ORIGINAL ARTICLE

doi:10.1111/j.1558-5646.2007.00129.x

EVOLUTIONARY DIVERGENCE IN DIRECTIONS OF HIGH PHENOTYPIC VARIANCE IN THE OSTRACODE GENUS *POSEIDONAMICUS*

Gene Hunt

Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, MRC 121, P.O. Box 37012, Washington, DC, 20013-7012

E-mail: hunte@si.edu

Received October 2, 2006
Accepted February 24, 2007

Trait variation and covariation are understood to influence the response of populations to natural selection on generational time scales, but their role, if any, in shaping long-term macroevolutionary divergence is still unclear. The present study uses the rich fossil record of the ostracode genus *Poseidonamicus* to reconstruct in great detail the evolutionary history of a set of landmark-based morphometric characters. This reconstruction included two kinds of evolutionary inferences: ancestor–descendant transitions among populations repeatedly sampled at the same location and divergence between lineages measured as independent contrasts on a phylogeny. This reconstructed history was then used to test if evolutionary changes were concentrated in directions (traits or combinations of traits) with high phenotypic variance. Two different statistics of association between evolution and variation tested the null hypothesis that evolutionary changes occur in random directions with respect to trait variability. The first of these measured the similarity between the directions of evolutionary change and the axis of maximum variance, and the second measured the degree to which evolutionary changes were concentrated in directions of high phenotypic variation. Randomization tests indicated that both kinds of evolutionary inferences (ancestor–descendant and phylogenetic contrasts) occurred preferentially in directions of high phenotypic variance (and close to the axis of maximal variation), suggesting that within-population variation can structure long-term divergence. This effect decayed after a few million years, but at least for one metric, never disappeared completely. These results are consistent with Schluter's genetic constraints model in which evolutionary trajectories on adaptive landscapes are deflected by variation within and covariation among traits.

KEY WORDS: **Genetic constraints, independent contrasts, macroevolution, microevolution, *Poseidonamicus*, variation.**

Biologists have long known that variation is the raw material of evolution. On generational time scales, quantitative genetic approaches have successfully incorporated variation within, and covariation among traits into predictive models of evolution (Conner and Via 1992; Grant and Grant 1995; Falconer and Mackay 1996; Via and Shaw 1996). Despite considerable interest in the problem, the influence of variation patterns on evolutionary divergence over longer, macroevolutionary time scales is still actively debated.

Although the absence of heritable variation precludes evolution over all temporal scales, whether abundant within-population variation actually facilitates long-term evolutionary divergence is less clear. This issue has arisen in a number of different contexts over the past half century of evolutionary inquiry. Empirical studies testing the link between trait variability and evolutionary lability date back to Simpson (1953), who compared rates of evolution in fossil taxa that differed in overall phenotypic variability. Other workers followed with similar studies over several decades (Bader 1955; Guthrie 1965; Kluge and Kerfoot 1973; Kelley 1983). Although some studies found a tendency for traits that are highly variable within populations to also differ most between populations

(e.g., Guthrie 1965; Kluge and Kerfoot 1973; Sokal 1976; Rohlf et al. 1983 and references therein), Rohlf and colleagues (1983) cautioned that this result may be a statistical artifact of analyzing imprecise and heterogeneous datasets.

Fully multivariate and process-based approaches to this problem were made possible by the development of a mathematical framework for phenotypic evolution on a multivariate adaptive landscape (Lande 1979). In this framework, a population evolves according to the equation $\mathbf{z} = \mathbf{G\beta}$, where $\mathbf{z}$ is the vector of evolutionary response across a generation, $\mathbf{G}$ is the additive genetic variance–covariance matrix, and $\boldsymbol{\beta}$ is the vector of selection gradients. Natural selection pushes a population directly up adaptive hills, but this push is deflected by genetic constraints embodied by the $\mathbf{G}$ matrix. One important result from this model is the prediction that under most conditions, a population evolving in the vicinity of an adaptive peak will reach the optimum (Lande 1979; Via and Lande 1985). Trait variances and covariances can deflect the initial response of a population toward directions of higher genetic variance, but eventually the population will evolve directly uphill and come to rest on the optimum. Thus, although patterns of genetic variance and covariance are important over the short term, the net evolutionary change in a population is determined only by the position of the selective peak (Zeng 1988). Variation patterns may have a more lasting influence if variance is lacking in any morphological direction (Barton and Turelli 1989), some or all traits are selectively neutral (Lande 1979; Bürger 1986; Wagner 1988; Price and Langen 1992), or if the adaptive surface has multiple peaks (Kirkpatrick 1982; Price et al. 1993).

Even the temporary channeling by genetic constraints as a population climbs an adaptive peak has the potential to endure for geologically significant periods of time. The predicted duration of genetic constraints in this scenario depends on factors, such as the shape and temporal constancy of the fitness surface, that are never completely known for natural populations. Schluter (1996) was the first to suggest on empirical grounds that the effects of genetic constraints may in fact be long-lasting enough to have a detectable macroevolutionary influence. In analyzing divergence in several vertebrate taxa, he found that evolution had occurred preferentially in morphological directions close to the axis of maximal genetic variation ($\mathbf{g}_{\mathbf{max}}$). Because evolutionary responses are channeled by genetic constraints toward $\mathbf{g}_{\mathbf{max}}$, Schluter's (1996) dubbed this axis a line of "least evolutionary resistance." This effect was temporary, lasting a few million years, but persistent enough to leave a detectable imprint on diversification within clades (Schluter 1996, 2000).

Other empirical studies followed, and by now there are several that have found that evolution occurs preferentially close to the direction of maximal variation (Mitchell-Olds 1996; Badyaev and Foresman 2000; Bégin and Roff 2003, 2004; Marriog and Cheverud 2005; Renaud et al. 2006), or more generally that a correspondence exists between variation within populations and divergence within species and clades (Ackermann and Cheverud 2002; Blows and Higgie 2003; Hansen et al. 2003). However, there are other examples that show more complex relationships between variation and divergence (Lofsvold 1988; Venable and Búrquez 1990; Merilä and Björklund 1999; Badyaev and Hill 2000; Ackermann and Cheverud 2004; Renaud et al. 2006), and further empirical studies of different clades and different kinds of traits are needed to delimit the scope and magnitude of this effect.

The purpose of the present study is to test if evolutionary changes have been concentrated in directions of high phenotypic variance in the ostracode genus *Poseidonamicus*. The rich deep-sea fossil record of this genus permits the inference of evolutionary transitions at two scales of resolution: ancestor–descendant sequences within lineages, and transitions between lineages based on a phylogenetic hypothesis. The inference of a large number of evolutionary transitions has allowed for a detailed reconstruction of the history of this genus, and a correspondingly powerful test of the hypothesis that evolutionary changes should be concentrated in directions of high morphological variance.

Materials and Methods

Testing the hypothesis that within-population variation structures long-term evolution proceeded in five basic steps: (1) geometric morphometric data were obtained from specimens from multiple populations; (2) trait phenotypic variance–covariance matrices were estimated for a subset of populations with high sample sizes; (3) evolutionary transitions between populations were inferred, including both ancestor–descendant relationships and contrasts between phylogenetically related populations; (4) the concordance between evolutionary changes and patterns of phenotypic variation was measured using two different metrics; and (5) the observed concordance was compared with that expected if evolutionary changes were random with respect to phenotypic variation.

STUDY CLADE AND SAMPLED POPULATIONS

The ostracode genus *Poseidonamicus* has been broadly distributed in deep marine environments since the Eocene. This taxon has a robustly mineralized, bivalved carapace that is abundantly preserved in deep-sea sediments. In addition to its rich fossil record, *Poseidonamicus* is otherwise well suited for evolutionary analysis: its phylogenetic relationships have been studied recently (Hunt 2007), its carapace morphology is complex with many landmarks suitable for morphometric analysis, and it has sexes and growth stages that can be discriminated, even in fossils.

The present study includes 915 individuals from 51 populations scattered throughout the 40-million year history of this genus (Table 1). Most populations include individuals from a

Table 1. Populations used in the current analysis. Species assignments are from Hunt (2007), ages are in millions of years before the present, and *N* indicates the number of measured individuals. Population labels correspond to those in Figure 2.

Label	Species	Age	*N*
ant-E1	*dinglei*	37.06	18
ant-E2	*dinglei*	34.11	16
din-O1	*dinglei*	31	10
gr1-H1	species 3	0	12
gr3-H1	species 4	0	14
gr3-H2	species 4	0	21
gr3-P1	species 4	3.21	25
gr3-P2	species 4	2.67	12
maj-H2	*major*	0	29
maj-P2	*major*	2.4	20
maj-P3	*major*	2.19	22
maj-Q1	*major*	3.33	12
maj-Q2	*major*	1.9	23
maj-Q3	*major*	1.68	20
maj-Q4	*major*	0.89	25
maj-Q5	*major*	0.14	20
min-H1	*minor*	0	9
min-Q1	*minor*	1.09	9
mio-M1	*miocenicus*	8.67	9
mio-M2	*miocenicus*	5.62	25
mio-M3	*miocenicus*	7.6	12
mio-P1	*miocenicus*	4.46	19
mio-P2	*miocenicus*	3.96	11
mio-P4	*miocenicus*	3.58	9
mio-P5	*miocenicus*	5.49	20
mio-PA	*miocenicus*	5.41	10
mio-PB	*miocenicus*	5.1	14
mio-PC	*miocenicus*	4.76	25
mio-PD	*miocenicus*	4.54	18
mio-PE	*miocenicus*	4.26	17
mio-PF	*miocenicus*	4.07	26
mio-PG	*miocenicus*	3.85	25
pin-H1	*pintoi*	0	10
pin-H2	*pintoi*	0	11
pin-H3	*pintoi*	0	17
pin-Q2	*pintoi*	0.11	25
pin-Q3	*pintoi*	0.09	40
pin-Q4	*pintoi*	0.08	32
rio-M2	*riograndensis*	21.53	17
rio-M3	*riograndensis*	19.09	32
rio-M4a	*riograndensis*	11.71	11
rio-M4b	*riograndensis*	11.61	24
rio-M5	*riograndensis*	10.26	15
rio-M7	*riograndensis*	10.82	19
rio-P1	*pintoi*	3.96	13
rud-M2	*rudis*	15.7	28
rud-O1	*rudis*	26.9	10
spE-O1	species 2	28.06	15
spE-O2	species 2	26.13	13
spF-E1	species 1	34.33	9
spF-O1	species 1	32.09	17

single dredge or core sample; a few lump individuals from closely spaced core samples. To standardize with respect to ontogeny, only individuals from the last juvenile instar were analyzed. In *Poseidonamicus*, instars can be recognized on the basis of size clustering and a suite of instar-specific morphological traits (Hunt 2007). The last juvenile instar was chosen for analysis because it is usually the most abundant in deep-sea samples and lacks the sexual dimorphism of the adult stage (which would reduce sample sizes by requiring separate analysis of males and females). Ostracodes have gradual development (Cohen and Morin 1990), and individuals from the last juvenile instar are morphologically very similar to adults.

Geological ages of populations were determined by plotting the age of known biostratigraphic datums against the depth at which they appear in the core. In the few cases in which they were available, published age-depth models were used (Cronin et al. 1996, 1999). For the remaining cores, new age models were created, mostly using biostratigraphic occurrence data from the Initial Reports of the Deep-sea Drilling Program, updated to recent syntheses of Cenozoic chronostratigraphy (Berggren et al. 1995a, 1995b). Stratigraphers calibrate the absolute ages of biostratigraphic events using cores with good magnetostratigraphic control; this magnetostratigraphic framework is in turn tied to the absolute time scale by radiometrically dated tiepoints (Berggren et al. 1995b). Because of the practical importance of deep marine sediments in paleoclimatic studies, the deep sea has been subjected to extensive chronostratigraphic investigation. Consequently, age control is usually quite good, relative to other depositional environments. Although some cores are better constrained than others, I estimate that population ages in the present study are typically accurate to about 10%, which amounts to several million years for the oldest samples, but less than one million years for the youngest. All age models used in the present study are available on request from the author.

MORPHOMETRIC DATA

Each specimen was digitally imaged using a scanning electron microscope. Ten landmarks were chosen to cover the major regions of the valve (Fig. 1). Eight of these landmarks were defined by the location of homologous ridges, and two were based on the location of identifiable pores (Table 2). Although observable in the mineralized valves, both kinds of landmarks are closely associated with underlying soft anatomy. The pores are perforations in the valves through which sensory setae protrude (Okada 1982b; Maddocks 1992). In addition, anatomical work on modern ostracodes with similar surface ornament to *Poseidonamicus* has found that the skeletal ridges are patterned by an underlying field of epidermal cells (Okada 1981, 1982a; Keyser 1995). The ridges correspond to the junctions of these epidermal cells, and thus each polygon in the skeletal mesh outlines a single

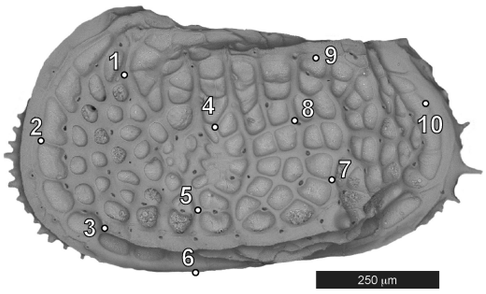

Figure 1. Left valve of *Poseidonamicus pintoi* showing location of 10 landmarks used in morphometric analysis. See Table 2 for information about landmark definitions.

epidermal cell (Fig. 1). Because these features often differ between closely related species, but have not been observed to vary systematically with environmental conditions, the conventional wisdom among ostracode workers is that they are genetically determined (e.g., Liebau 1971; Benson 1972; Maddocks 1992; Irizuki 1993), although no breeding studies have been performed to estimate their heritabilities quantitatively. Because the 10 landmarks are arrayed over the whole valve, landmark positions likely reflect overall carapace shape, in addition to more local processes that influence the development of cells in the vicinity of each morphometric landmark.

Digitizing error was assessed by reacquiring landmark data for a subset of 10 specimens from a single population. The resulting error was low, constituting on average less than 4% of the within-population variation (Table 2). Because there were no systematic differences between left and right valves according to a multivariate analysis of variance (MANOVA), data from both valves were lumped within populations after first reflecting right-valve landmarks over the vertical axis.

All measured individuals were subjected to a generalized full Procrustes superimposition (Rohlf and Slice 1990; Dryden and Mardia 1998). This procedure translates, scales, and rotates all landmark configurations such that they best fit (in a least-squares sense) a mean form. After superimposition, all information about specimen shape is retained in the residuals around this mean form. Principal components analysis (PCA) was performed on the Procrustes residuals, and the resulting PC scores were used as shape variables.

Although Procrustes residuals are themselves shaped variables, they have two disadvantages in the study of trait variation and covariation. First, variances and covariances among the residuals depend on the orientation of the mean form after fitting, but there is generally no basis for preferring any particular orientation. Second, the process of fitting removes four directions of variation among the residuals, one for each parameter that is estimated (scale, rotation, and translation in horizontal and vertical directions). This induced absence of variation in four multivariate directions has the potential to confound attempts to dissect the relationship between variation and evolution in a multivariate space. PC scores have neither of these shortcomings of raw Procrustes residuals. They are invariant to changing the orientation of the mean form, and because the last four PC axes have no variation, they can be safely dropped from the analysis. For 10 landmarks, PCA yields 16 orthogonal axes of variation that span the space of the original Procrustes residuals (Dryden and Mardia 1998).

Table 2. Landmark definitions. Most of the landmarks in the present study (Fig. 1) are defined as the intersection of specific ridges that are identifiable in all analyzed species of *Poseidonamicus*. The exceptions are landmarks 6 and 10, which are based on the position of a pore, projected to the edge of a nearby ridge, and landmark 9, which is defined as the midpoint of the ventral edge of a specific ridge. Error is the digitizing error, expressed as a percentage of within-population phenotypic variance.

No.	Landmark basis	Error
1	Ridge intersection	6.2%
2	Ridge intersection	1.3%
3	Ridge intersection	6.4%
4	Ridge intersection	2.7%
5	Ridge intersection	4.5%
6	Pore location	6.0%
7	Ridge intersection	1.8%
8	Ridge intersection	1.8%
9	Ridge midpoint	4.4%
10	Pore location	1.9%

PHENOTYPIC COVARIANCE PATTERNS

Covariance matrices of PC scores were estimated for the five populations represented by at least 25 measured individuals (Table 3). In addition, another eight covariance matrices were estimated by

Table 3. Estimated phenotypic covariance matrices, which were estimated for the five populations drawn from single samples with 25 or more measured individuals (first column). In addition, eight additional covariance matrices were estimated from pooling two or three closely spaced samples from the same core (second column). For each, the label for the pooled covariance matrix is given followed by the its included populations.

Single populations	Pooled estimates
gr3-P1	ant-E1.2 = ant-E1, ant-E2
maj-H2	maj-P2.3 = maj-P2, maj-P3
mio-M2	maj-Q2.3 = maj-Q2, maj-Q3
rio-M3	mio-P1.2 = mio-P1, mio-P2
rud-M2	pin-Q2.3.4 = pin-Q2, pin-Q3, pin-Q4
	rio-M4 a.b = rio-M4a, rio-M4b
	rio-M5.7 = rio-M5, rio-M7
	spE-O1.2 = spE-O1, spE-O2

pooling the covariance matrices of two or three closely related populations from the same deep-sea core (Table 3). A covariance matrix that is "pooled" over multiple populations is equal to the average of the covariance matrices computed separately, weighted by the degrees of freedom of each matrix:

$$\frac{1}{N-k}\sum_{i=1}^{k}(n_i-1)\,\mathbf{P}_i$$

where $\mathbf{P}_i$ is the phenotypic covariance matrix and n_i is the sample size associated with the ith population, and N is the total number of individuals across all pooled populations.

When two populations share the same underlying phenotypic covariance matrix, pooling provides the best estimate of this matrix. Even when the true covariance patterns are different in the two populations, their pooled covariance matrix provides a reasonable estimate for the ancestral covariance structure (see Bégin and Roff 2004). Although a complete analysis of the evolution of phenotypic covariance matrices in *Poseidonamicus* will be presented elsewhere, for the present purposes, it is sufficient to note that (1) none of the matrices combined into a pooled covariance matrix (Table 3) are demonstrably unequal (i.e., their dissimilarity is no greater than that expected from sampling error, as indicated by simulating random samples from their pooled covariance matrix); and (2) Mantel tests (see Renaud et al. 2006) and Common Principal Components analysis (Flury 1988) indicate that covariance matrices across the whole genus are generally conserved (significantly similar, with substantial shared principal component structure), and thus are stable enough to potentially influence long-term evolutionary divergence (Steppan et al. 2002).

INFERRING EVOLUTIONARY TRANSITIONS

Two different approaches to estimate evolutionary changes are possible in paleontological studies. The first considers a series of populations arrayed over time and infers ancestor–descendant relationships between them on the basis of morphological continuity and relative stratigraphic position. The second approach uses a phylogenetic hypothesis of relationships to infer the nature of evolutionary changes. These two kinds of inference are complementary in that they capture different types of evolutionary transitions. Ancestor–descendant transitions, when correctly inferred, reflect anagenetic (within lineage) evolution. In contrast, comparisons drawn from a phylogeny often include lineage splitting, in addition to phenotypic divergence within lineages. In an effort to reconstruct evolutionary changes in as much detail as possible, the present study analyzes phenotypic divergences from both kinds of evolutionary transitions.

ANCESTOR–DESCENDANT TRANSITIONS

Populations sampled from the same locality over time were treated as a single evolving lineage if there appeared to be morphological continuity through the sequence. These preliminary judgments were then tested against a phylogenetic analysis of the genus (Hunt 2007), with the prediction that putative ancestor–descendant populations should be closely related to each other (generally monophyletic, but conceivably paraphyletic) and that ancestors should not be characterized by many derived features that are not present in descendants (Smith 1994). These conditions could be met for a total of 25 ancestor–descendant population pairs, some of which are arrayed in series of up to 10 populations (Table 4; Fig. 2). The amount of time elapsed between ancestor and descendant pairs ranged from 14,000 to 7.38 million years (median = 290,000 years; Table 4).

Only sequences of populations at the same locality were considered candidate ancestor–descendant pairs. Deep-sea ostracode species often have large geographic ranges (Coles et al. 1990), but the genetic connectivity of widely spaced ostracode populations is not yet known. Cytheroidean ostracodes such as *Poseidonamicus* lack a dispersal stage and cannot swim, and so it is possible that some widespread species may contain deep phylogeographic structure. For this reason, ancestor–descendant relationships were

Table 4. Ancestor–descendant evolutionary transitions. For each ancestor–descendant pair, the reference phenotypic covariance matrix ($\mathbf{P}_{ref}$, Table 3) and elapsed time between ancestor and descendant (Δt, in million years) are given.

Ancestor	Descendant	$\mathbf{P}_{ref}$	Δt
ant-E1	ant-E2	ant-E1.2	2.95
pin-Q2	pin-Q3	pin-Q2.3.4	0.03
pin-Q3	pin-Q4	pin-Q2.3.4	0.01
maj-P2	maj-P3	maj-P2.3	0.21
maj-P3	maj-Q2	maj-P2.3	0.29
maj-Q2	maj-Q3	maj-Q2.3	0.22
maj-Q4	maj-Q5	maj-P2.3	0.75
rio-M2	rio-M3	rio-M3	2.44
rio-M3	rio-M4a	rio-M3	7.38
rio-M4a	rio-M4b	rio-M4a.b	0.10
rio-M4b	rio-M7	rio-M4a.b	0.82
rio-M7	rio-M5	rio-M5.7	0.56
mio-M1	mio-M3	mio-M2	1.07
mio-M3	mio-P5	mio-M2	2.11
mio-M2	mio-PA	mio-P1.2	0.21
mio-PA	mio-PB	mio-P1.2	0.31
mio-PB	mio-PC	mio-P1.2	0.34
mio-PC	mio-PD	mio-P1.2	0.22
mio-PD	mio-P1	mio-P1.2	0.08
mio-P1	mio-PE	mio-P1.2	0.20
mio-PE	mio-PF	mio-P1.2	0.20
mio-PF	mio-P2	mio-P1.2	0.11
mio-P2	mio-PG	mio-P1.2	0.11
spE-O1	spE-O2	spE-O1.2	1.93
spF-E1	spF-O1	spE-O1.2	2.24

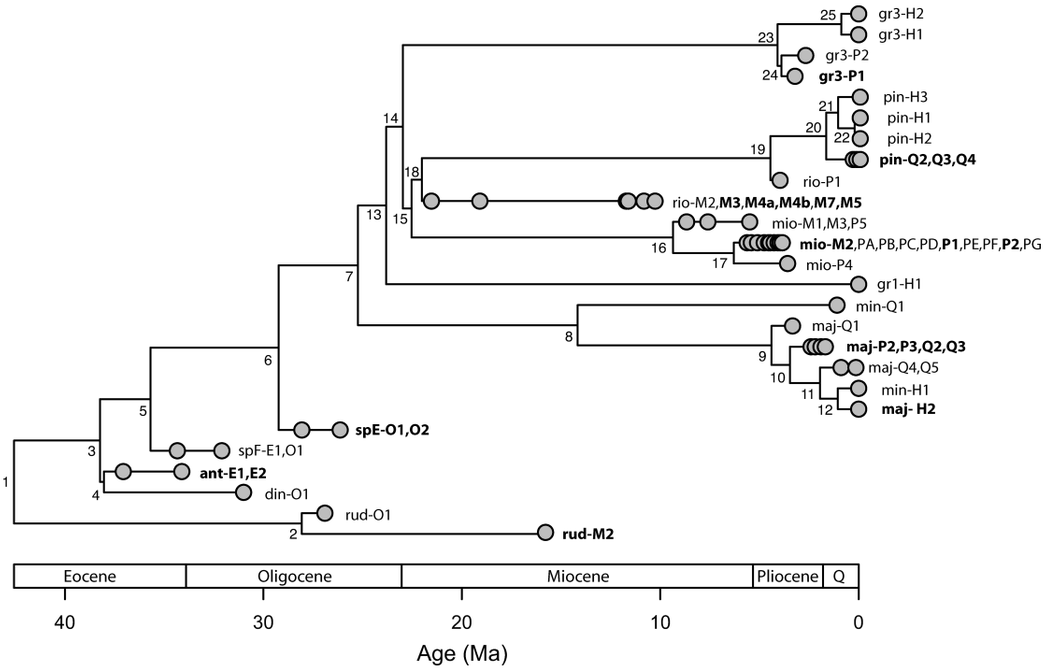

Figure 2. Phylogenetic hypothesis for the genus *Poseidonamicus*, with populations plotted according to their stratigraphic age. Population labels correspond to Table 1; population labels for ancestor–descendant sequences (Table 4) are listed consecutively with prefixes omitted after the first population. Internal nodes are numbered starting from the root and correspond to the phylogenetic contrasts in Table 5. Bold typeface indicates populations for which phenotypic covariance matrices were estimated (Table 3).

only inferred where the inference was strongest: for repeated samples at the same location on the ocean floor.

The vector of evolutionary change for ancestor–descendant pairs ($\mathbf{z}_{AD}$) is computed as $\mathbf{x}_D - \mathbf{x}_A$, where $\mathbf{x}_D$ and $\mathbf{x}_A$ are vectors representing the multivariate phenotypic means of the descendant and ancestral populations, respectively.

PHYLOGENETIC TRANSITIONS

Inferring evolutionary transitions among phylogenetically related populations requires a hypothesis of relationships and adoption of an algorithm to infer changes on a tree with unsampled ancestors. Phylogenetic relationships among populations of *Poseidonamicus* were taken from a recent parsimony analysis of 42 morphological characters (Hunt 2007). This study drew on all aspects of the skeletal morphology of this genus, including the presence/absence and location of homologous pores, and details of carapace shape and ornament. None of the phylogenetic characters corresponds directly to the 10 digitized landmarks, although a few (∼5) characterize morphological variation in the vicinity of the landmarks and could be indirectly related to the morphometric variables. Nevertheless, because only a few characters are potentially involved, the influence of the morphometric variables on the phylogenetic topology is likely to be small.

The major features of this phylogenetic hypothesis are well supported by character data, with most of the deeper divisions in the tree supported by one or more unambiguous synapomorphies (Hunt 2007). The inferred phylogeny receives some independent corroboration in that the branching sequence of resolved nodes generally agrees with the order of appearance of species and clades in the fossil record; more basal taxa appear before later differentiating clades (Hunt 2007). Although many nodes are well supported, some relationships, especially near the tips of the tree, are unresolved or have only limited support. Because this phylogenetic study was performed with populations (not species) as the operational units, it included multiple terminal taxa from the same Linnaean species. As a result, some of the lack of resolution near the tips of the tree likely reflects intraspecific geographical and temporal variation, which is not necessarily resolvable by phylogenetic methods.

For the purpose of analysis, phylogenetic uncertainties were resolved using stratigraphic information. Polytomies were transformed into pectinate clades with geologically older populations basal to geologically younger populations. Although there has been contentious debate as to the appropriateness of temporal data for phylogenetic inference (see, e.g., Smith 2000; Alroy 2002; Fisher et al. 2002; Smith 2002; Wagner 2002), the role of stratigraphic information here is limited only to relationships not resolved by character data, which is a fairly uncontroversial practice. For populations that are truly part of the same evolving lineage, which is likely true for many of the unresolved nodes

near the tips of the tree, resolving by geological age is particularly reasonable. Regardless, it is likely that the general findings of this study are robust to minor inaccuracies in the phylogenetic hypothesis. Phylogenetic analysis is often an iterative process, and I have performed similar analyses on two different previous phylogenetic hypotheses generated by somewhat different sets of cladistic characters. These topologies were similar in general structure to that used in the present study, but differed in some details, especially toward the tree tips. Despite these differences, analyses using these previous phylogenetic hypotheses produced results not substantially different from those presented here.

In addition to topology, most methods for inferring evolutionary changes on a tree require information on divergence times in the form of branch length estimates. Because most of the populations in the present study are extinct, the usual branch lengths of genetic distance cannot be obtained. Instead, the branch lengths were calibrated on the basis of the geologic age of the terminal taxa. To fully compute branch lengths, it is necessary to assign ages to internal nodes as well as terminal taxa, but there is some ambiguity as to how internal nodes are best dated. The minimum age of a node is simply the age of its oldest included terminal taxon, but such minimum ages are underestimates because it is extraordinarily unlikely that the oldest sampled terminal taxon in a clade is actually that clade's founding population. It is difficult, however, to place an upper bound on node ages because all nodes may be pushed backwards in time indefinitely. One could use sampling considerations to extend observed ages backwards in time similar to the computation of confidence intervals on stratigraphic ranges (e.g., Strauss and Sadler 1989; Marshall 1990; Solow 2003), but the terminal taxa are populations, not species, and therefore do not have any stratigraphic range. In addition, the taxon sampling in the present study is highly nonrandom in time and space, which complicates approaches that rely on simplified models of fossil preservation and recovery. Because of these difficulties, a crude but simple alternative was used: each branch of the tree was constrained to be at least 0.5 million years long. Essentially, this procedure assumes that each clade in the tree exists for at least 0.5 million years before it is sampled. The value of 0.5 million years is of course arbitrary, but the results are not sensitive to the magnitude of branch length extension; all reasonable values tested, from as little as 0.1 million years to as much as 10 million years, yielded qualitatively the same results. If branch lengths are not extended, zero length branches result between terminal taxa and many internal nodes. When this occurs, reconstructed ancestral states of internal nodes become identical to terminal taxa to which they are connected by zero-length branches, effectively forcing geologically old populations to be ancestors, even when character data are inconsistent with this interpretation.

Evolutionary transitions were calculated on the basis of a phylogenetic hypothesis of 26 terminal populations (Fig. 2). Evolutionary divergence among these populations was decomposed into 25 phylogenetically independent contrasts (Felsenstein 1985) (Table 5). Each contrast represents an evolutionary transition ($\mathbf{z}_{\mathbf{phylo}}$) between two nodes (terminal or internal) on the phylogenetic tree (Fig. 2). Only the oldest population in each ancestor–descendant sequence was included among the contrasts (Table 5) so that each lineage segment in the phylogeny would be included exactly once as an ancestor–descendant transition or as part of a phylogenetic contrast, but not both.

Table 5. Phylogenetic contrasts. Terminal taxa are indicated by their label (Table 1), and internal nodes are numbered starting at the root of the tree (Fig. 2). For each contrast, the reference covariance matrix ($\mathbf{P}_{\mathbf{ref}}$, Table 3) and divergence time between tips (Δt, measured as the sum of the intervening branch lengths in million years) are given. Because the minimum length for each branch was constrained to be 0.5 million years (see text), the minimum branch length separating a phylogenetic comparison is 1.0 years.

Node	tip 1	tip 2	$\mathbf{P}_{\mathbf{ref}}$	Δt
1	2	3	rud-M2	18.83
2	rud-O1	rud-M2	rud-M2	11.71
3	4	5	ant-E1.2	3.04
4	ant-E1	din-O1	ant-E1.2	6.69
5	6	spF-E1	spE-O1.2	6.96
6	7	spE-O1	spE-O1.2	4.50
7	8	13	spE-O1.2	13.69
8	9	min-Q1	maj-P2.3	21.79
9	10	maj-Q1	maj-P2.3	1.43
10	11	maj-P2	maj-P2.3	2.01
11	12	maj-Q4	maj-P2.3	1.40
12	min-H1	maj-H2	maj-H2	1.00
13	14	gr1-H1	spE-O1.2	22.05
14	15	23	rio-M3	18.83
15	16	18	rio-M3	13.36
16	17	mio-M1	mio-M2	3.56
17	mio-M2	mio-P4	mio-M2	2.54
18	19	rio-M2	rio-M3	18.07
19	20	rio-P1	rio-M3	4.35
20	21	pin-Q2	pin-Q2.3.4	1.00
21	22	pin-H3	pin-Q2.3.4	1.00
22	pin-H1	pin-H2	pin-Q2.3.4	1.00
23	24	25	gr3-P1	3.71
24	gr3-P1	gr3-P2	gr3-P1	1.04
25	gr3-H1	gr3-H2	gr3-P1	1.00

CONCORDANCE BETWEEN EVOLUTION AND PHENOTYPIC VARIATION

Given phenotypic covariance patterns and a set of evolutionary changes, the goal is to test whether the latter bear any relationship to the former. Because phenotypic covariance patterns are conserved but not constant across *Poseidonamicus*, it is necessary to determine for each evolutionary transition which covariance

matrix to use as a reference. Ideally, one would compare evolutionary changes to the ancestral covariance matrix of each transition. This is not practical, however, because sample size limitations in many terminal taxa and because most ancestors are unsampled. As a consequence, each evolutionary transition ($\mathbf{z}_{AD}$ and $\mathbf{z}_{phylo}$) was compared with the reference covariance matrix (Table 3) phylogenetically closest to that transition. If more than one covariance matrix was equally related, analyses were repeated using the alternative reference covariance matrix, but results never differed appreciably between alternative reference covariance matrices.

TEST STATISTICS

Two different test statistics were used to measure the association between evolutionary changes and covariance patterns. The first follows Schluter (1996), who measured the angle (θ) between the axis of maximal additive genetic variance ($\mathbf{g}_{max}$) and the observed evolution vector $\mathbf{z}$ (Fig. 3). This metric will often correlate with the amount of variation in the direction of evolutionary change; generally speaking, the smaller the angle, the higher the variance in the direction of the transition. In addition, $\mathbf{g}_{max}$ is particularly important in deflecting a population's response to selection as it climbs an adaptive peak (Schluter 1996, 2000), especially when the axis of maximum variation accounts for a large proportion of total variance. Because the current study considered phenotypic, rather than additive genetic variation, the axis of maximum variation is referred to as $\mathbf{p}_{max}$, and was computed as the leading eigenvector of the reference phenotypic covariance matrix. The angle (θ) between $\mathbf{p}_{max}$ and a vector of evolutionary divergence $\mathbf{z}$ was calculated as the arccosine of the inner product of $\mathbf{p}_{max}$ and $\mathbf{z}$ after both were scaled to unit length.

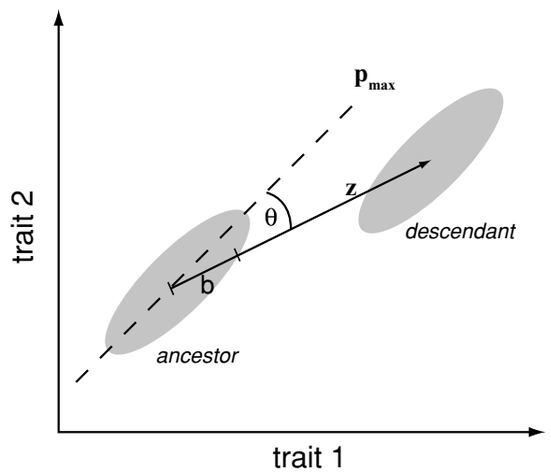

Figure 3. Schematic showing the measurement of an evolutionary change and its correspondence to within-population variation. Ancestor and descendant populations are represented as grey ellipses. Mean morphology is indicated by the location of the ellipses, and variation patterns are represented by the shape and orientation of the ellipse (for bivariate normal distributions, ellipses represent lines of equal probability density). Ancestor and descendant populations differ by a vector of evolutionary change (z), which forms an angle (θ) with the axis of maximum phenotypic variance ($\mathbf{p}_{max}$). The amount of variance in the direction of evolution (v_{proj}) is proportional to the square of the length of z that is inside the covariance ellipse (b).

Although adopted in several subsequent studies, comparing evolutionary divergence to the line of maximum variation has been criticized for placing undue emphasis on just one axis (Blows and Higgie 2003). The channeling effects of the $\mathbf{G}$ matrix are not limited to $\mathbf{g}_{max}$, and measuring θ ignores what maybe important effects operating in other dimensions. This is especially true when, as in the present study, the axis of maximal variation is not particularly dominant; $\mathbf{p}_{max}$ accounts for 27% of within-population variation on average (range: 22%–34%). As a result, a second test statistic was designed to more directly measure the correlation between evolutionary changes and phenotypic variance. This statistic measures the amount of variance present in an observed direction of evolution by projecting the data points onto the axis defined by the axis of net evolutionary divergence ($\mathbf{z}$), and then calculating the variance of these projections (Fig. 3). This quantity may be called the "projected variance" and is calculated as $v_{proj} = \mathbf{z}^T\mathbf{P}\mathbf{z}$, where $\mathbf{P}$ is the reference covariance matrix, $\mathbf{z}$ is the evolution vector ($\mathbf{z}_{AD}$ or $\mathbf{z}_{phylo}$) scaled to unit length, and a superscript "T" indicates transpose. For convenience, the value of v_{proj} was normalized to be a proportion of the total variance of all traits in the population. This normalization facilitates interpretation of v_{proj}, but has no substantial effects on the analysis because total variance does not differ much among populations.

NULL DISTRIBUTION OF TEST STATISTICS

The null distribution of each statistic was generated by simulating the behavior of these metrics when evolutionary changes occur in random directions. In this scenario, the null distribution of the angle between $\mathbf{p}_{max}$ and $\mathbf{z}$ is simply the expected distribution of the angle between two randomly chosen vectors (or between one random vector and a fixed reference direction). This null distribution was generated by computing the angle between 10,000 pairs of random vectors using the algorithm of Knuth (1969) for drawing vectors equally from all possible directions. Note that the broken stick distribution, which has been used previously for this purpose (e.g., Schluter 1996; Marriog and Cheverud 2005), does not generate a distribution of truly random directions. Instead, it disproportionately favors angles close to the coordinate axes, although the effect is not very large (see Appendix).

The null distribution of v_{proj} is somewhat more complicated to generate because it depends on the shape of the reference covariance matrix. If the covariance matrix is spherical with equal variance in all directions, the null distribution of v_{proj} will be uniform.

In contrast, if most of the population variation is dominated by a few axes, the distribution of v_{proj} will be very skewed, with most morphological directions having relatively little variance and a few with much higher variance. As a result of this dependence, it is necessary to generate the null distribution of v_{proj} with respect to a specific set of reference covariance matrices.

The null distribution of v_{proj} was generated by slightly altering the procedure for estimating v_{proj} in the observed set of evolutionary transitions. Instead of calculating the variance in the reference covariance matrix in the observed direction of evolution ($\mathbf{z}$), the variance was calculated in randomly chosen directions. For a set of K evolutionary transitions, this produces K variances in randomly chosen directions. This procedure was repeated 10,000 times to create separate null distributions for the two types of evolutionary transitions ($\mathbf{z_{AD}}$ and $\mathbf{z_{phylo}}$). Observed distributions of v_{proj} and θ were compared with their respective null distributions using Wilcoxon signed rank tests for differences in median and Kolmogorov-Smirnov tests for differences in distribution shape.

PATTERNS IN THE ORIGINAL LANDMARK SPACE

In addition to this hypothesis testing procedure performed with the PCA-transformed Procrustes coordinates, it was heuristically useful to examine the patterns of variation and evolution with respect to the original landmarks. Specifically, the variability of each landmark within populations was compared to the typical magnitude of evolutionary change occurring at that landmark using a Spearman rank correlation coefficient. Landmark variability was computed as the variance of the fitted Procrustes coordinates, summed over the horizontal and vertical directions. This landmark variability was then averaged over all populations for which variation patterns were estimated (Table 3). Total evolutionary change at each landmark was computed similarly, as the average of the sum of the squared differences in the horizontal and vertical directions between ancestor and descendant (or phylogenetic contrasts) at each landmark.

Some caution is required in interpreting these results because statistical inference of variation patterns at specific landmarks has only just begun to be explored (Walker 2000), and the process of fitting can induce some variation and covariation structure into the fitted coordinates (Adams et al. 2004). To explore this latter possibility, a large dataset ($n = 10{,}000$) of simulated landmark configurations was generated assuming circular Gaussian variation around the mean landmark configuration. The magnitude of simulated variation was the same for each landmark, and set as the average landmark variance in the empirical data. These simulated individuals were subject to generalized Procrustes superimposition, and the covariance matrix among the resulting fitted coordinates was computed. Differences between this estimated matrix and the spherical covariance matrix that was simulated can be attributed to the fitting procedure, and thus give some idea of the nature and magnitude of the bias induced by Procrustes fitting.

All analyses were performed using code written in the R statistical programming language (R Development Core Team 2005). Independent contrasts were calculated using the Analysis of Phylogenetics and Evolution package for R (Paradis et al. 2004).

Results

Evolutionary divergence in *Poseidonamicus* has occurred preferentially in morphological directions with relatively abundant phenotypic variation. This trend holds for both kinds of evolutionary transitions (ancestor–descendant and phylogenetic contrasts), and for both means of measuring the concordance between evolution and phenotypic variation patterns (θ and v_{proj}).

Transitions are disproportionately concentrated in trajectories that are close to the axis of maximal phenotypic variation, $\mathbf{p_{max}}$, relative to the expectation of randomly directed divergences (Table 6; Fig. 4). Similarly, evolutionary differences have occurred preferentially in multivariate directions of high phenotypic variance (Table 6; Fig. 5). Moreover, the two kinds of evolutionary transitions are similar in the degree to which directions of high variance are favored; the distributions of θ and v_{proj} do not differ significantly between ancestor–descendant and phylogenetic changes (θ: $D = 0.32$, $P = 0.16$; v_{proj}: $D = 0.32$, $P = 0.16$). As expected, the two metrics are not independent; evolutionary directions with high variance also tend to have small angles with $\mathbf{p_{max}}$, although the correlation is stronger for ancestor–descendant ($r_s = -0.91$, $P = 2 \times 10^{-6}$) than phylogenetic transitions ($r_s = -0.51$, $P = 0.01$).

Table 6. Differences between the observed and null distributions for testing for concordance between evolutionary divergence and variation structure within populations. Metrics include the angle between the vector of evolutionary change and $\mathbf{p_{max}}$ (θ), and the variance in the observed direction of evolution (v_{proj}). Statistical tests compare the observed distribution of the test statistic to the null generated by simulation (see text). Wilcoxon tests evaluate the null hypothesis of equal medians, and Kolmogorov–Smirnov (KS) tests evaluate the null hypothesis that the observed and null metrics were drawn from the same continuous distribution.

Transition type	Metric	Test	Result
Ancestor–	θ	Wilcoxon	V=63, P=0.006
descendant	θ	KS	D=0.43, P=0.0002
	v_{proj}	Wilcoxon	V=281, P=0.0008
	v_{proj}	KS	D=0.46, P=5.9×10^{-5}
Phylogenetic	θ	Wilcoxon	V=87, P=0.042
contrasts	θ	KS	D=0.29, P=0.028
	v_{proj}	Wilcoxon	V=307, P=1.5×10^{-5}
	v_{proj}	KS	D=0.49, P=1.5×10^{-5}

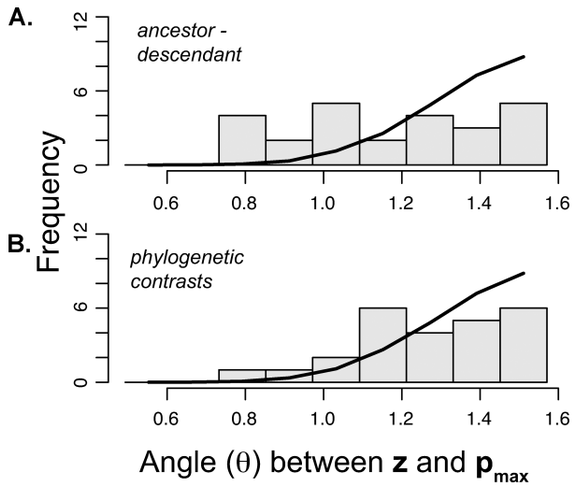

Figure 4. Angle between the direction of evolutionary change and the axis of maximum phenotypic variation. Histograms show observed values for ancestor–descendant transitions (A) and phylogenetic contrasts (B). Solid line shows the null distribution expected if evolutionary changes are random with respect to $\mathbf{p}_{max}$; in this high-dimensional space, most randomly chosen vectors are distant (nearly orthogonal) to $\mathbf{p}_{max}$. Both empirical distributions have a significant excess of evolutionary changes in directions close to $\mathbf{p}_{max}$.

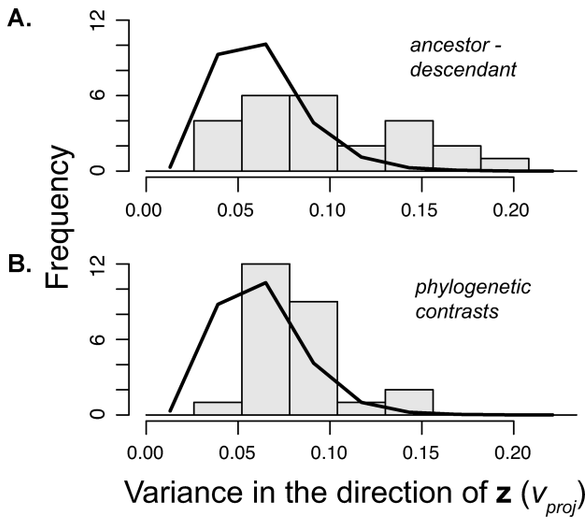

Figure 5. Variance in the direction of evolutionary changes. Histograms show observed values for ancestor–descendant transitions (A) and phylogenetic contrasts (B). Solid line shows the null distribution expected if evolutionary changes are random with respect to within-population variation. Both empirical distributions have an excess of changes in directions of high variance and a deficiency of changes in directions of low variance.

The correspondence between trait variation and evolutionary divergence diminishes over time. Nearly all evolutionary transitions that are very close to the $\mathbf{p}_{max}$ axis span relatively brief intervals of time (< 2 million years), and evolutionary transitions that span longer intervals are approximately centered around the random expectation for θ (Fig. 6). A similar decay is seen in the metric v_{proj}, in which most evolutionary changes in directions of highest variability span a few million years or less (Fig. 6). However, unlike θ, the influence of v_{proj} never completely disappears—even evolutionary divergences over the longest intervals do not scatter around the random expectation. Of the 22 evolutionary transitions that span more than two million years, all but two occur in directions with more phenotypic variance than the median of the null distribution (Fig. 6). Thus, although evolutionary changes in *Poseidonamicus* have frequently occurred in directions distant from $\mathbf{p}_{max}$, they have seldom occurred in morphological directions of very low variation.

In addition to their disproportionate frequency, evolutionary changes are generally faster in directions that are close to $\mathbf{p}_{max}$ and have ample phenotypic variance (Table 7). Here the pace of change is measured as the sum of the squared differences for each variable, divided by time (Marriog and Cheverud 2005). This trend is in the same direction for both kinds of evolutionary transitions, but is only significant for phylogenetic contrasts (Table 7).

The positive relationship between phenotypic variation and propensity for evolutionary divergence also exists in the original space of the fitted landmark coordinates. Landmarks differed markedly in their variability, with a greater than twofold range from the least to the most variable landmarks (Fig. 7). It is important to note that those landmarks that varied more within populations also experienced larger evolutionary changes between populations (Fig. 7). Again, this pattern held for both ancestor–descendant transitions ($r_s = 0.75$, $P = 0.009$) and phylogenetic contrasts ($r_s = 0.59$, $P = 0.04$).

These correlations are probably best viewed as heuristic because, unlike the PCA axes, landmark variances do not take into account correlations within and among landmark residuals. Moreover, it is necessary to make sure that the observed covariance structure is not an artifact of the fitting process (Adams et al. 2004). Simulation of isotropic error around each landmark does indeed show some spurious covariance structure: even when all landmarks are equally variable, fitted coordinates from landmarks located farther away from the center of form are estimated to be less variable than more central landmarks (Fig. 8). This pattern is likely produced because the same rotational deviation causes greater misfit the more distant the landmark, and so more peripheral landmarks have greater influence in the fitting process. Nevertheless, the effect is not large—the most distant landmark is on average only 35% more variable than the most central. Even more important, the bias is the opposite of the empirical pattern, which finds that peripheral landmarks (such as #2 and #10) are the most variable and many central landmarks (such as #4, #5, and #8)

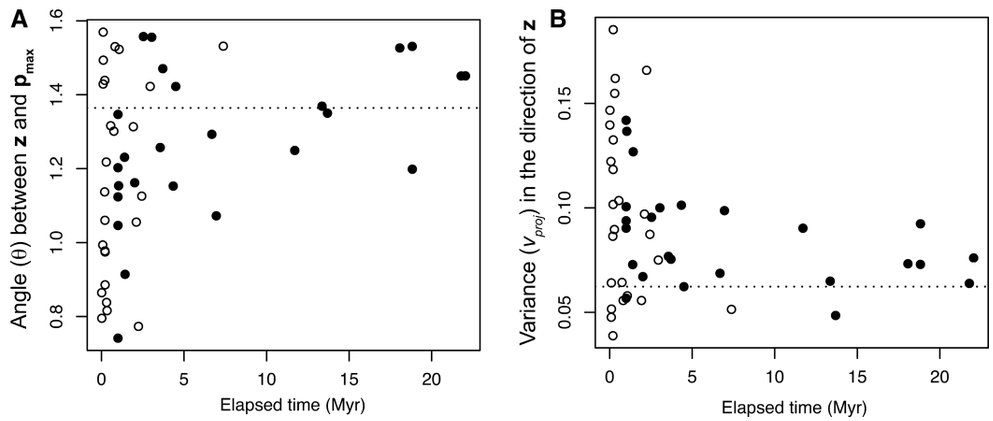

Figure 6. Relationship between elapsed time and the propensity for evolutionary changes to be in directions close to p_{max} (A) or in directions of high phenotypic variance (B). Open circles are ancestor–descendant transitions, filled circles are phylogenetic contrasts. Dotted lines indicate the median of the null distribution for each metric.

are the least variable. Thus, although the Procrustes fitting does induce some covariance structure, the effect is small, and in this dataset, overwhelmed by opposing biological variability.

Discussion

Evolution in the measured features of *Poseidonamicus* has preferentially occurred in directions of high phenotypic variance. Thus, at least for this clade and these traits, patterns of trait variation and covariation leave a detectable imprint on evolutionary divergence over geological time scales. That this trend is potentially applicable to all sets of traits that vary on the same scale makes it an unusually general macroevolutionary hypothesis.

EVOLUTIONARY MECHANISMS

Several aspects of phenotypic divergence in *Poseidonamicus* are consistent with Schluter's (1996; 2000) genetic constraints model. First, evolutionary changes have disproportionately occurred in phenotypic directions that are both highly variable and close to the axis of maximal variation. Second, this channeling effect decays over time, reducing to a lesser (v_{proj}) or negligible (θ) influence after a few million years. For a population climbing an adaptive peak, genetic constraints deflect strongly at first, but decreasingly so until the population eventually proceeds directly up the adaptive surface (Zeng 1988; Schluter 1996). Although a few million years may seem rather long to ascend an adaptive peak, a similar effect may be achieved when there are multiple peaks in the neighborhood of a population (Schluter 1996) or when

Table 7. The correspondence between evolutionary rate and the direction of change. Spearman rank correlations (r_s) are given between the rate of change and the degree to which changes are in directions of high phenotypic variance (measured by θ and v_{proj}). Rate is measured as the sum of squared differences over all variables between ancestor and descendant or phylogenetic contrasts.

	θ	v_{proj}
Ancestor–descendant	r_s=−0.27, P=0.20	r_s=0.29, P=0.16
Phylogenetic contrasts	r_s=−0.45, P=0.03	r_s=0.48, P=0.02

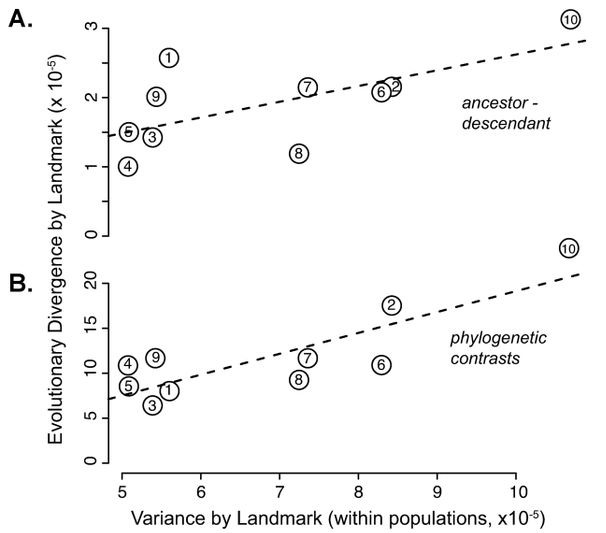

Figure 7. Trait variation and evolutionary divergence in the fitted landmark coordinates. The horizontal axis indicates for each landmark its mean variance within populations; numbers correspond to landmark labels (Fig. 1). The vertical axis shows for each landmark its mean evolutionary divergence between populations for ancestor–descendant transitions (A) and phylogenetic contrasts (B). Dotted lines indicate least-squares regression; spearman rank correlations are significant for both kinds of evolutionary transitions. Note different vertical scale for phylogenetic contrasts, which are on average larger than the ancestor–descendant transitions.

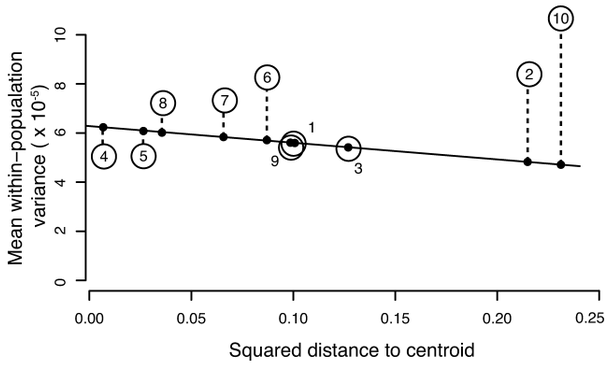

Figure 8. Landmark variability as a function of the distance from the center of form. Large open circles indicate observed mean variance of each landmark; numbers correspond to landmark labels. Smaller filled circles show the expected variability induced by the Procrustes fitting (see text).

a population tracks an adaptive peak that fluctuates in position over time. In either case, genetic constraints channel the adaptive climbs, but their long-term influence declines as more and more peaks are climbed, assuming that the position and/or movement of peaks are unrelated to variation patterns within the population. This last assumption—that the direction of natural selection is unrelated variation patterns—is an important one, and will be considered further below. Regardless, it is interesting to note that the temporal duration of genetic constraints in the present study is similar to that reported by Schluter (1996). If these two studies are a fair indication, the channeling effects of genetic constraints typically last a few million years. The third and final aspect that is consistent with Schluter's model is that evolutionary changes were generally faster in directions of high phenotypic variance, although this relationship was not always significant. This difference in pace is consistent with inhibition of evolutionary changes in directions with little variation and facilitation in directions with ample variation.

Genetic constraints are not the only mechanism that can cause variation to structure long-term evolution. A complete absence of genetic variation in some morphological directions precludes evolutionary changes in those directions, causing a commensurate excess of change along more variable axes. Because nearly all morphological traits artificially selected have evolved in response (Charlesworth et al. 1982), it seems unlikely that such traits commonly lack genetic variance. In the only test thus far conducted with geometric morphometric variables, Mezey and Houle (2005) found evidence for additive genetic variance in all multivariate directions in a *Drosophila* wing shape dataset. Even if some traits or combinations of traits lack genetic variance in the present study, it is unlikely that this effect accounts for the observed correlation between variation and divergence. This effect would likely be limited to at most a few of the least variable axes, and its influence is likely to be small. In addition, the metric θ, because it is based only on the axis of maximum variation, should be particularly insensitive to a total lack of variance in the least variable dimensions. It is possible, however, that the paucity of observed evolutionary changes in directions of very low variance, even over very long (> 2 Myr) durations of time is related to an absence (complete or nearly so) of genetic variation in these directions.

Neutral genetic drift is another mechanism that causes correspondence between within-population variation and evolutionary divergence. Drift has been implicated as a potential influence in several empirical studies of divergence (Roff et al. 1999; Marriog and Cheverud 2001; Ackermann and Cheverud 2002, 2004) and is expected to produce evolutionary divergence that is exactly proportional to the within-population additive genetic covariance matrix (Lande 1979; Arnold et al. 2001). Under drift, however, the correlation between variation and divergence should not diminish over time as it does in the present study. It is possible that drift may play a role in more complicated scenarios, for example, if neutral evolution is interrupted periodically by periods of directional selection. As intervals of selection accumulate, they will cause increasing divergence from the drift-imposed correlation between variation and evolutionary change (again assuming that the direction of selection is random with respect to phenotypic variation).

Of course, genetic drift is only plausible as a mechanism to the extent that the measured traits are likely to be selectively neutral. Although there are good reasons to believe that the presence of skeletal ridges is biomechanically and functionally important (Benson 1975), the selective relevance of more subtle differences in ridge structure and location is less clear. Although I am aware of no relevant direct measurements of fitness in ostracodes, indirect evidence suggests that the traits measured in the present study are not selectively neutral. Quantitative morphological traits evolving under mutation-drift balance are expected to evolve as a random walk with a rate parameter that is proportional to the amount of genetic variance added by mutation each generation (Lynch 1990). Mutation studies in modern organisms find that under neutral drift, the expected rate of change (Δ) ranges from roughly 5×10^{-5} to 5×10^{-3} (Lynch 1990). This rate parameter can be best estimated from the most intensively sampled lineages, *P. miocenicus* and *P. riograndensis*, with 10 and six sampled populations, respectively (Table 4; Fig. 2). The 16 principal component axes from these two lineages yield 32 separate rates, each estimated using to the maximum-likelihood method developed by Hunt (2006) for analyzing fossil sequences. (In the terminology of Hunt (2006), Δ is equal to the step variance, measured per generation, divided by the phenotypic variance.) Cold-water ostracodes generally live a few years (Cohen and Morin 1990), and so I assume a generation time of five years (generation times of one year to 10 years give the same results). Of the 32 rate estimates

computed in this manner, every single one was lower than the neutral evolution prediction (all $\Delta < 5 \times 10^{-5}$). Such systematically low rates are inconsistent with neutral genetic drift and instead implicate conservative evolutionary mechanisms such as stabilizing natural selection.

One potential alternative mechanism that may be difficult to distinguish from genetic constraints is related to the recurring assumption that natural selection is random with respect to within-population variation. This assumption may be violated if, for example, peak shifts in the adaptive landscape match patterns of genetic variation because the latter are shaped by natural selection to mirror the selective regime experienced by populations (Lande 1980; Cheverud 1984; Arnold et al. 2001). Arnold and colleagues (2001) refer to this effect as evolution along lines of least selective resistance, and these authors note that selective least resistance may be difficult to distinguish from genetic least resistance.

This selective least resistance model may be particularly relevant for suites of traits correlated with overall body size. For these traits, body size usually accounts for a large proportion of within-population variation, and the axis of maximal variation is usually similar to the vector of isometric or allometric growth. Size may dominate variation patterns because of the molding effects of natural selection (mentioned earlier), or perhaps because size is a larger target for mutation than shape (Houle 1998). Moreover, because body size is so ecologically and functionally important, it is very commonly a target of selection (LaBarbera 1986; Kingsolver and Pfennig 2004). Thus, evolutionary changes may be disproportionately close to $\mathbf{g}_{\mathbf{max}}$ or $\mathbf{p}_{\mathbf{max}}$ because of frequent natural selection on body size, even if genetic constraints have no effect whatsoever. Indeed, most published datasets that show a macroevolutionary effect of $\mathbf{g}_{\mathbf{max}}$ analyze traits that are correlated with body size (Schluter 1996; Bégin and Roff 2003, 2004; Marriog and Cheverud 2005). Recognizing this possibility, Marriog and Cheverud (2005) note that although the structure of evolutionary divergence in new world monkeys was consistent with the effects of genetic constraints, it could also be explained nearly as well by chronic natural selection on body size.

In this context, it may be especially interesting that the correspondence between divergence and variation holds in the present analysis of fitted landmark coordinates, which are generally uncorrelated with body size ($r^2 < 0.05$). Without a strong influence of body size, the model of selective least resistance is still plausible but probably less likely, strengthening the case for genetic constraints as a significant force structuring long-term phenotypic divergence. A similar result was also reported by Renaud and colleagues (2006), who found evidence for genetic constraints in the evolution of characters uncorrelated with body size (Fourier descriptors of molar shape in rodents).

THE ROLE OF FOSSIL DATA

The current study, along with recent work by Renaud et al. (2006), differs from most previous in this subject by relying heavily on the fossil record to infer evolutionary changes. This strategy has both advantages and limitations. On the positive side, fossil data allow for much more detailed reconstruction of evolutionary history than would be possible from only extant populations. In addition to permitting ancestor–descendant inferences, some of the sampled fossil populations are geologically old and closely related to deep nodes that would otherwise be estimated very imprecisely by comparative methods (Cunningham et al. 1998). Of course, these advantages are most realized in taxa such as *Poseidonamicus* that have a fossil record that is relatively complete and geographically extensive.

Nevertheless, there are some concerns particular to the use of fossil populations to study variation. First, it is possible that fossilization may distort variation patterns of biological populations. Although compaction and shearing of sedimentary rocks can influence variation patterns in fossils (Webster and Hughes 1999), all samples in the present study are from unlithified deep-sea oozes that preserve fossils without distortion. Potentially more of a concern is time-averaging, which is the inclusion in a fossil sample of individuals of differing ages. Nearly all paleontological samples are time-averaged to some degree (Kidwell and Behrensmeyer 1993); sedimentation rates and stratigraphic spans of samples in the present study indicate that individual samples generally represent several thousand years of accumulated time. Evolutionary changes occurring over these several thousand years have the potential to alter patterns of trait variance and covariance. Fortunately, a number of empirical studies have found that trait variation is seldom much affected by this scale of time averaging (Bell et al. 1987; Bush et al. 2002; Hunt 2004a,b). This finding confirms other paleontological evidence indicating that rates of evolution are generally quite slow on geological time scales (Gould and Eldredge 1977; Gingerich 1983), and, importantly for the present study, suggests that variation structure in fossil samples approximates standing phenotypic patterns.

A second consequence of relying on fossil samples is that genetic parameters cannot be measured, and thus patterns of phenotypic variation must be used instead. This limitation holds for any populations that cannot not be analyzed genetically, not just those that are extinct (e.g., living populations of *Poseidonamicus*, which cannot feasibly be bred in the laboratory). Although a strong case can be made for the evolutionary importance of phenotypic variation in its own right (Steppan 1997; Arnold and Phillips 1999), much of the theory linking variation to mechanistic models of evolution involves the additive genetic variance–covariance matrix. Although there are factors that can cause differences (Willis et al. 1991), empirically there is usually moderate

to high correspondence between genetic and phenotypic variation patterns (Cheverud 1988; Roff 1995, 1997; Steppan et al. 2002). Thus, it is reasonable to suggest that directions of high phenotypic variation are often also directions of high genetic variation, especially for morphological traits (Roff 1997). In addition, although differences between genetic and phenotypic covariance matrices could cause one to miss a correspondence between variation and evolutionary divergence, it is more difficult to account for positive evidence of genetic constraints unless phenotypic patterns truly reflect underlying genetic parameters.

CONTINUITY BETWEEN MICRO- AND MACROEVOLUTION

Since proponents of punctuated equilibria argued that modern microevolutionary theory offers an incomplete accounting of macroevolution (Eldredge and Gould 1972; Gould and Eldredge 1977; Stanley 1979), there has been continued disagreement as to the proper relationship between evolution observed at these two scales (Charlesworth et al. 1982; Erwin 2000; Leroi 2000; Arnold et al. 2001; Gould 2002). The fact that microevolutionary models make predictions about macroevolutionary divergence (Lande 1979; Hansen and Martins 1996; Schluter 1996), and that the present and previous studies find support for some of these predictions argue for a close connection between long-term phenotypic evolution and standard microevolutionary processes.

A related aspect of the micro- versus macroevolution debate is the distinction between changes occurring within lineages and changes accounting for differences between lineages. Under the standard punctuational view, differences between lineages dominantly arise from geologically rapid bursts of change associated with cladogenesis (Gould and Eldredge 1977). Furthermore, processes governing within-lineage evolution (anagenesis) may be different from those important in cladogenetic-associated divergence, and if so, microevolutionary processes shaping evolution within lineages are not relevant for understanding long-term, large-scale macroevolutionary divergence (Stanley 1979; Gould 2002).

In a limited way, divergence in *Poseidonamicus* can be brought to bear on the distinction between anagenetic and cladogenetic evolution. Both kinds of evolutionary modes are represented among the evolutionary transitions studied here: ancestor–descendant transitions represent within-lineage evolution, and phylogenetic contrasts capture differences between lineages. In reality, this distinction is not quite so simple; ancestor–descendant relationships may be incorrectly inferred, and some of the phylogenetic contrasts in the present study likely represent geographic or temporal variation within lineages. Nevertheless, it is probable that ancestor–descendant differences are at least mostly anagenetic, and that phylogenetic contrasts mostly reflect differences between lineages. Thus, if different processes are responsible for anagenetic and cladogenetic evolution, then these two kinds of transitions should have different characteristics. This prediction is not supported in the present study, at least with respect to the propensity of evolutionary changes to occur in directions of high phenotypic variance. Both measures of the correspondence between variation and divergence (θ and v_{proj}) do not differ significantly between the two kinds of evolutionary transitions (ancestor–descendant vs. phylogenetic contrasts). Such similarity argues for continuity of process between anagenesis and cladogenesis, and across phenotypic evolution regardless of the scale of observation.

Summary

1. Evolutionary divergence in the deep-sea ostracode genus *Poseidonamicus* has occurred preferentially in morphological directions of high phenotypic variance.
2. The correspondence between evolution and variation decays over time such that its effects are much diminished, but not totally absent after a few million years.
3. Results are consistent with Schluter's genetic constraints model in which evolutionary trajectories on the adaptive landscape are temporarily deflected by trait variances and covariances. An alternative model in which natural selection disproportionately favors changes in directions of high variance is also plausible, but may be less applicable for suites of traits that are uncorrelated with overall body size such as the geometric morphometric variables in the present study.
4. Ancestor–descendant transitions and phylogenetic contrasts are similar in preferentially occurring in directions of high phenotypic variance. This shared pattern does not support views in which evolutionary processes operating within populations are distinct from those responsible for differences between lineages.

ACKNOWLEDGMENTS

M. Foote, B. Chernoff, D. Jablonski, M. LaBarber, and T. Cronin provided many useful comments and suggestions throughout the completion of this project. For constructive reviews of the manuscript, I thank D. Ayre and two anonymous reviewers. I thank G. Miller, T. Cronin, R. Norris, S. Schellenberg, and the late R. Benson for providing access to samples and specimens; G. Miller, S. Whittaker, and A. Davis assisted with scanning electron microscopy. This work was financially supported by NSF Doctoral Dissertation Improvement Grant DEB-0104960, a student grant from Sigma Xi, a predoctoral fellowship from the Smithsonian Institution, and the Harper Fellowship at the University of Chicago.

LITERATURE CITED

Ackermann, R. R., and J. M. Cheverud. 2002. Discerning evolutionary processes in patterns of tamarin (genus *Saguinus*) craniofacial variation. Am. J. Phys. Anthropol. 117:260–271.

———. 2004. Detecting genetic drift versus selection in human evolution. Proc. Natl. Acad. Sci. USA. 101:17946–17951.

Adams, D. C., F. J. Rohlf, and D. Slice, E. 2004. Geometric morphometrics: ten years of progress following the 'revolution'. Ital. J. Zool. 71:5–16.

Alroy, J. 2002. Stratigraphy in phylogeny reconstruction–reply to Smith (2000). J. Paleont. 76:587–589.

Arnold, S. J., M. E. Pfrender, and A. G. Jones. 2001. The adaptive landscape as a conceptual bridge between micro- and macroevolution. Genetica 112/113:9–32.

Arnold, S. J., and P. C. Phillips. 1999. Hierarchical comparison of genetic variance-covariance matrices. II. Coastal-inland divergence in the garter snake, *Thamnophis elegans*. Evolution 53:1516–1527.

Bader, R. S. 1955. Variability and evolutionary rate in the oreodonts. Evolution 9:119–140.

Badyaev, A. V., and K. R. Foresman. 2000. Extreme environmental change and evolution: stress-induced morphological variation is strongly concordant with patterns of evolutionary divergence in shrew mandibles. Proc. R. Soc. Lond. B. 267:371–377.

Badyaev, A. V., and G. E. Hill. 2000. The evolution of sexual dimorphism in the house finch. I. Population divergence in morphological covariance structure. Evolution 54:1784–1794.

Barton, N. H., and M. Turelli. 1989. Evolutionary quantitative genetics: how little do we know? Annu. Rev. Genet. 23:337–370.

Bégin, M., and D. A. Roff. 2003. The constancy of the G matrix through species divergence and the effects of quantitative genetic constraints on phenotypic evolution: a case study in crickets. Evolution 57:1107–1120.

Bégin, M., and D. A. Roff. 2004. From micro- to macroevolution through quantitative genetic variation: positive evidence from field crickets. Evolution 58:2287–2304.

Bell, M. A., M. S. Sadagursky, and J. V. Baumgartner. 1987. Utility of lacustrine deposits for the study of variation within fossil samples. Palaios 2:455–466.

Benson, R. H. 1972. The *Bradleya* problem, with descriptions of two new psychrospheric ostracode genera, *Agrenocythere* and *Poseidonamicus* (Ostracoda: Crustacea). Smithson. Contrib. Paleobiol. 12:1–138.

Benson, R. H. 1975. Morphologic stability in Ostracoda. Am. Paleontol. Bull. 65:13–45.

Berggren, W. A., F. J. Hilgen, C. G. Langereis, D. V. Kent, J. D. Obradovich, I. Raffi, M. E. Raymo, and N. J. Shackleton. 1995a. Late Neogene chronology: new perspectives in high-resolution stratigraphy. GSA Bull. 107:1272–1287.

Berggren, W. A., D. V. Kent, C. C. I. Swisher, and M.-P. Aubrey. 1995b. A revised Cenozoic geochronology and chronostratigraphy. Pp. 129–212 *in* W. A. Berggren, D. V. Kent, M.-P. Aubrey, and J. Hardenbol, eds. Geochronology, Time Scales and Global Stratigraphic Correlation. Society for Sedimentary Geology, Tulsa, OK.

Blows, M. W., and M. Higgie. 2003. Genetic constraints on the evolution of mate recognition under natural selection. Am. Nat. 161:240–253.

Bürger, R. 1986. Constraints for the evolution of functionally coupled characters: a nonlinear analysis of a phenotypic model. Evolution 40:182–193.

Bush, A., M. G. Powell, W. S. Arnold, T. M. Bert, and G. M. Daley. 2002. Time-averaging, evolution and morphological variation. Paleobiology 28:9–25.

Charlesworth, B., R. Lande, and M. Slatkin. 1982. A neo-Darwinian commentary on macroevolution. Evolution 36:474–498.

Cheverud, J. M. 1984. Quantitative genetics and developmental constraints on evolution by selection. J. Theor. Biol 110:155–171.

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

Cohen, A. C., and J. G. Morin. 1990. Patterns of reproduction in ostracodes: a review. J. Crust. Biol. 10:184–211.

Coles, G., M. Ayress, and R. C. Whatley. 1990. A comparison of North Atlantic and Pacfic Cainozoic deep-sea Ostracoda. Pp. 287–305 *in* R. Whatley and C. Maybury, eds. Ostracoda and global events. Chapman and Hall, London.

Conner, J., and S. Via. 1992. Natural selection on body size in *Tribolium*: possible genetic constraints on adaptive evolution. Heredity 69: 73–83.

Cronin, T. M., D. M. DeMartino, G. S. Dwyer, and J. Rodriguez-Lazaro. 1999. Deep-sea ostracode species diversity: response to late quaternary climate change. Mar. Micropaleontol. 37:231–249.

Cronin, T. M., M. E. Raymo, and K. P. Kyle. 1996. Pliocene (3.2–2.4 Ma) ostracode faunal cycles and deep ocean circulation, North Atlantic Ocean. Geology 24:695–698.

Cunningham, C. W., K. E. Omland, and T. H. Oakley. 1998. Reconstructing ancestral character states: a critical reappraisal. Trends Ecol. Evol. 13:361–366.

Dryden, I. L., and K. V. Mardia. 1998. Statistical shape analysis. John Wiley & Sons, Chichester, U. K.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism *in* T. J. M. Schopf, ed. Models in paleobiology. Freeman, Cooper & Company, San Francisco, CA.

Erwin, D. H. 2000. Macroevolution is more than repeated rounds of microevolution. Evol. Dev. 2:78–84.

Falconer, D. S., and T. F. C. Mackay. 1996. Introduction to quantitative genetics. Addison Wesley Longman Limited, Essex, U.K.

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

Fisher, D. C., M. Foote, D. L. Fox, and L. R. Leighton. 2002. Stratigraphy in phylogeny reconstruction—comment on Smith (2000). J. Paleont. 76:585–586.

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

Gingerich, P. D. 1983. Rates of evolution: effects of time and temporal scaling. Science 222:159–161.

Gould, S. J. 2002. The structure of evolutionary theory. Belknap Press, Cambridge, MA.

Gould, S. J., and N. Eldredge. 1977. Punctuated equilibria: the tempo and mode of evolution reconsidered. Paleobiology 3:115–151.

Grant, P. R., and B. R. Grant. 1995. Predicting microevolutionary responses to directional selection on heritable variation. Evolution 49:241–251.

Guthrie, R. D. 1965. Variability in characters undergoing rapid evolution, an analysis of *Microtus* molars. Evolution 19:214–233.

Hansen, T. F., and E. P. Martins. 1996. Translating between microevolutionary process and macroevolutionary patterns: the correlation structure of interspecific data. Evolution 50:1404–1417.

Hansen, T. F., C. Pélabon, W. S. Armbruster, and M. L. Carlson. 2003. Evolvability and genetic constraint in *Dalechampia* blossoms: components of variance and measures of evolvability. J. Evol. Biol. 16:754–766.

Houle, D. 1998. How should we explain variation in the genetic variance of traits? Genetica 102/103:241–253.

Hunt, G. 2004a. Phenotypic variance inflation in fossil samples: an empirical assessment. Paleobiology 30:487–506.

———. 2004b. Phenotypic variation in fossil samples: modeling the consequences of time-averaging. Paleobiology 30:426–443.

———. 2006. Fitting and comparing models of phyletic evolution: random walks and beyond. Paleobiology 32:578–601.

———. 2007. Morphology, ontogeny, and phylogenetics of the genus *Poseidonamicus* (Ostracoda: Thaerocytherinae). J. Paleont. 81. *In press*.

Irizuki, T. 1993. Morphology and taxonomy of some Japanese hemicytherine Ostracoda with particular reference to ontogenetic changes of marginal pores. Trans. Proc. Palaeont. Soc. Japan, N.S. 170:186–211.

Kelley, P. H. 1983. The role of within-species differentiation in macroevolution of Chesapeake group bivalves. Paleobiology 9:261–268.

Keyser, D. 1995. Structural elements on the surface of ostracod shells. Pp. 5–10 *in* J. Riha, ed. Ostracoda and Biostratigraphy. Balkema, Rotterdam.

Kidwell, S. M., and A. K. Behrensmeyer, eds. 1993. Summary: estimates of time-averaging. Pp. 301–302 Taphonomic approaches to time resolution in fossil assemblages. The Paleontological Society, Knoxville, TN.

Kingsolver, J. G., and D. W. Pfennig. 2004. Individual-level selection as a cause of Cope's Rule of phyletic size increase. Evolution 58:1608–1612.

Kirkpatrick, M. 1982. Quantum evolution and punctuated equilibria in continuous genetic characters. Am. Nat. 119:833–848.

Kluge, A. G., and W. C. Kerfoot. 1973. The predictability and regularity of character divergence. Am. Nat. 107:426–442.

Knuth, D. E. 1969. The art of computer programming. Addison Wesley, Reading, MA.

LaBarbera, M. 1986. The evolution and ecology of body size. Pp. 69–98 *in* D. M. Raup and D. Jablonski, eds. Patterns and processes in the history of life. Springer-Verlag, Heidelberg.

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

———. 1980. The genetic covariance between characters maintained by pleiotropic mutations. Genetics 94:203–215.

Leroi, A. M. 2000. The scale independence of evolution. Evol. Dev. 2:67–77.

Liebau, A. 1971. Homologous sculpture patterns in trachyleberididae and related ostracodes. Nolit Publishing House, Belgrade.

Lofsvold, D. 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.

Maddocks, R. F. 1992. Ostracoda. Pp. 415–441 *in* F. W. Harrison and A. G. Humes, eds. Crustacea. Wiley-Liss, Inc., New York.

Marriog, G., and J. M. Cheverud. 2001. A comparison of phenotypic variation and covariation patterns and the role of phylogeny, ecology, and ontogeny during cranial evolution of New World monkeys. Evolution 55:2576–2600.

———. 2005. Size as a line of least resistance: diet and adaptive morphological radiation in New World monkeys. Evolution 59:1128–1142.

Marshall, C. R. 1990. Confidence intervals on stratigraphic ranges. Paleobiology 16:1–10.

Merilä, J., and M. Björklund. 1999. Population divergence and morphometric integration in the green-finch (*Carduelis chloris*)—evolution against the trajectory of least resistance? J. Evol. Biol. 12:103–112.

Mezey, J. G., and D. Houle. 2005. The dimensionality of genetic variation for wing shape in *Drosophila melanogaster*. Evolution 59:1027–1038.

Mitchell-Olds, T. 1996. Pleiotropy causes long-term constraints on life-history evolution in *Brassica rapa*. Evolution 50:1849–1858.

Okada, Y. 1981. Development of cell arrangement in ostracod carapaces. Paleobiology 7:276–280.

———. 1982a. Structure and cuticle formation of the reticulated carapace of the ostracode *Bicornucythere bisanensis*. Lethaia 15:85–101.

———. 1982b. Ultrastructure and pattern of the carapace of *Bicornucythere bisanensis* (Ostracoda, Crustacea). Pp. 229–267 *in* T. Hanai, ed. Studies on Japanese Ostracoda. Univ. Tokyo Press, Tokyo.

Paradis, E., J. Claude, and K. Strimmer. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20:289–2–90.

Price, T., and T. Langen. 1992. Evolution of correlated characters. Trends Ecol. Evol. 7:307–310.

Price, T., M. Turelli, and M. Slatkin. 1993. Peak shifts produced by correlated response to selection. Evolution 47:280–290.

R Development Core Team. 2005. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Renaud, S., J.-C. Auffray, and J. Michaux. 2006. Conserved phenotypic variation patterns, evolution along lines of least resistance, and departure due to selection in fossil rodents. Evolution 60:1701–1717.

Roff, D. A. 1995. The estimation of genetic correlations from phenotypic correlations: a test of Cheverud's conjecture. Heredity 74:481–491.

———. 1997. Evolutionary quantitative genetics. Chapman and Hall, New York.

Roff, D. A., T. A. Mousseau, and D. J. Howard. 1999. Variation in genetic architecture of calling song among populations of Allonemobius socius, A. fasciatus, and a hybrid population: drift or selection? Evolution 53:216–224.

Rohlf, F. J., A. J. Gilmartin, and G. Hart. 1983. The Kluge-Kerfoot phenomenon—a statistical artifact. Evolution 37:180–202.

Rohlf, F. J., and D. E. Slice. 1990. Extensions of the Procrustes method for the optimal superimposition of landmarks. Syst. Zool. 29:40–59.

Schluter, D. 1996. Adaptive radiation along genetic lines of least resistance. Evolution 50:1766–1774.

———. 2000. The ecology of adaptive radiation. Oxford Univ. Press, Oxford, U.K.

Simpson, G. G. 1953. The major features of evolution. Columbia Univ. Press, New York.

Smith, A. 1994. Systematics and the fossil record. Blackwell Scientific, Oxford, U.K.

Smith, A. B. 2000. Stratigraphy in phylogeny reconstruction. J. Paleont. 74:763–766.

Smith, A. B. 2002. Stratigraphy in phylogeny reconstruction–response. J. Paleont. 76:594–595.

Sokal, R. R. 1976. The Kluge-Kerfoot phenomenon reexamined. Am. Nat. 110:1077–1091.

Solow, A. R. 2003. Estimation of stratigraphic ranges when fossil finds are not randomly distributed. Paleobiology 29:181–185.

Stanley, S. M. 1979. Macroevolution. The Johns Hopkins Univ. Press, Baltimore, MD.

Steppan, S. J. 1997. Phylogenetic analysis of phenotypic covariance structure. II. Reconstructing matrix evolution. Evolution 51:587–594.

Steppan, S. J., P. C. Phillips, and D. Houle. 2002. Comparative quantitative genetics: evolution of the **G** matrix. Trends Ecol. Evol. 17:320–327.

Strauss, D., and P. Sadler. 1989. Classical confidence intervals and Bayesian probability estimates for ends of local taxon ranges. Math. Geol. 21:411–427.

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

Via, S., and R. Lande. 1985. Genotype-environment interaction and the evolution of phenotypic plasticity. Evolution 39:505–522.

Via, S., and A. J. Shaw. 1996. Short-term evolution in the size and shape of pea aphids. Evolution 50:163–173.

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.

Wagner, P. J. 2002. Testing phylogenetic hypotheses with stratigraphy and morphology—a comment on Smith (2000). J. Paleont. 76:590–593.

Walker, J. A. 2000. Ability of geometric morphometric methods to estimate a known covariance matrix. Syst. Biol. 49:686–696.

Webster, M., and N. C. Hughes. 1999. Compaction-related deformation in Cambrian olenelloid trilobites and its implications for fossil morphometry. J. Paleont. 73:355–371.

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

Zeng, Z.-B. 1988. Long-term correlated response, interpopulation covariation, and interspecific allometry. Evolution 42:363–374.

Associate Editor: D. Ayre

Appendix

In generating the null distribution of the test statistics used in this study, Knuth's (1969) algorithm was used to generate random directions in multivariate space. This method draws a random vector from a multivariate normal distribution in which all variables are uncorrelated and equally variable. This vector is then divided by its magnitude to scale to unit length. Knuth's method essentially picks a random point along the circumference of a unit circle (or, in higher dimensions, along the surface of a unit spheroid). In contrast, the broken stick distribution—which breaks a unit length "stick" into two pieces at a randomly chosen cut point—essentially chooses a random point along the line $y = 1 - x$ (Fig. A1). These two strategies are not equivalent—equal angular increments (grey dotted lines) result in equal-sized arcs on the unit circle, but not equal-sized segments on the broken stick line. Accordingly, Knuth's method generates random directions that are uniformly distributed in the quadrant, whereas the broken stick method disproportionately favors angles close to the coordinate axes (Fig. A1).

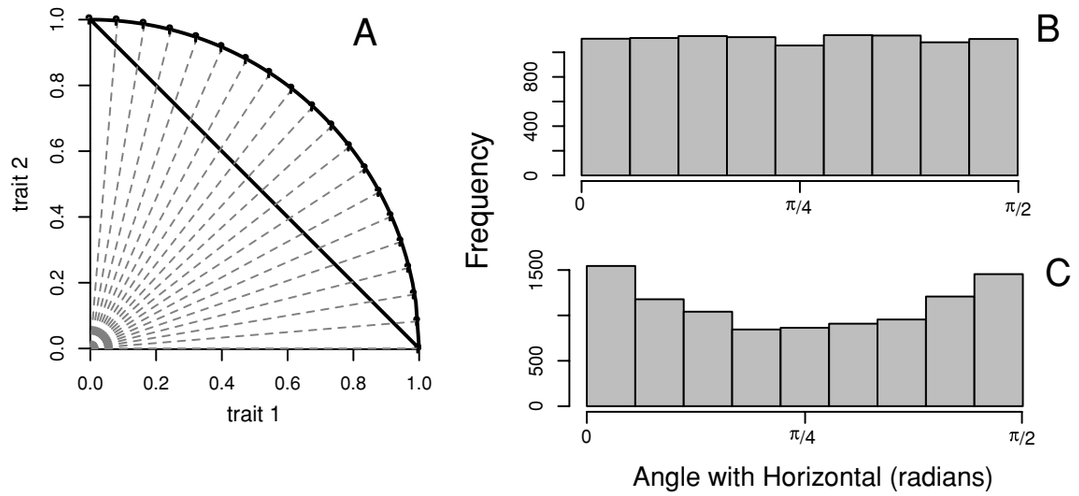

Figure A1. **Generation of random directions using Knuth's method and the broken stick distribution. (A) Upper right quadrant for two traits. The broken stick distribution draws vectors evenly from the line $y = 1 - x$ (solid black line), whereas Knuth's method draws random points on a unit circle. Points and dashed grey lines delimit equally spaced angles; equal angular increments are equally spaced on the unit circle but not on the line corresponding to the broken stick distribution. (B, C) Histograms of the angle between the positive *x*-axis and random vectors, generated according to Knuth's method (B) and the broken stick distribution (C). Note that Knuth's method produces the desired uniform distribution of angles, but the broken stick yields too many vectors close to the horizontal and vertical axes, and too few of intermediate angles.**