Phylogenetic Analysis of Phenotypic Covariance Structure. I. Contrasting Results from Matrix Correlation and Common Principal Component Analysis

Scott J. Steppan

Evolution, Volume 51, Issue 2 (Apr., 1997), 571-586.

Stable URL:
http://links.jstor.org/sici?sici=0014-3820%28199704%2951%3A2%3C571%3APAOPCS%3E2.0.CO%3B2-Z

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Evolution is published by Society for the Study of Evolution. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/ssevol.html.

Evolution
©1997 Society for the Study of Evolution

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

http://www.jstor.org/
Mon Mar 3 11:41:05 2003

Evolution, 51(2), 1997, pp. 571–586

PHYLOGENETIC ANALYSIS OF PHENOTYPIC COVARIANCE STRUCTURE. I. CONTRASTING RESULTS FROM MATRIX CORRELATION AND COMMON PRINCIPAL COMPONENT ANALYSES

SCOTT J. STEPPAN[1]
Committee on Evolutionary Biology, The University of Chicago, Chicago, Illinois 60637 and Division of Mammals, The Field Museum, Chicago, Illinois 60605
E-mail: steppan@onyx.si.edu

Abstract.—Applications of quantitative techniques to understanding macroevolutionary patterns typically assume that genetic variances and covariances remain constant. That assumption is tested among 28 populations of the *Phyllotis darwini* species group (leaf-eared mice). Phenotypic covariances are used as a surrogate for genetic covariances to allow much greater phylogenetic sampling. Two new approaches are applied that extend the comparative method to multivariate data. The efficacy of these techniques are compared, and their sensitivity to sampling error examined. Pairwise matrix correlations of correlation matrices are consistently very high (> 0.90) and show no significant association between matrix similarity and phylogenetic relatedness. Hierarchical decomposition of common principal component (CPC) analyses applied to each clade in the phylogeny rejects the hypothesis that common principal component structure is shared in clades more inclusive than subspecies. Most subspecies also lack a common covariance structure as described by the CPC model. The hypothesis of constant covariances must be rejected, but the magnitudes of divergence in covariance structure appear to be small. Matrix correlations are very sensitive to sampling error, while CPC is not. CPC is a powerful statistical tool that allows detailed testing of underlying patterns of covariation.

Key words.—Common principal components, comparative method, cranial morphology, macroevolution, phenotypic covariance structure, *Phyllotis*, quantitative genetics.

Received May 20, 1996. Accepted October 17, 1996.

The evolution of patterns of phenotypic covariation is central to two major approaches to understanding macroevolution in terms of microevolutionary process. One is that of quantitative genetics, where evolutionary change can be modeled as natural selection acting on the additive genetic variances and covariances. Solving equations for a large number of generations extends this typically microevolutionary technique to macroevolutionary phenomena. Phenotypic covariation is the expression of these underlying genetic parameters and is the attribute from which genetic covariances are calculated. The other approach is a set of techniques that can be grouped under the field of evolutionary constraints, which views phenotypic covariances as the expressions of underlying constraints, whether genetic, developmental, or other.

The two approaches address the issue from different directions. Quantitative genetics has built a body of explicit theory and analytical models for microevolutionary change, and requires a series of assumptions to extrapolate these models to understand macroevolution. On the other hand, an evolutionary constraints approach infers underlying mechanisms from macroevolutionary patterns.

There is ample information that covariance patterns are fundamental to quantitative models of phenotypic divergence over short time periods (Lande 1979; Lande and Arnold 1983; Arnold and Wade 1984; Maynard Smith et al. 1985). However, if such models are to be successfully applied to macroevolution, patterns of genetic covariation must either remain constant or behave predictably over macroevolutionary time scales (Lande 1979; Price et al. 1984; Lande 1986; Cheetham et al. 1993). Theory has not provided models to predict the evolution of genetic covariances, and such capabilities are not likely to be immediately forthcoming (Turelli 1988). Lande and others (Lande 1979; Turelli 1988; Wagner 1988; Emerson and Arnold 1989; Arnold 1992) have called for more empirical studies of covariance patterns within a phylogenetic context. Within quantitative genetics, the question has been primarily whether genetic covariances remain constant among related taxa. With an evolutionary constraints approach, a principal question is whether the patterns of phenotypic covariances remain constant. An underemphasized issue, once one moves conceptually from simple comparisons of pairs of taxa to phylogenies, is how do the covariance patterns evolve?

The underlying, although often unstated, objective of these approaches is to bridge the fields of microevolution and macroevolution. During speciation, the accumulated effects of microevolutionary processes become recognized as macroevolutionary patterns. The transition from populations to divergent species is fundamental to evolutionary theory. This is also the transition between population genetics and systematics. Additionally, macroevolutionary patterns and even trends can be achieved through common anagenetic changes among a group of lineages, independent of speciation. Applying quantitative genetic approaches to understanding macroevolutionary patterns requires knowledge of the patterns of covariation among traits ($\Delta\bar{z} = \mathbf{G}\beta$, where $\Delta\bar{z}$ is the evolutionary change in a vector of trait means, β is the vector of selection gradients, and $\mathbf{G}$ is the additive genetic variance/covariance matrix [Lande 1979]).

Solving the iterated version of the above equation for β allows for the estimation of the selective forces that account for the historical divergence of mean phenotypes (Lande 1979). Modifications allow testing whether the amount of selection estimated from observed phenotypic evolution is

[1] Present address: Laboratory of Molecular Systematics, MRC 534, Smithsonian Institution, Washington, DC 20560.

© 1997 The Society for the Study of Evolution. All rights reserved.

necessary to explain that change, or if drift is a sufficient null hypothesis (Lande 1986; Lynch 1990; Cheetham et al. 1993). The historical reconstruction has been used in a variety of studies (Price et al. 1984; Schluter 1984; Lofsvold 1988; Cheverud 1996). One can also extrapolate the effects of genetic constraints to predict the dynamics of future evolutionary changes (e.g., Boulding and Hay 1993), although confidence limits can be considerable (McCulloch et al. 1996). These approaches all share the unproven assumption of constant or predictable genetic covariances, in addition to other simplifying assumptions.

Covariance structure as used in this paper refers to the patterns of covariation among traits. Pattern is present in any variance/covariance matrix with nonuniform variances and covariances. These patterns are nonrandom to the extent that they reflect underlying mechanisms including genetic organization, pleiotropy, developmental programs, or environmental effects. One of the first attempts to describe covariance structure was Berg (1960), who called sets of highly correlated traits "correlation pleiades." Covariance structure is a more general concept subsuming correlation pleiades. It does not imply any specific mechanism producing that structure.

Previous investigations into the evolution of covariance structure have been hampered not by insufficiently robust estimation of **G**-matrices, but by their lack of phylogenetic resolution. Previous tests of covariance evolution have typically employed two or three taxa (Arnold 1981; Atchley et al. 1981, 1992; Lofsvold 1986; Kohn and Atchley 1988; Shaw and Billington 1991; Brodie 1993; Paulsen 1996) or with larger numbers of populations, little or no phylogenetic resolution (Riska 1985; James et al. 1990; Venable and Búrquez M. 1990; Voss et al. 1990). Either way, the internal branches of a phylogeny were poorly sampled. In the two extreme sampling designs, single populations from each of two species or multiple populations of a single species, only one clade or internal node is sampled. Even comparisons involving two or three nodes in a phylogeny give very little information regarding the evolution of covariance patterns. For instance, one cannot determine the phylogenetic level at which covariance patterns diverge. Nor can one determine whether existing differences reflect actual phylogenetic divergence or could be accounted for by intraspecific variation. Those questions require more detailed sampling of the phylogenetic hierarchy. Furthermore, prior studies have used a diversity of hypotheses and test methods, with the result that conclusions have been mixed.

An additional impediment to synthesizing the results of previous studies is the different life histories of the study organisms (e.g., determinate vs. indeterminate growth). Some studies involve a single ontogenetic stage (static allometry), whereas others span multiple stages (ontogenetic allometry) (Klingenberg and Zimmermann 1992). This study cannot synthesize these results directly because *Phyllotis* has its own characteristic life history. Instead, I address the principal shortcomings of previous studies by developing techniques to test for the conservation of covariance structure across a diverse clade of taxa. These techniques are applied to six species of South American leaf-eared mice, *Phyllotis*. More broadly applied, these techniques can allow the comparative study of matrices in a variety of other contexts.

This study more thoroughly samples the evolutionary history of a group, but there are trade-offs to this benefit. Because the number of populations is large (28) it is impractical to determine the genetic covariances. Replicate sets of relatives with genealogical information are needed to estimate genetic covariances, and these require a controlled captive breeding program. It would not be practical to maintain 28 separate captive breeding populations, with the populations established from wild-caught individuals and collectively representative of the systematic diversity for a clade. Instead, phenotypic covariances are calculated from population samples already residing in museum collections. Phenotypic covariance matrices may not be equivalent to genetic covariance matrices (Lofsvold 1986; Kohn and Atchley 1988; Willis et al. 1991), but are typically similar (Cheverud 1988; Arnold 1992). If they are similar, variation in phenotypic covariance structure should give insights into the variation in genetic covariance structure. Constant phenotypic covariances would be a strong indication of common genetic covariances. Independent of the relationship between genetic and phenotypic covariances, the evolution of phenotypic covariances has been poorly documented, and is as important a subject of study as univariate aspects of the phenotype that are commonly and routinely studied in comparative biology. The phenotypic patterns are also directly informative about the degree of variation available to natural selection.

Study Organisms

Phyllotis is a member of the taxonomically and morphologically diverse muroid subfamily Sigmodontinae and its members are commonly known as the leaf-eared mice. Its species are frequently the most abundant mammal species at a locality and they have been described as the South American equivalents of *Peromyscus* (Pearson 1958), although they average two to four times the mass. Most species of *Phyllotis* live in rocky or brushy habitats in the Andes and nearby areas (Fig. 1). *Phyllotis* has semideterminate growth as size increases well into adulthood before asymptoting. Other New World muroid rodent genera have figured prominently in studies of the evolution of covariance structure, including *Peromyscus* (Lofsvold 1986, 1988), *Sigmodon* (Zelditch and Carmichael 1989; Zelditch et al. 1990), and *Zygodontomys* (Voss et al. 1990; Voss and Marcus 1992).

The *P. darwini* species group was chosen as the subject of this research because most species are locally abundant and several thousand individual specimens reside in museum collections. Species-level systematics (i.e., the alpha level taxonomy of the group), has been sufficiently well documented (Pearson 1958; Hershkovitz 1962; Pearson and Patton 1976; Walker et al. 1984), so that individuals could be assigned to species with confidence. Subsequent morphology-based phylogenetic analyses were able only to provide minimal resolution of relationships within *Phyllotis* (Steppan 1993; Steppan 1995b). DNA sequence data from a 973-bp region of the mitochondrial cytochrome *b* gene increased resolution within the genus (Steppan 1995a, in press) and provides the phylogenetic hypothesis for this study.

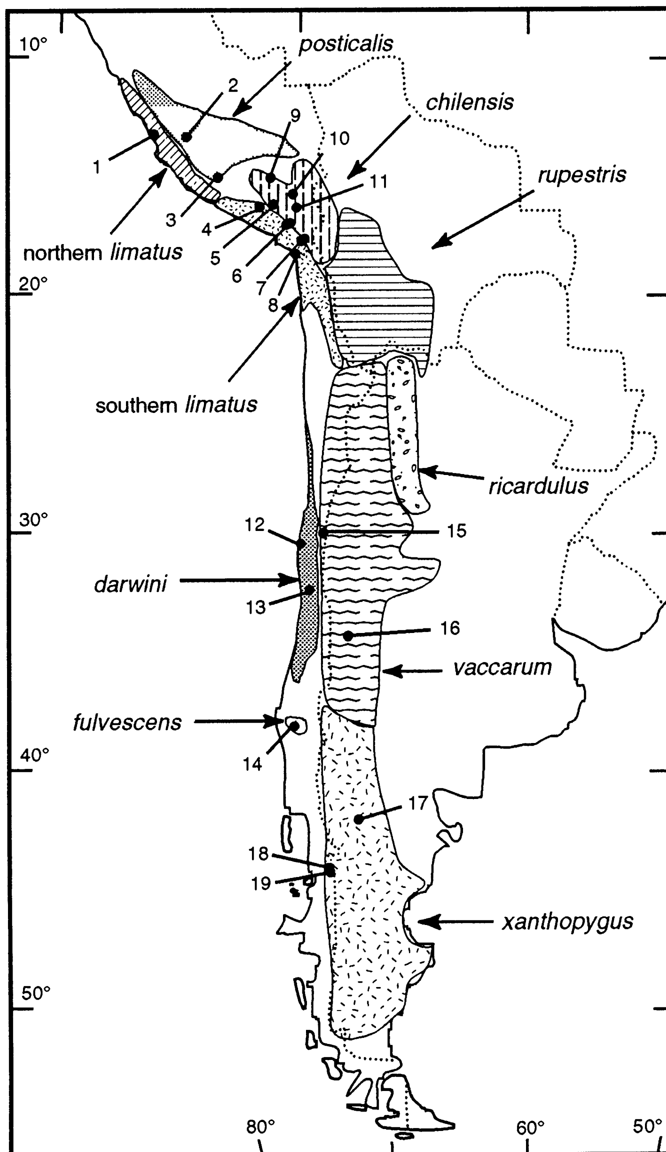

FIG. 1. Distribution of subspecies belonging to *Phyllotis darwini* (*darwini, fulvescens*), *P. limatus* (northern, southern), and *P. xanthopygus* (*posticalis, chilensis, rupestris, ricardulus, vaccarum, xanthopygus*) and locations of sampled populations. (1) Rimac Valley (Lima, Peru); (2) Huancavelica (Huancavelica, Peru); (3) Puquio (Ayacucho, Peru); (4, 5) Arequipa (region, Peru); (6) Torata (Moquegua, Peru); (7) Tarata (Tacna, Peru); (8) Moro Sama (Tacna, Peru); (9) Cailloma (Puno, Peru); (10) Puno (Puno, Peru); (11) Santa Rosa (Puno, Peru); (12) Fray Jorge (Coquimbo, Chile); (13) Santiago (region, Chile); (14) Angol (Malleco, Chile); (15) Paiguano (Coquimbo, Chile); (16) Rt. 40, Mendoza (Argentina); (17) Huanuluan (Río Negro, Argentina); (18) Puerto Ibañez (Aisén, Chile); (19) Chile Chico (Aisén, Chile).

Review of Previous Studies

The conservation of genetic correlations or covariances has been tested in several studies, most of which employed only two operational taxonomic units. The first explicit tests were by Arnold (1981) and Atchley et al. (1981). Visual examination of factor loadings led Arnold (1981) to conclude that genetic correlations for response to prey were the same in two geographic populations of the garter snake, *Thamnophis elegans*. Using similar techniques, Atchley et al. (1981) concluded that genetic correlations for skull dimensions were the same between an inbred laboratory strain of rats (*Rattus norvegicus*) and an inbred strain of mice (*Mus musculus*). While Atchley et al. (1981) found an overall similarity in the genetic and phenotypic correlation patterns, they also noted several differences as well. For example, the first pooled principal components from the genetic correlations in each of the two species appeared to represent different biological properties: a general size factor in rats, a shape vector in mice. This result indicates divergence in genetic covariances, but to an unknown degree and at an undetermined phylogenetic level.

Statistical tests were first employed by Lofsvold (1986), and later by Kohn and Atchley (1988), Atchley et al. (1992), and Paulsen (1996). Nonparametric matrix permutation tests (Dietz 1983) were used by both Kohn and Atchley (1988) and Atchley et al. (1992) for the same samples of rats and mice. The two studies examined different morphological structures, and arrived at different conclusions about the conservation of genetic correlations. Kohn and Atchley (1988) failed to find significant divergence between the two species for genetic correlations for the pelvis. In contrast, Atchley et al. (1992) could not reject the null hypothesis of independent matrices for mandible traits. This difference may indicate that the pelvis is more highly constrained than the jaw. Certainly, the pelvis is much less variable among muroid rodents than are the jaw or skull. However, the general significance of these results should be interpreted cautiously, because they do not address the hypothesis of primary evolutionary interest, that the covariance matrices have remained constant (Turelli 1988). Instead, the matrix permutation tests determine whether correlation matrices are significantly similar. Nonetheless, the similarity of correlation matrices is an empirically valuable insight.

A more significant criticism of comparative studies with only two operational taxonomic units (OTUs) is the inability to generalize the conclusions, and this is exacerbated by the use of laboratory and inbred strains. Divergence (or convergence) could occur at any point in the phylogeny since the two lineages split, from approximately 10 M.Y.B.P. for rats and mice (Catzeflis et al. 1992), to the populations from which the strains were derived, or even since inbreeding. In effect, there are only two data points (or just one independent contrast, sensu Felsenstein 1985), and no statistically significant pattern can be determined from just two data points.

Lofsvold (1986) conducted the only test of genetic covariances in more than two taxa. Correlation and covariance matrices of external and cranial measurements were compared in two subspecies of *Peromyscus maniculatus* and in *P. leucopus*. Matrix permutation tests indicated that the null hypothesis of independent covariance matrices could not be rejected for the two species, and that covariance matrices were significantly similar in the two *P. maniculatus* subspecies. All correlation matrices were significantly similar. Kohn and Atchley (1988) later reanalyzed Lofsvold's data using a different algorithm and more permutations. They found significant similarity among all comparisons of covariance matrices. Since matrix similarity is a less restrictive test than

matrix equality (Turelli 1988; Shaw 1991), it is unclear whether there has been significant divergence in covariance structure among these species. Based on Lofsvold's (1986) results, covariance matrices of subspecies do appear to be more similar than species matrices.

More recently, Paulsen (1996) found no significant difference in **G**-matrices between two sibling butterfly species in *Precis*. She did find significant differences between heritabilities for some traits in pairwise element comparisons. Paulsen interpreted this as a warning against assuming the constancy of **G**.

These genetic studies suggest that genetic covariance structure may diverge over an indefinite time period, perhaps at phylogenetic levels above subspecies. Covariance structure for some features may be under tighter constraints than others. Divergence in genetic covariance structure has been demonstrated in as few as 20 generations for fruitflies under strong directional selection (Wilkinson et al. 1990).

All studies employing larger numbers of OTUs have been limited to phenotypic correlations or covariances. Three of the five largest were entirely intraspecific and lacked an evolutionary hierarchy. Riska (1985) examined 34 local populations of the aphid *Pemphigus populicaulis*. There was no explicit phylogenetic structure to the data, although geographic proximity may be associated with genetic relatedness. Significant heterogeneity of correlation elements among populations was observed using a jackknifing procedure. Despite this heterogeneity, and in contrast to observations of population means, no geographic pattern was detected. Local samples appeared to diverge significantly, but randomly from each other. Riska interpreted this result as possibly indicating that geographic patterns in the means was the result of regional similarity in natural selection, but that the genetic changes responsible for this convergence on phenotype differed among the populations. Each population presented a different set of genetic variants for selection to act upon, and thus although the phenotypic response may be similar, the underlying genetic patterns may differ. The random variations in phenotypic correlations were interpreted as evidence of that genetic organization.

Voss et al. (1990) examined 15 populations of the Neotropical cane mouse, *Zygodontomys brevicauda*. No estimate of relationship among populations by geography or other biological measures was included. Analysis focused on normalized coefficients of the first principal component because characteristic roots after the first were seldom distinct. The coefficients were very similar among all samples. Voss et al. (1990) hypothesized that the populations shared a latent growth factor producing parallel allometric trajectories. Differences in adult morphology among populations would be due to developmental changes before weaning that result in an offset of allometric trajectories.

Voss and Marcus (1992) expanded the taxonomic scope of their earlier study to include 14 species; two from each of seven Neotropical muroid (sigmodontine) genera. The expanded study confirmed their earlier results. Phenotypic divergence between congeneric pairs was produced primarily by morphological shifts early in development, followed by nearly parallel postweaning allometric trajectories. The degree of conservation or divergence among genera was not discussed. A similar observation was made by Björklund (1993), who found that three species of finch shared tightly constrained growth trajectories.

James et al. (1990) measured bill shape, and wing and tarsus lengths in the red-winged blackbird. They divided samples across the geographic range into 17 blocks, sized 2° in latitude and longitude on a side. They found significant heterogeneity among the blocks for variances, resulting in significant differences among covariances, but not correlations. However, the significance of this finding is difficult to interpret because the correlations/covariances were calculated only between a single estimate of shape (the first principal component derived from three "shape" ratios) and a single estimate of size (wing length).

The only study with significant phylogenetic structure is Goodin and Johnson (1992). They studied 32 populations of land snails in a dataset that was both geographically and phylogenetically structured, spanning four hierarchical levels. They calculated an index of similarity in covariance patterns based on the number of traits having their maximum factor loadings on comparable factors from independent factor analyses. That index is a modification of a technique used by Gould et al. (1974) to recognize sets of highly correlated characters. Based on this index from pairwise comparisons, covariance similarity was found to be unaffected by the hierarchical level of the comparison. There was a suggestion that covariance patterns were most similar between populations of the same subspecies, and lowest among subspecies of the same biological species, but these differences are probably not statistically significant. Multidimensional scaling of the matrix of similarity indices segregated the subspecies from each other (i.e., distinct patterns of covariation), but there was nearly total overlap among species. Although there was little evidence for phylogenetic structure to the data, indices of covariance similarity among species were strongly correlated with genetic similarity based on allozymes.

All previous studies have lacked at least one of two critical elements. They either lack sufficient phylogenetic information and/or a statistical test of the null hypothesis of equivalent matrices. In this paper, I introduce two new phylogeny-based comparative techniques for the study of the evolution of covariance structure: one heuristic, the other statistical. Each one addresses different difficulties with the comparative analysis of multivariate data and thereby give complementary insights.

Objectives

In this paper, I test whether phenotypic covariance structure is conserved throughout a species group phylogeny. In conjunction with this test, I examine the efficacy of two new approaches for quantifying and testing the evolution of covariance structure. The effect of sample size on estimates of matrix similarity/identity is also tested.

Methods

Materials and Measurements

Systematics.—Twenty-eight populations representing 13 diagnosable taxa (subspecies) belonging to six biological spe-

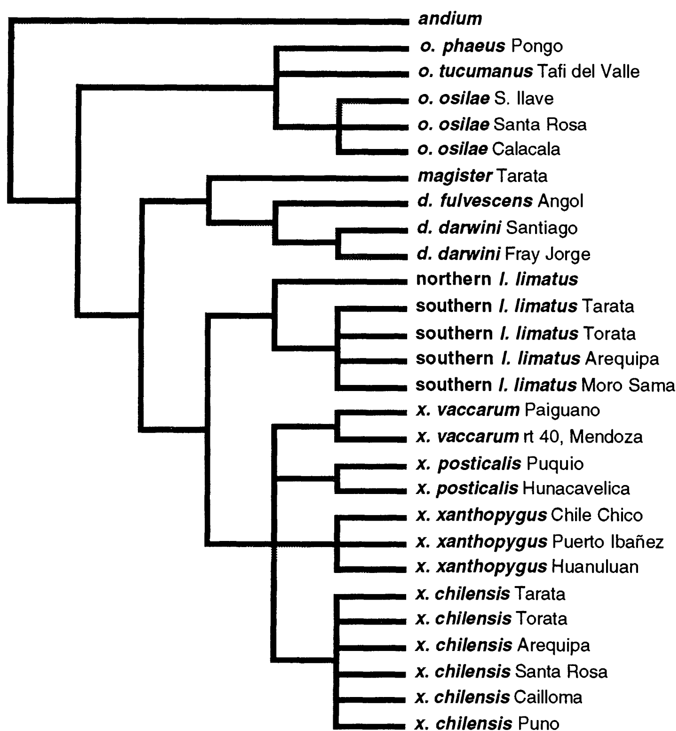

FIG. 2. Phylogenetic hypothesis used for the analyses of covariance structure within *Phyllotis*.

cies were included in the population analyses. The species (with total number of individuals analyzed) are *P. andium* ($n = 71$), *osilae* ($n = 405$), *darwini* ($n = 246$), *magister* ($n = 159$), *limatus* ($n = 414$), and *xanthopygus* ($n = 1250$). The taxonomy for the assignment of individuals, taxon boundaries, and phylogenetic relationships follows that presented in Steppan (1995a, in press) including the separation of *P. limatus* as a species distinct from *P. xanthopygus*. The revised geographic distributions for *P. darwini, P. limatus*, and *P. xanthopygus* are presented in Figure 1 and localities for all populations are given in the Appendix. The preferred phylogeny from an analysis of a 973-bp region of the mitochondrial cytochrome *b* gene (Steppan, in press) was the standard model for the hierarchical analyses of covariance structure. The working phylogenetic hypothesis for the OTUs is presented in Figure 2. The tree in Figure 2 groups all populations of a single subspecies into an unresolved polytomy, reflecting the best cladistic representation of relationships among interbreeding populations with undetermined amounts of gene flow. Subspecies are retained as separate lineages within a species hierarchy because they are roughly equivalent to phylogenetic species, and are diagnosably distinct, at least morphometrically. They are grouped into polytomies at the species level because there is circumstantial evidence for at least some gene flow among putative subspecies, and significant gene flow would render a strictly dichotomous branching hierarchy inappropriate. These nested pairs of polytomies reflect the hypothesized discontinuities in interbreeding and evolutionary history, with relatively high levels of gene flow among populations of the same subspecies, limited gene flow among subspecies of putatively biological species, and effectively no gene flow among biological species. Higher-level branching is strictly dichotomous.

Age criteria.—Individuals were assigned to age classes based on molar eruption and wear. Molar eruption criteria were given priority because tooth wear can be affected by differences in diet. The age classes and definitions are modified from common criteria (e.g., Voss et al. 1990) to better reflect *Phyllotis* dental morphology and ontogeny. The definitions for the age classes are as follows.

Age class 0: M3 unerupted.

Age class 1: M3 erupted, but not reaching the occlusal plane of M1 and M2. Associated with juvenile body size.

Age class 2: Mesoflexus of M2 (labial fold) partially obscured by the maxilla. M3 slightly worn. Associated characters include small testes, juvenile pelage, body size approximately that of adult.

Age class 3: All labial flexi of upper molars fully exposed, M3 worn to a "J" or "S" shape. Roots unexposed. Associated characters include sexual maturity, adult pelage and body size.

Age class 4: Individual roots become distinguishable above alveolus, mesoflexus of M3 cut off by wear to form lake.

Age class 5: All molars heavily worn, folds and flexi reduced to shallow indentations.

Morphometric analyses included individuals in age classes 2–5. In *Phyllotis*, growth continues through to age class 4, but age class 5 individuals are no larger on average than those in age class 4, even for features like the rostrum. Trait mean size changes from age class 2 to 5 ranged from +29% (incisor depth) to −1% (interorbital breadth). Age classes 0 and 1 were excluded because many populations had little or no representation from these classes. The rapid growth at these ages can significantly increase the correlation and covariances among traits, such that populations including individuals in these age classes will appear to exhibit significantly higher covariances and overall morphological integration (sensu Cheverud 1982) than populations lacking such representation.

Measurements.—Where possible, measurements were chosen to address the criticisms of Bookstein et al. (1985) regarding traditional distance-based schemes. Landmarks were chosen to represent homologous points, such as suture junctions, rather than extremal points, such as maximum length or maximum width. Long measurements that traverse multiple structures were avoided, and most measurements were of individual bones. Long measurements will obscure true patterns of variation if multiple elements subsumed by the measurement do not covary isometrically. Selected measurements did not repeatedly span the same element, which would produce redundancy and obscure subtle internal variations. Element redundancy is a greater problem with longer measurements. For example, total skull length subsumes all longitudinal measures along the midline, and thus would be partially redundant with all other longitudinal measures. Furthermore, highly variable structures may have undue influence over long or overlapping measurements. A common example from mammalian morphometric studies involves the rostrum, which exhibits much greater postnatal growth and variation than the cranium. Three traditional measurements

used in mammalian taxonomic studies are nasal length, condylobasal length, and total length. As these all subsume the rostrum, they are effectively redundant measures of a single element, and even in combination, provide little information about the less variable braincase. Using the same landmark for multiple measurements was avoided except for sequential structures along the midline, because any error in locating the landmark, such as due to damage, would be compounded in the several measurements. Including longitudinal measurements on both dorsal and ventral sides of the skull (e.g., diastema and nasal length) better describes shape changes like arching of the skull without confounding multiple elements. While not uniformly distributed, measurements were chosen to be relatively evenly distributed around the skull. Some measurements were chosen that were not necessarily optimal from an analytical standpoint, but that are commonly used in systematic studies, allowing some comparison across studies. Two measurements, jaw length and moment arm of the masseter, were used because of their biomechanical importance in chewing and thus their adaptive significance (Radinsky 1985).

Twenty-four measurements were taken from skulls and jaws using digital calipers precise to 0.01 mm (Fig. 3). A total of 2861 *Phyllotis* skulls were measured, although not all were included in the among-population analyses of covariance evolution. Because the OTUs in the analyses were meant to represent local populations as much as possible, individual specimens were included if they could be aggregated into population samples collected within a radius of 25 km and temporal range of 30 years. The minimum sample size was 24. These temporal or spatial criteria were relaxed in some instances to allow the inclusion of taxa that otherwise lacked sufficiently aggregated local sampling. For example, the OTU *P. andium* represents a composite of several local populations from adjacent provinces because no single population was large enough for robust estimation of correlation or variance/covariance matrices. Similarly, *P. osilae phaeus* included individuals from outside the locality Pongo, Bolivia, because despite a sample size of 40, the Pongo population showed an unusually limited age distribution. As a result of these minimum requirements for population samples, some taxa belonging to the *darwini* species group were not represented in the covariance analyses, including *P. osgoodi, P. x. rupestris,* and *P. x. ricardulus*. A list of the specimen numbers and locality information for all populations is included in the Appendix.

For the comparisons among taxa, data were pooled among all representatives of a taxon, not just among the populations from the population level analyses. By pooling across geographic samples, the subspecies *P. x. rupestris* could be included in some analyses. The species *P. caprinus*, which appears to be a member of the *darwini* species group (Steppan 1995b), was not analyzed because insufficient DNA sequence data was acquired to incorporate it into the species group phylogeny (Steppan 1995a).

Analysis: Correlation Matrices

Correlation matrices for 24 variables were calculated for each OTU from raw, untransformed data. Logarithmic transformations were not made because the raw data were already normally distributed and an allometric model was not being applied to the interpretation of the results. Matrix correlations were then calculated for all possible pairwise comparisons of OTUs. The degree of matrix correlation was found to be correlated with the sample sizes of the populations being compared (Fig. 4). The harmonic mean of the sample sizes was the best predictor of the matrix correlations, as compared to arithmetic means, geometric means, or the smaller of the two sample sizes. If the reduction in matrix correlations is due to sampling error in estimating the correlation matrix, the smaller sample will account for a greater proportion of the error. Harmonic means give greater weight to the smaller values.

A rarefaction analysis was performed on selected matrix correlations to estimate the impact of sampling error. Rarefaction analysis involves randomly subsampling an original dataset to successively smaller sample sizes to estimate an expected parameter value for a given sample size. The largest geographic population sample, *P. x. chilensis* from Tarata ($n = 135$; Fig. 5), and the two subspecies with the largest samples, *P. x. chilensis* ($n = 501$) and southern populations of *P. limatus* ($n = 315$), were subjected to rarefaction analyses. Matrix correlations were calculated among all subsamples for each subspecies, and between subsamples of the two taxa. The matrix correlations were regressed against the inverse of the harmonic mean of the sample sizes, which resulted in a linear relationship, and slopes of the least-squares regression lines determined (Fig. 6). The mean of the three slopes was used in a correction factor. The population sample was analyzed separately.

A partial correction for sampling error was achieved by incorporating the information from the rarefaction analysis. The reduction in the mean observed matrix correlation from the maximum value was a linear function of the inverse of the harmonic mean sample size. Thus, given the sample sizes, one can estimate the expected reduction in matrix correlation below the true matrix correlation. An "adjusted correlation" was calculated as the sum of the observed correlation and the correction factor, by the formula

$$r_c = r_m - b/\mu_h, \tag{1}$$

where r_c is the "adjusted correlation," r_m is the observed matrix correlation, b is the slope of the regression line of r_m on $1/\mu_h$(the slope is negative), and $\mu_h$is the harmonic mean of the sample sizes. However, this correction cannot also correct for the heteroscedasticity of the sampling error, in which the variance in matrix correlations increases with decreasing sample sizes. Use of residuals from a regression of the population by population comparisons for the complete dataset produces statistically indistinguishable results. That fact suggests that rarefaction analysis is an appropriate method to estimate the average effect of sampling error. Rarefaction was preferred over "matrix repeatability" (Cheverud 1996) because that technique compares observed correlations to the theoretical maximum, not the expected mean, and may be biased to underestimate matrix correlations at small sample sizes (< 40; Cheverud 1996).

Adjusted matrix correlations were assigned to categories reflecting the taxonomic level of the pairwise comparison.

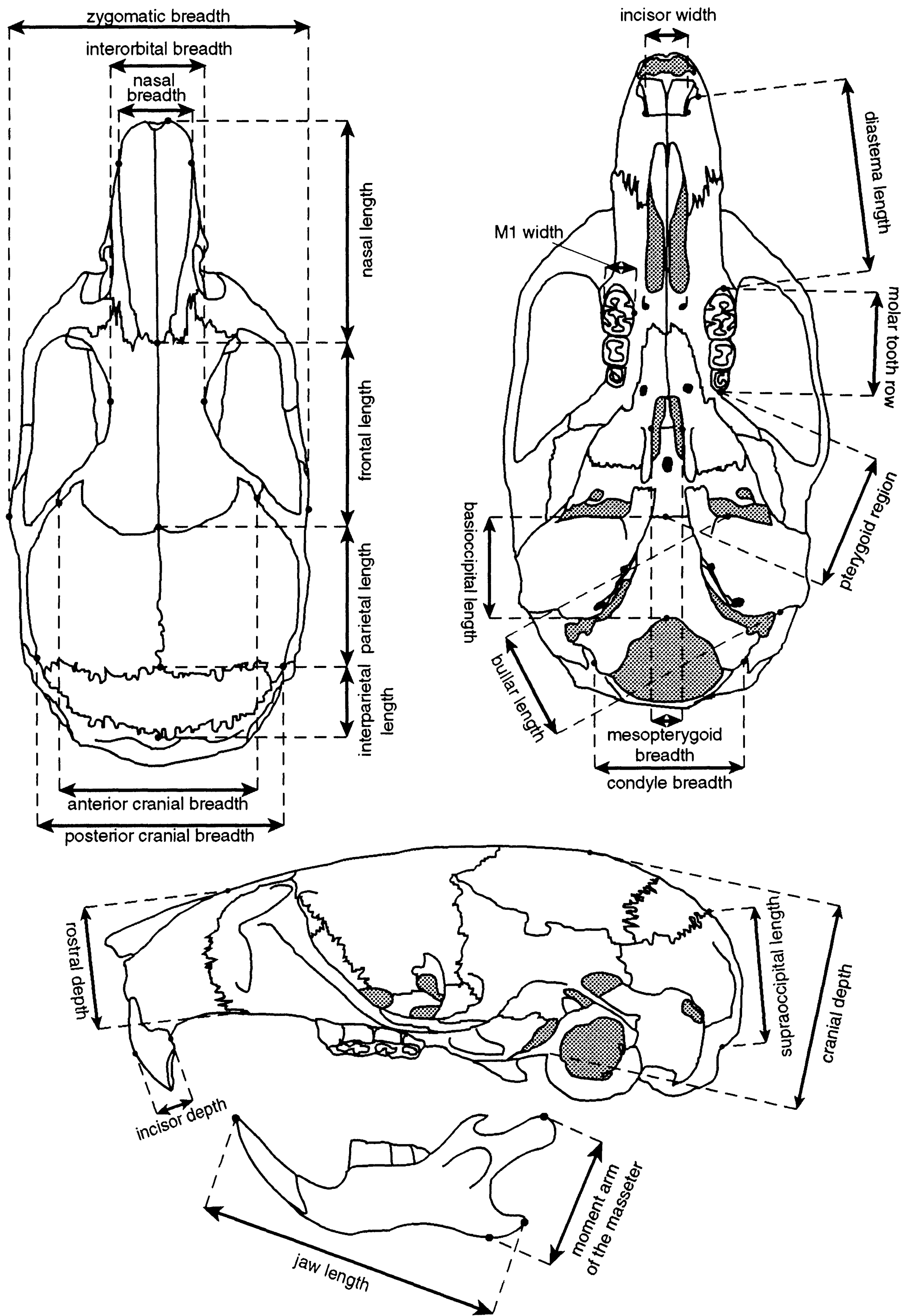

FIG. 3. Twenty-four skull and jaw measurements.

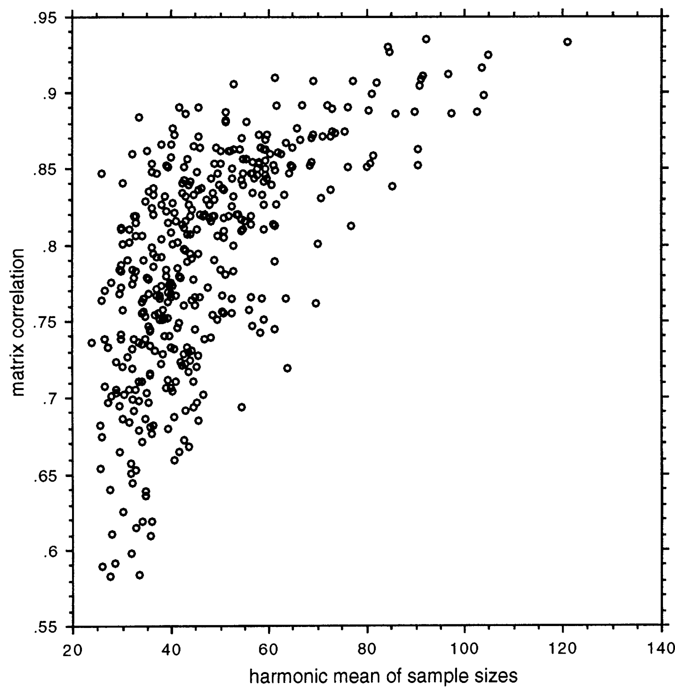

FIG. 4. Plot of all possible pairwise matrix correlations among population correlation matrices for 24 variables.

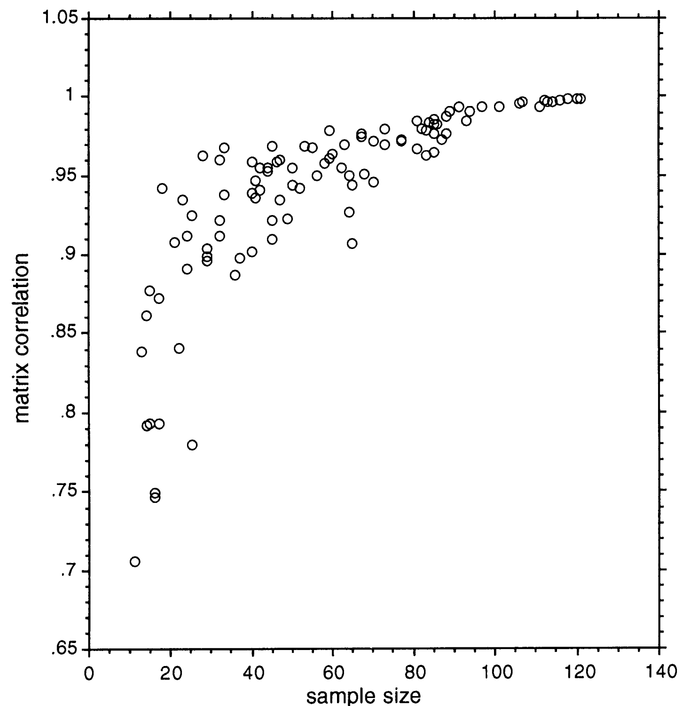

FIG. 5. Matrix correlations of random subsamples of the population *chilensis* Tarata with the original dataset. Plotted sample size is that of the subset.

For example, the adjusted matrix correlation between *P. d. darwini* from Fray Jorge and *P. d. darwini* from Santiago was classified as between populations of the same subspecies. Higher level categories were comparisons between populations belonging to different subspecies of the same biological species, between populations belonging to sister-species, between populations belonging to sister-clades (one node higher than sister-species), and so on. All possible pairwise matrix correlations were calculated and each matrix correlation was assigned to an hierarchical category. An analysis of variance was conducted on the adjusted matrix correlations, grouped by the hierarchical level of the comparison. The redundancy of populations among all possible pairwise comparisons means that the matrix correlations are not all independent of each other, and that the degrees of freedom will be significantly inflated. However, as will be explained below, a correction for this inflation was found to be unnecessary.

A second set of correlation analyses was conducted to reduce the effect of sampling error. Correlation matrices were calculated for each subspecies, pooling all specimens that matched the age criteria. This pooled multiple populations, and also allowed the inclusion of specimens belonging to populations too small to be treated individually.

Analysis: Common Principal Components Model

Statistical tests of variance/covariance matrices typically test a single hypothesis: are two matrices equal? However, Flury (1987) and Airoldi and Flury (1988) pointed out that two matrices can be associated by a hierarchical series of relationships. For example, one matrix may simply be a scalar multiple of another. In that case, they would be proportional, with identical patterns of covariation but differing only in magnitude by some scalar. Similarly, two matrices may share a common principal component structure (same eigenvectors), but differ in the eigenvalues for those components. In other words, the orientation of axes (individual principal

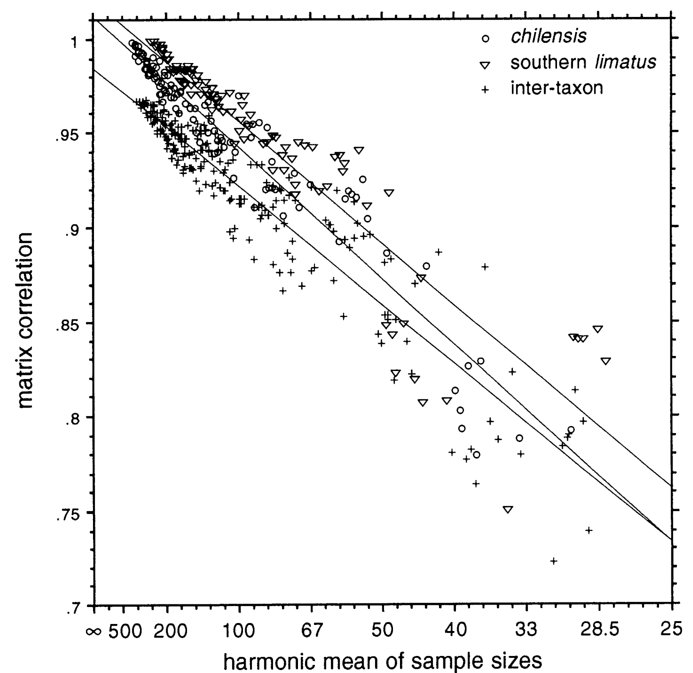

FIG. 6. Rarefaction plots for matrix correlations among random subsamples of the two largest taxon datasets and between subsamples of each taxon.

components) may be the same but the variance associated with each may vary. Furthermore, two matrices may have eight of 10 components in common, but two of those components may differ in orientation. Here, a partial common principal component model would apply. A nested series of models thus exists. The least inclusive model is that matrices share a single principal component. This nested hierarchy of matrix associations is amenable to the analysis of phylogenetically structured data, which is also a nested hierarchy of associations. Flury's approach provides finer resolution to the evolution of covariance matrices than single tests of matrix equality or similarity.

The common principal component approach, hereafter referred to as CPC, proceeds by building up each level in the hierarchy, from arbitrary to equality, and testing the significance of each level against the next lower level. The likelihood of a given model (level) is calculated assuming that all deviations in individual matrices from this model are due to sampling error. Differences in likelihoods between levels in the hierarchy are distributed as a χ^2, assuming multivariate normality (Flury 1988). This tests the likelihood that each level is true given that the lower level is true, resulting in the decomposition of the log-likelihood ratio statistic.

CPC also has the significant advantage in being applicable to sets of matrices, rather than being limited to pairwise comparisons. Shaw (1991) proposed a technique for testing equality of genetic covariance matrices, which like CPC is a maximum-likelihood method, but which is only applicable to pairs of matrices. CPC is a generalization of the principal component model from a single group to multiple groups. This generalization eliminates the problem of nonindependence of data points and the inflation of degrees of freedom in multiple pairwise comparisons. All pertinent populations or taxa can be examined simultaneously for shared covariance structure. CPC does not directly account for nonindependence due to phylogenetic history. For a different application of CPC to evolutionary constraints and allometry, see Klingenberg and Zimmermann (1992) and Klingenberg and Spence (1993).

Variance/covariance matrices were calculated for each population from data transformed into natural logarithms. The logarithmic transformation was used in this analysis because it improved the stability in jackknifed analyses. Thirteen variables were analyzed, rather than the 24 in the correlation analyses, because matrix equality could not be concluded when more than 14 variables were included in nonoverlapping random subsets of a single population (*chilensis* from Tarata, $n = 70, 65$). Thirteen variables also reduces the number of parameters that need to be estimated by the model, which is of greater concern in a hypothesis testing approach like CPC than in the correlation analyses. If there are p variables in an analysis, then there are $p(p + 1)/2$ parameters for variance/covariance matrices (also the minimum number of parameters in CPC analyses under matrix identity). Additionally, the computation time for the CPC analysis increases approximately as 2^p. Thus, each additional variable doubles the computation time. The thirteen variables, chosen to represent a range of correlation and covariation values as well as major cranial features are: diastema length, molar toothrow length, pterygoid region, basioccipital length, frontal length, nasal length, nasal breadth, interorbital breadth, anterior cranial breadth, posterior cranial breadth, incisor depth, bullar length, and cranial depth.

Common principal component analyses were conducted using the program CPC, made available by the author Patrick Phillips (1994), which performs the analyses outlined by Flury (1988). The OTU sets included in each analysis constituted the members of each clade in the phylogeny. The analyses were repeated for each more inclusive clade, until in the largest analysis, all 28 OTUs were analyzed simultaneously (i.e., 28 separate matrices analyzed as a set, rather than pooling the data to form a single matrix). The more inclusive analyses contained more OTUs, but did not group by subclades. The hierarchical level of matrix association (equality, proportionality, common principal component, partial common principal component, etc.) was then mapped onto the phylogeny. The level of matrix association was defined as the highest level in the hierarchy for which the null hypothesis of common structure among matrices could be accepted at the $P = 0.05$ significance level.

A supplemental criterion for choosing the optimal hierarchical level is the Akaike Information Criterion (AIC) (Akaike 1973). The AIC balances the fit of the model against the number of parameters required. Thus, if successive levels provide an equal fit to the data (as determined by the log-likelihood decomposition of χ^2s), then the lower level with fewer parameters has a smaller AIC value and is preferred (Flury 1988). In some cases, a high level may have the lowest AIC value even if lower levels of association are rejected by the log-likelihood ratio test. Because most of the variation is contained in the first few components, nonequality of later components may not be biologically significant. If there were only one or two rejected components and they were beyond the fourth partial common principal component (PCPC 4), then the best fitting model by the AIC criterion was still accepted as the best description of common structure. The program CPC ranks PCPCs by the total variance associated with each (Phillips 1994).

Phenotypic covariance matrices from Lofsvold (1986) were reanalyzed using the CPC analysis. This allows a direct comparison of covariance evolution in three taxa of *Peromyscus* and the 28 OTUs in *Phyllotis*, as well as simplifying interpretation of those published results in light of the different methods used in this study.

Results

Matrix Correlation Analyses

Sample sizes and rarefaction analysis.—When matrix correlations are plotted against sample size, a strong, nonlinear relationship is apparent (Fig. 4). The highest matrix correlation was between southern *limatus* from Tarata and *chilensis* from Tarata ($n = 74, 139$) while the lowest matrix correlation was between *osilae* from Calacala and *posticalis* from Huancavelica ($n = 27, 28$). The populations with the largest sample sizes all had high correlations, and the smallest correlations all belonged to populations with the smallest sample sizes. The variance in matrix correlation was also much greater for small sample sizes. This pattern is very similar to that found by Cheverud (1988) in comparing genetic and phenotypic

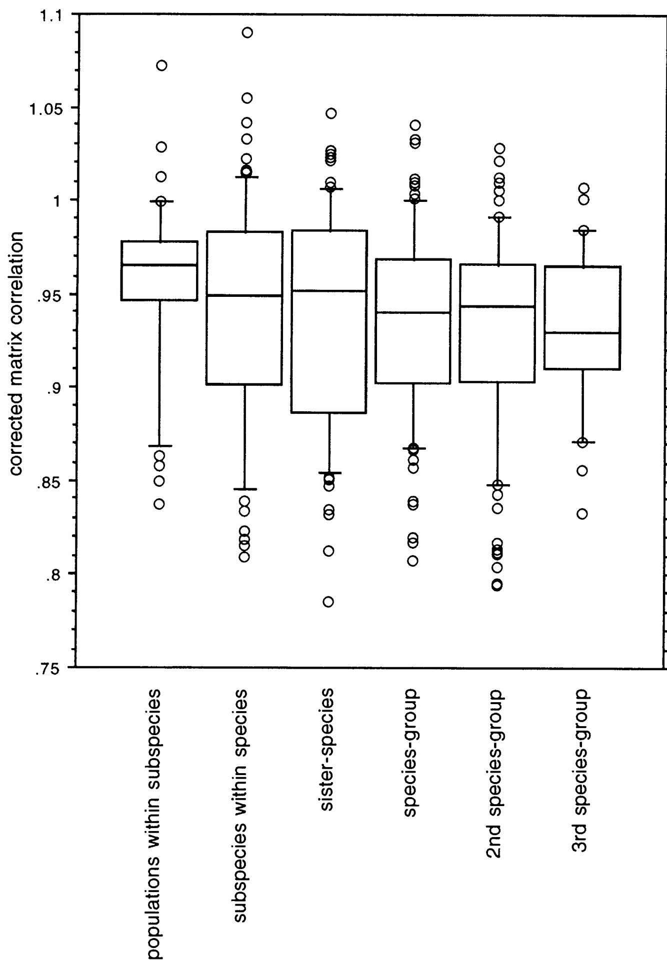

FIG. 7. Matrix correlations, partially corrected for sample size, of all possible pairwise comparisons, categorized by the phylogenetic relationship of the population pairs. Bars indicate the 10, 25, 50, 75, and 90 percentiles. Circles represent individual outliers.

correlation matrices. It is clear that sampling error in estimating correlation matrices could overwhelm any pattern produced by phylogeny. Mean sample sizes did not vary significantly among the phylogenetic levels of the comparison, with average μ_h ranging from 38 (between third level species-group) to 49 (between subspecies).

To test the effect of sampling error, a rarefaction analysis was conducted. The largest population sample, *P. xanthopygus chilensis* from Tarata, showed the same drop in matrix correlations with small sample sizes, although with higher overall values (Fig. 5). The distribution of points in Figure 5 is higher than that in Figure 4 because a single population was sampled. Also, each subsample was correlated with the original dataset rather than with other subsamples. The sample sizes indicate those of the subsamples, not the harmonic means, which would be higher. The asymptotic relationship evident in Figure 5 indicated that matrix correlations may be appropriately described by the inverse of the sample size. Subsampling a single small population (n = 120) results in repeated representation of individuals in many subsamples, thus inflating the matrix correlations in the rarefaction analysis, while different field samples do not result in repeated

TABLE 1. Mean matrix correlations among all pairwise comparisons within specified phylogenetic levels. Student's t is calculated for the difference in means between the corresponding phylogenetic level and the next higher level, and is not corrected for inflation in degrees of freedom due to multiple comparisons.

Phylogenetic level	n	Matrix correlation, mean	Student's t
Population	38	0.957	1.50
Subspecies	75	0.939	0.29
Sister-species	74	0.936	0.11
Species-group	95	0.935	0.56
2nd species-group	96	0.931	0.11
3rd species-group	30	0.932	—

representation of individuals. Field samples effectively represent subsamples from an infinite sized population. In Figure 5, subsamples of size n = 60 share on average half of their individuals in common, while different populations of n = 60 share no individuals in common.

Rarefaction analysis was conducted on pooled subspecies samples to more closely approximate the natural situation. Matrix correlations within and between each of the two largest subspecies samples are plotted in Figure 6. At least down to n = 30, the relationship between matrix correlation and the inverse of μ_h appears linear. The slopes for each group are not statistically different from each other, and the dispersions differ only in their intercepts. Again, there is significant heteroscedasticity, with the variance increasing with smaller samples. The mean slope of the least-squares regression lines from the rarefaction analysis (intertaxon: b = −6.27; *chilensis*: b = −6.95; southern *limatus*: b = −6.45) were used to partially correct for sampling error, as described in the methods section.

Hierarchical matrix correlations.—Adjusted matrix correlations were categorized by the phylogenetic level of the comparisons. Analysis of variance indicated that phylogenetic relatedness had a significant impact on matrix correlations (P = 0.0006). However, this significance value can not be accepted because it is based on inflated degrees of freedom owing to repeated comparisons by each population. Additionally, the low mean matrix correlations of second species-group comparisons (*P. osilae* populations vs. members of the *darwini* species group) is largely due to the effect of a single population, *osilae phaeus* from Pongo. Ninety-six percent of the individuals in the Pongo population are in age class 3 or young age class 4, with the smallest overall size variation of any population. As a result, matrix correlations are unusually low. Because *P. osilae* includes only five populations, comparisons involving *phaeus* from Pongo constitute 20% of all second species-group comparisons.

The mean adjusted correlations for each category are all very high, greater than 0.93 (Fig. 7, Table 1). However, excluding the Pongo population results in a nonsignificant effect due to phylogeny (P = 0.27), even without correcting for inflated degrees of freedom. Corrections for inflated degrees of freedom were unnecessary because they would only further weaken the statistical significance. This single population affects the results so significantly because of the limited number of data points in phylogenetically structured data, and

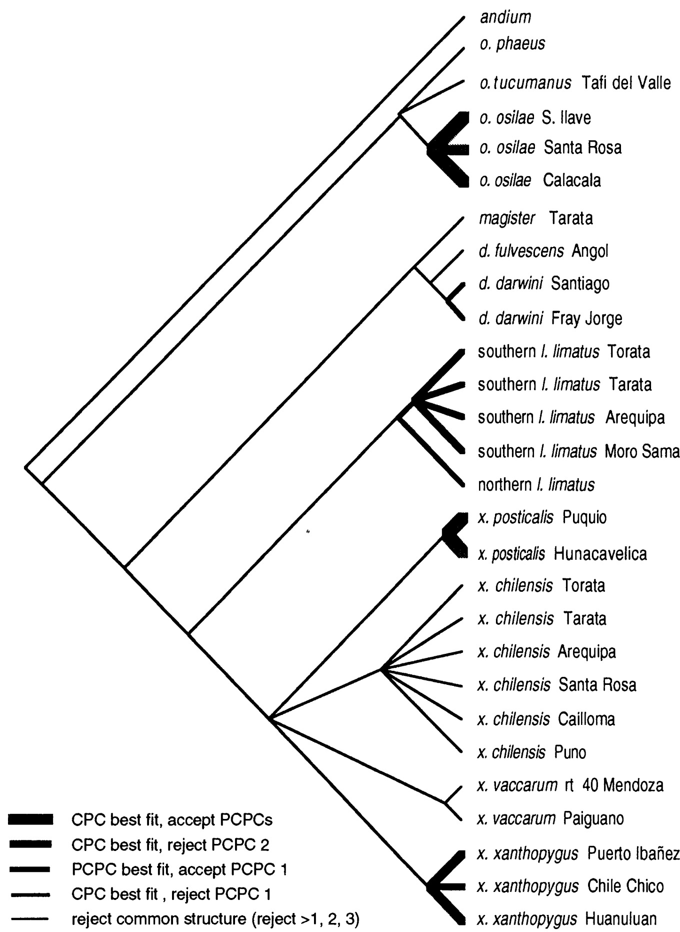

FIG. 8. Common principal components analysis of shared hierarchical structure in covariance matrices. Thicker lines indicate clades sharing common structure higher in the hierarchy of matrix associations. For most clades, one must reject at least the first three partial common principal components.

the lack of independence in the data when one groups by, but does not correct for, the shared phylogenetic history of populations. Correlation matrices of populations belonging to the same subspecies are on average no more similar to each other than are populations separated by several speciation events. However, there is some indication that populations of the same species are more similar to each other than are populations compared at the higher taxonomic levels. The difficulty lies in estimating the effective sample sizes (i.e., number of comparisons) in calculating the test statistic and the appropriate degrees of freedom. Population level comparisons would be significantly different from species group and higher levels of comparison (Student's t = 2.15–2.53), provided the degrees of freedom exceeded 13 and five, respectively (Table 1). The largest difference in mean correlations between adjacent phylogenetic levels is between population level comparisons and subspecies comparisons (r = 0.957 vs. 0.939, a difference of 0.018). No other adjacent pairs of levels differ by more than 0.004. This value is not significant, even with uncorrected degrees of freedom, but it is unclear to what degree that nonsignificance might be due to sampling error in the estimation of correlation matrices. This transition from population to subspecies may be significant when only larger samples are included ($\mu_h > 60$, t = 2.65, uncorrected P = 0.019), but the number of populations involved is small.

Comparisons between pooled subspecies samples were conducted to partially compensate for sampling error. This resulted in the elimination of the population-level category. The smallest subspecies sample was n = 50 for *darwini fulvescens*. Mean sample size was n = 169. Phylogenetic level of comparison did not have a significant effect on matrix correlation values (uncorrected P = 0.095). Excluding *osilae phaeus*, which is supplemented beyond the Pongo population, further lowers the significance to P = 0.36.

Common Principal Components

The CPC analyses indicated a greater degree of differentiation among taxa than did the correlation analyses. For all clades, equality and proportionality of covariance matrices is rejected. The hierarchical level of matrix association is indicated on the phylogenetic tree in Figure 8. Thicker branches indicate a greater degree of shared structure. Only two clades share a CPC structure including all components with the largest eigenvalues: *osilae osilae* and *xanthopygus posticalis*. The *xanthopygus xanthopygus* clade is best described by a CPC model according to the AIC, but the partial CPC model (referred to as PCPC), with two components is rejected by the log-likelihood ratio test. The southern *limatus* clade is best described by a PCPC model with the first component (PCPC 1) accepted as common, but PCPC 2 is rejected. The first principal component is primarily a size vector. Only one clade above the subspecies level is found to share common structure, that being *P. limatus*. In *P. limatus*, the CPC model is the best fit, but the first component, PCPC 1, must be rejected. The same condition holds for *darwini darwini*. At all higher levels, minimally the first three PCPCs must be rejected, demonstrating statistically significant divergence in covariance structure. Common structure must also be rejected within the subspecies *xanthopygus chilensis* and *xanthopygus vaccarum*.

CPC analyses were also applied to the pooled subspecies covariance matrices, again to address the issue of sampling error in estimating covariance matrices. However, because the CPC model takes into account sample size in deriving the χ^2 probabilities, the principal effect of pooling the data is to increase the power of the method to reject common structure. Common structure is rejected for all clades among the pooled subspecies samples.

Phenotypic covariance matrices for *Peromyscus leucopus* and the two subspecies of *Peromyscus maniculatus* (Lofsvold 1986) were analyzed using CPC. The results are in agreement with those from the pooled subspecies analysis of *Phyllotis*. Common covariance structure is rejected among the subspecies and among the species. In both clades, the AIC indicates that the best fitting model is that of arbitrary or unrelated matrices.

DISCUSSION

Efficacy of the Methods

Because the two approaches employed here use different techniques, each has its own strengths and weaknesses. They

use different methods to estimate the degree of shared covariance structure, different null hypotheses, and different methods to account for phylogeny. For example, the matrix correlation analyses incorporate phylogeny by categorizing population pairs by the degree of phylogenetic relatedness. All comparisons possessing the same degree of cladistic relatedness are considered members of the same class and are grouped. Classes are defined by the number of nodes subtended via the common ancestors and by the definition of those nodes; for example, biological species. This approach is similar to that employed by Goodin and Johnson (1992) in how phylogeny is incorporated into the analysis, but uses matrix correlation as the measure of similarity in covariance structure rather than their index of factor loading similarity. In contrast, the CPC approach analyzes each clade separately, does not group by class, and is not limited to pairwise comparisons.

Matrix correlations is the more heuristic approach and has the advantage of providing an easily interpretable statistic that is also continuously distributed. That value is a description of overall similarity in covariance structure. It reflects patterns of covariation and is insensitive to differences in overall magnitudes of correlation/covariation. Unequal age distribution among population samples could result in proportional differences in the magnitude of correlations and covariances without modifying the patterns of correlations/covariances if the actual covariance structures were the same. Matrix correlation has several disadvantages as well. The principal shortcoming is that it does not provide a direct statistical test. Applying statistical tests to the resulting data is complicated by the inflation in degrees of freedom from the multiple comparisons in addition to the phylogenetic nonindependence among members of clades (Felsenstein 1985). Grouping by cladistic classes is a relatively crude adjustment for phylogeny. It also does not intrinsically distinguish the contributions of individual taxa or clades to the overall pattern. As an estimate of matrix similarity, matrix correlation is sensitive to sampling error, which results in a high degree of scatter among comparisons with small to moderate sample sizes.

The application of CPC to comparative studies as advocated here provides a direct statistical test of the null hypothesis of matrix identity. In an advance over other tests between matrices (e.g., matrix permutation) it decomposes matrix associations into a nested series of hypotheses regarding covariance structure. This then allows more detailed analyses of which components are shared within and among populations or clades and thus the nature of deviations in covariance structure. It also allows multiple groups to be analyzed, which is invaluable whenever polytomies are present. This is true whether the polytomies represent our ignorance of the true branching pattern, or if as in the case of interbreeding populations, relationships can not be accurately expressed as a dichotomously branching tree. In comparison to matrix correlation, CPC more explicitly accounts for phylogeny by analyzing each clade independently. CPC also has some weaknesses that must be remembered. As employed here, some phylogenetic information is ignored in analyses of the more inclusive clades. In effect, it treats each clade in turn as a polytomy or bush without any phylogenetic structure. This problem could be remedied if we could estimate effective sample sizes for matrices calculated at internal nodes (i.e., hypothetical ancestors) and then restrict analyses to only the direct descendants of a particular node. In the case of a fully dichotomous phylogeny, this would be a comparison between sister taxa, whether actual populations or estimated hypothetical ancestors. That approach is conceptually similar to Felsenstein's (1985) independent contrasts and would better isolate the phylogenetic location of significant divergence. One limitation of CPC is that it does not describe the nature of matrix association with a simple, continuously distributed variable that can then be used in further analyses.

Sample sizes and the number of characters affect the two approaches differently. Smaller sample sizes can greatly reduce the matrix correlation, suggesting less shared structure, while small sample sizes would only result in accepting a higher degree of shared covariance structure with CPC. There is no objective way to determine minimum sample size for matrix correlation analyses, but with this dataset sample sizes greater than 100 would be strongly preferred. Increasing the number of characters should increase the accuracy of matrix correlations as an estimate of overall structure, but it appears to increase the likelihood of incorrectly rejecting common structure with CPC.

The selection of characters included in the analyses can also affect the strength of matrix correlations. This study included traits with a wide range of correlations, from nearly 1.0 in the case of diastema and nasal length to slightly negative correlations involving interorbital breadth with nearly everything else. Including fewer very high and very low correlations would be expected to decrease the overall matrix correlations.

Comparison of Results

The two techniques applied to covariance evolution in this study present different conclusions. The uniformly high matrix correlations indicate a high degree of shared covariance structure and no significant decline with decreasing relatedness. In contrast, the CPC analyses show statistically significant divergence in covariance structure at all phylogenetic levels with fewer shared components at more inclusive levels. How can these conflicting results be reconciled?

The combined results indicate that while there is statistically significant divergence even among conspecific populations, the magnitudes of those divergences is small. Furthermore, there is no evidence for a long-term directionality to divergence in covariance structure. The pattern is more consistent with random variation in covariance structure among populations, perhaps by a Brownian motion model, although that model of evolution is not tested here. In contrast, long-term persistence of directional selection can result in clear divergence in phenotypic means.

An argument could be made that the cause of reduced shared structure with greater phylogenetic inclusivity may be due to the difficulty of recovering shared structure for any large set of randomly varying matrices. In that interpretation, the decreasing common structure at more inclusive clades would be a computational artifact of the number of OTUs

rather than accumulated interspecific divergence. However, sensitivity analyses indicate that with CPC analysis, number of matrices alone does not increase Type I error. Increasing the number of random subsamples (50% each) of a single population, up to 12 subsamples, did not prevent the recovery of shared structure, in that case accepting the hypothesis of matrix equality (unpubl. data). Therefore, the higher level divergences appear to be real and represent at least a small amount of phylogenetic structure present in the data.

There are at least three levels of sampling that should be considered in interpreting the results. First is genetic sampling (e.g., genetic drift) at the population level due finite sample sizes. That would result in variation in the genetic covariances. Second is environmental sampling by the populations. Each population experienced a different environment by virtue of existing in different places at different times. That would result in variation in the environmental covariances, which may then cause incongruence between genetic and phenotypic covariances (Willis et al. 1991). The third is sampling of the population at the level of the collector, which only affects the ability to estimate the actual phenotypic covariances. This last level is of course the font of statistics, and can be partially compensated for through rarefaction in the matrix correlation analyses or explicitly in the likelihood estimates in CPC.

Conclusions

CPC analyses demonstrate statistically significant divergence in phenotypic covariance structure at all levels of the phylogeny. Only within a minority of subspecies can even the common principal component model be accepted. The hypothesis of constant covariance matrices must be rejected. However, the magnitudes of those divergences appear to be small and are not associated with phyletic distances. Pairwise matrix correlations are very high, even for the most distantly related species. Neither analytic approach demonstrates a strong phylogenetic pattern, although both suggest greater shared covariance structure within subspecies than among them or higher taxa.

The phylogenetic level of the comparison did not have a significant effect on the pairwise similarity of covariance patterns. The greatest similarities were among populations of the same subspecies, while higher levels did not show any pattern. This phylogenetic pattern of covariance similarity seen in *Phyllotis* using matrix correlations was also apparent in the land snail *Partula* (Goodin and Johnson 1992). Goodin and Johnson (1992) also showed that subspecies were relatively more divergent from each other than are higher categories, an observation that finds only weak support in the two analysis presented here. There are no other studies fully comparable to the CPC analyses.

The matrix correlations approach is an effective heuristic technique, but is limited to describing the general pattern of matrix similarity. CPC is a powerful statistical tool that allows detailed examination of underlying patterns of covariation. It would be particularly promising for data with greater phylogenetic structure or with less population-level divergence.

Because of the conflicting nature of the results from these two approaches, and because the methods only partially account for phylogeny, a more explicitly phylogenetic method for comparing matrices is needed. A minimum evolution approach for reconstructing the evolution of covariance matrices and the broader implications of these results will be addressed in an accompanying paper (Steppan 1997).

Acknowledgments

I thank my committee members B. Patterson, B. Chernoff, S. Arnold, D. Jablonski, and J. Brown for their support, time, and input thoughout this project. For assistance with loans or making specimens available for study, I thank B. Stanley, J. Patton, B. Stein, G. Musser, R. Voss, S. Anderson, M. Carleton, P. Myers, C. Carmichael, L. Abraczinskas, M. Rutzmoser, S. McLaren, J. Braun, J. Yañez, J. C. Torres-Mura, A. Spotorno, and R. Ojeda. J. Phelps in particular processed a great many loans for me promptly and with great care. I am sincerely grateful to H. Nuñez, and S. and J. Anderson for opening their homes to me while I was visiting museums. I am grateful to A. Mayumi Kawamoto for her technical advice on graphics. B. Patterson provided important advise and input thoughout this project. I thank S. Arnold for pointing me toward common principal component analysis. P. Phillips generously supplied a version of his CPC program and provided input on interpretation of results. C. P. Klingenberg and an anonymous reviewer gave valuable suggestions to improve the paper. This research was supported by grants from the Hinds Fund of the University of Chicago, Center for Latin American Studies at the University of Chicago, American Society of Mammalogists Grants-in-Aid-of-Research, Sigma Xi, the Scientific Research Society, an American Museum of Natural History Collections Study Grant, the Searle Fellowship of the University of Chicago, and the Rowley Fellowship of the Field Museum.

Literature Cited

Airoldi, J., and B. K. Flury. 1988. An application of common principal component analysis to cranial morphometry of *Microtus californicus* and *M. ochrogaster* (Mammalia, Rodentia). J. Zool. (Lond.) 216:21–36.

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pp. 267–281 *in* B. N. Petrov and F. Csaki, eds. Second international symposium on information theory. Akademiai Kiado, Budapest, Hungary.

Arnold, S. J. 1981. Behavioral variation in natural populations. I. Phenotypic, genetic and environmental correlations between chemoreceptive responses to prey in the garter snake, *Thamnophis elegans*. Evolution 35:489–509.

———. 1992. Constraints on phenotypic evolution. Am. Nat. 140: S85–S107.

Arnold, S. J., and M. J. Wade. 1984. On the measurement of natural selection: Theory. Evolution 38:709–719.

Atchley, W. R., J. J. Rutledge, and D. E. Cowley. 1981. Genetic components of size and shape. II. Multivariate covariance patterns in the rat and mouse skull. Evolution 35:1037–1055.

Atchley, W. R., D. E. Cowley, C. Vogl, and T. McLellan. 1992. Evolutionary divergence, shape change, and genetic correlation structure in the rodent mandible. Syst. Biol. 41:196–221.

Berg, R. L. 1960. The ecological significance of correlation pleiades. Evolution 14:171–180.

Björklund, M. 1993. Phenotypic variation of growth trajectories in finches. Evolution 47:1506–1514.

Bookstein, F., B. Chernoff, R. Elder, J. Humphries, G. Smith, and R. Strauss. 1985. Morphometric in evolutionary biology.

Spec. Publ. Academy of Natural Sciences of Philadelphia, Philadelphia, PA.

BOULDING, E. G., AND T. K. HAY. 1993. Quantitative genetics of shell form of an intertidal snail: Constraints on short-term response to selection. Evolution 47:576–592.

BRODIE, E. D., III. 1993. Homogeneity of the genetic variance-covariance matrix for antipredator traits in two natural populations of the garter snake *Thamnophis ordinoides*. Evolution 47: 844–854.

CATZEFLIS, F. M., J.-P. AGUILAR, AND J.-J. JAEGER. 1992. Muroid rodents: Phylogeny and evolution. Trends Ecol. Evol. 7:122–126.

CHEETHAM, A. H., J. B. C. JACKSON, AND L.-A. C. HAYEK. 1993. Quantitative genetics of bryozoan phenotypic evolution. I. Rate tests for random change versus selection in differentiation of living species. Evolution 47:1526–1538.

CHEVERUD, J. M. 1982. Phenotypic, genetic, and environmental morphological integration in the cranium. Evolution 36:499–516.

———. 1988. A comparison of genetic and phenotypic correlations. Evolution 42:958–968.

———. 1996. Quantitative genetic analysis of cranial morphology in the cotton-top (*Sanguinus oedipus*) and saddle-back (*S. fuscicollis*) tamarins. J. Evol. Biol. 9:5–42.

COCK, A. G. 1966. Genetical aspects of metrical growth and form in animals. Q. Rev. Biol. 41:131–190.

DIETZ, E. J. 1983. Permutation tests for association between two distance matrices. Syst. Zool. 32:21–26.

EMERSON, S. B., AND S. J. ARNOLD. 1989. Intra- and interspecific relationships between morphology, performance, and fitness. Pp. 295–314 *in* D. B. Wake and G. Roth, eds. Complex organismal functions: Integration and evolution in vertebrates. Wiley, New York.

FELSENSTEIN, J. 1985. Phylogenies and the comparative method. Am. Nat. 125:1–15.

FLURY, B. 1988. Common principal components and related multivariate models. Wiley, New York.

FLURY, B. K. 1987. A hierarchy of relationships between covariance matrices. Pp. 31–43 *in* A. K. Gupta, ed. Advances in multivariate statistical analysis. Reidel, Dordrecht, The Netherlands.

GOODIN, J. T., AND M. S. JOHNSON. 1992. Patterns of morphological covariation in *Partula*. Syst. Biol. 41:292–304.

GOULD, S. J., D. S. WOODRUFF, AND J. P. MARTIN. 1974. Genetics and morphometrics in *Cerion* at Pongo Carpet: A new systematic approach to this enigmatic land snail. Syst. Zool. 25:518–535.

HERSHKOVITZ, P. 1962. Evolution of neotropical cricetine rodents (Muridae) with special reference to the phyllotine group. Fieldiana Zool. 46:1–524.

JAMES, F. C., C. NESMITH, AND R. LAYBOURNE. 1990. Geographic differentiation in the red-winged blackbird: A check of one of Lande's assumptions. Pp. 2454–2461 *in* B. D. Bell, R. O. Cossee, J. E. C. Flux, B. D. Heather, R. A. Hitchmough, C. J. R. Robertson, and M. J. Williams, eds. Acta XX Congressus Internationalis Ornithologici, New Zealand Ornithological Congress Trust Board, Wellington, New Zealand.

KLINGENBERG, C. P., AND J. R. SPENCE. 1993. Heterochrony and allometry: Lessons from the water strider genus *Limnoporus*. Evolution 47:1834–1853.

KLINGENBERG, C. P., AND M. ZIMMERMANN. 1992. Static, ontogenetic, and evolutionary allometry: A multivariate comparison in nine species of water striders. Am. Nat. 140:601–620.

KOHN, L. A. P., AND W. R. ATCHLEY. 1988. How similar are genetic correlation structures? Data from mice and rats. Evolution 42: 467–481.

LANDE, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. Evolution 33:402–416.

———. 1986. The dynamics of peak shifts and the pattern of morphological evolution. Paleobiology 12:343–354.

LANDE, R., AND S. J. ARNOLD. 1983. The measurement of selection on correlated characters. Evolution 37:1210–1226.

LOFSVOLD, D. 1986. Quantitative genetics of morphological differentiation in *Peromyscus*. I. Tests of homogeneity of genetic covariance structure among species and subspecies. Evolution 40:559–573.

———. 1988. Quantitative genetics of morphological differentiation in *Peromyscus*. II. Analysis of selection and drift. Evolution 42:54–67.

LYNCH, M. 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. Am. Nat. 136: 727–741.

MAYNARD SMITH, J., R. BURIAN, S. KAUFFMAN, P. ALBERCH, J. CAMPBELL, B. GOODWIN, R. LANDE, D. RAUP, AND L. WOLPERT. 1985. Developmental constraints and evolution. Q. Rev. Biol. 60:265–287.

MCCULLOCH, C. E., M. D. BOUDREAU, AND S. VIA. 1996. Confidence regions for evolutionary trajectories. Biometrics 52:184–192.

PAULSEN, S. M. 1996. Quantitative genetics of the wing color pattern in the buckeye butterfly (*Precis coenia* and *Precis evarete*): Evidence against the constancy of **G**. Evolution 50:1585–1597.

PEARSON, O. P. 1958. A taxonomic revision of the rodent genus *Phyllotis*. Univ. Calif. Publ. Zool. 56:391–477.

PEARSON, O. P., AND J. L. PATTON. 1976. Relationships among South American phyllotine rodents based on chromosome analysis. J. Mammal. 57:339–350.

PHILLIPS, P. 1994. CPC—common principal component analysis program. Univ. of Texas, Arlington.

PRICE, T. D., P. R. GRANT, AND P. T. BOAG. 1984. Genetic changes in the morphological differentiation of Darwin's ground finches. Pp. 50–66 *in* K. Wöhrmann and V. Loescheke, eds. Population biology and evolution. Springer, Berlin.

RADINSKY, L. B. 1985. Approaches in evolutionary morphology: A search for patterns. Annu. Rev. Ecol. Syst. 16:1–14.

RISKA, B. 1985. Group size factors and geographic variation of morphometric correlation. Evolution 39:792–803.

SCHLUTER, D. 1984. Morphological and phylogenetic relations among the Darwin's finches. Evolution 38:921–930.

SHAW, R. G. 1991. The comparison of quantitative genetic parameters between populations. Evolution 45:143–151.

SHAW, R. G., AND H. L. BILLINGTON. 1991. Comparison of variance components between two populations of *Holcus lanatus*: A reanalysis. Evolution 45:1287–1289.

STEPPAN, S. J. 1993. Phylogenetic relationships among the Phyllotini (Rodentia: Sigmodontinae) using morphological characters. J. Mammal. Evol. 1:187–213.

———. 1995a. Phylogenetic relationships of the phyllotine rodents (Sigmodontinae) and the evolution of phenotypic patterns of covariation in *Phyllotis*. Ph.D. diss. Univ. of Chicago, Chicago.

———. 1995b. Revision of the leaf-eared mice Phyllotini (Rodentia: Sigmodontinae) with a phylogenetic hypothesis for the Sigmodontinae. Fieldiana Zool. 80:1–112.

———. In press. Phylogenetic relationships and species limits within *Phyllotis* (Rodentia; Sigmodontinae): concordance between mtDNA sequence and morphology. J. Mammal.

———. 1997. Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. Evolution 51.

TURELLI, M. 1988. Phenotypic evolution, constant covariances, and the maintenance of additive variance. Evolution 42:1342–1347.

VENABLE, D. L., AND A. BÚRQUEZ M. 1990. Quantitative genetics of size, shape, life-history, and fruit characteristcs of the seed heteromorphic composite *Heterosperma pinnatum*. II. Correlation structure. Evolution 44:1748–1763.

VOSS, R. S., AND L. F. MARCUS. 1992. Morphological evolution in muroid rodents II. Craniometric factor divergence in seven Neotropical genera, with experimental results from *Zygodontomys*. Evolution 46:1918–1934.

VOSS, R. S., L. F. MARCUS, AND P. P. ESCALANTE. 1990. Morphological evolution in muroid rodents I. Conservative patterns of craniometric covariance and their ontogenetic basis in the neotropical genus *Zygodontomys*. Evolution 44:1568–1587.

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.

WALKER, L. I., A. E. SPOTORNO, AND J. ARRAU. 1984. Cytogenetic and reproductive studies of two nominal subspecies of *Phyllotis*

darwini and their experimental hybrids. J. Mammal. 65:220–230.

Wilkinson, G. S., K. Fowler, and L. Partridge. 1990. Resistance of genetic correlation structure to directional selection in *Drosophila melanogaster*. Evolution 44:1990–2003.

Willis, J. H., J. A. Coyne, and M. Kirkpatrick. 1991. Can one predict the evolution of quantitative characters without genetics? Evolution 45:441–444.

Zelditch, M. L., and A. C. Carmichael. 1989. Ontogenetic variation in patterns of developmental and functional integration in skulls of *Sigmodon fulviventer*. Evolution 43:814–824.

Zelditch, M. L., D. O. Straney, D. L. Swiderski, and A. C. Carmichael. 1990. Variation in developmental constraints in *Sigmodon*. Evolution 44:1738–1747.

Corresponding Editor: L. Leamy

Appendix

Specimens Examined in Population Samples

Measurements were made of skulls and jaws from specimens of *Phyllotis* belonging to the following collections: American Museum of Natural History (AMNH); Carnegie Museum of Natural History (CMNH); Field Museum of Natural History (FMNH); Laboratorio de Citogenética Mamíferos, Facultad de Medicina, Universidad de Chile (LCM); Museum of Comparative Zoology, Harvard (MCZ); Museo Nacional de Historia Natural, Santiago, Chile (MNHN); Museum of Southwestern Biology, University of New Mexico (MSB); The Museum, Michigan State University (MSU); Museum of Vertebrate Zoology, University of California, Berkeley (MVZ); Oklahoma Museum of Natural History (OMNH); University of Michigan, Museum of Zoology (UMMZ); and National Museum of Natural History, Smithsonian Institution (USNM).

andium (n = 71). PERU. Ancash: Huaraz (FMNH 81209–81215, 81217–81219, 81221–81227); Mácate (FMNH 20914–20918, 20923–20925, 20927, 20928, 20933–20935, 20937–20939, 20943, 20945, 20946, 20949, 20952–20958); Recuay (FMNH 81229–81233); Yungay (FMNH 81242–81247). Cajamarca: Cajamarca (FMNH 19466–19468, 19471, 19472); Balsas, Hac. Limon (FMNH 19476, 19477, 19836). Lima: Matucana (FMNH 23740, 23742); Surco (FMNH 53061); Santa Eulalia Canyon (FMNH 107357, 107360–107362).

osilae osilae Calacala (n = 28). PERU. Puno: Alta Calacala, 40 mi NNE Juliaca (MVZ 139565–139570); Arapa, 3 mi NE (MVZ 116683, 116684, 116686, 116687); Hac. Calacala, 7 mi SW Putina, 37 mi NNE Juliaca (MVZ 114692–114694, 114707, 116688, 116689, 139571); Hac. Chicayani, 20 km NE Azangaro (MVZ 136337–136339); Putnia, 6 km N (MVZ 171534–171541).

osilae osilae Santa Rosa (n = 50). PERU. Puno: Asillo, 10 km W (MVZ 171522, 171529–171531); Munani, 3.6 km W (MVZ 172767–172769, 173167–173171, 173291); Pucara, 6 km S (MVZ 172763–172765, 173165, 173166, 173288, 173290); San Anton, 4.5 km NE (MVZ 172749–172762); Santa Rosa, 12 km S (MVZ 171532, 171533, 172730–172732, 172735, 172737, 172738, 172742–172747, 173163, 173164).

osilae osilae South Ilave (n = 72). PERU. Puno: Caccachara, W Ilave (FMNH 42850); Hac. Ontave, 40 km S Ilave (MVZ 141567–141578); Hac. Pairumani, 40 km S Ilave (FMNH 42896, MVZ 114696–114703, 114710–114714, 114716, 114717, 115855, 115856); Ilave, 35 km S, 5 km W (FMNH 107828, 107831–107833, 107835, 107843, 107844, 107846, 107847, 107849, 107851, 107854, 107855, 107859, 107870–107872, 107874, 107881, 107885, 107888, 107891, 107895); Pomata, 4 km NW(MVZ 115865–115869, 115871, 115872); Yunguyo (FMNH 51265, 51267, 51269, 51270, 51272–51277).

osilae phaeus Pongo (n = 56). BOLIVIA. La Paz: Pongo (AMNH 72916, 72918–72926, 72928–72931, 72933–72940, 72942, 72943, 72945, 72947–72954, 72958–72961, 72963, 72966–72972, 72974–72977, 72984–72986, 72988, 72992, 72993, 81280).

osilae tucumanus Tafi del Valle (n = 60). ARGENTINA. Tucumán: Tafi del Valle, El Infernillo (ARG 1272, 1282, 1300–1303, 1305, 1306, 1318–1320, 1332, 1352, 1356; CMNH 43996–44002, 44004–44008, 86452, 86453, 86530–86532, 86534; MSU 19216, 19217; OMNH 15266–15274, 15279, 15287, 15284, 19375–19381, 44011–44015, 44017, 44018).

magister Tarata (n = 103). PERU. Tacna: Tarata (FMNH 107611–107613, 107616, 107620, 107622, 107623, 107625, 107629, 107631, 107653, 107654; MVZ 141513–141534, 141536, 141537, 115874–115884, 139384–139432, 139438–139445).

darwini fulvescens Angol (n = 48). CHILE. Malleco: Angol (AMNH 93256, 93258, 93259, 93262, 93265, 93268, 93271, 93274, 93277–93285, 93288–93290; MSU 7430, 7432, 7434, 7436); Los Alpes, 5 km S Angol (MSU 5677, 7019, 7021, 7413, 7415, 7416, 7419-7423, 7426, 7427, 7429, 7528; MCZ 10540, 10541); Angol, 10–18 km WNW (MSU 7411, 7412); Collipulli, 14 km NW (MSU 7438–7440); Sierra Nahuelbuta (FMNH 50550; holotype).

darwini darwini Santiago (n = 71). CHILE. Santiago: Bocatoma (AMNH 391807); 2.5 km NE Cerro Manquehue (FMNH 119491–119493, 119496–119500, 119504, 119505); Farellones (LCM 737, 741, 746); Fundo Santa Laura (MVZ 150061–150063, 150065, 150066); La Dehesa (LCM 318); Las Condes (AMNH 391805); Los Dominicos (MNHN 535, 544, 548); Quebrada de la Plata, Maipu (MNHN 644, 646, 665, 677, 679, 680, 747, 756); Rinconada de Maipu (AMNH 541761); San Cristobal (FMNH 35901, 35902); Til Til (FMNH 119505, 119506; MNHN 977). Valparaíso: Buen Retiro, Calera (FMNH 23889–23890); 4.5 km N Caleta Los Molles (MNHN 573); Cuesta Las Chilcas, Com. Llay Llay (MNHN 1234, 1235, 1273, 1456); La Rojas (FMNH 23882, 23883); Olmue (FMNH 22347, 22348); Palmilla (FMNH 24394); Peñuelas (MNHN 1166, 1167, 1169, 1364, 1365, 1370); 7 km SSE Papudo (MNHN 708); Quilpue (MSU 2102); Reserva Forestal Peñuelas (MNHN 1090, 1094, 1102, 1131, 1133, 1172, 1174, 1183, 1186).

darwini darwini Fray Jorge (n = 38). CHILE. Coquimbo: Asentamiento Ceeres, La Serena (MNHN 973, 974); La Serena, Rio Limon (LCM 365); Las Palmas, 95 km N Los Vilos (MVZ 150060); Las Tacas (LCM 220, 221, 233, 234, 258, 263, 790); Parque Nacional Fray Jorge (FMNH 119512, 133874, 133877, 133881, 133894; MNHN 319; 1212, 1214–1220, 1222, 1223, 1227; MVZ 118662); Romero (FMNH 22325–22329).

northern *limatus* (n = 75). PERU. Arequipa: Bella Union, 8 mi NNW (MVZ 145562, 145563–145581); La Planta (MVZ 145582, 145583); Ayacucho: Nazca, 35 mi ENE (MVZ 138091–138094). Huancavelica: Ticrapo, Pisco Valley (MVZ 136309–136311, 136313–136319). Ica: Hac. San Jacinto (FMNH 53162, 53163); Humay, 3 km E, Pisco Valley (MVZ 136320–136325); Pisco, 10 km SSE (MVZ 136326–136328). Lima: Chilca, 8 mi SE (MVZ 138089, 138090); Chillon Valley (MVZ 136329, 136330); Chosica (FMNH 53164, 53169, 53170; MVZ 120058–120063, 120066); Oscolla (FMNH 53056); Rimac Valley (MVZ 120067–120071); San Bartolome Station (MVZ 120026, 120072, 120073); Santa Eulalia (MVZ 120074–120076, 120078); Yangas (MVZ 136332).

southern *limatus* Tarata (n = 85). PERU. Tacna, Tarata, 16 km S to 8 km NE (FMNH 107574, 107575, 107598, 107602, 107603, 107605, 107609, 107610, 107614, 107615; MVZ 115837, 115838, 115840–115846, 116788, 139313–139333, 139335–139341, 139348, 139365, 139369, 139370, 139376, 139378, 141423, 141634–141636, 141639–141651, 141653–141659, 141661–141665).

southern *limatus* Torata (n = 42). PERU. Moquegua, Torata, 3 km N to 20 km NE (FMNH 107403–107407, 107410, 107411, 107413–107415, 107418, 107420–107427, 107429, 107430, 107432, 107435, 107437, 107438, 107440, 107442, 107443, 107444, 107476, 107477, 107482, 107484; MVZ 115790, 115791, 115793, 115796–115800).

southern *limatus* Arequipa (n = 73). PERU. Arequipa: Arequipa, 7 km E (MVZ 115779–115785), 12 km E (MVZ 136300–136307), 12 km SSW (MVZ 115786–115789), 15 km E (MVZ 139560); Balneario de Jesus (FMNH 50991–51002, 53161); Salinas (FMNH 49480–49488, 49637); Yura (FMNH 49451–49479, 49608, 49748).

southern *limatus* Moro Sama (n = 46). PERU. Tacna: Moro Sama, 65 km W Tacna (MVZ 141492, 141493, 141498–141500, 141502–141509, 143713–143716, 143718–143738, 143741, 143742, 143744, 143750, 143751, 143757, 143758).

xanthopygus posticalis Huancavelica (*n* = 28). PERU. Huancavelica: Huancavelica (FMNH 75427, 75441, 75443–75455); Lircay (FMNH 75456–75459); Locroja, Hac. Piso (FMNH 75425, 75426); Mayoc (FMNH 75436, 75439); San Jenaro (FMNH 75437, 75460–75462, 75465).

xanthopygus posticalis Puquio (*n* = 24). PERU. Ayacucho: Puquio, 2 km E (MVZ 139306–139309), 18 km E (MVZ 174042), 35 km E (MVZ 115808), 15–21 km NE (MVZ 115809, 116024), 10–15 km WMW (MVZ 138098, 138099–138113).

xanthopygus chilensis Tarata (*n* = 139). PERU. Tacna: Tarata, 2.6 mi N (MVZ 139342-139347, 139349, 139351, 139353, 139355-139363, 139366–139368, 139371–139375, 139377, 139379–139381), 4 km N (MVZ 115832, 115833, 115835, 115836, 115847, 115848), 2 km NE (MVZ 141638), 5–8 km NE (FMNH 107550-107552, 107556-107560, 107562–107566, 107571, 107572, 107576, 107577, 107582, 107588, 107590, 107636–107638, 107595, 107596, 107604, 107606–107608, 107635, 107648; MVZ 141422, 141424–141427, 141429–141431, 141433–141436, 141438–141440, 141442-141446, 141448, 141451–141458, 141460, 141461, 141553–141555, 141557, 141564–141566), 13 km NE (MVZ 141463, 141465–141477, 141480–141482, 143753–143756), 20-25 km NE (MVZ 115849–115854), 10 km S (MVZ 141652); Livine, 21 mi NE Tarata (MVZ 114687–114690); Nevado Livine, 2 km NW (MVZ 115839).

xanthopygus chilensis Torata (*n* = 46). PERU. Moquegua: Torata, 19 km NE (FMNH 107494–107496), 24–27 km NNE (FMNH 107445–107452, 107454–107456, 107458, 107459, 107471, 107473, 107474, 107485–107487, 107489, 107490), 31–35 km NNE (FMNH 107498–107508, 107510, 107512–107515, 107519, 107520, 107524, 107527, 107529, 107537, 107545).

xanthopygus chilensis Arequipa (*n* = 31). PERU. Arequipa: Cailloma, 16–20 km SE (FMNH 107757, 107759, 107775) Callalli, 15 km S (MVZ 174028–174031); Chiguata, 8 km E (FMNH 107778–107782, 107778–107790, 107796, 107798, 107800, 107802); Chivay (FMNH 107668, 107672, 107677, 107682, 107685, 107686, 107695, 107697, 107706, 107708, 107715, 107730).

xanthopygus chilensis Santa Rosa (*n* = 42). PERU. Puno: Santa Rosa, 3–12 km W (FMNH 107898, 107900–107905, 107907, 107909–107911, 107914–107916, 107924, 107927, 107929, 107934–107939, 107942, 107945, 107947, 107948, 107955, 107957, 107958, 107961, 107963, 107965, 107966, 107972, 107973, 107977, 107978, 107980, 107986, 107992, 107994).

xanthopygus chilensis Cailloma (*n* = 43). PERU. Arequipa: Cailloma (FMNH 49507–49515, 49519–49525, 49533, 49626–49630, 49632–49635); Sibayo (FMNH 39508, 39509, 39511–39513); Sumbay (FMNH 49501–49506, 49516–49518, 49609, 49610, 49612, 49625).

xanthopygus chilensis Puno (*n* = 92). PERU. Puno: Arapa (MVZ 116685); Asillo, (FMNH 51245–51248), 5 km S (MVZ 116130); Chicuito (FMNH 52584, 52586, 52588–52592, 52595); Hac. Collacachi (FMNH 49526–49532, 49534, 49535, 49776); Huacullani (FMNH 52576, 52577, 52580, 52581); Rio Huanque (MVZ 136333, 139382); Ilave, 15 km S (MVZ 115816, 115817); Imata, 6 mi S (MVZ 116131, 116132); Juli (MCZ 39502, 39503, 39506; MVZ 115818–115821, 115885); Juliaca, (FMNH 49489, 49490, 49621), 6 km NNW (136334, 136340–136342), 3 mi SW (MVZ 116690); Mazocruz, 30 mi S (MVZ 116786); Occomani (FMNH 53182); Hac. Pairumani, 24 mi S Ilave (MCZ 39518; MVZ 114684, 115815, 115887); Pampa de Ancomarca (MVZ 115822-115824); Pisacoma (FMNH 52601–52604); Pomata (MVZ 115873, 115970); Pucara (MVZ 172776); Puno (AMNH 213613, 213615, 213617–213619, 213623, 213625; MVZ 116133, 116146), 5–15 km W (MVZ 115825, 115826, 115888, 115889, 115891), 82 km W (MVZ 115827, 115828); San Antonio de Esquilache (FMNH 49638, 49709); San Ignacio (MCZ 39507, 39514, 39515); Hac. Umayo, 15 mi S Juliaca (MVZ 116691, 116694–116696); Yunguyo (FMNH 51266).

xanthopygus vaccarum Paiguano (*n* = 42). CHILE. Coquimbo: Paiguano (AMNH 259593, 259594; FMNH 22251, 22260, 22262–22267, 22269, 22272, 22273–22288, 22292–22296, 22298-22301, 25278–25281, 25284).

xanthopygus vaccarum rt. 40, Mendoza (*n* = 40). ARGENTINA. Mendoza: Cerro Medio, 35 km WNW 25 de Mayo (MVZ 181569, 181570); El Manzano Monument, 10 km W (CMNH 43977–43979); 10 km W old rt. 40, road to Lago Diamante (IADZA 2194, 2199, 2206, 2250; OMNH 15259–15264, 15318–15321, 15325, 15329–15331, 15362–15366, 15367–15378).

xanthopygus xanthopygus Chile Chico (*n* = 121). CHILE. Aisén: Chile Chico (FMNH 133934, 133935, 133938, 133940, 133944, 133946–133950, 133952, 133953, 133955–133959, 133961, 133962, 133964–133968, 133970, 133971–133977, 133980, 133983–133986, 133988, 133989, 133991–133998, 134001–134004, 134006, 134007, 134009–134013, 134070, 134073–134076, 134103, 134105, 134106, 134108–134112, 134114, 134115, 134117, 134118, 134120, 134121, 134124, 134126–134128, 134130, 134132, 134133, 134135–134139, 134141, 134142, 134145–134147; MNHN 469, 1405, 1414, 1423, 1433, 1436, 1442, 1445, 1491, 1494, 1506, 1509, 1515, 1560, 1584, 1628, 1631, 1637, 1644, 1662, 1900, 3629, 3955, 3961, 3964, 3970, 3973, 3990, 3999).

xanthopygus xanthopygus Puerto Ibañez (*n* = 83). CHILE. Aisén: Puerto Ibañez (FMNH 134015, 134021, 134022, 134024, 134025, 134028–134031, 134034, 134036–134040, 134042, 134045–134049, 134051, 134052, 134154–134056, 134058, 134060, 134061, 134063–134067, 134069, 134079, 134081, 134082, 134090–134093, 134097, 134099, 134100, 134102, 134148, 134150, 134151, 134153–134157, 134159, 134160, 134163–134166, 134168, 134169; MNHN 1411, 1665, 1705, 1708, 1752, 1755, 1761, 1774, 1827, 1849, 1852, 1897, 3476, 3593, 3623, 4007, 4071, 4074, 4080, 4099, 4115).

xanthopygus xanthopygus Huanuluan (*n* = 42). ARGENTINA. Río Negro: Bariloche, 15–25 km ENE (MVZ 151029, 158469, 159412, 163810, 171145–171148); Comallo, 8 km WSW (MVZ 164024–164027, 165848, 169012, 179307); Huanuluan (MCZ 18996–19000, 19026–19035, 19038, 19040, 19042–19045, 29114; MNHN 4118); Pilcaniyeu (MCZ 23837, 29115; MVZ 21944).