Are the most durable shelly taxa also the most common in the marine fossil record?

Anna K. Behrensmeyer, Franz T. Fürsich, Robert A. Gastaldo,
Susan M. Kidwell, Matthew A. Kosnik, Michal Kowalewski, Roy E. Plotnick,
Raymond R. Rogers, and John Alroy

Abstract.—This paper tests whether the most common fossil brachiopod, gastropod, and bivalve genera also have intrinsically more durable shells. Commonness was quantified using occurrence frequency of the 450 most frequently occurring genera of these groups in the Paleobiology Database (PBDB). Durability was scored for each taxon on the basis of shell size, thickness, reinforcement (ribs, folds, spines), mineralogy, and microstructural organic content. Contrary to taphonomic expectation, common genera in the PBDB are as likely to be small, thin-shelled, and unreinforced as large, thick-shelled, ribbed, folded, or spiny. In fact, only six of the 30 tests we performed showed a statistically significant relationship between durability and occurrence frequency, and these six tests were equally divided in supporting or contradicting the taphonomic expectation. Thus, for the most commonly occurring genera in these three important groups, taphonomic effects are either neutral with respect to durability or compensated for by other factors (e.g., less durable taxa were more common in the original communities). These results suggest that biological information is retained in the occurrence frequency patterns of our target groups.

Anna K. Behrensmeyer. Department of Paleobiology, National Museum of Natural History, Smithsonian Institution, Post Office Box 37012, NHB MRC 121, Washington, D.C. 20013-7012.
  E-mail: Behrensa@si.edu
Franz T. Fürsich. Institut für Paläontologie, Universität Würzburg, 97070 Würzburg, Germany.
  E-mail: franz.fuersich@mail.uni-wuerzburg.de
Robert A. Gastaldo. Department of Geology, Colby College, Waterville, Maine 04901.
  E-mail: ragastal@colby.edu
Susan M. Kidwell. Department of the Geophysical Sciences, University of Chicago, Chicago, Illinois 60637.
  E-mail: skidwell@uchicago.edu
Matthew A. Kosnik. Centre for Coral Reef Biodiversity, School of Marine Biology and Aquaculture, James Cook University, Townsville 4811, Australia. E-mail: mkosnik@alumni.uchicago.edu
Michal Kowalewski. Department of Geosciences, Virginia Polytechnic Institute and State University, Blacksburg, Virginia 24061. E-mail: michalk@vt.edu
Roy E. Plotnick. Department of Earth and Environmental Sciences, University of Illinois at Chicago, Chicago, Illinois 60607. E-mail: plotnick@uic.edu
Raymond R. Rogers. Geology Department, Macalester College, St. Paul, Minnesota 55105.
  E-mail: rogers@macalester.edu
John Alroy. National Center for Ecological Analysis and Synthesis, University of California, Santa Barbara, California 93101. E-mail: alroy@nceas.ucsb.edu

Introduction

Taxa vary widely in their representation in the fossil record, ranging from singleton occurrences to widespread and frequent. But why are some fossil taxa more common than others? To what degree does this reflect original biological frequency versus artifacts of preservation, driven by inherent differences in postmortem durability or in environments of deposition? The reliability of fossil-occurrence frequency has profound implications for paleobiological analyses at many scales, from studies of the structure of ancient communi-ties preserved within individual beds to tests of large-scale macroevolutionary and macroecological diversity patterns.

At coarse taxonomic levels, durability—such as the contrast between mineralized and less mineralized forms—clearly biases fossil occurrences; jellyfish have lower preservation potential than clams, and insects are less preservable than dinosaurs. Less clear is the net effect of relative durability on the occurrence frequency of taxa within and among the major mineralized metazoan groups—mollusks, brachiopods, echinoderms, corals, etc.—

which constitute the bulk of the known marine fossil record and are the focus of most quantitative macroevolutionary analyses. For the purposes of this paper, we define durable taxa as having large, thick and/or reinforced shells and diagenetically resistant mineralogies (e.g., calcite versus aragonite) or shell microstructures with low ratios of organic matrix to mineral crystallites (e.g., fibrous versus prismatic calcite and cross-lamellar versus nacreous aragonite). A priori, one might expect that taphonomic processes by themselves would lead to overrepresentation of more durable shelly taxa relative to less durable shelly taxa. If more durable organisms comprise a disproportionately large number of fossils within their taxonomic groups, then this will affect measures of past taxonomic richness, ecological dominance, and ecomorphic variability as well as potentially obscure underlying large-scale secular trends.

However, do durable taxa actually dominate the known fossil record? Extrapolating from actualistic studies within particular groups of shelly benthos suggests that this should be the case (e.g., various lab and field experiments going back to Chave 1964; Driscoll 1970; for reviews see Martin 1999; Zuschin and Oliver 2003). In actualistic tests of time-averaging, Cummins et al. (1986) found much shorter taphonomic half-lives among small-bodied versus large-bodied bivalve individuals, and Krause et al. (2002) showed that calcitic brachiopods have equal or greater time-averaging than similar-sized aragonitic bivalves. In contrast, Martin et al. (1996) found that high Mg-calcitic benthic foraminifera had maximum postmortem ages comparable to those of large-bodied aragonitic bivalves, indicating similar net durabilities (i.e., more durable shell mineralogy can compensate for small body size).

The potential impact of taphonomic processes on the preservability of an individual organism, that is, the per capita probability of preservation, is illustrated in Figure 1. It can be argued that this per capita preservation probability should be greater for large organisms than for small organisms, for organisms with thick and/or reinforced shells compared to those with thin and smooth shells, and for
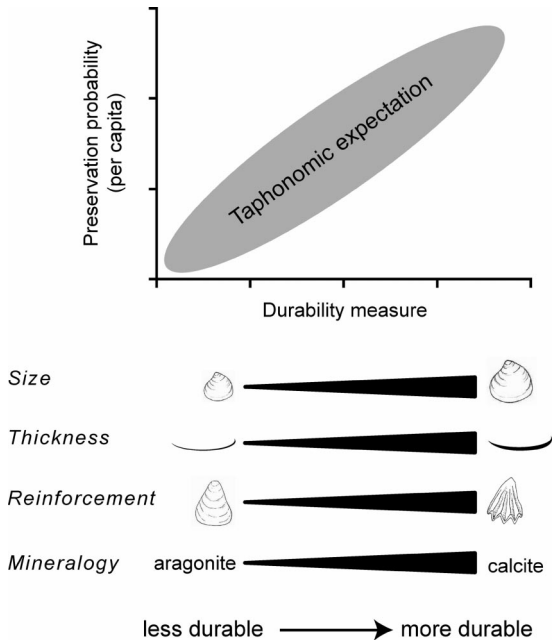

FIGURE 1. Schematic representation of the generally accepted taphonomic expectation that physically and chemically durable taxa are more common than less durable taxa in the fossil record.

those with calcitic rather than aragonitic shells (see ''Methods: Taphonomic Data Set'' for reasoning and supporting references). The taphonomic expectation is that shells that are more durable according to these criteria should be disproportionately represented in fossil assemblages, all else being equal. Of course, the actual representation of a taxon in a fossil assemblage depends not only on its per capita preservation probability (resulting from all durability factors), but also on the supply side of original biological populations (their abundance and turnover rates). For example, a genus with higher durability but low abundance in the living assemblage might still be rare at a fossil locality, whereas an originally abundant taxon with a high turnover rate might still be abundant as a fossil despite low durability. In either case, these taxa would be preserved and recorded as present in a collection captured by the Paleobiology Database (PBDB).

In this paper we test whether the most common fossil brachiopod, gastropod, and bivalve genera have intrinsically durable shells—with commonness measured by occurrence fre-

quency in the PBDB (http://paleodb.org). These analyses constitute the first steps toward developing a conceptual model that incorporates both biological reasoning and alternative impacts of taphonomic filters relating to shell durability. We used the PBDB to test whether the taphonomic filter defined above is directional enough and strong enough to result in a statistical correlation between durability and occurrence frequency for fossil marine brachiopods, bivalves, and gastropods. Our ability to infer a taphonomic filter from such a correlation assumes either that the original (living) occurrence frequencies of durable and less durable taxa were comparable or that less durable taxa had higher occurrence frequencies (see reasoning below). In September 2003, the three target groups represented 43% of the genera and 66% of the occurrences in PBDB, and they arguably include the most common and best preserved higher taxa found in the marine fossil record (Alroy et al. 2001). The 450 genera we investigated (the top 150 of each group) also constitute more than half of all occurrences for their respective groups. The top 150 genera include a wide range of occurrence frequencies, from PBDB-common taxa in more than 200 formations down to PBDB-infrequent taxa that occur in as few as 14 formations. Representation of the three groups through time is uneven (Fig. 2), with brachiopods dominating in the Paleozoic and mollusks in the post-Paleozoic, but the combined data set provides coverage for most of the Phanerozoic.

Conceptual Model

Although our goal is to test whether the most common taxa in the PBDB are characterized by high shell durability, we realize that many other taphonomic and biological factors could contribute to such a pattern. Figure 3 presents a range of scenarios that would lead to different proportions of durable and less durable taxa in the sampled fossil record, e.g., in the PBDB.

Conceptually, we can divide organisms in a living assemblage into two categories: those with less than average durability and those with greater than average durability. Here we
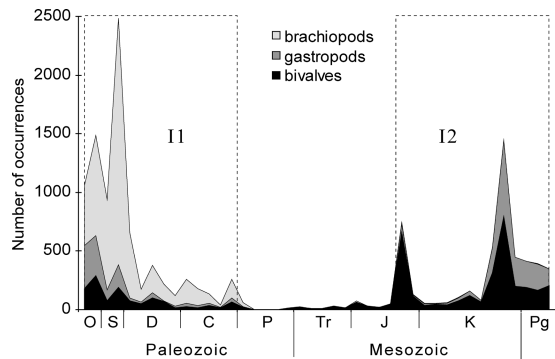

FIGURE 2. Overview of the compiled data on the three target groups, showing their distribution through time (10-Myr bins) based on the PBDB records for the total number of formation-occurrences in each bin. Dashed boxes show the data subsets used in the durability analysis. The peak in I1 for brachiopods indicates an unusually large number of formations with common brachiopod genera; likewise for bivalves and gastropods in last Cretaceous bins. Note that the I1 fossil record is better sampled than I2, on the basis of formation-occurrences. Total invertebrate marine genera: 16,380; total occurrences (collections): 214,428.

consider two possible initial (biological) frequency distributions of more and less durable taxa (left column of Fig. 3). In the upper example, the frequency distributions of the two groups in the living community are equal, and this is the simplest assumption. In the lower example, the less durable taxa (black outline) are more common on average than the more durable taxa (shaded area), and this may be biologically plausible. For example, several studies suggest that the local abundance of marine invertebrate species can peak at either small or intermediate body sizes (e.g., Warwick and Clarke 1996; McClain 2004). Thus, if we use body size as a proxy for durability (i.e., large-bodied taxa have more durable shells), then less durable taxa are more abundant in the original biological communities and have a supply-side advantage to compensate for their intrinsically low durability. A converse relationship in which large-bodied taxa are more abundant represents a third possible condition in the living community, but this is less biologically plausible and is not considered further here. Note that although the illustrated taxon frequency distributions (histograms) are normal, the argument is independent of the shape of the distribution.
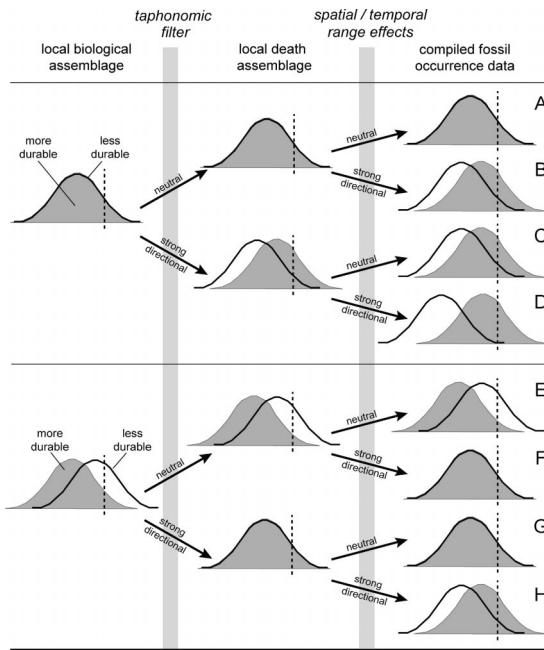
Taphonomic modifications of the original

FIGURE 3. Given differing initial frequency distributions of more durable (shaded area) and less durable (area under solid black line) taxa in a local life assemblage (left column), local taphonomic processes and relationships between durability and temporal-spatial distributions (filters represented by shaded vertical bars) can result in different frequency distributions in the fossil record (right column). Eight qualitatively different scenarios (A–H) are illustrated here. The taphonomic filter is always present but may be neutral or directional and strong, that is, causing the local death assemblage (middle column) to be more or less similar to the initial frequency distribution (operationally, a strong filter causes the fossil pattern to differ qualitatively from the original biological signal). Similarly, the spatial-temporal filter may be neutral or directional and strong, causing the ultimate fossil frequency distributions to be less or more offset from what is observed in local death assemblages. See text for further detail. Key: Individual frequency curves are normal and depict idealized distributions of taxa; the x-axis represents occurrences (grouped as sequential categories) and the y-axis represents frequencies of taxa, i.e., resulting in a histogram of numbers of taxa in each occurrence category. Thus, a shift to the right on the x-axis indicates absolutely more occurrences. Dashed vertical lines ("veil lines" of Preston 1948) show the lower limit of a sample comparable to our most common 150 taxa of each major group; only taxa to the right of the lines are sampled.

frequency distributions are shown in the middle column of Figure 3. A "neutral" taphonomic filter is random with respect to durability; more durable and less durable individuals have on average the same per capita preservation probability, and thus their relative representation in the local death assemblage

remains unchanged from the living assemblage. A "directional" taphonomic filter preferentially removes less durable individuals, causing the local death assemblage (middle column) to differ qualitatively from the original biological signal, e.g., by changing the rank order occurrence frequencies of taxa. For our model, such modifications would be the result of a "strong" taphonomic filter, and this is what we focus on in Figure 3. If the two distributions were originally equal, the fossil distribution becomes biased, with durable forms more common than less durable forms (second row). If the less durable forms were originally more common than the durable ones, the filter results in bias by causing their distributions to coincide, or even to become offset in the opposite direction (fourth row). A "mild" directional taphonomic filter would produce a range of patterns intermediate to the neutral and strong filters (not depicted in Fig. 3).

Occurrence frequencies in the actual fossil record reflect not only taphonomic filters acting at single localities, but also a variety of broader spatial and temporal range effects. For example, studies of terrestrial vertebrates suggest a relationship between geographic range and body size (Brown 1995), with larger animals generally having larger geographic ranges. Unfortunately, there is little macroecological information to indicate what these relationships might be for marine shelly macrobenthos; available evidence suggests that bivalve genera with small body size can have small or large geographic ranges, but large-bodied taxa only have large ranges (Roy et al. 2001; unpublished data of Roy et al. 2000). Durability-related variables may be linked not only with local abundance and geographic range, as considered above, but also taxon duration, stenotypy, and other factors that might increase or decrease occurrence frequency in the fossil record.

The right column in Figure 3 shows different combined influences of stratigraphic and geographic range on the total number of occurrences. In the right column, rows 1, 3, 5, and 7 reflect the absence of a relationship between range and durability; rows 2, 4, 6, and 8 show the resulting pattern if this effect is

strongly positive. The predicted scenarios in the right column of Figure 3 are not exhaustive but explore the consequences of macroecological relationships such as body size and spatial-temporal ranges. In order to quantitatively assess the importance of such relationships for the composition of the marine fossil record, more information will have to be collected for modern benthos. Only then will we be able to assess rigorously whether durability-related variables (e.g., large body size, resistant shell mineralogy) are positively, negatively, or neutrally related to other biological factors that contribute to high occurrence levels in the fossil record (e.g., wide geographic range, long evolutionary range, high local abundance).

This thought experiment produces eight possible idealized occurrence frequencies of durable versus less-durable taxa (right column of Fig. 3). In each scenario, the vertical dashed line represents the division between unsampled taxa on the left ("rare" taxa) and sampled taxa on the right ("veil line" of Preston 1948). Focusing only on the sampled taxa, three scenarios (A, F, and G) predict that more durable and less durable forms will be equally represented. Four scenarios (B, C, D, and H) predict a dominance of more durable forms among the sampled taxa; this is how many paleontologists would visualize taphonomic bias that is linked to shell durability. Only one scenario (E) predicts that less durable forms would dominate.

Methods

We used the PBDB to test the relationship between durability and occurrence frequency among the most common fossil marine brachiopods, bivalves and gastropods, corresponding to taxa to the right of the veil line in Figure 3. Our analyses were based on three data sets: (1) *Occurrence Data Set*, initiated by a download of the 150 most commonly occurring brachiopod, bivalve, and gastropod genera (~450 genera total; we included genera with tied rankings) in the PBDB. This data set was continuously updated until 15 September 2003; (2) *Taphonomic Data Set*, consisting of multivariate data on the durability characteristics of the genera present in the occurrence

data set—the taphonomic data set was generated from original specimens or from published literature; and (3) *Lithologic Data Set,* downloaded from the PBDB lithology field for each collection.

In this section, we briefly describe variables and data-collecting strategies that we used to assemble these three data sets (See the supplementary material online at http://dx.doi.org/10.1666/04023.S1).

Occurrence Data Set.—Our measure of fossil-occurrence frequency is based on the number of occurrences recorded in the PBDB, where an occurrence is a record of a genus from a specific formation. Occurrence frequency does not reflect the abundance of *specimens* of a genus in individual collections, but rather it refers to the number of *collections* in which the genus is known to be present. An individual occurrence may be based on thousands of complete specimens or on a single fragment. Collections entered in the PBDB are as specific as possible to particular sites or strata. Thus, there may be multiple collections from a single measured section or from a single lithostratigraphic formation. For our analysis we binned occurrence data by formation ("formation-occurrences") to minimize the effect of multiple records within single well-studied formations (see Appendix for discussion of monographic effects).

Two time segments of the Phanerozoic, referred to as "I-1" (Late Ordovician through Late Carboniferous) and "I-2" (Late Jurassic through Paleogene), respectively (Alroy et al. 2001), were best represented in the PBDB at the time of our study, thus we focused on these intervals (Fig. 2). We downloaded occurrence data by genus, both for the database as a whole (i.e., total numbers of occurrences over the entire span of geologic time represented by the combined I-1 and I-2 time intervals) and for data binned into 10-myr intervals.

Taphonomic Data Set.—Information relevant to postmortem durability was gathered for the top 150 taxa in the PBDB for each of our three target groups (see Appendix for more details about sample sizes, taxa studied, and procedures, and Supplementary Materials for a list of these taxa). Durability variables for speci-

TABLE 1. Measures of durability used in this study. n = total number of genera with taphonomic data for that variable.

Durability measure	n	Variables	Units
Size	266	X: maximum dimension Y: intermediate dimension perpendicular to X Z: smallest dimension perpendicular to X	Millimeters
Thickness	114	From interior to exterior surface of shell; semi-quantitative score relative to shell size	Ranks: 0 = thin 1 = intermediate 2 = thick
Skeletal reinforcement	249	Spines Folds Ribs	Ranks: 0 = absent 1 = present 2 = prominent
Mineralogy	309	Dominant shell layers, based on published literature for the genus or family	Categories: High-Mg calcite Low-Mg calcite Aragonite Bimineralic Organophosphate
Organic content	217	Dominant shell layers in mollusks only, based on published data for microstructural types	Categories: High organic = ≥1% Low organic = <1% Bimineralic-organic

mens representing each target genus were scored by one or more of the authors, using a set of durability scales suitable for statistical analysis (described below; Table 1) and either museum or teaching/research collections or the published literature (e.g., monographs, treatises). Choice of specimens was determined by availability, but an effort was made to include a wide range of species and specimens for each genus.

We define taphonomically durable taxa as those having shells with high intrinsic resistance to destruction by physical, chemical, and biological processes. A wealth of evidence from manipulative and observational experiments over the last several decades indicates that, in modern environments, shell robustness generally increases in some positive fashion with shell size, presence of prominent shell reinforcements (ribs, folds, spines), relative shell thickness standardized against the overall shell size, and the inclusion of calcite and/or low-organic microstructures (Fig. 1) (for reviews, see Parsons and Brett 1991; Kidwell and Bosence 1991; Briggs 1995; Martin 1999; Harper 2000; Sanders 2003; Wright et al. 2003; Zuschin et al. 2003). In addition, we sus-

pect that (1) larger shells are more likely to be seen and recovered (especially if specimens are collected from outcrop surfaces rather than derived from bulk samples) and identified at the genus or species level; and (2) shells with more or stronger ribs, folds, and spines will leave more identifiable fragments than unornamented or unreinforced shells (Kowalewski et al. 2003). Both of these factors would further increase the likelihood that intrinsically durable genera would be included in fossil-taxon compendia such as the PBDB.

1. *Shell size.* We measured the X, Y, and Z linear dimensions of shells (mm), with X the longest dimension, Y the longest dimension perpendicular to X, and Z the length perpendicular to X and Y. When information from multiple individual specimens or congeneric species was available to estimate the body size of a genus, mean dimensions were used. All size variables were transformed as natural logs (ln) to normalize distributions. Given the difficulty of obtaining all three X-Y-Z measurements for some specimens and the fact that the three variables are highly correlated ($r^2 > 0.65$ in all cases), we restricted final analyses to the maximum dimension (X) only. The

median size of genera within each of the major groups was used to categorize each genus as either ''large-bodied'' or ''small-bodied.'' This median value of X varies depending on the temporal and taxonomic scope of the particular analysis; median values used as thresholds are reported in subsequent tables.

2. *Relative shell thickness.* On the basis of a pilot study that included tests for operator error (see Appendix), we assigned genera to three ordinal categories based on a visual assessment: 0 = thin, 1 = intermediate, 2 = thick. Thick shells have a high ratio of shell thickness relative to body size; generally, taxa assigned to this category had shells of ≥3 or 4 mm thickness regardless of shell length (except for brachiopods, where a shell thickness of ~2 mm qualified as thick). Thin shells have a low ratio of thickness relative to body size, which generally meant an absolute shell thickness ≤1 mm regardless of shell length (all groups).

3. *Shell reinforcements.* Each of three characters (ribs, folds, spines) was scored in ordinal ranks (0 = absent, 1 = present, 2 = prominent) and these were summed to obtain the total Shell Reinforcement score (SR). Irregularities or comarginal banding in shells caused by growth lines were not considered as reinforcements; however, concentrically ridged ornament (including axial sculpture on gastropods due to apertural thickening during shell growth) was considered along with radial or spiral ridges as ''ribs.'' ''Folds'' are relatively large invaginations of the commissure (bivalves, brachiopods) or aperture (gastropods) that deflect the interior as well as exterior surface of the shell (if such deflections are filled, they were considered ''ribs'').

4. *Shell mineralogy.* To characterize mineralogy we used five categorical variables: (1) high-Mg calcite, (2) low-Mg calcite, (3) aragonite, (4) bimineralic calcite-aragonite, and (5) organophosphate. Shell mineralogy for each genus was assigned according to literature for that genus (if possible) or the lowest-rank higher taxon (usually the family for mollusks, order for brachiopod) that includes that genus. Primary sources were Taylor et al. (1969, 1973), Carter (1990), and Williams et al. (2000). Mollusks containing both aragonitic and low-Mg calcitic shell layers were categorized as ''bimineralic.''

5. *Organic content of microstructure.* All genera were scored in terms of shell organic content as either ''high'' (≥1% organic content by weight), ''low'' (≤1%), or ''mixed'' (different organic contents in different shell layers). This follows the system of Kidwell and Brenchley (1996), drawing on published measurements in particular from Taylor et al. (1969, 1973), Harper (2000), Jope (1965), and Peck et al. (1987).

For each of the five durability metrics, the durability of genus [D] was computed as follows:

$$D = \sum d_i / n_t \qquad (1)$$

where d_i is a value of the durability measure for the ith specimen representing this genus and n_t is the total number of specimens from that genus included in the analysis.

To analyze changes in durability between I1 and I2, we merged data for all measured taxa in 10-Myr intervals within I-1 and I-2 to obtain characteristic durability values for each interval. Two methods were used to obtain these values: the unweighted average interval durability (D_U) was calculated from

$$D_U = \sum D_i / N_t \qquad (2)$$

where D_i is the durability of the ith genus and N_t is a total number of genera in a given time interval. This calculation counts all genera equally, independent of the number of genus occurrences. We also calculated a weighted average interval durability:

$$D_W = \frac{\sum (Q_i \times D_i)}{\sum O_t} \qquad (3)$$

where O_i is the number of formation-occurrences of the ith genus in that time interval, and O_t is the total number of formation-occurrences of all genera in that time interval. Equation (2) weights all genera equally when computing average durability per time interval; equation (3) assigns more importance to genera with higher formation-occurrence frequencies.

Lithologic Data Set.—In addition, we tested the relationship of different lithologies and oc-

TABLE 2. Summary of formation-occurrence data for genera grouped into two size categories: ''small'' (genus size < median size) and ''large'' (genus size > median size). Data analyzed for pooled data as well as separately for the three higher taxa: brachiopods, bivalves, and gastropods. Differences in median size and shape of the size distribution were tested using the Wilcoxon and K-S Tests, respectively; neither shows any significant difference. Pearson and Spearman Rank Tests show that the formation-occurrence patterns for the two size groups are not significantly different.

| Taxon | Median (mm) | Small (<median) | | Large (median) | | Wilcoxon Two-sample Test | | Kolmogorov-Smirnov Test | | Correlation between taxon size and no. of formation-occurrences | |
		No. of genera	Median no. of occurrences	No. of genera	Median no. of occurrences	Z	p	D	p	Pearson*	Spearman Rank
Total	24.4	133	25	133	24	0.26	0.8	0.12	0.29	$r = -0.0008$ $p = 0.99$	$r = -0.016$ $p = 0.8$
Brachio-pods	21.5	39	51	39	32	−1.29	0.098	0.23	0.25	$r = -0.12$ $p = 0.30$	$r = -0.13$ $p = -0.24$
Bivalves	28.6	44	25	43	28	0.29	0.58	0.13	0.86	$r = 0.18$ $p = 0.10$	$r = 0.12$ $p = 0.26$
Gastro-pods	28.6	51	16	50	15	−0.61	0.53	0.13	0.80	$r = -0.009$ $p = 0.93$	$r = -0.06$ $p = 0.52$

* Pearson correlation coefficients and corresponding *p*-values computed for log-transformed size data.

currence frequency in our target groups, for a total of 453 taxa (including a few that had tied scores or shifted rank into and out of the top 150 during the course of our study). This addressed the question of whether taxa might be more frequent than expected on the basis of durability because they occur in commonly preserved marine environments or in a wider range of lithologies.

Using the three independently derived data sets, we evaluated the influence of shell durability on the occurrence frequency of fossil genera. Concurrently, we examined data reliability issues including sample size, operator error, and other causes of variability in the durability measurements, taxonomic representation in the I1 and I2 time intervals, binning effects on occurrence data, and monographic biases (Appendix). The univariate statistical tests employed in this study were restricted to nonparametric, rank-based techniques (Wilcoxon two-sample median test, Kruskal-Wallis multi-sample test, etc.). These tests were deemed more appropriate than standard parametric tests (e.g., *t*-test, ANOVA) because (1) the data include unbalanced sampling designs; (2) obvious departures from normality affect some of our variables; (3) some of the samples may be insufficient in terms of sample size to overcome those departures from normality; and (4) the selected rank-based

tests would offer nearly as much power as their parametric counterparts, even when the assumptions of the latter methods were not violated (e.g., Zar 1999). The significance level of α = 0.05 has been used in all our analyses.

Results

General Characteristics of the Data.—We compared the number of more durable versus less durable taxa in terms of their occurrence frequency in the PBDB for the whole data set, for I1 and I2, and for each of the major groups. We also used our data set to examine secular variation in shell mineralogy of these groups through the Phanerozoic; additional investigation of secular patterns in durability is beyond the scope of this paper. The durability variables (shell size, thickness, reinforcement, mineralogy, organic content) behaved independently, according to bivariate tests that failed to show any tendency for covarying relationships. For example, there is no tendency for genera with calcitic shells to be large bodied, thick, highly reinforced, and/or low-organic, or for genera with low-durability mineralogies also to have smaller body sizes, thinner shells, etc.

Body Size.—Small-bodied and large-bodied taxa, which are separated into two groups at 24.4 mm (median maximum dimension of 266 genera) (Table 2, Fig. 4), show no statistical
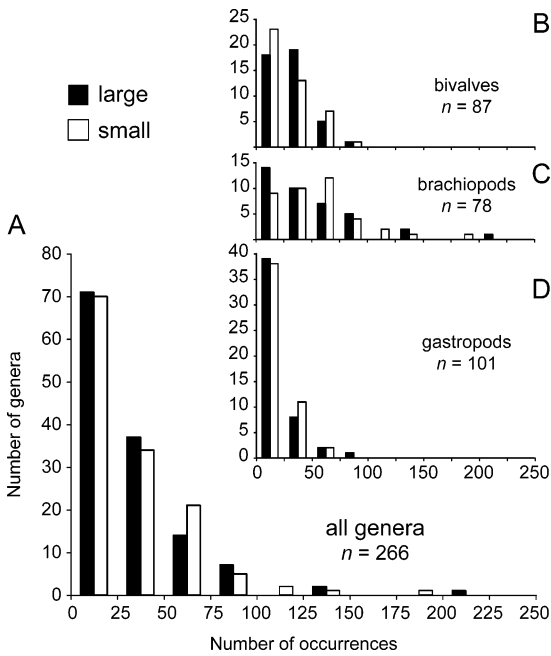
FIGURE 4. Formation-occurrence frequency distributions for genera grouped into ''large'' (>median) and ''small'' (<median) shells. A, Combined data. B, Bivalves. C, Brachiopods. D, Gastropods. See Table 2 for data summary and statistical tests.
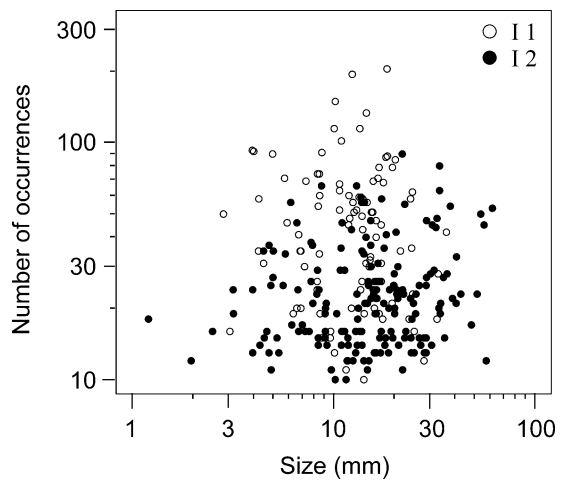

FIGURE 5. Scatter plot of average per-genus maximum size versus number of formation-occurrences on $\log_{10}$ axes, showing that shell size is not correlated with formation-occurrence frequency in the combined sample of brachiopod, bivalve, and gastropod genera. Data plotted separately for I-1 (empty points) and I-2 (filled points) 10-Myr time intervals. See Table 2 for data summary and statistical tests.

difference overall or within any of the major taxonomic groups in the frequency distribution of occurrences. The distributions do not differ in central tendency or shape, and ln-body size and frequency of formation-occurrences are not correlated significantly (Table 2). Comparison of I1 and I2 (Fig. 5) shows that there is no difference in the median size preserved, although the range for I2 extends to both larger and smaller genera than that for I1.

Shell Thickness.—Thin-shelled and thick-shelled genera (N_a = 114 genera; Table 3, Fig. 6; ''intermediate thickness'' taxa were excluded from the analysis) show no statistical difference overall or within any of the major groups in the frequency distribution of occurrences. Exceptionally, brachiopods do show a significant difference, but the relationship is opposite to the taphonomic expectation: thin-shelled genera have a higher median occurrence than thick-shelled genera.

Shell Reinforcement (SR).—The shell reinforcement index shows no statistical difference overall or for any of the major groups in

TABLE 3. Summary of formation-occurrence data for genera grouped into thickness categories: ''thin'' (average per-genus thickness < 0.67) and ''thick'' (average per-genus thickness > 1.34). Genera with ''intermediate thickness'' (0.67–1.34; a large proportion of the scored sample) were excluded from the analysis to highlight the thick-thin differences. Data analyzed for pooled data as well as separately for the three higher taxa: brachiopods, bivalves, and gastropods.

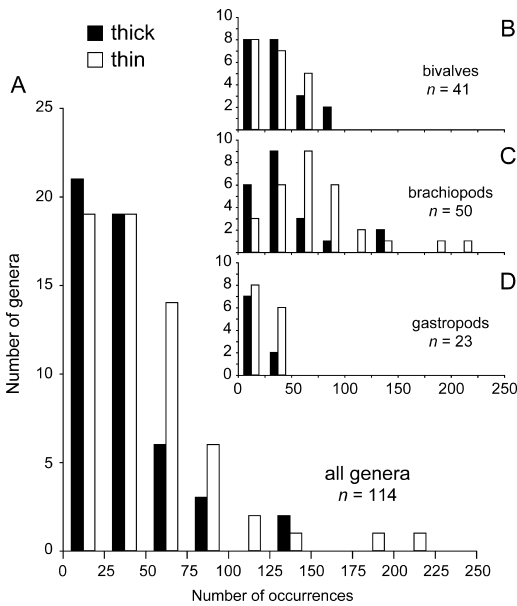|	Thin (score < 0.67)		Thick (score > 1.34)		Wilcoxon Two-sample Test	
Taxon	No. of genera	Median no. of formation-occurrences	No. of genera	Median no. of formation-occurrences	Z	p
Total data	63	43	51	30	−1.85	0.065
Brachiopods	29	71	21	34	−2.55	0.01*
Bivalves	20	28	21	31	−0.31	0.75
Gastropods	14	22	9	17	−0.99	0.32

* Tests significant at α = 0.05 level.

FIGURE 6. Formation-occurrence frequency distributions for genera grouped into "thick" (thickness > 1.34) and "thin" (thickness < 0.67) shells. Genera categorized as "intermediate" were excluded from the analysis. A, Pooled data. B, Bivalves. C, Brachiopods. D, Gastropods. See Table 3 for data summary and statistical tests.

the frequency distribution of occurrences (N_a = 249 genera; Table 4, Fig. 7). Differences in SR distribution for I1 and I2 appear to result from a somewhat higher degree of shell reinforcement in common I1 brachiopods than in I2 mollusks.

Shell Mineralogy.—Only a few genera among the targeted 450 have organo-phosphatic shells (n = 5, all linguliform brachiopods), and these genera were excluded from the analysis. Low-Mg calcitic genera (i.e., more chemically stable shells) do have a significantly higher number of occurrences than bimineralic or aragonitic genera in the overall data set.
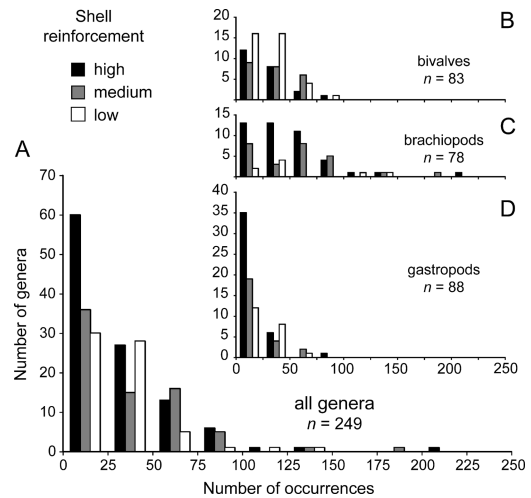

FIGURE 7. Formation-occurrence frequency distributions for genera grouped into three shell reinforcement categories: "high" (SR ≥ 2); "medium" (1 < SR < 2); and "low" (SR ≤ 1). A, Pooled data. B, Bivalves. C, Brachiopods. D, Gastropods. See Table 4 for data summary and statistical tests.

(N_a = 309 genera, p = 0.0001; Table 5, Fig. 8). Significant differences appear within the Paleozoic subset of data, but not within the post-Paleozoic, where genera are overwhelmingly aragonitic (p = 0.12, Table 5, Fig. 9). Within major groups, bivalves show no significant differences in median occurrences of calcitic versus aragonitic genera, brachiopods are not relevant (calcitic genera only), and I2 gastropod genera differ significantly but opposite to taphonomic expectation (aragonitic genera have more occurrences than bimineralic genera; Table 5). Where significant differences exist in the overall data set, they are driven by the exceptionally high occurrences of very thin-shelled calcitic brachiopods.

Approximately the same range of body siz-

TABLE 4. Summary of formation-occurrence data for genera grouped into three shell reinforcement (SR) categories: "low" (average per-genus SR score ≤ 1); "medium" (1 < average per-genus SR score < 2), and "high" (average per-genus SR score ≥ 2). SR scores computed as a sum of raw scores for folds, ribs, and spines. Data analyzed for pooled data as well as separately for the three higher taxa: brachiopods, bivalves, and gastropods.

	Low		Medium		High		Kruskal-Wallis Test	
Taxon	No. of genera	Median no. of occurrences	No. of genera	Median no. of occurrences	No. of genera	Median no. of occurrences	χ	p
Total data	74	26	109	24	66	26.5	2.06	0.36
Brachiopods	8	38	26	57.5	44	35.5	1.26	0.53
Bivalves	37	26	23	34	23	25	1.72	0.42
Gastropods	21	23	25	18	42	15	4.89	0.09

TABLE 5. Summary of formation-occurrence data for genera grouped into three mineralogic categories: ''low Mg-calcite,'' ''bimineralic,'' and ''aragonite.'' The other two categories in the durability data set (''high Mg-calcite'' and ''phosphate'') are not included because of small sample sizes. Note that for seven out of the nine groups analyzed here (taxic groups listed in the first column from the left), the total number of genera for the entire Phanerozoic (the second column from the left) is smaller than the sum of the genera computed by adding the I1 (the fourth column from the left) and I2 (the sixth column from the left) genera. This is because some genera range through I1 and I2: they are included as observations in both time intervals, while representing only one observation in the pooled data. However, these long-ranging genera represent <5% of genera for each mineralogical group (1.7% for low Mg-calcite taxa, 2.8% for bimineralic taxa, and 4.8% for aragonite taxa, respectively). Their complete exclusion does not change the outcome of any of the tests reported in the table. Significant results are in boldface.

	Phanerozoic (I1 + I2)		I1		I2	
Taxon	No. of genera	Median no. of formation-occurrences	No. of genera	Median no. of formation-occurrences	No. of genera	Median no. of formation-occurrences
All taxa–low Mg-calcite	116	35.5	94	40.5	24	23
All taxa–bimineralic	35	28.0	18	35.0	18	18.5
All taxa–aragonite	158	21.5	26	28.0	140	20.5
Kruskal-Wallis Test	**$\chi^2 = 49.5$** **$p < 0.0001$****		**$\chi^2 = 9.6$** **$p = 0.008$***		$\chi^2 = 4.3$ $p = 0.12$	
Wilcoxon Two-sample Rank Test (bimineralic taxa excluded)***	**$Z = 7.05$** **$p < 0.0001$****		**$Z = 2.93$** **$p = 0.003$****		**$Z = 2.05$** **$p = 0.04$***	
Bivalves–low mg-calcite	22	27	1	40	21	25
Bivalves–bimineralic	15	33	3	41	12	28
Bivalves–aragonite	63	26	15	21	55	25
Kruskal-Wallis Test	$\chi^2 = 0.36$ $p = 0.83$		$\chi^2 = 3.34$ $p = 0.19$		$\chi^2 = 0.06$ $p = 0.97$	
Wilcoxon Two-sample Rank Test (bimineralic taxa excluded)***	$Z = 0.10$ $p = 0.92$		*Insufficient sample size (n = 1 for low Mg-calcite)*		$Z = 0.14$ $p = 0.89$	
Brachiopods–low Mg-calcite	94	40	93	41	3	16
Gastropods–bimineralic	20	24	15	35	6	12.5
Gastropods–aragonite	95	16	11	33	85	16
Wilcoxon Test	$Z = 1.17$ $p = 0.24$		$Z = 0$ $p = 1$		**$Z = -2.22$** **$p = 0.03$***	

* Tests significant at $\alpha = 0.05$ level.
** Tests significant at $\alpha = 0.005$ level.
*** All p-values for Wilcoxon Test are based on a two-tailed test with normal approximation. Other approximations, based on the t and χ^2 distributions (not reported here), have yielded in all cases outcomes consistent with those reported here.

es is represented by calcite, aragonite, and bimineralic shell mineralogies (Fig. 10). However, within each major group, the number of occurrences increases with body size up to about ln(length) = 3.5 (33 mm). Above this size there is no further trend or difference between calcitic and aragonitic shells.

Shell Organic Content.—We tested this variable for mollusks only (Table 6). Among bivalves, differences are significant ($p = 0.02$) but opposite to those expected: genera with high-organic shells have a higher number of occurrences than other genera. Gastropods show no significant differences as a function of organic content.

Lithology.—Occurrence frequency is not correlated with either carbonate or siliciclastic lithology for any of the groups. Thus, durability and general lithology are independent in our data set. Taxa with greater occurrence frequencies tend to be represented in a broader array of facies (Fig. 11; Supplementary Materials 4).

Discussion

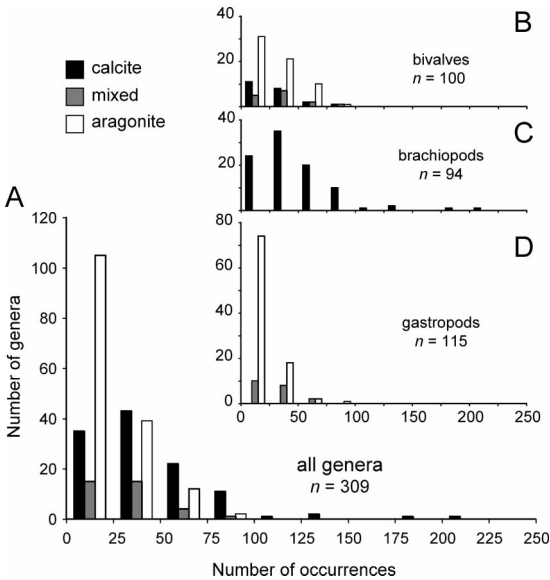Our results, based on formation-occurrence frequencies of the 450 most common brachio-

FIGURE 8. Formation-occurrence frequency distributions for genera grouped into mineralogical categories. A, Pooled data. B, Bivalves. C, Brachiopods. D, Gastropods. See Table 5 for data summary and statistical tests.
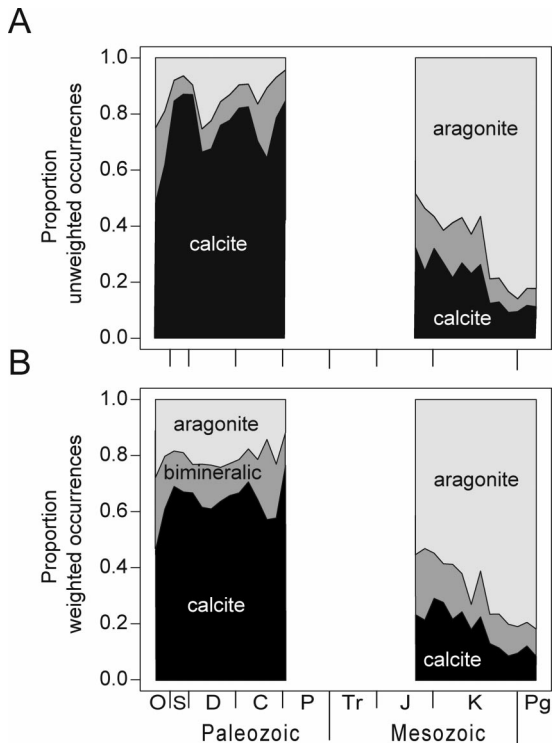

FIGURE 9. Plots of mineralogical patterns through the two target-time intervals (I1 and I2) with data binned into 10-Myr time intervals. The data are plotted separately for unweighted (A) and weighted (B) formation-occurrence data (see text for equations 2, 3). The 10-Myr time bins offered the highest possible resolution in which reasonable sample sizes (>75 formation-occurrences per interval) were maintained for all time intervals included in the analysis (Alroy et al. 2001).

pods, bivalves, and gastropods in the PBDB, generally show that commonness (i.e., occurrence frequency in the PBDB) is not correlated with shell durability (Table 6). Of the 30 tests we performed, only six show a significant relationship of shell durability with occurrence frequency, of which three are concordant with and three contrary to the taphonomic expectation of a positive relationship. Although taphonomic processes undoubtedly acted upon these durability factors and modified the pattern of fossil occurrences to some degree, our results suggest that durability-related taphonomic filters are either (1) neutral to mild or (2) compensated for by other factors, e.g., high supply-side input of less durable shells and/or spatial-temporal range effects. Testing these other factors explicitly will require basic macroecological data to generate realistic null expectations for supply-side input of durable and less durable shells in living communities.

Returning to Figure 3, almost none of our results match the most common taphonomic expectations (Fig. 3B,C,D,H), which are that durable genera should be more common in the fossil record (on the basis of occurrence frequency). Body size, shell thickness, shell re-

inforcement, and shell organic content show no significant relationships in this expected direction (i.e., that larger, thicker, more reinforced, lower organic content should be more common). The only significant positive relationships between occurrence and durability are found in mineralogy, but even these were not consistent: six out of nine either were not significant or were significant but contrary to taphonomic expectation. The three significant positive relationships occurred both for pooled data and when I1 and I2 were considered separately.

Reconsidering the alternative conceptual scenarios (Fig. 3), the bulk of our observations are consistent with the three pathways (A, F, G) that show no significant difference in the occurrences of more or less durable taxa. On the basis of our analyses, we cannot empiri-
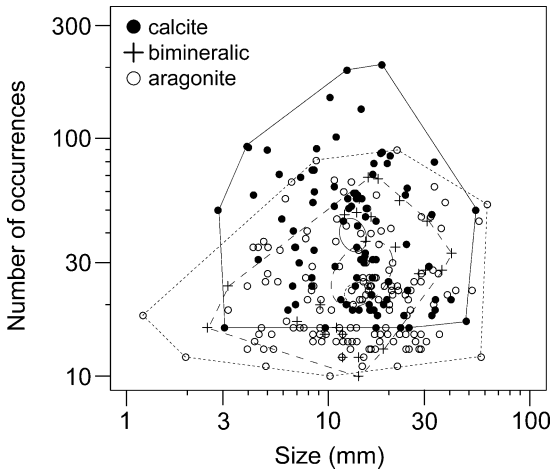
FIGURE 10. Scatter plot of the number of per-genus formation-occurrence versus shell size for bivalves with different shell mineralogies on $\log_{10}$ axes. Convex hulls for each type enclose the point scatters and the ovals show the 95% confidence intervals of the centroids. Key: Filled points and solid line for calcite, open points and short-dash line for aragonite, crosses and long-dash line for bimineralic mineralogies.

cally evaluate the relative importance of these alternatives, and it is possible that all three could hold for different taxonomic groups, times, or geographic locales. Given initially identical frequency distributions of more and less durable taxa, scenario A supposes no relative change in the fossil record, either from local taphonomic biases (the first filter in Fig.

3) or from spatial-temporal distribution effects (the second filter in Fig. 3); although these processes may have been operating, they do not affect less and more durable taxa differentially. Scenarios F and G suppose a different, biologically realistic initial distribution with less durable taxa more common than durable ones. In F, local taphonomic bias is neutral but spatial-temporal effects are directional and strong, whereas in G, local taphonomic effects are directional and strong and spatial-temporal effects are neutral.

It is possible that the top 150 genera in each group may not reflect the durability trends of other, less common genera in the database. However, we have reason to believe that our current results are more broadly representative of the groups we examined. As noted earlier, the 450 most common PBDB taxa that we considered include a large number of taxa with only moderate numbers of occurrences (down to 14 formation-occurrences; see Supplementary Materials); thus our sample covers more than just the extreme tail of the occurrence frequency distribution.

Recording durability variables and testing all genera for the three groups are beyond the scope of this project, but we can examine the robustness of our results using bivalve shell mineralogy. Mineralogy was the one param-

TABLE 6. Summary of information regarding tests of durability vs. occurrence-frequencies. ''O'' = no statistically significant relationship; ''Yes'' = significant on the basis of the taphonomic expectation of a positive relationship between durability and occurrence-frequency; ''No'' = significant but contrary to the taphonomic expectation of a positive relationship between durability and occurrence-frequency. Blank cells were not tested in this study.

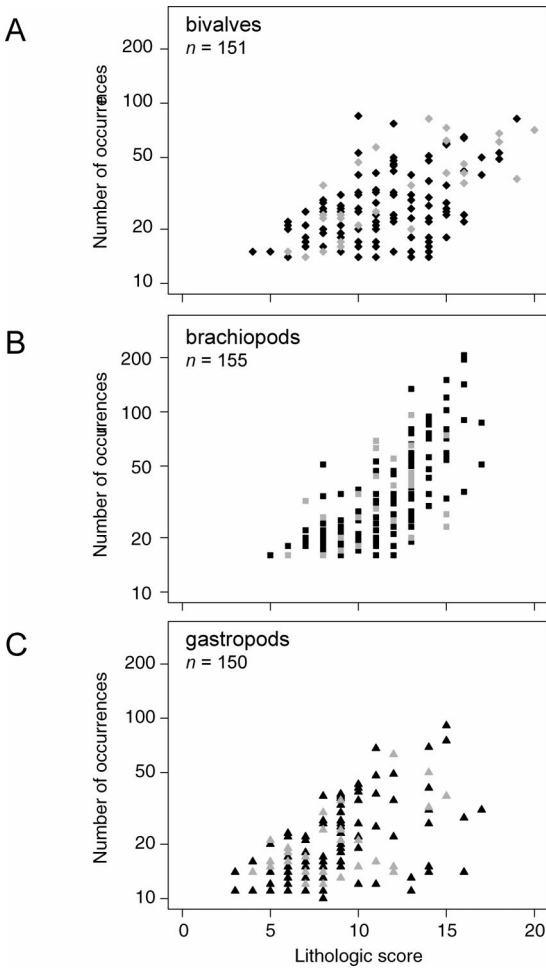| Tests for differences in median no. of formation-occurrences for durability categories	Ln body size (smaller vs. larger than median value)	Thin vs. thick shells (intermed. excluded)	Low vs. medium vs. high shell reinforcement	Aragonite vs. calcite (bimineralic excluded)	High- vs. low-organic micro-structures (hetero-organic excluded)	Carbon-ate vs. siliciclastic facies
n genera w/data	266	114	249	309	207	453
All data	O	O	O	Yes		O
All brachiopods	O	No	O			O
All bivalves	O	O	O	O	No	O
All gastropods	O	O	O	O	O	O
I1 Paleozoic	O			Yes		
I2 post-Pz	O			Yes		
I1 Paleozoic brachiopods						
I2 post-Pz brachiopods						
I1 Paleozoic bivalves				O		
I2 post-Pz bivalves				O		
I1 Paleozoic gastropods				O		
I2 post-Pz gastropods				No		

FIGURE 11. Plot of number of per-genus formation-occurrences versus lithological score. A, Bivalves. B, Brachiopods. C, Gastropods. Lithologies were taken from the PBDB, which provides a list of standardized terms and definitions for dominant lithology associated with a fossil collection. Lithologic score is determined by the number of different lithologies in which a genus is recorded. The plots show a general tendency for more common genera in each major group to occur in a larger number of lithologies. See Supplementary Materials 4 for data.

eter that showed a significant relationship for the pooled data consistent with taphonomic expectations, although this relationship did not hold for bivalves alone. An additional data set compiled of one of us (Kidwell 2005) permitted a test for consistency between the top 150 bivalve genera in the PBDB and all recorded bivalve genera. In comparing the rarest taxa in the PBDB ($n = 238$ genera having single documented occurrences) to the top 150 bivalves in our data set, we find no significant

difference in the proportion of taxa that are aragonitic (less chemically durable) versus calcite-bearing (69% of the singletons versus 63% of the top taxa; $p = 0.08$ according to the G-test). This implies that we would have come to the same conclusion with regard to mineralogy if we had scored the entire PBDB bivalve data set.

Additional research could focus on distinguishing the alternatives outlined in Figure 3. For example, actualistic research could test whether, in fact, taphonomic filters do or do not have a net effect on occurrence frequencies in the three target groups as a function of durability. Two scenarios, A and F, posit that filters do have a net effect, whereas one scenario (G) posits that they do not. Macroecological research could test whether spatial-temporal effects are linked to shell durability; one scenario (F) posits that they do have a net effect, whereas two scenarios (A and G) posit that they do not. Despite the remaining unresolved alternatives, this study eliminated five out of eight predicted scenarios, thus successfully reducing the number of initial working hypotheses regarding the relationship of durability and occurrence frequency in the fossil record.

Conclusions

The central goal of this research project was to test, for bivalves, gastropods, and brachiopods, whether genera with durable shells are more common in the fossil record than genera with less durable shells. Using occurrence frequencies in the Paleobiology Database (PBDB), we conclude that durability is not significantly related to occurrence frequencies for the top 150 genera of each group. The most frequently occurring genera include as many taxa having small, thin-shelled, and unreinforced shells as taxa having large, thick-shelled, ribbed, folded, or spiny shells. Exceptionally, there is an effect of mineralogy: common genera are more likely to be calcitic than aragonitic when all groups are pooled, but this does not hold up when they are considered separately. Consequently, because (1) so few of the 30 tests we performed show any significance and (2) the few that were significant were equally divided in supporting or rejecting the taphonomic ex-

pectation, we conclude that taphonomic effects relating to durability are either neutral with respect to the shell durability factors in the groups we examined or compensated for by other biological factors (e.g., less durable taxa were more abundant in the original communities). If the former is true and the taphonomic filter is indeed neutral, then the occurrence frequency patterns of our target groups in the PBDB retain biological information.

The tests of shell durability in this paper provide an example of how a compendium such as the Paleobiology Database can be used to explore questions that relate to the quality of the fossil record. Fossil occurrences are complex in origin, resulting from multiple biologic, taphonomic, spatial-temporal, and methodological factors, which at this point remain difficult to parse. Conceptually (Fig. 3) we found it useful to differentiate between alternative biological starting points, with the frequency distributions of durable versus less durable taxa being modified by taphonomic filters, spatial-temporal effects, or both. Through this exercise, we became aware of what will be needed to evaluate these effects rigorously, namely data on (1) the original biological distributions of durable versus less durable taxa (supply-side issues), (2) the macro-ecological properties of taxa (linkages among species-level characters such as shell type, body size, geographic range), and (3) more information on taphonomic filters affecting shell preservation in modern benthic environments. The approach initiated with this study should be useful in exploring the nature of the fossil record for other marine groups, as well as for terrestrial organisms such as tetrapods and plants. Defining more precisely the taxonomic or morphological levels at which durability does affect large-scale (or any-scale) patterns in the preserved fossil record will be critical to the goal of differentiating taphonomic and biological signals in the fossil record.

Acknowledgments

Literature Cited

Adrain, J. M., and S. R. Westrop. 2000. An empirical assessment of taxic paleobiology. Science 289(5476):110–112.

Alroy, J., C. R. Marshall, R. K. Bambach, K. Bezusko, M. Foote, F. T. Fürsich, T. A. Hansen, S. M. Holland, L. C. Ivany, D. Jablonski, D. K. Jacobs, D. C. Jones, M. A. Kosnik, S. Lidgard, S. Low, A. I. Miller, P. M. Novack-Gottshall, T. D. Olszewski, M. E. Patzkowsky, D. M. Raup, K. Roy, J. J. Sepkoski, M. G. Sommers, P. J. Wagner, and A. Webber. 2001. Effects of sampling standardization on estimates of Phanerozoic marine diversification. Proceedings of the National Academy of Sciences USA 98:6261–6266.

Allison, P. A., and D. E. G. Briggs. 1993. Paleolatitudinal sampling bias, Phanerozoic species diversity, and the end-Permian extinction. Geology 21:65–68.

Briggs, D. E. G. 1995. Experimental taphonomy. Palaios 10:539–550.

Brown, J. H. 1995. Macroecology. Chicago University Press, Chicago.

Carter, J. G. 1990. Skeletal biomineralization: patterns, processes, and evolutionary trends. Van Nostrand Reinhold, New York.

Chave, K. E. 1964. Skeletal durability and preservation. Pp. 377–387 *in* J. Imbrie and N. D. Newell, eds. Approaches to paleoecology. Wiley, New York.

Cummins, H., E. N. Powell, R. J. Stanton Jr., and G. Staff. 1986. The size-frequency distribution in palaeoecology; effects of taphonomic processes during formation of molluscan death assemblages in Texas bays. Palaeontology 29:495–518.

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

Harper, E. M. 2000. Are calcitic layers an effective adaptation against shell dissolution in the Bivalvia? Journal of Zoology 251:179–186.

Jackson, J. B. C., and K. G. Johnson. 2001. Measuring past biodiversity. Science 293(5539):2401.

Jope, H. M. 1965. Composition of brachiopod shell. Pp. 156–164 *in* A. Williams et al. Brachiopoda, Vol. 1. Part H of R. C. Moore, ed. Treatise on invertebrate paleontology. Geological Society of America, New York, and University of Kansas Press, Lawrence.

Kidwell, S. M. 2005. Shell composition has no net impact on large-scale evolutionary patterns in molluscs. Science 307:914–917.

Kidwell, S. M., and D. W. J. Bosence. 1991. Taphonomy and time-averaging of marine shelly faunas. Pp. 115–209 *in* P. A. Allison and D. E. G. Briggs, eds. Taphonomy. Plenum, New York.

Kidwell, S. M., and P. J. Brenchley. 1994. Patterns in bioclastic accumulation through the Phanerozoic: changes in input or in destruction. Geology 22:1139–1143.

———. 1996. Evolution of the fossil record: thickness trends in marine skeletal accumulations and their implications. Pp. 290–336 *in* D. Jablonski, D. H. Erwin, and J. H. Lipps, eds. Evolutionary paleobiology: essays in honor of James W. Valentine. University of Chicago Press, Chicago.

Koch, C. F. 1978. Bias in the published fossil record. Paleobiology 4:367–372.

Kowalewski, M., M. Carroll, L. Casazza, N. Gupta, B. Hannisdal, A. Hendy, R. A. Krause Jr., M. LaBarbera, D. G. Lazo, C. Messina, S. Puchalski, T. A. Rothfus, J. Sälgeback, J. Stempien, R. C. Terry, and A. Tomasovych. 2003. Quantitative fidelity of brachiopod-mollusk assemblages from modern subtidal environments of San Juan Islands, USA. Journal of Taphonomy 1:43–65.

Krause, R. A., J. Stempien, M. Kowalewski, and A. I. Miller. 2002. Differences in size of early Paleozoic bivalves and brachiopods: the influence of intrinsic and extrinsic factors on body size evolution. Geological Society of America Abstracts with Programs 34(6):33.

Martin, R. E. 1999. Taphonomy: a process approach. Cambridge University Press, Cambridge.

Martin, R. E., J. F. Wehmiller, M. S. Harris, and W. D. Liddell. 1996. Comparative taphonomy of bivalves and foraminifera from Holocene tidal flat sediments, Bahía la Choya, Sonora, Mexico (Northern Gulf of California): taphonomic grades and temporal resolution. Paleobiology 22:80–90.

McClain, C. R. 2004. Connecting species richness, abundance, and body size in deep sea gastropods. Global Ecology and Biogeography 13:327–334.

Parsons, K. M., and C. E. Brett. 1991. Taphonomic processes and biases in modern marine environments: an actualistic perspective on fossil assemblage preservation. Pp. 22–65 *in* S. K. Donovan, ed. The processes of fossilization. Columbia University Press, New York.

Patterson, C., and A. B. Smith. 1987. Is the periodicity of extinctions a taxonomic artefact? Nature 330:248–251.

Peck, L. S., A. Clarke, and L. J. Holmes. 1987. Size, shape and the distribution of organic matter in the Recent Antarctic brachiopod *Liothyrella uva*. Lethaia 20:33–40.

Peters, S. E., and M. Foote. 2001. Biodiversity in the Phanerozoic: a reinterpretation. Paleobiology 27:583–601.

Plotnick, R. E. and P. Wagner. In press. Round up the usual suspects: occurrence distribution and wastebasket taxa in the fossil record. Paleobiology.

Plotnick, R. E., G. T. Ventura, M. Medved, and C. Hudson. 2002. Round up the usual suspects: ubiquitous taxa and systematic inertia. Geological Society of America Abstracts with Programs 34(6):283–13.

Preston, F. W. 1948. The commonness, and rarity, of species. Ecology 29:254–283.

Raup, D. M. 1972. Taxonomic diversity during the Phanerozoic. Science 177:1065–1071.

———. 1976. Species diversity in the Phanerozoic: an interpretation. Paleobiology 2:289–297.

Roy, K., D. Jablonski, and K. K. Martien. 2000. Invariant size-frequency distributions along a latitudinal gradient in marine bivalves. Proceedings of the National Academy of Sciences USA 97:13150–13155.

Roy, K., D. Jablonski, and J. W. Valentine. 2001. Climate change, species range limits, and body size in marine bivalves. Ecology Letters 4:366–370.

Sanders, D. 2003. Syndepositional dissolution of calcium carbonate in neritic carbonate environments: geological recognition, processes, potential significance. Journal of African Earth Sciences 36:99–134.

Sepkoski, J. J., Jr. 1993. Ten years in the library: new data confirm paleontological patterns. Paleobiology 19:43–51.

Smith, A. B. 2001. Large-scale heterogeneity of the fossil record: implications for Phanerozoic biodiversity studies. Philosophical Transactions of the Royal Society of London B 356:351–367.

Smith, A. B., A. S. Gale, and N. E. A. Monks. 2001. Sea-level change and rock-record bias in the Cretaceous: a problem for extinction and biodiversity studies. Paleobiology 27:241–253.

Sohl, N. F., and C. F. Koch. 1983. Upper Cretaceous (Maestrichtian) Mollusca from the *Haustator bilira* Assemblage Zone in the East Gulf Coastal Plain. U.S. Geological Survey Open-File Report 83–451:1–239.

Taylor, J. D., W. J. Kennedy, and A. Hall. 1969. The shell structure and mineralogy of the Bivalvia. Introduction. Nuculacea-Trigonacea. Bulletin of the British Museum (Natural History), Zoology 3(Suppl.):1–125.

———. 1973. The shell structure and mineralogy of the Bivalvia. II. Lucinacea—Clavagellacea. Conclusions. Bulletin of the British Museum (Natural History), Zoology 22(Suppl.):253–294.

Valentine, J. W. 1970. How many marine invertebrate fossil species? A new approximation. Journal of Paleontology 44:410–415.

Wagner, P. J. 1995. Stratigraphic tests of cladistic hypotheses. Paleobiology 21:153–178.

Warwick, R. M., and K. R. Clark. 1996. Relationships between body-size, species abundance, and diversity in marine benthic assemblages: facts or artifacts. Journal of Experimental Marine Biology and Ecology 202:63–71.

Williams, A., S. J. Carlson, and C. H. C. Brunton. 2000. Brachiopod classification. Pp. 1–27 *in* W. Williams et al. Brachiopoda (revised), Vols. 2, 3. Part H of R. L. Kaesler and R. C. Moore, eds. Treatise on invertebrate paleontology. Geological Society of America, Boulder, Colo., and University of Kansas Press, Lawrence.

Wright, P., L. Cherns, and P. Hodges. 2003. Missing molluscs: field testing taphonomic loss in the Mesozoic through early large-scale aragonite dissolution. Geology 31:211–214.

Zar, J. H. 1999. Biostatistical analysis, 4th ed. Prentice-Hall, Englewood Cliffs, NJ.

Zuschin, M., and P. G. Oliver. 2003. Fidelity of molluscan life and

death assemblages on sublittoral hard substrata around granitic islands of the Seychelles. Lethaia 36:133–149.

Zuschin, M., M. Stachowitsch, and R. J. Stanton. 2003. Patterns and processes of shell fragmentation in modem and ancient marine environments. Earth-Science Reviews 63:33–82.

Appendix

Data Quality

Most of the taxa in our data set were recognized as common prior to the creation of the database, supporting the assumption that PBDB fairly represents the known molluscan and brachiopod fossil record in this respect. It is clear from previous work that large segments of the fossil record are not represented in the scientific literature because of historical and geographic contingencies, with potential impact on reconstructions of evolutionary and ecological trends through the Phanerozoic. For example, it is known that sampling intensity biases (Raup 1972) distort these trends, including changes in the number of available fossil samples, and indirect proxies for sampling such as variation in the amount of available outcrop (Raup 1972, 1976; Smith 2001) and/or the number of named formations (Peters and Foote 2001). Sampling intensity biases can be accounted for (Alroy et al. 2001), but additional problems include changes through time in the latitudinal distribution of fossil samples (Allison and Briggs 1993) changes in beta diversity driven by paleobiogeography and the geographic concentration of samples (Valentine 1970; Sepkoski 1993) changes in sampling of environments (Smith et al. 2001) monographic effects such as variation in the quality of taxonomy (Patterson and Smith 1987; Wagner 1995; Adrain and Westrop 2000) and the general fact that rare taxa are underrepresented in the record (Koch 1978; Jackson and Johnson 2001). The durability versus occurrence-frequency patterns we found for the most common shelly taxa could be seriously affected by these issues if there were many unsampled regions and time periods that have durability-occurrence relationships contrary to what is documented in the available fossil record.

Methodology

Evaluating Variability in Durability Scores

To evaluate potential problems related to the combined sampling and operator errors, all of the authors measured and scored the durability variables for the same subset of 15 genera. The authors shared responsibility for collecting taphonomic data from the sources available to them. Because each operator used a different set of specimens (often from congeneric species, time intervals, and depositional systems), this estimate provides the worst-case scenario for the amount of variability in our genus-level estimates of durability. The test indicates that variation was low except for maximum shell size, where the median coefficient on a natural log scale was 0.47. This reflects the fact that congeneric species can vary considerably in size, but it is a small proportion of the total range of size values in our data set (Fig. 5; 0.47 would span only one-tenth of the range of values on the x-axis).

Validity of Taxa and Time Intervals as Proxies

The two selected time intervals represent substantial portions of the Paleozoic and post-Paleozoic fossil record. In terms of PBDB coverage, these time intervals were, and still are, the best-represented segments of the Phanerozoic. The three selected groups represent 43% of the genera and 66% of the occurrences in PBDB as of September 2003, representing a large proportion of the higher taxa recorded in the marine fossil record (Alroy et al. 2001). Thus, whereas the results and interpretations presented below primarily apply to mollusks and brachiopods from I-1 and I-2 time intervals, these groups and time intervals represent a substantial portion of the entire Phanerozoic marine record.

Monographic Biases I: Variable Research Intensity

Occurrence frequency estimates may be heavily biased by variable research intensity across formations. Genera may appear frequent simply because they happened to occur in formations that have been extensively sampled (e.g., Providence Formation [Sohl and Koch 1983]) and are therefore reported in a large number of collection lists. To minimize this effect of multiple (and potentially duplicative) records, we binned occurrence data by formation (''formation-occurrences''). Such binning should dampen the effects of uneven field and monographic effects among geologic periods and regions. For the genera in our analysis, the coefficient of variation for unbinned occurrences is 146%, whereas for the binned data it is 114%. Nevertheless, the correlation between the number of unbinned and binned occurrences per time bin is quite high, 0.71, suggesting that the two metrics yield consistent results. Indeed, with few exceptions ''formation-occurrence'' estimates and ''raw occurrence'' estimates yielded consistent results. Further support for consistency is the observation that the *ratio* of binned to non-binned occurrences does not suggest any dramatic secular trends through time (i.e., I-1 and I-2 ratios vary over comparable range of values). Thus, there is no evidence that our occurrence data set is severely affected by long-term biases in geographic and temporal coverage of the fossil record. Because ''formation-occurrence'' estimates should be less sensitive to monographic biases, all figures and tables graphing results use ''formation-occurrence'' estimates.

Monographic Biases II: Garbage Can Taxa

An implicit assumption of this analysis is that the individual genera recorded in the database are monophyletic. However, it is possible that taxa in the top 150 for each group have a large number of occurrences as a result of taxonomic practices, with many of these common genera representing ''garbage can taxa'' (Plotnick et al. 2002). Garbage can taxa can be both paraphyletic (comprising related forms grouped together because of generally primitive morphology) and/or polyphyletic (comprising mainly unrelated forms, often poorly preserved, but with some distinctive morphologic feature) and can arise as a result of either taxonomic practice or taphonomy (Plotnick and Wagner in press).

Lithology

We downloaded lithologic information for each raw occurrence to evaluate lithofacies-level effects on occurrence frequency data. Twenty lithologic terms are available for scoring PBDB occurrences with respect to sedimentary matrix (e.g., among siliciclastics, six terms ranging from claystone through conglomerates; among carbonates, eight terms ranging from lime mudstone to framestone and bindstone; plus chert, phosphorite, etc.; see PBDB website: http://paleodb.org). These were used to classify genera in terms of their occurrence in different lithologies.