Time-Averaging and Morphology: Variability in Modern Populations and Fossil Assemblages of <u>Mercenaria</u> (Bivalvia)

Andrew M. Bush

Thesis submitted to the Faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Master of Science
in
Geological Sciences

Richard K. Bambach, Chair
Michal Kowalewski
Stephen E. Scheckler

August 9, 1999
Blacksburg, Virginia

Keywords: Evolution, Stasis, Punctuated Equilibrium, Procrustes Analysis, Taphonomy

Time-Averaging and Morphology: Variability in Modern Populations and Fossil Assemblages of Mercenaria (Bivalvia)

Andrew M. Bush

(Abstract)

The morphologic variability of a fossil assemblage is of interest in many paleontological studies. However, many fossil assemblages are time-averaged; that is, many generations of non-contemporaneous organisms are mixed into the same fossil bed. Assemblages of robust mollusk shells deposited in nearshore marine environments are often time-averaged over 100's to 1000's of years. Mixing many generations of a taxon can increase measured morphologic variability over that of a single generation if morphology is changing during the interval of time-averaging. If morphology is changing, time-averaging can also alter observed correlations between morphologic variables, as well as allometric growth patterns. If morphology is static, then time-averaging will not increase variability or otherwise obscure patterns of morphologic variability. Testing the effects of time-averaging on morphology will help determine the reliability of information derived from the fossil record.

In this study, morphologic variability was compared between 6 standing crop, living populations of Mercenaria campechiensis (Bivalvia) and two fossil assemblages of M. campechiensis and M. permagna. One fossil sample was collected as a series of superposed units that could be analyzed individually or in aggregate. The x,y coordinates of 13 landmarks and pseudolandmarks were recorded on over 600 valves, and variability was calculated using Least Squares Procrustes Analysis. Once corrections were made for allometry, the variabilities of the samples drawn from single time-averaged fossil beds were indistinguishable from the variabilities of the recent samples. For this data set, the variabilities of the fossil samples could be used without reservation to

estimate the variability of the standing crop populations from which they formed. Morphology was quite stable over the 100's to 1000's of years that likely passed as the assemblages accumulated.

A small amount of analytical time-averaging of the samples increases variability slightly, but additional analytical time-averaging causes no further increase. Very slight morphologic fluctuations are evident at time spans exceeding 100's to 1000's of years. Lumping geographically separated samples and samples of different species also increases variability.

Morphologic stasis is evident in Mercenaria over 100's to 1000's of years, but previous studies have indicated that evolutionary rates over this time frame are typically high. These studies are based on colonization events, however, and are biased towards high rates. Data gathered here and in previous studies suggest that local populations may evolve rapidly at their founding, but that stasis follows this initial burst of change. This model describes a pattern similar to Punctuated Equilibrium at a lower level of the genealogical hierarchy, and is here termed "Punctuated Equilibrium, Jr." This model can be further tested in empirical studies and should aid in determining the causes of species-level evolutionary patterns.

Dedication

To my parents and grandparents, with love.

Acknowledgements

I am deeply grateful to my committee chair, Richard Bambach, for his guidance and friendship, from the inception of this project to its completion (and beyond). Special thanks go to my committee members, Michal Kowalewski and Stephen Scheckler, for help and advice that have greatly improved the quality of my work. My most sincere thanks go to Gwen Daley, who took me under her wing and taught me some of those unwritten lessons I never knew I needed. My thanks go to Matt Powell for the long hours we shared over (large) boxes of clam shells. Thanks are due to Bill Arnold, whose wonderful samples allowed this project to grow beyond any of my initial expectations.

Many people and organizations have provided logistical and financial support that have made this project possible. Thanks to William Casey and the Caloosa Shell Corporation, PCS Phosphate, Quality Aggregates Inc., Roger Portell, Amy Tobias, Jennifer Wheaton, Brian Coffey, Bret Bennington, Tom Waller, Warren Blow, the Geological Society of America, the Byron Cooper Geoscience Fellowship, and the Department of Geological Sciences at Virginia Tech.

Thanks to the VT geology students who have shared these past few years with me. Special thanks to Rod Brame, Susan Barbour, Brooke Wilborn, Robin Guynn, Jake Beale, Vicki Suchanek, the ERHS posse, and my Cellar dinner pals for their friendship. Finally, my thanks and love to my family, whose support has never once wavered over the years.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Most macroinvertebrate fossil deposits are time-averaged; that is, fossils occurring in the same bed differ in age, often by hundreds to thousands of years (Flessa et al. 1993; Flessa and Kowalewski 1994; Kowalewski et al. 1998). The effects of time-averaging on paleoecologic analysis have been extensively tested, but the effects on morphology have received less attention. If the morphology of a taxon remains constant during the interval of time-averaging, then the morphologic variability of the resulting fossil assemblage should equal that of the standing-crop populations from which it formed. If morphology is changing, however, then the fossils should be more variable, since populations with different mean morphologies are being mixed. Morphologic change can be caused by evolution, ecophenotypy, or migration.

The effect of time-averaging on morphologic variability is of great importance to many paleobiological and geobiological studies. A variety of paleontological, microevolutionary, macroevolutionary, and geological questions require estimation of the variability of standing-crop, biologic populations. This has been termed the "biologic component of variation" by Bell et al. (1987), who distinguish it from the taphonomic component. Emiliani (1950) suggested using the morphological variability of fossils to infer physical paleoenvironmental parameters. Charlesworth (1984) presented a method for studying evolutionary rates in which the variability in fossil samples is assumed to be equivalent to the variability of standing-crop populations. Lande (1976) also suggests a method for studying evolution from fossil material that requires knowledge of the biological component of variability. Smith (1998) examined variability in Lower Paleozoic trilobites to test the hypothesis that changes in developmental regulation influenced the pattern of origination of higher taxa. He attempted to control taphonomic and methodological factors that might increase variability.

The relationship between phenotypic variability and evolutionary rate has been discussed by a number of paleontologists. Bader (1954) found an inverse

relationship between evolutionary rate and morphologic variability in two subfamilies of the oreodonts. However, Williamson (1981a, 1981b, 1986) described a positive correlation between evolutionary rate and morphologic variability in Cenozoic mollusks, which he attributed developmental instability in the evolving populations. Guthrie (1965) observed a positive correlation between rate and variability in the teeth of Cenozoic microtine rodents.

Determining the effect of time-averaging on morphologic variability is important so that fossil data can be appropriately interpreted and utilized. Moreover, an examination of variability can also test the morphologic stability of biological populations through time. A time-averaged fossil assemblage contains individuals drawn from populations that lived at a particular location through hundreds or thousands of years, a span of time rarely resolvable for time series data in the geologic record. Increased variability in fossil assemblages over conspecific living populations indicates probable morphologic change (for whatever reason); a lack of increase in variability indicates general morphologic stasis during the interval of time mixing. Debates on the morphologic stability of organisms have focused at the level of the species, but competing explanations for species-level morphologic stasis make varying predictions about the evolutionary stability of the constituent local populations. Patterns of evolution over thousands of years may help explain how many species persist apparently unchanged for millions of years.

In this study, the effect of time-averaging on the morphologic variability of Mercenaria (Bivalvia) was tested using samples collected from six standing crop, modern populations and two fossil assemblages from the southeastern United States. The coordinates of 13 landmarks on each shell were recorded, and the variability of each sample was determined using general least-squares Procrustes analysis. The variability of six samples of living M. campechiensis established the range of variability in living populations. These were compared to two fossil samples, one sample of M. campechiensis from Florida and one of M. permagna from North Carolina. Both samples should be time-averaged, but

the M. campechiensis from Florida were collected through 2.4 m of section and must represent a significant amount of time. If significant morphologic evolution (or any other type of morphologic change) was occurring during the deposition of these samples, then they should be more variable than the living populations.

Chapter 2:
Time-Averaging and Morphology

The Law of Superposition allows the relative dating of sedimentary rocks and the fossils they contain. Simply put, younger strata and fossils lie on top of older ones in an undeformed stratigraphic package. Geologists and paleontologists have used this law to reconstruct the geological and biological history of the earth, and twentieth century workers have verified its predictions using isotopic dating.

However, the Law of Superposition does not always apply when examining narrow stratigraphic intervals. Younger fossils do not always occur stratigraphically higher than older ones; fossils of varying ages are often mixed in the same horizon. An older fossil can even occur slightly above a younger one, a phenomenon known as "stratigraphic disorder" (Flessa et al. 1989; Cutler and Flessa 1990). Walker and Bambach (1971) coined the term "time-averaging" for the mixing of differently-aged fossils in a single bedding plane. More generally, time-averaging is "the process by which events that happened at different times appear to be synchronous in the geological record" (Kowalewski 1996).

The Process of Time-Averaging

The duration of time-averaging in fossil assemblages varies greatly (Fig 2.1). Kidwell and Bosence (1991) define four general categories of time-averaging in fossil deposits. A "census assemblage" results from the sudden burial or mass mortality of a community and contains little time-averaging. "Within-habitat time-averaged assemblages" contain fossils mixed over years to thousands of years in coastal settings; longer periods of mixing are expected farther offshore. They contain fossils drawn from a single, yet not necessarily static, environment. "Environmentally condensed assemblages" form when environmental changes occur fast enough that organisms from different habitats

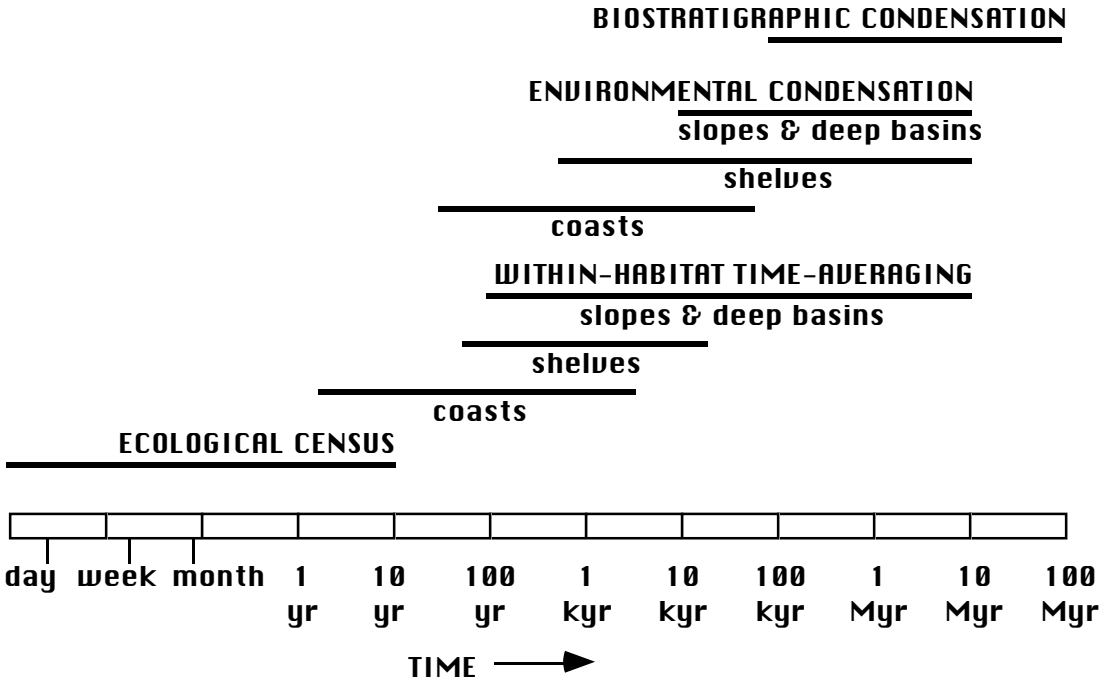

Figure 2.1. Varying scales of time-averaging in marine macroinvertebrate deposits. See text for definitions. Redrawn from Kidwell and Bosence (1991, p. 176).

are mixed in a single horizon. The time represented in these assemblages is typically somewhat greater than that of within-habitat time-averaged assemblages. "Biostratrigraphically condensed assemblages" contain taxa with non-overlapping geologic ranges. They can encompass hundreds of thousands to many millions of years. Within-habitat and environmentally condensed assemblages are common in the fossil record, and many studies have tested the effects of intermediate-level time-averaging on paleobiologic patterns (e.g., Miller 1988; Staff and Powell 1988). Census assemblages and biostratigraphically condensed assemblages, which contain very small or very large amounts of time-averaging, can generally be recognized by field observations and knowledge of the geologic ranges of the enclosed fossils.

The absolute amount of time-averaging in marine macroinvertebrate assemblages is of great interest, since many stratigraphic and paleobiologic studies are based upon these deposits. Many methods have been used to estimate the absolute duration of time-averaging (see Flessa [1993] for a review),

but dating of young fossil and subfossil shells by radiocarbon or amino acid racemization provides the most direct information. Dead shells co-occurring on the sediment surface are available for permanent burial at the same stratigraphic level, so the range of ages in a surface collection of dead shells from an active sedimentary environment is also a measurement of the duration of time averaging. Many studies of this sort have shown that time-averaging on the order of 100's to 1000's of years is a dominant feature of marine macrofossil deposits (Table 2.1). However, the generation of a time-averaged assemblage is complex, and many factors influence the duration of time-averaging in a particular deposit.

When Walker and Bambach (1971) first proposed that fossil deposits were time-averaged, they invoked the slowness of sediment accumulation relative to the lifespan of fossilizable organisms as the cause. If many generations of a species can live and die during the time required to bury their skeletons, they reasoned, then non-contemporaneous individuals should be preserved in the same stratum. However, experimental taphonomy has shown that bivalve shells sitting on the seafloor are rapidly destroyed by biological and physical attack, at least in shallow water (Cutler 1994). Dead bivalve shells can be severely degraded in a single year. Low sedimentation rates relative to the generation length of organisms cannot alone cause time-averaging greater than a few years.

In fact, shells are much more persistent than expected because they often do not remain on the sediment surface. They can become shallowly buried, where the surrounding sediment prevents abrasion, boring, and other destructive processes. Once shallowly buried, the shells can become more deeply and permanently buried (facilitated by rapid sedimentation rates), or they may be exhumed and mixed with younger shells. The upper layer of sediment in which shells are subject to reworking and exhumation has been called the mixing zone or "taphonomically active zone" (TAZ), since shells in this zone are still available for taphonomic alteration (Davies et al. 1989). Once a shell passes below the base of the mixing zone or TAZ, it is relatively difficult to disturb.

Table 2.1. Several estimates of time-averaging determined from recent marine sedimentary environments.

Paper	Type of collection	Species	Duration (years)
Flessa et al. 1993	17 shells from surface of tidal flat on a sediment-starved shelf	Chione fluctifraga & Chione californiensis	3,569
	13 shells from tidal channel on a sediment-starved shelf		1,752
Flessa and Kowalewski 1994	literature survey, surface collections, nearshore	varies	Median 1,250, Range 0 to 45,000
	literature survey, surface collections, shelf		Median 9,190, Range 0 to 40,400
	literature survey, beach ridge deposits		Median 1,390, Range 0 to 35,460
	literature survey, nearshore deposits		Median 830, Range 0 to 19,950
Meldahl et al. 1997	24 shells from fan-deltas in a faulted rift basin	Chione fluctifraga & Chione californiensis	800—1170
	24 shells from fan-deltas in a faulted rift basin		920—1260
	24 shells from pocket bays in a faulted rift basin		280—530
Kowalewski et al. 1998	Shells sampled from single levels in beach ridges, sample size from 7 to 21	Chione fluctifraga	279, 193, 1063, 516, 763, 848, 730, 494, 1063

Reworking of sediment and shells near the sediment-water interface has many causes. The role of bioturbation is somewhat unclear (Kidwell and Bosence 1991). Callianassid shrimp and some polychaete worms concentrate shells in subsurface beds by cycling sediment (Meldahl 1987). They ingest sediment at the bottom of their burrow, then excrete in onto the sediment

surface. This causes the sediment to cycle downwards, transporting shells and other coarse particles. Because the shells are not ingested and transferred to the surface, they form subsurface concentrations (Meldahl 1987). However, Kershaw et al. (1988) observed that the shells of Turritella communis in the Irish Sea were not concentrated in the subsurface despite the presence of a callianassid and a deep-burrowing worm. In fact, they found that T. communis shells increased steadily with depth, whereas the finer sediments around them were extensively time-mixed (Kershaw et al. 1988). Computer models suggest that mixing by non-cyclic bioturbation is inefficient at disordering shells, although reworking from older deposits can easily cause disorder (Cutler and Flessa 1990). Reworking of shells and sediment by storms, waves, tides, and currents can also mix shells of different ages.

Despite these complications, it is still widely accepted that sedimentation rate is a primary control on the duration of time-averaging (Kidwell and Bosence 1991, Meldahl et al. 1997). Greater rates of sedimentation yield lower durations of time-averaging, since older shells get buried below the taphonomically active zone more quickly and cannot be mixed with younger shells. In fact, biostratigraphically condensed assemblages, which contain the greatest amount of time-averaging, typically occur where sedimentation rates are particularly slow (Kidwell and Bosence 1991).

Flessa and Kowalewski (1994) suggest that time-averaging in nearshore environments is less than in offshore environments because sedimentation is typically more rapid in nearshore environments. Offshore environments are farther from clastic input and have lower rates of carbonate production. Modern offshore environments have been sediment starved since the Holocene rise in sea level. In a literature survey of radiocarbon dates on shells in active sedimentary environments, they found that the median duration of time-averaging in nearshore environments was 1,250 years, whereas the median duration in offshore environments was 9,190 years (see Table 2.1). The age ranges of shells in inactive beach ridges (median 1,390 years) and other nearshore deposits (830

years) are similar to those obtained for active nearshore sedimentary environments. There is, however, great variation in the estimates of time-averaging among different locations. For example, the estimates of the amount of time-averaging in active, nearshore environments ranged from zero to 45,000 years. Because of their methodology, these numbers are minimum estimates, so a measured interval of zero does not mean time-averaging was absent.

Different amounts of time-averaging are also expected in different tectonic environments, since tectonic activity significantly affects sedimentation rates. Meldahl et al. (1997) measured the amount of time-averaging in sediments in two different tectonic environments. Shells at Bahia Concepcion, a faulted rift basin that has high rates of sedimentation due to the proximity of highlands, varied in age by 800 to 1260 years. At Bahia la Choya, an intertidal bay located on a sediment-starved shelf, shells were time-averaged over 1310—1690 years in the tidal channel and over 2970—3450 years on the tidal flats (Meldahl et al. 1997).

The resistance of a particular taxon's skeleton to destruction is also a controlling factor in time-averaging (Kowalewski 1996). A type of shell that is easily destroyed cannot undergo many cycles of burial and exhumation; it will be destroyed before it can be mixed with shells that are significantly younger. A durable shell, however, can survive longer in the TAZ and be more thoroughly mixed. The resistance of a shell to destruction depends on size, thickness, and microstructure (organic matrix as well as mineral). Driscoll (1970) also suggests that thicker, more robust shells are buried more rapidly than thinner shells, removing them from the destructive sediment surface. Different taxa in the same bed may be time-averaged over different lengths of time ("disharmonious time-averaging," Kowalewski 1996).

Other factors influence the length of time it takes to destroy a shell. Meldahl et al. (1997) note that time-averaging is probably greater in intertidal environments than in continually submerged sites, since both bioerosion and shell abrasion are slowed significantly by areal exposure. Although some

biological activities (e.g., boring) accelerate shell destruction, others help preserve shells. Shells coated by coralline algae, for example, seem to partially protected from boring (Kidwell and Bosence 1991). Organic algae may protect shells by shielding them from chemical corrosion; however, this protection is dependent on the density of dead shells. In areas with many dead shells, grazing gastropods feed on attached algae and expose the shells to erosion; if a critical density of dead shells is not reached, the gastropod population cannot be supported (Cutler 1994).

Opinions are varied on the temporal distribution of shell ages within a time-averaged accumulation. Some have reported that shells in time-averaged deposits and surface collections are not uniformly distributed in age; young shells dominate and old shells are relatively few (Meldahl et al. 1997; Flessa 1993). Plots of shell abundance vs. age in surface collections resemble exponential decay curves (Fig 2.2). However, Kowalewski et al. (1998) report that the age distributions of shells taken from single levels in Colorado Delta cheniers do not significantly differ from uniform.

Collection technique and post-collection treatment of a sample can also influence the amount of time represented. If fossils from different stratigraphic horizons are mixed during or after collection, the amount of time-averaging can increase. Such "analytical time-averaging" can significantly reduce the temporal resolution of fossil data (Fürsich and Aberhan 1990).

Effects on Paleobiological Analysis

Walker and Bambach (1971) acknowledge that time-averaged fossil accumulations cannot reveal the structure of an ancient community at a single instant, and cannot be used to track short-term (e.g., yearly) fluctuations in population sizes and ecosystem dynamics. However, they suggest that fossil communities better represent long-term ecologic patterns because time-averaging removes short-term ecologic "noise" that is observed by neontologists.

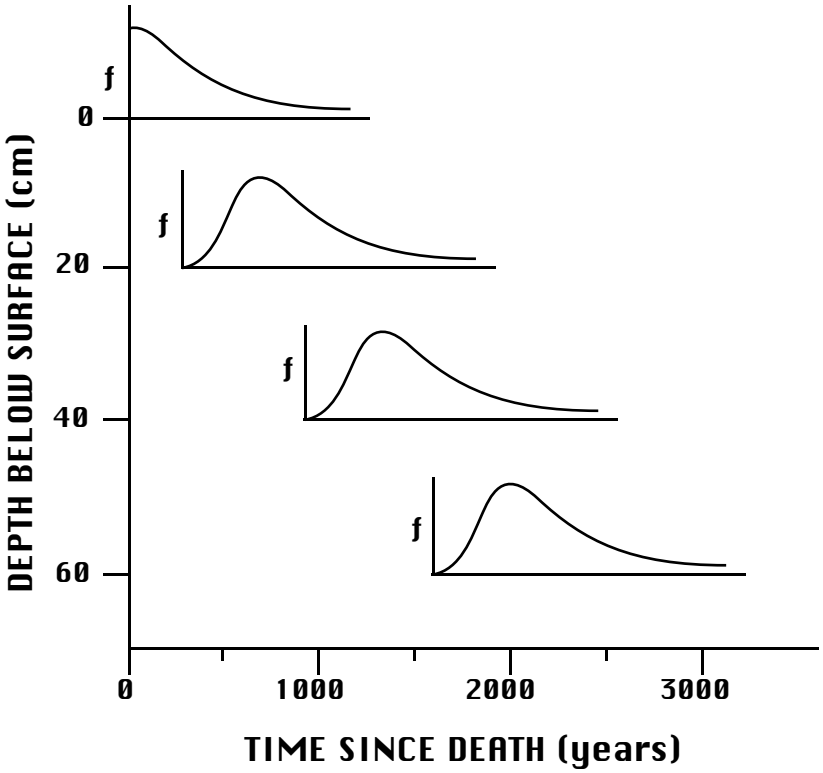

Figure 2.2. Schematic distribution of shell ages at varying levels in a core. Note the temporal overlap of successive samples, and the skewed frequency (*f*) of shell ages at each level. Based on shell ages in surface collections and shallow cores from tidal flats in Bahia la Choya, Mexico. Redrawn from Flessa et al. (1993).

The effect of time-averaging on paleobiological analysis depends on the relative scale of time-averaging and the process being studied (Kowalewski 1996). If the process being studied occurs over a shorter time span than the duration of time-averaging, then time-averaging will obscure observation of the process. For example, yearly fluctuations in the population size of a taxa will not be resolvable in a bed averaged over a thousand years. However, if the process being observed (e.g., long-term community structure) occurs on a greater time scale than that of time-averaging, then noise (short-term fluctuations) in the signal will be eliminated, as suggested by Walker and Bambach (1971). In order to properly interpret paleobiologic data, the time scale of both the process and of time-averaging must be known (Kowalewski 1996).

Research into the effects of time-averaging on the interpretation of paleobiological patterns has focused almost exclusively on the (non-) correspondence between ecologic and paleoecologic data. In fact, the concept of time-averaging originated in an explicitly paleoecologic context. In the original paper on time-averaging, Walker and Bambach (1971) assert that fossil assemblages "do not reflect instantaneous community structure, as do grab samples of a modern community, but instead reflect community structure as it persisted through time." Subsequent papers describing the ecologic resolution of the fossil record are many (e.g., Fürsich 1978; Fürsich and Aberhan 1990; Kidwell and Bosence 1991; Kowalewski 1996); two studies from modern sedimentary environments are described next as examples.

Miller (1988) sampled the dead mollusks along a 360-meter transect in a back-barrier lagoon in St. Croix. He found that the death assemblages nicely reflected a transition from vegetated to non-vegetated bottom along the transect, demonstrating that small-scale horizontal variations can be preserved in the fossil record. Staff and Powell (1988) found that the inverse relationship between species-richness and environmental variability, well-known in ecology, does not always hold in paleoecologic analysis. They collected samples of living benthos from several habitats periodically for 24 to 30 months. The standing diversity of preservable organisms was generally higher in the stable, euhaline environment than in the variable estuarine habitat, but the variability of the estuarine habitat allowed different organisms to migrate in and out over time. Cumulatively, the diversities were comparable between the habitats (Staff and Powell 1988).

Time-averaging can also complicate evolutionary studies of the fossil record. Evolutionary changes in a lineage cannot be traced at a finer scale of resolution than the local magnitude of time-averaging. Stratigraphic disordering of fossils further complicates microeveolutionary studies by causing temporal overlap of samples collected at very fine stratigraphic intervals (Fig. 2.2). For example, Flessa et al. (1993) suggest that spacing samples by 50 cm is necessary to avoid temporal overlap in tidal deposits similar to their study area, though

they speculate that smaller intervals would be acceptable in many other environments. Unusual depositional systems such as varved lakes can provide better time resolution, but the typical marine section will not.

Time-Averaging and Morphology

Time-averaging can potentially confuse studies of the fossil record by increasing the amount of morphologic variability in a set of fossils (Fig. 2.3). Time-averaging mixes a series of generations of a species; if the average morphology changes from generation to generation, then morphologic variability in the time-mixed fossil sample will be greater than that in a standing-crop population. If, however, the average morphology is static, then the individuals being mixed come from identical distributions and morphologic variability is unchanged by time-averaging (Fig 2.3).

Many factors can cause mean morphology to change through time. Morphologic evolution, shifts in a population's genome that impact the phenotype, can cause great changes over tens or hundreds of years. For example, Losos et al. (1997) found that Anolis lizards introduced onto several islands evolved rapidly over 10 to 14 years. Gingerich (1983) cites other examples of high evolutionary rate in colonizing populations over hundreds of years. Migration of organisms can also cause a change in the morphology observed at a location. If a local population is replaced (entirely or in part) by another that differs in morphology, the taphonomic mixture of the populations will increase the observed variability.

Many organisms vary ecophenotypically, exhibiting different morphologies under different environmental conditions. For example, modern mytilid bivalves and Ordovician Ambonychia sp. tend to grow larger in onshore environments than in offshore ones (Daley 1999). Patton and Brylski (1987) relate variations in pocket gopher body size to nutrition, and Robinson and Wilson (1995) induced guppies to assume different body shapes under different

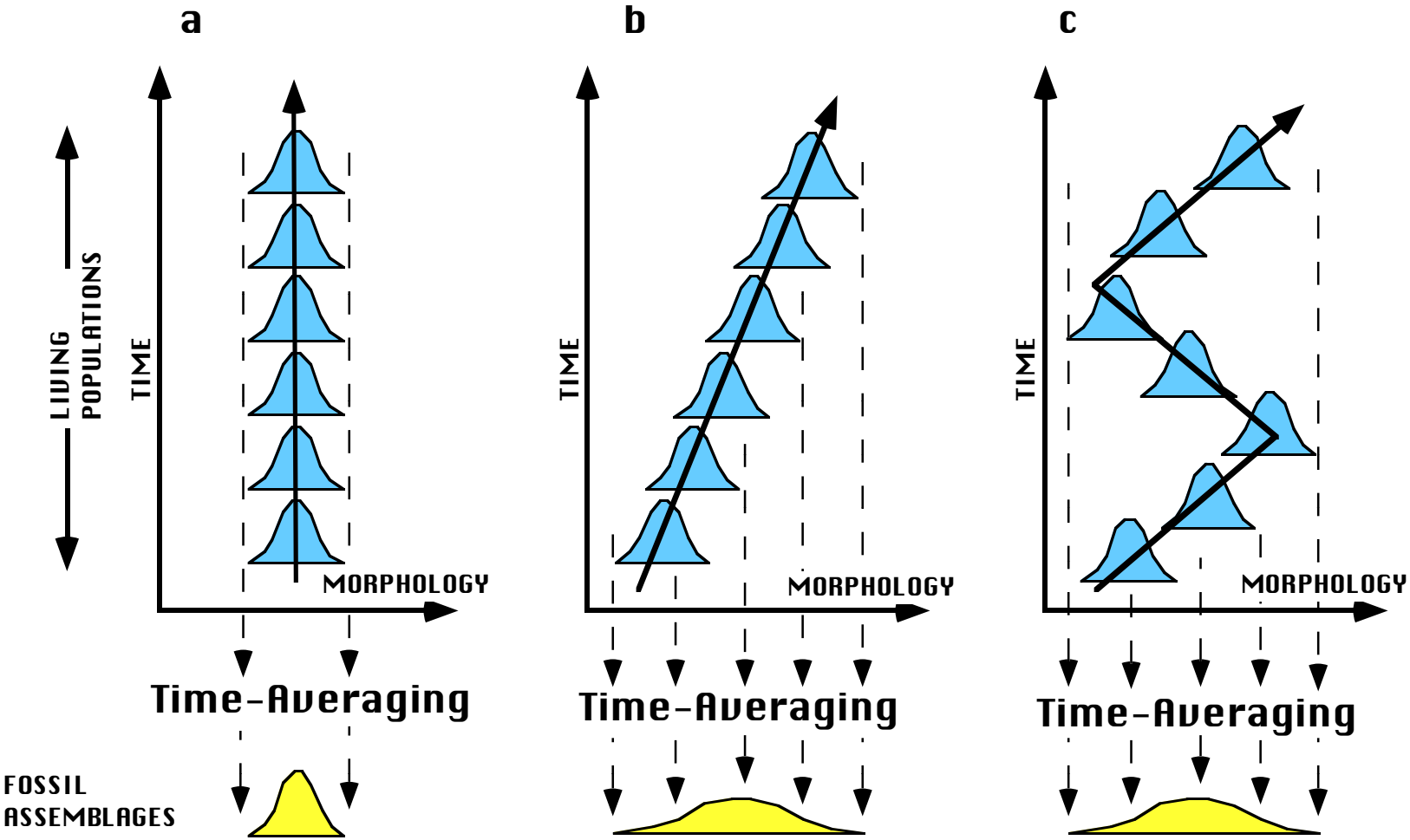

Figure 2.3. The potential effects of time-averaging on morphologic variability. In (a), stasis causes no increase in variabililty in the time-averaged assemblage. In (b) and (c), morphologic change causes an increase in variability.

feeding conditions. Environmental changes, which may be unobservable in the geologic record, can thus drive morphologic change and increase variability. Using a computer model, Chapman (1994) simulated the effects of time-averaging on morphologic variance in an ecophenotypically varying population. He found that, in theory, variance can increase in such a situation.

In addition to time-averaging, other taphonomic and diagenetic processes can affect the variability of a fossil assemblage. Selective sorting of skeletal elements by currents can decrease variability by removing certain morphs. On the other hand, currents can transport individuals from other environments, potentially increasing variability. Chipping, abrasion, boring, dissolution, and other physical, biological, and chemical agents can alter the morphology of skeletal elements, increasing variability. Post-depositional recrystallization and dissolution can obscure morphology. Tectonic deformation can stretch, compress, or shear a fossil (Bambach 1973).

Kidwell (1986) and Kidwell and Aigner (1985) discuss the interplay of evolution, ecophenotypy, and taphonomy in the formation of complex shell beds. They discuss the case where microevolution or ecophenotypic change is occurring in a taxon during the time it takes for a shell bed to accumulate. Kidwell (1986) proposes that the pattern of morphologic variability depends on the sedimentary history of the bed. If sedimentation rate decreases during the deposition of the bed, it will grade from sparsely fossiliferous and lightly time-averaged to fossiliferous and heavily time-averaged, since sedimentation rate is a prime control on time-averaging. In this case, variability could increase up-section since more generations are being mixed at any given level. However, lowering the sedimentation rate up-section increases the residence time of skeletal elements in the taphonomically active zone, potentially decreasing variability by selectively destroying certain morphs (Kidwell 1986). Of course, taphonomic processes could also increase variability by altering but not destroying organic hard parts. Shell beds that form during an increase in sedimentation rate will display the opposite pattern (Kidwell 1986).

Time-Averaging and Morphology: Theoretical Models

The effects of time-averaging on morphologic variability can be described by a series of simple models. In the examples below, variability is assumed to be measured as variance, standard deviation, mean squared distance to centroid, or some other measure of average distance to a mean value. The effect of using the range of values is discussed later.

The amount by which time-averaging increases the variability of a fossil assemblage depends on two factors: distribution of specimens in time and distribution of population means in morphospace. First, consider a situation where both distributions are uniform. The distribution in time will be uniform when the time-averaging process provides a continuous record through the interval of time-averaging with the fossils in even temporal distribution. Kowalewski et al. (1998) report such continuous, even temporal coverage in bivalve shells in Colorado Delta cheniers. A uniform probability distribution of means can be achieved by a simple linear trend with a constant rate, but it can also occur when the mean fluctuates evenly within a range of values (Fig 2.3 b-c).

If the distribution of means in morphospace deviates from uniformity, the amount of variability inflation can either increase or decrease. Consider several situations in which time-averaging is uniform and the morphology remains within a given range. A linear trend with constant rate will produce a uniform distribution of means and a certain increase in variability (Fig 2.4a). However, if a population exhibits a pattern resembling punctuated equilibrium, quickly switching between two stable configurations, the extreme values of morphology are more heavily sampled (Fig 2.4b). Compared to a uniform distribution within the same range of morphology, the average distance (or squared distance) to the mean morphology is greater since intermediate values are undersampled. If the path of the mean through morphospace contains a sequence of two closely-

spaced reversals (Fig 2.4c), the central values will be oversampled, and the magnitude of variability increase will be reduced.

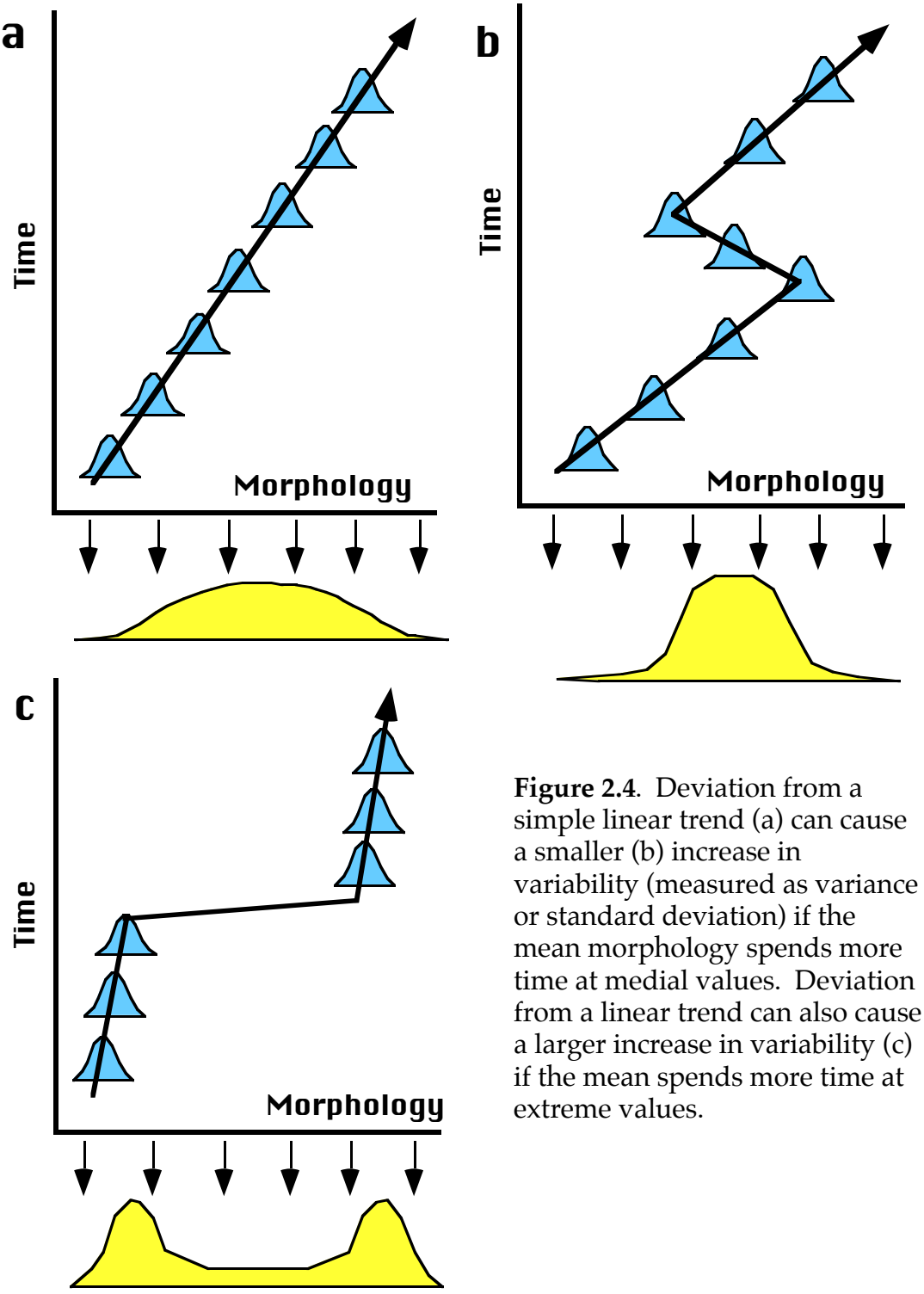

Figure 2.4. Deviation from a simple linear trend (a) can cause a smaller (b) increase in variability (measured as variance or standard deviation) if the mean morphology spends more time at medial values. Deviation from a linear trend can also cause a larger increase in variability (c) if the mean spends more time at extreme values.

Likewise, deviations from uniformity in age distribution of the time-averaging process will oversample some portions of the morphologic transition and undersample others. This can either increase or decrease the variability of the resulting fossil assemblage relative to an assemblage produced by uniform time-averaging (2.5a-b, Figs 2.4a). For example, time-averaging in a deposit could be non-continuous, only sampling individuals at the beginning and end of a linear trend (Fig. 2.5a). The oversampling of extreme values will inflate the average distance to the mean. Conversely, a normal distribution of ages (or a skewed distribution, e.g. Fig 2.2) will yield less of an increase in variability relative to a uniform distribution when superimposed on a constant linear trend in morphology (Fig 2.5b). The normal or skewed distribution will heavily sample the middle portion of the morphologic range while neglecting the more extreme values. Evidence for skewed distributions of ages in some environments, like that provided by Flessa et al. (1993), makes this scenario plausible and important.

Finally, different results will be obtained if range is used to measure variability. If time-averaging is continuous, the probability distribution of the fossils in time no longer affects the actual range of values in the population. (Population is used here in the statistical sense). If the record is discontinuous, variability will be underestimated if the span of time containing the most extreme morphologies is missing. The path through morphospace likewise becomes irrelevant to the actual range of values. However, the distribution of the fossils in time and morphospace will affect the ability to estimate the range from a sample. If the distribution of morphologies in a sample has long, thin tails (Figs 2.4c and 2.5b), then realistic estimates of the range will be hard to obtain without very large samples. If the distribution in morphospace emphasizes extreme values (Figs 2.4b and 2.5a), then the range will be more reliably estimated at a particular sample size.

This discussion has centered so far on situations where the variability of a fossil assemblage is greater than that of the living populations from which it

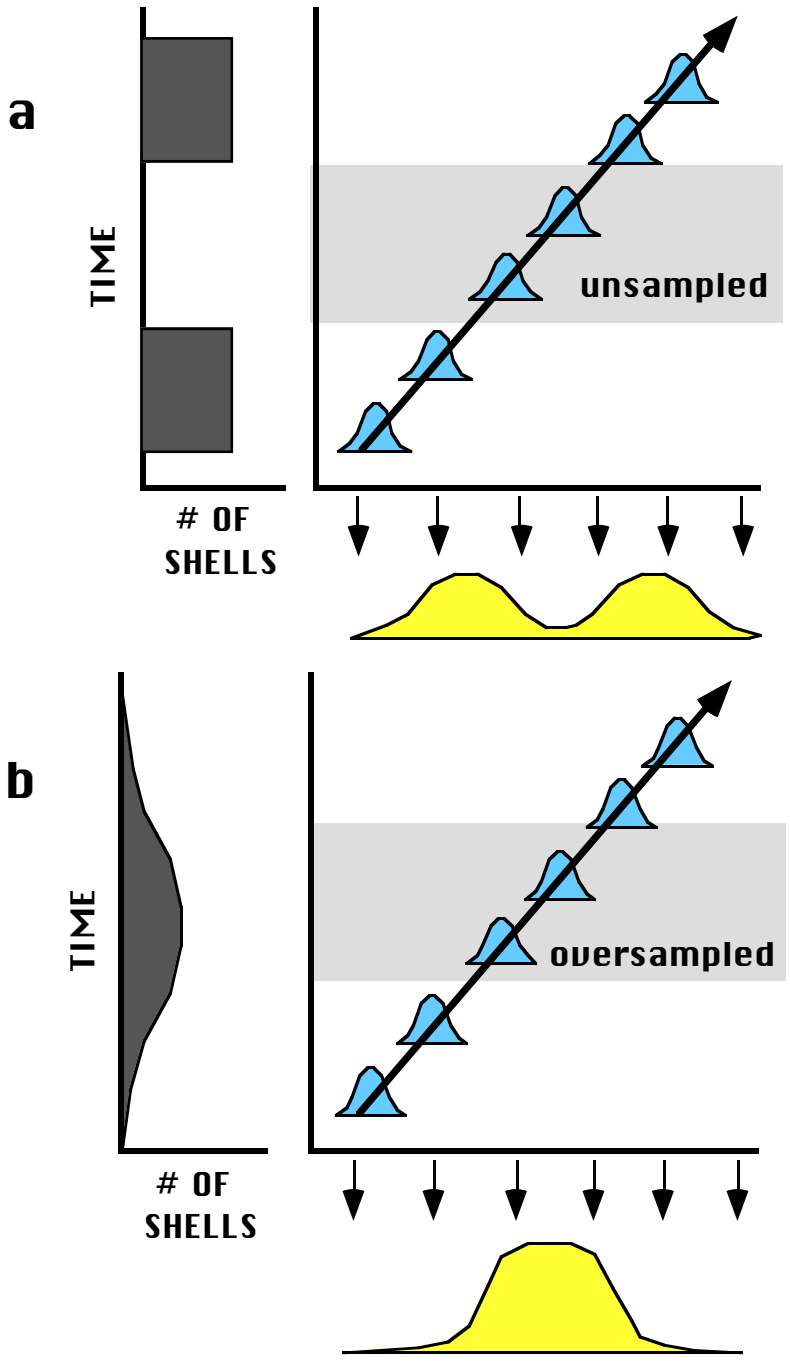

Figure 2.5. Non-uniform time-averaging can cause a larger variability increase by removing values near the mean. However, non-uniformity can also cause a smaller variability increase by oversampling values near the mean.

forms. Commensurate variability, evidence of morphological stasis, is of course a viable scenario (Fig 2.3a); less discussion is given to this situation since less is needed. If little or no morphologic change occurs, then the probability distribution of the means is very tight, and the temporal probability distribution has no effect on the variability of the fossil assemblage.

Stasis is, however, a useful model to demonstrate one final scenario: fluctuation in standing-crop population variability through time (Fig. 2.6). These changes can be considered "noise" that the time-averaging process filters out; the variability of the fossil assemblage will be an average value. The fossils will be more variable than some living population and less variable than others. However, the fossil sample will be more variable than most of the populations that produced it. This can be demonstrated simply with a simulation. Three samples of mean zero and sample size 1000 were generated from a normal distribution. The standard deviations were 1.00, 2.00, and 3.00. When the samples were combined, the standard deviation was 2.16. If the variability of a population does vary through time, then a single standing-crop population will be unreliable for comparisons with time-averaged fossil material. Multiple living populations should be sampled to determine the range of standing-crop population variability.

Actual time-averaged assemblages will be more complex than these simple heuristic models. The temporal distribution within a time-averaged assemblage can vary (Flessa et al. 1993; Kowalewski et al. 1998), especially between robust and fragile taxa (Kowalewski 1996), and may never be known for a particular sample. The path through morphospace followed by a taxon while a single time-averaged bed accumulates is unobservable.

In addition to increasing variability, time-averaging can also alter the correlation structure of a set of variables if morphology is changing through time. Two characters that are completely uncorrelated in a series of standing-crop populations can appear correlated in a time-averaged assemblage if both

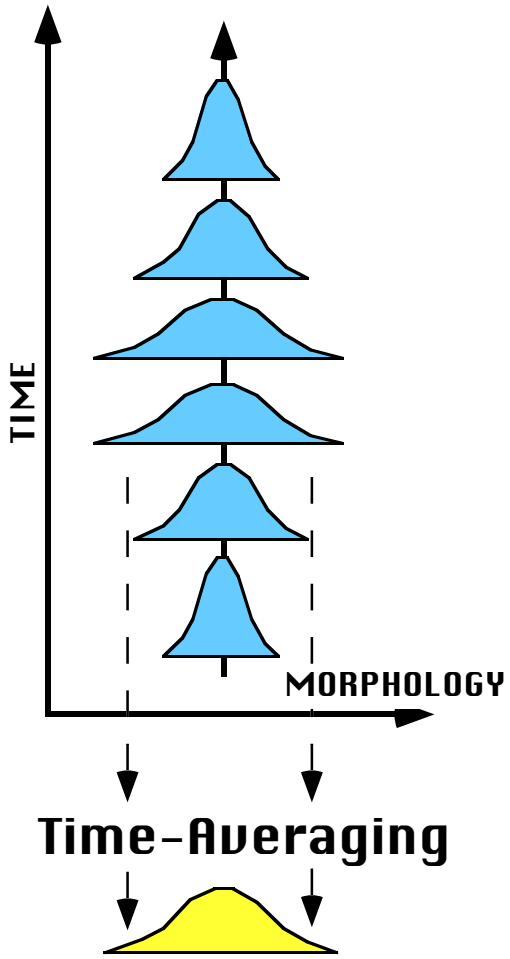

Figure 2.6. Influence of fluctuating variability in living populations on the variability of the resulting fossil assemblage. The fossils will have an intermediate variability, but it will be higher than the average variability of the living populations.

characters are evolving (Fig 2.7a-b). The correlation produced can be positive or negative. Time-averaging can also mask the correlations present in standing-crop populations. Two characters that are positively correlated in a series of populations can appear uncorrelated if one character increases in magnitude through time while the other decreases (Fig 2.7c-d). A negative correlation can be masked if the two characters both increase or both decrease.

Time-averaging also has the potential to mask patterns of allometric growth. For example, consider a morphologic character (such as the height-to-length ratio of a shell) that shows positive allometry in a series of populations

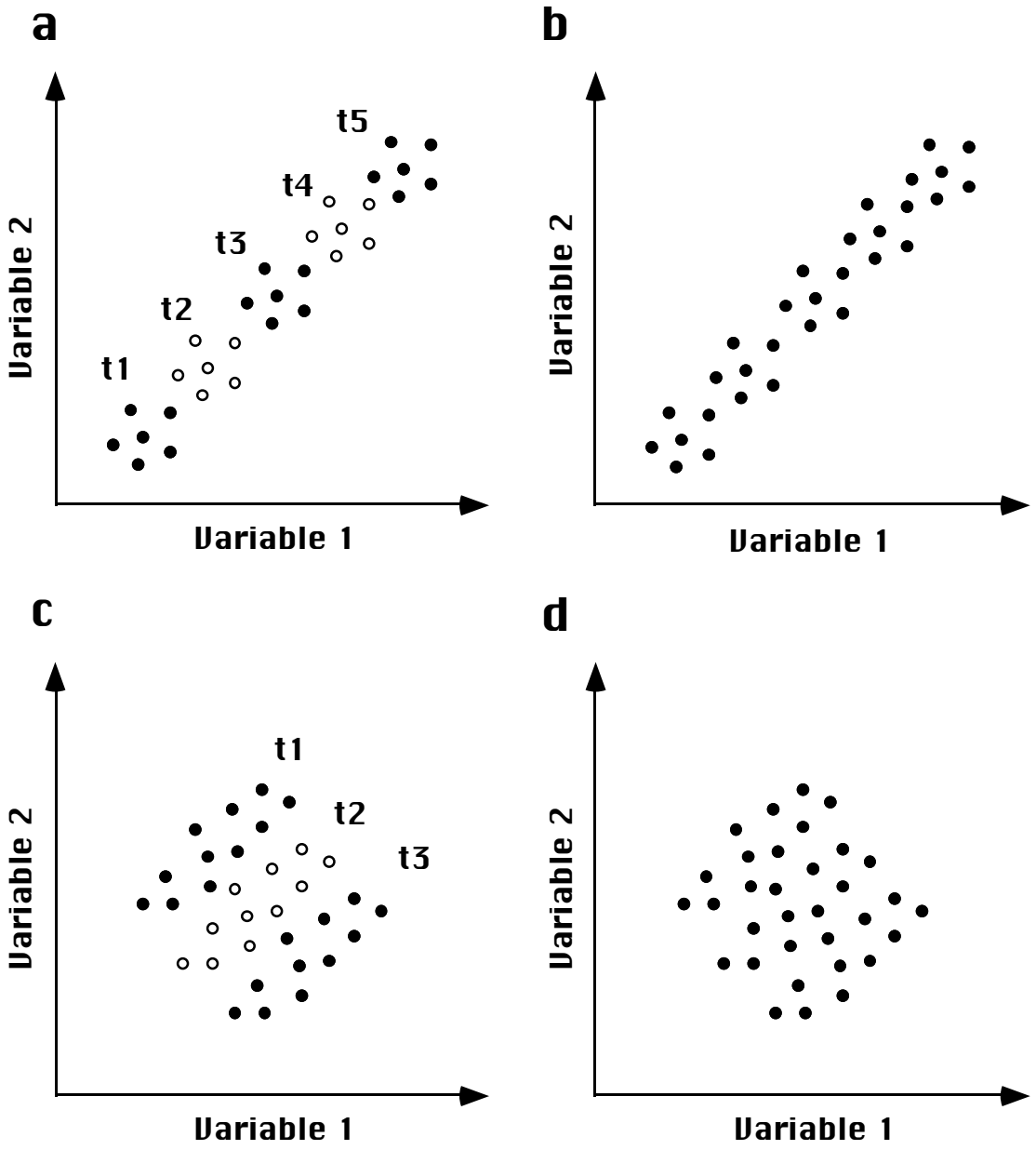

Figure 2.7. The potential effects of time-averaging on morphologic correlations. Five populations at five times (t1 to t5) are shown in (a). Variables 1 and 2 are uncorrelated in each population, but both variables are evolving. When the populations are time-averaged (b), the two variables appear correlated. Time-averaging can also obscure correlations. In (c), two variables are correlated in each of three successive populations. However, when time-averaged, no correlation is observed.

but decreases in absolute magnitude through time (Fig 2.8a). In the standing crop populations at times t1 and t2, there is a strong correlation between size and height to length ratio. When the populations are mixed, however, no relationship exists between the two variables, and the distribution is a random scatter (Fig 2.8b). No allometric relationship is observable, though allometry was always present. A change in the rate of allometric change can also obscure the allometric relationships within standing-crop populations (Fig 2.8c-d). In Fig 2.8d, there is still an observable relationship between size and height-to-length ratio, but there is much more scatter than in a biologic population. Time-averaging can also introduce false allometry into fossil accumulations (Fig 2.8e-f).

Other Sources of Morphologic Variability

Variability of fossil samples is affected by collection and selection procedures. Bell et al. (1987) warn against using museum samples that may not have been randomly collected. Daley (1993) found that museum samples of the articulate brachiopod Onniella labeled as "bulk collected" from a particular locality contained only the largest size-class available at that outcrop. Lumping of cryptic sister species that are slightly differentiated in morphology can increase variability, and exclusion of morphologic outliers incorrectly identified as a different species can lower variability. Analytical time-averaging or lumping of samples from different localities can exaggerate variability (Bell et al. 1987).

The variability of a standing-crop population, termed the "biologic component of variation" by Bell et al. (1987), has several sources. Phenotypic variability that is based in genetic variability is generally the factor of greatest interest in evolutionary studies, even if it cannot be separated from other factors. If members of a local population inhabit a variety of microenvironments, then ecophenotypic variability may add to the genetic component. Sub-optimal environmental conditions can increase the morphologic variability of some organisms; Emiliani (1950) suggests that variability of fossil assemblages could

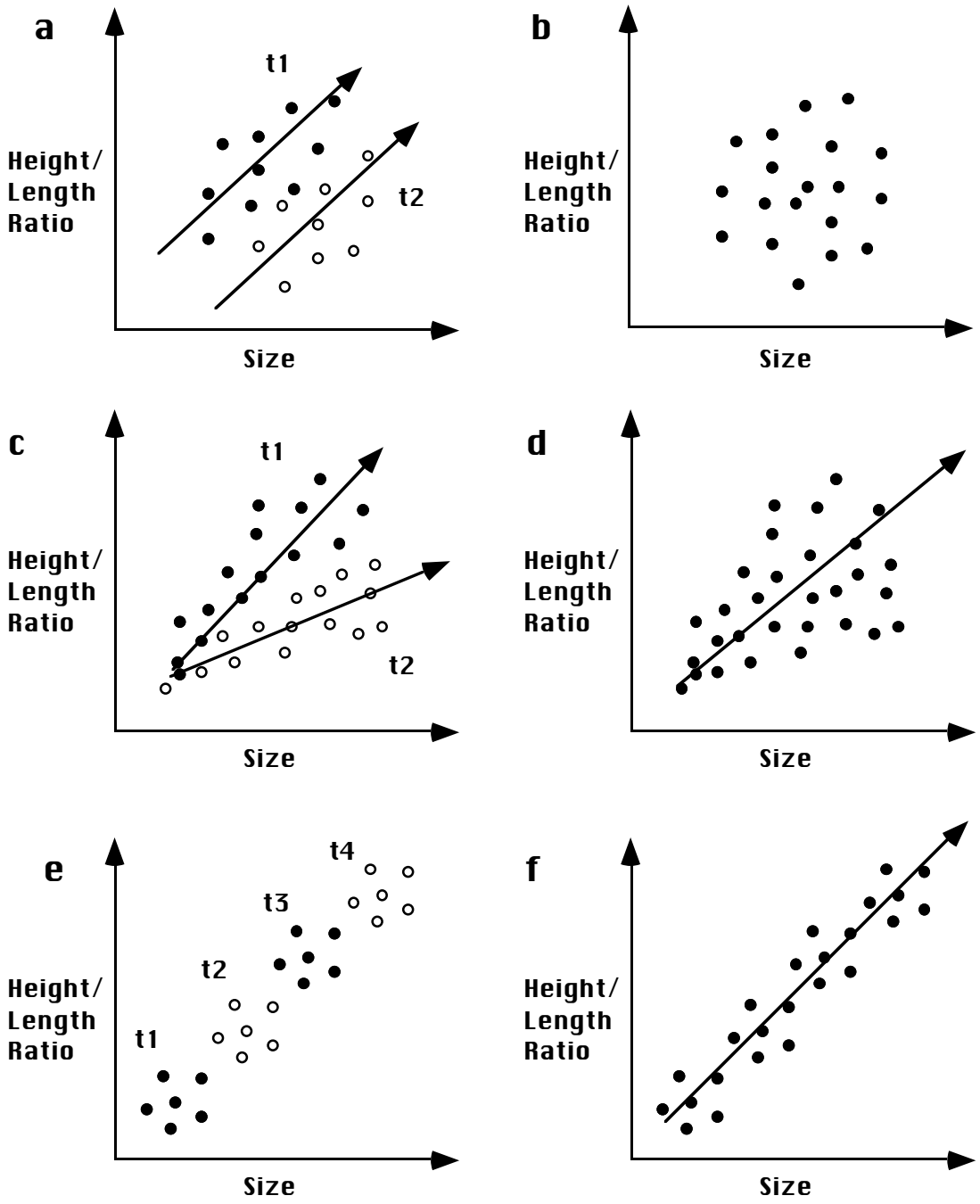

Figure 2.8. The potential effects of time-averaging on allometry. At times t1 and t2 in (a), there exists a definite allometric trajectory (shown with an arrow): the height to length ratio increases with size. When time-averaged (b), no allometry is detectable. In (c), the slope of the allometric trajectory changes from t1 to t2; allometry is still detectable after time-averaging (d), but the apparent relationship is much weaker. In (e), there is no allometry in any of a series of standing crop populations t1 to t4. Both size and shape evolve, however, and after time-averaging there is an apparent (but false) allometry.

therefore be used to infer paleoenvironmental condition. However, his method does not account for the effects of time-averaging. If a taxon grows allometrically, changing in shape during its growth history, then mixing of different size/age classes will introduce "ontogenetic" variation. The effect of ontogenetic averaging can be avoided by examining a single size class, or it can be removed statistically.

Various opinions have been advanced regarding the effect of evolutionary change on morphologic variability. Some (e.g., Bader 1955) have argued that variability is inversely correlated with evolutionary rate. Others (Williamson 1981a, 1981b, 1986) have argued that evolution, concentrated at speciation events, is accompanied by developmental instability that is manifested in an increase in phenotypic variability. Because time-averaging will preferentially increase the apparent variability of rapidly evolving populations, positive correlations between evolutionary rate and variability are questionable if based upon time-averaged fossil material.

A Previous Study

Bell et al. (1987) tested the effects of time-averaging on the morphologic variability of a single character in stickleback fish. They collected 10 samples of living Gasterosteus doryssus and counted the number of dorsal-fin rays on 100 specimens from each sample to determine the variance of standing-crop populations. They also determined the variance of this character in 27 samples of fossil Gasterosteus aculeatus, a similar stickleback. These fossils were collected from a varved Miocene lake deposit, and the temporal span of each collection was known to within 50 years.

The duration of time-averaging varied from 1 year to almost 3000 years, but no increase in variance was detected. When a regression was performed, the interval of time-averaging only explained 4% of the variance in the fossil samples; no fossil sample was more variable than the most variable recent

sample (though two were less variable than any recent sample). Bell et al. (1987) found it "surprising that not even time-averaging increased variance of the fossil samples." Though they discuss factors that may have obscured an increase, such as the uncertainty involved in estimating variance, they admit that "it is possible that there simply was not enough change through time in dorsal-fin ray number within the samples to detect an effect of pooling-time interval on sample variance."

Bell et al. (1987) demonstrate that time-averaging is not always an impediment to the study of variation in fossil organisms. The excellent time-resolution and multiple samples allow a well-controlled study. However, Bell et al. (1987) only examined a single character, and the variability of other anatomical features may well have been inflated by time-averaging. In this study, many variables are used to obtain a more complete measure of organismal variability.

Chapter 3: Mercenaria Samples

The Genus Mercenaria

Mercenaria Schumacher, 1817 are shallow infaunal, filter feeding, venerid bivalves (Fig. 3.1). Two species (M. mercenaria and M. campechiensis) currently inhabit the Atlantic coasts of Canada and the United States and the Gulf of Mexico coasts of the United States and Mexico, and one species (M. stimpsoni) is recognized from the Western Pacific. In addition, Harte (1992) reclassified Chione kellettii of the tropical east Pacific as Mercenaria kellettii. The genus dates to the Oligocene, when it may have evolved from the Pacific genus Securella (Harte 1998).

Along the Atlantic coast of the United States, modern Mercenaria mercenaria live in estuaries and inshore embayments, while M. campechiensis inhabit nearshore, open ocean environments from New Jersey southwards. Mercenaria campechiensis also extends into the shallow marine waters of the Gulf of Mexico. M. campechiensis has traditionally been distinguished from M. mercenaria by having a wider lunule, no purple nacre on the shell interior, more prominent comarginal ribs, and a heavier shell (Abbott 1974). Mercenaria from Texas lack the thicker ribs of M. campechiensis and have been classified variously as M. mercenaria texana and M. campechiensis texana; on the basis of morphology and genetics, Dillon and Manzi (1989) classify the Texas form as M. campechiensis texana.

For the most part, Mercenaria mercenaria and M. campechiensis do not co-occur, but both inhabit the Indian River Lagoon on Florida's Atlantic coast. Here, they produce hybrids that are fertile with both parent species. Hybrids suffer an increased occurrence of gonadal neoplasia, a tumorous disease (Bert et al. 1993; Eversole and Heffernan 1995), and they experience shorter and less intense spawning peaks than the pure species (Eversole 1997). Despite selection against hybrids throughout the lagoon, the hybrids seem to have advantages in

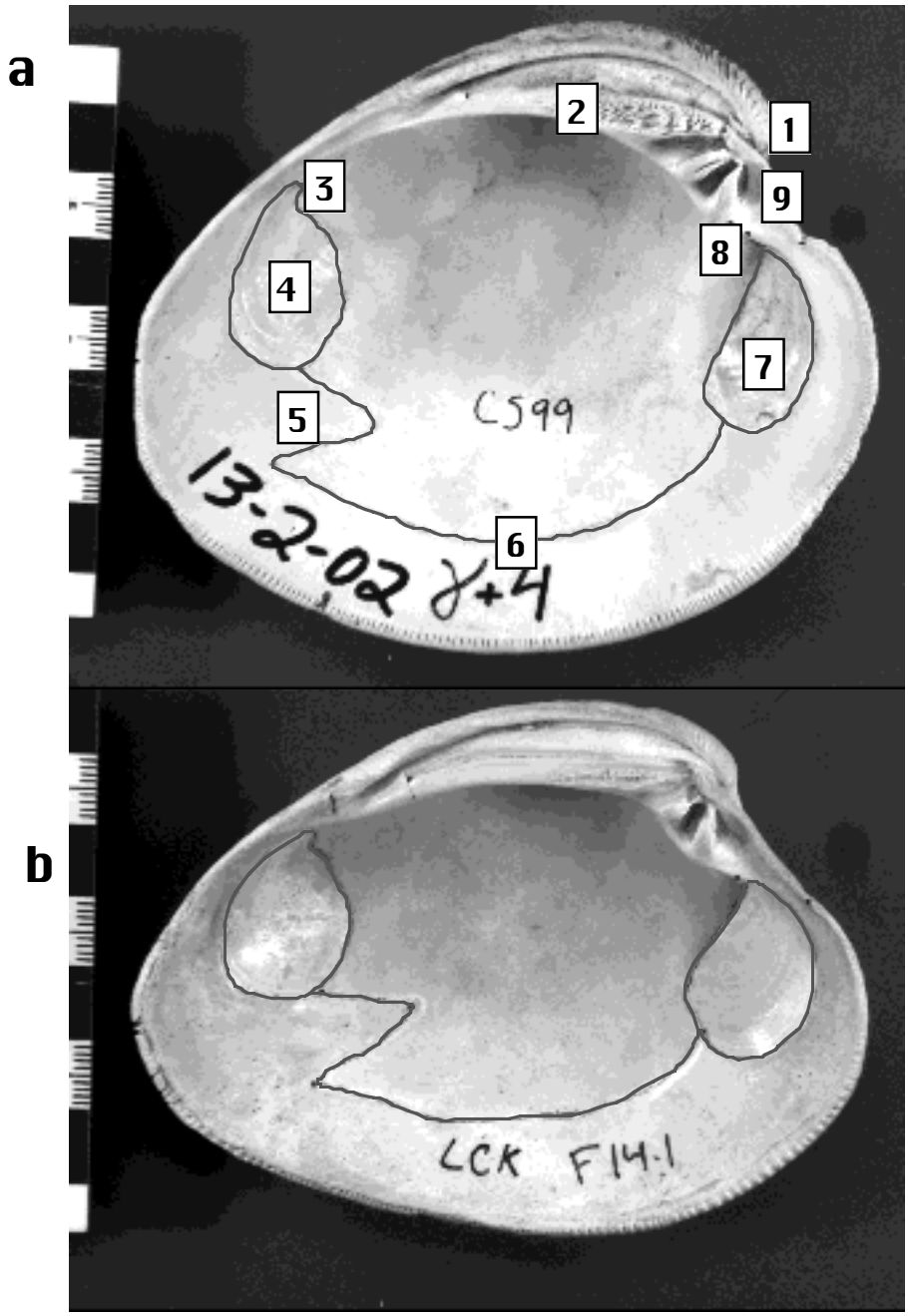

Figure 3.1. Interior surfaces of left valves of Mercenaria campechiensis (a) and M. permagna (b). The following features are numbered: (1) umbo, (2) dorsal hinge ligament, (3) posterior pedal retractor muscle scar, (4) posterior adductor muscle scar, (5) pallial sinus, (6) pallial line, (7) anterior adductor muscle scar, (8) anterior pedal retractor muscle scar, (9) lunule. Note the more elongate shape of M. permagna. Scales in cm and mm.

certain microhabitats (Bert and Arnold 1995; Arnold et al. 1996). Many species of fossil Mercenaria have been described. The only extinct species of importance to this study is M. permagna from the Lee Creek Mine in North Carolina (Fig. 3.1a,c). It has strong ribs like M. campechiensis, but it is more elongate, less inflated, and has a longer ligament and lunule.

Recent Samples

The six collections of modern Mercenaria campechiensis used in this study are housed at the Florida Marine Research Institute in St. Petersburg and were made available for study by William Arnold. Arnold and his associates collected several of the samples, and others were sent to him along with locality information. Collection sites are indicated on Fig. 3.2 and in Table 3.1, and the locality information is summarized in the text below.

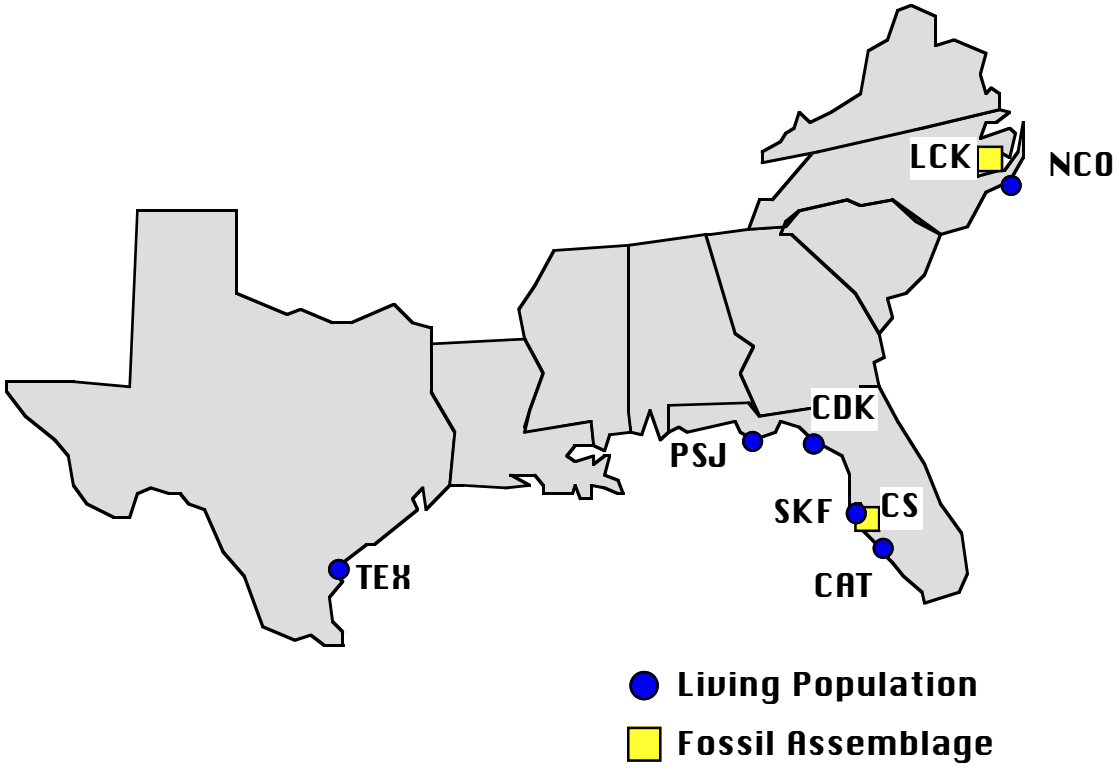

Figure 3.2. Collection localities of recent and fossil Mercenaria samples.

Table 3.1. Collection localities and dates of 6 recent populations of Mercenaria campechiensis. Sample size indicates the number of individuals measured for this study, which may be less than the number originally collected.

Sample	Location	Date	Sample Size
TEX	Corpus Christi Bay, TX	4/87	64
PSJ	St. Joseph Bay, FL	3/87	53
CDK	Cedar Key, FL	6/89—10/91	43
SKF	Tampa Bay, FL	10/86—6/89	44
CAT	Charlotte Harbor, FL	8/89	52
NCO	Cape Lookout, NC	6/87	52

The Corpus Christi, Texas sample (TEX) was collected by Teri Nelson on April 26, 1987 from Ingleside Cove on the north side of Corpus Christi Bay. The clams were living in a muddy bottom with scattered patches of Halodule wrightii; water depth was 0 to 10 cm at low tide. The sample from St. Joseph Bay, Florida (PSJ) was mechanically harvested from 26 feet of water on March 30, 1987. Harry C. Lawder arranged the collection and reported a substrate of 80% mud and 20% sand, with a top layer of broken shell and no vegetation. In June 1989, William Arnold collected a sample (CDK) from Suwanee Reef, just north of Cedar Key, Florida. An additional sample was taken in August/October 1991. The Tampa Bay, Florida sample (SKF) was collected by Arnold on October 22, 1986, June 19, 1987, March 25, 1988, and June 19, 1989. The M. campechiensis inhabit a sand flat near the Skyway Bridge south of St. Petersburg. A nearby oyster reef is inhabited by M. mercenaria. The Charlotte Harbor, Florida sample (CAT) was taken from Catfish Creek, a tidal creek on Charlotte Harbor, by Arnold in August 1989. A sample was taken from the open ocean off of Cape Lookout, North Carolina sample (NCO) in June 1987.

All samples except CDK and SKF were collected at a single site on a single day. CDK and SKF were collected at single sites over 3 to 4 year periods, but this should not introduce any additional variability into those samples. These samples contain large individuals that would have been alive at the same time, so the clams are still contemporaneous.

These 6 samples of Mercenaria campechiensis were previously used by Bert et al. (1993), Bert and Arnold (1995), and Arnold et al. (1996) to determine the range of genetic variability of the species at several loci. The same genetic analysis was performed on a number of populations of M. mercenaria, and these results were collectively used to genetically classify clams in the Indian River as M. campechiensis, M. mercenaria, or a hybrid of the two. Only the M. campechiensis samples from Texas, St. Joseph Bay, and Cedar Key were chosen to represent the genotype of that species; the other samples had some genetic outliers that were suspected as hybrids. The discriminant function used to classify the Indian River clams was also applied to the other samples, and the genetic classifications of the six M. campechiensis samples used in this study were supplied by William Arnold.

Previous studies that used these samples address a number of questions about the Mercenaria hybrid zone in the Indian River. Bert et al. (1993) found that hybrids from the Indian River showed higher rates of gonadal neoplasia than either pure species. Bert and Arnold (1995) used the genotyped clams to examine the selection regime in the hybrid zone and test which of two theories on hybrid zones applies to the Indian River. The tension-zone model predicts that hybrids are selected against across a hybrid zone, independent of environment. The ecotone model predicts that selection varies with environment, so that hybrids are as fit as the parental species in certain environments. They found that aspects of both models apply to the Indian River: selection does occur against all hybrids, but its severity varies with environment. Arnold et al. (1996) found that the growth rates of M. mercenaria, M. campechiensis, and their hybrids in the Indian River vary with environment; in different microhabitats, each of the three genotype classes displays a superior growth rate.

Fossil Samples

Two samples of fossil Mercenaria were collected for this study from Pleistocene deposits of the Atlantic and Gulf Coastal Plains (Fig. 3.2, Table 3.2).

Table 3.2. Collection localities and dates of the 2 fossil Mercenaria samples. Sample size indicates the number of individuals measured for this study, which may be less than the number collected.

Sample	Species	Location	Date	Sample Size
CS	Mercenaria campechiensis	Caloosa Shell Quarry Ruskin, FL	1/99	215
LCK	Mercenaria permagna	Lee Creek Mine Aurora, NC	1/97	98

Caloosa Shell, Florida

LOCATION:
3939 Cockroach Bay Rd.
Ruskin, FL 33570
813-645-3068

MAILING ADDRESS:
P.O. Box 7240
Sun City, FL 33586
Fax: 813-645-1696

Caloosa Shell Corporation, formerly Leisey Shell, is located in southwestern Hillsborough County, Florida. Tampa Bay is located less than a mile to the north, and Ruskin, FL lays nearby to the northeast. Three fossiliferous Pleistocene beds assigned to two formations are exposed at this quarry (Fig. 3.3): the "lower shell bed" and "bone bed" belong to the Bermont Formation, and the "upper shell bed" belongs to the Fort Thompson Formation (Hulbert and Morgan 1989, Portell et al. 1995). The quarry is best known for the Pleistocene vertebrate fauna derived from the bone bed (see papers in Hulbert et al. 1995a, 1995b), but Portell et al. (1992, 1995) analyzed the diverse invertebrate fauna, identifying 248 species, including 217 mollusks.

Erosional disconformities separate the lower shell bed from the bone bed and the bone bed from the upper shell bed (Jones et al. 1995). The age of the

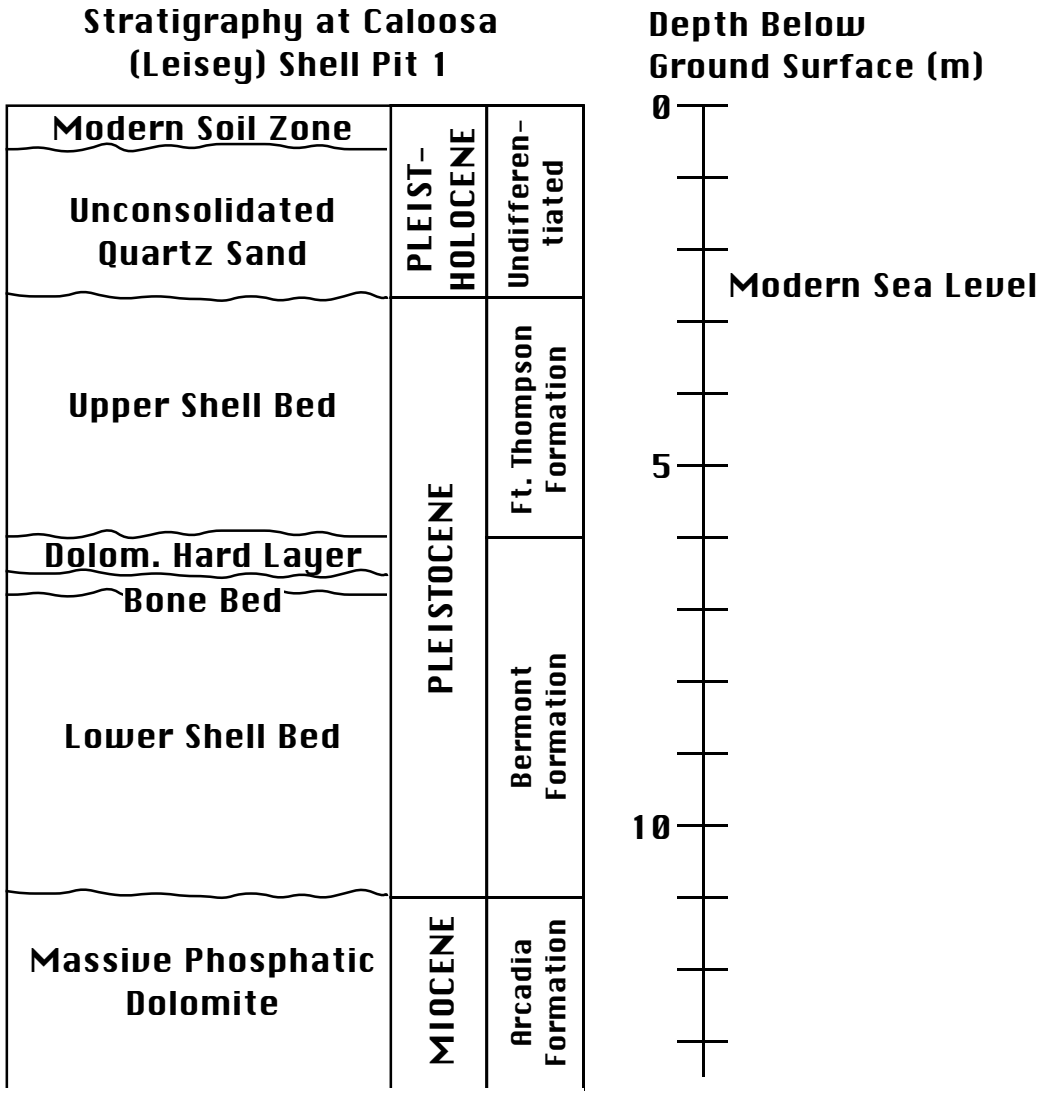

Figure 3.3. Stratigraphy at Caloosa (Leisey) Shell. No unit below the upper shell bed was exposed at the pit where the Mercenaria were collected; this section was measured at an earlier pit. See text for descriptions of the units. Modified from Portell et al. (1995).

bone bed has been a subject of debate. The vertebrate fauna has been assigned to the late early Irvingtonian Land Mammal Age, 1.5 to 1.0 Ma (Jones et al. 1995, Hulbert and Morgan 1989). However, the bone bed has been correlated with the Bermont Formation of southern Florida on the basis of marine mollusks (Hulbert and Morgan 1989, Webb et al. 1989, Portell et al. 1995); many have considered the Bermont to be less than .5 Ma (DuBar 1974, Blackwelder 1981). This opinion is not universal, however, because Lyons (1991) and Zullo and Harris (1992) argue

an early Pleistocene age for the Bermont. In any case, strontium isotope data (Webb et al. 1989, Jones et al. 1995) and paleomagnetic data (MacFadden 1995) support an early Pleistocene age for the bone bed at Caloosa Shell.

The age of the upper shell bed has been of less interest. Using strontium isotopes, Jones et al. (1995) suggest an age of early to middle Pleistocene, perhaps around 0.9 to 0.6 Ma. However, DuBar (1974) considered the Fort Thompson Formation to be .14 to .12 Ma. As with the bone bed, the age determined for the upper shell bed at Caloosa Shell does not match age estimates for the formation to which it is biostratigraphically correlated. The coastal plain stratigraphy of Florida deserves further study.

The lower shell bed (Bermont Fm.) is a bluish, massive shell bed with a matrix of fine quartz sand (Portell et al. 1995). About four meters thick, it unconformably overlies the Miocene Arcadia Formation. The unit is dominated by marine mollusks, with bivalves slightly more diverse than gastropods (79 marine and 5 nonmarine bivalves compared to 64 marine and 10 nonmarine gastropods). Articulated bivalves are in high abundance. Other invertebrates from the lower shell bed include 1 type of sponge, 2 corals, 14 bryozoans, 1 annelid, 3 arthropods, 1 echinoid, and 2 scaphopods. Dark, fine-grained, organic sediments lay at the base of the unit as irregular lenses or as a distinct layer. The organic sediments contain estuarine, freshwater, and terrestrial invertebrate remains, as well as vertebrates and plants (Portell et al. 1995). About 3% of marine mollusks in the lower shell bed are extinct. Portell et al. (1995) suggest that the basal, organic sediments containing non-marine fossils were deposited in a freshwater spring, a river, or a pond. The rest of the unit is interpreted as a soft-bottom, low energy environment in 3 to 6 m of water, perhaps in a protected embayment.

Overlying the lower shell bed, the bone bed (Bermont Fm.) contains most of the extensively-studied vertebrate fauna (e.g., papers in Hulbert et al. 1995a and 1995b). It is a 5 to 30 cm thick, unconsolidated mix of fine sand, silt, mud,

bones, and shells (Portell et al. 1995). Molluscan diversity is dominated by 93 species of gastropods, 10 of which are nonmarine. However, the 68 species of bivalves (all of which are marine) dominate the gastropods in abundance. Other taxa include 1 species of sponge, 2 corals, 8 bryozoans, 1 annelid, 3 arthropods, one echinoid, 2 scaphopods, 3 polyplacophorans, and plant material. Hulbert and Morgan (1989) and Portell et al. (1995) report that the unit is a mixture of transported terrestrial and freshwater organisms and locally-derived marine fossils. A calcareous marl (the "hard layer") overlies the bone bed in some parts of the quarry. Three percent of recognized marine mollusks are extinct. Portell et al. (1995) interpret the depositional environment as a soft-bottomed marine embayment or lagoon of 2 m or less water depth. Along with the well-preserved shells, the silts and clays suggest a low energy setting, and the mollusk fauna indicates the likely presence of seagrass. A large river emptying nearby could have imported the freshwater and terrestrial remains (Portell et al. 1995).

The upper shell bed (Ft. Thompson Fm.) is a sandy, buff, massive shell bed that ranges from 1.5 to 3 (or more) meters thick. The quartz sand contains little silt or clay, and the shells show a greater degree of fragmentation, abrasion, and disarticulation than the shells in the underlying units. Bivalves and gastropods have approximately the same diversity (51 to 53 respectively, including 5 nonmarine gastropod species), though Portell et al. (1992) note that bivalves greatly exceed gastropods in abundance. Ninety-nine percent of the marine mollusks are extant. The fauna also includes a sponge, a coral, 6 bryozoans, 5 arthropods, and an echinoid. The upper shell bed is overlain by undifferentiated, non-fossiliferous Pleistocene/Holocene sediment. The lack of fine sediments and the increase in taphonomic wear led Portell et al. (1995) to suggest that the upper shell bed was deposited in a higher energy environment than the underlying strata. The mollusks suggest water depths of no more than 2 m and the possible presence of seagrass. This unit may have been deposited in a bay, lagoon, tidal channel, or embayment inlet (Portell et al. 1995).

Several hundred valves of Mercenaria campechiensis were collected from the upper shell bed at Caloosa Shell in January 1999 with the assistance of Gwen Daley and her spacious Isuzu Rodeo. Every specimen that could be located (and reached) was collected through several meters of section at a vertical cut through the shell bed. The outcrop extended for several hundred feet; undulations in the floor of the pit allowed collection at different vertical levels of the cut over this distance. Mercenaria were evenly but widely distributed through the unit, though they were easier to locate in more weathered parts of the outcrop. Many hardy valves had weathered out and lay at the foot of the cut, but only *in situ* specimens were collected.

The outcrop was divided into lateral and vertical segments ("columns" and "strata"), and the location from which each shell was collected was recorded (Fig 3.4 and Table 3.3). The four lateral segments were demarcated by the locations of collections made by Gwen Daley for paleoecologic analysis. The vertical levels were 30 cm thick, horizontal layers numbered from 1 to 14 moving up section. The boundaries were sited along the outcrop with a Brunton compass. There was no evidence of stratification within the shell bed, so these strata are largely arbitrary. However, the formation boundaries are generally flat lying, and horizontal divisions are reasonable. Neighboring strata certainly overlap temporally; stratigraphic disorder (see Fig. 2.2) is likely, and the large valves of Mercenaria often straddled strata boundaries. These valves were assigned to the stratum that enclosed the majority of the valve, but shells virtually at the same level could be assigned to different strata. Still, the average shell age in each of the strata should decrease up section. The paleoecologic samples were taken at vertical levels corresponding to the Mercenaria intervals (Fig 3.4).

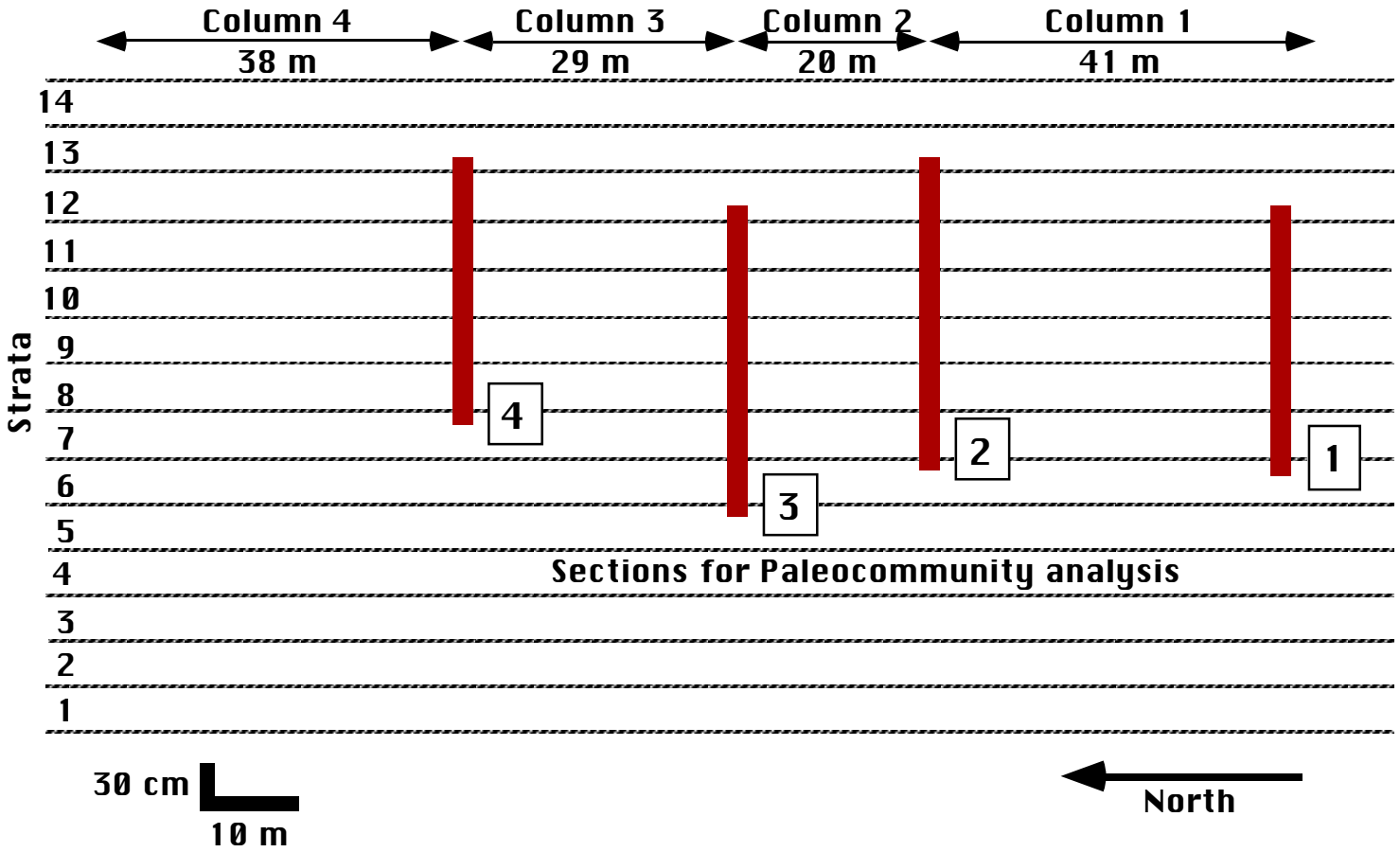

Figure 3.4. Divisions of the Caloosa Shell outcrop. The outcrop was divided into horizontal units that were 30 cm thick (strata). It was also divided into 4 lateral units (columns) bounded by four paleoecologic collection sites and the ends of the outcrop.

Table 3.3. Distribution of measurable Mercenaria campechiensis from the Caloosa Shell outcrop.

Strata	Column 4	Column 3	Column 2	Column 1
13			2	5
12			10	5
11	1	13	25	7
10	7	23	6	3
9	12	10		1
8	37	5	1	4
7	7	4	1	6
6	3	3		6
5		1		2
4				3
3				1
2				
1				1

About 80% of the individuals measured were encrusted or bored on the shell's inner surface, indicating some amount exposure on the sea floor after death. Traces of barnacles, bryozoans, sponges, annelids, and corals were common. Some altered shells showed only minor encrustation or boring, but some were so riddled with sponge borings that they crumbled when removed from the outcrop. Others were in such poor condition that they could not be measured. Of those that were encrusted or bored and could be measured, the average degree of alteration was moderate to high. Excluding the most degraded valves could potentially decrease the amount of time-averaging, but several studies in modern sedimentary environments have found that shell degradation is poorly related to shell age (Flessa et al. 1993; Meldahl et al. 1997).

Lee Creek Mine, North Carolina

PCS Phosphate
Highway 306 N
Aurora, NC 27806
252-322-4111

The Lee Creek Mine, located near Aurora in eastern North Carolina, exposes strata ranging from the lower Miocene to the Pleistocene (Fig. 3.5) (Ward and Blackwelder 1987). The lower Pleistocene James City Formation, the uppermost fossiliferous unit at Lee Creek, is about 2.5 to 3.5 m thick at Lee Creek. It sits unconformably upon the upper Pliocene Chowan River formation and is unconformably overlain by non-fossiliferous beds tentatively assigned to the Flanner Beach Formation by Ward and Blackwelder (1987).

Plio-pleistocene stratigraphic units in North Carolina and Florida have been correlated in various ways. Ward and Blackwelder (1987) correlate the James City with the upper Caloosahatchee in Florida, but Zullo and Harris (1992) correlate it with the Bermont, which they believe to be lower Pleistocene. They would equate the entire Caloosahatchee with the Chowan River. Ward and Blackwelder (1987) correlate the Flanner Beach with the Fort Thompson (Fig. 3.5).

Ward and Blackwelder (1987) studied the depositional environments and molluscan fauna of the James City and Chowan River; the following summary is derived from their work. The James City at Lee Creek has been divided into three units: C (lowest), D, and E (highest). Collections from six measured sections, most of which did not pass through all three units, yielded 77 species of bivalve, 91 gastropods, and no other mollusks. The fauna indicates subtropical conditions, though some warm-temperate species are also present.

Unit C appears locally where the underlying Chowan River has been channeled, and is a "greenish gray iron-stained clayey fine sand containing abundant large mollusks" (Ward and Blackwelder 1987). It is interpreted as a shallow shelf environment in 15 to 20 m of water. Unit D is a fine to coarse light

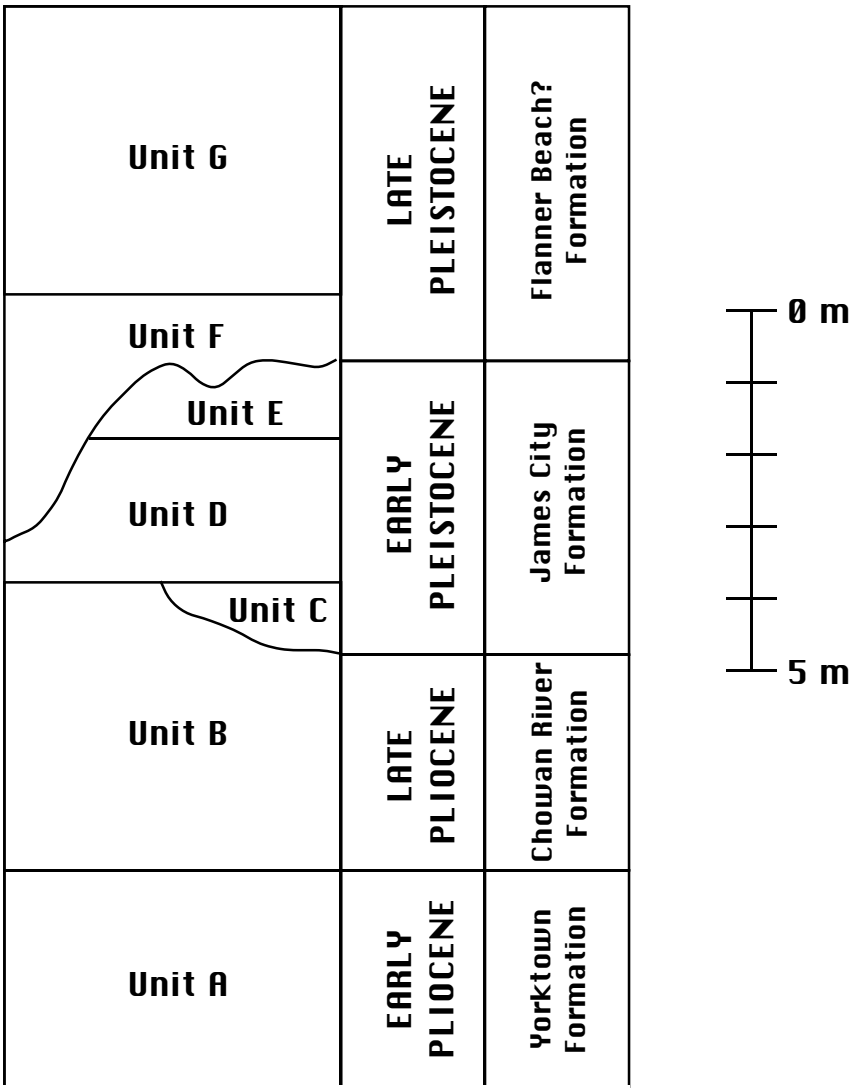

Figure 3.5. Generalized stratigraphy of the Pliocene and Pleistocene units at Lee Creek Mine. See text for descriptions of units. Modified from Ward and Blackwelder (1987).

gray sand that also contains many mollusks. On the north side of the pit (as it existed in 1972-73), there are unidirectional, high angle (34°) current beds, and unit D is horizontally bedded elsewhere. It is interpreted as part of an offshore bar system, with the main portion of the bar to the east of the pit. Deposition is thought to be rapid. The light gray, clayey, medium fine sands of Unit E contain many small mollusk shells and are thought to have been deposited in a lower energy, shallower environment than D. There may have been a bar further offshore at this time (Ward and Blackwelder 1987).

Over 100 Mercenaria permagna were collected from the James City Formation at Lee Creek in January, 1997 with assistance from Brian Coffey. The clams were concentrated in a zone 30 to 40 cm thick and were randomly oriented. Articulated pairs as well as single valves were collected. The outer surface of many specimens suffered from dissolution, but the inner surfaces were generally well preserved. About half of the shells were bored or encrusted to some extent on their inner surface. The amount of encrustation or boring ranged from little to the total coverage of the interior surface, but was typically fairly low. Three specimens of M. mercenaria were found with the M. permagna; this classification is based upon the smooth exterior and more triangular shape of these shells. These three shells are morphologically indistinguishable from modern M. mercenaria and were excluded from all analyses.

Chapter 4: Methods

Data Collection

A digital image of each clam was captured using a Watech WAT-902 camera attached to a Macintosh Quadra 660AV. Each shell was oriented so that a line connecting the anterior and posterior retractor landmarks (landmarks 8 and 3, Fig. 4.1, Table 4.1) was horizontal in the image, and the camera was oriented perpendicular to the plane of commissure of the shell. A scale bar was placed next to the shell, and the distance between the camera and the shell was varied so that each shell appeared to be the same size in the image. The amount of error introduced by digitizing the coordinates was thus held constant regardless of shell size.

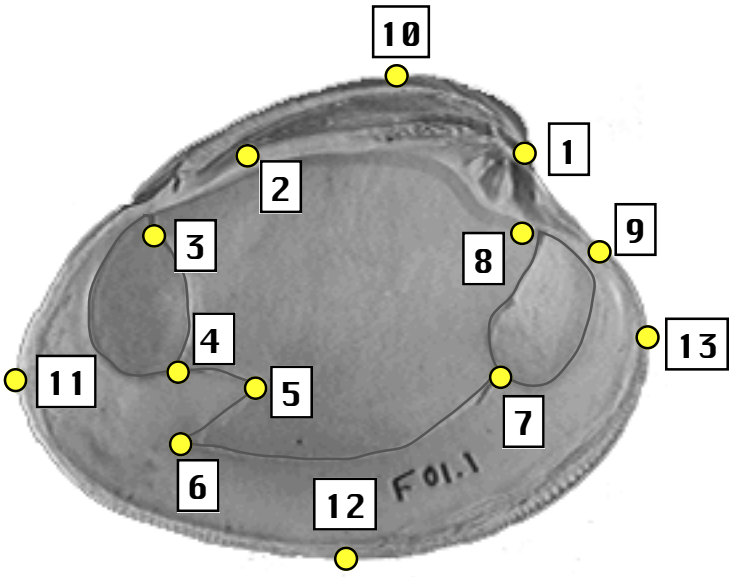

Figure 4.1. The 9 landmarks (1-9) and 4 pseudolandmarks (10-13) measured on each Mercenaria valve. See Table 4.1 for landmark descriptions and Fig. 3.1 for the anatomical features of the valve.

Any valve that was damaged in a way that obscured the landmarks was not imaged for measurement. One recent sample (PSJ) consisted solely of right valves, but all other recent samples consisted of right and left valves. Fossil samples consisted of both left and right valves, some of which comprised articulated pairs. In samples that contained left and right valves, some of each

were measured. All features measured were presumed symmetric between the valves (this assumption is tested in Chapter 5). When both valves were available from a single individual and both were measurable, one was selected randomly for measurement. Images of right valves were reflected before measurement so that the data would be comparable to the left valves.

Table 4.1. The landmarks (1—9) and pseudolandmarks (10—13) measured on each Mercenaria valve.

Number	Abbreviation	Description
1	UMBO	Umbo
2	LIG	Posterior end of the dorsal hinge ligament
3	P_RET	Junction of the posterior pedal retractor muscle scar with the posterior adductor muscle scar.
4	P_ADD	Junction of the pallial line with the posterior adductor muscle scar
5	PS_IN	Maximum of curvature of the pallial line inside the pallial sinus
6	PS_OUT	Maximum of curvature of the pallial line outside the pallial sinus
7	A_ADD	Junction of the pallial line with the anterior adductor muscle scar
8	A_RET	Anterior pedal retractor scar
9	LUN	End of the lunule
10	DOR	Dorsal-most point on the shell
11	POST	Posterior end of the shell
12	VENT	Ventral-most point on the shell
13	ANT	Anterior end of the shell

The x and y coordinates of 13 landmarks and pseudolandmarks were recorded for each Mercenaria valve using NIH Image 1.55. A landmark is a specific anatomical point that is homologous between all organisms in a study, and pseudolandmarks are points that are treated as landmarks even though

biological homology is less well established (see "Types of Landmarks," below). These landmarks are listed in Table 4.1 and are illustrated in Fig. 4.1. Landmarks 1 through 9 are true landmarks, while 10 through 13 are pseudolandmarks. Certain landmarks were highlighted with a small dot of ink before imaging. The data for each sample were assembled in Microsoft Excel.

Analytic Methods

Variability can be construed as an average difference between individuals in a set, or as the maximum difference between those individuals, but either definition requires some measure of "difference" between entities. In a typical quantitative analysis of difference, a number of measured variables are used to construct a simulated, variously-dimensioned space where the measured properties are assigned orthogonal axes (Bookstein 1991, p. 55). Each entity, be it a physical object, a personality, a demographic group, or whatever, can be plotted as a point in this "feature space" (Rohlf 1990) according to its measured values. Patterns of variability are examined by examining this distribution, and the difference or distance between the entities can be measured as the Euclidean distance between these points (Bookstein 1991). In traditional analyses of organic form, a variety of lengths, areas, angles, and meristic characters are used to construct a feature space. Traditional univariate statistics, such as mean and variance, can be computed, and the data can be analyzed with traditional multivariate procedures such as principal component analysis and cluster analysis.

More recently, techniques have been developed to analyze data in the form of 2- or 3-dimensional coordinate data. This sub-field of biostatistics has been termed "geometric morphometrics." Coordinates based on landmarks, points biologically homologous between all objects in a study, can be analyzed by a number of methods including Procrustes superposition and relative warp analysis. The shape space that underlies landmark methods is theoretically unlike the traditional feature space outlined above, and deserves some

description. More complete reviews of landmark theory and methods are given in Dryden and Mardia (1998), Bookstein (1991), Marcus et al. (1996), Rohlf and Marcus (1993), and Rohlf and Bookstein (1990).

Types of Landmarks

A landmark is a point that is homologous between all organisms in a study. Bookstein (1991) classifies landmarks into three types of varying biological import. Type I landmarks, for which homology is most definite, are those that involve "discrete juxtaposition of tissues." Points where three structures meet and centers of small, distinct patches of tissue, such as the vertebrate eye, are Type I landmarks. Type II landmarks are maxima of curvature, like the tip of a claw or the valley of an invagination. Though they may be functionally equivalent between individuals, they provide weaker evidence for biological homology than do Type I landmarks. Type III landmarks are referred to as deficient landmarks or pseudolandmarks because they are meaningful only in a single dimension. They include extremal points on an outline and endpoints of diameters through a structure (Bookstein 1991).

Landmark Coordinates: Size and Shape

The theory underlying the analysis of landmark data is complex, mathematical, and difficult to visualize. A comprehensive understanding requires knowledge of statistics and geometry that can only be acquired from the original literature on geometric morphometrics. The references provided in this chapter are useful starting points.

The standard size measure in landmark studies is Centroid Size, which is the square root of the summed squared distances from each landmark on an organism to their common centroid (Fig. 4.2). Other possible size metrics include the distance between two specified landmarks or the area of the convex hull of the configuration, but centroid size is more commonly used in geometric morphometrics (Dryden and Mardia 1998, p. 24).

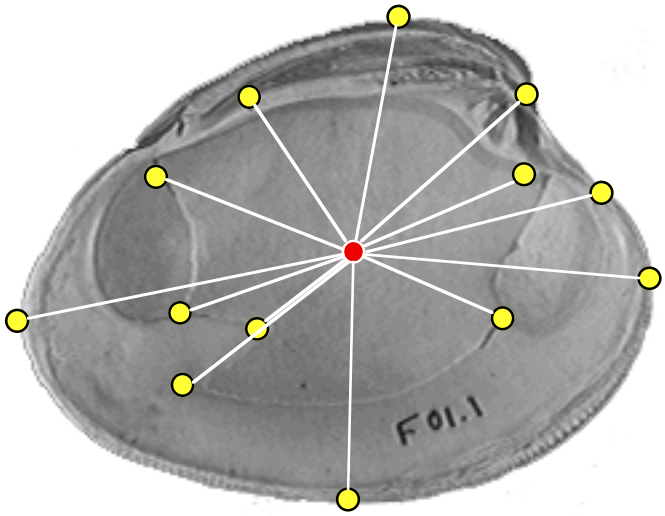

Figure 4.2. Centroid Size is the square root of the sum of squared distances from the landmarks to their common centroid.

Two configurations of landmarks have the same shape if they differ only in location, size, and rotation (Goodall 1991). In a proper "shape space" describing the shape of a set of landmarks, all configurations that differ only by translation, scaling, or rotation should plot as coincident points. The space shape used in landmark analyses has an additional constraint. If a single landmark in a configuration is moved a specified distance, the distance in shape space between the original and modified configurations will be the same regardless of the direction in which the landmark was moved (Bookstein 1996). This shape space is called Kendall's shape space, and was developed by Kendall (1984, 1986).

Kendall's space shape provides an *a priori* space in which coordinate data can be ordered. It differs from other feature spaces in that it does not consist of perpendicular axes along which the data are plotted. For example, the space shape for 3 landmarks in 2 dimensions (k=3, m=2, i.e., triangles) is a sphere in 3 dimensions with a radius of 1/2; any triangle of points corresponds to a point on the surface of the sphere (Dryden and Mardia 1998, p. 36). For more than three coordinates, the shape space is a complex projective space upon which each configuration of coordinates is plotted as a point (Dryden and Mardia 1998, p. 59).

"Pre-shape" space is also useful in discussing coordinate data. The pre-shape of a configuration is independent of its location and scaling, but rotation. If a pre-shape is rotated, the result is a different pre-shape, but a rotated shape is still the same shape. For k landmarks in m dimensions, the pre-shape space is a hypersphere of unit radius in (k-1)m dimensions (Dryden and Mardia 1998, p. 55). Every pre-shape (i.e., every separate rotation of a shape) plots as a different point on the surface of the pre-shape sphere; pre-shapes that differ only by a rotation fall along a line on the surface, a fiber (Fig. 4.3a). Each fiber in pre-shape space corresponds to a point in shape space (Dryden and Mardia 1998, p. 56). It is easier to discuss various landmark distance measurements with reference to pre-shape space since a sphere is easier to visualize than a complex projective space.

There are several ways in which distance can be measured between two configurations in shape/pre-shape space (Dryden and Mardia 1998, p. 61). For all these distances, the two pre-shapes are rotated until they are at the points on their fibers that are closest together (Fig. 4.3a). The Procrustes distance ρ is the shortest distance between two points along the curved surface of the pre-shape space (Fig. 4.3b). The partial Procrustes distance d_P is the length of a chord drawn between the two points, and the full Procrustes distance d_F is the minimum distance between one point and the radius connecting the other to the center of the pre-shape space (Fig. 4.3b). When two configurations are very similar, these three distances will be nearly identical. As configurations become more dissimilar and plot further apart, the 3 measures will diverge. When shape variability is small in a sample, as is commonly the case in biological studies, it makes little difference which distance is used. Total shape variability is the root mean squared (RMS) distance of the points to the average shape (Dryden and Mardia 1998, p. 46).

Tangent Spaces and Coordinates

Because landmark shapes are plotted in a curved, non-Euclidean space, normal statistical procedure cannot be used directly. First, the shapes must be

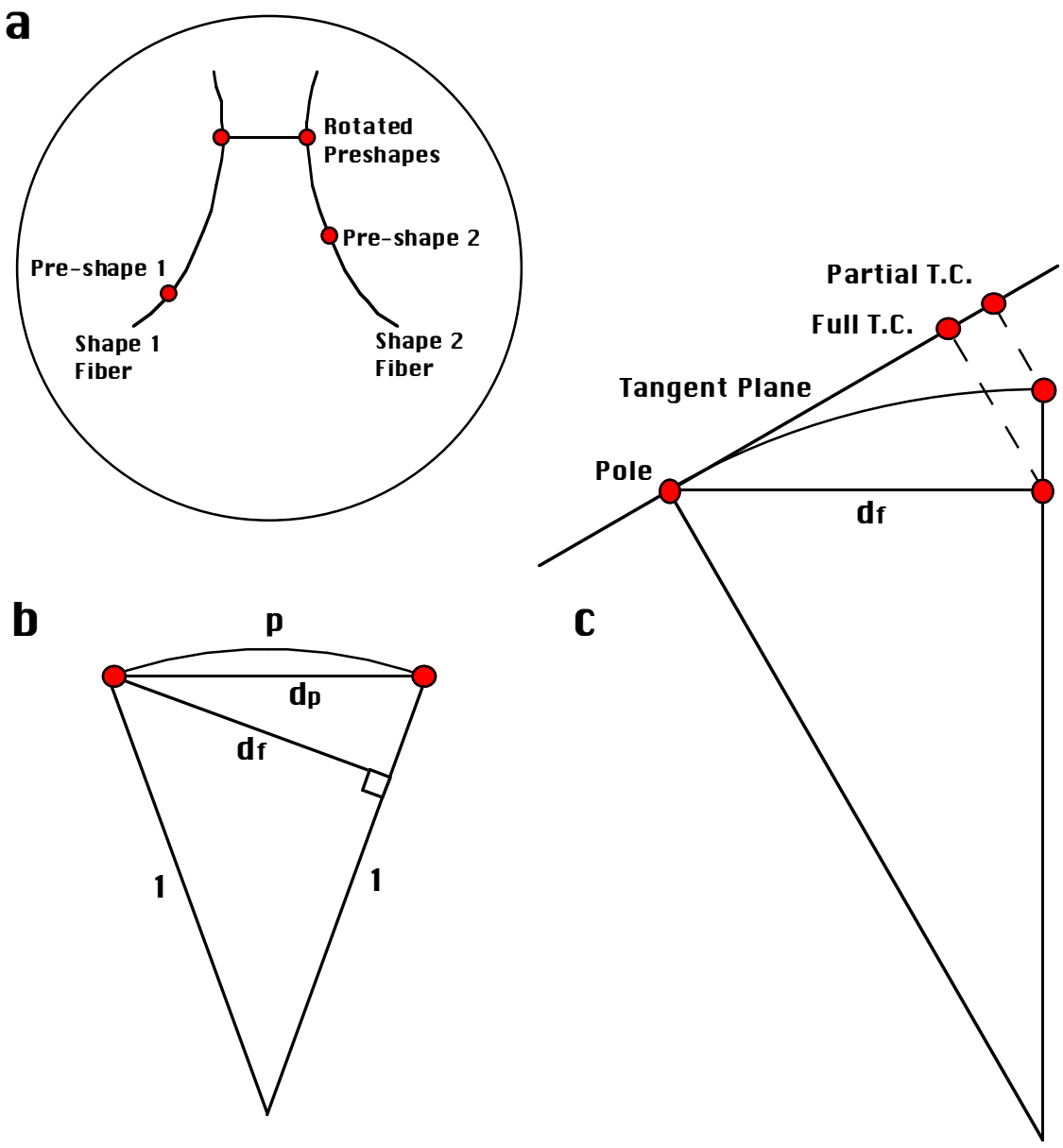

Figure 4.3. (a) Simplistic representation of the pre-shape sphere. The various rotation of a single shape lay along fibers. Preshapes are rotated to be as close as possible. (b) Cross-section of the pre-shape sphere, showing the three distance measures between two configurations of landmarks. (c) Cross-section of the pre-shape sphere showing a projection into a tangent space, and the difference between full and partial tangent coordinates. Note that the distance from the partial tangent coordinate to the pole (mean) equals the full Procrustes distance (d_f). All figures modified from Dryden and Mardia (1998).

transferred into a tangent space, a linearized vector space (Dryden and Mardia 1998). A tangent space is constructed by finding the mean configuration and using it as a "pole." At this pole, the tangent space touches the shape space. In pre-shape space, the observations are rotated to be as close to the pole as possible, then projected into the tangent space, yielding tangent coordinates (Fig. 4.3c). The distances between observations in the tangent space are very nearly the distances (either ρ, d_F, or d_P) in the shape/pre-shape space, as long as the observations are near the mean (Dryden and Mardia 1998). If a new set of observations with a new mean is used, a new tangent space should be constructed. Tangent coordinates can be entered into a principal components analysis or other multivariate statistical procedures, so they are very useful for practical applications.

Procrustes Analysis

Procrustes superposition analysis is a useful technique for measuring the Procrustes distances (ρ, d_F, or d_P) between configurations, for constructing tangent spaces, and for producing graphical displays of shape variability. Procrustes analysis superimposes one configuration of coordinates upon another, matching them as closely as possible according to some criterion. Least-squares Procrustes, which was used in this study, performs the match by minimizing the sum of the squared distances between corresponding landmarks. Resistant-fit Procrustes is useful in examining shape variability when the differences between configurations are localized at only a few landmarks. Resistant-fit Procrustes has been used in many studies (e.g., Hughes and Chapman 1995), but it does not conform to the shape space described above and Bookstein (1996) advises against using the residuals in further statistical manipulations. Procrustes analysis can also remove the affine or uniform component of shape variation (simple stretching or shearing), but this is not necessary when a global measure of variability is needed.

Dryden and Mardia (1998) distinguish between partial and full Procrustes superimposition. In partial Procrustes analysis, one configuration is translated

and rotated to optimally match another, but it is not scaled. In full Procrustes analysis, the shape is rescaled using Centroid Size to further reduce the summed squared distances between landmarks. Dryden and Mardia (1998) also distinguish between ordinary and generalized Procrustes analysis (OPA and GPA). OPA superimposes one landmark configuration on another by translation, rotation, and possibly scaling. GPA takes two or more shapes, finds the average, and optimally matches every configuration in the sample to that average. The average shape (or full Procrustes mean shape) is an average shape of unit centroid size. It is calculated such that the sum of squared full Procrustes distances from all configurations to the mean is minimized. GPA can be performed using either partial or full Procrustes matching. In partial GPA, all configurations are scaled to unit centroid size and optimally matched. In full GPA, the configurations are rescaled away from unit centroid size to achieve a closer fit. The mean configuration will still have unit centroid size, but the observed configurations will be slightly smaller (Bookstein 1996).

Several types of tangent coordinates can be generated by Procrustes superimposition. Fig. 4.3c illustrates the difference between partial and full tangent coordinates. The distance of a point in partial Procrustes tangent space from the mean is exactly equal to the full Procrustes distance between the original configuration and the mean (oddly enough). The residuals from a Procrustes superimposition are the deviations of each observation from the mean at each landmark. They are approximate tangent coordinates, and are frequently used in morphometric analyses (Dryden and Mardia 1998). If the variability in a data set is slight, then the choice of tangent coordinate matters little.

The amount of variability in a data set can be visually assessed by examining the spread of the landmarks after superimposition. The root mean squared Procrustes distance will increase as the spread increases, so tight clusters of points indicate lack of variability. If the spread of points at a particular landmark is non-circular, then it is likely that some factor (such as allometry) is causing shape change.

Procrustes superimpositions were performed by a computer program written in SAS IML by myself (Appendix B). The program calculates the full Procrustes mean shape, performs a full generalized least-squares Procrustes superimposition, and calculates the results, residuals, and partial tangent coordinates. From the tangent coordinates and residuals, the program calculates the variabilities of the samples. The total variability of each sample was calculated as the square root of the mean squared distance of each data point to the average shape, which for the tangent coordinates is equal to the RMS full Procrustes distance (Dryden and Mardia 1998, p. 72). Because ontogenetic averaging can inflate the variability of a sample, the effects of allometry were removed by regressing each tangent coordinate (or residual) variable on centroid size and retaining the residuals. (Linear regression fit these data well, see Chapter 5.) The root mean squared (RMS) distance to the mean was calculated again after the allometric correction to determine "allometry-free" variability. This measure is intended to make the different size ranges of the samples irrelevant to the observed variability.

Principal Component Analysis

Principal component analysis (PCA) is a commonly-used multivariate technique that can be applied to a data set consisting of multiple measurements on objects in a sample. It creates new variables as linear combinations of the original variables; the data values can be plotted along these new axes, essentially rotating the cloud of data points in multidimensional space. The original variables will often be heavily intercorrelated; the rotation yields uncorrelated variables. PCA can be performed using either the covariances or the correlations between the variables. Using the covariances leaves distances between points unchanged, but using the correlations is useful when the measurements are not commensurate in scale. All PCA's performed in this study used the covariances.

When PCA is performed on k variables, k new variables (the principal components) are created. The variances of the principal components are the eigenvalues of the covariance or correlation matrix; the first eigenvalue is the greatest and the k^{th} is the smallest. Each eigenvalue corresponds to an eigenvector which describes the linear relationship between that principal component and the original variables. The first principal component is the linear combination of variables that explains the maximum possible proportion of the variance in the data set. The second principal component is the linear combination orthogonal to the first that explains the maximum remaining variance, and so on. The principal component scores are the values of the observations on the new principal component axes. They can be plotted just like ordinary variables. PCA can be used to reduce the dimensionality of a data set and to explore its covariance or correlation structure.

PCA was applied to the tangent coordinates of various Procrustes superimpositions to explore the distribution of Mercenaria in morphospace. When PCA is applied to the residuals or tangent coordinates of a Procrustes superimposition, the last 4 eigenvalues equal zero (within calculation error) due to linear dependence introduced by the superimposition (Bookstein 1996). These zero eigenvalues are mathematical artifacts and should not be interpreted as meaningful in the data analysis.

Canonical Variate Analysis

Canonical variate analysis, sometimes called canonical discriminant analysis, is another common multivariate statistical technique. Like PCA, it produces new, uncorrelated variables that are linear combinations of the original variables (Marcus 1990). The data must be arranged in *a priori* groups, and the canonical variate analysis (CVA) constructs new variables that show the maximum separation between group centroids. The first new variable, or canonical variate, shows the maximum separation between groups, and separation decreases with each canonical variate. The number of canonical variates produced equals the number of groups minus one or the number of

original variables, whichever is less. There is an eigenvalue to match each canonical variate; the eigenvalues indicate the amount of group separation described by the axes. The data are rotated and rescaled in a CVA, and this can alter distance relationships between points. The distance measure between points in the canonical variate space is Mahalanobis distance, D, often reported as D^2.

A canonical variate analysis can only be run on a data set in which the variables are linearly independent. Because the Procrustes superimposition introduces dependence, the tangent coordinates were first entered into a PCA and the last four principal components were dropped. The remaining principal components are linearly independent. PCA using the covariance matrix only rotates the data, without changing the relationships between the observations.

Bootstrapping

Bootstrapping is a set of techniques useful in estimating parameters of a data set when the underlying statistical distribution is unknown or poorly understood. The data set is treated as a population and is randomly sampled with replacement many times; the parameter of interest is calculated from each new random sample. The resulting probability distribution can be used to estimate a confidence interval or perform a hypothesis test.

Confidence intervals were obtained for variability estimates using a bootstrapping routine written in SAS IML by M. J. Kowalewski and myself (Appendix B). Each sample was resampled with replacement 10,000 times, and for each iteration a Procrustes superimposition was performed and the variability metrics calculated. From the resulting distribution, the 99% and 95% confidence intervals were estimated using the 0.5, 2.5, 97.5, and 99.5 percentiles. The estimates of these CI values were calculated after every 100 iterations so that the stability of the estimates could be gauged. The means of the bootstrapped distributions were lower than the estimates calculated from the original sample by 0.2% to 2.5%. This problem is often encountered when the bootstrap is

applied to parameters not distributed normally (Kowalewski et al. 1998). To correct for this bias, the difference between the bootstrapped estimate and the real estimate was added to all bootstrapped values. More advanced methods have been developed to more accurately make this correction (DiCiccio and Romano 1988), but the error introduced is minimal and does not affect the results.

Summary

All Mercenaria valves used in this study were digitally imaged, and the x,y coordinates of 13 landmarks and pseudolandmarks were recorded. Generalized least-squares Procrustes analysis was used to superimpose the landmark configurations in each sample and to generate partial tangent coordinates and Procrustes residuals. From the tangent coordinates and the residuals, variability was calculated as the root mean squared distance of each observation to the sample mean. Variability was calculated again after regression on centroid size to remove the effects of allometry. Confidence intervals around the variability metrics were estimated using bootstrapping. Exploratory multivariate methods (PCA and CVA) were applied to the partial tangent coordinates and Procrustes residuals.

Chapter 5: Results

Analysis of Variability: Living Populations and Single Fossil Beds

The Caloosa Shell sample was split into 6 subsamples based on the stratigraphic level of collection (Table 5.1); most subsamples are restricted to a single 30 cm layer, but strata 6 and 7 and strata 12 and 13 were combined to increase sample sizes. In all analyses, strata 6/7 and 12/13 are treated as single layers. Due to limited exposure, only 8 specimens were collected from strata 1 through 5, so these were excluded from all variability calculations. The resulting CS sample contains 207 specimens spanning 2.4 m of section.

Table 5.1. The number of Mercenaria in the 6 stratal subsample of CS. Each single stratum spans 30 cm of section; combined strata span 60 cm. See Table 3.3 and Fig. 3.4.

Strata	Total	Column 4	Column 3	Column 2	Column 1
12/13	22			12	10
11	46	1	13	25	7
10	39	7	23	6	3
9	23	12	10		1
8	47	37	5	1	4
6/7	30	10	7	1	12

Procrustes Superimpositions: Introduction

Generalized Procrustes analysis superimposes two or more sets of landmarks, minimizing the sum of the squared deviations between the configurations and the mean. The Procrustes superimposition of the 64 Mercenaria in the CAT sample is shown in Fig. 5.1a, with the outline of a valve sketched in for reference. The image is rotated so that the mean values for landmarks 3 and 8 have the same y-coordinate. There is obvious scatter at each of the landmarks, indicating either shape variability, measurement error, or both. To assess the importance of measurement error, a single valve (CAT 14) was

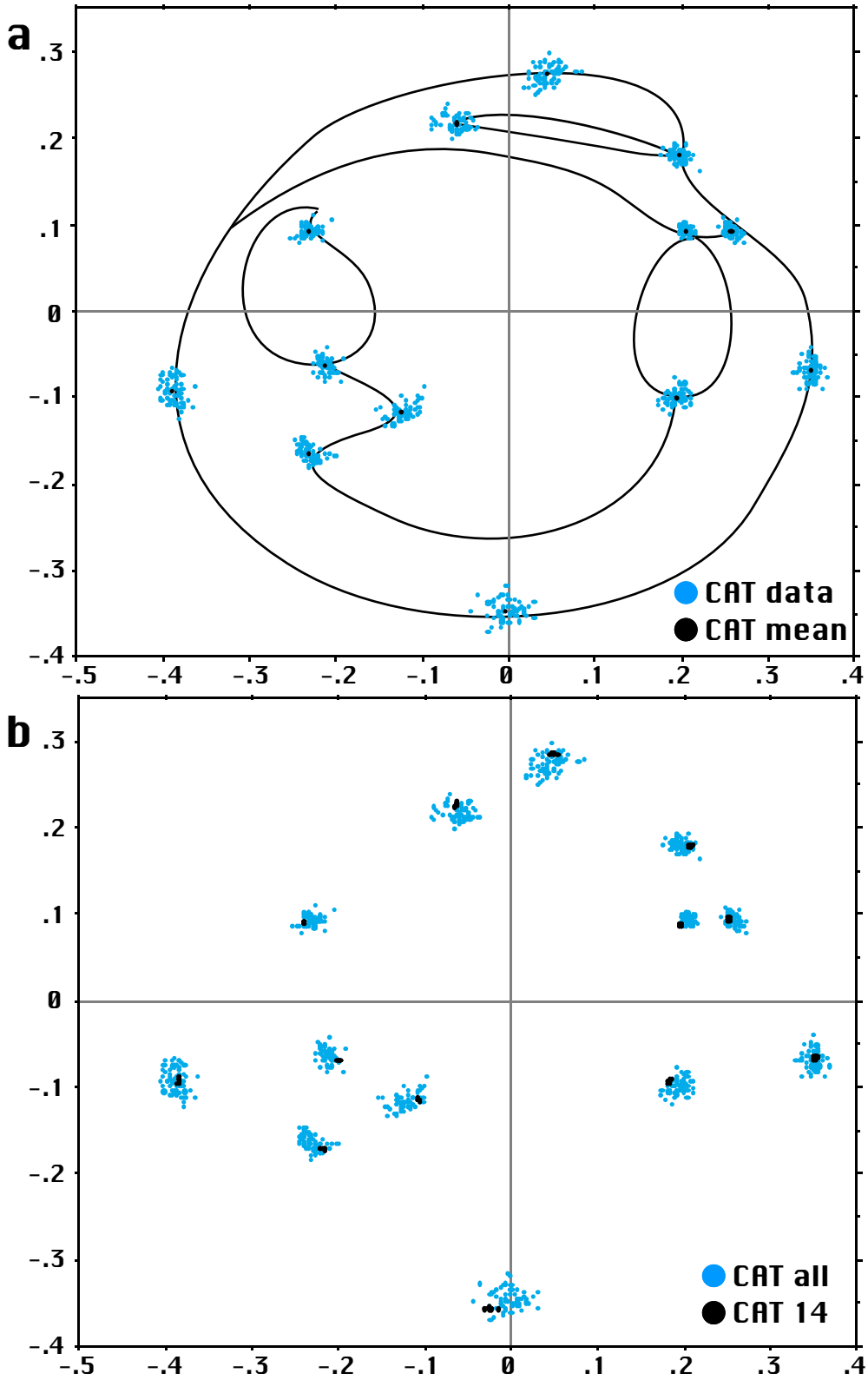

Figure 5.1. (a) Full Procrustes fit of the CAT sample, with a clam valve sketched for reference. (b) CAT plus 10 replicate measurements of CAT 14 showing digitizing error.

imaged and measured 10 times. These replicates were combined with the full CAT sample and another Procrustes fit was performed (Fig. 5.1b). The replicates of CAT 14 form tight, overlapping clusters at each landmark, and the spread is minor compared to the variability in the full sample. Measurement error is slight, and true shape variability is being observed in CAT. For many landmarks, CAT 14 plots near the edge of the distribution, indicating that its shape is relatively atypical for this sample.

Total Variability

The full generalized Procrustes superimpositions of the 6 recent samples and the 2 fossil samples are shown in Figs. 5.2—5.4. Total shape variability can be roughly gauged by examining the amount of spread in the Procrustes fit. The CS samples appear to have greater spread than the recent samples. Total variability was calculated as the root mean squared distance of the observations to the sample mean; the total variabilities of the recent samples and single fossil beds (LCK and the CS strata) are listed in Table 5.2. Variability was calculated from both the partial tangent coordinates and the residuals; for the tangent coordinates, the result is equal to the root mean squared full Procrustes distance from each observation to the mean. The results obtained from the tangent coordinates and the residuals are nearly identical (some values are different by .0001), as expected for a data set with little variability. All further discussions will focus on the tangent coordinates, though the residuals yield the same results.

The total variability of each sample (as calculated from the partial tangent coordinates) is shown in Fig. 5.5 with bootstrapped 95% and 99% confidence intervals (confidence interval values are listed in Appendix A). The samples are generally similar in variability, though several of the fossil samples are more variable than any recent sample. No fossil sample is less variable than the least variable recent sample. Though the width of the confidence intervals makes firm conclusions difficult, the fossil samples as a group show a very slight increase in

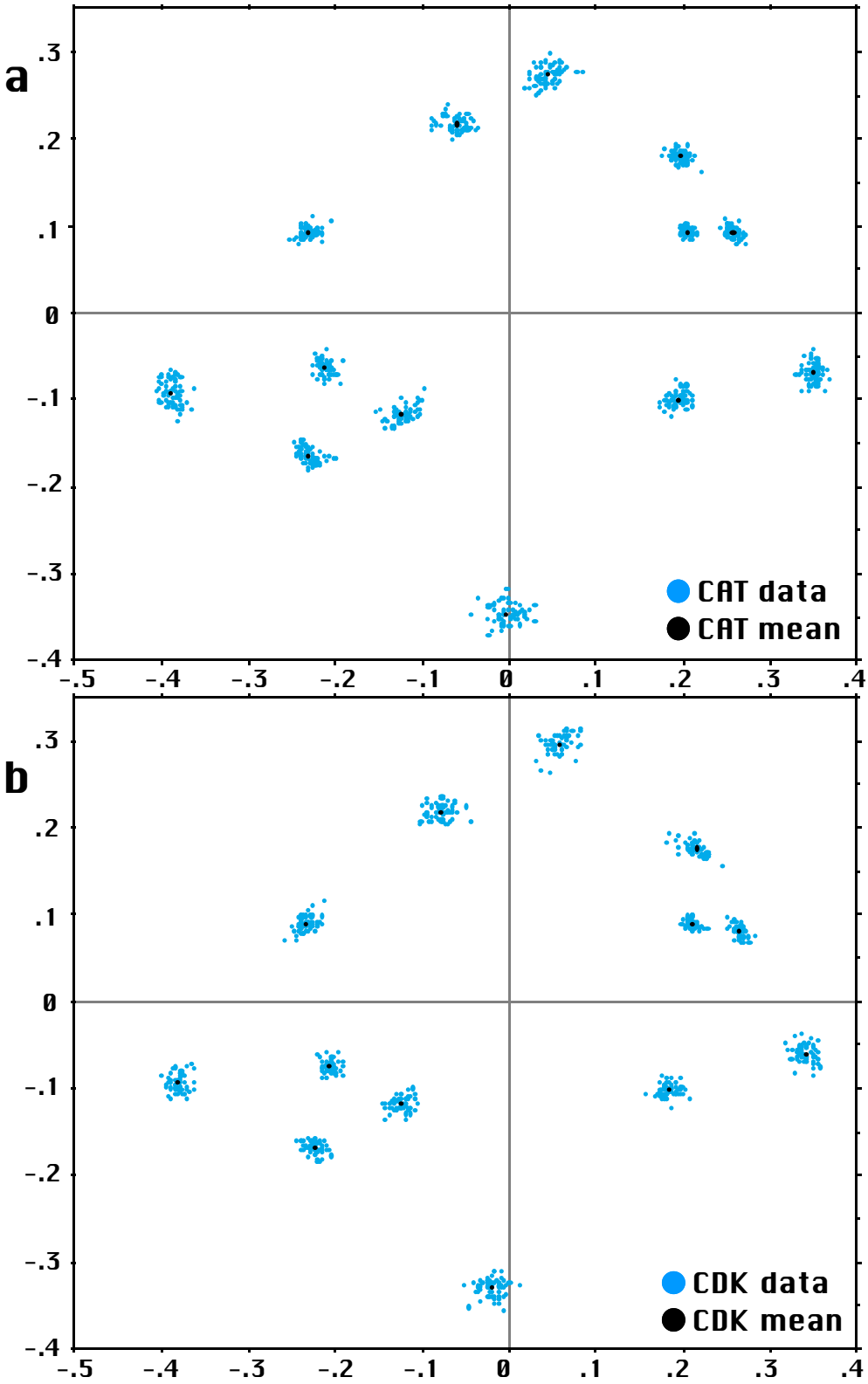

Figure 5.2a-b. (a) Full Procrustes fit of CAT. (b) Full Procrustes fit of CDK.

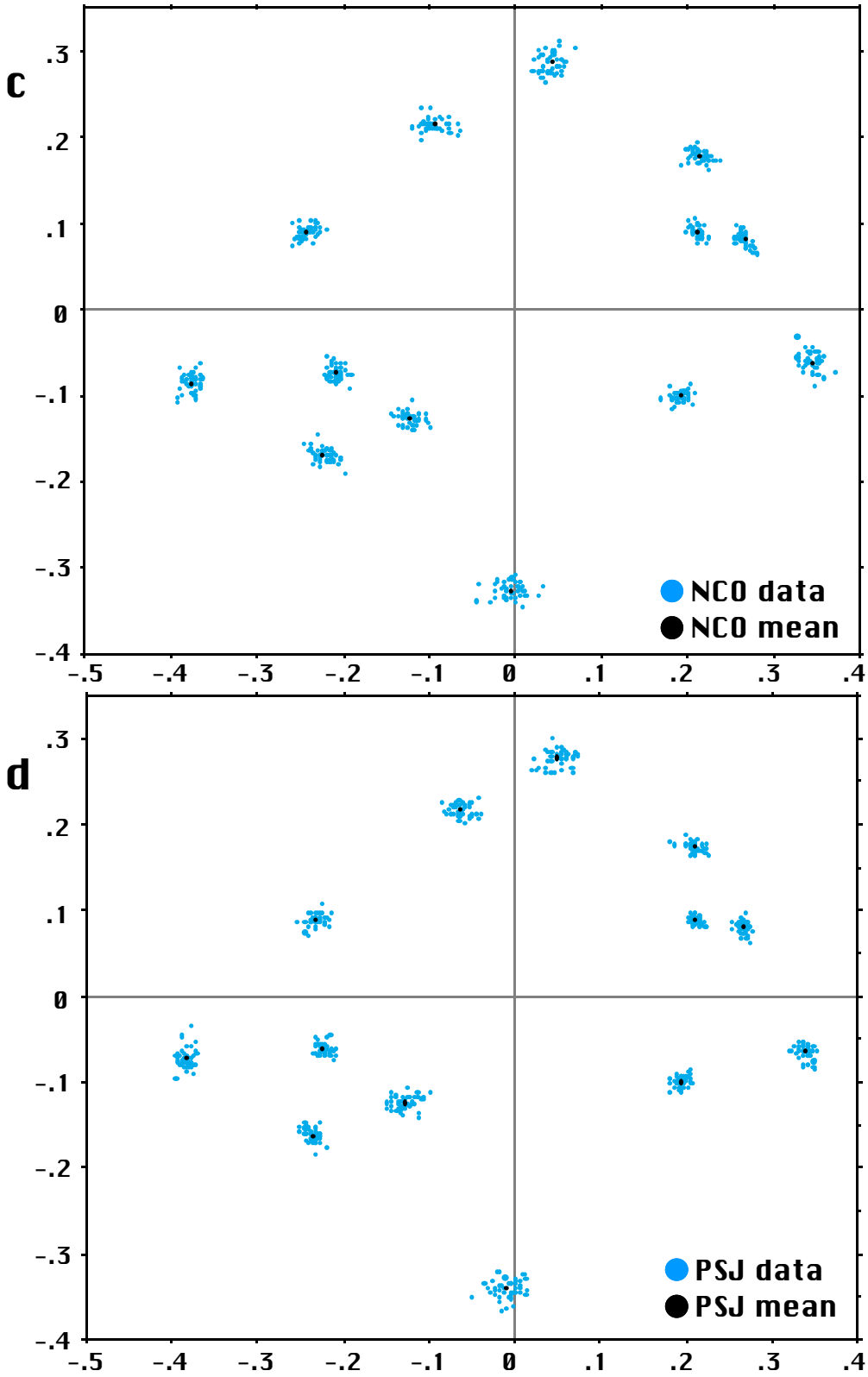

Figure 5.2c-d. (c) Full Procrustes fit of NCO. (d) Full Procrustes fit of PSJ.

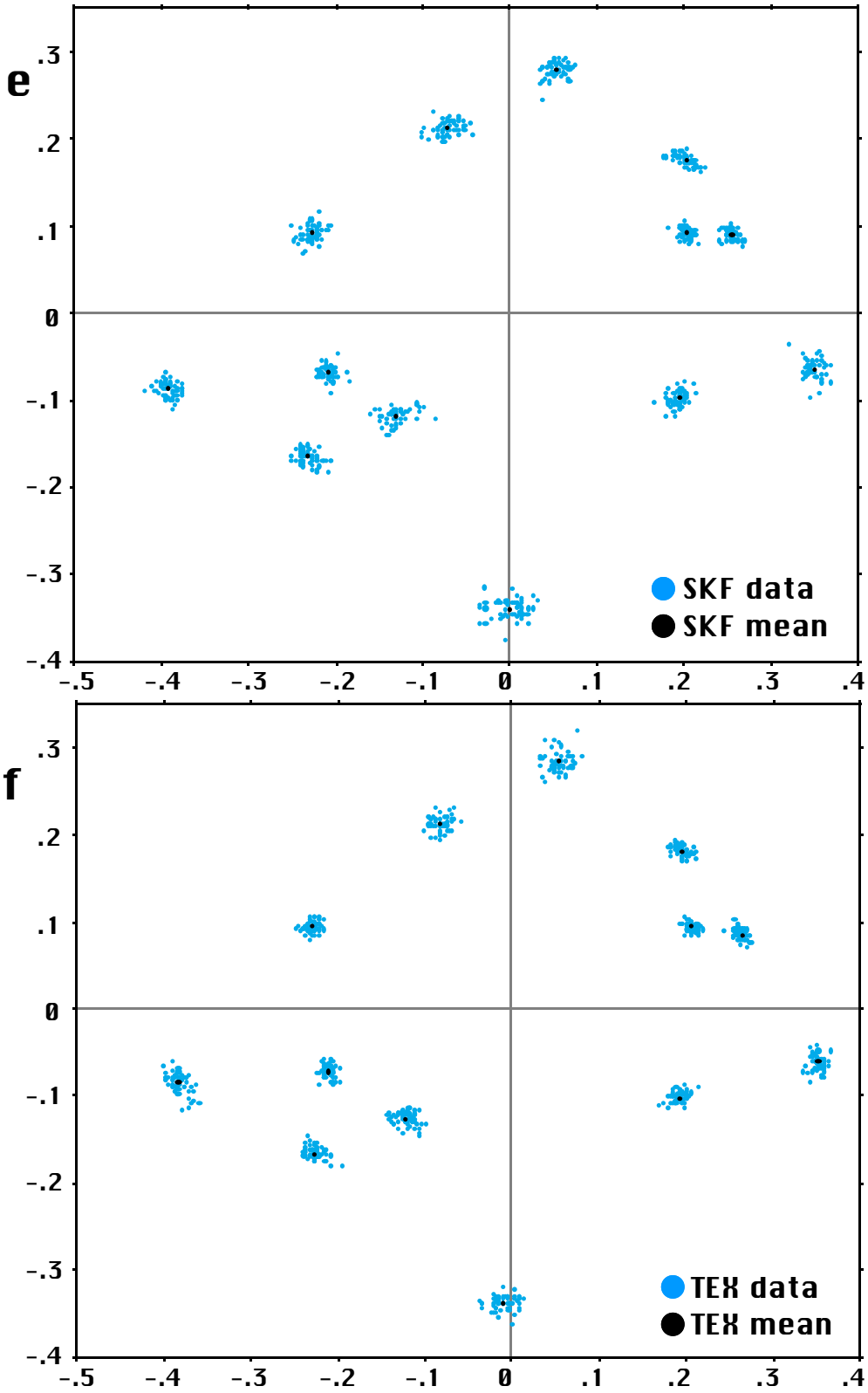

Figure 5.2e-f. (e) Full Procrustes fit of SKF. (f) Full Procrustes fit of TEX.

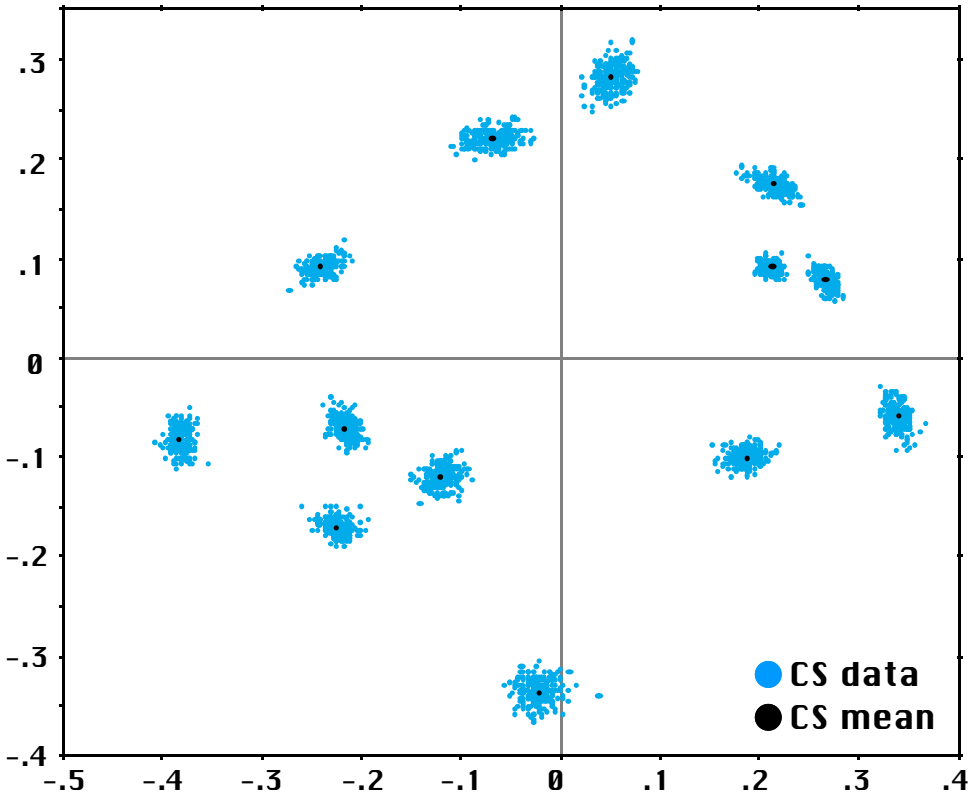

Figure 5.3. Full Procrustes fit of CS.

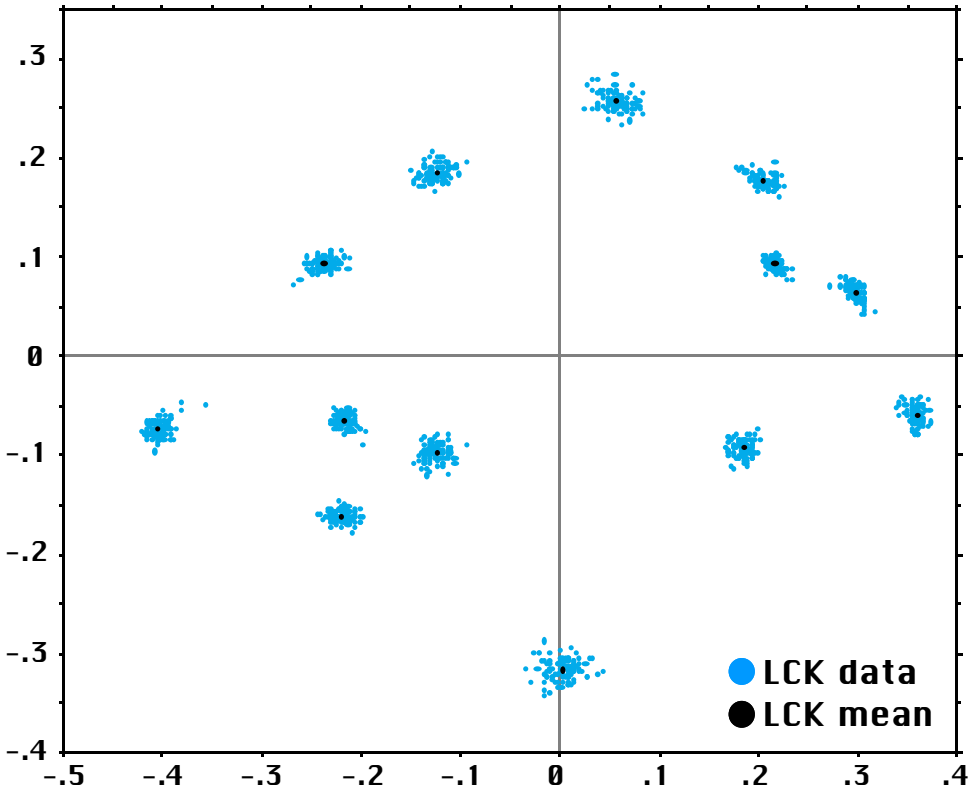

Figure 5.4. Full Procrustes fit of LCK.

total variability over the living populations. The causes of this increase in variability are explored below.

Table 5.2. Variability of recent samples, LCK, and Caloosa Shell stratal subsamples. "Tan. Coor." indicates Partial Tangent Coordinates.

	Sample	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
RECENT SAMPLES	CAT	0.0477	0.0477	0.0426	0.0425
	CDK	0.0466	0.0466	0.0448	0.0448
	NCO	0.0465	0.0465	0.0459	0.0458
	PSJ	0.0430	0.0430	0.0406	0.0405
	SKF	0.0491	0.0491	0.0472	0.0471
	TEX	0.0421	0.0421	0.0408	0.0408
SINGLE FOSSIL BEDS	LCK	0.0442	0.0442	0.0414	0.0414
	CS/13	0.0453	0.0453	0.0404	0.0404
	CS 11	0.0492	0.0492	0.0472	0.0471
	CS 10	0.0520	0.0520	0.0476	0.0475
	CS 9	0.0493	0.0493	0.0432	0.0432
	CS 8	0.0486	0.0486	0.0428	0.0428
	CS 6/7	0.0509	0.0509	0.0455	0.0454

Partitioning Total Variability

Numerous factors could be contributing to the total variability observed in the Mercenaria samples. First of all, some variability is inherent in any population due to the non-identity of individuals. Individuals vary phenotypically because of differences in both genetics and random external influences. Variability can be introduced by allometry if different size or age classes are mixed in a sample. Variability can conceivably change with size, even when the effects of allometry are removed. That is, the variability of a single size class of large individuals may be more (or less) variable than a single size class of small individuals, even though both are normalized to the same size.

Taphonomic alteration could increase the variability of the fossil samples by increasing measurement error in damaged valves. Pooling left and right valves could increase variability if left and right valves are not identical. Finally, evolution, ecophenotypic change, or migration could increase the variability of the fossil samples.

Allometry and Size

The size of each Mercenaria was measured as the centroid size, the square root of the summed squared distances from each landmark on a valve to their common centroid (Fig 4.2). The samples vary in both mean size and size range (Table 5.3, Figs. 5.6 to 5.8). The total Caloosa Shell sample had the greatest centroid size range; this range corresponds to a range of 32 to 150 mm in shell length. LCK and CAT also had wide ranges, and PSJ had the smallest range, one half that of LCK and CAT and a third that of CS. The stratal subsamples of CS are similar in centroid size distribution to the total sample.

The variations in centroid size could have several causes. The living samples were collected by a number of different people, some of whom may not have been collecting randomly with respect to shell size. Some populations may have been dominated by a small number of cohorts due to the irregularity of larval settlement. When the fossil samples were collected, all size classes were targeted, and time-averaging smoothes out larval settlement patterns. The differences in centroid size between the samples are meaningful to this study only because they may complicate the results.

Table 5.3. Centroid sizes of Mercenaria samples and subsamples. Units are mm.

	SAMPLE	N	MEAN	RANGE	MIN	MAX
RECENT SAMPLES	CAT	64	113.77	114.43	65.03	179.46
	CDK	53	163.18	87.43	103.16	190.59
	NCO	43	135.28	64.52	105.04	169.56
	PSJ	44	101.82	55.08	79.19	134.27
	SKF	52	140.85	75.34	92.76	168.10
	TEX	52	126.11	58.23	99.95	158.18
SINGLE FOSSIL BEDS	LCK	98	122.15	113.71	64.65	178.36
	CS 12&13	22	136.64	107.32	69.63	176.95
	CS 11	46	139.10	117.16	85.85	203.01
	CS 10	39	131.37	141.82	53.30	195.12
	CS 9	23	124.94	124.56	62.34	186.90
	CS 8	47	122.85	124.28	49.13	173.41
	CS 6&7	30	127.15	134.65	43.66	178.31
2.4 m BED	Total CS	207	130.39	159.35	43.66	203.01

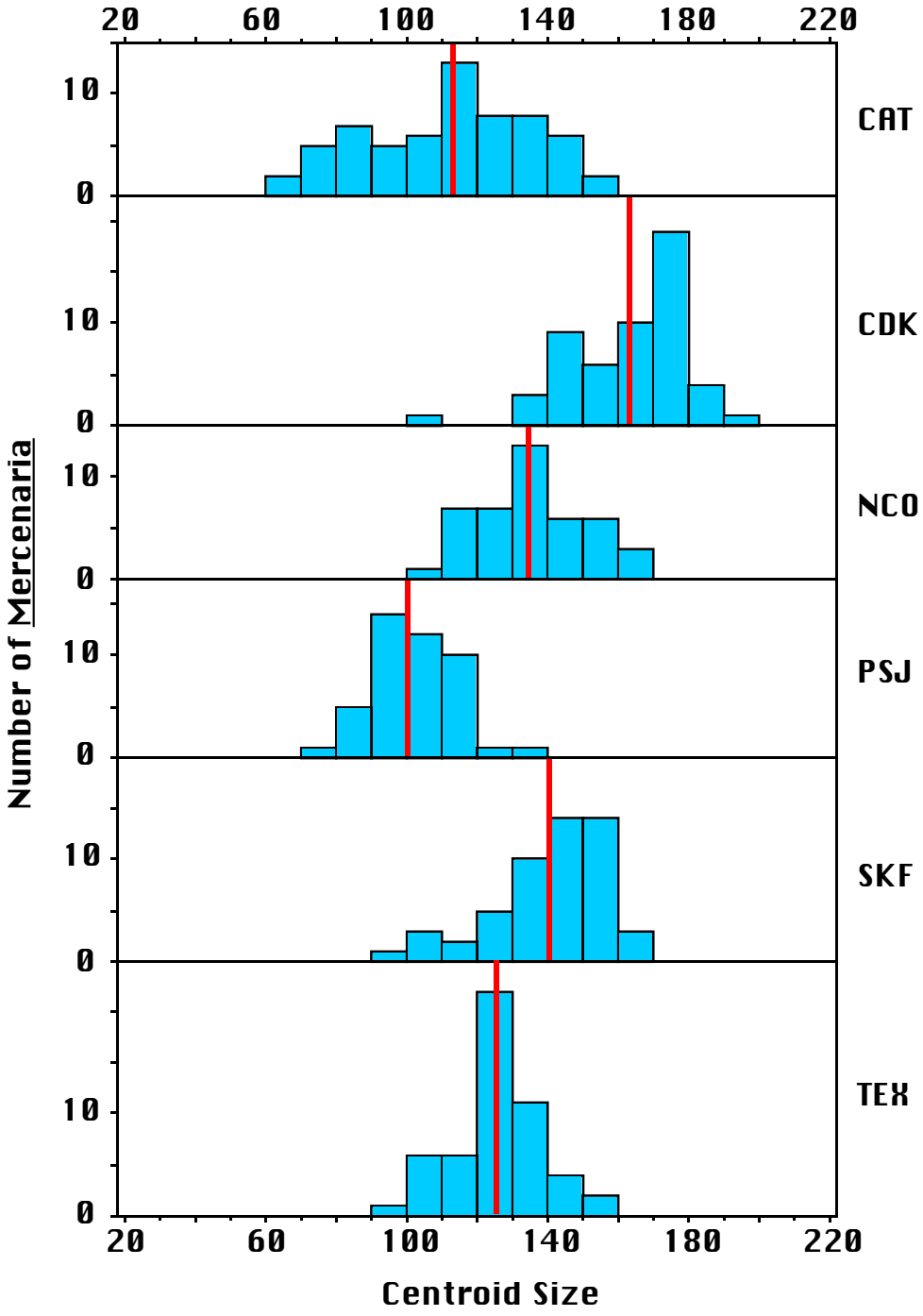

Figure 5.6. Centroid Size distribution of recent samples. Vertical lines mark the mean centroid size.

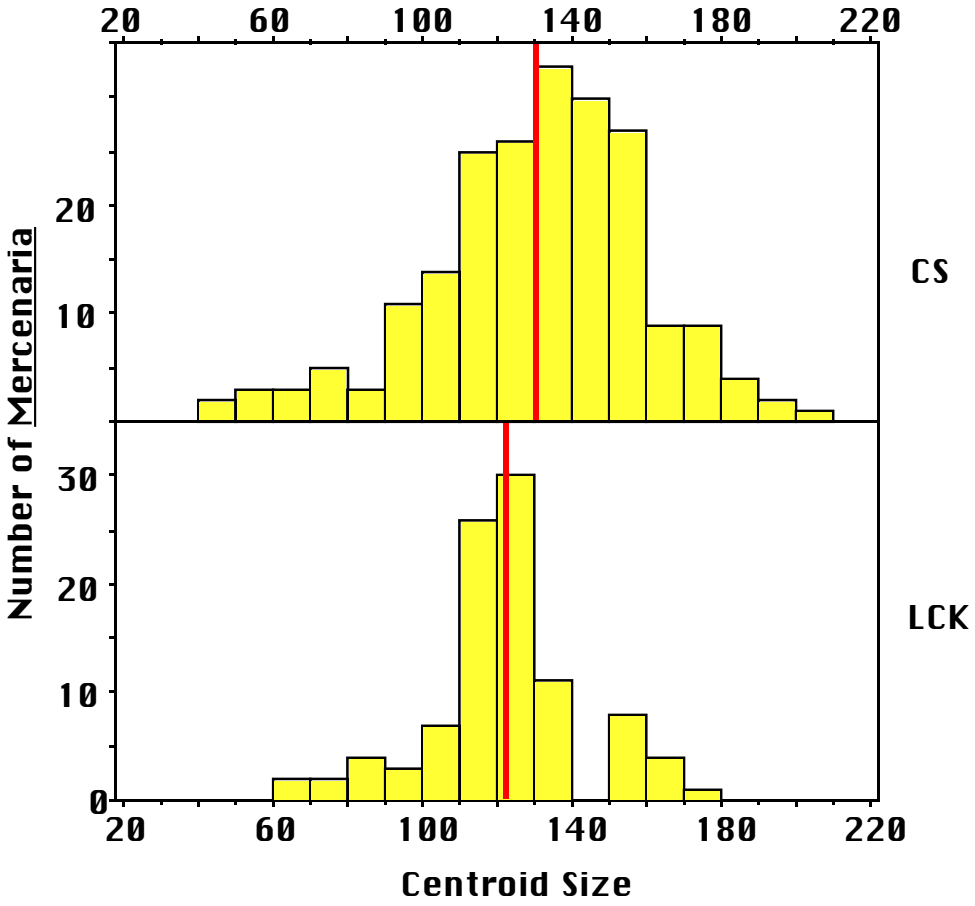

Figure 5.7. Centroid Size distribution of fossil samples. Vertical lines mark the mean centroid size.

If Mercenaria grow allometrically, changing shape as size increases, then the variability of the fossil samples will be preferentially inflated, since these samples have the greatest size ranges (Table 5.3, Figs. 5.6—5.8). In the Procrustes superimpositions of several samples, a number of landmarks are distributed elliptically. For example, in CS (Fig. 5.3a) the umbo and end of the ligament (landmarks 1 and 2) have elongate fields; this is often a sign of allometry or some other biological factor causing shape change. It is obvious from visual inspection of the specimens that the length of the ligament (distance between landmarks 1 and 2) shows positive allometry. The allometric origin of the ellipticity can be verified by plotting the tangent coordinates or residuals at the landmark against

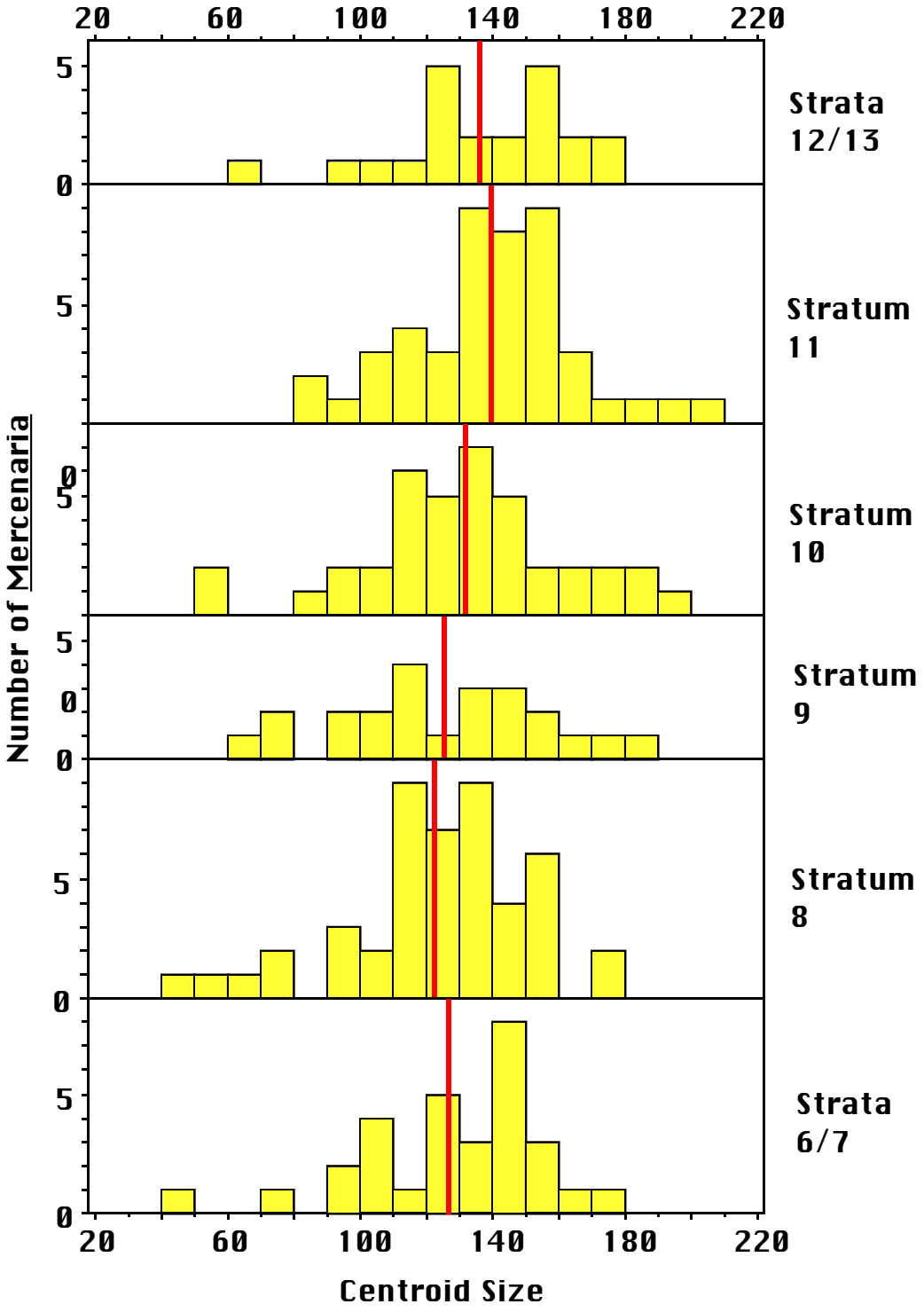

Figure 5.8. Centroid Size distribution of Caloosa Shell stratal subsamples. Vertical lines mark the mean centroid size.

centroid size (e.g., Fig. 5.9). The allometric signal is less obvious in the samples with smaller centroid size ranges.

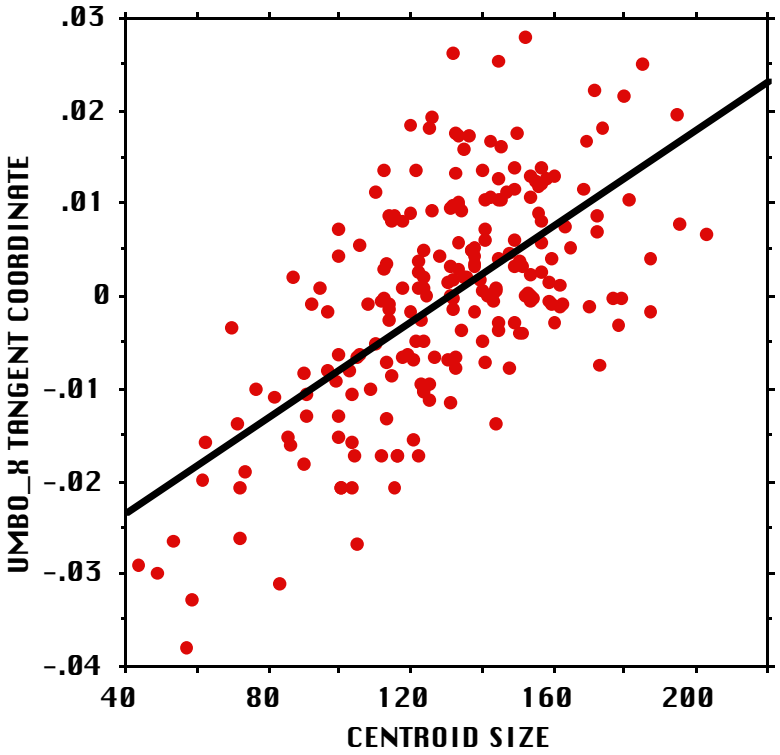

Figure 5.9. Relationship between the UMBO_X tangent coordinate and centroid size in the CS sample. The clear positive correlation (r^2=.411) reflects allometric growth. The residuals from the regression line are retained for allometry-free analysis.

The effects of allometry were removed by regressing the tangent coordinates and Procrustes residuals at each landmark on centroid size. "Allometry-free" variability was calculated as the root mean squared distance of the data to the mean using the residuals from the regressions (see Fig 5.9). The allometry-free variabilities of the recent samples and CS stratal subsamples are listed in Table 5.2, and the allometry-free variabilities from the tangent coordinates are plotted in Fig. 5.10. Fig. 5.11 shows the difference between total and allometry-free variability for each sample. Little variability was removed from the 5 recent samples with the smallest size ranges (CDK through TEX), but the greater size range of CAT led to greater ontogenetic averaging. Most of the fossil samples had a relatively large amount of variability removed (0.004 to

0.006), a result of the wide range of sizes present, but LCK and CS strata 11 were similar to the recent samples.

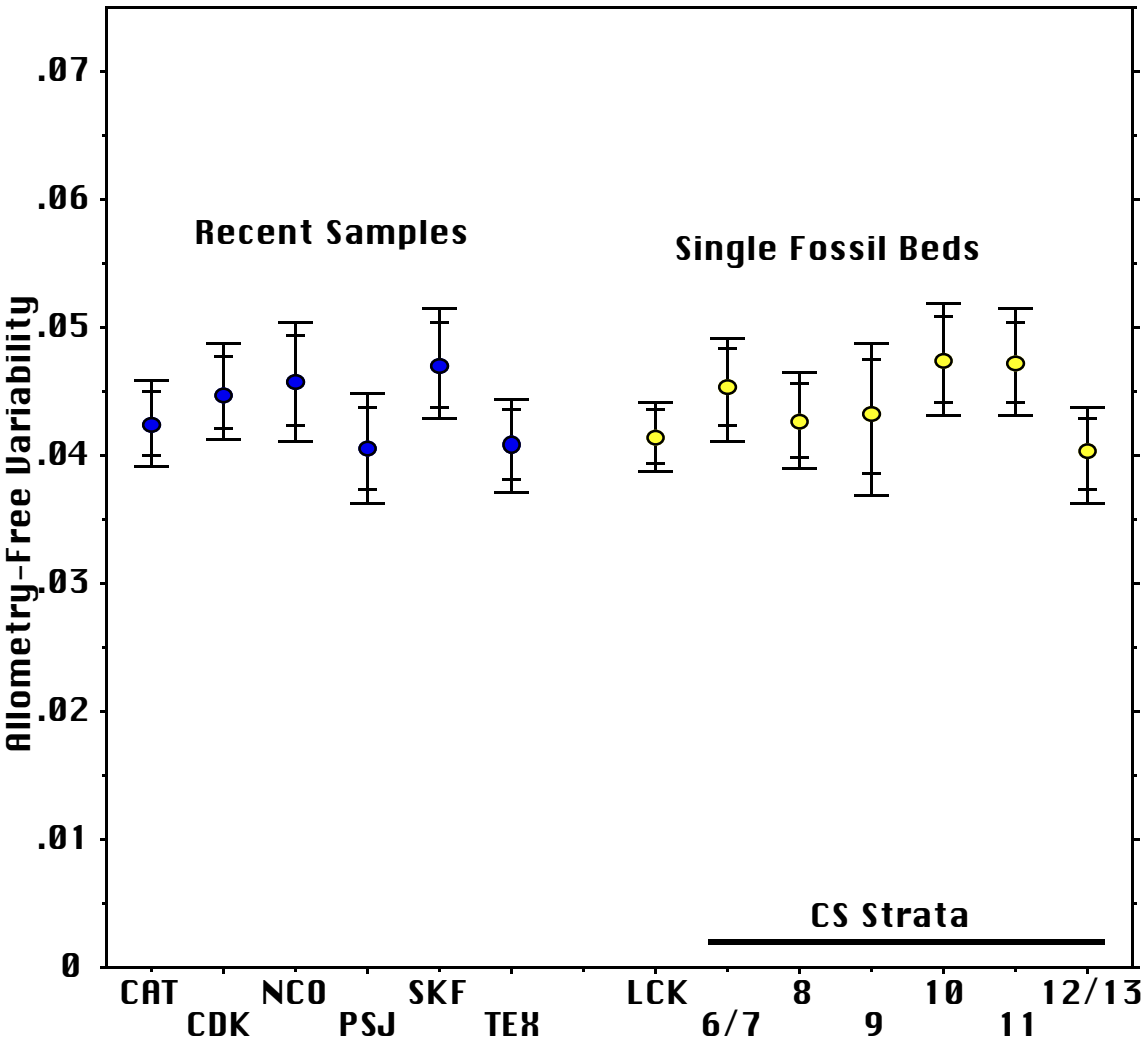

Figure 5.10. Allometry-free variability of the recent samples and samples from single fossil beds. 95% and 99% confidence intervals are shown.

There is no evidence that the fossil samples have higher allometry-free variabilities than the recent samples (Fig 5.10). One fossil sample (CS 10) is more variable than any recent sample, but another (CS 12 & 13) is less variable than any recent sample. All samples from single fossil beds fall easily within the 95% confidence intervals of the 6 living samples. Once the greater centroid size ranges of the fossil samples are taken into account, the fossil and recent samples are identical in variability.

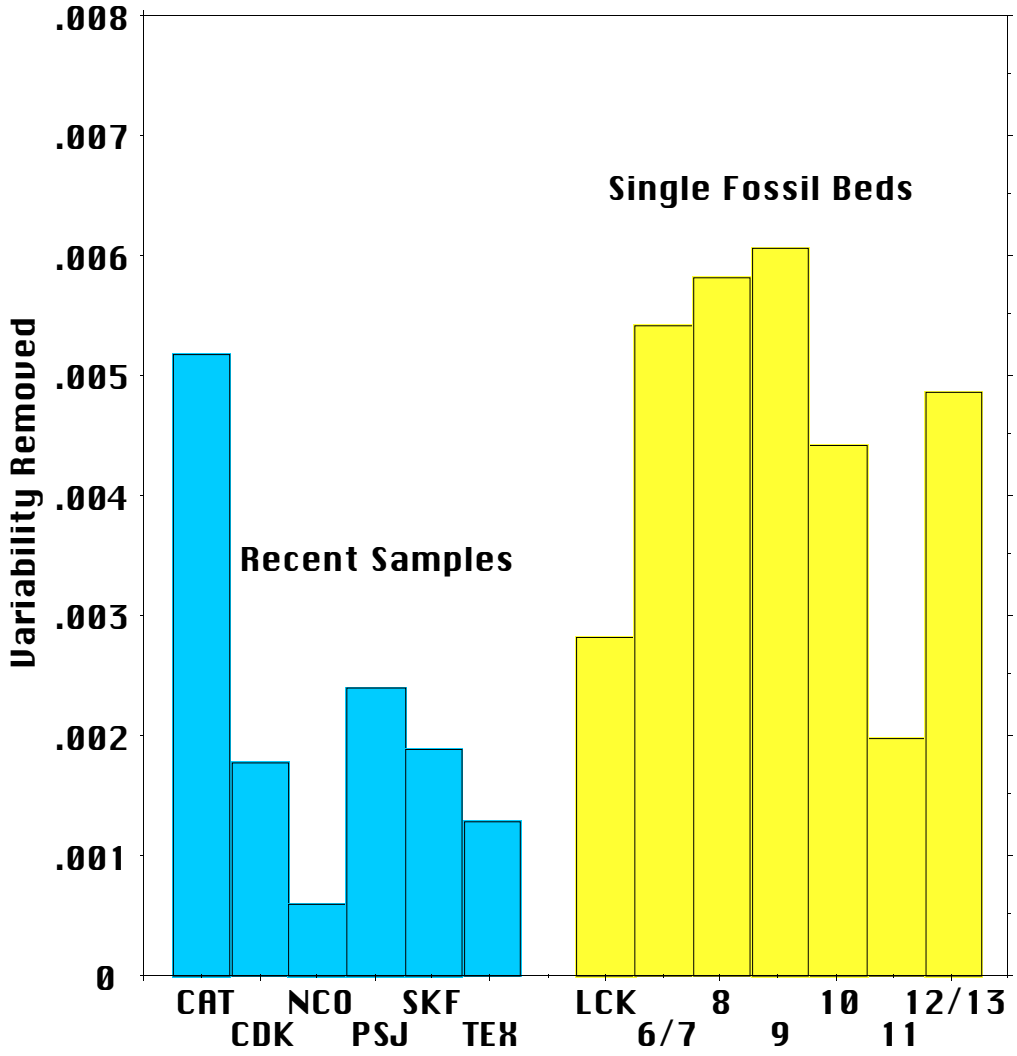

Figure 5.11. The amount of variability removed in the allometry correction.

Even after allometry is removed from the variability calculation, the sizes of the shells in a sample may affect the amount of shape variability. For example, consider a sample in which all shells have centroid size 50 and one in which all centroid sizes are 100. Since there is no size variability within samples, allometry does not affect shape variability. However, the larger shells might be more variable in shape simply because they are larger. (This corresponds to an increase in the coefficient of variation of traditional morphometrics.) This would be reflected in a plot like Fig 5.9 as an increase in scatter around the regression line.

To assess the effect of shell size on measured variability, the full CS sample was divided into 4 size classes, each consisting of one quarter of the specimens (Table 5.4). CSa contains the smallest 25% of the specimens by centriod size, and CSd contains the largest 25%. Two additional samples were constructed by combining the smallest and largest size classes (CSa&d) and the two intermediate size classes (CSb&c). LCK was divided into two halves by centroid size. These samples were chosen for this analysis because they have large sample sizes (the quarters of CS and halves of LCK are similar in sample size to the recent samples) and a wide enough centroid size range that the subsamples are somewhat different in mean size (Table 5.4).

Table 5.4. Centroid sizes of Mercenaria size subsamples. Units are mm.

	SAMPLE	N	MEAN	RANGE	MIN	MAX
CENTROID SIZE SUB-SAMPLES	CSa	52	92.82	69.97	43.66	113.63
	CSb	52	123.32	18.67	113.65	132.32
	CSc	52	140.81	17.82	132.51	150.33
	CSd	51	165.26	52.35	150.66	203.01
	CSa&d	103	128.69	159.35	43.66	203.01
	CSb&c	104	132.07	36.68	113.65	150.33
	LCKa	49	107.76	57.41	64.65	122.06
	LCKb	49	136.55	56.11	122.25	178.36

The size classes of CS show a progressive increase in allometry-free variability as shell size increases, as do the two halves of LCK (Table 5.5, Fig. 5.12). This is not the effect of ontogenetic averaging—allometry is factored out in the allometry-free variability measure. That allometry is being effectively removed is shown by the variabilities of CSa&d and CSb&c. CSa&d contains the largest and smallest size classes in CS and CSb&c contains shells of intermediate size; the two subsamples have nearly identical mean centroid sizes but vastly different ranges of centroid size. If allometric variability was not being removed, the sample with the greater size range should have a significantly higher

variability. However, the allometry-free variabilities of these two subsamples are nearly identical to each other and to the full CS sample. Further evidence that variability increases with size is provided by the recent samples. The three recent samples with the lowest average centroid size (CAT, PSJ, and TEX) have lower size-free variabilities than the 3 recent samples with larger shells (CDK, NCO, and SKF), (Fig. 5.10).

Table 5.5. Variability of Mercenaria size subsamples. "Tan. Coor." indicates Partial Tangent Coordinates.

	Sample	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CENTROID SIZE SUB-SAMPLES	CSa	0.0454	0.0454	0.0424	0.0423
	CSb	0.0445	0.0445	0.0439	0.0439
	CSc	0.0469	0.0469	0.0463	0.0462
	CSd	0.0503	0.0503	0.0493	0.0493
	CSa&d	0.0529	0.0529	0.0467	0.0466
	CSb&c	0.0470	0.0470	0.0458	0.0458
	LCKa	0.0395	0.0395	0.0370	0.0370
	LCKb	0.0461	0.0461	0.0410	0.0410

The increase in variability with shell size could have several causes. As the Mercenaria grow, injuries, errors, and eccentricities in the calcification process may accumulate, lending a greater random component to shape. Alternatively, young Mercenaria may start with the same basic shape but follow slightly different allometric trajectories, so that they become more disparate with age. Because the average centroid sizes of the fossil samples are well within the range of values of the recent samples, the samples can be compared without a further correction. However, the effect of size on the variability measurements should be kept in mind.

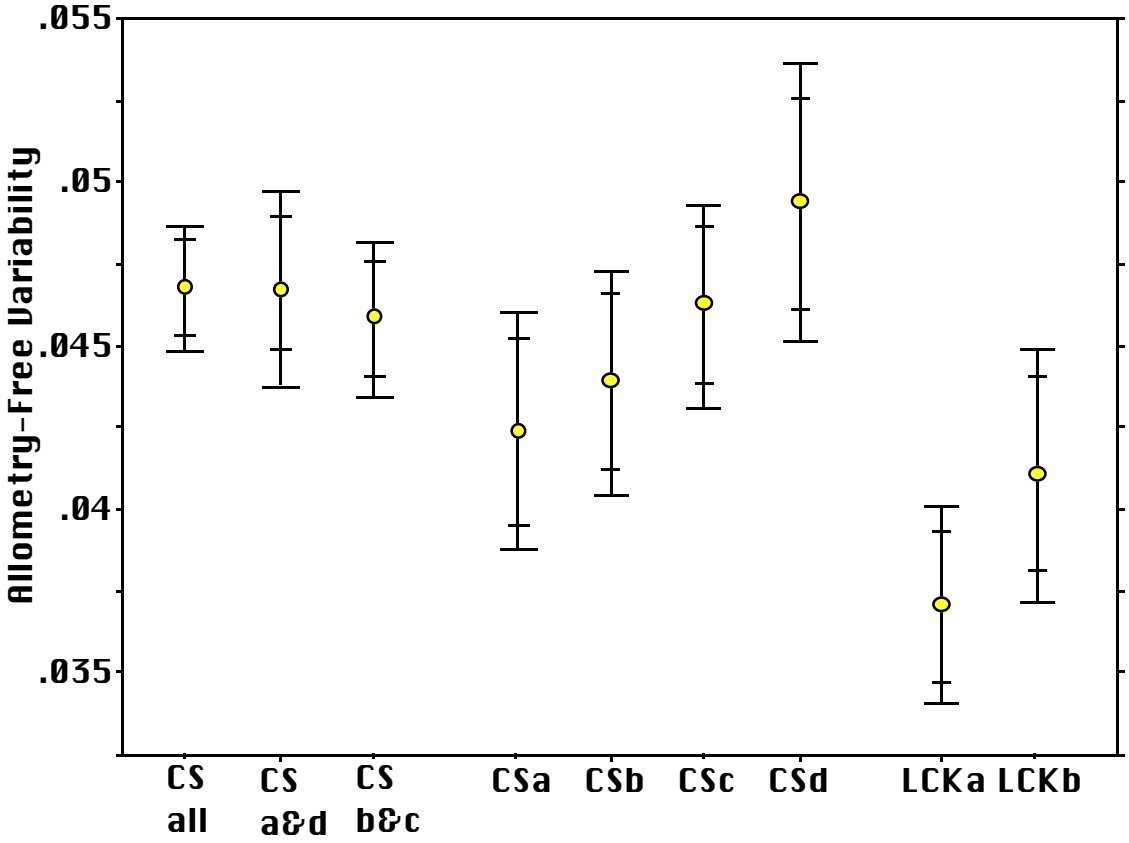

Figure 5.12. Allometry-free variability of CS and LCK size subsamples. CSa contains the 25% of CS that have the smallest centroid size, and CSd likewise contians the largest specimens. CSa&d contains these two samples pooled, and CSb&c contains the two intermediate size classes pooled. LCKa and LCKb are the smaller and larger halves of LCK. 95% and 99% confidence intervals are shown.

Taphonomy

The ability to reliably locate and measure the nine true landmarks (landmarks 1—9, see Fig. 4.1) was little affected by taphonomic alteration. However, the 4 pseudolandmarks (landmarks 10—13) were more difficult to reliably locate on shells that were chipped or abraded around the edges. To eliminate the potential effects of taphonomy, the variabilities of the original samples and the CS stratal subsamples were calculated a second time after eliminating the four pseudolandmarks at the dorsal, ventral, anterior, and posterior ends of the shells (Table 5.6, Fig. 5.13). Eliminating the 4 pseudolandmarks does not significantly change the results of the variability

analysis; the recent and fossil samples are still comparable in allometry-free variability.

Table 5.6. Variability of Mercenaria samples using the first 9 landmarks. "Tan. Coor." indicates Tangent Coordinates.

	Sample	Tan. Coor. Total	Tan. Coor. Allometry-free	Residuals Total	Residuals Allometry-free
RECENT SAMPLES	CAT	0.0438	0.0393	0.0438	0.0393
	CDK	0.0444	0.0424	0.0444	0.0424
	NCO	0.0415	0.0409	0.0415	0.0409
	PSJ	0.0387	0.0367	0.0387	0.0366
	SKF	0.0476	0.0459	0.0476	0.0458
	TEX	0.0387	0.0377	0.0387	0.0377
SINGLE FOSSIL BEDS	LCK	0.0418	0.0398	0.0418	0.0397
	CS 12/13	0.0468	0.0412	0.0468	0.0412
	CS 11	0.0496	0.0479	0.0496	0.0479
	CS 10	0.0513	0.0460	0.0513	0.0459
	CS 9	0.0485	0.0402	0.0485	0.0402
	CS 8	0.0483	0.0406	0.0483	0.0406
	CS 6/7	0.0512	0.0437	0.0512	0.0437

Pooling Left and Right Valves

All samples except for PSJ contained both left and right valves; all right valves were reflected before measurement and the data from left and right valves were pooled. Though all landmarks measured should be symmetric between the valves, a bootstrap hypothesis test was run on the other 5 recent samples to determine whether left and right valves are different. The bootstrap test is a non-parametric equivalent of the MANOVA and was run at 10,000 iterations (SAS code written by Eric Dyerson, University of Arizona, see Appendix B). The Procrustes partial tangent coordinates for each sample were first entered into a PCA, and the last four PC's were dropped to eliminate linear dependence. After

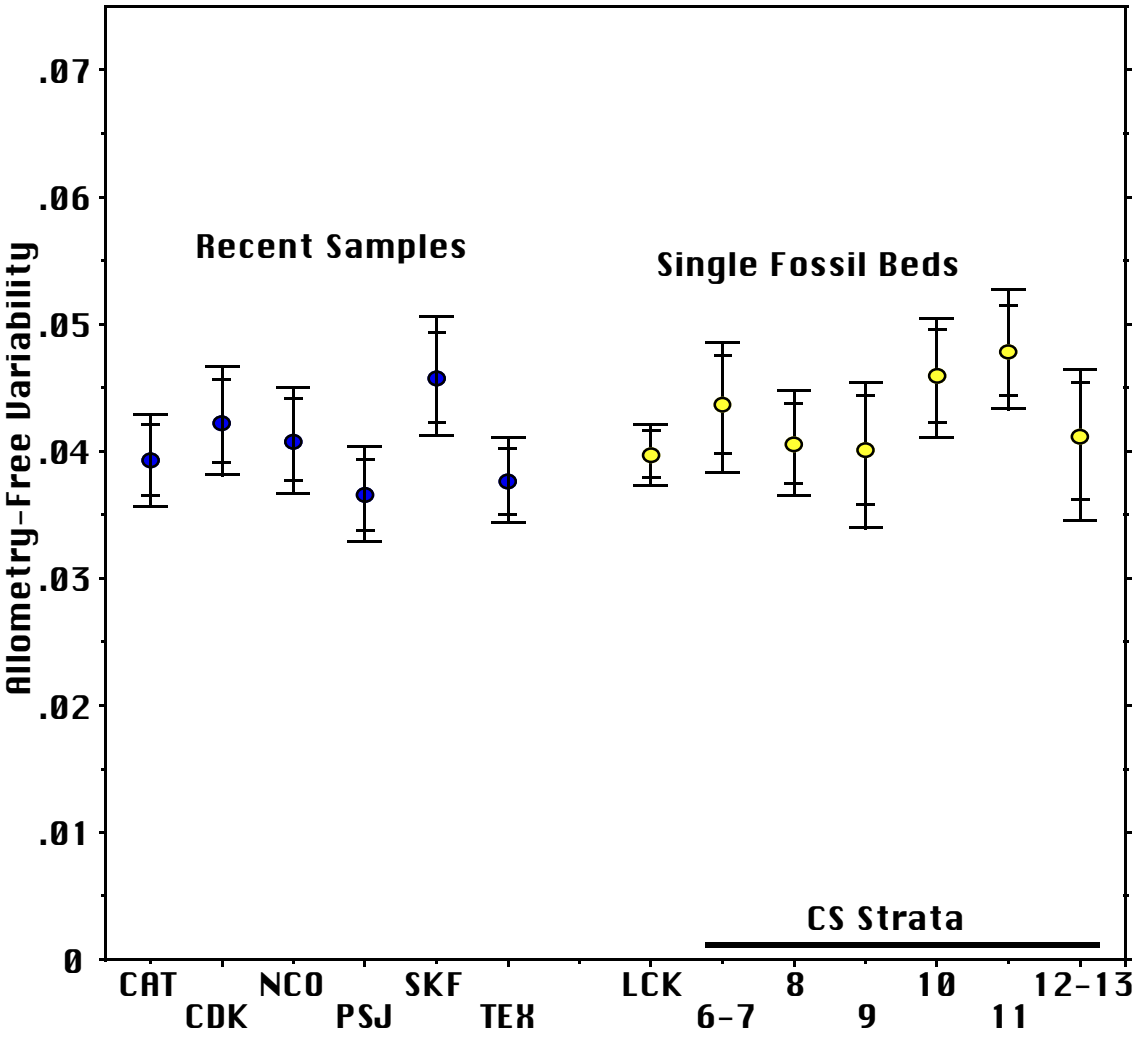

Figure 5.13. Allometry-free variability calculated only using the 9 true landmarks.

a Bonferroni correction for multiple tests, no significant differences were detected between the left and right valves of any sample (total α=0.05, pairwise α=0.01, see Table 5.7). Without the Bonferroni correction, the left and right valves of NCO are significantly different; however, the range of centroid sizes differs between the right and left valves in NCO, so allometry complicates the comparison. Pooling left and right valves has little, if any, effect on variability.

Table 5.7. Results of bootstrap test for significant differences between left and right valves. Significant differences detected without a Bonferroni correction are underlined; no significant differences were detected after the Bonferroni correction.

Sample	Bootstrapped p-value
CAT	0.070
CDK	0.253
NCO	0.014
SKF	0.219
TEX	0.156

Change in Variability Through Time

The variability of the Caloosa Shell Mercenaria differs from one stratum to the next, exceeding slightly the range of variabilities of the 6 recent samples (Fig. 5.10). This cannot be explained by differences in centroid size; the size-frequency distributions of the strata are quite similar, especially relative to differences between the 6 recent samples (Figs. 5.6 and 5.8). The centroid sizes of the CS strata were not significantly different by ANOVA ($p=0.088$). The differences in variability could be partly explained by estimation error, but the 95% confidence intervals for CS 11 and CS 12/13 do not overlap. Different amounts of variability could be caused by different amounts of time-averaging if morphology was fluctuating very rapidly (Kidwell 1986). The more time-averaged samples would be more variable. However, the variability of the samples does not increase or decrease up section as described by Kidwell (1986), nor do field observations support an increase or decrease in time-averaging up section. The shell bed likely had a more complex depositional history than the models presented by Kidwell (1986), and it is possible that time-averaging varies between strata. However, the range of variabilities of the different CS strata is similar to the range of the six living samples, so significant fluctuations in morphology during the interval of time-averaging in a single 30 cm layer are not evident. Rather, the population's variability fluctuated through time at a coarse (between-layer) time scale.

Analytical Time-Averaging

To assess the effects of analytical time-averaging, the 6 Caloosa Shell strata were pooled into 5 adjacent pairs, 4 adjacent triples, and so on (Table 5.8). The allometry-free variabilities are plotted in Fig. 5.14 with the 6 recent samples and 7 single beds for reference. While the single fossil beds have a range of variabilities indistinguishable from the recent samples, the adjacent pairs are restricted to the high end of this range. Additional analytic time-averaging does not, however, increase morphologic variability further. The scatter of values produced when two beds are combined is simply averaged out when more than two adjacent beds are pooled.

Table 5.8. Variabilities of various combinations of Caloosa Shell strata. "Tan. Coor." indicates Tangent Coordinates.

Number of Adjacent Strata Pooled	Strata	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-Free	Residuals Allometry-Free
2	11 to 12/13	0.0485	0.0485	0.0464	0.0464
	10 to 11	0.0510	0.0510	0.0483	0.0483
	9 to 10	0.0517	0.0517	0.0471	0.0470
	8 to 9	0.0491	0.0491	0.0437	0.0436
	6/7 to 8	0.0500	0.0500	0.0446	0.0446
3	10 to 12/13	0.0501	0.0501	0.0474	0.0474
	9 to 11	0.0510	0.0510	0.0479	0.0478
	8 to 10	0.0505	0.0505	0.0458	0.0458
	6/7 to 9	0.0501	0.0501	0.0449	0.0448
4	9 to 12/13	0.0502	0.0502	0.0472	0.0471
	8 to 11	0.0504	0.0504	0.0468	0.0467
	6/7 to 10	0.0510	0.0510	0.0463	0.0462
5	8 to 12/13	0.0499	0.0499	0.0464	0.0464
	6/7 to 11	0.0508	0.0508	0.0471	0.0470
6	6/7 to 12/13	0.0504	0.0504	0.0467	0.0467

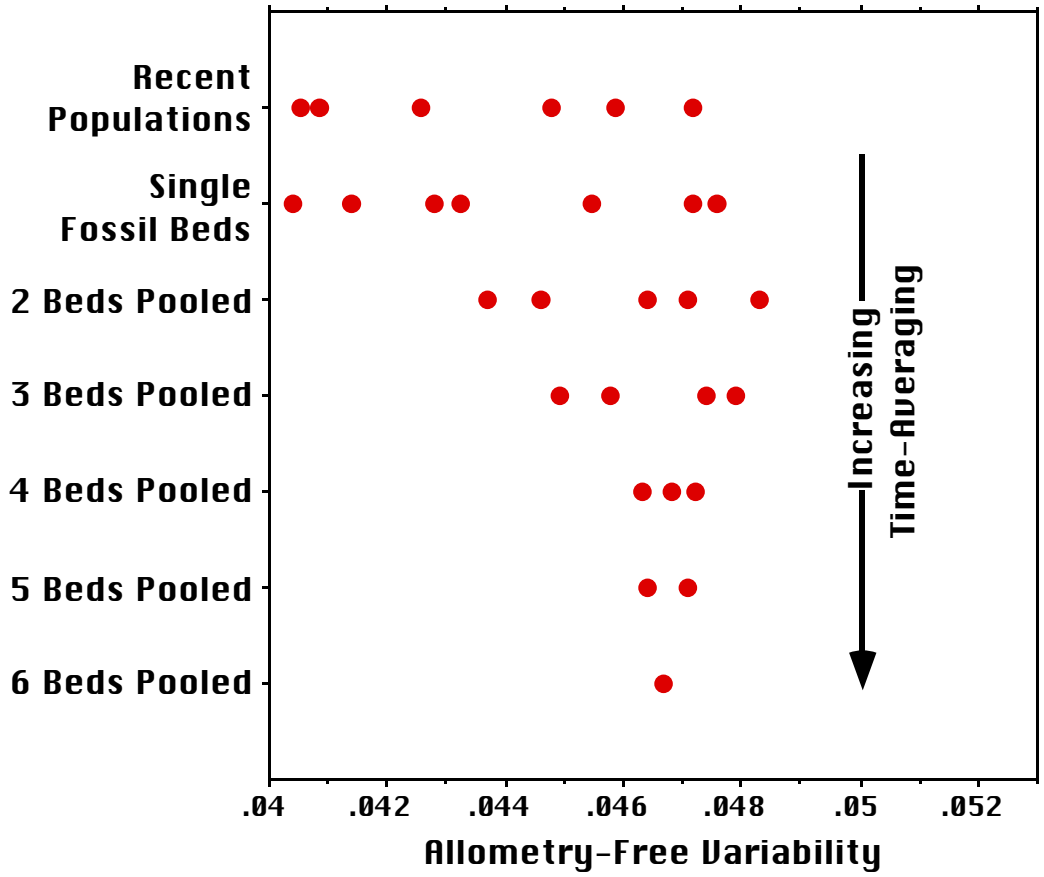

Figure 5.14. The effects of analytical time averaging on variability. Analytical time-averaging removes the low values observed for some recent samples and single fossil beds, but does not increase variability much beyond the range seen in living populations.

Distribution of Caloosa Shell Strata in Morphospace

To aid in interpreting the pattern of variability in analytically time-averaged samples, the Caloosa Shell Mercenaria were ordered in a canonical variate (canonical discriminant) analysis. The Procrustes partial tangent coordinates for the combined CS strata were first regressed on centroid size to remove size differences caused by allometry. The residuals from the regressions were entered into a principal components analysis and the last four PC's were discarded. (Due to the mathematics of the Procrustes analysis, these last four principal components have scores of zero. Removing the linear dependence in the data allows the matrix inversions required in a canonical variate analysis.)

The remaining 22 principal components were entered into a canonical variate analysis grouped by stratum.

The first two canonical variates contain 68.1% of the separation (Table 5.9), and the strata plot in broadly overlapping fields (Fig. 5.15a, c). The third and fourth variates contain 23.2% of the separation, and overlap between the samples is great (Fig. 5.15b, d). The means and ranges of the canonical variate scores are also plotted in stratigraphic sequence in Fig 5.16. The population seems to exhibit general stasis, though the means differ slightly between strata. With only six data points and probable temporal overlap between samples (Flessa et al. 1993), it is difficult to analyze the time-series pattern. However, ordinary hypothesis tests can be performed to test for differences in mean values of the canonical variates.

Table 5.9. Summary statistics of canonical variate analysis on the Caloosa Shell strata.

Canonical Variate	Eigenvalue	Proportion	Cumulative
1	0.463	0.400	0.400
2	0.325	0.281	0.681
3	0.146	0.126	0.807
4	0.123	0.106	0.913
5	0.100	0.087	1.000

Hypothesis Tests

Statistically significant differences in mean shape were identified using MANOVA and the non-parametric, bootstrap equivalent run at 30,000 iterations. The results of the MANOVA and the bootstrap test are identical (Table 5.10). When a Bonferroni correction for multiple comparisons is applied, 5 pairwise differences are detected (total α=0.05, pairwise α=0.0033). Only strata 6/7 and 10 are significantly different from other stratum. Even when the Bonferroni correction is not applied, inflating the probability of incorrectly identifying a difference, still only strata 6/7 and 10 are significantly different from any other

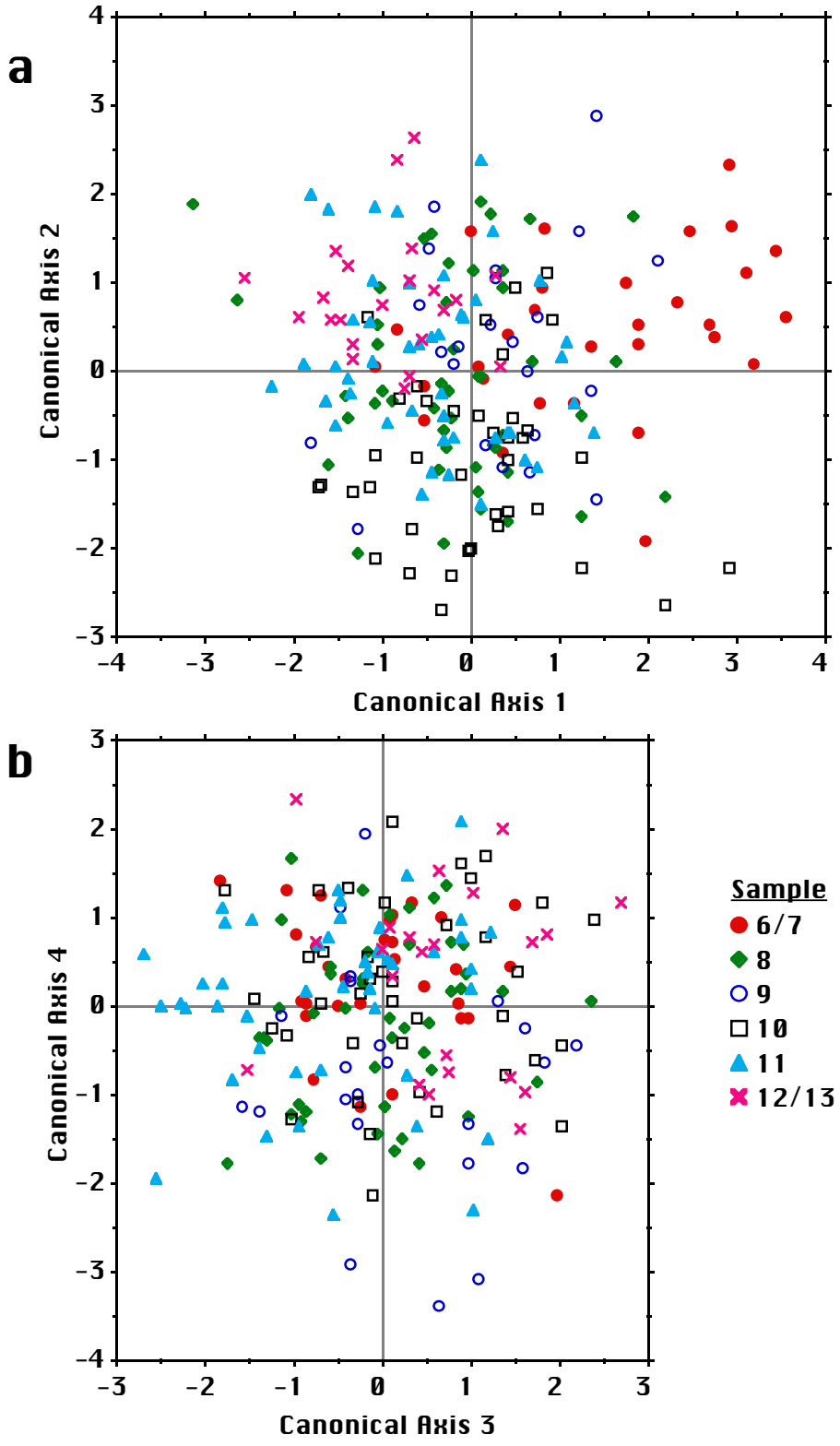

Figure 5.15a-b. Canonical Variate Analysis of the Caloosa Shell strata. (a) Canonical variates 1 and 2, (b) 3 and 4.

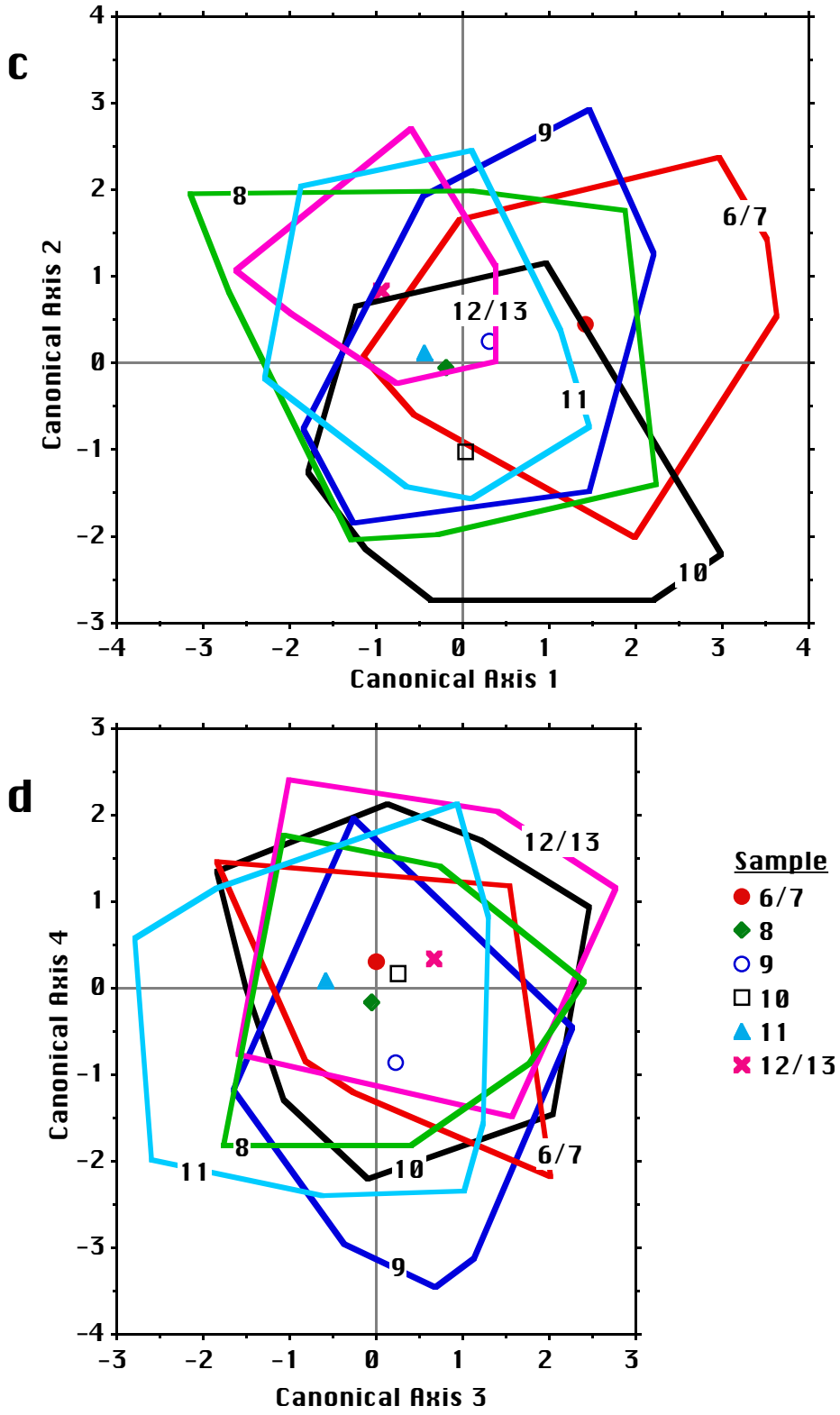

Figure 5.15c-d. Convex hulls of the canonical scores of the CS strata. (c) Canonical variates 1 and 2, (d) 3 and 4.

straum (Table 5.10). No statistically significant differences were detected in the pairwise combinations of strata 8, 9, 11, and 12. The reasons for the detected differences between the strata are unknown; any combination of evolution, ecophenotypy, and migration is possible. It is apparent, however, that there is no trend in morphologic change; there are at most slight fluctuations. Successive populations broadly overlap in morphology even when plotted in canonical variate space, which maximizes separation (Figs. 15—16).

Table 5.10. Mahalanobis distances between strata and tests of significance. The bootstrap test was run for 30,000 iterations. Pairwise comparisons that are significant at α=0.05 are underlined; those that are significant after a Bonferroni correction are also in bold type (total α=0.05, pairwise α=0.0033).

Samples		Mahalanobis Distance (D^2)	Bootstrap p-value	MANOVA p-value
6/7	8	3.347	**0.0008**	**0.0005**
6/7	9	2.816	0.0797	0.0810
6/7	10	4.232	**0.0001**	**0.0001**
6/7	11	4.010	< **0.0001**	**0.0001**
6/7	12/13	6.231	**0.0001**	**0.0001**
8	9	1.673	0.4062	0.4049
8	10	1.672	0.0969	0.0966
8	11	1.046	0.4811	0.4791
8	12/13	2.573	0.0578	0.0578
9	10	2.788	0.0426	0.0417
9	11	2.124	0.1586	0.1597
9	12/13	3.572	0.0430	0.0430
10	11	2.272	0.0090	0.0092
10	12/13	4.665	**0.0003**	**0.0002**
11	12/13	2.409	0.0926	0.0932

Slight, slow, directionless fluctuation is the apparent morphologic pattern. Change is not rapid enough to inflate the variability of individual strata, yielding variabilities similar to those of living populations. Slight differences do

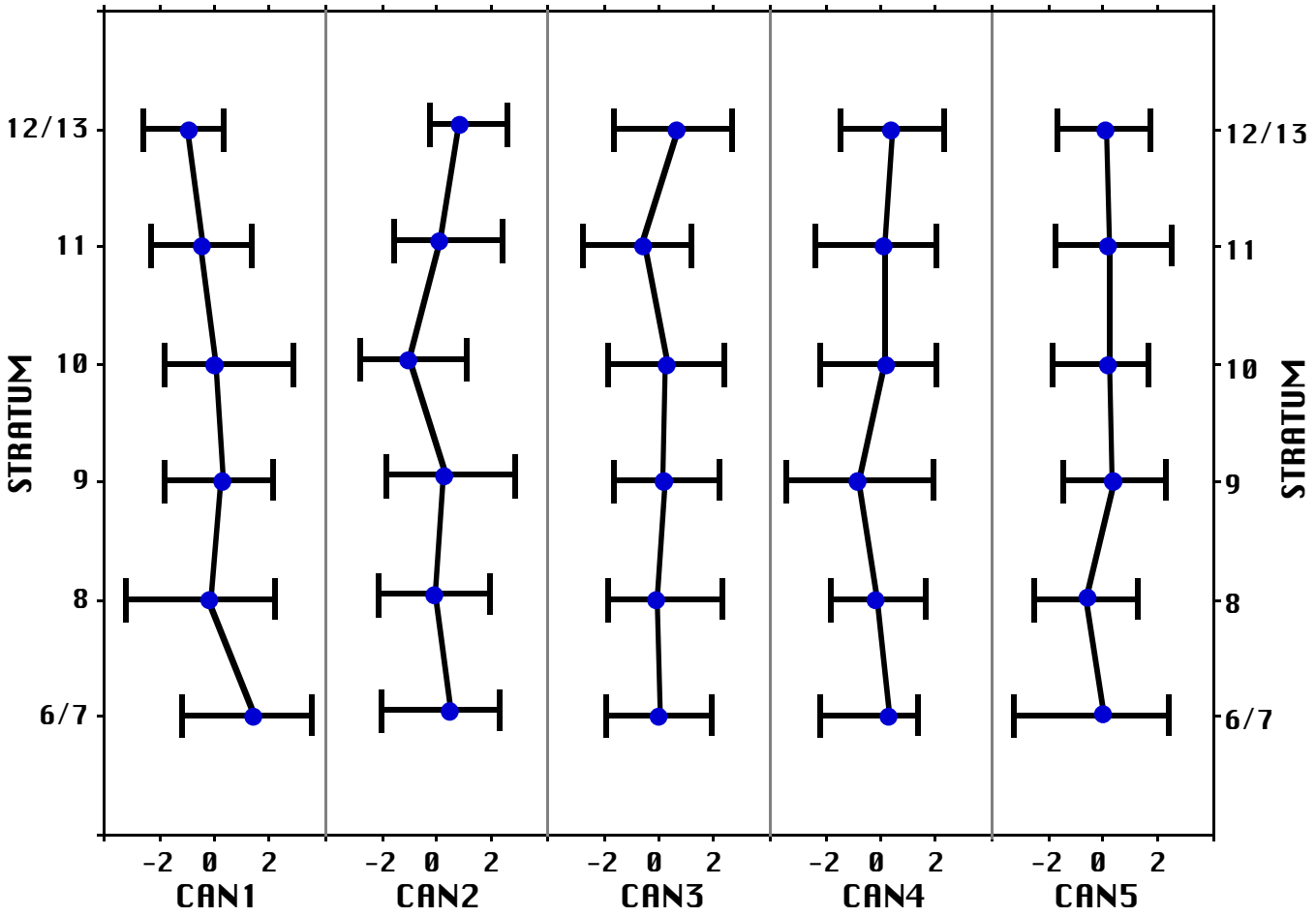

Figure 5.16. Morphology through the CS section. Means and ranges of the canonical scores are plotted.

accumulate between some strata, but not enough to inflate the variability much beyond the range seen in living populations. Because there is no consistent direction to change, additional analytic time-averaging does not tend to further inflate variability. It simply produces variability comparable to the more variable living populations.

Geographic Variability

To assess the effects of geographic pooling on morphologic variability, the six recent samples were combined in various ways and analyzed. Table 5.11 presents the variabilities of the 15 pairwise combinations of the 6 samples, the 4 Florida samples combined, and all 6 samples combined. The allometry-free variabilities are grouped into categories of roughly increasing geographic separation in Fig. 5.17. First are the pairwise combinations of the four Florida

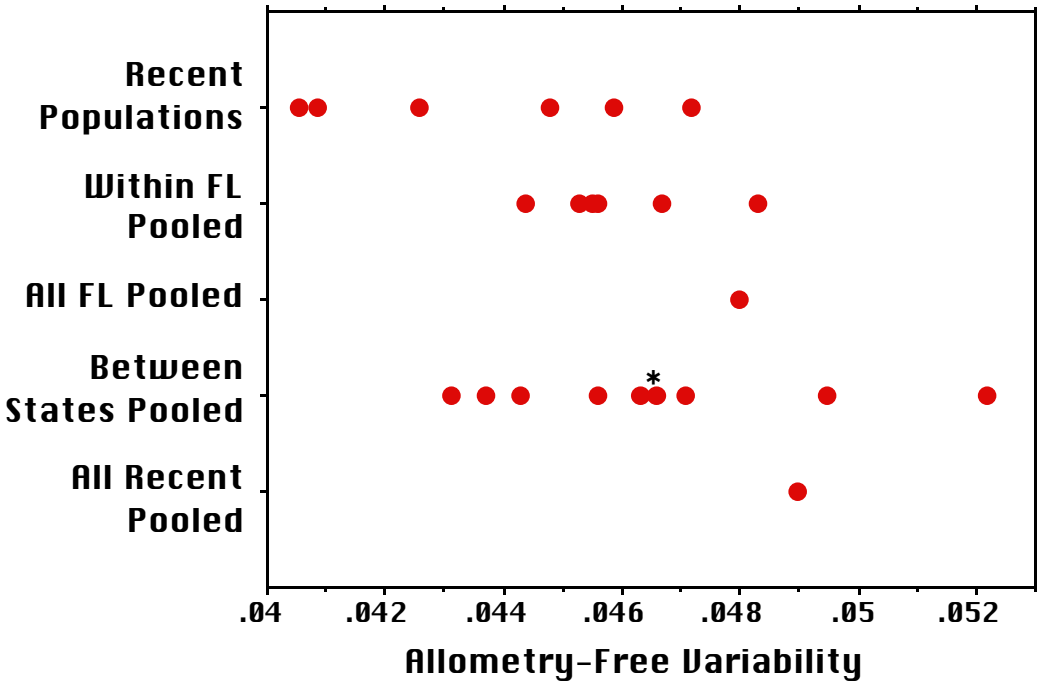

Figure 5.17. Variability of the pairwise pooled recent samples. The TEX-NCO pooled sample is marked by "*".

Table 5.11. Variability of pooled recent samples. "Tan. Coor." stands for Partial Tangent Coordinates.

Samples			Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
PAIRWISE POOLED SAMPLES	CAT	CDK	0.0539	0.0539	0.0453	0.0452
	CAT	NCO	0.0543	0.0543	0.0495	0.0494
	CAT	PSJ	0.0484	0.0484	0.0456	0.0456
	CAT	SKF	0.0498	0.0498	0.0455	0.0454
	CAT	TEX	0.0483	0.0483	0.0443	0.0443
	CDK	NCO	0.0491	0.0491	0.0471	0.0471
	CDK	PSJ	0.0505	0.0505	0.0444	0.0443
	CDK	SKF	0.0523	0.0523	0.0483	0.0483
	CDK	TEX	0.0477	0.0477	0.0437	0.0437
	NCO	PSJ	0.0494	0.0494	0.0456	0.0455
	NCO	SKF	0.0528	0.0528	0.0522	0.0521
	NCO	TEX	0.0474	0.0474	0.0466	0.0465
	PSJ	SKF	0.0489	0.0489	0.0467	0.0466
	PSJ	TEX	0.0455	0.0455	0.0431	0.0431
	SKF	TEX	0.0476	0.0476	0.0463	0.0463
ALL FLORIDA SAMPLES			0.0526	0.0526	0.0480	0.0479
ALL RECENT SAMPLES			0.0523	0.0523	0.0490	0.0489

samples, which are located relatively close to each other on Florida's west coast and panhandle. The variabilities of these pairwise combinations fall in the upper half of the range of the single populations, without much exceeding the range of those samples. When all 4 Florida samples are pooled, the variability slightly exceeds those of the single populations. The Texas and North Carolina samples are relatively distant geographically from the Florida samples (Fig 3.2). The pairwise combinations of TEX and NCO with the Florida samples and each other produce a disperse range of variabilities that broadly overlaps the single populations, though higher values also result. Pooling the two most distant

samples, TEX and NCO, produces only an intermediate variability (marked by "*" in Fig 5.17). The variability of all recent samples combined is somewhat higher than any single living population.

The relationship between variability and scale of geographic pooling is weak at best. Pooling geographically separated samples does, on average, increase variability; no low values of variability are produced. However, the distance between the pooled populations has at best a weak relationship with the amount of variability increase. Most samples pooled between states are similar to those pooled within Florida, though a few are more variable. Pooling TEX and NCO, the most distant populations, produces only an average variability compared to the other pooled samples. The recent populations appear to be differentiated morphologically, but not in a way that varies systematically with distance.

There is a technical difficulty with calculating allometry-free variability in these samples. If the samples being combined show different allometric trajectories, then the allometry correction by regression may work poorly (see Fig 2.8). Since the recent samples differ in centroid size, false allometry may even be detected, and too much variability may be removed. If the allometry of each population could be examined, then the severity of this complication could be assessed. However, the size ranges of the six recent samples are generally too small to detect and interpret the slight allometry of Mercenaria. If a morphometric analysis is being performed on geographically pooled samples, then this allometry complication should be kept in mind. Of course, studying allometry can be similarly complicated in time-averaged samples.

Distribution of Recent Samples in Morphospace

A canonical variate analysis was performed to explore the distribution of the recent populations in morphospace, following the procedure outlined for the Caloosa Shell strata. The six recent samples are more differentiated morphologically than are the CS strata (Fig. 5.18, cf. Fig. 5.15). Many pairs of

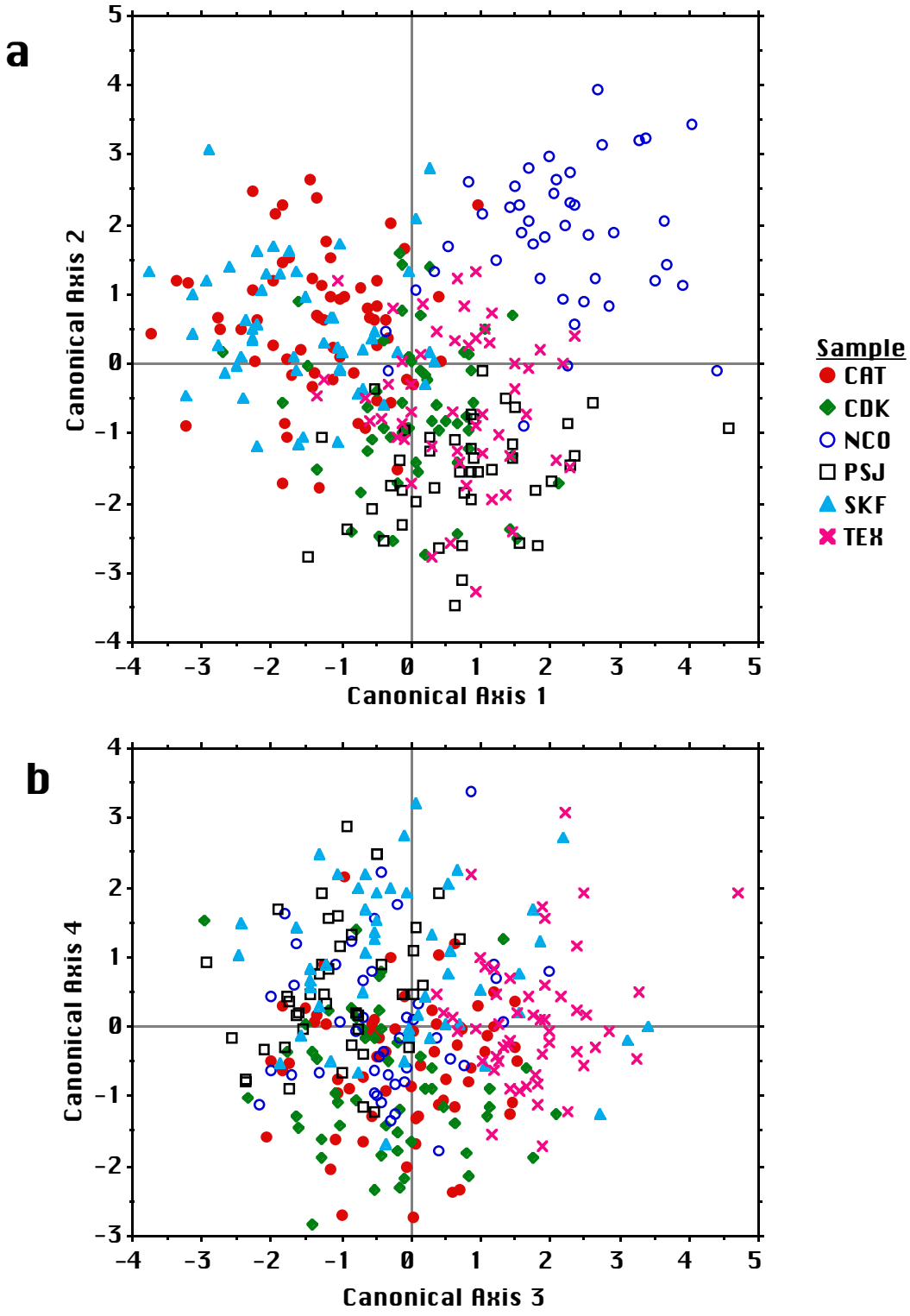

Figure 5.18a-b. Canonical Variate Analysis of the recent samples. (a) Canonical variates 1 and 2, (b) 3 and 4.

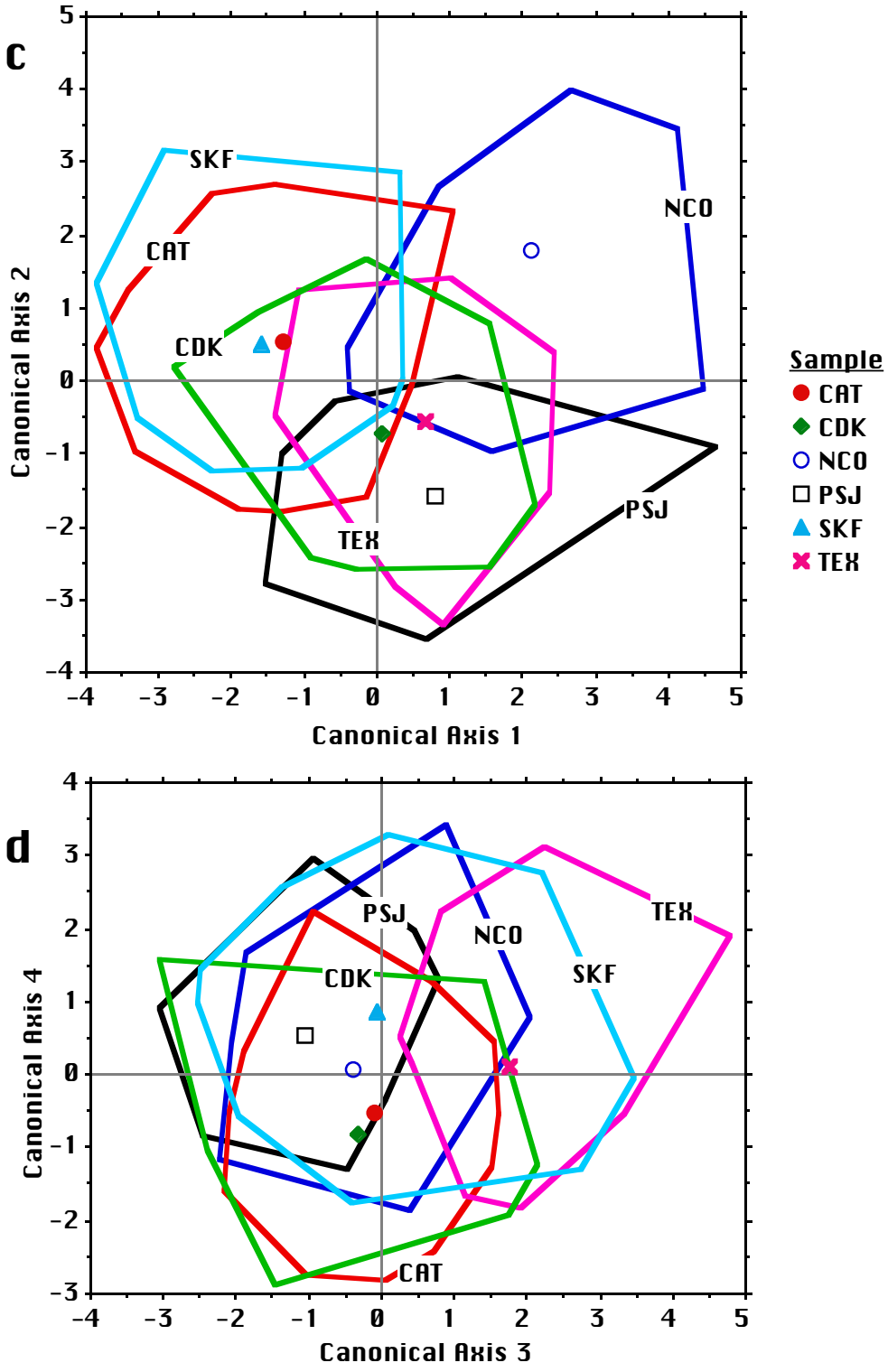

Figure 5.18c-d. Convex hulls of the canonical scores of the recent samples. (c) Canonical variates 1 and 2, (d) 3 and 4.

recent samples show virtually no overlap in canonical variate space, whereas the CS strata overlap heavily. When the canonical variates are plotted against geographic position (Fig. 5.19), no clear trends emerge. The eigenvalues from the canonical variate analysis are presented in Table 5.12.

Table 5.12. Summary statistics of canonical variate analysis on the 6 recent samples.

Canonical Variate	Eigenvalue	Proportion	Cumulative
1	1.576	0.411	0.411
2	1.066	0.278	0.689
3	0.740	0.193	0.882
4	0.352	0.092	0.973
5	0.103	0.027	1.000

Hypothesis Tests

Using the same MANOVA and bootstrap tests as above, highly significant differences were detected between all pairs of recent samples, even after application of the Bonferonni correction (Table 5.13). The lack of correspondence between geographic separation and morphologic divergence is demonstrated by the Mahalanobis distances between populations (Table 5.13). Some samples that are closely spaced geographically are not too dissimilar from each other (e.g., CAT and CDK), but others are relatively quite divergent (e.g., CDK and PSJ). The distance between the samples from Texas and North Carolina is fairly average (10.874), only one above the median.

Morphologic Differentiation: Time vs. Geography

There is greater morphologic divergence between the 6 recent populations from North Carolina, Florida, and Texas than between the 6 Caloosa Shell strata. The recent samples overlap less in canonical variate space (Fig. 5.18, cf. Fig 5.15), they are all significantly different, and the Mahalanobis distances between populations are greater on average (Table 5.13, cf. Table 5.10). The pooled recent samples can be more variable than the pooled fossil strata, but it is surprising

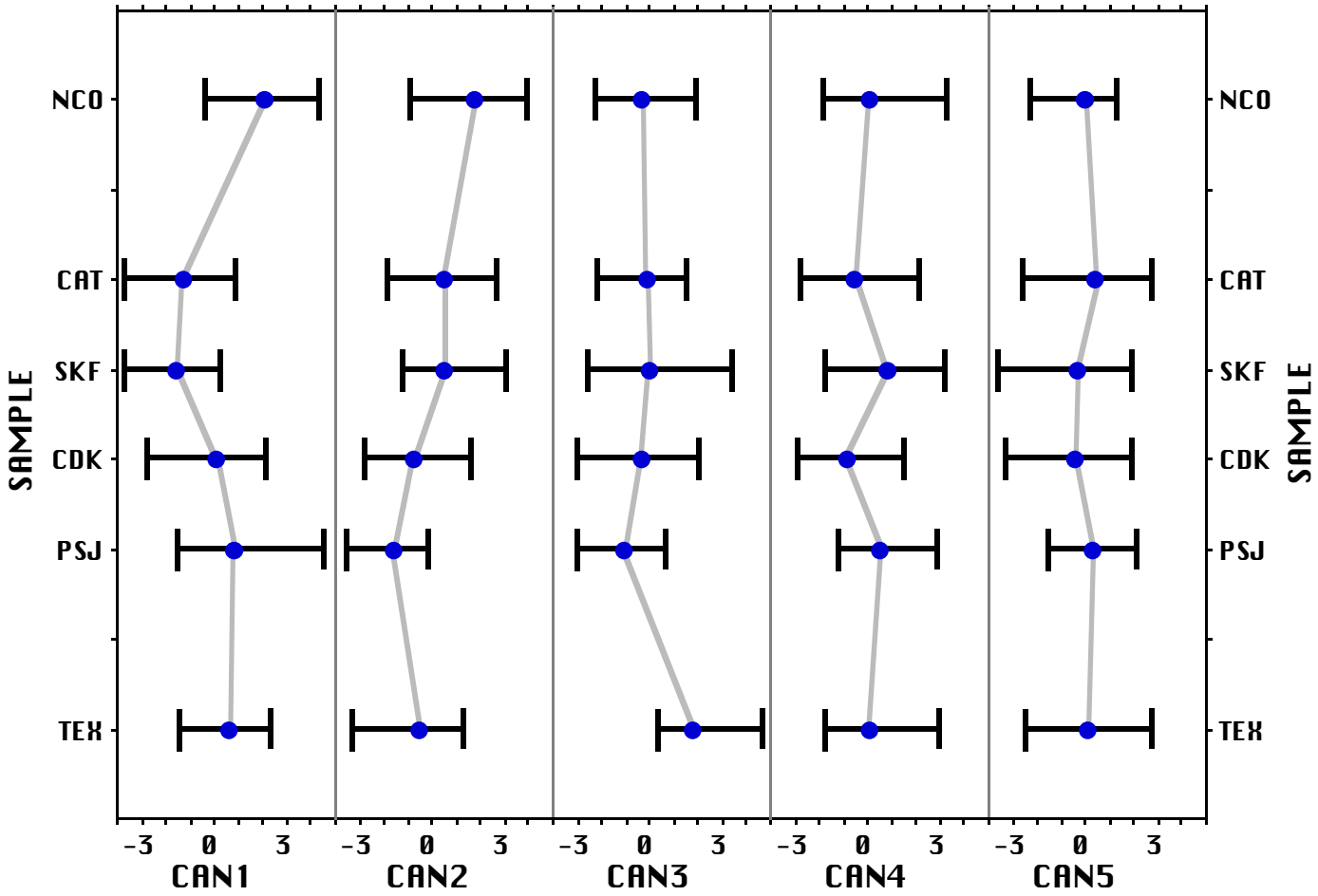

Figure 5.19. Means and ranges of the canonical scores of the recent samples plotted against geographic position.

that there is not more of a difference in the distributions. Perhaps differences in allometry between the recent samples are complicating the analysis.

Table 5.13. Mahalanobis distances between recent samples and tests of significance. The bootstrap test was run for 30,000 iterations. Pairwise comparisons that are significant at α=0.05 are underlined; those that are significant after a Bonferroni correction are also in bold type (total α=0.05, pairwise α=0.0033.

Samples		Mahalanobis Distance (D^2)	Bootstrap p-value	MANOVA p-value
CAT	CDK	4.272	< **0.0001**	**0.0001**
CAT	NCO	13.636	< **0.0001**	**0.0001**
CAT	PSJ	10.818	< **0.0001**	**0.0001**
CAT	SKF	2.512	< **0.0001**	**0.0001**
CAT	TEX	8.924	< **0.0001**	**0.0001**
CDK	NCO	11.378	< **0.0001**	**0.0001**
CDK	PSJ	4.26	< **0.0001**	**0.0001**
CDK	SKF	7.161	< **0.0001**	**0.0001**
CDK	TEX	5.945	< **0.0001**	**0.0001**
NCO	PSJ	13.861	< **0.0001**	**0.0001**
NCO	SKF	16.089	< **0.0001**	**0.0001**
NCO	TEX	12.255	< **0.0001**	**0.0001**
PSJ	SKF	11.451	< **0.0001**	**0.0001**
PSJ	TEX	9.227	< **0.0001**	**0.0001**
SKF	TEX	10.127	< **0.0001**	**0.0001**

Species-Level Morphologic Patterns

The full Generalized Procrustes fit of all specimens is shown in Figs. 5.20. The 6 samples of recent Mercenaria campechiensis and the fossil M. campechiensis from Caloosa Shell are indistinguishable based on the 13 landmarks studied (Fig 5.20c). Not only is morphology stable over hundreds to thousands of years, it has also remained unchanged in M. campechiensis over

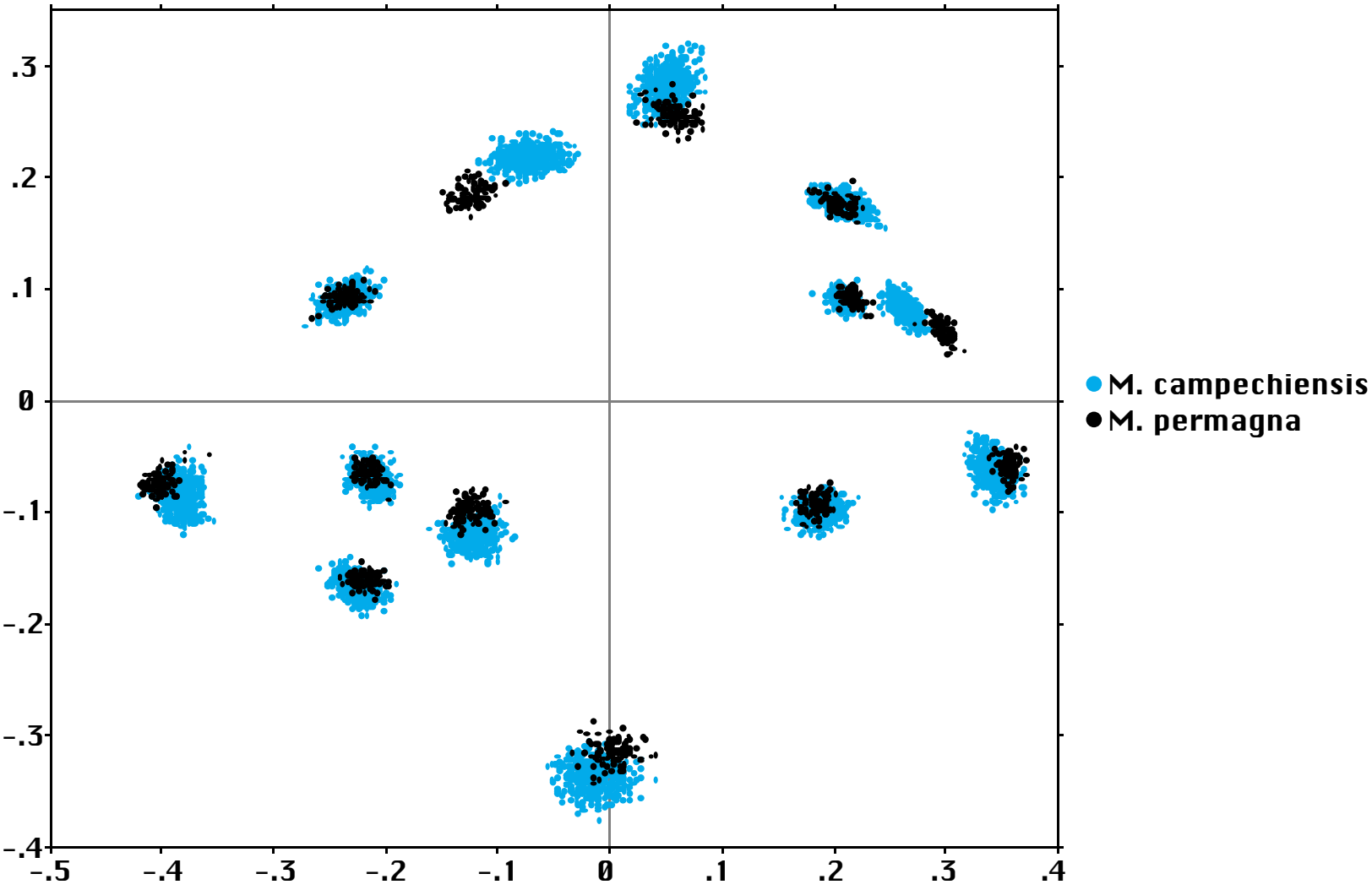

Figure 5.20a. Full Procrustes fit of all samples.

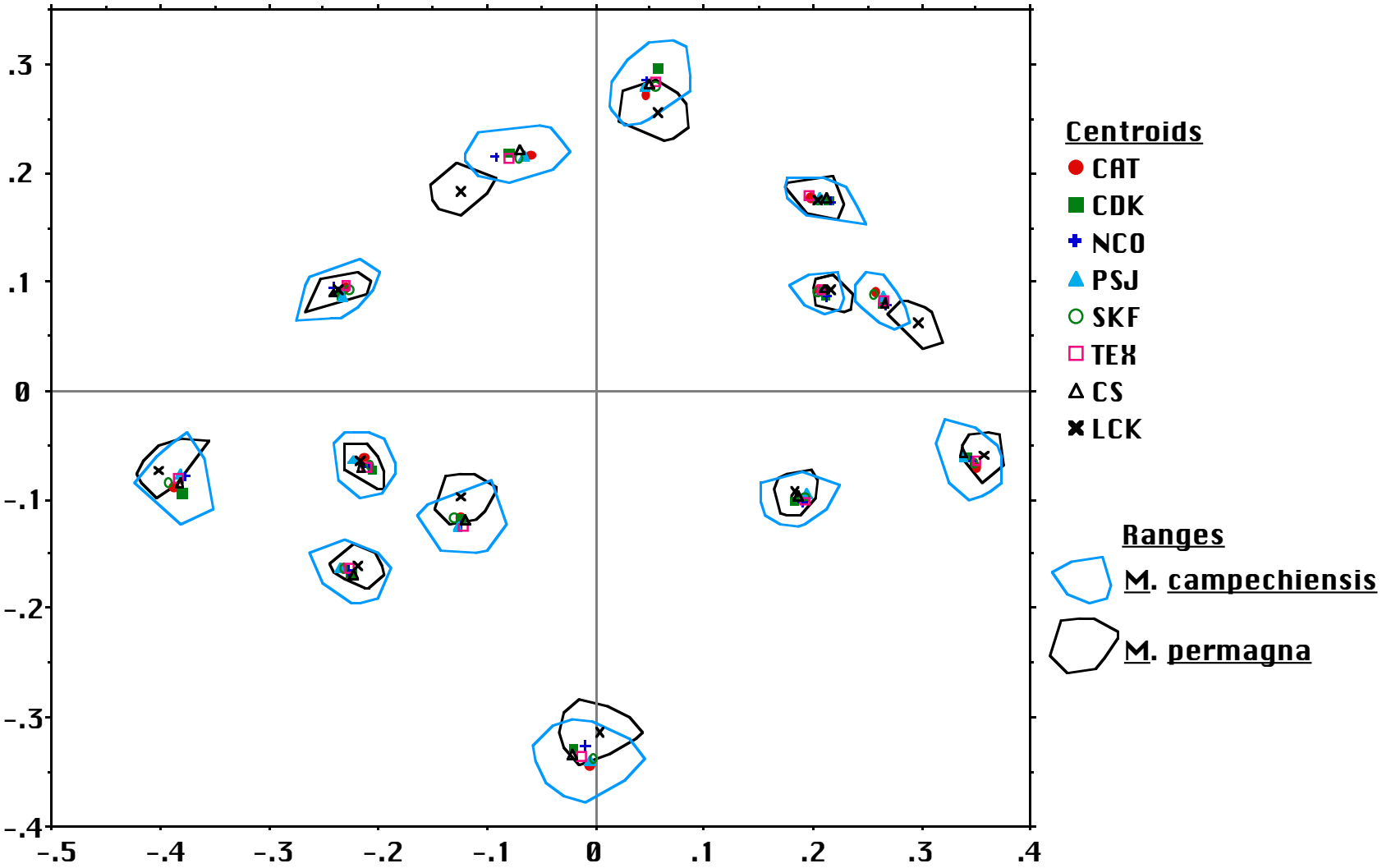

Figure 5.20b. Centroids of each sample's Procrustes coordinates, with convex hulls of each species.

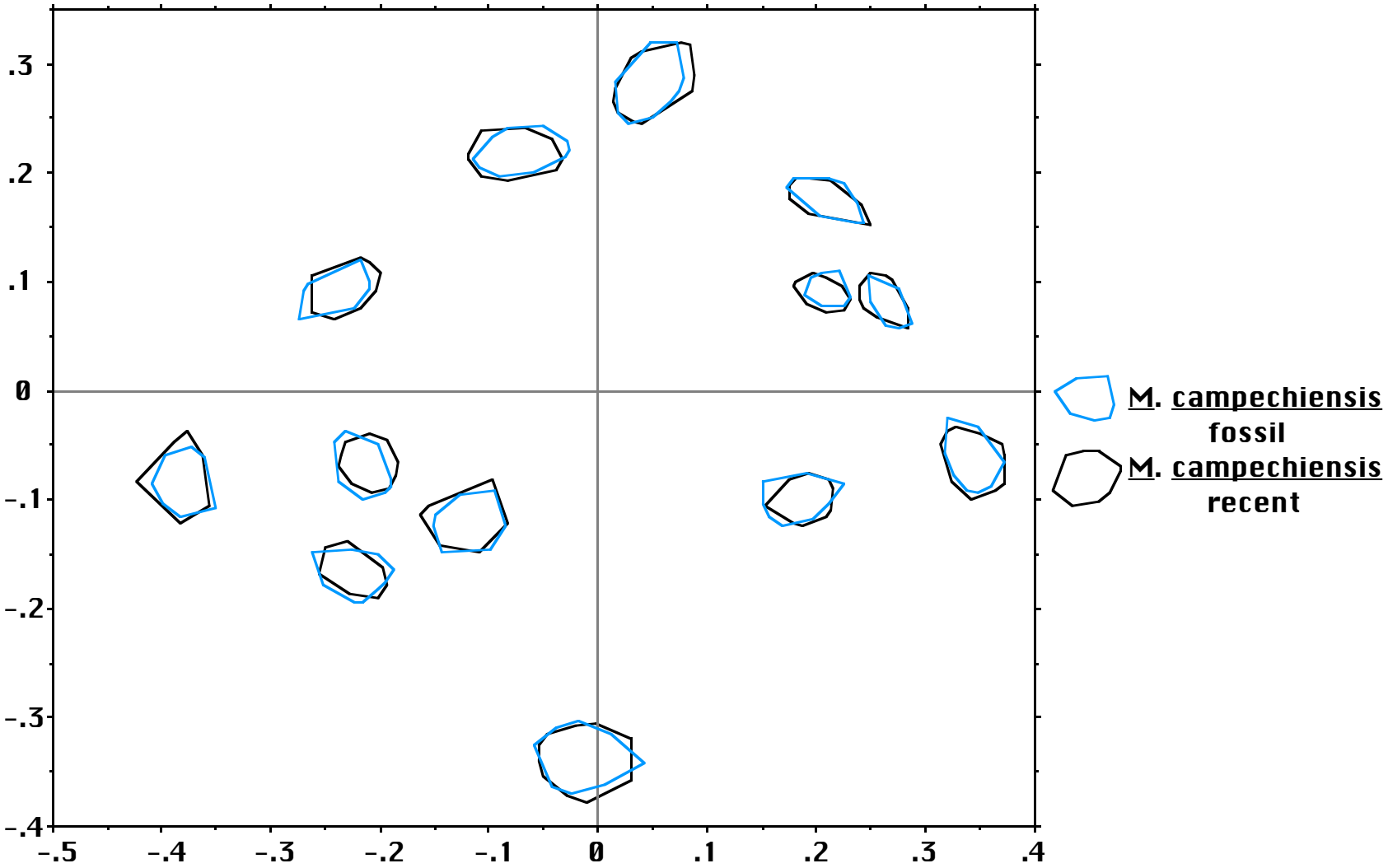

Figure 5.20c. Convex hulls of the Procrustes coordinates of the recent and fossil M. campechiensis, from the fit that included LCK.

half a million years or more. M. permagna, however, is morphologically distinct from M. campechiensis. At the end of the ligament and the end of the lunule, the M. permagna show virtually no overlap with the M. campechiensis samples. The longer ligament and longer lunule M. permagna were noted earlier (Chapter 3). The more elongate form of M. permagna can also be seen in the Procrustes fit: the dorsalmost and ventralmost points on the shell are closer together than on M. campechiensis, and the pseudolandmarks at the anterior and posterior ends of the shell are further apart.

When the partial tangent coordinates for all 621 Mercenaria are run through a principal components analysis, the two species are separated by the first principal component, the axis of maximum variance (Fig. 5.21). A canonical variate analysis (and thus prior knowledge of group membership) is not needed to establish morphologic separation between these two species. M. permagna plots well outside the range of morphology established by multiple populations of M. campechiensis across its modern geographic range; the division of M. permagna and M. campechiensis into separate species is strongly supported.

The M. permagna (LCK) sample was combined pairwise with the 6 recent samples to test the effects of pooling between species. Only half of LCK was used so that sample size would be comparable. The interspecies pooled samples have high variabilities, well beyond the range of standing crop or time-averaged samples (Table 5.14).

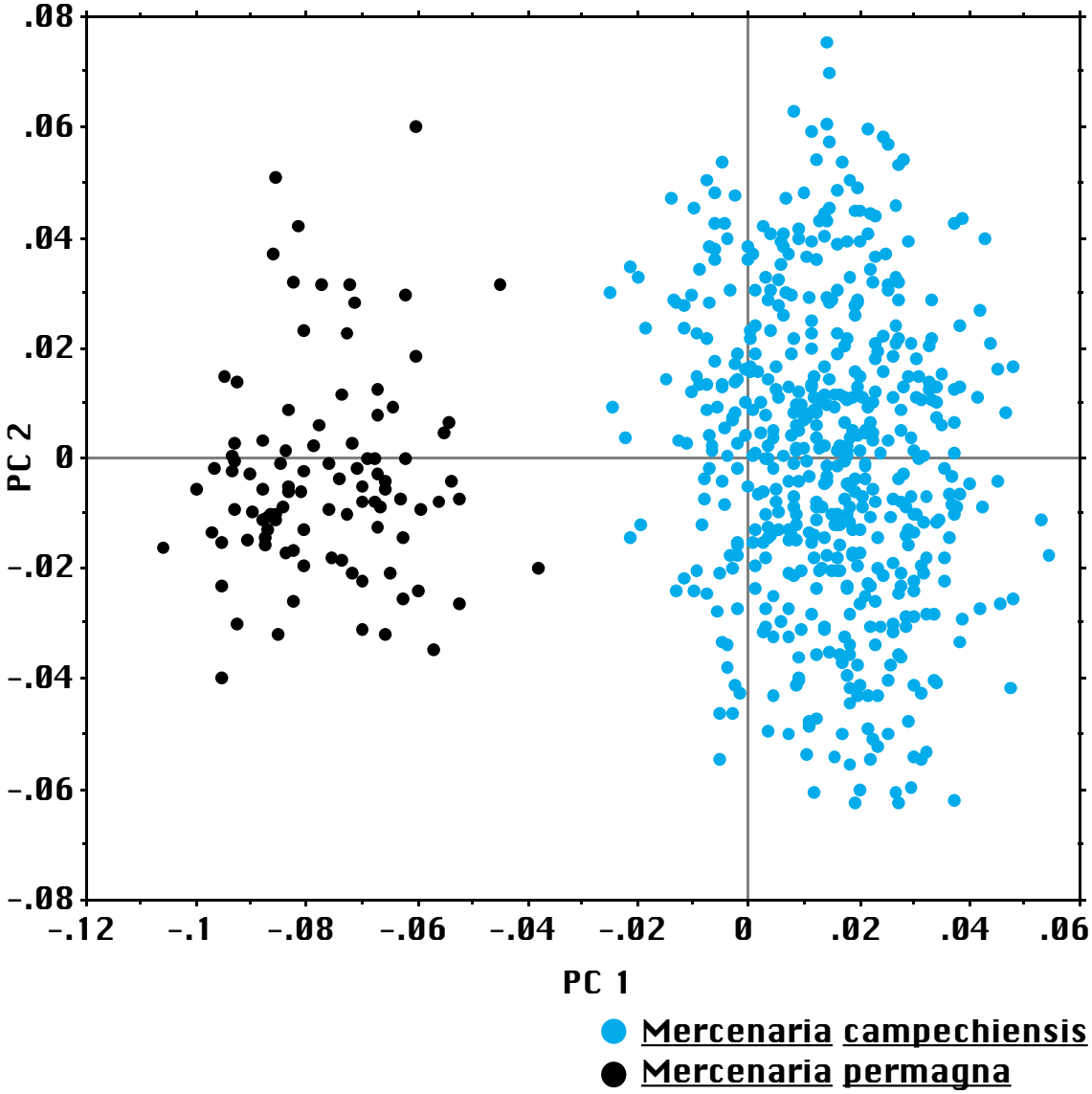

Figure 5.21. Principal components analysis of the partial tangent coordinates of all samples.

Table 5.14. Variability of interspecific pooled samples. Only half of LCK was used in the pooling so that sample sizes were comparable. "Tan. Coor." stands for partial tangent coordinates.

Samples		Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CAT	LCK	0.0684	0.0649	0.0684	0.0647
CDK	LCK	0.0654	0.0527	0.0654	0.0526
NCO	LCK	0.0602	0.0576	0.0602	0.0575
PSJ	LCK	0.0646	0.0558	0.0646	0.0557
SKF	LCK	0.0652	0.0612	0.0652	0.0611
TEX	LCK	0.0609	0.0602	0.0609	0.0600

Summary of Variability Analyses

Morphologic variability at various scales of temporal and spatial pooling is summarized in Fig. 5.22. Samples taken from single living populations have allometry-free variabilities between 0.040 and 0.048 root mean squared Procrustes distance, as do fossil samples collected from single layers at an outcrop. When adjacent layers are analytically time-averaged, the variability values are restricted to the upper half of the range established by the living populations and single beds. Analytical time-averaging never inflates variability much beyond the range of living populations, though it eliminates low values. Analytically time-averaging over greater intervals does not further inflate variability since the slight changes through time in the population were non-directional.

Morphologic differentiation in geographically separated modern samples is greater than differentiation through time at Caloosa Shell. All pairs of modern sample are highly significantly different, and there is less overlap between populations in canonical variate space than was seen with the CS strata. The variabilities of pairwise combinations of the recent samples are somewhat similar to the variabilities of analytically time-averaged samples, but some higher values

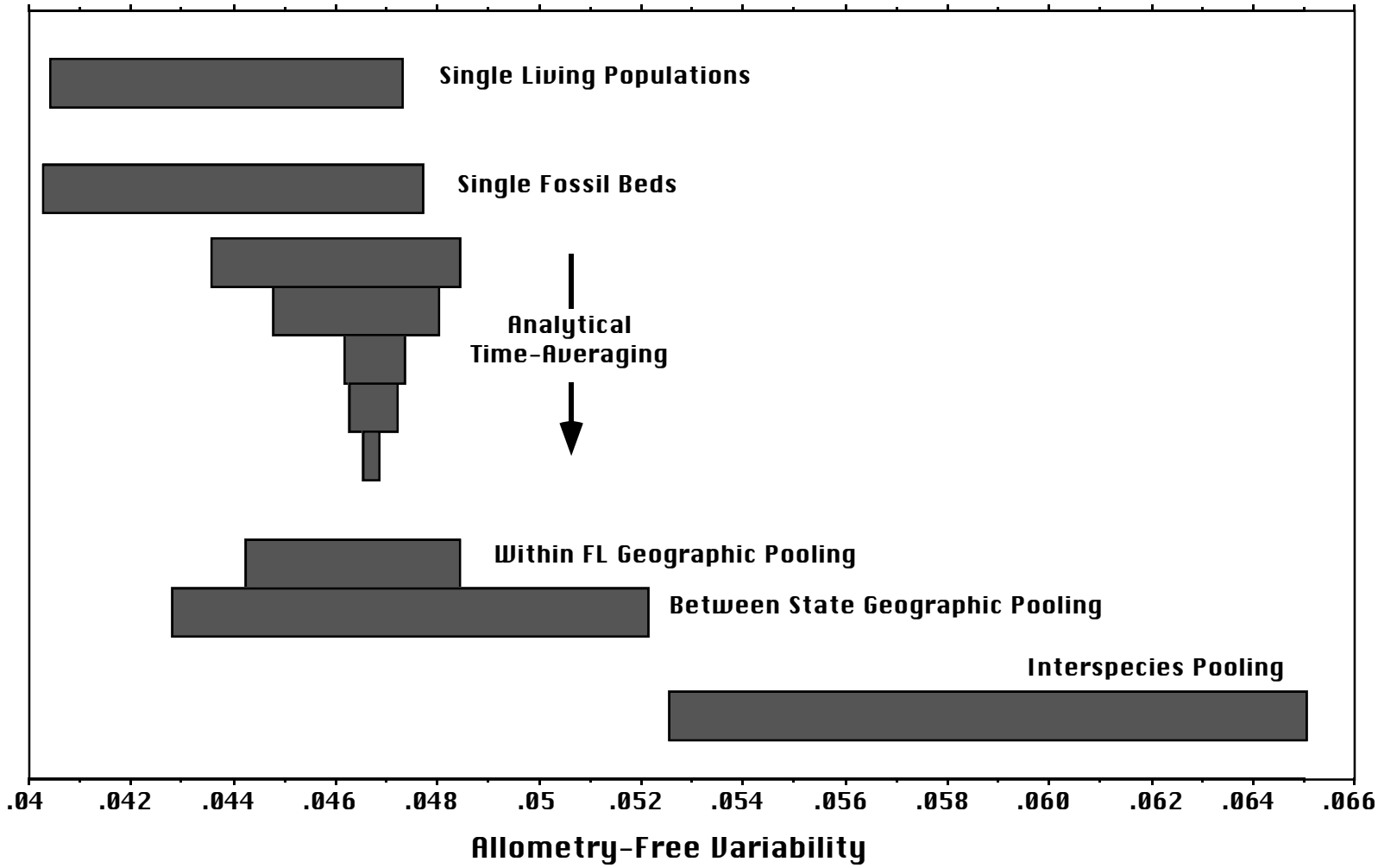

Figure 5.22. Ranges of variability in Mercenaria at various temporal, spacial, and taxonomic scales.

are produced. No aspect of morphology was clearly correlated with the amount of geographic separation.

Mercenaria campechiensis is a morphologically conservative taxon. The fossil Caloosa Shell sample cannot be distinguished from the 6 modern populations, which are quite similar to each other. M. permagna falls well outside the range of morphospace occupied by M. campechiensis. When samples of the two species are mixed, they are more variable than any combination of same-species samples.

Evaluation of Bootstrap Methods

To assess the stability of the bootstrapped confidence intervals, the confidence interval estimates were plotted against iteration number (Fig. 5.23). The estimates are quite unreliable during the first thousand iterations, but are stable after several thousand iterations. Similar plots were examined for every variability estimate on each sample, and all were stable before ten thousand iterations. The bootstrapped confidence interval values are listed in Appendix A.

The bootstrapped estimates of variability were consistently less than the value calculated from the actual data, so the variability estimates and the confidence intervals were adjusted upwards so that the bootstrapped and actual variability estimates matched. The adjustment was small, between 0.25% and 2.5% of the actual values (Table 5.15). Other methods could be used to give a more accurate adjustment, but the improvement in accuracy would have been insignificant (DiCiccio and Romano 1988).

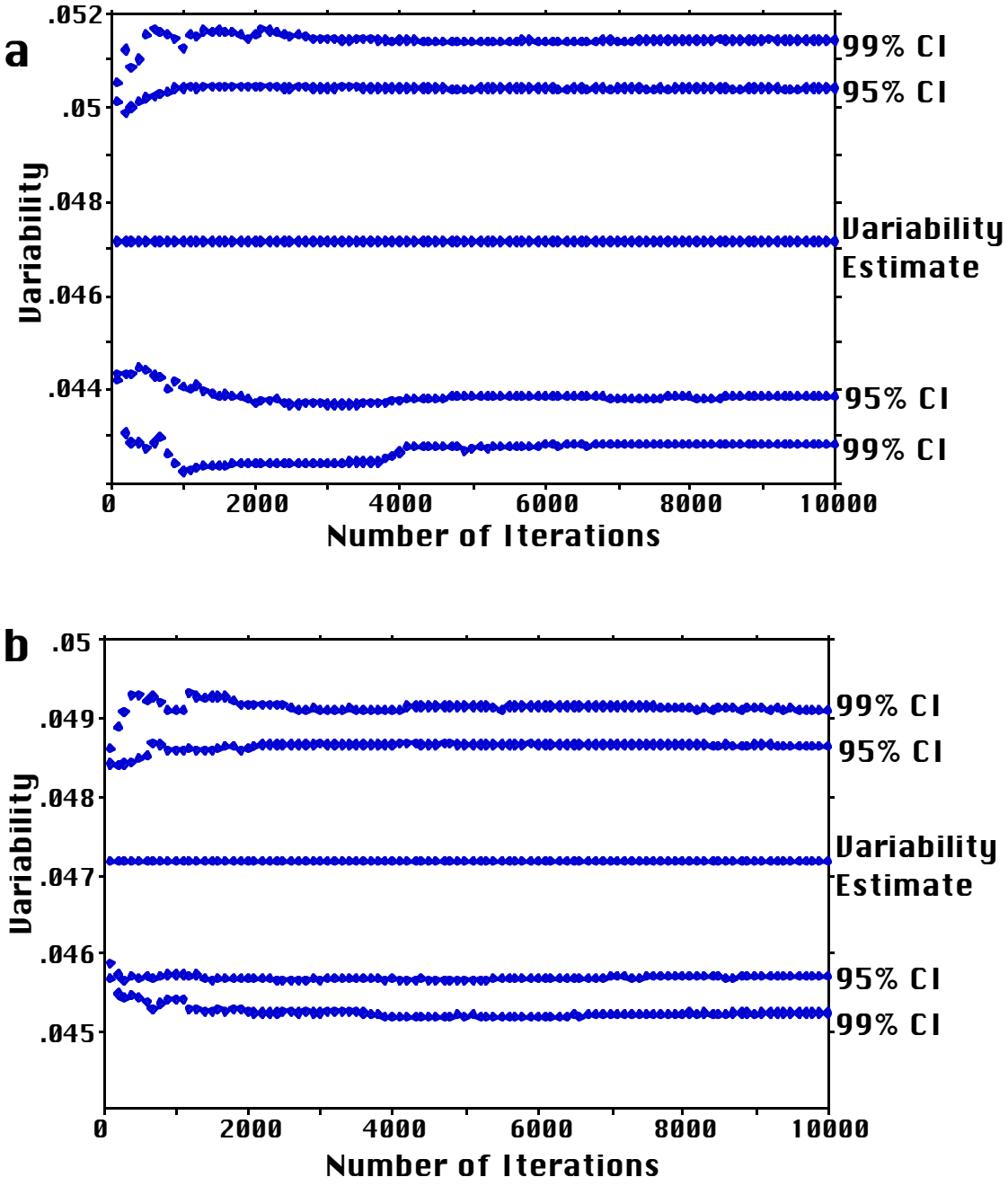

Figure 5.23. Stabilization of bootstrapped confidence intervals. (a) SKF allometry-free tangent coordinates; (b) CS allometry-free tangent coordinates.

Table 5.15. The difference between the bootstrapped and actual estimates of variability, as a percentage of actual variability.

Sample	Tan. Coor. Total	Tan. Coor. Allometry-free	Residuals Total	Residuals Allometry-free
CAT	0.83%	1.60%	0.83%	1.60%
CDK	1.00%	1.85%	1.00%	1.85%
NCO	1.29%	2.42%	1.29%	2.42%
PSJ	1.19%	2.48%	1.19%	2.48%
SKF	0.95%	1.93%	0.95%	1.93%
TEX	1.04%	1.97%	1.04%	1.97%
CS	0.25%	0.47%	0.25%	0.47%
LCK	0.52%	1.14%	0.52%	1.23%

Chapter 6: Discussion

Time-Averaging and Variability

For a single morphologic character, Bell et al. (1987) detected no variability increase in fossil assemblages that were time-averaged over thousands of years. Using a multivariate data set to obtain a comprehensive variability metric, I have likewise found that time-averaged fossil assemblages of Mercenaria drawn from single layers at an outcrop are identical in variability to living populations (Fig. 5.10). The variability of these fossil assemblages could be used without reservation to represent standing-crop, biologic variability.

Very slight changes in the morphology of Mercenaria are evident between some levels at Caloosa Shell (Fig. 5.16), but there were no rapid fluctuations in morphology over the hundreds to thousands of years contained in a single layer. If any changes at all were occurring over this short time-span, then the single fossil layers would not be so similar in variability to the standing crop populations. The fossil samples would be restricted to the upper half of the range of variabilities found in the living populations, with no low values. This is the pattern seen when pairs of adjacent strata are mixed (Fig. 5.14), *even though some pairs are not demonstrably different in morphology*. Changes occurring during the time-averaging of single levels must be even less than the undetectable differences between strata for such low variability values to occur.

The effects of analytical time-averaging on the variability of Mercenaria assemblages are somewhat counterintuitive: a little bit of analytical time-averaging increases the amount of variability in the average sample, but additional time-averaging contributes no additional variability. In fact, the additional time-averaging merely averages the scatter of variability values found in assemblages with less analytical time-averaging. The explanation of this pattern lies in the path of morphologic change through time in the Mercenaria. Morphologic change is too slow to inflate the variability of individual layers, but

it is occurring, and slight changes accumulate between layers in an outcrop. These slight changes are enough that samples pooled from pairs of adjacent strata never have variabilities that are as low as the less variable living populations of the species. The changes are not enough, however, to inflate variability much beyond the more variable standing-crop populations. If the direction of slow morphologic change was constant through time, then additional analytic time-averaging would add more variability. Reversals in the direction of change are common, however, so that fossils a meter apart stratigraphically may be more alike than those separated by half a meter. Additional time-averaging simply picks up additional small fluctuations, without adding variability (Fig. 6.1). Even when fossils from 2.4 m of section are mixed, variability does not exceed that of the most variable living population.

If these results can be generalized to other taxa and other times, then the utility of analytically time-averaged samples for estimating standing-crop variability depends on the accuracy required. If a researcher wanted to test whether two geographically separated, conspecific fossil populations had different variabilities, then analytical time-averaging is a serious problem. Not only may the fossil populations have changed in variability through time, but analytical time-averaging will also obscure the fine-scale differences between populations. If the estimate of variability need only fall within the range of variability values found in living populations, then analytical time-averaging will not significantly distort the results.

As a general rule, can the variability of a time-averaged fossil assemblage be used to represent standing-crop, biological variability? Poorly preserved or deformed fossils are the least reliable source of variability information (Bambach 1973). Biological variability should not be casually inferred from organisms known to display significant ecophenotypy; species that have mobile, ephemeral populations and significant interpopulational variability are more likely to show increases due to migration. Lumping of geographically separated samples is especially risky, as shown here by pooling recent samples. Fossils should be

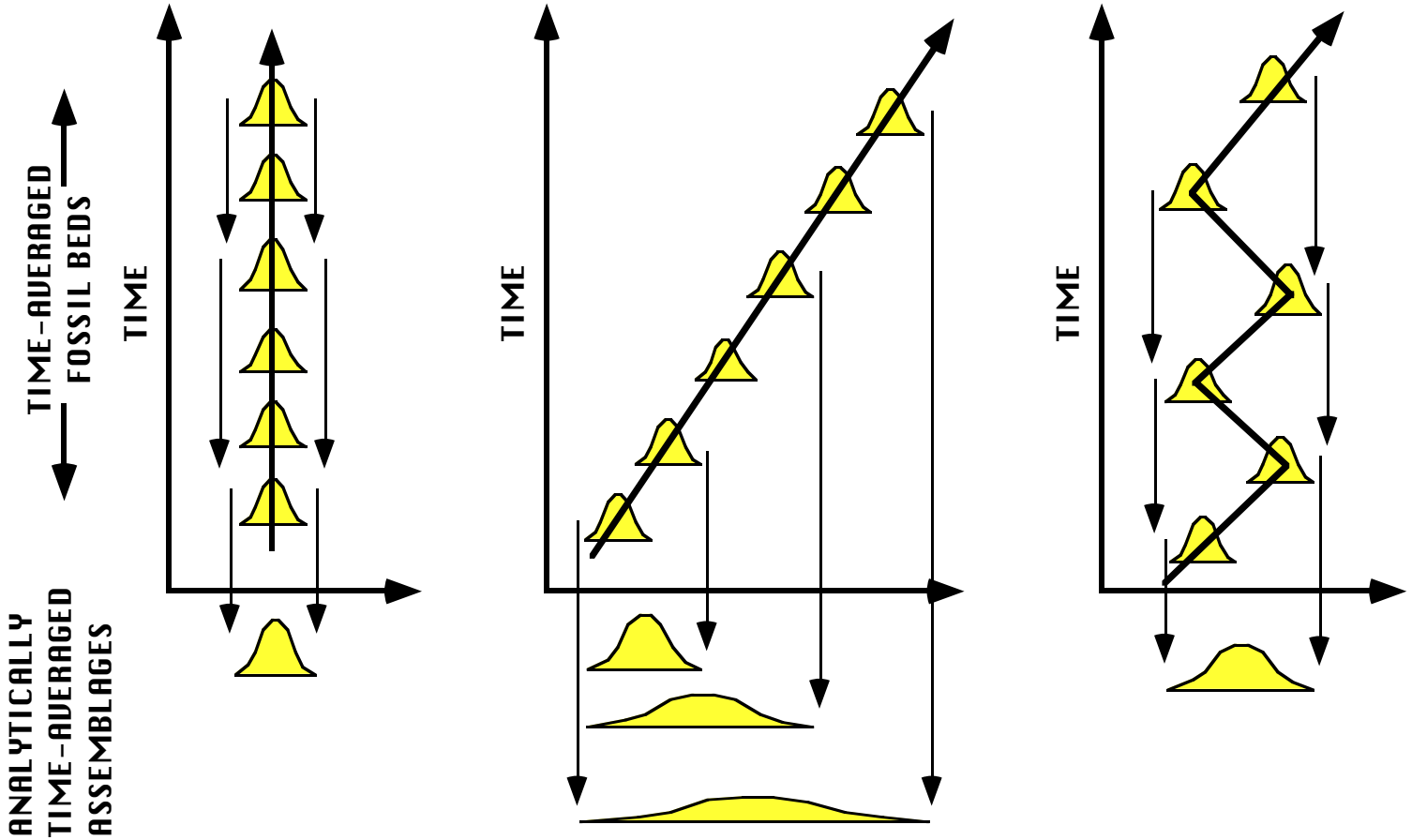

Figure 6.1. The effects of analytical time-averaging on variability. If a population is in perfect stasis (a), then an analytically time-averaged sample will have the same variability as a sample from a single bed, regardless of the amount of sample mixing. A long-term trend in morphologic change (b) will produce more variable samples as the amount of analytical time-averaging increases. If morphology fluctuates, then analytically time-averaged samples will have more variability than single bed samples, but variability will not increase further as analytical time-averaging increases.

collected randomly with respect to the features of interest. The severity of the above factors can be assessed by careful collection, knowledge of a taxon's biology, and common sense. If independent evidence indicates that a lineage is rapidly evolving (as in Williamson 1981a, 1981b, 1986), then the variability of fossils should not be equated with biologic variability if time-averaging is likely. Unfortunately, attempts to relate rate of evolution to variability in fossil material will always be complicated if time-averaged assemblages are used.

Barring ecophenotypy, migration, preservation, and collection effects, the impact of time-averaging on morphologic variability depends on the typical rate of evolution within local populations over hundreds to thousands of years. My results and those of Bell et al. (1987) indicate that evolution can be slow over this time frame, slow enough that variability is not inflated. These results contradict a common expectation that time-averaging should lead to an increase in variability. For example, Kidwell and Flessa (1995) suggest that "variation in a single population of hardparts is likely to be less than among the hardparts derived from many populations." Bell et al. (1987) themselves find it "surprising" that time-averaging did not increase the variability of their fossil stickleback fish. In a study of evolution in land snails, Chiba (1998) states, "If substantial age mixture is present, [time-averaged] samples should show unusually high morphologic variability." The lack of inflated variability is cited as evidence for little time mixing, even though the lineages being studied are mostly static through time.

Gingerich's (1983) influential survey of evolutionary rates is at least partly responsible for the assumption that *some* change should occur in organisms over hundreds to thousands of years. In a literature survey, Gingerich showed that measured rates of morphologic change are high over geologically short time frames and decrease as the measurement interval increases (Fig 6.2). An inverse relationship is actually expected when all samples exhibit stasis, since a constant estimation error term in the numerator is being divided by increasing time intervals. However, for thousands of years or less, rates are high and do not fall

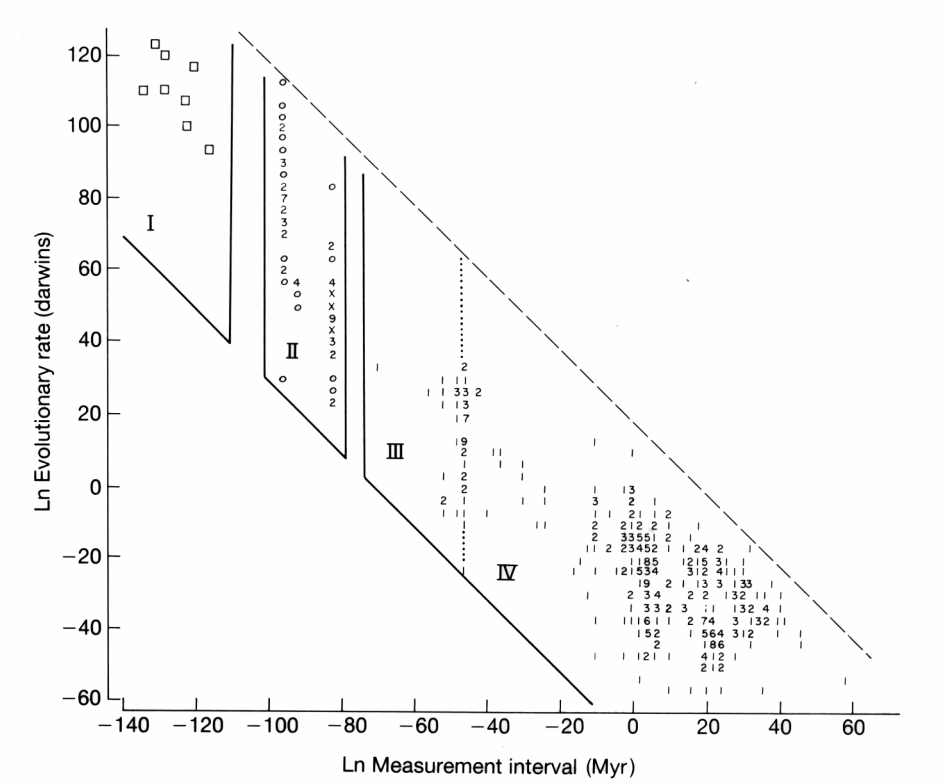

Figure 6.2. Gingerich's (1983) data set. Natural log of evolutionary rate is inversely correlated with natural log of measurement interval. From Ridley (1993).

into the field labeled "No Measurable Change" on Gingerich's (1983) plot. In Gould and Gingerich (1984), Gould criticized the striking inverse relationship as an artifact of human psychology, and asserts that low, short-term rates are missing simply because no one bothers to measure them. Gingerich himself (1983) makes a similar statement: "Organisms differing by more (or less) than 1.2 are so different (or so similar) that they are rarely compared in calculating rates." A careful examination of the data used by Gingerich (1983) reveals the extent of the bias: only populations that would be expected to evolve rapidly are sampled.

Gingerich's (1983) Data Set

The very short term data (1.5 to 10 y) used by Gingerich (1983) are derived from laboratory selection experiments. They are useful as a measure of how fast evolution can occur, but should not be construed as a sample (random or

otherwise) of natural rates. To characterize rates over 70 to 300 y, Gingerich (1983) uses data from several papers describing historical colonization events. In the most applicable of these studies, the original authors compared the morphology of a colonizing population with the morphology of the parent population from which it evolved. In other studies, the source population was not sampled or specifically identified, so the rates calculated are probably less accurate. The times of divergence were established from historical records.

Berry (1964) examined the evolution of a population of Mus musculus introduced onto Stockholm, a small island off the coast of Wales. Berry cites an account of the colonization provided by Lockley (1943, 1947), in which several mice escape onto the island from a boat. Though Berry (1964) considers the details of this story apocryphal, the colonizing population must have been transported on a small boat and must have been very small, though the mice numbered a quarter of a million by 1957. The variables measured were the percent incidence of 35 skeletal characters. Though mainland British mice are fairly uniform in morphology, the Stockholm mice are skeletally distinct from their mainland relatives. Mice sampled from a Scottish island, the Isle of May, were also differentiated from nearby mainland populations. Berry (1964) concludes "that the peculiar features of these island races stem from the chance characteristics of their founder members."

In another study used by Gingerich (1983), Johnston and Selander (1971) studied the morphology of Passer domesticus, the house sparrow. This bird was introduced into North America between 1852 and 1860 from Europe. Samples were taken from across North America and the lengths of 16 skeletal elements were measured. Differentiation of the North American populations was correlated with local environmental conditions. Larger birds are found in locales with lower winter temperatures, as predicted by Bergmann's ecogeographic rule. Bergmann's rule is a well-supported, empirical generalization; by Mayr's (1963) definition it states that "races from cooler climates tend to be larger in species of warm-blooded vertebrates than races of the same species living in warmer

climates." Changes in size alter the surface-area-to-volume ratio of an organism, and Bergmann's rule is often explained by adaptation to retain body heat in cooler climates and shed it more rapidly in warmer areas. The details of this explanation have been debated—for example, James (1970) discusses the role of humidity—but the pattern is supported "very convincingly" (Mayr 1963). Johnston and Selander (1971) also found that sparrows with small bodies and long limbs were found in warm regions, and birds with large bodies and shorter limbs were found in cooler areas. This variation also affects area-to-volume relationships, enhancing thermoregulation according to Allen's ecogeographic rule. The evidence is strong that the sparrows were adapting to new local environments as they spread across the continent.

Johnston and Selander (1971) also compared the North American sparrows to the ancestral European stock. Presumably, this is the data used by Gingerich (1983) to compute evolutionary rates. Johnston and Selander recognize that the most rigorous analysis of the evolution of American populations necessitates comparison with their actual ancestors, European sparrows from 1852. Lacking such samples, Johnston and Selander (1971) use European samples from 1962, but they argue that this substitution matters little:

> Such recent samples of course can have the disadvantage of being phenetically different from ones that might have been taken in 1852, but we suppose that such differences will be minimal, and for two reasons.
>
> First, samples of both sexes of house sparrows from Renthendorf, Germany, taken by Brehm in 1815-1854 and by Piechocki in 1963-1965, have been shown to be statistically identical in lengths of wings and of tails (Niethammer, 1969). Second, house sparrows, "P.d. italiae," on Malta and Crete have been shown (Johnston, 1969) to be essentially invariant in color and pattern over periods of from 35 to about 100 years. We therefore assume that if these aspects of phenotype have remained constant over the time span approximating 1852-1962, other aspects of gross phenotype could also have remained reasonably constant for the same period.
>
> Johnston and Selander (1971)

Though they are investigating evolutionary divergence of sparrows in America, Johnston and Selander (1971) argue for stability in established European populations.

The other data sets used by Gingerich (1983) to describe evolution over 70-300 y also involve historical colonization of islands. Ashton and Zuckerman (1950, 1951) studied the divergence of cranial and dental features of green monkeys from St. Kitts in the West Indies. Descendants of West African monkeys, the St. Kitts monkeys had been isolated for about 300 years; they were compared with modern green monkeys from West Africa. Ashton and Zuckerman (1950, 1951) postulate that the observed differences between the populations result from the actions of natural selection, possibly acting upon general body size.

Patton et al. (1975) investigated divergence of populations of Rattus rattus introduced into the Galapagos Islands. (I'm almost certain that Gingerich used this source, but his citation is incorrect.) The history of roof rat colonization is complex—in a section labeled "Hypothetical Origins," Patton et al. (1975) propose three waves of colonization. The first rats probably came in the late 1600's, possibly from England. Another group of islands was colonized in the 1800's, possibly from South America, and other islands were colonized in the 1930's or 1940's, potentially from California, Hawaii, or Panama. In addition to the Galapagos samples, 20 rats from California were examined. It is unclear how evolutionary rates could be calculated from these data—the ancestor-descendant relationships are foggy, and the putative parental populations were generally not sampled. California is a possible source population for one set of islands, but a maximum of 45 years passed between this colonization event and the publication of Patton et al.'s paper. Gingerich (1983) indicates that all historical colonizations spanned at least 70 years, so this comparison was evidently not made. In any case, this paper describes yet another colonization of a new environment.

The high rates of evolution documented by Gingerich (1983) for periods of 70 to 300 y result from the introduction of small populations into novel environments (Berry 1964; Johnston and Selander 1971; Ashton and Zuckerman 1950, 1951; Patton et al. 1975). Founder effects are likely in some of these introductions, but selection can also operate quite rapidly in such cases (Losos et al. 1997). The sparrows studied by Johnston and Selander (1971) appear to have undergone adaptive evolution as they spread across North America in a series of colonization events. The other studies were introductions onto small islands, where selective pressures are likely to be quite different than on the mainland (which in some cases was thousands of miles away). These introductions provide the opportunity for change to occur, and would likely appear as punctuations in the geologic record.

To characterize evolutionary rates over 1,000 to 10,000 y, Gingerich uses data on "faunal change following Pleistocene glaciation." One of the studies accessed by Gingerich (1983) is a study of the tri-colored squirrel (Callosciurus prevosti) of Southeast Asia. Heaney (1978) measured body size of the squirrels from a variety of islands and found a distinct non-linear pattern of variation between body size and island area. Large squirrels occupy medium-sized islands, but small and large islands are home to smaller squirrels (Heaney 1978, Fig. 2). Heaney postulates that the variation in size is caused by changes in predation, food resources, interspecific competition, and optimum physiological efficiency that correlate with island size.

Calculation of evolutionary rates from these data is problematic. The presumed time of divergence is the last sea level lowstand, when islands were connected, but it is not established that these populations actually diverged at that time. Isolation of the populations for the past 5,000 years is asserted, but it is not established that some of the animals did not diverge much earlier. In addition, the data consist only of scattered modern populations, with no indication of ancestor-descendant relationships. Even if these qualms are ignored, the squirrels were colonizing new island habitats. The pattern of size

variation suggests that body size was evolving selectively to fit the environmental parameters of the new habitat.

Kurtén (1959) presents evolutionary rates in land mammals of Northern Europe, comparing modern to fossil and subfossil remains. Various size variables was measured on over 10 taxa; most showed a decrease in size over several thousand to 20,000 years. Some comparisons were between samples collected at different locations (e.g., 10,000 year old wolverines from "Middle Europe" compared to modern Fennoscandian forms), making the analysis questionable. This aside, Northern Europe has undergone drastic changes in environment and climate over the past ten to twenty thousand years (Donner 1995). Mayr (1963, p. 320) cites changes in European land mammal size since glacial times as an example of Bergmann's rule, and states, "The adaptive nature of these consistent changes cannot be doubted, even though in some of the cases it is not certain whether the changes from the Pleistocene to the present are due to a genetic reconstitution of local populations or to a latitudinal displacement of populations (through migration)." This study does not describe an island colonization event. Rather, it describes an intense climatic change driving organic evolution and/or migration in a predictable way.

These quickly evolving populations are extremely interesting and say much about the process of evolution, but they are hardly a random sample of populations. Selection can operate rapidly when small populations colonize new habitats; in several cases, the original author documented the probable selective pressure that caused differentiation. Selection is not the only cause of evolution in small, colonizing populations; founder effects and drift are effective at small population sizes, and the former was explicitly implicated in one study. The rapid evolution documented by Kurtén (1959) is not induced not by the movement of animals to a new environment, but rather by rapid change of the environment in a particular area. If one wished deliberately to document rapid evolution in natural populations, the above populations are the ones that would be targeted.

How fast is the evolution of large, established populations that are already adapted to local environmental conditions? Johnston and Selander (1971) believed that the established European sparrows that gave rise to the American colonists were changing little. Berry (1964) reports that mainland British mice are fairly homogeneous across the island; are they stable through time as well? Are the mice on the island of Stockholm still changing rapidly, or are they the same as they were in the early 1960's, when Berry's collections were made? Rates in small, colonizing populations that are adapting to new environments cannot be generalized to large populations that may already be optimally adapted to their environments. More research must be done to determine the typical rate of morphologic change within these populations. If rates are typically slow, then time-averaging should often have little influence on morphologic variability.

Lamarck, Cuvier, and Egyptian Mummies

My results indicate that morphology can be stable over hundreds to thousands of years, but this is hardly a new finding: the same result was achieved 200 years ago in the first study of evolutionary rate. Georges Cuvier and Jean-Baptiste Lamarck compared 3,000 year old mummified animals brought back from Egypt by Etienne Geoffroy Saint-Hillaire to similar modern animals. The collection contained cats, dogs, monkeys, ibis (Fig. 6.3), birds of prey, crocodiles, and other animals (Rudwick 1997). Cuvier and Lamarck both judged the mummies identical to modern animals, but were quite at odds in their interpretations. Lamarck (1963) wrote:

> I do not refuse to believe in the close resemblance of these animals with individuals of the same species living to-day. Thus, the birds that were worshipped and embalmed by the Egyptians two or three thousand years ago are still exactly like those which now live in that country.
>
> It would indeed be very odd if it were otherwise; for the position and climate of Egypt are still very nearly what they were in those times. Now the birds which live there, being still in the

> same conditions as they were formerly, could not possibly have been forced into a change of habits.
>
> Lamark (1963)

Lamarck argued that the earth is much older than several thousand years, and that the "imperceptible changing of species" resembles stability when observations are restricted to time spans too narrow (Lamarck 1963).

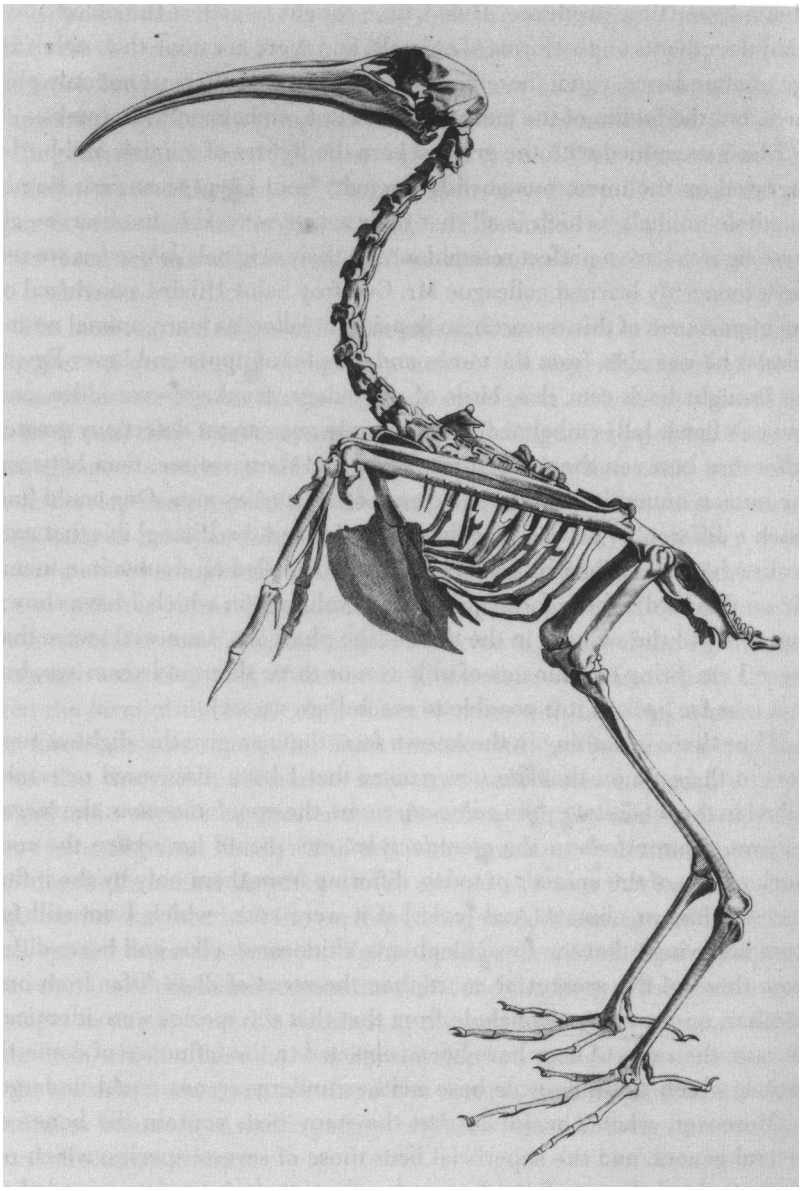

Figure 6.3. The Egyptian ibis (from Rudwick 1997).

Cuvier, on the other hand, took the similarity between the ancient and modern animals at face value. "There is nothing in the known facts that can give the slightest support to the opinion that the new genera that I have discovered or established in the fossil state—the palaeotheriums, the anoplotheriums, the megalonyxes, the mastodons, the pterodactyles, etc.—could have been the root stock of any of the animals today" (Rudwick 1997). To the argument that the history of the earth is long, so that very slow change would appear as stasis, Cuvier replied, "I know that some naturalists rely a lot on the thousands of centuries that they pile up with a stroke of a pen; but in such matters we can hardly judge what a long time would produce, except by multiplying in thought what a lesser time produces." Even if the history of the earth were so long, Cuvier argued, a rate of zero multiplied is still zero.

Short Term Evolution and Evolutionary Theory

The pace of evolution over geologically short periods of time has important implications for evolutionary theory. Despite some counter-examples, it is unambiguous that some species remain unchanged throughout their history (e.g., Stanley and Yang 1987; Cheetham 1986). The speciation aspect of punctuated equilibrium has benefited from solid theoretic explanations since its original proposal, although the importance of particular speciation mechanisms have been debated. That a species can form in a geologically short time span has always been compatible with evolutionary theory. Evolutionary theory did not, however, predict that species would persist unchanged for millions of years. Stasis was simply an empirical observation based on the fossil record, and it lacked the theoretical underpinnings enjoyed by rapid speciation. Ever since, the cause(s) of evolutionary stasis have been debated, with stabilizing selection and developmental homeostasis among the candidates (e.g., Stebbins and Ayala 1981; Eldredge and Gould 1972).

Futuyma (1987) proposes an alternative mechanism: stasis is caused by division of a species into local interbreeding populations, or demes. A deme can evolve and diverge from the rest of its species, Futuyma argues, but eventually the deme will die out or interbreed with another deme, losing its novelties. Since different demes live under different environmental conditions, they will evolve in different directions, and when gene mixing occurs no net change will accrue. In this view, speciation and morphologic change are linked because genetic isolation imparts permanence to the novel features of a population:

> Thus, speciation can facilitate morphologic change not by liberating a population from genetic homeostasis or accelerating the response to selection, but by enabling a gene pool to remain subject to consistent selection pressures even as it moves about in space. By isolating gene pools from other gene pools that they encounter as they move about, speciation enables them to retain characters that evolved in a localized context.
>
> Futurma (1987)

Eldredge (1989) states that "geographically based variability that is seldom cumulatively and significantly modified in any particular direction throughout the subsequent history of an entire species" is "probably the best single theoretical reason to expect net, specieswide stasis" (Eldredge 1989). Gould and Eldredge (1993) also believe that Futuyma's idea may explain why morphological change is concentrated at speciation events.

These mechanisms of stasis entail radically different view of evolution. Stabilizing selection and homeostasis both produce stasis by the *repression of evolutionary change*; the former is an external bridle, the latter an internal property of the organisms. Futuyma's stasis can be described as "laissez-faire stasis," for it allows populations to evolve rapidly and constantly, or not as all, as local conditions allow. The properties of the spacio-temporal architecture of the species are the only factor preventing constant, unchecked divergence.

These theories are also quite different when viewed from a hierarchical perspective on evolution. This view recognizes that evolutionary processes happen on a variety of levels, and that the levels may have emergent properties (properties not explained or predicted by the properties of the smaller-scale levels). Eldredge (1996) calls this hierarchy the "Genealogical Hierarchy," and identifies its levels as germ-line genome, organisms, demes (local interbreeding population), species, and monophyletic taxa (Fig. 6.4). Repressive stasis (stabilizing selection, homeostasis) implicates the cause of stasis at the level of the organism and deme; individual organisms tend not to deviate greatly from the norm because it is disadvantageous. In small populations in unusual environments, selection or homeostatic breakdown allows change to occur. Futuyma's stasis is an emergent property of the species; nothing about the organisms or demes require them to be static, but their arrangement into species compels stasis.

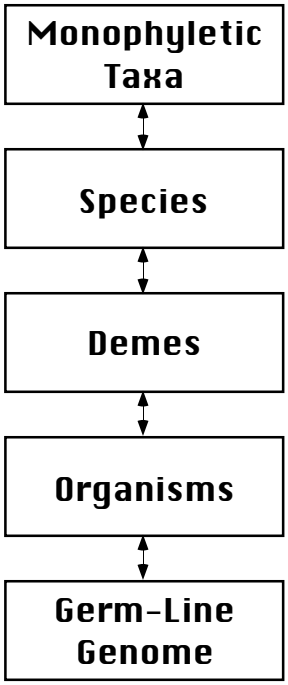

Figure 6.4. The Genealogical Hierarchy. Modified from Eldredge (1996).

Knowing the patterns of deme-level evolution is essential to evaluating the causes of evolutionary stasis. If the deme-level evolutionary rates summarized by Gingerich (1983) are the norm, then laissez-faire stasis is

certainly favored. If these rates simply represent the initial random or adaptational adjustments of small populations to new environments, and demes are typically static, then species-level evolutionary stasis may be caused by more than the simple division of the species into demes. Futuyma (1987) laments that "The fossil record, except where it is unusually detailed, probably records little of the evolution that occurs within populations, because these evolutionary events are spatially highly localized and are obliterated by extinction and interbreeding with other populations over the course of even a few thousand generations." This is unfortunately true, but exceptions like the varved lakes used by Bell et al. (1987) do exist. An attractive source for studying within-deme evolution is the sub-fossil record, where radiocarbon and amino acid racemization permit the necessary time resolution. Anderson et al. (1998) studied paleoecologic changes over thousands of years on the Louisiana coast by collecting from a series of beach ridges. Archeological samples, such as those used by Lamarck and Cuvier, provide another source of datable remains. Studies of variation in time-averaged assemblages can also provide information, but the absolute amount of time involved can be difficult to assess.

With more data on deme-level evolution, a better picture of the spacio-temporal architecture of the species can be established. Do demes evolve constantly, or is rapid geographic differentiation followed by relative morphologic stasis? Rapid change at the founding of a deme followed by stasis resembles the bimodality of evolutionary rate proposed for species by Punctuated Equilibrium; the deme-level model should be called "Punctuated Equilibrium, Jr." If this is indeed the common pattern of evolution, then some truth can be mined from the comments of both Cuvier and Lamarck. Lamarck would be correct that the Egyptian animals did not change because they had no need to change, being already well suited to their environment. But value would also be found in Cuvier's statements—the appearance of new morphologies is not caused by the simple extrapolation through time of the evolutionary rates measured on established populations. Cuvier's assumption that low rates are

universal led to his denial of evolution as the engine of organismic diversity, but modern research clearly negates this assumption.

If established demes do evolve, then many interesting questions can be asked. Do deme-level evolutionary patterns vary from group to group? Can these differences be related to the biology or life history of the species, or to species- or clade-level properties of the group? Are particular aspects of a species' morphology more plastic than others, and are these the aspects that differentiate demes and species? These questions can only be answered when deme-level evolution is better understood.

In Gould and Gingerich (1984), Gingerich asks whether the two classes of evolutionary rate described by punctuated equilibrium are truly different. He suggests that the high "punctuation" rates observed over the short term and the low "stasis" rates observed over the long term are not any different once temporal scaling is considered. When the evolutionary hierarchy is considered, however, short and long term evolutionary rates belong to fundamentally distinct classes. Short term evolutionary rates measure the behavior of single populations following their unique trajectories. Long term rates measure the behavior of an anastomosing complex of multiple populations, and as such are a function of both individual population behavior and the interactions between populations. Understanding patterns of species evolution through geologic time requires the adequate characterization of both short and long term patterns of evolution.

Chapter Seven: Conclusions

1. Time-averaging should inflate the morphologic variability of fossil assemblages over that of living populations if morphologic change is occurring during the interval of time-averaging. Time-averaging can also alter the correlations between morphologic characters and obscure allometric relationships.

2. Collections of fossil Mercenaria from single levels at an outcrop are no more variable than samples drawn from standing-crop, living populations. The rate of morphologic change during the intervals of time-averaging was very nearly zero. The single-level collections can be used to represent standing-crop population variability.

3. Analytical time-averaging of two adjacent layers eliminates the low variability values found in some individual layers, but does not inflate variability significantly beyond the range of values typical of single layers or living populations. Further analytical time-averaging does not further increase variability, since the small changes in morphology between some layers are non-directional. In Mercenaria from Caloosa Shell, variability changes through time and analytically time-averaged collections cannot be used to represent standing-crop variability. However, analytically time-averaged samples still yield variability values that are typical for the species, which for some research questions may be completely adequate.

4. Allometry slightly inflates variability in samples with large size ranges, though this can be corrected statistically. Samples with large shells have greater variability than those with small shells, even when the data are rescaled to the same size. Measurement error, taphonomy, and the pooling of right and left valves were not important influences on variability.

5. Some slight morphologic differences in Mercenaria existed between the strata at Caloosa Shell, but these were exceeded by the differences between geographically separated modern populations. At Caloosa Shell, morphological difference between samples is not correlated with stratigraphic separation. Increased geographic separation in the modern samples is also not correlated with morphologic divergence.

6. Mercenaria campechiensis from the Pleistocene of Florida fall within the range of variability of 6 modern populations, demonstrating stasis over half a million years or more. M. permagna from the Pleistocene of North Carolina are a distinct morphotype from all measured populations of M. campechiensis, so a separate species designation is reasonable.

7. Studies of evolutionary rate over hundreds to thousands of years have been focused on colonization events. As such, our database of these rates is heavily biased towards situations where high rates would be expected. Short term morphologic stasis is evident in the time-averaged samples of Mercenaria, and more research may reveal that established populations in stable environments tend not to evolve. This model, in which the rapid evolution of novelty in a population is followed by relative stasis, is the deme-level equivalent of punctuated equilibrium. The two differ considerably, since demes are not isolated genetic entities. Further study of deme-level evolutionary rates should shed light on the causes of species-level stasis, and may lead to many other questions about the process and pattern of evolutionary change.

References Cited

Abbott, R. T. 1974. American Seashells. Van Nostrand Reinhold, New York.

Anderson, L. C., R. A. McBride, M. J. Taylor, and M. R. Byrnes. 1998. Late Holocene record of community replacement preserved in time-averaged molluscan assemblages, Louisiana chenier plain. Palaios 13:488-499.

Arnold, W. S., T. M. Bert, D. C. Marelli, H. Cruz-Lopez, and P. A. Gill. 1996. Genotype-specific growth of hard clams (genus Mercenaria) in a hybrid zone: variation among habitats. Marine Biology 125:129-139.

Ashton, E. H., and S. Zuckerman. 1950. The influence of geographic isolation on the skull of the green monkey (Cercopithecus aethiops sabaeus). I. Proceedings of the Royal Society of London. 137B:211-238.

Ashton, E. H., and S. Zuckerman. 1951. The influence of geographic isolation on the skull of the green monkey (Cercopithecus aethiops sabaeus). II, III, IV. Proceedings of the Royal Society of London. 138B:204-218, 354-374.

Bader, R. S. 1954. Variability and evolutionary rate in the oreodonts. Evolution 9:119-140.

Bambach, R. K. 1973. Tectonic deformation of composite-mold fossil Bivalvia (Mollusca). American Journal of Science 273-A:409-430.

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

Berry, R. J. 1964. The evolution of an island population of the house mouse. Evolution 18:468-483.

Bert, T. M., D. M. Hesselman, W. S. Arnold, W. S. Moore, H. Cruz-Lopez, D. C. Marelli. 1993. High frequency of gonadal neoplasia in a hard clam (Mercenaria spp.) hybrid zone. Marine Biology 117:97-104.

Bert, T. M., and W. S. Arnold. 1995. An empirical test of predictions of two competing models for the maintenance and fate of hybrid zones: both models are supported in a hard-clam hybrid zone. Evolution 49:276-289.

Blackwelder, B. W. 1981. Late Cenozoic stages and molluscan zones of the U.S. middle Atlantic Coastal Plain. Paleontological Society Memoir 12:1-34.

Bookstein, F. L. 1991. Morphometric tools for landmark data: geometry and biology. Cambridge University Press, New York.

Bookstein, F. L. 1996. Combining the tools of geometric morphometrics. Pp. 131-151 *in* L. F. Marcus, M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. Advances in Morphometrics. NATO ASI Series A: Life Sciences vol. 284. Plenum, New York.

Chapman, R. E. 1994. Morphometric variation in fossil populations: the effect of environmental variability and taphonomic time-averaging. Abstracts with Programs-Geological Society of America 26:486.

Charlesworth, B. 1984. Some quantitative methods for studying evolutionary patterns in single characters. Paleobiology 10:308-318.

Cheetham, A. H. 1986. Tempo of evolution in a Neogene bryozoan: rates of morphologic change within and across species boundaries. Paleobiology 12:190-202.

Chiba, S. 1998. Synchronized evolution in lineages of land snails in oceanic islands. Paleobiology 24:99-108.

Cutler, A. H., and K. W. Flessa. 1990. Fossils out of sequence: computer simulations and strategies for dealing with stratigraphic disorder. Palaios 5:227-235.

Cutler, A. H. 1994. Taphonomic implications of shell surface textures in Bahia la Choya, northern Gulf of California. Palaeogeography, Palaeoclimatology, Palaeoecology 114:219-240.

Daley, G. M. 1999. Environmentally controlled variation in shell size of Ambonychia Hall (Mollusca: Bivalvia) in the type Cincinnatian (Upper Ordovician). Palaios In press.

Daley, G. M. 1993. Morphological variability in Onniella, Rafinesquina (Brachiopoda: Articulata), and Ambonychia (Mollusca: Pelecypoda) through changing environments in the Cincinnatian Series (Upper Ordovician). Unpublished M.S. Thesis, University of Cincinnati.

Davies, D. J., E. N. Powell, and R. J. Stanton, Jr. 1989. Relative rates of shell dissolution and net sediment accumulation—a commentary: can shell beds form by the gradual accumulation of biogenic debris on the sea floor? Lethaia 22:207-212.

DiCiccio, T. J., and J. P. Romano. 1988. A review of bootstrap confidence intervals. Journal of the Royal Statistical Society B (Methodological) 50:338-354.

Dillon, R. T. Jr., and J. J. Manzi. 1989. Genetics and shell morphology of hard clams (genus Mercenaria) from Laguna Madre, Texas. Nautilus 103:73-77.

Driscoll, E. G. 1970. Selective bivalve shell destruction in marine environments, a field study. Journal of Sedimentary Petrology 40:898-905.

Dryden, I. L., and K. V. Mardia. 1998. Statistical Shape Analysis. John Wiley & Sons, Chichester.

DuBar, J. R. 1974. Summary of the Neogene stratigraphy of southern Florida. Pp. 206-231 *in* R. Q. Oaks, Jr., and J. R. DuBar, eds. Post-Miocene stratigraphy, central and southern Atlantic Coastal Plain. Utah State University Press, Logan.

Eldredge, N. 1989. Macroevolutionary dynamics: species, niches, and adaptive peaks. McGraw Hill, New York.

Eldredge, N. 1996. Hierarchies in macroevolution. Pp. 42-61 *in* D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. Evolutionary Paleobiology. University of Chicago Press, Chicago.

Eldredge, N., and S. J. Gould. 1972. Punctuated equilibria: an alternative to phyletic gradualism. Pp. 82-115 *in* T. J. M. Schopf, ed. Models in Paleobiology. Freeman, Cooper & Company, San Francisco.

Emiliani, C. 1950. Introduction to a method for determining the physical characters of fossil environments. Journal of Paleontolgy 24:485-491.

Eversole, A. G. 1997. Gametogenesis of Mercenaria mercenaria, M. campechiensis, and their hybrids. Nautilus 110:107-110.

Eversole, A. G., and P. B. Heffernan. 1995. Gonadal neoplasia in northern Mercenaria mercenaria (Linnaeus, 1758) and southern M. campechiensis

(Gmelin, 1791) quahogs and their hybrids cultured in South Carolina. Journal of Shellfish Research 14:33-39.

Flessa, K. W. 1993. Time-averaging and temporal resolution in Recent marine shelly faunas. Pp. 9-33 *in* S. M. Kidwell and A. K. Behrensmeyer, eds. Taphonomic approaches to time resolution in fossil assemblages. Paleontological Society Short Courses in Paleontology No. 6. University of Tennessee, Knoxville.

Flessa, K. W., A. H. Cutler, K. H. Meldahl, and F. T. Fürsich. 1989. Taphonomic processes and stratigraphic disorder. Abstract, 8th International Geological Congress, Washington, D.C., p. 1-493—1-494.

Flessa, K. W., A. H. Cutler, and K. H. Meldahl. 1993. Time and taphonomy: quantitative estimates of time-averaging and stratigraphic disorder in a shallow marine habitat. Paleobiology 19:266-286.

Flessa, K. W., and M. Kowalewski. 1994. Shell survival and time-averaging in nearshore and shelf environments: estimates from the radiocarbon literature. Lethaia 27:153-165.

Fürsich, F. T. 1978. The influence of faunal condensation and mixing on the preservation of fossil benthic communities. Lethaia 11:243-250.

Fürsich, F. T., and M. Aberhan. 1990. Significance of time-averaging for paleocommunity analysis. Lethaia 23:143-152.

Futuyma, D. J. 1987. On the role of species in anagenesis. American Naturalist 130:465-473.

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

Goodall, C. 1991. Procrustes methods in the statistical analysis of shape. Journal of the Royal Statististical Society B 53:285-339.

Gould, S. J., and P. D. Gingerich. 1984. Smooth curve of evolutionary rate: a psychological and mathematical artifact: discussion and reply. Science 226:994-996.

Gould, S. J., and N. Eldredge. 1993. Punctuated equilibrium comes of age. Nature 366:223-227.

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

Harte, M. E. 1992. An eastern Pacific Mercenaria and notes on other chionine genera (Bivalvia: Veneridae). Veliger 35:137-140.

Harte, M. E. 1998. The evolution of Mercenaria Schumacher, 1817 (Bivalvia: Veneridae). Pp. 305-315 *in* P. A. Johnston and J. W. Haggart, eds. Bivalves: an eon of evolution—paleobiological studies honoring Norman D. Newell. University of Calgary Press, Calgary.

Heaney, L. R. 1978. Island area and body size of insular mammals: evidence from the tri-colored squirrel (Callosciurus prevosti) of southeast Asia. Evolution 32:29-44.

Hughes, N. C., and R. E. Chapman. 1995. Growth and variation in the Silurian proetide trilobite Aulacopleura konincki and its implications for trilobite palaeobiology. Lethaia 28:333-353.

Hulbert, R. C., Jr., and G. S. Morgan. 1989. Stratigraphy, paleoecology, and vertebrate fauna of the Leisey Shell Pit Local Fauna, early Pleistocene

(Irvingtonian) of southwestern Florida. Papers in Florida Paleontology 2:1-19.

Hulbert, R. C., Jr., G. S. Morgan, and S. D. Webb (eds). 1995a. Paleontology and geology of the Leisey Shell Pits, Early Pleistocene of Florida. Part I. Bulletin of the Florida Museum of Natural History 37 Pt. 1.

Hulbert, R. C., Jr., G. S. Morgan, and S. D. Webb (eds). 1995b. Paleontology and geology of the Leisey Shell Pits, Early Pleistocene of Florida. Part II. Bulletin of the Florida Museum of Natural History 37 Pt. 2.

James, F. C. 1970. Geographic size variation in birds and its relationship to climate. Ecology 51:365-390.

Johnston, R. F. 1969. Taxonomy of house sparrows and their allies in the Mediterranean basin. Condor 71:129-139.

Johnston, R. F., and R. K. Selander. 1971. Evolution in the house sparrow. II. Adaptive differentiation in North American populations. Evolution 25:1-28.

Jones, D. S., P. A. Mueller, T. Acosta, and R. D. Shuster. 1995. Strontium isotopic stratigraphy and age estimates for the Leisey Shell Pit faunas, Hillsborough County, Florida. Bulletin Florida Museum of Natural History 37:93-105.

Kendall, D. G. 1984. Shape manifolds, Procrustean metrics, and complex projective spaces. Bulletin of the London Mathematical Society 16:81-121.

Kendall, D. G. 1986. Comment on F. L. Bookstein, Size and shape spaces for landmark data in two dimensions. Statistical Science 1:222-226.

Kent, J. T. 1994. The Complex Bingham Distribution and Shape Analysis. Journal of the Royal Statistical Society Series B 56:285-299.

Kershaw, P. J., D. J. Swift, and D. C. Denoon. 1988. Evidence of recent sedimentation in the eastern Irish Sea. Marine Geology 85:1-14.

Kidwell, S. M. 1986. Models for fossil concentrations: paleobiologic implications. Paleobiology 12:6-24.

Kidwell, S. M., and T. Aigner. 1985. Sedimentary dynamics of complex shell beds: implications for ecologic and evolutionary patterns. Pp. 382-395 *in* U. Bayer and A. Seilacher, eds. Sedimentary and Evolutionary Cycles. Springer Verlag; Berlin.

Kidwell, S. M., and D. W. J. Bosence. 1991. Taphonomy and time-averaging of marine shelly faunas. Pp. 115-209 *in* P. Allison and D. E. G. Briggs, eds. Taphonomy: releasing the data locked in the fossil record. Plenum, New York.

Kidwell, S. M., and K. W. Flessa. 1995. The quality of the fossil record: populations, species, and communities. Annual Review of Ecology and Systematics 26:269-99.

Kowalewski, M. 1996. Time-averaging, overcompleteness, and the geological record. Journal of Geology 104:317-326.

Kowalewski, M., G. A. Goodfriend, and K. W. Flessa. 1998. High-resolution estimates of temporal mixing within shell beds: the evils and virtues of time-averaging. Paleobiology 24:287-304.

Kurtén, B. 1959. Rates of evolution in fossil mammals. Cold Spring Harbor Symposia on Quantitative Biology 25:205-215.

Lamarck, J. B. 1963. Zoological Philosophy. (English translation by H. Elliot). Hafner Publishing Company, New York.

Lande, R. 1976. Natural selection and random genetic drift in phenotypic evolution. Evolution 30:314-334.

Lockley, R. M. 1943. Dream island days: a record of the simple life. Witherby, London.

Lockley, R. M. 1947. Letters from Stockholm. Dent, London.

Losos, J. B., K. I. Warheit, and T. W. Schoener. 1997. Adaptive differentiation following experimental island colonization in Anolis lizards. Nature 387:70-73.

Lyons, W. G. 1991. Post-Miocene species of Latirus Montfort, 1810 (Mollusca: Fasciolariidae) of southern Florida, with a review of regional marine biostratigraphy. Bulletin of the Florida Museum Natural History, Biological Science 35:131-308.

MacFadden, B. J. 1995. Magnetic Polarity stratigraphy and correlation of the Leisey Shell Pits, Tampa Bay, Hillsborough County, Florida. Bulletin of the Florida Museum of Natural History 37:107-116.

Marcus, L. 1990. Traditional Morphometrics. Pp. 77-122 *in* F. J. Rohlf and F. L. Bookstein, eds. Proceedings of the Michigan Morphometrics Workshop. University of Michigan Museum of Zoology Special Publication 2.

Marcus, L. F., M. Corti, A. Loy, G. J. P. Naylor, and D. E. Slice, eds. 1996. Advances in Morphometrics. NATO ASI Series A: Life Sciences vol. 284. Plenum, New York.

Mayr, E. 1963. Animal species and evolution. Harvard University Press, Cambridge, Massachusetts.

Meldahl, K. H. 1987. Sedimentologic and taphonomic implications of biogenic stratification. Palaios 2:350-358.

Meldahl, K. H., K. W. Flessa, and A. H. Cutler. 1997. Time-averaging and postmortem skeletal survival in benthic fossil assemblages: quantitative comparisons among Holocene environments. Paleobiology 23:207-229.

Miller, A. I. 1988. Spatial resolution in subfossil molluscan remains: implications for paleobiological analyses. Paleobiology 14:91-103.

Niethammer, G. 1969. Vergleich der Renthendorfer Haussperlinge von heute mit einer von C. L. Brehm von 110 Jahren gesammelten Serie. J. f. Ornith. 110:205-208.

Patton, J. L., and P. V. Brylski. 1987. Pocket gophers in alfalfa fields: causes and consequences of habitat-related body size variation. American Naturalist 130:493-506.

Patton, J. L., S. Y. Yang, and P. Myers. 1975. Genetic and morphologic divergence among introduced rat populations (Rattus rattus) of the Galapagos Archipelago, Ecuador. Systematic Zoology 24:296-310.

Portell, R. W., K. S. Schindler, and G. S. Morgan. 1992. The Pleistocene molluscan fauna from Leisey Shell Pit 1, Hillsborough County, Florida. Pp. 181-194 *in* T. M. Scott and W. D. Allmon, eds. The Plio-Pleistocene stratigraphy and paleontology of southern Florida. Florida Geological Survey, Special Publication 36.

Portell, R. W., K. S. Schindler, and D. Nicol. 1995. Biostratigraphy and paleoecology of the Pleistocene invertebrates from the Leisey Shell Pits, Hillsborough County, Florida. Bulletin of the Florida Museum of Natural History 37:127-164.

Ridley, M. 1993. Evolution. Blackwell Science, Cambridge, Mass.

Robinson, B. W., and D. S. Wilson. 1995. Experimentally induced morphological diversity in Trinidadian guppies (Poecilia reticulata). Copeia 1995:294-305.

Rohlf, F. J. 1990. Morphometrics. Annual Review of Ecology and Systematics 21:299-316.

Rohlf, F. J., and F. L. Bookstein, eds. 1990. Proceedings of the Michigan Morphometrics Workshop. University of Michigan Museum of Zoology Special Publication 2.

Rohlf, F. J., and L. F. Marcus. 1993. A revolution in morphometrics. Trends in Ecology and Evolution 8:129-132.

Rudwick, M. J. S. 1997. Georges Cuvier, fossil bones, and geological catastrophes. University of Chicago Press, Chicago.

Smith, L. H. 1998. Species level phenotypic variation in lower Paleozoic trilobites. Paleobiology 24:17-36.

Stanley, S. M., and X. Yang. 1987. Approximate evolutionary stasis for bivalve morphology over millions of years: a multivariate, multilineage study. Paleobiology 13:113-139.

Staff, G. M., and E. N. Powell. 1988. The paleoecologic significance of diversity: the effect of time averaging and differential preservation on macroinvertebrate species richness in death assemblages. Palaeogeography, Palaeoclimatology, Palaeoecology 63:73-89.

Stebbins, G. L., and F. J. Ayala. 1981. Is a new evolutionary synthesis necessary? Science 213:967-971.

Walker, K. R., and R. K. Bambach. 1971. The significance of fossil assemblages from fine-grained sediments: time-averaged communities. Geological Society of America Abstracts with Programs 3:783-784.

Ward, L. W., and B. W. Blackwelder. 1987. Late Pliocene and early Pleistocene Mollusca from the James City and Chowan River Formation at the Lee Creek Mine. Pp. 113-283 *in* C. E. Ray, ed. Geology and Paleontology of the Lee Creek Mine, North Carolina, II. Smithsonian Contributions to Paleobiology 61.

Webb, S. D., G. S. Morgan, R. C. Hulbert, Jr., D. S. Jones, B. J. MacFadden, and P. A. Mueller. 1989. Geochronology of a rich early Pleistocene vertebrate fauna, Leisey Shell Pit, Tampa Bay, Florida. Quaternary Researcy 32:96-110.

Williamson, P. G. 1981a. Paleontological documentation of speciation in Cenozoic molluscs from Turkana Basin. Nature 293:437-443.

Williamson, P. G. 1981b. Morphological stasis and developmental constraint: real problems for neo-Darwinism. Nature 294:214-215.

Williamson, P. G. 1986. Selection or constraint?: a proposal on the mechanism for stasis. Pp. 129-142 *in* K. S. Campbell, ed. Rates of Evolution. Allen and Unwin; London.

Zullo, V. A., and W. B. Harris. 1992. Sequence stratigraphy of the marine Pliocene and lower Pleistocene deposits in southwestern Florida: preliminary assessment. Pp. 27-40 *in* T. M. Scott and W. D. Allmon, eds. The Plio—Pleistocene stratigraphy and paleontology of southern Florida. Florida Geological Survey, Special Publication 36.

Appendix A:
Confidence Intervals Around Variability Estimates

Table A.1. Confidence intervals around variability estimates. "RMS dist." indicates actual root mean squared distance to the mean, "Tan. Coor." indicates partial Procrustes Tangent Coordinates, and "Residuals" indicates Procrustes residuals.

Sample	Percentile	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CAT	RMS dist.	0.04774	0.04774	0.04257	0.04253
	0.5%	0.04406	0.04406	0.03919	0.03915
	2.5%	0.04487	0.04487	0.04006	0.04002
	97.5%	0.05062	0.05062	0.04508	0.04503
	99.5%	0.05163	0.05163	0.04587	0.04582
CDK	RMS dist.	0.04657	0.04657	0.04480	0.04475
	0.5%	0.04269	0.04269	0.04131	0.04128
	2.5%	0.04347	0.04347	0.04207	0.04203
	97.5%	0.04959	0.04959	0.04770	0.04764
	99.5%	0.05057	0.05057	0.04873	0.04867
NCO	RMS dist.	0.04647	0.04647	0.04587	0.04582
	0.5%	0.04189	0.04189	0.04112	0.04108
	2.5%	0.04302	0.04302	0.04232	0.04228
	97.5%	0.04991	0.04991	0.04934	0.04928
	99.5%	0.05099	0.05099	0.05037	0.05030
PSJ	RMS dist.	0.04296	0.04296	0.04056	0.04053
	0.5%	0.03811	0.03811	0.03634	0.03631
	2.5%	0.03918	0.03918	0.03731	0.03728
	97.5%	0.04695	0.04695	0.04384	0.04379
	99.5%	0.04806	0.04806	0.04486	0.04481

Sample	Percentile	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
SKF	RMS dist.	0.04906	0.04906	0.04716	0.04710
	0.5%	0.04418	0.04418	0.04281	0.04277
	2.5%	0.04548	0.04548	0.04383	0.04379
	97.5%	0.05266	0.05266	0.05041	0.05034
	99.5%	0.05365	0.05365	0.05142	0.05135
TEX	RMS dist.	0.04213	0.04213	0.04085	0.04081
	0.5%	0.03833	0.03833	0.03704	0.03701
	2.5%	0.03927	0.03927	0.03804	0.03801
	97.5%	0.04489	0.04489	0.04358	0.04354
	99.5%	0.04578	0.04578	0.04438	0.04434
LCK	RMS dist.	0.04421	0.04421	0.04140	0.04136
	0.5%	0.04075	0.04075	0.03867	0.03864
	2.5%	0.04154	0.04154	0.03929	0.03926
	97.5%	0.04707	0.04707	0.04360	0.04356
	99.5%	0.04794	0.04794	0.04420	0.04415
CS 6/7	RMS dist.	0.05087	0.05087	0.04546	0.04541
	0.5%	0.04587	0.04587	0.04106	0.04102
	2.5%	0.04696	0.04696	0.04219	0.04215
	97.5%	0.05492	0.05492	0.04842	0.04836
	99.5%	0.05602	0.05602	0.04919	0.04912
CS 8	RMS dist.	0.04861	0.04861	0.04280	0.04275
	0.5%	0.04351	0.04351	0.03887	0.03884
	2.5%	0.04480	0.04480	0.03985	0.03981
	97.5%	0.05233	0.05233	0.04562	0.04557
	99.5%	0.05335	0.05335	0.04641	0.04635

Sample	Percentile	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CS 9	RMS dist.	0.04931	0.04931	0.04324	0.04319
	0.5%	0.04300	0.04300	0.03686	0.03682
	2.5%	0.04459	0.04459	0.03859	0.03856
	97.5%	0.05406	0.05406	0.04751	0.04745
	99.5%	0.05547	0.05547	0.04876	0.04869
CS 10	RMS dist.	0.05202	0.05202	0.04760	0.04754
	0.5%	0.04634	0.04634	0.04308	0.04304
	2.5%	0.04742	0.04742	0.04420	0.04416
	97.5%	0.05651	0.05651	0.05083	0.05076
	99.5%	0.05788	0.05788	0.05177	0.05170
CS 11	RMS dist.	0.04916	0.04916	0.04719	0.04714
	0.5%	0.04506	0.04506	0.04307	0.04303
	2.5%	0.04598	0.04598	0.04406	0.04402
	97.5%	0.05252	0.05252	0.05038	0.05032
	99.5%	0.05343	0.05343	0.05141	0.05134
CS 12/13	RMS dist.	0.04528	0.04528	0.04042	0.04039
	0.5%	0.04059	0.04059	0.03629	0.03625
	2.5%	0.04159	0.04159	0.03734	0.03731
	97.5%	0.04900	0.04900	0.04298	0.04294
	99.5%	0.05003	0.05003	0.04365	0.04360
CS all	RMS dist.	0.05036	0.05036	0.04673	0.04668
	0.5%	0.04816	0.04816	0.04478	0.04474
	2.5%	0.04868	0.04868	0.04528	0.04523
	97.5%	0.05201	0.05201	0.04820	0.04814
	99.5%	0.05249	0.05249	0.04866	0.04860

Sample	Percentile	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CS a	RMS dist.	0.04542	0.04237	0.04542	0.04232
	0.5%	0.04162	0.03878	0.04162	0.03874
	2.5%	0.04239	0.03953	0.04239	0.03950
	97.5%	0.04839	0.04520	0.04839	0.04515
	99.5%	0.04933	0.04603	0.04933	0.04597
CS b	RMS dist.	0.04450	0.04394	0.04450	0.04389
	0.5%	0.04095	0.04042	0.04095	0.04038
	2.5%	0.04176	0.04122	0.04176	0.04118
	97.5%	0.04726	0.04658	0.04726	0.04653
	99.5%	0.04812	0.04729	0.04812	0.04724
CS c	RMS dist.	0.04687	0.04630	0.04687	0.04625
	0.5%	0.04352	0.04306	0.04352	0.04302
	2.5%	0.04437	0.04386	0.04437	0.04382
	97.5%	0.04923	0.04864	0.04923	0.04858
	99.5%	0.04994	0.04930	0.04994	0.04924
CS d	RMS dist.	0.05030	0.04935	0.05030	0.04929
	0.5%	0.04580	0.04513	0.04580	0.04509
	2.5%	0.04684	0.04611	0.04684	0.04606
	97.5%	0.05382	0.05256	0.05382	0.05248
	99.5%	0.05495	0.05359	0.05495	0.05351
CS a&d	RMS dist.	0.05290	0.04666	0.05290	0.04660
	0.5%	0.04959	0.04371	0.04959	0.04366
	2.5%	0.05031	0.04438	0.05031	0.04432
	97.5%	0.05548	0.04896	0.05548	0.04889
	99.5%	0.05630	0.04970	0.05630	0.04964

Sample	Percentile	Tan. Coor. Total	Residuals Total	Tan. Coor. Allometry-free	Residuals Allometry-free
CS b&c	RMS dist.	0.04701	0.04585	0.04701	0.04580
	0.5%	0.04461	0.04341	0.04461	0.04337
	2.5%	0.04523	0.04409	0.04523	0.04405
	97.5%	0.04876	0.04760	0.04876	0.04754
	99.5%	0.04931	0.04814	0.04931	0.04808
LCK a	RMS dist.	0.03946	0.03703	0.03946	0.03700
	0.5%	0.03651	0.03401	0.03651	0.03398
	2.5%	0.03723	0.03468	0.03723	0.03466
	97.5%	0.04162	0.03934	0.04162	0.03931
	99.5%	0.04230	0.04011	0.04230	0.04007
LCK b	RMS dist.	0.04613	0.04106	0.04613	0.04102
	0.5%	0.04116	0.03713	0.04116	0.03710
	2.5%	0.04237	0.03810	0.04237	0.03807
	97.5%	0.05008	0.04405	0.05008	0.04400
	99.5%	0.05135	0.04491	0.05135	0.04486

Appendix B: SAS Programs

Root Mean Squared Procrustes Variability with Bootstrapped Confidence Intervals

```
*****************;
%let sizemax=9999;          /*  These variables are used to select the data
%let sizemin=0;                 to be analyzed */
%let sam='CS';
%let strmax=20;
%let strmin=-1;
%let colmax=20;
%let colmin=-1;
%let step=1;                /*  Interval for recording CI estimates */
%let times=1;               /*  Number of iterations for bootstrap */
*****************;

data merc1;
     infile 'MercRaw.dat';
     input num sample $ strata col size UMBOX UMBOY LIGX LIGY PRETX
               PRETY PADDX PADDY PSINX PSINY PSOUTX PSOUTY AADDX
               AADDY ARETX ARETY LUNX LUNY DORX DORY POSTX
               POSTY VENTX VENTY ANTX ANTY;
     if sample=&sam;
     if &strmin <= strata <= &strmax;
     if &colmin <= col <= &colmax;
     if &sizemin <= size <= &sizemax;
     run;

proc print;
     var sample num size;
     run;
```

```
data merc2;
      set merc1;
      drop num sample strata col size;
      run;

 proc IML;

/*****************************************************************************
        Shape1R arranges a matrix into one row (eg x1, y1, x2, y2,..., xk, yk)
*****************************************************************************/
start Shape1R(input,output);
      output=shape(input,1,0);
finish Shape1R;

/*****************************************************************************
        Shape2C arranges a matrix into two columns (eg, an X and a Y column)
*****************************************************************************/
start Shape2C(input,output);
      output=shape(input,0,2);
finish Shape2C;


/*****************************************************************************
        Vectorize takes the landmark data in Input and arranges it into a vertical
vector of the form (x1, x2,..., xk, y1, y2,..., yk)T.  Input can be arranged as a single
row (x1, y1, x2, y2,..., xk, yk) or as an x column and a y column.  See Dryden and
Mardia (1998, p. 76).
*****************************************************************************/
start Vectoriz(Input, Output);
      run Shape2C(input,X);
      do i=1 to ncol(X);
```

```
      NewVec=NewVec//X[,i];
    end;
    Output=NewVec;
finish Vectoriz;


/*****************************************************************************
      InvVect does the inverse of Vectorize.  Output is a row.  See Dryden and
Mardia (1998, p. 76).
*****************************************************************************/
start InvVect(Input, Output);
    XInput=Input;
    n=nrow(XInput)/2;
    Y=XInput[1:n]||XInput[(n+1):2*n];
    run Shape1R(Y, Output);
finish InvVect;


/*****************************************************************************
      Center takes the configurations in Shapes and places the centroid of each
at the origin.  Both Shapes and Outshaps are a matrix with one configuration in
each row, arranged as x1, y1, x2, y2,..., xk, yk. See Dryden and Mardia (1998, p.
24, eq. 2.3).
*****************************************************************************/
start Center(Shapes, Outshaps);
    X=Shapes;
    k=ncol(X)/2;
    C=i(k)-(1/k)*j(k,k,1);
    do i=1 to nrow(X);
      Y1=shape(X[i,],0,2);
      Y2=C*Y1;
      Y3=shape(Y2,1,0);
```

```
      Ymat=Ymat//Y3;
    end;
    Outshaps=Ymat;
finish Center;


/*****************************************************************************
     CSize calculates the Centroid Size of the configurations in Shapes. Shapes
is a matrix with one configuration in each row, arranged as x1, y1, x2, y2,..., xk,
yk.  Sizes is a vertical array containing the sizes.  Shapes must be centered.  See
Dryden and Mardia (1998, p. 24).
*****************************************************************************/
start CSize(Shapes, Sizes);
    XShapes=Shapes;
    Do i=1 to nrow(XShapes);
      S=sqrt(ssq(XShapes[i,]));
      Svec=Svec//S;
    end;
    Sizes=Svec;
finish CSize  ;

/*****************************************************************************
     CentScal centers the configurations in Rawcoor and scales them to unit
size.  Rawcoor and Preshaps (the output) are matrices with one configuration in
each row, arranged as x1, y1, x2, y2,..., xk, yk.  CentScal calls the subroutines
Center and CSize.
*****************************************************************************/
start CentScal(Rawcoor, Preshaps, Sizes);
    XRawcoor=Rawcoor;
    run Center(XRawcoor, Centcoor);
    run CSize(Centcoor, Sizes);
    Do i=1 to nrow(Centcoor);
```

```
      Y=Centcoor[i,]/Sizes[i];
      Yvec=Yvec//Y;
   end;
   Preshaps=Yvec;
finish CentScal;


/*****************************************************************************
      ProcMean finds the Procrustes mean configuration for the configurations
in Shapes.  Shapes is a matrix with one configuration in each row, arranged as x1,
y1, x2, y2,..., xk, yk.  Shapes must first be run through CentScal before ProcMean
will work.  Mean is the mean configuration and is a horizontal array in the same
form as Shapes.  The mean configuration is rotated so that Ant and Post
Retractors have the same y coordinate.  See Dryden and Mardia (1998, p. 44).
The actual math used comes from Kent (1994, 289-290).
*****************************************************************************/
start ProcMean(Shapes, Mean);
   XShapes=Shapes;            /* XShapes is a dummy variable */
   nshapes=nrow(XShapes);     /* nshapes = number of shapes being fitted  */
   k=ncol(XShapes)/2;         /* k = number of landmarks */
   VZ=t(XShapes);             /* vertical vectors (x1, y1,...,xk, yk) */
   T=j(2*k,2*k,0);            /* initialize T matrix */

   do i=1 to nshapes;
      Y1=shape(VZ[,i], 0, 2);
      Y2=-1*Y1[,2] || Y1[,1];
      ViZ=shape(Y2,0,1);
      Ttemp=VZ[,i]*t(VZ[,i])+ViZ*t(ViZ);
      T=T+Ttemp;
   end;
   Tevecs=eigvec(T);
   A=shape(Tevecs[,1],0,2);
```

```
    Angle=-atan((A[8,2]-A[3,2])/(A[8,1]-A[3,1]));
    If A[3,1] > A[8,1]
      then Angle = Angle - arcos(-1);
    c=cos(Angle);
    s=sin(Angle);
    ns=-sin(Angle);
    top=c||s;
    bottom=ns||c;
    Rotate=top//bottom;
    B=A*Rotate;
    Mean=shape(B,1,0);
finish ProcMean;


/*****************************************************************************
        OPA performs an full ordinary procrustes analysis fit between
configurations Move and Fixed.  Move and Fixed are centered configurations
that are entered as horizontal arrays of the form x1, y1, x2, y2,..., xk, yk.  Move is
rotated and scaled to fit Fixed.  Output is in the same form as the input.  The
algorithm should work for data in higher dimensions, though the "shape" lines
would have to be altered.  Also returns partial*** Procrustes tangent coordinates
of Move with Fixed as the pole, though Fixed must be scaled to unit size for this
to be meaningful.  See Dryden & Mardia (1998) p. 84-85 for OPA, 76-77 for
tangent coordinates.
*****************************************************************************/
start OPA(Move,Fixed,Result,TanCoor);
    run Shape2c(Move,X1);
    run Shape2c(Fixed,X2);
    NormX1=sqrt(trace(X1`*X1));
    NormX2=sqrt(trace(X2`*X2));
    A=(X2`*X1)/(NormX1*NormX2);
    Call SVD(u,q,v,A);
```

```
    Rot=v*u`;
    Scale=trace(X2`*X1*Rot)/trace(X1`*X1);
    FNewX1=Scale*X1*Rot;
    PNewX1=X1*Rot;
    run Shape1R(FNewX1,Result);
    run Vectoriz(X2,Pole);
    run Vectoriz(PNewX1,VMoved);
    Vf=(I(nrow(pole))-pole*pole`)*VMoved;
    run InvVect(Vf, TanCoor);
finish OPA;


/*****************************************************************************
      GPA performs a full General Procrustes Analysis on Shapes.  Shapes is a
set of 2-dimensional landmark configurations, with one configuration in each
row, arranged as x1, y1, x2, y2,..., xk, yk.  Mean is the centered, unit-scaled full
Procrustes mean configuration.  Results are the landmark configurations rotated
and scaled to fit Mean.  Resids are the Procrustes residuals.  TanCoor are the full
Procrustes tangent coordinates.  Must run through CentScal first.
*****************************************************************************/
start GPA(Shapes,Sizes,Resids,TanCoor);
    CShapes=Shapes;
    run ProcMean(CShapes,Mean);
    do i=1 to nrow(CShapes);
      run OPA(CShapes[i,], Mean, tempresu, temptanc);
      Resitot=Resitot//(tempresu-mean);
      Tanctot=Tanctot//temptanc;
    end;
    Resids=Resitot;
    TanCoor=Tanctot;
finish GPA;
```

```
/*****************************************************************************
```

FindVar finds the total variance Input, both as the sum of the variances of the columns and as the geometric mean. It also calculates these statistics after regressing each variable on size, so that variation due to linear dependence on size is removed.

```
*****************************************************************************/
start FindVar(Input, size, AddVar, RAddVar);
n=nrow(Input);
k=ncol(Input);
onevec=j(n,1,1);
X=onevec||size;
invxs=inv(X`*X)*X`;
totadd=0;
resadd=0;
do i=1 to k;
   y=Input[,i];
   totvar=ssq(y)/(n);
   totadd=totadd+totvar;
   b=invxs*y;
   yhat=x*b;
   resid=y-yhat;
   resvar=ssq(resid)/(n);
   resadd=resadd+resvar;
end;
AddVar=sqrt(totadd);
RAddVar=sqrt(resadd);
finish FindVar;

/*****************************************************************************
*****************************************************************************/
start random(X,S,XR,SR);
```

```
      k=nrow(X);
      m_index=X;
      n_index=S;
      do i=1 to k;
        rand=floor(k*ranuni(0))+1;
        if rand=k+1 then rand=k;
        m_rand=m_rand//m_index[rand,];
        n_rand=n_rand//n_index[rand];
      end;
      XR=m_rand;
      SR=n_rand;
finish random;

/*****************************************************************************
*****************************************************************************/
start boot(Y,times,step,outCIs,Mact,Mout);
      run CentScal(Y,X,sizes);
      run GPA(X,sizes,resids,tancoor);
      run FindVar(tancoor, sizes, Ttotadd, Tsfadd,);
      run FindVar(resids, sizes, Rtotadd, Rsfadd,);
      Mact=Ttotadd||Tsfadd||Rtotadd||Rsfadd;
      print Mact;
      labels={50,.5,2.5,97.5,99.5};
      do k=step to times by step;
            do i=1 to step;
               run random(X,sizes,XR,SR);
               run GPA(XR,SR,rresids,rtancoor);
               run FindVar(rtancoor, SR, RTtotadd, RTsfadd);
               run FindVar(rresids, SR, RRtotadd, RRsfadd);
               Mrand=RTtotadd||RTsfadd||RRtotadd||RRsfadd;
               Mout=Mout//Mrand;
            end;
```

```
            tempM=Mout;

            corr=j(nrow(tempM),1,9999);
            do i=1 to ncol(tempM);
               ave=sum(tempM[,i])/nrow(tempM);
               c=ave-Mact[1,i];
               cvec=j(nrow(tempM),1,c);
               corr=corr||cvec;
            end;
            corr=corr[,2:ncol(corr)];
            corrdata=tempM-corr;
            run ConfInt(corrdata,CIs);
            a=Mact//CIs;

            b=j(5,1,k)||labels||a;
            totCIs=totCIs//b;
      end;
      outCIs=totCIs;
finish boot;


/*****************************************************************************
*****************************************************************************/
start ConfInt(Input,Output);
      c=ncol(Input);
      r=nrow(Input);
      pctiles=j(4,c);
      do i=1 to c;
         tempcol=Input[,i];
         b=tempcol;
         tempcol[rank(tempcol),]=b;
         CIvals={.005 .025 .975 .995};
```

```
      NumCIs=ncol(CIvals);
      do k=1 to NumCIs;
        np=r*CIvals[k];
        j=floor(np);
        g=np-j;
        if g=0 then y=(tempcol[j]+tempcol[j+1])/2;
          else y=tempcol[j+1];
        pctiles[k,i]=y;
      end;
    end;
    Output=pctiles;
finish ConfInt;


/*****************************************************************************
  Main Program
*****************************************************************************/
use merc2;
read all into Y;

run boot(Y,&times,&step,CI,Mact,M);

print CI;

filename out 'bootres.dat';
file out;

    do i=1 to nrow(CI);
        put (CI[i,1]) 5.0 +2 @;
        put (CI[i,2]) 4.1 +2 @;
        do k=3 to ncol(CI);
            put (CI[i,k]) 10.8 +2 @;
```

```
            end;
            put;
        end;
closefile out;


quit;
```

Bootstrap Version of MANOVA

```
*************************************************************************;
Program to do a bootstrap test similar to a MANOVA.
Input:  Data is in X, group info is in groups.

Code by Eric Dyerson, University of Arizona.
*************************************************************************;

**********(Enter these values)*****************************************;
%let TIMES= 10000;          *--number of times to randomize           ;
*************************************************************************;



********************MODULE DEFINITIONS*****************************;
groups=char(groups,1);
*---------MODULE THAT RANDOMIZES group labels------------------------;
           start randomiz(group,rgroup);
            igroup=group;
            k=nrow(group);
            do i=1 to k;
              rand=floor((k-i+1)*ranuni(0) + 1);
              kgroup=kgroup//igroup[rand];
              igroup=remove(igroup,rand);
            end;
            rgroup=kgroup;
          finish randomiz;

*-MODULE THAT COMPUTES  MAH distances-------------------------;
  start mah(X,group,MAH);
     index=unique(group);              *---symbols for groups----;
     m=ncol(index);                  *---number of groups----;
```

```
    n=nrow(X);                    *---number of specimens--;
    h=ncol(X);                    *---number of characters--;
    G=design(group);              *---design matrix---;
    mean_w = inv(G`*G)*G`*X;      *---within-group means----;
    Cp=j(h,h,0);                  *---initialize pooled cov----;
     do k=1 to ncol(index);
      Y=X[loc(index[k]=group),];  *---choose one group---;
      Y1=Y-j(nrow(Y),1,1)*Y[:,];  *---zero center-----;
      C=(1/(nrow(Y)-1))*Y1`*Y1;   *---covariance------;
      Cp = ((nrow(Y)-1) * C) + Cp; *---augment to pooled---;
     end;
    Cp=Cp/(n-m);                  *---final pooled cov-----;
    inC=inv(Cp);                  *---inverse of pooled cov---;
     do i=1 to m-1;
      do j=i+1 to m;
       diff= (mean_w[i,]-mean_w[j,]);
       MAH1 = MAH1 // diff * inC * diff`;
      end;
     end;
     MAH=MAH||MAH1;
  finish MAH;

*------MODULE THAT RUNS MAH OVER AND OVER---------------;
      start ranmah(X,group,times,MAH);
       do i=1  to times;
        run randomiz(group,rgroup);
        run mah(X,rgroup,MAH);
       end;
      finish ranmah;

*--------MODULE THAT CREATES LABELS FOR MAH DISTANCES-----------;
        start labels(groups,labels);
```

```
          index=unique(groups);
          m=ncol(index);
           do i=1 to m;
             do j=i+1 to m;
               labels=labels//(index[i]||'_'||index[j]);
             end;
           end;
          labels=rowcatc(labels);
        finish labels;
*****************END OF MODULE DEFINITIONS**********************;


****RUN MODULES THAT COMPUTE MAHALONOBIS DISTANCES*****;
       run mah(X,groups,MAH0);
       run ranmah(X,groups,&times,MAH);
*------CALCULATE PERCENTAGE ABOVE OBSERVED VALUE OF
DISTANCE--------;
      vt=t(shape(MAH0,ncol(MAH),nrow(MAH)));
      prob=MAH>vt;
      prob=prob[,+]/ncol(MAH);
*---OUTPUT RESULTS-------------;
    run labels(groups,dists);
    print "MAHALANOBIS DISTANCES",
        "(probabilities based on distribution with &times elements)",
       dists MAH0[format=8.3]  prob[format=8.3];
*--------STORE MATRIX OF MAH DISTANCES IN SASdataset CALLED
glot.NAHAL----;
   MAH=t(MAH);
   create glot.bsfdo from MAH [colname=dists];
   append from MAH;
   close glot.bsfdo;
*-----------------NOTIFY OPERATOR THAT PROGRAM IS FINISHED--------------;
```

```
            notes=400#(2##do(0,1,1/12));
            call sound(notes,0.2);

quit;
run;
```

Andrew M. Bush

Andrew M. Bush was born in Yonkers, NY on December 11, 1976, and he grew up in Cheverly, MD. While attending Eleanor Roosevelt High School, he interned at the Naturalist Center in the Natural History Museum in Washington, D. C. He received his B.S. in 1997 and his M.S. in 1999, both from the Department of Geological Sciences at Virginia Polytechnic Institute and State University. He has also taken classes through Rice University, the University of Washington, the University of Akron, and Miami University of Ohio. In the fall of 1999 he began working on his Ph.D. in the Department of Earth and Planetary Sciences at Harvard University.