SI APPENDIX

Materials and Methods

Habitat Focus of Collections. To limit body size variability caused by environmental heterogeneity, we compiled brachiopod collections from a single, well defined habitat type—thedeep-subtidal, soft-substrate habitat—well known for its plentiful sedimentary record, excellent taphonomical preservation, and relatively homogeneous environmental conditions (1, 2). We inferred depth using lithological and biological criteria: fine-grained siliciclastic and carbonate shales and mudstones with occasional interbeds of limestone, siltstone, and calcisiltite and an absence of shallow water sedimentary structures and calcareous algae (3-5). Because reduced oxygenation is often linked to small body size, we removed those collections indicating lowoxygen levels based on lithological and biological criteria—presence of pyrite/sulfides and lowoxygen specialists and absence of trace fossils (6, 7).

Instances of Size Measurement Approximation. We sought to obtain brachiopod size measurements for taxa in our database at the lowest taxonomic level possible using specimens in monographic illustrations, disregarding taxa for which we could not make reliable size measurements. When all three dimensions were not available in an illustration, we estimated non-illustrated measurements assuming isometric dimensions with contemporaneous (i.e., from the same geological period), phylogenetically related, and morphologically similar analogs. Biases caused by such approximations are unlikely because of the wide range of sizes observed. Nearly 80% of the brachiopods in the database were coded at genus-or-finer level (Table 2);

brachiopods measured at higher levels included indeterminate genera in families (e.g., Chonetidae indet.) displaying less than one order of volumetric variation in the database.

Additional Details on Time-Series Analyses of Size Evolution. Rarefaction (8, 9) of mean and minimum size trends was conducted at a standard sample size of 10 genera per bin and 2000 replicates. Within-clade trend dynamics were evaluated using maximum-likelihood-based model selection (10, 11), as explained in text. These analyses were conducted separately for the entire phylum and for all constituent clades (classes, orders, and families) with time series spanning at least five bins and ten total occurrences. Within-order and within-family trends spanning at least two bins were further evaluated using a test of the behavior of minimum and maximum size (12), as explained in text. To mitigate against statistical artifacts, subset analyses of this second test were conducted on orders and families with changing minimum and maximum sizes, with at least ten and five total occurrences, respectively, and that did not fall on the 1:1 line. These 1:1 cases occurred when the first and last occurrences within a clade were each represented by a single genus, unfairly biasing against observing instances of increased or decreased variance (quadrants 2 and 4). The ten best sampled orders included Acrotretida, Atrypida, Lingulida, Orthida, Orthotetida, Pentamerida, Productida, Rhynchonellida, Spiriferida, and Strophomenida; the 20 best sampled families included Acrotretidae, Amphistrophiidae, Anazygidae, Anoplothecidae, Atrypidae, Atrypinidae, Chilidiopsidae, Clorindidae, Cyrtiidae, Dalmanellidae, Delthyrididae, Eatoniidae, Obolidae, Orthidae, Rafinesquinidae, Rhynchotrematidae, Sowerbyellidae, Strophodontidae, Strophomenidae, and Triplesiidae. All statistics, quantitative analyses, and figures were prepared using R 2.5.0 for Windows (13).

Macroevolutionary Sorting Among Families. Although high taxonomic levels within the brachiopods have been well substantiated using cladistics (14-17), similar phylogenies are lacking for most brachiopod families and lower levels. Fortunately, recent taxonomic practice at these levels has benefited from a considerable degree of standardization and collaboration among brachiopod systematists (15-17), allowing us to assume that families were defined using similarly standardized criteria. In the absence of family-level cladograms, we used a series of resampling tests to evaluate whether brachiopod families displayed body size-related biases in the geological duration (a proxy for extinction rate), genus richness (a proxy for speciation rate), and body size at origination of new families. Analyses were restricted to the well represented 87 families spanning at least two bins to mitigate against sampling biases and to allow measurement of family duration (Table 6). Genus richness was measured as the number of unique genera here in each family; duration was measured as the difference between first and last occurrences, with dates set as the midpoint for each bin. Because of the similar duration of bins, dating errors should have limited effects on results.

To evaluate whether duration, genus richness, and size at origination were size-biased among families, we assigned candidate ancestor-descendent pairs at random based on order of stratigraphic occurrence. This technique has been shown to give similar—and in many cases, improved—phylogenetic reconstructions as cladistics (18-20), and has been used in other analyses of Cope's rule in fossil taxa (21, 22). Furthermore, stratigraphic order of high taxonomic levels in brachiopods corresponds well with cladistic analyses (14). Candidate ancestor-descendent pairs were chosen in two ways, at the level of families and genera. We first matched each family with a candidate ancestral family chosen at random from con-ordinal families in the previous time bin; restricting ancestral assignments to the previous bin is

reasonable given the exceptional preservation potential of brachiopods (23). The second analysis matched a single genus occurring in each family's first bin with a candidate ancestral genus chosen at random from con-ordinal genera in the previous bin. We calculated the relationship between body size and the factor of interest for each pool of ancestor-descendent pairs, and then repeated this process for 2000 iterations. For evaluation of size-biased origination of families, we used the one-sided, paired *t*-test to compare whether the body sizes of newly originating (e.g., those in their first bin) families (or an individual genus in each family) were significantly larger than candidate ancestral (e.g., those in the previous bin) families (or an individual genus); for each iteration we recorded the corresponding *p*-value. If there is a significantly positive size-bias in size at origination, the first members of new families consistently will be larger than their ancestors. For evaluation of size-biased relationships with duration or genus richness, we made similar ancestor-descendent matches, measuring the magnitudes of size change and change in family duration or genus richness from ancestor to descendent. We then measured the relationship between size and these factors using linear regression, assembling corresponding *p*values. If there is a significantly positive size-bias in these two factors, new families that are larger than their ancestors should persist longer and contain more genera, while new families that are smaller should have briefer durations and contain fewer genera. We then used the proportion of *p*-values less than 0.05 to evaluate the significance of size-bias in these factors, with 95% confidence intervals calculated using the Wilson score method with continuity correction (24). In essence, these resampling tests evaluate the sensitivity of results to differing phylogenies, in an analogous way as bootstrap support in cladistics (25). Those results that occur in most phylogenetic reconstructions are likely to withstand future cladistic analyses.

References

- 1. Brett CE (1990) in *Palaeobiology: A Synthesis*, eds Briggs DEG, Crowther PR (Blackwell Scientific, London, U.K.), pp. 239-243.
- 2. Schieber J, Zimmerle W, Sethi P eds (1998) *Shales and Mudstones* (E. Schweizerbart'sche Verlagsbuchhandlung, Stuttgart).
- 3. Holmes RW (1957) in *Treatise on marine ecology and paleoecology. Volume 1: Ecology*, ed Hedgpeth JW *GSA Mem.* 67:109-128.
- 4. Jennette DC, Pryor WA (1993) Cyclic alternation of proximal and distal storm facies: Kope and Fairview Formations (Upper Ordovician), Ohio and Kentucky. *J. of Sediment. Petrolog.* 63:183-203.
- 5. O'Brien NR, Brett CE, Taylor WL (1994) Microfabric and taphonomic analysis in determining sedimentary processes in marine mudstones: Examples from Silurian of New York. *J. of Sediment. Res.* A64:847-852.
- 6. Schovsbo NH (2001) Why barren intervals? A taphonomic case study of the Scandinavian Alum Shale and its faunas. *Lethaia* 34:271-285.
- 7. Gaines RR, Droser ML (2003) Paleoecology of the familiar trilobite *Elrathia kingii*: An early exaerobic zone inhabitant. *Geol.* 31:941-944.
- 8. Sanders HL (1968) Marine benthic diversity: A comparative study. *Am. Nat.* 102:243-282.
- 9. Hurlbert SH (1971) The nonconcept of species diversity: A critique and alternative parameters. *Ecol.* 52:577-586.
- 10. Hunt G (2006) Fitting and comparing models of phyletic evolution: Random walks and beyond. *Paleobiology* 32:578-601.
- 11. Hunt G (2007) *paleoTS: Modeling evolution in paleontological time-series*, Version 0.1-3. (http://cran.r-project.org/src/contrib/Descriptions/paleoTS.html).
- 12. Jablonski D (1997) Body-size evolution in Cretaceous mollusks and the status of Cope's rule. *Nature* 385:250-252.
- 13. R Development Core Team (2007) *R: A Language and Environment for Statistical Computing*, Version 2.5.0. (R Foundation for Statistical Computing, Vienna).
- 14. Williams A, Carlson SJ, Bruton CHC, Holmer LE, Popov L (1996) A supra-ordinal classification of the Brachiopoda. *Philos. Trans. R. Soc. London Ser. B* 351:1171-1193.
- 15. Williams A, Bruton CHC, Carlson SJ (1997) *Brachiopoda 1. Part H, Revised. Treatise on Invertebrate Paleontology* (Geological Society of America, New York, and Univ. of Kansas, Lawrence, KS).
- 16. Williams A, Bruton CHC, Carlson SJ (2000) *Brachiopoda 2 and 3. Part H, Revised. Treatise on Invertebrate Paleontology* (Geological Society of America, New York, and Univ. of Kansas, Lawrence, KS).
- 17. Williams A, Bruton CHC, Carlson SJ (2002) *Brachiopoda 4. Part H, Revised. Treatise on Invertebrate Paleontology* (Geological Society of America, New York, and Univ. of Kansas, Lawrence, KS).
- 18. Wagner PJ (1995) Stratigraphic tests of cladistic hypotheses. *Paleobiology* 21:153-178.
- 19. Fox DL, Fisher DC, Leighton LR (1999) Reconstructing phylogeny with and without temporal data. *Science* 284:1816-1819.
- 20. Angielczyk KD, Fox DL (2006) Exploring new uses for measures of fit of phylogenetic hypotheses to the fossil record. *Paleobiology* 32:147-165.
- 21. Alroy J (1998) Cope's rule and the dynamics of body mass evolution in North American mammals. *Science* 280:731-734.
- 22. Alroy J (2000) Understanding the dynamics of trends within evolving lineages. *Paleobiology* 26:319-329.
- 23. Foote M, Sepkoski JJ, Jr (1999) Absolute measures of the completeness of the fossil record. *Nature* 398:415-417.
- 24. Newcombe RG (1998) Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat. Med.* 17:857-872.
- 25. Felsenstein J (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.

Table 1. Summary of brachiopod database used in analyses

Taxon occurrences are the sum of all individual genus (or genus-equivalent) occurrences across all time bins; taxa occurring in more than one bin will have multiple occurrences. All other taxonomic totals are measures of total unique richness in the database.

Table 2. Taxonomic resolution of size data in the brachiopod database

N is the number of genera (or genus equivalents) coded at a particular taxonomic level. Most are coded at genus-or-finer level. Brachiopods measured at higher levels included indeterminate genera in families (e.g., Chonetidae indet.) displaying less than one order of volumetric magnitude variation in the database; the size of such indeterminate taxa was estimated using the size of a contemporaneous (same geological period) type species in each indeterminate taxon.

Table 3. Results of maximum likelihood comparison of evolutionary models for brachiopod clades during the Cambrian-

Devonian

Table 3 continued

Only clades (represented here by primarily cladistically defined taxonomic groups) with minimally ten total occurrences over five intervals are included here. N and bins are the number of genus occurrences and bins, respectively, included in each analysis; AIC_C is the small-sample unbiased Akaike Information Criterion (1, 2). Models analyzed using refs (3, 4). For the DRW model, the parameters include a mean rate ($\mu_{step-DRW}$) and variance ($\sigma^2_{step-DRW}$) for the normally distributed, within-clade distribution of size transitions. For the URW model, the normal distribution of size transitions is centered over a mean of zero, with only a single parameter for its variance (σ^2 _{step-URW}). The stasis model includes mean (θ_{stasis}) and variance (ω_{stasis}) parameters for a normal distribution of sizes in which sizes are drawn independently at any time interval; that is, size transitions do not accumulate during the time series because the stasis model is nonautocorrelated. Therefore, the mean and variance define the optimal normal distribution of sizes for each clade. Relative model support is provided as a probability using Akaike weights (2, 5); model weights with substantial (>0.05) support are highlighted in bold. Variances within each clade were not pooled across bins because of significant heterogeneity. Variance parameter values listed as zero here are very small (1×10^{-9}) but not zero.

- 1. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans. on Automat. Control* 19:716-723.
- 2. Anderson DR, Burnham KP, Thompson WL (2000) Null hypothesis testing: problems, prevalence, and an alternative. *J. of Wildl. Manag.* 64:912-923.
- 3. Hunt G (2006) Fitting and comparing models of phyletic evolution: random walks and beyond. *Paleobiology* 32:578-601.
- 4. Hunt G (2007) *paleoTS: Modeling evolution in paleontological time-series*, Version 0.1-3. (http://cran.r-project.org/src/contrib/Descriptions/paleoTS.html).
- 5. Posada D, Buckley TR (2004) Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. *Syst. Biol.* 53:793-808.

Table 4. Minimum-change/maximum-change transitions for brachiopod clades during the

	Cambrian-Devonian

Table 4 continued

Table 4 continued

N is the number of genus occurrences included in analysis of each clade; only clades (represented here by primarily cladistically defined taxonomic groups) with occurrences spanning at least two time intervals are included here. ∆Min and ∆Max are changes in minimum and maximum sizes (log_{10} ml) from first to last occurrence within each clade; sign is the direction of these transitions. Quadrant is the quadrant these transitions plot in Figure 3; quadrant zero means that the transition falls on the origin or along the axes (in other words, at least one size extreme was unchanging in this clade).

Table 5. Wilcoxon ranked-sum and *t***-test statistical results for tests of whether**

directionality parameters are positively biased

One-sided tests of whether the directionality parameters in Figure 3 have positive tendencies at each clade level, consistent with a driven mechanism to size increase (i.e., Cope's rule). Distributions for larger, more inclusive clades (classes and orders) have positive tendencies whereas that for smaller, constituent clades (families) is not statistically different from zero tendency. The marginal significance using the Wilcoxon test for classes is caused by small sample size (four classes) that diminishes the power of this test.

	Initial mean size		
Family	$(log_{10} ml)$	Duration (Myr)	Genus richness
Acrotretidae	-2.544	84.9	10
Ambocoeliidae	-0.291	42.2	6
Amphistrophiidae	0.101	19.9	4
Anazygidae	-0.448	19.5	3
Anoplothecidae	-1.316	30.7	3
Athyrididae	0.050	20.1	$\overline{2}$
Atrypidae	0.419	44.6	8
Atrypinidae	-0.497	44.6	5
Bimuriidae	-0.742	17.9	$\overline{2}$
Camarotoechiidae	0.446	55.5	\overline{c}
Camerellidae	-0.416	51.5	$\overline{2}$
Centronellidae	0.279	20.1	\overline{c}
Chilidiopsidae	-0.356	24.5	4
Chonetidae	-0.165	42.2	3
Christianiidae	-0.315	8.6	$\mathbf 1$
Chrustenoporidae	-0.601	10.9	$\overline{2}$
Clorindidae	0.417	44.6	3
Craniidae	-0.221	55.5	3 $\overline{2}$
Craniopsidae	-2.886	64.1 10.9	$\overline{2}$
Cryptonellidae	-0.450	10.9	1
Cyclospiridae Cyrtiidae	-0.716	44.6	5
Cyrtinidae	0.313 0.173	20.1	1
Dalmanellidae	-0.358	75.6	17
Delthyrididae	0.051	44.6	12
Dicoelosiidae	-1.231	35.4	$\overline{2}$
Discinidae	-0.623	64.1	1
Douvillinidae	-0.387	42.2	6
Draboviidae	0.399	44.0	6
Eatoniidae	-0.072	19.9	5
Glassiidae	-0.674	44.6	1
Glyptomenidae	-0.558	8.6	3
Glyptorthidae	0.181	8.6	1
Gypidulidae	0.527	10.6	3
Harknessellidae	0.621	8.6	\overline{c}
Hebetoechiidae	0.458	10.6	$\overline{2}$
Hesperorthidae	-0.005	44.0	5
Heterorthidae	0.571	9.3	1
Leiorhynchidae	0.254	20.1	\overline{c}
Leptaenoideidae	0.069	10.6	$\mathbf 1$
Leptestiidae	-1.047	24.8	\overline{c}
Leptostrophiidae	-0.627	67.0	8
Lingulidae	-0.225	115.6	$\overline{\mathbf{c}}$
Lissatrypidae	0.079	13.9	3
Machaerariidae	0.085	24.5	3
Meristellidae	0.592	55.5	6
Mucrospiriferidae	0.705	11.5	$\overline{\mathbf{c}}$
Mutationellidae	-1.001	10.9	$\overline{\mathbf{c}}$
Nanorthidae	-1.045	9.9	\overline{c}
Nucleospiridae	-0.432	44.6	$\mathbf{1}$
Obolidae	-0.145	60.2	9
Orthidae	0.138	62.1	5
Parastrophinidae	-0.213	44.0	3
Paterinidae	-1.489	14.5	$\overline{2}$
Paurorthidae	-0.315	37.7	1

Table 6. Summary data used in analyses of sorting among brachiopod families

Table 6 continued

Only families with occurrences spanning at least two time intervals are included here.

Initial mean size is for each family's first bin only; duration is the difference between each

family's first and last occurrence, with dates taken from bin midpoints.

Table 7. Results of size-biased sorting analyses among brachiopod families

Ancestor-descendent pairs were matched from the bin in which each family originates to its previous bin, choosing an ancestor at random among con-ordinal taxa; analyses were conducted choosing either a single genus (genus-genus) or all genera in a family (family-family) for each ancestor and descendent in their corresponding bins. The first pair of analyses used the one-sided paired *t*-test to evaluate whether originating families were preferentially larger than candidate ancestral families. The remaining two pairs used linear regression to evaluate whether the magnitude of ancestor-descendent size change was correlated with the magnitude of change in family duration or genus richness. This process was repeated 2000 times, tallying the proportion of instances in which the relationship, measured by the *p*-value, was less than alpha = 0.05. 95%-confidence intervals are calculated using the Wilson score method with continuity correction (1). Distributions of *p*-values are in Figure 7.

1. Newcombe RG (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat. Med.* 17:857-872.

Fig. 5. Sample-standardized size trends in mean and minimum (dashed line) size of Cambrian-Devonian brachiopod genera based on rarefaction (2000 replicates) to 10 genera (or genus-equivalents). Data points are the observed genus body volumes, as in Fig. 1. Standard error bars around means are one standard deviation from the distribution of 2000 bootstrap replicates. There were too few genera during the C2, C4, O1, and D5 intervals to yield standardized estimates.

Fig. 6. Mean size trends in individual Cambrian-Devonian brachiopod clades: (*A*) classes, (*B*) orders, and (*C*) families. Unlike in Fig. 2, only consecutive mean sizes within each clade are connected with a line; this limits potentially false visual inferences of size evolution. Only clades with a minimum of ten occurrences over five intervals are figured, except for the addition of order Paterinida included to visualize similarity of Cambrian trends. All brachiopod genus sizes are plotted to allow visual assessment of the widespread taxonomic nature of the overall size increase. Other plot details are same as in Fig. 1. Maximum-likelihood trend statistics are available in Table 4.

Fig. 7. Distributions of *p*-values from size-biased sorting analyses among brachiopod families. (*A* and *B*) Distribution of *p*-values for resampled *t*-tests comparing ancestor-descendent body sizes of originating families at levels of families (*A*) and individual genera (*B*). (*C* and *D*) Distribution of *p*-values for linear regression between change in body size and change in family

duration from ancestral to descendent families (*C*) and individual genera (*D*). (*E* and *F*) Distribution of *p*-values for linear regression between change in size and change in family genus richness from ancestral to descendent families (*E*) and individual genera (*F*). Dashed lines denote alpha = 0.05. Each resampling test included 2000 iterations. Distributions with substantial phylogenetic iterations falling under alpha = 0.05 (i.e., Fig. 7 *A* and *B*) are likely to remain significant in the face of future phylogenetic studies. Proportions of significant *p*values falling below alpha = 0.05 are reported in Table 7.