* can be computed on this basis using

$$\delta_{ij} = \sum_k \text{cov}(x_{ik}, x_{jk})$$

where x_{ik} represents the species scores of trait *i* on the *k*th principal component of the compromise space, and the summation is carried out over all principal components. If estimates of ancestral phenotypes were available, they could be readily incorporated in this procedure and used instead of the traits centroid or mean in the computation of covariances.

Permutation tests (Dietz, 1983) were used to test the null hypothesis that co-disparity values were produced by chance alone. In this case, rows (i.e. specimens), but not columns (i.e. principal components), were permuted from the matrices of species scores projected onto the compromise space. In each pairwise comparison, only one randomly chosen trait matrix was permuted, thereby randomly relocating each species trait value in trait space. Multivariate covariance was computed among traits in each permutation to generate a null distribution; statistical significance of co-disparity required observed values exceed 95% of permuted values. In these tests, all possible permutations were used for the *Indigenes* group (i.e. 6! permutations from six species), and 9999 random permutations for the *Montanus* group.

Correlated trajectories in trait space

Two-block partial least squares (2B-PLS) is used here to produce visualizations of species ordinations and trait changes associated with those pairs of traits for which a significant correlation was found. 2B-PLS is often used to

estimate and test the significance of linear associations among multivariate traits (Bookstein, 1986; Bookstein *et al.*, 1996; Klingenberg & Zaklan, 2000; Rohlf & Corti, 2000; Bastir & Rosas, 2005; Marugán-Lobón & Buscagliani, 2006). 2B-PLS is an ordination technique, similar to CCA, in which the covariances among two sets of variables are used to obtain two sets of orthogonal vectors, as linear combinations of the original variables, which successively account for higher portions of the covariance among the two sets. However, a limitation of this technique, as with other eigen decomposition methods such as principal component analysis (PCA), is that only the first extracted vector is unambiguously interpretable; additional axes are constrained to be orthogonal to the previous ones. If two or more PLS axes are shown to be significant, the dimensionality of the association cannot be unambiguously equated to the number of significant axes. In this case, correlation coefficients between PLS vectors will be poor approximations of the overall multivariate association between traits because of the embedding of a low-dimensional (and possibly nonlinear) structure onto a higher dimensional subspace (Rohlf & Corti, 2000). By contrast, if only the first PLS axis accounts for a significantly higher proportion of the covariation (i.e. the other axes do not differ from expectations based on random associations), it suggests that traits occupy a unidimensional trajectory in morphospace. This trajectory may indicate the presence of constraining factors on trait co-divergence; however, the position of species along the trajectory does not imply a particular temporal sequence of evolution.

The hypothesis that pairwise trait co-divergence is essentially one-dimensional, and hence divergence among species has occurred along a single trajectory, was addressed by determining whether more than the first PLS axis explained a significant amount of covariance among traits. Significance was assessed by determining whether the magnitude of covariance accounted for by PLS axes could be attained by mere chance. The two highest singular values from a singular value decomposition of the matrix of cross-covariances among two characters (i.e. the first two PLS axes) were computed from 9999 random permutations of each pairwise comparison of traits (Rohlf & Corti, 2000). Tests in which the first (highest), but not the second singular, value was significantly higher than expected by chance were interpreted as supporting a single one-dimensional trajectory capturing most of the covariation between the two traits. Projection of the original variables onto the significant PLS axis was then used to identify the proportion of character divergence among species explained by the single trajectory. This is important because even if a trajectory accounts for a large portion of the covariance between two traits, this covariance may in turn represent a relatively insignificant portion of trait variation. New sets of 9999 permutations were produced to test whether these proportions were higher than

expected by chance. Visualizations of trait divergence along the trajectory were generated using representatives along the first PLS axis that were converted into shape outlines using inverse Fourier transforms, as implemented in the program *SHAPE* (Iwata & Ukai, 2002). The 'PrinPrint' module of this software was used to draw the resulting outline deformations.

Testing for shared magnitude of divergence among traits

Estimation of multivariate 'co-disparity'

Disparity measures the overall magnitude of divergence observed in a group of taxa (Foote, 1997; Eble, 2002; Stayton, 2006), and is expected to be higher in clades containing more extreme forms. Co-disparity, in turn, measures the extent to which two traits show similar magnitudes of departure, irrespective of the direction of such departure, across sampled taxa. In the absence of estimates of ancestral trait values that serve as a reference value to measure magnitude of deviations, co-disparity can be obtained from departures from a fixed quantity, such as the centroid or origin of a centred trait distribution. Here, co-disparity between the *i*th and *j*th traits was computed as the Pearson correlation coefficient between the squared Euclidean distances of individual taxa for each trait from its centroid, so that, for instance, a high positive co-disparity would result from taxa deviating by similar magnitudes from their centroid in both traits simultaneously (squared distances can be obtained as the diagonal elements of matrix S_i in Appendix 1). Thus, congruence compares interspecific distances, whereas co-disparity compares species distances to a fixed value. Standardization of co-disparity values by sample size was done using the same approach used for congruence; for significance testing, observed co-disparities were compared against values obtained from 9999 random permutations.

Results

Shared directionality of divergence among traits

Patterns of trait congruence

Significant trait congruence was detected; however, the pattern of trait congruence differed somewhat between the two *Melanoplus* species groups (Table 2). Standardized congruence is significantly higher than expected by chance among intromittent structures ($P \leq 0.001$), unlike the nonintromittent genitalic structures (i.e. cerci and furculae), in the Montanus species group. This clear difference between intromittent and nonintromittent genitalic traits in degree of congruence was not observed among taxa in the Indigens species group. In these taxa, divergence in the furcula (a nonintromittent trait) was congruent with an intromittent trait (the VAV) ($P = 0.015$) (Table 1).

Table 2 Standardized congruence values among the genitalic structures (below diagonal) and corresponding *P*-values from permutation tests (above diagonal); significant congruence values ($P < 0.05$) are highlighted.

	Cercus	Furcula	DAV	LFA	VAV
Montanus					
Cercus	–	0.283	0.401	0.295	0.301
Furculae	0.028	–	0.104	0.496	0.314
DAV	0.007	0.088	–	0.001	0.0002
LFA	0.032	–0.008	0.288	–	0.0005
VAV	0.030	0.019	0.425	0.289	–
Indigens					
Cercus	–	0.168	0.596	0.568	0.756
Furcula	0.286	–	0.970	0.581	0.015
DAV	–0.134	–0.438	–	0.070	0.932
LFA	–0.102	–0.084	0.463	–	0.466
VAV	–0.173	0.706	–0.350	0.001	–

DAV, dorsal aedeagal valve; LFA, lateral flexion of the aedeagus; VAV, ventral aedeagal valve. Negative standardized congruence values result from raw congruence being lower than the mean congruence from random permutations.

Multivariate covariance among traits

The intromittent traits generally exhibit higher multivariate covariances compared with nonintromittent traits as predicted, although this difference is more pronounced among species from the Indigens group (Table 3). In the Montanus species group, multivariate covariances are also higher among the intromittent traits, although the furcula–dorsal aedeagal valve pair also shows a relatively high covariance (Table 3); all genitalic traits show their highest covariance with the dorsal valve – a result that might be linked to this character exhibiting considerably more disparity than any other trait in the Montanus group.

Table 3 Multivariate covariance values among the genitalic structures (below diagonal), and corresponding *P*-values from permutation tests (above diagonal); significant values ($P < 0.05$) are highlighted.

	Cercus	Furcula	DAV	LFA	VAV
Montanus					
Cercus	–	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Furcula	0.038	–	< 0.0001	< 0.0001	< 0.0001
DAV	0.063	0.136	–	< 0.0001	< 0.0001
LFA	0.045	0.088	0.180	–	< 0.0001
VAV	0.046	0.089	0.199	0.122	–
Indigens					
Cercus	–	< 0.001	< 0.001	< 0.001	< 0.001
Furculae	0.035	–	< 0.001	< 0.001	< 0.001
DAV	0.070	0.071	–	< 0.001	< 0.001
LFA	0.074	0.080	0.227	–	< 0.001
VAV	0.076	0.102	0.174	0.193	–

DAV, dorsal aedeagal valve; LFA, lateral flexion of the aedeagus; VAV, ventral aedeagal valve.

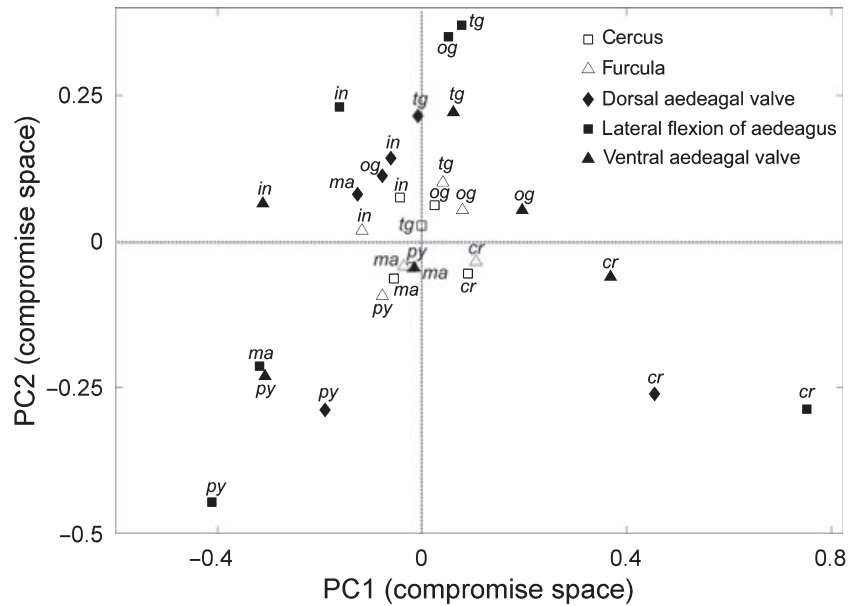


Fig. 4 Projection of genitalic traits of Indigens species group onto the first two principal components of compromise space; see Table 2 for species acronyms.

To explore the relationship between trait congruence and multivariate covariance, trait values for each species are plotted on the compromise space axes for the Indigens group (Fig. 4). This plot shows that patterns of interspecific distances differ among traits. For example, comparison of the three intromittent traits among *Melanoplus oregonensis* (*og*), *Melanoplus triangularis* (*tg*), and *Melanoplus indigens* (*in*) shows that the LFA (black squares) is quite similar between *tg* and *og* but dissimilar to *in*, whereas the DAV (black diamonds) is more similar between *og* and *in*, and rather dissimilar to *tg*; the VAV (black triangles) are approximately equally dissimilar among the three taxa. Patterns such as this contribute to the low congruence observed in the Indigens group, which differs markedly from species in the Montanus group where congruence values are high and significant among intromittent structures. On the other hand, Fig. 4

clearly demonstrates how intromittent organs (represented by black symbols) depart in the same general direction with respect to the centroid of the plot in most species, in agreement with the high covariance estimated for these traits.

Correlated trajectories in trait space

The first PLS axes account for a higher squared covariance than expected by chance among all three intromittent traits in the Montanus group, but only between one nonintromittent trait (the remarkably conserved furcula) and one intromittent trait (VAV) in the Indigens group (Table 4). The squared covariance accounted for by second PLS axes is nonsignificant for all pairs of traits ($P > 0.05$), indicating that significant patterns of correlated change across trait space essentially correspond to divergence along a linear multivariate trajectory. The first

Table 4 Results from pairwise 2B-PLS analyses showing that the first axis explained a significant percentage of the squared covariance of the trait pairs shown; the second axis was nonsignificant in all cases.

Trait pair in 2B-PLS	Explained covariance (%)	P-value	r	P-value	Trait	Explained disparity (%)	P-value
Indigens species group							
Furcula-VAV	81.46	0.0078	0.97	0.0290	Furcula	50.67	0.0096
					VAV	54.52	0.0273
Montanus species group							
DAV-LFA	78.30	0.0007	0.81	0.0065	DAV	47.84	0.2027
					LFA	28.02	0.0142
DAV-VAV	86.23	0.0003	0.80	0.0005	DAV	47.28	0.0994
					VAV	43.48	0.1204
LFA-VAV	81.16	0.0003	0.81	0.0087	LFA	28.51	0.0087
					VAV	43.52	0.2983

Pearson correlation coefficients (*r*) between projections of the original traits onto their corresponding PLS axes and percentage of the disparity of each trait in the set explained by the corresponding PLS axis are also shown.

PLS axes account for about 80% of the squared covariance among the respective traits showing evidence of significant co-divergence (Fig. 5), and correlation coefficients among PLS vectors are relatively high and significant ($r > 0.80$, $P < 0.03$) (Table 4). A significant correlation was not observed among pairs of traits with no significant PLS vectors (results not shown).

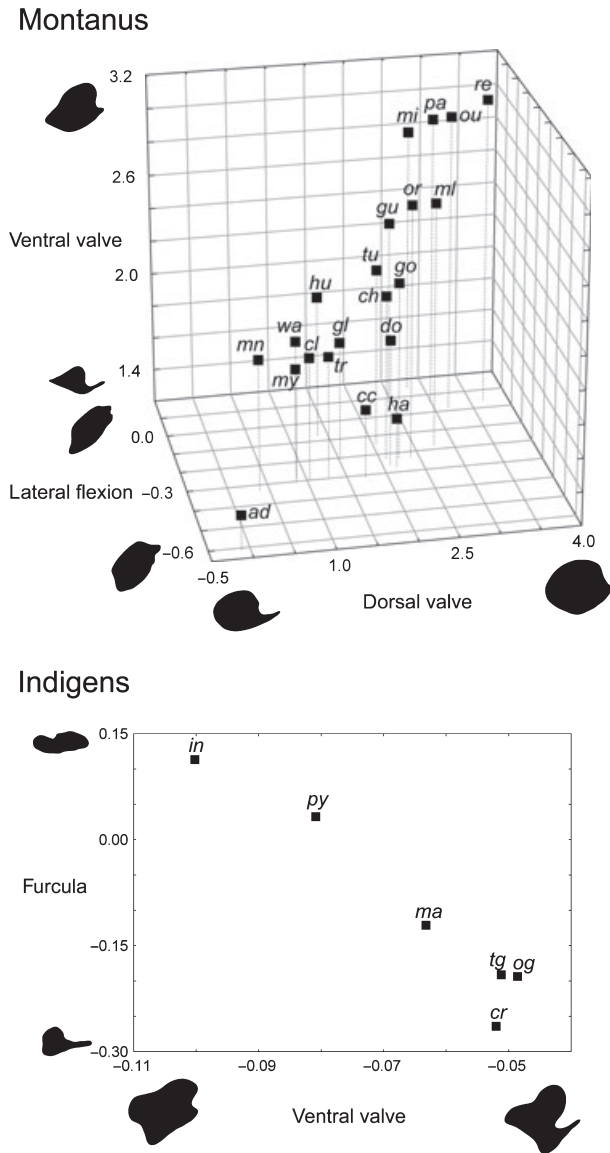


Fig. 5 Plot of significant first partial least squares (PLS) axes for the Montanus and Indigens species groups, where shape silhouettes by each axis represent the extreme shapes of the continuum implied by the axis. Each PLS axis represents the direction of interspecific variation that covaries the most with the other plotted trait(s). In Montanus, results from a combined analysis of the three aedeagal traits using three-block PLS are shown (Bookstein *et al.*, 2003); ordinations obtained from 3B-PLS were identical to those found in individual 2B-PLS.

Trajectories captured by PLS axes account for 51–55% and 28–48% of the disparity of the involved traits in the Indigens and Montanus groups respectively. However, the amount of disparity in an individual trait explained by the linear association with other traits is not significantly higher ($P < 0.05$) than expected for random vectors (Table 4), except for the furcula–ventral valve association in the Indigens group, and the associations with the lateral aedeagal valve in Montanus. For the remaining traits, this result indicates that even if some pairwise associations share more variation than expected by chance alone, the direction of such shared variation may capture a relatively low portion of the variation of each trait separately. In some of these cases (e.g. DAV and VAV in the Montanus group) random vectors can account for over 43% of their disparity, suggesting a remarkably high dimensionality for these genitalic traits.

Testing for shared magnitude of divergence among traits

As predicted, intermittent traits show a remarkably higher disparity than nonintermittent traits (Fig. 6). Co-disparity values, however, do not suggest that intermittent traits generally deviate by similar or proportional magnitudes in individual species (Fig. 7). Permutation tests indicate that in most cases co-disparity values are not higher than expected by chance (Table 5). Exceptions to this result are observed in the Indigens group, where a significant positive co-disparity between the DAV and the LFA ($r = 0.96$, $P = 0.007$), and a significant negative co-disparity between these two traits and the furcula ($r_{F,DAV} = -0.79$, $P = 0.02$, $r_{F,LFA} = -0.81$, $P = 0.01$) were obtained.

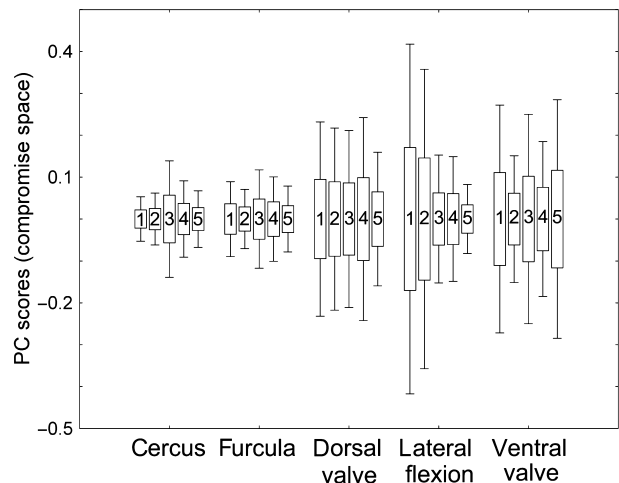


Fig. 6 Total trait variation (i.e. disparity) on first five principal components of compromise space. Boxes represent standard errors of means; whiskers represent standard deviations on each PC. Proportions of variance accounted for each PC are 31.67%, 20.8%, 19.2%, 16.82% and 11.52% respectively.

Fig. 7 Squared Euclidean distances for each trait of each species location in the compromise space from the centroid. Height of a bar segment is proportional to a species departure from the trait mean, and total bar height indicates how much a species deviates from the origin in compromise space. Traits are: cercus (C), furcula (F), dorsal aedeagal valve (DAV), lateral flexion of the aedeagus (LFA) and ventral aedeagal valve (VAV); see Table 2 for species acronyms.

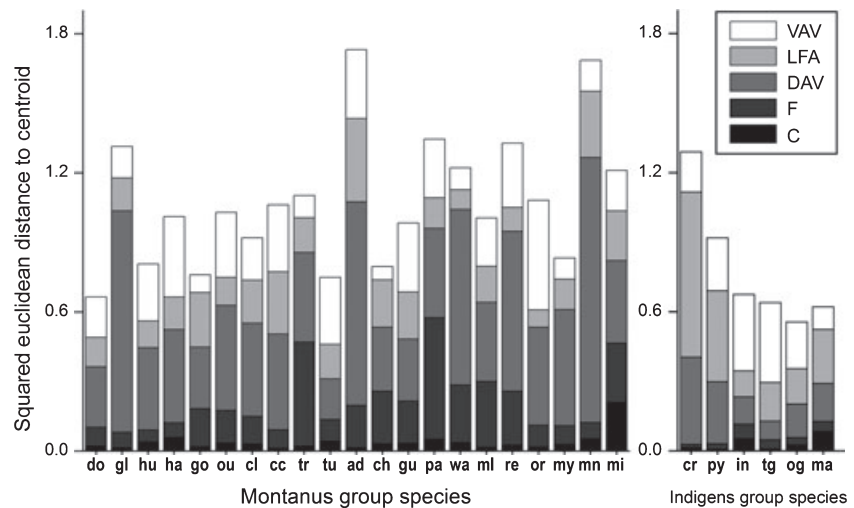


Table 5 Standardized co-disparity values among the genitalic structures (below diagonal), and corresponding *P*-values from permutation tests (above diagonal); significant values ($P < 0.05$) are highlighted.

	Cercus	Furcula	DAV	LFA	VAV
Montanus					
Cercus	–	0.185	0.395	0.316	0.423
Furcula	0.153	–	0.291	0.394	0.194
DAV	–0.101	–0.146	–	0.120	0.306
LFA	0.071	–0.080	0.271	–	0.294
VAV	–0.055	–0.200	–0.122	–0.129	–
Indigens					
Cercus	–	0.110	0.302	0.234	0.199
Furculae	0.612	–	0.020	0.010	0.195
DAV	–0.325	–0.791	–	0.007	0.162
LFA	–0.369	–0.814	0.965	–	0.214
VAV	–0.426	0.432	–0.485	–0.408	–

DAV, dorsal aedeagal valve; LFA, lateral flexion of the aedeagus; VAV, ventral aedeagal valve. Co-disparity values between two traits are computed as the Pearson correlation coefficient between squared Euclidean distances of each species to the centroid, after projecting both traits onto compromise space (see text for details).

Discussion

The question of whether two or more traits diverge in a correlated manner has received considerable attention in both macro- and microevolutionary studies (Olson & Miller, 1958; Piggliucci & Preston, 2004). For example, such tests are relevant to addressing whether there are recurrent patterns of trait association in phenotypic space among species (e.g. Zelditch *et al.*, 1990; Armbruster & Schwaegerle, 1996; Stepan, 1997; Marroig & Cheverud, 2001; Baker & Wilkinson, 2003), or examining the effects of covariation structure on indirect (correlated) phenotypic responses to directional selection within a species (e.g. Lande & Arnold, 1983; Cowley & Atchley, 1990; Klingenberg & Leamy, 2001; Merilä & Björklund, 2004).

The methodology illustrated in this study applies to both, and indeed to any questions requiring the estimation of trait covariation patterns. It provides a framework to analyse inherently multidimensional aspects of the phenotype as multivariate variables, and extends analyses of trait association beyond the traditional test of shared directionality of variation to also testing for shared magnitude (or rates) of co-divergence.

The distinction between directionality and magnitude of co-divergence is important because each of these attributes reflects different aspects of the evolutionary associations among the sampled traits. Consistent patterns of deviations among specimens (congruence) or with respect to a fixed reference (covariance and co-disparity) may represent evidence of relatively strong constraints resulting in a common evolutionary trajectory, whereas a shared magnitude of divergence (i.e. co-disparity) may reflect similar rates of divergence among the interacting traits. Although it is possible that traits evolving in concert may diverge along common trajectories and at similar rates, trait associations are generally complex and involve a large number of interactions that may prevent such a correlation pattern from being observed. Furthermore, in functionally coupled characters, different attributes of a trait may change in a compensatory fashion as an evolutionary response to constraint on other attributes imposed by the functional associations among traits. The ability of a phenotypic trait embedded in a complex network of interacting traits to evolve by changing different attributes becomes manifest as an increase in the dimensionality of variation of each trait (Mezey & Houle, 2005). Consequently, a variety of approaches, each assessing a distinct aspect of character associations, may also increase the detectability of correlated evolutionary divergence among multivariate traits, given that the number of ways in which two traits can correlate correspondingly increases with the dimensionality of the traits.

Application of these methods to the analysis of co-divergence among traits of the genitalic complex in the *Melanoplus* grasshoppers illustrates that patterns of correlated evolution span multiple dimensions (Figs 4–6). Although the covariation structure differs among traits and among the constituent taxa in the two grasshopper clades, the most closely interacting characters of the genitalic complex (i.e. the intromittent characters) show a strong tendency to co-diverge in both clades, as expected for functionally coupled traits (Wagner, 1988; Armbruster & Schwaegerle, 1996). This association was observed as a nearly linear trajectory with consistent trait differences among taxa in one of the clades. However, this trajectory only occurs in multidimensional space, and its detection required explicit consideration of the inherent multivariate nature of the phenotypic traits (i.e. the shape of the traits, Fig. 3). The lack of significant patterns of co-disparity among these traits also indicates that despite similar directions of divergence, genitalic characters have diverged at different rates. These results highlight the usefulness of a combined use of different approaches to understand the complex interactions among multivariate traits, and specifically, to detect and characterize patterns of correlated evolutionary divergence.

Multivariate approaches for detecting co-divergence across multiple dimensions

As computed in this study, congruence and multivariate covariance are analogous to metric scaling (Krzanowski, 2000) – analyses are all based on pairwise taxon distances instead of the original phenotypic data. This allows computing correlations and covariances irrespective of the number and type of variables used to define each trait. Nevertheless, the same caveats that apply to distance measures in general (Perry *et al.*, 2002), also apply to the present analyses. For example, collapsing location information of two points onto a single scalar measure (the distance between them) results in a loss of information that may affect correlation estimates because equidistant pairs of taxa may occur in disparate regions of trait space. As this would only influence congruence estimates, the complementary approaches used here should provide a robust test of correlated evolution. Likewise, each of the major quantities estimated in the present approach (congruence, multivariate covariance and co-disparity) are nondimensional descriptors for capturing different aspects of trait co-divergence, which result in a loss of information with respect to the full geometry of multivariate correlations. However, their combined use is potentially more informative in practice because of the difficulties of analysing and visualizing correlation patterns across multiple dimensions.

Estimations of congruence, multivariate covariance and co-disparity are based on divergence from a reference value, whether explicit or implicit. Incorporating phylogenetic information into the estimation of these

statistics – namely, by estimating direction and magnitude of divergence from a common ancestor rather than an arbitrary reference value – accommodates the potential for the magnitude of divergence to differ along individual species lineages, thereby potentially improving the accuracy of tests of correlated evolutionary change. Unfortunately, methodologies for reliably estimating ancestral trait values for shape outline data are currently limited (see Rohlf, 2001).

The lack of phylogenetic independence in tests of correlated trait evolution in comparative analyses has been extensively documented: phylogenetic proximity can induce interspecific correlations among otherwise independently diverging traits, an issue usually dealt with by computing phylogenetic contrasts (Felsenstein, 1985; Martins & Hansen, 1997; Rohlf, 2001, 2006). However, such methods are not available for shape data. Such phylogenetic inertia is also not expected to dominate species differences if the traits under study are subject to strong diversifying selection (e.g. Rheindt *et al.*, 2004), as is most likely the case for *Melanoplus* male genitalia.

Implications of the patterns of correlated evolution in *Melanoplus*

The pronounced difference in the shape of the genitalic complex among *Melanoplus* species is not observed among nongenitalic structures (Deyrup, 1996; Squitier *et al.*, 1998; Knowles & Otte, 2000; Knowles, 2000), suggesting an important role for sexual selection (Eberhard, 1985, 2004; Arnqvist, 1998; Arnqvist & Rowe, 2002; Hosken & Stockley, 2004). Although the functional significance of shape variation of the genitalic traits has not been investigated in *Melanoplus* specifically, studies have identified that the shape of the genitalia in insects relates to various aspects of copulatory behaviour. This includes the transfer and manipulation of spermatophores to the female reproductive tract and physical retention or stabilization of the female body during copula, as in the acridoid genus *Taeniopoda* (Whitman & Lohrer, 1984), and in some cases, the actual ejection of other males' spermatophores from the female's sperm storage organs (the spermatheca), as in calopterygid damselflies (Córdoba-Aguilar, 2005).

Consideration of the results of the study, and specifically differences in the pattern of covariation among the genitalic traits and between the two species groups, provides some insights into how sexual selection may operate. The higher disparity among the intromittent compared with the nonintromittent characters (Fig. 6) suggests that the strength of selection may relate to the function of the traits in the genitalic complex (Arnqvist, 1998). The components comprising the intromittent organ (Scudder, 1898) are involved in intricate internal interactions with females and these traits also show corresponding evidence of integration (Fig. 5). This

contrasts with traits that interact with external features of the females (the nonintromittent characters of the genitalic complex) which may function in the stabilization of the sexes during copula (or possibly tactile stimulation; see Eberhard, 1996), and do not tend to show strong evidence of correlated divergence. These results indicate the genitalic complex is not subject to selection as a single unit, but rather depends on the form of engagement of the traits during interactions with females. One implication of this study is that the intromittent traits may experience stronger selection (i.e. intromittent traits exhibit greater disparity), and that the selection appears to be divorced from the selection operating on the external organs (i.e. there is no consistent strong pattern of correlated evolution between intromittent and nonintromittent organs).

Although the intromittent characters tend to evolve in concert (i.e. diverge in the same direction) in both the Montanus and Indigenus species groups (Table 2, Figs 4 and 5), divergence does not appear to be highly constrained. For example, co-divergence among species in the Montanus group occurs along a single linear trajectory (suggesting the pattern of correlated evolution is constrained). However, the magnitudes of divergence in these traits are not correlated – there is no significant co-disparity, despite a common direction of shape change (Fig. 7). In fact, despite the significant congruence of trait divergence across species (Table 4), the vectors do not explain a significant amount of variance observed in many of the traits. This suggests that there is large component of variation contributing to observed species differences in these traits unrelated to their functional integration with other genitalic traits. The apparent contradiction between independent trait divergence (Table 4) with 2B-PLS and congruence analyses that suggest correlated evolution (Fig. 4) is consistent with a scenario where genitalic traits occupy a highly multidimensional space containing numerous nonorthogonal trajectories of co-divergence. Such a pattern could also be expected if the traits had multiple functional associations with characters not included in the present study (Hansen, 2003).

The extreme diversity of insect genitalia is generally interpreted as evidence for the sexual selection. Irrespective of whether divergence is caused by cryptic female choice, sexual conflict or some form of sperm competition (Simmons, 2001), the adaptive landscape of these sexually selected traits is arguably characterized by directional selection (Blows *et al.*, 2004), albeit it may be a constantly shifting landscape given that sexual selection on genitalia inevitably involves interactions between the sexes. The multivariate covariance analyses (Figs 3 and 4; Table 3) generally support the expectation of concerted evolution in the functionally coupled traits (Cheverud, 1982; Wagner, 1988), with much higher covariance values among the intromittent than non-intromittent traits in *Melanoplus*. However, the analyses

reveal clear differences between the two grasshopper clades in the pattern of trait covariation. Despite shared directionality of co-divergence in genitalic characters in both groups (Tables 2 and 3), the diverging species in the Indigenus group occupy a proportionally greater extent of morphospace (data not shown) compared with taxa in the Montanus group whose divergence has been more constrained along a linear multivariate trajectory (Fig. 5). The particular mechanisms underlying the differing levels of constraint between the Montanus and Indigenus species groups are not clear and require further investigation. However, this difference raises some intriguing hypotheses that could be explored and has broad implications for the role of sexual selection in promoting rapid divergence in these grasshoppers. Namely, the presence of constraints could produce a morphological 'corridor' or ridge (Cheverud, 1982; Wagner, 1988) along which species might diverge. Such ridges could facilitate and speed up divergence by providing readily accessible novel morphological combinations (Burger, 1986) and by providing a link between the 'evolvability' of these characters and other differences between the species groups, such as the differing number of species among clades of the *Melanoplus* grasshoppers.

We approached a test of correlated evolution considering the multi-faceted nature of concerted evolution among multivariate traits – considering covariance, congruence and co-disparity of genitalic traits. It is noteworthy that the complement of approaches was requisite to detecting the different features of trait covariation. This emphasizes the importance of considering how the dimensionality of interactions needs to be considered with tests of correlated evolution (including approaches that would extend the aspects of concerted evolution considered here).

Conclusions

The study of correlated evolution requires careful consideration of the dimensionality of character variation given the inherent multivariate nature of traits. Otherwise, nonsignificant associations among traits may be due to a failure to sample the relevant dimensions. Incomplete sampling of trait dimensions, for instance, may be one possible explanation for the apparent instability of genetic covariance matrices during evolutionary divergence (Steppan *et al.*, 2002). Because co-divergence of multivariate traits can occur across multiple dimensions, combined analyses that can capture different aspects of trait correlation (e.g. detecting shared directionality and/or similar magnitudes of co-divergence) complement more traditional multidimensional approaches to the study of shape covariation (e.g. Klingenberg & Zaklan, 2000; Rohlf & Corti, 2000). Tests of correlated evolution in the genitalic complex in *Melanoplus* grasshoppers reveal differences in covariation among traits and between clades. These comparative

analyses identified different levels of integration, as well as constraint, with interesting implications for how divergence has proceeded in the genitalic traits. Specifically, divergence in these sexually selected traits not only has to conform to the constraints imposed by the functional integration of the genitalic traits, but there is also a large component of variation underlying the rapid divergence of these traits unrelated to their concerted evolution.

Acknowledgments

We thank D. Otte for generously providing some of the specimens for the study as a loan from the museum collections of the Academy of Natural Sciences, Philadelphia, and M. Zelditch and two anonymous reviewers for their helpful comments. The research was funded by the following awards: to LLK: a National Science Foundation grant (DEB-04-47224), the Elizabeth Caroline Crosby Fund, NSF ADVANCE Project, University of Michigan, and a grant from the University of Michigan (Office of the Vice President for Research); to EJM: an NSF Doctoral Dissertation Improvement Grant (F010527-047887).

References

- Abdi, H., Valentin, D., O'Toole, A.J. & Edelman, B. 2005. DISTATIS: The analysis of multiple distance matrices. In: *Proceedings of the IEEE Computer Society: International Conference on Computer Vision and Pattern Recognition, San Diego*, pp. 42–47.
- Armbruster, W.S. & Schwaegerle, K.E. 1996. Causes of covariation of phenotypic traits among populations. *J. Evol. Biol.* **9**: 261–276.
- Armbruster, W.S., Pélabon, C., Hansen, T.F. & Mulder, C.P.H. 2004. Floral integration, modularity, and accuracy. In: *Phenotypic Integration* (M. Pigliucci & K. Preston, eds), pp. 23–49. Oxford University Press, Oxford.
- Arnqvist, G. 1998. Comparative evidence for the evolution of genitalia by sexual selection. *Nature* **393**: 784–786.
- Arnqvist, G. & Rowe, L. 2002. Correlated evolution of male and female morphologies in water striders. *Evolution* **56**: 936–947.
- Baker, R.H. & Wilkinson, G.S. 2003. Phylogenetic analysis of correlation structure in stalk-eyed flies (*Diasemopsis*, Diopsidae). *Evolution* **57**: 87–103.
- Bastir, M. & Rosas, A. 2005. Hierarchical nature of morphological integration and modularity in the human posterior face. *Am. J. Phys. Anthropol.* **128**: 26–34.
- Berg, R.L. 1960. The ecological significance of correlation Pleiades. *Evolution* **14**: 171–180.
- Blows, M.J., 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.
- Bookstein, F.L. 1986. The elements of latent variable models: a cautionary lecture. In: *Advances in Developmental Psychology*, Vol. 4 (M. E. Lamb, A. L. Brown & B. Rogoff, eds), pp. 203–230. Laurence Erlbaum, Hillsdale, NJ.
- Bookstein, F.L., Sampson, P.D., Streissguth, A.P. & Barr, H.M. 1996. Exploiting redundant measurement of dose and developmental outcome: new methods from the behavioral teratology of alcohol. *Dev. Psychol.* **32**: 404–415.
- Bookstein, F.L., Gunz, P., Mitteroecker, P., Prossinger, H., Schaefer, K. & Seidler, H. 2003. Cranial integration in *Homo*: singular warps analysis of the midsagittal plane in ontogeny and evolution. *J. Hum. Evol.* **44**: 167–187.
- Burger, R. 1986. Constraints for the evolution of functionally coupled characters: a nonlinear analysis of a phenotypic model. *Evolution* **40**: 182–193.
- Carstens, B.C. & Knowles, L.L. 2007a. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from *Melanoplus* grasshoppers. *Syst. Biol.* **56**: 400–411.
- Carstens, B.C. & Knowles, L.L. 2007b. Shifting distributions and speciation: species divergence during rapid climate change. *Mol. Ecol.* **16**: 619–627.
- Cheverud, J.M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. *Evolution* **36**: 499–516.
- Córdoba-Aguilar, A. 2005. Possible coevolution of male and female genital form and function in a calopterygid damselfly. *J. Evol. Biol.* **18**: 132–137.
- Cowley, D.E. & Atchley, W.R. 1990. Development and quantitative genetics of correlation structure among body parts of *Drosophila melanogaster*. *Am. Nat.* **135**: 242–268.
- Deyrup, M. 1996. Two new grasshoppers from relict uplands of Florida (Orthoptera: Acrididae). *Trans. Am. Entomol. Soc.* **122**: 199–211.
- Dietz, E.J. 1983. Permutation tests for association between two distance matrices. *Syst. Zool.* **32**: 21–26.
- Eberhard, W.G. 1985. *Sexual Selection and Animal Genitalia*. Harvard University Press, Cambridge.
- Eberhard, W.G. 1996. *Female Control: Sexual Selection by Cryptic Female Choice*. Princeton University Press, Princeton, NJ.
- Eberhard, W.G. 2004. Rapid divergent evolution of sexual morphology: comparative tests of antagonistic coevolution and traditional female choice. *Evolution* **58**: 1947–1970.
- Eble, G.J. 2002. Multivariate approaches to development and evolution. In: *Human Evolution through Developmental Change* (N. Minugh-Purvis & K. J. McNamara, eds), pp. 51–78. John Hopkins University Press, Baltimore, MD.
- Escoufier, Y. 1973. Le traitement des variables vectorielles. *Biometrics* **29**: 751–760.
- Felsenstein, J. 1985. Phylogenies and the comparative method. *Am. Nat.* **125**: 1–15.
- Foote, M. 1997. The evolution of morphological diversity. *Annu. Rev. Ecol. Syst.* **28**: 128–152.
- Gower, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika* **53**: 325–338.
- Hansen, T.F. 2003. Is modularity necessary for evolvability? Remarks on the relationship between pleiotropy and evolvability. *BioSystems* **69**: 83–94.
- 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. (Mol. Dev. Evol.)* **296B**: 23–39.
- Heo, M.S. & Gabriel, K.R. 1998. A permutation test of association between configurations by means of the RV coefficient. *Comm. Stat. – Simul. Comput.* **27**: 843–856.

- Hosken, D.J. & Stockley, P. 2004. Sexual selection and genital evolution. *Trends Ecol. Evol.* **19**: 87–93.
- Iwata, H. & Ukai, Y. 2002. SHAPE: a computer program package for quantitative evaluation of biological shapes based on elliptic Fourier descriptors. *J. Hered.* **93**: 384–385.
- Klingenberg, C.P. & Leamy, L.J. 2001. Quantitative genetics of geometric shape in the mouse mandible. *Evolution* **55**: 2342–2352.
- Klingenberg, C.P. & Zaklan, S.D. 2000. Morphological integration between developmental compartments in the *Drosophila* wing. *Evolution* **54**: 1273–1285.
- Klingenberg, C.P., Mebus, K. & Auffray, J.-C. 2003. Developmental integration in a complex morphological structure: how distinct are the modules in the mouse mandible? *Evol. Dev.* **5**: 522–531.
- Knowles, L.L. 2000. Tests of Pleistocene speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of Western North America. *Evolution* **54**: 1337–1348.
- Knowles, L.L. 2001. Genealogical portraits of speciation in montane grasshoppers (genus *Melanoplus*) from the sky islands of the Rocky Mountains. *Proc. R. Soc. Lond. Ser. B.* **268**: 319–324.
- Knowles, L.L. & Otte, D. 2000. Phylogenetic analysis of montane grasshoppers from Western North America (genus *Melanoplus*, Acrididae: Melanoplinae). *Ann. Entomol. Soc. Am.* **93**: 421–431.
- Krzyszowski, W.J. 2000. *Principles of Multivariate Analysis: A User's Perspective*. Oxford University Press, Oxford.
- Kuhl, F.P. & Giardina, C.R. 1982. Elliptic Fourier features of a closed contour. *Comput. Graph. Image Process.* **18**: 236–258.
- Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210–1226.
- Lavit, C., Escoufier, Y., Sabatier, R. & Traissac, P. 1994. The ACT (STATIS method). *Comput. Stat. Data Anal.* **18**: 97–119.
- Leamy, L.J., Routman, E.J. & Cheverud, J.M. 1999. Quantitative trait loci for early- and late-developing skull characters in mice: a test of the genetic independence model of morphological integration. *Am. Nat.* **153**: 201–214.
- Lestrel, P.E. 1997. *Fourier Descriptors and their Applications in Biology*. Cambridge University Press, Cambridge.
- Magwene, P.M. 2001. New tools for studying integration and modularity. *Evolution* **55**: 1734–1745.
- Marroig, G. & Cheverud, J.M. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of new world monkeys. *Evolution* **55**: 2576–2600.
- Martins, E.P. & Hansen, T.F. 1997. Phylogenies and the comparative method: a general approach to incorporating phylogenetic information into the analysis of interspecific data. *Am. Nat.* **149**: 646–667.
- Marugán-Lobón, J. & Buscaglioni, A.D. 2006. Avian skull morphological evolution: exploring exo- and endocranial covariation with two-block partial least squares. *Zoology* **109**: 217–230.
- Merilä, J. & Björklund, M. 2004. Phenotypic integration as a constraint and adaptation. In: *Phenotypic Integration* (M. Pigliucci & K. Preston, eds), pp. 107–129. Oxford University Press, Oxford.
- Meyners, M., Kuners, J. & Qannari, E.M. 2000. Comparing generalized procrustes analysis and STATIS. *Food Qual. Pref.* **11**: 77–83.
- Mezey, J.G. & Houle, D. 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. *Evolution* **59**: 1027–1038.
- Monteiro, L.R., Bonato, V., dos Reis, S.F. 2005. Evolutionary integration and morphological diversification in complex morphological structures: mandible shape divergence in spiny rats (Rodentia, Echimyidae). *Evol. Dev.* **7**: 429–439.
- Olson, E.C. & Miller, R.L. 1958. *Morphological Integration*. Chicago University Press, Chicago, IL.
- Perry, J.N., Liebhold, A.M., Rosenberg, M.S., Dungan, J., Miriti, M., Jakomulska, A. & Citron-Pousty, S. 2002. Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data. *Ecography* **25**: 578–600.
- Pigliucci, M. & Preston, K. 2004. *Phenotypic Integration*. Oxford University Press, Oxford.
- Rao, C.R. & Rao, M.B. 1998. *Matrix Algebra and Its Applications to Statistics and Econometrics*. World Scientific, London.
- Rheindt, F.E., Grafe, T.U. & Abouheif, E. 2004. Rapidly evolving traits and the comparative method: how important is testing for phylogenetic signal? *Evol. Ecol. Res.* **6**: 377–396.
- Robert, P. & Escoufier, Y. 1976. A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Appl. Stat.* **25**: 257–265.
- Rohlf, F.J. 2001. Comparative methods for the analysis of continuous variables: geometric interpretations. *Evolution* **55**: 2143–2160.
- Rohlf, F.J. 2006. A comment on phylogenetic correction. *Evolution* **60**: 1509–1515.
- Rohlf, F.J. & Corti, M. 2000. Use of two-block partial least-squares to study covariation in shape. *Syst. Biol.* **49**: 740–753.
- Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. *Evolution* **50**: 1766–1774.
- Scudder, S.H. 1898. Revision of the orthopteran group Melanopli (Acrididae) with special reference to North American forms. *Proc. U.S.A. Nat. Mus.* **20**: 1–421.
- Simmons, L.W. 2001. *Sperm Competition and its Evolutionary Consequences in the Insects*. Princeton University Press, Princeton, NJ.
- Squittier, J.M., Deyrup, M. & Capinera, J.L. 1998. A new species of *Melanoplus* (Orthoptera: Acrididae) from an isolated upland in peninsular Florida. *Fla. Entomol.* **81**: 451–460.
- Stanimirova, I., Walczak, B., Massart, D.L., Simeonov, V., Saby, C.A. & Di Crescenzo, E. 2004. STATIS, a three-way method for data analysis. Application to environmental data. *Chemom. Intell. Lab. Syst.* **73**: 219–233.
- Stayton, C.T. 2006. Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. *Evolution* **60**: 824–841.
- Steppan, S.J. 1997. Phylogenetic analysis of phenotypic covariance structure. I. Contrasting results from matrix correlation and common principal component analyses. *Evolution* **51**: 571–586.
- Steppan, S.J., Philips, P.C. & Houle, D. 2002. Comparative quantitative genetics: evolution of the G matrix. *Trends Ecol. Evol.* **17**: 320–327.
- The Mathworks, Inc. 2006. *Matlab R2006a*. The Mathworks, Inc., Natick, MA.
- Wagner, G.P. 1988. The influence of variation and of developmental constraints on the rate of multivariate phenotypic evolution. *J. Evol. Biol.* **1**: 45–66.
- Whitman, D.G. & Loher, W. 1984. Morphology of male sex organs and insemination in the grasshopper *Taeniopoda eques* (Burmeister). *J. Morphol.* **179**: 1–12.

Zelditch, M.L., Straney, D.O., Swiderski, D.L. & Charnichael, A.C. 1990. Variation in developmental constraints in Sigmoidon. *Evolution* **44**: 1738–1747.

Received 16 February 2007; revised 17 May 2007; accepted 18 June 2007

Appendix 1: Estimation of congruence matrix \mathbf{R}

To compute the matrix of congruence values between pairs of multivariate traits, a matrix of Euclidean or similar distances \mathbf{D} is computed for each trait. Each of the $n \times n$ distance matrices \mathbf{D}_i is transformed into a sum of squares and cross-products (SSCP) $n \times n$ matrix \mathbf{S}_i using

$$\mathbf{S}_i = -\frac{1}{2}\mathbf{C}\mathbf{D}_i\mathbf{C}'$$

where $i = 1, \dots, m$ indexes the m traits sampled for each of the n species, \mathbf{C} is an $n \times n$ centring matrix with off-diagonal elements equal to $-1/K$ and diagonal elements equal to $1-1/K$, where K equals the number of sampled taxa, and prime is the transpose symbol (Lavit *et al.*, 1994; Abdi *et al.*, 2005). This transformation is also the central operation in principal coordinate analysis (Gower, 1966). Resulting SSCP matrices contain the same information as the original distance matrices, but can be analysed using eigen decomposition. \mathbf{S}_i matrices were then standardized by total variance (i.e. trace of \mathbf{S}_i), so that congruence among distance patterns ignores differences in trait disparity. Alternatively, matrices could be standardized with respect to their first eigenvalue. Standardized SSCP matrices \mathbf{S}_i^* were then vectorized as columns each with n^2 elements, and concatenated to form an extended $n^2 \times m$ matrix \mathbf{X} , from which an SSCP matrix \mathbf{A} was obtained from

$$\mathbf{A} = \mathbf{X}'\mathbf{X}$$

Matrix \mathbf{A} , whose dimensions are $m \times m$, contains information about similarity among trait divergence patterns and was normalized as

$$\mathbf{R} = \mathbf{A}_{\text{diag}}^{-1/2} \mathbf{A} \mathbf{A}_{\text{diag}}^{-1/2}$$

where \mathbf{A}_{diag} has the same dimensions and diagonal elements as \mathbf{A} , and zeros elsewhere. Entries of the $m \times m$ congruence matrix \mathbf{R} (r_{ij} , also known as R_V coefficients, Robert & Escoufier, 1976) are analogous to Pearson's correlation coefficients between the vectorized forms of \mathbf{S}_i^* matrices for the i th and j th traits, and can range from -1 , indicating perfectly opposite trends between traits, to 1 , indicating perfect congruence (Abdi *et al.*, 2005).

If computed as described above, absolute congruence values (r_{ij}) show a negative logarithmic relationship with the number of taxa used to obtain distance matrices (data not shown). Therefore, congruence values must be standardized by sample size to compare clades with unequal numbers of taxa. In this study, standardized congruence values r_{ij}^* between traits i and j were obtained using the rescaling equation

$$r_{ij}^* = \frac{r_{ij} - r_{ij}^0}{1 - r_{ij}^0}$$

where r_{ij} is the observed congruence value from matrix \mathbf{R} , and r_{ij}^0 is the expected congruence value under the null hypothesis of no congruence. The latter value was obtained from the same permutation analyses used to test for significance of congruence (see Materials and methods), as the mean congruence of the randomly permuted data sets, for each pair of traits. Rarefaction of the largest (i.e. Montanus) data set showed that this standardization procedure does not alter congruence patterns among traits (Pearson correlation coefficient among the standardized congruence matrices of 19 rarified data sets containing from two to 20 taxa averages 0.96).

Appendix 2: Construction of compromise space for computation of multivariate covariances

As all elements in the unstandardized congruence matrix were positive in both species groups, coefficients in the first eigenvectors of these matrices are suitable for providing weights in the construction of the compromise space (Rao & Rao, 1998). Normalized coefficients of the first eigenvector of \mathbf{R} were used as weights w_i to produce the compromise SSCP matrix \mathbf{Z} , which is the best approximation to interspecific distances using the information contained in all traits simultaneously. The weights used for the cercus, furcula, DAV, LFA, and VAV were 0.182, 0.172, 0.212, 0.220, 0.213 for Montanus, and 0.198, 0.209, 0.199, 0.190, 0.204 for Indigenus respectively. The compromise was then computed using

$$\mathbf{Z} = \sum_i w_i \mathbf{S}_i^*$$

(Lavit *et al.*, 1994; Stanimirova *et al.*, 2004). \mathbf{Z} , an $n \times n$ matrix, can be interpreted as a weighted average of the standardized trait covariances, in which those traits that depart more from the rest are given a lower weight. \mathbf{Z} also provides a mathematically convenient construct for projecting the distance matrices from individual traits onto a common basis. To project each \mathbf{S}_i matrix onto this compromise space, $\mathbf{S}_i \mathbf{V} \mathbf{A}^{-1/2}$, where \mathbf{V} and \mathbf{A} are the normalized eigenvectors and the diagonal matrix of eigenvalues of \mathbf{Z} , respectively, was used (Abdi *et al.*, 2005). Note that the nonstandardized matrices, \mathbf{S}_i (i.e. those containing disparity information) were projected in this step. Estimates of multivariate covariance (δ_{ij}) are based on the species scores produced from this projection. To reiterate, these computations are possible only because traits were projected onto a common space, defined by the compromise or consensus matrix resulting from differentially weighting the original matrices. Alternatively, a common basis can also be obtained from a Procrustes rotation of the distance matrices (Krzanowski, 2000; Meyners *et al.*, 2000).